Secured Communication over Frequency-Selective Fading Channels: A Practical Vandermonde Precoding

Mari Kobayashi, Merouane Debbah, Shlomo Shamai

Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 386547, 19 pages
doi:10.1155/2009/386547

Research Article

Secured Communication over Frequency-Selective Fading Channels: A Practical Vandermonde Precoding

Mari Kobayashi,[1] Mérouane Debbah,[2] and Shlomo Shamai (Shitz)[3]

[1] *Department of Telecommunications, SUPELEC, 3 Rue Joliot-Curie, Gif-sur-Yvette, 91192, France*
[2] *Alcatel-Lucent Chair on Flexible Radio, SUPELEC, 3 Rue Joliot-Curie, Gif-sur-Yvette, 91192, France*
[3] *Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel*

Correspondence should be addressed to Mari Kobayashi, mari.kobayashi@supelec.fr

Received 2 February 2009; Accepted 16 June 2009

Recommended by H. Vincent Poor

We study the frequency-selective broadcast channel with confidential messages (BCC) where the transmitter sends a confidential message to receiver 1 and a common message to receivers 1 and 2. In the case of a block transmission of N symbols followed by a guard interval of L symbols, the frequency-selective channel can be modeled as a $N \times (N + L)$ Toeplitz matrix. For this special type of multiple-input multiple-output channels, we propose a practical Vandermonde precoding that projects the confidential messages in the null space of the channel seen by receiver 2 while superposing the common message. For this scheme, we provide the achievable rate region and characterize the optimal covariance for some special cases of interest. Interestingly, the proposed scheme can be applied to other multiuser scenarios such as the $K + 1$-user frequency-selective BCC with K confidential messages and the two-user frequency-selective BCC with two confidential messages. For each scenario, we provide the secrecy degree of freedom (s.d.o.f.) region of the corresponding channel and prove the optimality of the Vandermonde precoding. One of the appealing features of the proposed scheme is that it does not require any specific secrecy encoding technique but can be applied on top of any existing powerful encoding schemes.

1. Introduction

We consider a secured medium such that the transmitter wishes to send a confidential message to its receiver while keeping the eavesdropper, tapping the channel, ignorant of the message. Wyner [1] introduced this model named the wiretap channel to model the degraded broadcast channel where the eavesdropper observes a degraded version of the receiver's signal. In this model, the confidentiality is measured by the equivocation rate, that is, the mutual information between the confidential message and the eavesdropper's observation. For the discrete memoryless degraded wiretap channel, Wyner characterized the capacity-equivocation region and showed that a nonzero secrecy rate can be achieved [1]. The most important operating point on the capacity-equivocation region is the secrecy capacity, that is, the largest reliable communication rate such that the eavesdropper obtains no information about the confidential message (the equivocation rate is as large as the message rate). The secrecy capacity of the Gaussian wiretap channel was given in [2]. Csiszár and Körner considered a more general wiretap channel in which a common message for both receivers is sent in addition to the confidential message [3]. For this model known as the broadcast channel with confidential (BCC) messages, the rate-tuple of the common and confidential messages was characterized.

Recently, a significant effort has been made to opportunistically exploit the space/time/user dimensions for secrecy communications (see, e.g., [4–14] and references therein). In [4], the secrecy capacity of the ergodic slow fading channels was characterized and the optimal power/rate allocation was derived. The secrecy capacity of the parallel fading channels was given [6, 7] where [7] considered the BCC with a common message. Moreover, the secrecy capacity of the wiretap channel with multiple antennas has been studied in [8–13, 15] and references therein. In particular, the secrecy capacity of the multiple-input multiple-output (MIMO) wiretap channel has been fully

characterized in [5, 11, 12, 14] and more recently its closed-form expressions under a matrix covariance constraint have been derived in [15]. Furthermore, a large number of recent works have considered the secrecy capacity region for more general broadcast channels. In [16], the authors studied the two-user MIMO Gaussian BCC where the capacity region for the case of one common and one confidential message was characterized. The two-user BCC with two confidential messages, each of which must be kept secret to the unintended receiver, has been studied in [17–20]. In [18], Liu and Poor characterized the secrecy capacity region for the multiple-input single-output (MISO) Gaussian BCC where the optimality of the secret dirty paper coding (S-DPC) scheme was proved. A recent contribution [19] extended the result to the MIMO Gaussian BCC. The multireceiver wiretap channels have been also studied in [21–26] (and reference therein) where the confidential messages to each receiver must be kept secret to an external eavesdropper. It has been proved that the secrecy capacity region of the MIMO Gaussian multireceiver wiretap channels is achieved by S-DPC [24, 26].

However, very few work have exploited the frequency selectivity nature of the channel for secrecy purposes [27] where the zeros of the channel provide an opportunity to "hide" information. This paper shows the opportunities provided by the broad-band channel and studies the frequency-selective BCC where the transmitter sends one confidential message to receiver 1 and one common message to both receivers 1 and 2. The channel state information (CSI) is assumed to be known to both the transmitter and the receivers. We consider the quasistatic frequency-selective fading channel with $L + 1$ paths such that the channel remains fixed during an entire transmission of n blocks for an arbitrary large n. It should be remarked that in general the secrecy rate cannot scale with signal-to-noise ratio (SNR) over the channel at hand, unless the channel of receiver 2 has a null frequency band of positive Lebesgue measure (on which the transmitter can "hide" the confidential message). In this contribution, we focus on the realistic case where receiver 2 has a full frequency band (without null subbands) but operates in a reduced dimension due to practical complexity issues. This is typical of current orthogonal frequency division multiplexing (OFDM) standards (such as IEEE802.11a/WiMax or LTE [28–30]) where a guard interval of L symbols is inserted at the beginning of each block to avoid the interblock interference and both receivers discard these L symbols. We assume that both users have the same standard receiver, in particular receiver 2 cannot change its hardware structure. Studying secure communications under this assumption is of interest in general and can be justified since receiver 2 is actually a legitimate receiver which can receive a confidential message in other communication periods. Of course, if receiver 2 is able to access the guard interval symbols, it can extract the confidential message and the secrecy rate falls down to zero. Although we restrict ourselves to the reduced dimension constraint in this paper, other constraints on the limited capability at the unintended receiver such as energy consumption or hardware complexity might provide a new paradigm to design physical layer secrecy systems.

In the case of a block transmission of N symbols followed by a guard interval of L symbols discarded at both receivers, the frequency-selective channel can be modeled as an $N \times (N + L)$ MIMO Toeplitz matrix. In this contribution, we aim at designing a practical linear precoding scheme that fully exploits the degrees of freedom (d.o.f.) offered by this special type of MIMO channels to transmit both the common message and the confidential message. To this end, let us start with the following remarks. On one hand, the idea of using OFDM modulation to convert the frequency-selective channel represented by the Toeplitz matrix into a set of parallel fading channel turns out to be useless from a secrecy perspective. Indeed, it is known that the secrecy capacity of the parallel wiretap fading channels does not scale with SNR [7]. On the other hand, recent contributions [5, 11, 12, 14, 15] showed that the secrecy capacity of the MIMO wiretap channel grows linearly with SNR, that is, $r \log \mathrm{SNR}$ where r denotes the secrecy degree of freedom (s.d.o.f.) (to be specified). In the high SNR regime, the secrecy capacity of the MISO/MIMO wiretap channel is achieved by sending the confidential message in the null space of the eavesdropper's channel [10, 11, 14, 15, 18, 19]. Therefore, OFDM modulation is highly suboptimal in terms of the s.d.o.f.

Inspired by these remarks, we propose a linear Vandermonde precoder that projects the confidential message in the null space of the channel seen by receiver 2 while superposing the common message. Thanks to the orthogonality between the precoder of the confidential message and the channel of receiver 2; receiver 2 obtains no information on the confidential message. This precoder is regarded as a single-antenna frequency beamformer that nulls the signal in certain directions seen by receiver 2. The Vandermonde structure comes from the fact that the frequency beamformer is of the type $[1, a_i, a_i^2, \ldots, a_i^{N+L}]^T$ where a_i is one of the roots of the channel seen by receiver 2. Note that Vandermonde matrices [31] have already been considered for cognitive radios [32] and CDMA systems [33] to reduce/null interference but not for secrecy applications. One of the appealing aspects of Vandermonde precoding is that it does not require a specific secrecy encoding technique but can be applied on top of any classical capacity achieving encoding scheme.

For the proposed scheme, we characterize its achievable rate region, the rate-tuple of the common message, the confidential message, respectively. Unfortunately, the optimal input covariances achieving their boundary are generally difficult to compute due to the nonconvexity of the weighted sum rate maximization problem. Nevertheless, we show that there are some special cases of interest such as the secrecy rate and the maximum sum rate point which enable an explicit characterization of the optimal input covariances. In addition, we provide the achievable d.o.f. region of the frequency-selective BCC, reflecting the behavior of the achievable rate region in the high SNR regime, and prove that the Vandermonde precoding achieves

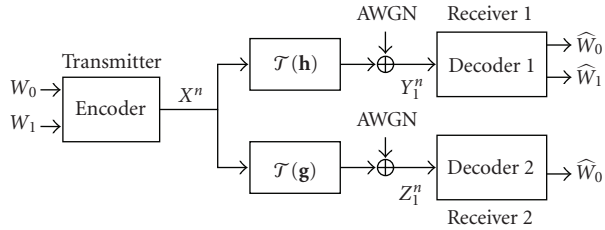

Figure 1: Frequency-selective broadcast channels with confidential messages.

this region. More specifically, it enables to simultaneously transmit l streams of the confidential message and $N-l$ streams of the common message for $l \le L$ simultaneously over a block of $N+L$ dimensions. Interestingly, the proposed Vandermonde precoding can be applied to multiuser secure communication scenarios: (a) a $K+1$-user frequency-selective BCC with K confidential messages and one common message, (b) a two-user frequency-selective BCC with two confidential messages and one common message. For each scenario, we characterize the achievable s.d.o.f. region of the corresponding frequency-selective BCC and show the optimality of the Vandermonde precoding.

The paper is organized as follows. Section 2 presents the frequency-selective fading BCC. Section 3 introduces the Vandermonde precoding and characterizes its achievable rate region as well as the optimal input covariances for some special cases. Section 4 provides the application of the Vandermonde precoding to the multiuser secure communications scenarios. Section 5 shows some numerical examples of the proposed scheme in the various settings, and finally Section 6 concludes the paper.

Notation. In the following, upper (lower boldface) symbols will be used for matrices (column vectors) whereas lower symbols will represent scalar values, $(\cdot)^T$ will denote transpose operator, $(\cdot)^\star$ conjugation, and $(\cdot)^H = ((\cdot)^T)^\star$ hermitian transpose. $\mathbf{I}_n$, $0_{n\times m}$ represent the $n\times n$ identity matrix, $n\times m$ zero matrix. $|\mathbf{A}|$, $\mathrm{rank}(\mathbf{A})$, $\mathrm{tr}(\mathbf{A})$ denote a determinant, rank, trace of a matrix $\mathbf{A}$, respectively. $\mathbf{x}^n$ denotes the sequence $(\mathbf{x}[1],\ldots,\mathbf{x}[n])$. w, u, v, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ denote the realization of the random variables W, U, V, X, Y, Z. Finally, "$\preceq$" denotes less or equal to in the positive semidefinite ordering between positive semidefinite matrices, that is, we have $\mathbf{A}\preceq\mathbf{B}$ if $\mathbf{B}-\mathbf{A}$ is positive semidefinite.

2. System Model

We consider the quasistatic frequency-selective fading BCC illustrated in Figure 1. The received signal $\mathbf{y}[t], \mathbf{z}[t] \in \mathbb{C}^{N\times 1}$ of receivers 1, 2 at block t is given by

$$\begin{aligned}\mathbf{y}[t] &= \mathcal{T}(\mathbf{h})\mathbf{x}[t] + \mathbf{n}[t],\\ \mathbf{z}[t] &= \mathcal{T}(\mathbf{g})\mathbf{x}[t] + \boldsymbol{\nu}[t], \quad t = 1,\ldots,n,\end{aligned} \tag{1}$$

where $\mathcal{T}(\mathbf{h}), \mathcal{T}(\mathbf{g})$ denote an $N\times(N+L)$ Toeplitz matrix with the $L+1$-path channel vector $\mathbf{h} = [h_L,\ldots,h_0]$ of user 1, $\mathbf{g} = [g_L,\ldots,g_0]$ of user 2, respectively, $\mathbf{x}[t] \in \mathbb{C}^{(N+L)\times 1}$ denotes the transmit vector, and finally $\mathbf{n}[t], \boldsymbol{\nu}[t] \sim \mathcal{N}_{\mathcal{C}}(0, \mathbf{I}_N)$ are mutually independent additive white Gaussian noise (AWGN). The input vector is subject to the power constraint given by

$$\frac{1}{n}\sum_{t=1}^{n}\mathbf{x}[t]^H\mathbf{x}[t] \le \overline{P}, \tag{2}$$

where we let $\overline{P} = (N+L)P$. The structure of $\mathcal{T}(\mathbf{h})$ is given by

$$\mathcal{T}(\mathbf{h}) = \begin{bmatrix} h_L & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h_L & \cdots & h_0 \end{bmatrix}. \tag{3}$$

We assume that the channel matrices $\mathcal{T}(\mathbf{h}), \mathcal{T}(\mathbf{g})$ remain constant for the whole duration of the transmission of n blocks and are known to all terminals. At each block t, we transmit $N+L$ symbols by appending a guard interval of size $L \ll N$ larger than the delay spread, which enables to avoid the interference between neighbor blocks.

The transmitter wishes to send a common message W_0 to two receivers and a confidential message W_1 to receiver 1. A $(2^{nR_0}, 2^{nR_1}, n)$ code consists of the following: (1) two message sets $\mathcal{W}_0 = \{1,\ldots,2^{nR_0}\}$ and $\mathcal{W}_1 = \{1,\ldots,2^{nR_1}\}$ with the messages W_0, W_1 uniformly distributed over the sets $\mathcal{W}_0, \mathcal{W}_1$, respectively; (2) a stochastic encoder that maps each message pair $(w_0, w_1) \in (\mathcal{W}_0, \mathcal{W}_1)$ to a codeword $\mathbf{x}^n$; (3) one decoder at receiver 1 that maps a received sequence $\mathbf{y}^n$ to a message pair $(\widehat{w}_0^{(1)}, \widehat{w}_1) \in (\mathcal{W}_0, \mathcal{W}_1)$ and another at receiver 2 that maps a received sequence $\mathbf{z}^n$ to a message $\widehat{w}_0^{(2)} \in \mathcal{W}_0$. The average error probability of a $(2^{nR_0}, 2^{nR_1}, n)$ code is defined as

$$P_e^n = \frac{1}{2^{nR_0}2^{nR_1}}\sum_{w_0\in\mathcal{W}_0}\sum_{w_1\in\mathcal{W}_1} P_e^n(w_0, w_1), \tag{4}$$

where $P_e^n(w_0, w_1)$ denotes the error probability when the message pair (w_0, w_1) is sent defined by

$$P_e^n(w_0, w_1) \triangleq \Pr\left(\left(\widehat{w}_0^{(1)}, \widehat{w}_1\right) \ne (w_0, w_1) \cup \widehat{w}_0^{(2)} \ne w_0\right). \tag{5}$$

The secrecy level of the confidential message W_1 at receiver 2 is measured by the equivocation rate R_e defined as

$$R_e \triangleq \frac{1}{n}H(W_1 \mid Z^n) \tag{6}$$

which is the normalized entropy of the confidential message conditioned on the received signal at receiver 2 and available CSI.

A rate-equivocation tuple (R_0, R_1, R_e) is said to be achievable if for any $\epsilon > 0$ there exists a sequence of codes $(2^{nR_0}, 2^{nR_1}, n)$ such that we have

$$\begin{aligned} P_e^n &\le \epsilon, \\ R_1 - R_e &\le \epsilon. \end{aligned} \tag{7}$$

In this paper, we focus on the perfect secrecy case where receiver 2 obtains no information about the confidential message W_1, which is equivalent to $R_e = R_1$. In this setting, an achievable rate region (R_0, R_1) of the general BCC (expressed in bit per channel use per dimension) is given by [3]

$$\begin{aligned} \mathcal{C}_s = \bigcup_{p(u,v,\mathbf{x})} \Bigg\{ &(R_0, R_1) : R_0 \le \frac{1}{N+L} \min\{I(U;Y), I(U;Z)\}, \\ &R_1 \le \frac{1}{N+L} [I(V;Y \mid U) - I(V;Z \mid U)] \Bigg\}, \end{aligned} \tag{8}$$

where the union is over all possible distribution U, V, X satisfying [20, Lemma 1]

$$U, V \longrightarrow X \longrightarrow Y, Z, \tag{9}$$

where U might be a deterministic function of V. Recently, the secrecy capacity region $\mathcal{C}_s$ of the two-user MIMO-BCC (1) was characterized in [16] and is given by all possible rate tuples (R_0, R_1) satisfying

$$\begin{aligned} R_0 &\le \frac{1}{N+L} \min\left\{ \log \frac{|\mathbf{I} + \mathbf{HSH}^H|}{|\mathbf{I} + \mathbf{HKH}^H|}, \log \frac{|\mathbf{I} + \mathbf{GSG}^H|}{|\mathbf{I} + \mathbf{GKG}^H|} \right\}, \\ R_1 &\le \frac{1}{N+L} \left[\log \left| \mathbf{I} + \mathbf{HKH}^H \right| - \log \left| \mathbf{I} + \mathbf{GKG}^H \right| \right] \end{aligned} \tag{10}$$

for some $0 \preceq \mathbf{K} \preceq \mathbf{S}$ with $\mathbf{S}$ denotes the input covariance satisfying $\mathrm{tr}(\mathbf{S}) \le \overline{P}$ and $\mathbf{H}$, $\mathbf{G}$ denotes the channel matrix of receiver 1, 2, respectively. Obviously, when only the confidential message is transmitted to receiver 1, the frequency-selective BCC (1) reduces to the MIMO flat-fading wiretap channel whose secrecy capacity has been characterized in [10–12, 14, 15]. In particular, Bustin et al. derived its closed-form expression under a power-covariance constraint [15]. Under a total power (trace) constraint, the secrecy capacity of the MIMO Gaussian wiretap channel is expressed as [19, Theorem 3]

$$C_s = \frac{1}{N+L} \bigcup_{\mathbf{S} \succeq 0;\ \mathrm{tr}(\mathbf{S}) \le \overline{P}} \sum_{j=1}^{r} \log \phi_j, \tag{11}$$

where $\{\phi_j\}_{j=1}^r$ are the generalized eigen-values greater than one of the following pencil:

$$\left(\mathbf{I} + \mathbf{S}^{1/2} \mathbf{H} \mathbf{H}^H \mathbf{S}^{1/2}, \mathbf{I} + \mathbf{S}^{1/2} \mathbf{G} \mathbf{G}^H \mathbf{S}^{1/2} \right). \tag{12}$$

(In [15, 19] the authors consider the real matrices $\mathbf{H}$, $\mathbf{G}$. Nevertheless, it is conjectured that for complex matrices the following expression without 1/2 in the prelog holds.) As explicitly characterized in [15, Theorem 2], the optimal input covariance achieving the above region is chosen such that the confidential message is sent over r subchannels where receiver 1 observes stronger signals than receiver 2. Moreover, in the high SNR regime the optimal strategy converges to beamforming into the null subspace of $\mathbf{G}$ [5, 11, 12, 14] as for the MISO case [14, 18]. In order to characterize the behavior of the secrecy capacity region in the high SNR regime, we define the d.o.f. region as

$$(r_0, r_1) \triangleq \lim_{P \to \infty} \left(\frac{R_0}{\log P}, \frac{R_1}{\log P} \right), \tag{13}$$

where r_1 denotes s.d.o.f. which corresponds precisely to the number r of the generalized eigenvalues greater than one in the high SNR.

3. Vandermonde Precoding

For the frequency-selective BCC specified in Section 2, we wish to design a practical linear precoding scheme which fully exploits the d.o.f. offered by the frequency-selective channel. We remarked previously that for a special case when only the confidential message is sent to receiver 1 (without a common message), the optimal strategy consists of beamforming the confidential signal into the null subspace of receiver 2. By applying this intuitive result to the special Toeplitz MIMO channels $\mathcal{T}(\mathbf{h})$, $\mathcal{T}(\mathbf{g})$ while including a common message, we propose a linear precoding strategy named *Vandermonde precoding*. Prior to the definition of the Vandermonde precoding, we provide some properties of a Vandermonde matrix [31].

Property 1. Given a full-rank Toeplitz matrix $\mathcal{T}(\mathbf{g}) \in \mathbb{C}^{N \times (N+L)}$, there exists a Vandermonde matrix $\widetilde{\mathbf{V}}_1 \in \mathbb{C}^{(N+L) \times l}$ for $l \le L$ whose structure is given by

$$\widetilde{\mathbf{V}}_1 = \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_l \\ a_1^2 & \cdots & a_l^2 \\ \vdots & \ddots & \vdots \\ a_1^{N+L-1} & \cdots & a_l^{N+L-1} \end{bmatrix}, \tag{14}$$

where $\{a_1, \ldots, a_l\}$ are the $l \le L$ roots of the polynomial $S(z) = \sum_{i=0}^{L} g_i z^{L-i}$ with $L+1$ coefficients of the channel $\mathbf{g}$. Clearly $\widetilde{\mathbf{V}}_1$ satisfies the following orthogonal condition:

$$\mathcal{T}(\mathbf{g}) \widetilde{\mathbf{V}}_1 = 0_{N \times l}, \tag{15}$$

and rank $(\widetilde{\mathbf{V}}_1) = l$ if $a_1, a_2, \ldots, a_l$ are all different.

It is well known that as the dimension of N and L increases, the Vandermonde matrix $\widetilde{\mathbf{V}}_1$ becomes ill-conditioned unless the roots are on the unit circle. In other

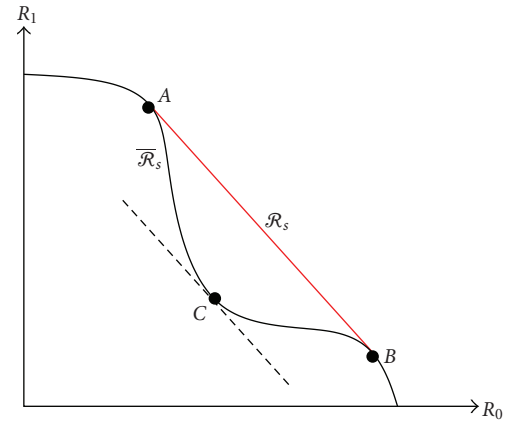

FIGURE 2: Achievable rate region $\mathcal{R}_s$ obtained by the convex hull on $\overline{\mathcal{R}}_s$.

words, the elements of each column either grow in energy or tend to zero [31]. Hence, instead of the brut Vandermonde matrix (14), we consider a unitary Vandermonde matrix obtained either by applying the Gram-Schmidt orthogonalization or singular value decomposition (SVD) on $\mathcal{T}(\mathbf{g})$.

Definition 1. We let $\mathbf{V}_1$ be a unitary Vandermonde matrix obtained by orthogonalizing the columns of $\widetilde{\mathbf{V}}_1$. We let $\mathbf{V}_0 \in \mathbb{C}^{(N+L)\times(N+L-l)}$ be a unitary matrix in the null space of $\mathbf{V}_1$ such that $\mathbf{V}_0^H\mathbf{V}_1 = 0$. The common message W_0, the confidential message W_1, is sent along $\mathbf{V}_0$, $\mathbf{V}_1$, respectively. We call $\mathbf{V} = [\mathbf{V}_0, \mathbf{V}_1] \in \mathbb{C}^{(N+L)\times(N+L)}$ *Vandermonde precoder*.

Further, the precoding matrix $\mathbf{V}_1$ for the confidential message satisfies the following property.

Lemma 2. *Given two Toeplitz matrices $\mathcal{T}(\mathbf{h})$, $\mathcal{T}(\mathbf{g})$ where $\mathbf{h}$, $\mathbf{g}$ are linearly independent, there exists a unitary Vandermonde matrix $\mathbf{V}_1 \in \mathbb{C}^{(N+L)\times l}$ for $0 \le l \le L$ satisfying*

$$\begin{aligned} \mathcal{T}(\mathbf{g})\mathbf{V}_1 &= 0_{N\times l}, \\ \operatorname{rank}(\mathcal{T}(\mathbf{h})\mathbf{V}_1) &= l. \end{aligned} \tag{16}$$

Proof. Appendix A. □

In order to send the confidential message intended to receiver 1 as well as the common message to both receivers over the frequency-selective channel (1), we consider the Gaussian superposition coding based on the Vandermonde precoder of Definition 1. Namely, at block t, we form the transmit vector as

$$\mathbf{x}[t] = \mathbf{V}_0\mathbf{u}_0[t] + \mathbf{V}_1\mathbf{u}_1[t], \tag{17}$$

where the common message vector $\mathbf{u}_0[t]$ and the confidential message vector $\mathbf{u}_1[t]$ are mutually independent Gaussian vectors with zero mean and covariance $\mathbf{S}_0$, $\mathbf{S}_1$, respectively. Under this condition, the input covariances subject to

$$\operatorname{tr}(\mathbf{S}_0) + \operatorname{tr}(\mathbf{S}_1) \le \overline{P} \tag{18}$$

satisfy the power constraint (2). We let $\mathcal{F}$ denote the feasible set $(\mathbf{S}_0, \mathbf{S}_1)$ satisfying (18).

Theorem 3. *The Vandermonde precoding achieves the following secrecy rate region:*

$$\begin{aligned} \mathcal{R}_s = \operatorname{cov} \bigcup_{(\mathbf{S}_0,\mathbf{S}_1)\in\mathcal{F}} \Bigg\{ (R_0, R_1) : R_0 &\le \frac{1}{N+L} \\ &\times \min\Bigg\{ \log \frac{\left|\mathbf{I}_N + \mathbf{H}_0\mathbf{S}_0\mathbf{H}_0^H + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|}{\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|}, \\ &\qquad \log\left|\mathbf{I}_N + \mathbf{G}_0\mathbf{S}_0\mathbf{G}_0^H\right| \Bigg\}, \\ R_1 &\le \frac{1}{N+L}\log\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right| \Bigg\}, \end{aligned} \tag{19}$$

where cov *denotes the convex hull and we let* $\mathbf{H}_0 = \mathcal{T}(\mathbf{h})\mathbf{V}_0$, $\mathbf{H}_1 = \mathcal{T}(\mathbf{h})\mathbf{V}_1$, $\mathbf{G}_0 = \mathcal{T}(\mathbf{g})\mathbf{V}_0$.

Proof. Due to the orthogonal property (16) of the unitary Vandermonde matrix, receiver 2 only observes the common message, which yields the received signals given by

$$\begin{aligned} \mathbf{y} &= \mathcal{T}(\mathbf{h})\mathbf{V}_0\mathbf{u}_0 + \mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{u}_1 + \mathbf{n}, \\ \mathbf{z} &= \mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{u}_0 + \boldsymbol{\nu}, \end{aligned} \tag{20}$$

where we drop the block index. We examine the achievable rate region $\mathcal{R}_s$ of the Vandermonde precoding. By letting the auxiliary variables $U = \mathbf{V}_0\mathbf{u}_0$, $V = U + \mathbf{V}_1\mathbf{u}_1$ and $X = V$, we have

$$\begin{aligned} &I(U;Y) \\ &= \frac{1}{N+L} \\ &\quad\times \log \frac{\left|\mathbf{I}_N + \mathcal{T}(\mathbf{h})\mathbf{V}_0\mathbf{S}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{h})^H + \mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{S}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right|}{\left|\mathbf{I}_N + \mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{S}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right|}, \\ &I(U;Z) = \frac{1}{N+L}\log\left|\mathbf{I}_N + \mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{S}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{g})^H\right|, \\ &I(V;Y \mid U) = \frac{1}{N+L}\log\left|\mathbf{I}_N + \mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{S}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right|, \\ &I(V;Z \mid U) = 0. \end{aligned} \tag{21}$$

Plugging these expressions to (8), we obtain (19). □

The boundary of the achievable rate region of the Vandermonde precoding can be characterized by solving the weighted sum rate maximization. Any point $(R_0^\star, R_1^\star)$ on the boundary of the convex region $\mathcal{R}_s$ is obtained by solving

$$\max_{(R_0,R_1)\in\mathcal{R}_s} \gamma_0 R_0 + \gamma_1 R_1 \tag{22}$$

for nonnegative weights γ_0, γ_1 satisfying $\gamma_0 + \gamma_1 = 1$. When the region $\overline{\mathcal{R}_s}$, obtained without convex hull, is nonconvex, the set of the optimal covariances $(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$ achieving the boundary point might not be unique. Figure 2 depicts an example in which the achievable rate region $\mathcal{R}_s$ is obtained by the convex hull operation on the region $\overline{\mathcal{R}_s}$, that is, replacing the non-convex subregion by the line segment A, B. For the weight ratio γ_1/γ_0 corresponding to the slope of the line segment A, B, there exist two optimal sets of the covariances yielding the points A and B (which clearly dominate the point C). These points are the solution to the weighted sum rate maximization (22). In summary, an optimal covariance set achieving (22) (might not be unique) is the solution of

$$\max_{(\mathbf{S}_0,\mathbf{S}_1)\in\mathcal{F}} \gamma_0 R_0 + \gamma_1 R_1 = \max_{(\mathbf{S}_0,\mathbf{S}_1)\in\mathcal{F}} \gamma_0 \min\{R_{01}, R_{02}\} + \gamma_1 R_1, \tag{23}$$

where we let

$$\begin{aligned} R_{01}(\mathbf{S}_0,\mathbf{S}_1) &= \frac{1}{N+L}\log\frac{\left|\mathbf{I}_N + \mathbf{H}_0\mathbf{S}_0\mathbf{H}_0^H + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|}{\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|}, \\ R_{02}(\mathbf{S}_0) &= \frac{1}{N+L}\log\left|\mathbf{I}_N + \mathbf{G}_0\mathbf{S}_0\mathbf{G}_0^H\right|, \\ R_1(\mathbf{S}_1) &= \frac{1}{N+L}\log\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|. \end{aligned} \tag{24}$$

Following [34, Section II-C] (and also [7, Lemma 2]), we remark that the solution to the max-min problem (23) can be found by hypothesis testing of three cases, $R_{01} < R_{02}$, $R_{02} < R_{01}$, and $R_{01} = R_{02}$. Formally, we have the following lemma.

Lemma 4. *The optimal* $(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$*, solution of* (23)*, is given by one of the three solutions.*

Case 1. $(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$ maximizes

$$\begin{aligned} f_1(\mathbf{S}_0,\mathbf{S}_1) = \gamma_0 \log\frac{\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H + \mathbf{H}_0\mathbf{S}_0\mathbf{H}_0^H\right|}{\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|} \\ + \gamma_1 \log\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right| \end{aligned} \tag{25}$$

and satisfies $R_{01}(\mathbf{S}_0^\star, \mathbf{S}_1^\star) < R_{02}(\mathbf{S}_1^\star)$.

Case 2. $(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$ maximizes

$$f_2(\mathbf{S}_0,\mathbf{S}_1) = \gamma_0 \log\left|\mathbf{I}_N + \mathbf{G}_0\mathbf{S}_0\mathbf{G}_0^H\right| + \gamma_1 \log\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right| \tag{26}$$

and satisfies $R_{02}(\mathbf{S}_1^\star) < R_{01}(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$.

Case 3. $(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$ maximizes

$$\begin{aligned} f_3(\mathbf{S}_0,\mathbf{S}_1) = \gamma_0 \Bigg[\theta \log\frac{\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H + \mathbf{H}_0\mathbf{S}_0\mathbf{H}_0^H\right|}{\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right|} \\ + (1-\theta)\log\left|\mathbf{I}_N + \mathbf{G}_0\mathbf{S}_0\mathbf{G}_0^H\right| \Bigg] \\ + \gamma_1 \log\left|\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right| \end{aligned} \tag{27}$$

and satisfies $R_{01}(\mathbf{S}_0^\star, \mathbf{S}_1^\star) = R_{02}(\mathbf{S}_1^\star)$ for some $0 < \theta < 1$.

Before considering the weighted sum rate maximization (23)*, one applies SVD to* $\mathbf{H}_1 \in \mathbb{C}^{N\times l}$, $\mathbf{G}_0 \in \mathbb{C}^{N\times(N+L-l)}$

$$\begin{aligned} \mathbf{H}_1 &= \mathbf{U}_{h1}\mathbf{\Lambda}_{h1}{\mathbf{V}_{h1}}^H, \\ \mathbf{G}_0 &= \mathbf{U}_{g0}\mathbf{\Lambda}_{g0}{\mathbf{V}_{g0}}^H, \end{aligned} \tag{28}$$

where $\mathbf{U}_{h1}, \mathbf{U}_{g0} \in \mathbb{C}^{N\times N}$, $\mathbf{V}_{h1} \in \mathbb{C}^{l\times l}$, *and* $\mathbf{V}_{g0} \in \mathbb{C}^{(N+L-l)\times(N+L-l)}$ *are unitary,* $\mathbf{\Lambda}_{h1}, \mathbf{\Lambda}_{g0}$ *contain positive singular values* $\{\sqrt{\lambda_i^{h1}}\}_{i=1}^{l}$, $\{\sqrt{\lambda_i^{g0}}\}_{i=1}^{N+L-l}$, *respectively. Following [7, Theorem 3], one applies Lemma 4 to solve the weighted sum rate maximization.*

Theorem 5. *The set of the optimal covariances* $(\mathbf{S}_0^\star, \mathbf{S}_1^\star)$*, achieving the boundary of the achievable rate region* $\mathcal{R}_s$ *of the Vandermonde precoding, corresponds to one of the following three solutions.*

Case 1. $(\mathbf{S}_0^\star, \mathbf{S}_1^\star) = (\mathbf{S}_0^1, \mathbf{S}_1^1)$, if $(\mathbf{S}_0^1, \mathbf{S}_1^1)$, solution of the following KKT conditions, satisfies $R_{01}(\mathbf{S}_0^1, \mathbf{S}_1^1) < R_{02}(\mathbf{S}_1^1)$

$$\begin{aligned} \gamma_0\mathbf{H}_0^H\mathbf{\Gamma}^{-1}\mathbf{H}_0 + \mathbf{\Psi}_0 &= \mu\mathbf{I}_{N+L-l}, \\ \gamma_0\mathbf{H}_1^H\mathbf{\Gamma}^{-1}\mathbf{H}_1 + (\gamma_1-\gamma_0)\mathbf{H}_1^H(\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H)^{-1}\mathbf{H}_1 + \mathbf{\Psi}_1 &= \mu\mathbf{I}_l, \end{aligned} \tag{29}$$

where $\mathrm{tr}(\mathbf{\Psi}_i\mathbf{S}_i) = 0$ with a positive semidefinite $\mathbf{\Psi}_i$ for $i = 0, 1$, $\mu \geq 0$ is determined such that $\mathrm{tr}(\mathbf{S}_0) + \mathrm{tr}(\mathbf{S}_1) = \overline{P}$, and we let $\mathbf{\Gamma} = \mathbf{I}_N + \mathbf{H}_0\mathbf{S}_0\mathbf{H}_0^H + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H$.

Case 2. $(\mathbf{S}_0^\star, \mathbf{S}_1^\star) = (\mathbf{S}_0^2, \mathbf{S}_1^2)$ if the following $(\mathbf{S}_0^2, \mathbf{S}_1^2)$ fulfills $R_{02}(\mathbf{S}_1^2) < R_{01}(\mathbf{S}_0^2, \mathbf{S}_1^2)$.

We let $\mathbf{S}_0^2 = \mathbf{V}_{g0}\widehat{\mathbf{S}}_0\mathbf{V}_{g0}^H$ and $\mathbf{S}_1^2 = \mathbf{V}_{h1}\widehat{\mathbf{S}}_1\mathbf{V}_{h1}^H$ where $\widehat{\mathbf{S}}_0, \widehat{\mathbf{S}}_1$ are diagonal with the ith element given by

$$\begin{aligned} p_{0,i} &= \left[\frac{\gamma_0}{\mu} - \frac{1}{\lambda_i^{g0}}\right]_+, \quad i = 1,\ldots,N+L-l, \\ p_{1,i} &= \left[\frac{\gamma_1}{\mu} - \frac{1}{\lambda_i^{h1}}\right]_+, \quad i = 1,\ldots,l, \end{aligned} \tag{30}$$

where $\mu \geq 0$ is determined such that $\sum_{i=1}^{N+L-l} p_{0,i} + \sum_{i=1}^{l} p_{1,i} = \overline{P}$.

Case 3. $(\mathbf{S}_0^\star, \mathbf{S}_1^\star) = (\mathbf{S}_0^3, \mathbf{S}_1^3)$, if $(\mathbf{S}_0^3, \mathbf{S}_1^3)$, solution of the following KKT conditions, satisfies $R_{02}^\theta(\mathbf{S}_1^3) = R_{01}^\theta(\mathbf{S}_0^3, \mathbf{S}_1^3)$ for some $0 < \theta < 1$

$$\begin{aligned}
&\theta \mathbf{H}_0^H \mathbf{\Gamma}^{-1} \mathbf{H}_0 + (1-\theta)\mathbf{G}_0^H \left(\mathbf{I}_N + \mathbf{G}_0 \mathbf{S}_0 \mathbf{G}_0^H\right)^{-1} \mathbf{G}_0 + \mathbf{\Psi}_0 \\
&\quad = \mu \mathbf{I}_{N+L-l}, \\
&\gamma_0 \theta \mathbf{H}_1^H \mathbf{\Gamma}^{-1} \mathbf{H}_1 + (\gamma_1 - \gamma_0\theta)\mathbf{H}_1^H \left(\mathbf{I}_N + \mathbf{H}_1 \mathbf{S}_1 \mathbf{H}_1^H\right)^{-1} \mathbf{H}_1 + \mathbf{\Psi}_1 \\
&\quad = \mu \mathbf{I}_l,
\end{aligned} \tag{31}$$

where $\mathrm{tr}(\mathbf{\Psi}_i \mathbf{S}_i) = 0$ with a positive semidefinite $\mathbf{\Psi}_i$ for $i = 0, 1$, $\mu \geq 0$ is determined such that $\mathrm{tr}(\mathbf{S}_0) + \mathrm{tr}(\mathbf{S}_1) = \overline{P}$.

Proof. Appendix B. □

Remark 6. Due to the non-concavity of the underlying weighted sum rate functions, it is generally difficult to characterize the boundary of the achievable rate region $\mathcal{R}_s$ except for some special cases. The special cases include the corner points, in particular, the secrecy rate for the case of sending only the confidential message ($\gamma_1 = 1$), as well as the maximum sum rate point for the equal weight case ($\gamma_0 = \gamma_1$). It is worth noticing that under equal weight the objective functions in three cases are all concave in $\mathbf{S}_0$, $\mathbf{S}_1$ since f_1 is concave if $\gamma_1 \geq \gamma_0$ and f_3 is concave if $\gamma_1 \geq \gamma_0\theta$ and $0 < \theta < 1$.

The maximum sum rate point $\gamma_0 = \gamma_1$ can be found by applying the following greedy search [7].

Greedy Search to Find the Maximum Sum Rate Point. (1) Find $\mathbf{S}_0$, $\mathbf{S}_1$ maximizing f_1 and check $R_{02} < R_{01}$. If yes stop. Otherwise go to (2).

(2) Find $\mathbf{S}_0$, $\mathbf{S}_1$ maximizing f_2 and check $R_{01} < R_{01}$. If yes stop. Otherwise go to (3).

(3) Find $\mathbf{S}_0$, $\mathbf{S}_1$ maximizing f_3 and check $R_{01}^\theta = R_{01}^\theta$ for some $0 < \theta < 1$.

For the special case of $\gamma_1 = 1$, Theorem 5 yields the achievable secrecy rate with the Vandermonde precoding.

Corollary 7. *The Vandermonde precoding achieves the secrecy rate*

$$\begin{aligned}
R_1^{\mathrm{vdm}} &= \max_{\mathbf{S}_1 : \mathrm{tr}(\mathbf{S}_1) \leq \overline{P}} \frac{1}{N+L} \log \det\left(\mathbf{I}_N + \mathcal{T}(\mathbf{h}) \mathbf{V}_1 \mathbf{S}_1 \mathbf{V}_1^H \mathcal{T}(\mathbf{h})^H\right) \\
&= \frac{1}{N+L} \sum_{i=1}^{L} \log\left(\mu \lambda_i^{h1}\right)_+,
\end{aligned} \tag{32}$$

where the last equality is obtained by applying SVD to $\mathbf{H}_1 = \mathcal{T}(\mathbf{h})\mathbf{V}_1$ *and plugging the power allocation of* (30) *with* $\gamma_0 = 0$, $\gamma_1 = 1$, μ *is determined such that* $\sum_{l=1}^{L} p_{1i} \leq \overline{P}$.

Finally, by focusing the behavior of the achievable rate region in the high SNR regime, we characterize the achievable d.o.f. region of the frequency-selective BCC (1).

Theorem 8. *The d.o.f. region of the frequency-selective BCC* (1) *with* $(N+L) \times L$ *Toeplitz matrices* $\mathcal{T}(\mathbf{h})$, $\mathcal{T}(\mathbf{g})$ *is given as a union of* $(r_0, r_1) = (1/(N+L))(l_0, l)$ *satisfying*

$$l \leq L, \tag{33}$$

$$l_0 + l \leq N, \tag{34}$$

where l_0, l *denote non-negative integers. The Vandermonde precoding achieves the above d.o.f. region.*

Proof. The achievability follows rather trivially by applying Theorem 3. By considering equal power allocation over all $N+L$ streams such that $\mathbf{S}_0 = P\mathbf{I}_{N+L-l}$, $\mathbf{S}_1 = P\mathbf{I}_l$, we obtain the rate tuple (R_0, R_1) where $R_0 \leq \min(R_{01}, R_{02})$

$$\begin{aligned}
&R_{01} \\
&= \frac{1}{N+L} \\
&\quad \times \log \frac{\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h})\mathbf{V}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{h})^H + P\mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right|}{\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right|}, \\
&R_{02} = \frac{1}{N+L} \log\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{g})^H\right|, \\
&R_1 \leq \frac{1}{N+L} \log\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right|.
\end{aligned} \tag{35}$$

We first notice that the prelog factor of $\log|\mathbf{I} + P\mathbf{A}|$ as $P \to \infty$ depends only on the rank of $\mathbf{A}$. From Lemma 2, we obtain

$$\begin{aligned}
&\mathrm{rank}\left(\mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right) \\
&\quad = \mathrm{rank}(\mathcal{T}(\mathbf{h})\mathbf{V}_1) = l,
\end{aligned} \tag{36}$$

$$\begin{aligned}
&\mathrm{rank}\left(\mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{g})^H\right) \\
&\quad = \mathrm{rank}(\mathcal{T}(\mathbf{g})\mathbf{V}_0) \\
&\quad \overset{(a)}{=} \mathrm{rank}(\mathcal{T}(\mathbf{g})[\mathbf{V}_0\mathbf{V}_1]) \\
&\quad \overset{(b)}{=} \mathrm{rank}(\mathcal{T}(\mathbf{g})) = N,
\end{aligned} \tag{37}$$

$$\begin{aligned}
&\mathrm{rank}\left(\mathcal{T}(\mathbf{h})\left(\mathbf{V}_0\mathbf{V}_0^H + \mathbf{V}_1\mathbf{V}_1^H\right)\mathcal{T}(\mathbf{h})^H\right) \\
&\quad = \mathrm{rank}\left(\mathcal{T}(\mathbf{h})\mathbf{V}\mathbf{V}^H\mathcal{T}(\mathbf{h})^H\right) \\
&\quad \overset{(b)}{=} \mathrm{rank}(\mathcal{T}(\mathbf{h})) = N,
\end{aligned} \tag{38}$$

where (a) follows from orthogonality between $\mathcal{T}(\mathbf{g})$ and $\mathbf{V}_1$, (b) follows from the fact that $\mathbf{V} = [\mathbf{V}_0\mathbf{V}_1]$ is unitary satisfying $\mathbf{V}\mathbf{V}^H = \mathbf{I}$. Notice that (36) yields $r_1 = l/(N+L)$. For the d.o.f. $r_0 = l_0/(N+L)$ of the common message, (36) and (38) yield

$$\begin{aligned}
l_0 &= \mathrm{rank}\left(\mathcal{T}(\mathbf{h})\left(\mathbf{V}_0\mathbf{V}_0^H + \mathbf{V}_1\mathbf{V}_1^H\right)\mathcal{T}(\mathbf{h})^H\right) \\
&\quad - \mathrm{rank}\left(\mathcal{T}(\mathbf{h})\mathbf{V}_1\mathbf{V}_1^H\mathcal{T}(\mathbf{h})^H\right) \\
&= N - l
\end{aligned} \tag{39}$$

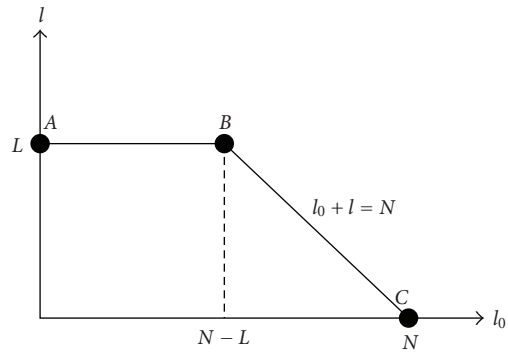

Figure 3: d.o.f. region (l_0, l_1) of frequency-selective BCC.

which is dominated by the pre-log of R_{02} in (37). This establishes the achievability.

The converse follows by noticing that the inequalities (33) and (34) correspond to trivial upper bounds. The first inequality (33) corresponds to the s.d.o.f. of the MIMO wiretap channel with the legitimate channel $\mathcal{T}(\mathbf{h})$ and the eavesdropper channel $\mathcal{T}(\mathbf{g})$, which is bounded by L. The second inequality (34) follows because the total number of streams for receiver 1 cannot be larger than the d.o.f. of $\mathcal{T}(\mathbf{h})$, that is, N. □

Figure 3 illustrates the region (l, l_0) of the frequency-selective BCC over $N + L$ dimensions. We notice that the s.d.o.f. constraint (33) yields the line segment A, B while the constraint (34) in terms of the total number of streams for receiver 1 yields the line segment B, C.

4. Multiuser Secure Communications

In this section, we provide some applications of the Vandermonde precoding in the multi-user secure communication scenarios where the transmitter wishes to send confidential messages to more than one intended receivers. The scenarios that we address are: (a) a $K+1$-user frequency-selective BCC with K confidential messages and one common message, (b) a two-user frequency-selective BCC with two confidential messages and one common message. For each scenario, by focusing on the behavior in the high SNR regime, we characterize the achievable s.d.o.f. region and show the optimality of the Vandermonde precoding.

4.1. $K + 1$-User BCC with K Confidential Messages. As an extension of Section 3, we consider the $K+1$-user frequency-selective BCC where the transmitter sends $K \leq L$ confidential messages $W_1, \ldots, W_K$ to the first K receivers as well as one common message W_0 to all receivers. Each of the confidential messages must be kept secret to receiver $K + 1$. Notice that this model, called multireceiver wiretap channel, has been studied in the literature ([20, 22–26] and reference therein). In particular, the secrecy capacity region of the Gaussian MIMO multireceiver wiretap channel has been characterized in [24, 26] for $K = 2$, an arbitrary K, respectively, where the optimality of the S-DPC is proved.

The received signal $\mathbf{y}_k$ of receiver k and the received signal $\mathbf{z}$ of receiver $K + 1$ at any block are given by

$$\mathbf{y}_k = \mathcal{T}(\mathbf{h}_k)\mathbf{x} + \mathbf{n}_k, \quad k = 1, \ldots, K, \tag{40}$$

$$\mathbf{z} = \mathcal{T}(\mathbf{g})\mathbf{x} + \boldsymbol{\nu}, \tag{41}$$

where $\mathbf{x}$ is the transmit vector satisfying the total power constraint and $\mathbf{n}_1, \ldots, \mathbf{n}_K$, $\boldsymbol{\nu}$ are mutually independent AWGN with covariance $\mathbf{I}$. We assume that the $K + 1$ vectors $\mathbf{h}_1, \ldots, \mathbf{h}_K$, $\mathbf{g}$ of length $L + 1$ are linearly independent and perfectly known to all the terminals. As an extension of the frequency-selective BCC in Section 2, we say that the rate tuple $(R_0, R_1, \ldots, R_K)$ is achievable if for any $\epsilon > 0$ there exists a sequence of codes $(2^{nR_0}, 2^{nR_1}, \ldots, 2^{nR_K}, n)$ such that

$$\begin{aligned} &P_e^n \leq \epsilon, \\ &\sum_{k \in \mathcal{K}} R_k - \frac{1}{n} H(W_{\mathcal{K}} \mid Z^n) \leq \epsilon, \quad \mathcal{K} \subseteq \{1, \ldots, K\}, \end{aligned} \tag{42}$$

where we denote $W_{\mathcal{K}} = \{\forall k \in \mathcal{K}, W_k\}$ and define

$$\begin{aligned} P_e^n = &\frac{1}{\prod_{k=0}^{K} 2^{nR_k}} \\ &\times \sum_{w_0 \in \mathcal{W}_0} \cdots \sum_{w_K \in \mathcal{W}_K} \Pr\left(\bigcup_{k=1}^{K} \left(\widehat{w}_0^{(k)}, \widehat{w}_k \right) \neq (w_0, w_k) \right). \end{aligned} \tag{43}$$

An achievable secrecy rate region (R_1, R_2) for the case of $K = 2$, when the transmitter sends two confidential messages in the presence of an external eavesdropper, is provided in [25, Theorem 1]. This theorem can be extended to an arbitrary K while including the common message. Formally we state the following lemma.

Lemma 9. *An achievable rate region of the $K+1$-user BCC, where the transmitter sends K confidential messages intended to the first K receivers as well as a common message to all users, is given as a union of all non-negative rate-tuple satisfying*

$$\begin{aligned} R_0 &\leq \min\left\{ I(U; Z), \min_k I(U; Y_k) \right\}, \\ R_k &\leq I(V_k; Y_k \mid U) - I(V_k; Z \mid U), \quad k = 1, \ldots, K, \\ \sum_{k \in \mathcal{K}} R_k &\leq \sum_{k \in \mathcal{K}} I(V_k; Y_k \mid U) \\ &\quad - \sum_{j=2}^{|\mathcal{K}|} I\left(V_{\pi(j)}; V_{\pi(1)}, \ldots, V_{\pi(j-1)} \mid U \right) \\ &\quad - I(V_{\mathcal{K}}; Z \mid U), \quad \forall \mathcal{K} \subseteq \{1, \ldots, K\}, \ \forall \pi, \end{aligned} \tag{44}$$

where π denotes a permutation over the subset $\mathcal{K}$, $|\mathcal{K}|$ denotes the cardinality of $\mathcal{K}$, we let $V_{\mathcal{K}} = \{\forall k \in \mathcal{K}, V_k\}$, and the random variables U, $V_1, \ldots, V_K$, X, $Y_1, \ldots, Y_K$, Z satisfy the Markov chain

$$U, V_1, \ldots, V_K \longrightarrow X \longrightarrow Y_1, \ldots, Y_K, Z. \tag{45}$$

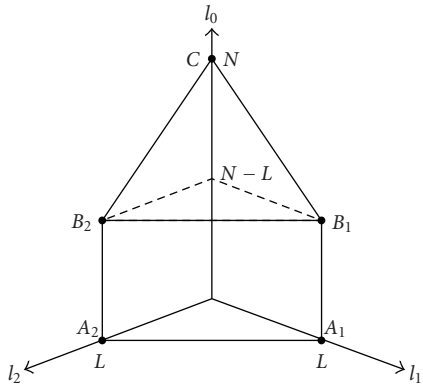

Figure 4: s.d.o.f. region (l_0, l_1, l_2) over $N + L$ dimensions of three-user frequency-selective BCC.

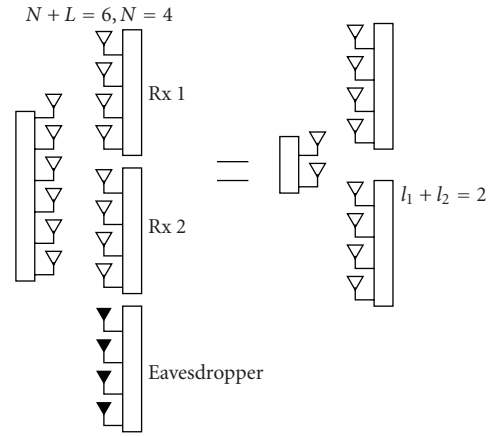

Figure 5: Equivalent MIMO interpretation for three-user frequency-selective BCC with two confidential messages.

Proof. Appendix C. □

Notice that the second term of the last equation in (44) can be also expressed by

$$\begin{aligned}\sum_{j=2}^{|\mathcal{K}|} I\left(V_{\pi(j)}; V_{\pi(1)}, \ldots, V_{\pi(j-1)} \mid U\right) &= \sum_{k\in\mathcal{K}} H(V_k \mid U) - H(V_{\mathcal{K}} \mid U), \\ &\forall \mathcal{K} \subseteq \{1, \ldots, K\}, \ \forall \pi.\end{aligned} \tag{46}$$

It can be easily seen that without the secrecy constraint the above region reduces to the Marton's achievable region for the general K-user broadcast channel [35].

In order to focus on the behavior of the region in the high SNR regime, we define the s.d.o.f. region as

$$r_0 = \lim_{P\to\infty} \frac{R_0}{\log P}, \qquad r_k = \lim_{P\to\infty} \frac{R_k}{\log P}, \quad k = 1, \ldots, K, \tag{47}$$

where r_0 denotes the d.o.f. of the common message and r_k denotes the s.d.o.f. of confidential message k. As an extension of Theorem 8, we have the following s.d.o.f. region result.

Theorem 10. *The s.d.o.f. region of the $K + 1$-user frequency-selective BCC (40) is a union of $(r_0, r_1, \ldots, r_K) = (1/(N + L))(l_0, l_1, \ldots, l_K)$ satisfying*

$$\sum_{k=1}^{K} l_k \le L, \tag{48}$$

$$l_0 + \sum_{k=1}^{K} l_k \le N, \tag{49}$$

where $\{l_0, l_1, \ldots, l_K\}$ are non-negative integers. The Vandermonde precoding achieves this region.

Proof. Appendix D. □

Figure 4 illustrates the region (l_0, l_1, l_2) for the case of $K = 2$ confidential messages. It can be easily seen that the constraint (49) in terms of the total number of streams for the virtual receiver yields the subspace C, B_1, B_2 while the s.d.o.f. constraint (48) for the virtual receiver yields the subspace A_1, A_2, B_2, B_1. We remark that for the special case of one confidential message and one common message ($K = 1$), the region reduces to Figure 3.

Remark 11. When only the K confidential messages are transmitted to the K intended receivers in the presence of the eavesdropper, the s.d.o.f. region has the equivalent MIMO interpretation [36]. More specifically, the frequency-selective BCC (40) is equivalent to the MIMO-BCC where the transmitter with $N + L$ dimensions (antennas) sends messages to K receivers with N antennas each in the presence of the eavesdropper with N antennas. The secrecy constraint (orthogonal constraint) consumes N dimensions of the channel seen by the virtual receiver and lets the number of effective transmit antennas be L. The resulting channel is the MIMO-BC without secrecy constraint with L transmit antennas and K receivers with N antennas each, whose multiplexing gain is $\min(L, KN) = L$ (we assume $L < N$). Figure 5 illustrates the example with $K = 2, N = 4, L = 2$.

4.2. Two-User BCC with Two Confidential Messages. We consider the two-user BCC where the transmitter sends two confidential messages W_1, W_2 as well as one common message W_0. Each of the confidential messages must be kept secret to the unintended receiver. This model has been studied in [17–19] for the case of two confidential messages and in [20] for the case of two confidential messages and a common message. In [19], the secrecy capacity region of the MIMO Gaussian BCC was characterized. The received signal at receivers 1, 2 at any block is given, respectively, by

$$\begin{aligned}\mathbf{y}_1 &= \mathcal{T}(\mathbf{h}_1)\mathbf{x} + \mathbf{n}_1, \\ \mathbf{y}_2 &= \mathcal{T}(\mathbf{h}_2)\mathbf{x} + \mathbf{n}_2,\end{aligned} \tag{50}$$

where $\mathbf{x}$ is the input vector satisfying the total power constraint and $\mathbf{n}_1$, $\mathbf{n}_2$ are mutually independent AWGN with

covariance $\mathbf{I}_N$. We assume the channel vectors $\mathbf{h}_1$, $\mathbf{h}_2$ are linearly independent.

We say that the rate tuple (R_0, R_1, R_2) is achievable if for any $\epsilon > 0$ there exists a sequence of codes $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ such that

$$P_e^n \le \epsilon,$$
$$R_1 - \frac{1}{n}H(W_1 \mid Y_2^n) \le \epsilon, \qquad R_2 - \frac{1}{n}H(W_2 \mid Y_1^n) \le \epsilon, \tag{51}$$

where we define the average error probability as

$$P_e^n = \frac{1}{\prod_{k=0}^{2} 2^{nR_k}} \times \sum_{w_0 \in \mathcal{W}_0} \sum_{w_1 \in \mathcal{W}_1} \sum_{w_2 \in \mathcal{W}_2} \Pr\left(\left(\widehat{w}_0^{(1)}, \widehat{w}_1\right) \ne (w_0, w_1) \cup \left(\widehat{w}_0^{(2)}, \widehat{w}_2\right) \ne (w_0, w_2)\right), \tag{52}$$

where $(\widehat{w}_0^{(1)}, \widehat{w}_1), (\widehat{w}_0^{(2)}, \widehat{w}_2)$ is the output of decoders 1, 2, respectively. A secrecy achievable rate region of the two-user BCC with two confidential messages and a common message is given by [20, Theorem 1]

$$\begin{aligned} R_0 &\le \min\{I(U; Y_1), I(U; Y_2)\}, \\ R_1 &\le I(V_1; Y_1 \mid U) - I(V_1; Y_2, V_2 \mid U), \\ R_2 &\le I(V_2; Y_2 \mid U) - I(V_2; Y_1, V_1 \mid U), \end{aligned} \tag{53}$$

where the random variables satisfy the Markov chain

$$U, V_1, V_2 \longrightarrow X \longrightarrow Y_1, Y_2. \tag{54}$$

We extend Theorem 8 to the two-user frequency-selective BCC (50) and obtain the following s.d.o.f. result.

Theorem 12. *The s.d.o.f. region of the two-user frequency-selective BCC* (50) *is a union of* $(r_0, r_1, r_2) = (1/(N + L))(l_0, l_1, l_2)$ *satisfying*

$$l_k \le L, \quad k = 1, 2, \tag{55}$$
$$l_0 + l_k \le N, \quad k = 1, 2, \tag{56}$$

where $\{l_0, l_1, l_2\}$ *are non-negative integers. The Vandermonde precoding achieves the region.*

Proof. Appendix F. □

Figure 6 represents the s.d.o.f. region (l_0, l_1, l_2) over $N + L$ dimensions of the two-user frequency-selective BCC. The per-receiver s.d.o.f. constraints (55) yield the subspace A_1, B_1, E, F for user 1 and the subspace A_2, B_2, E, F for user 2. The constraints (56) in terms of the total number of streams per receiver yield the subregion C, B_1, E for user 1 and the subregion C, B_2, E for user 2. For the special case of one confidential message and one common message, the region reduces to Figure 3.

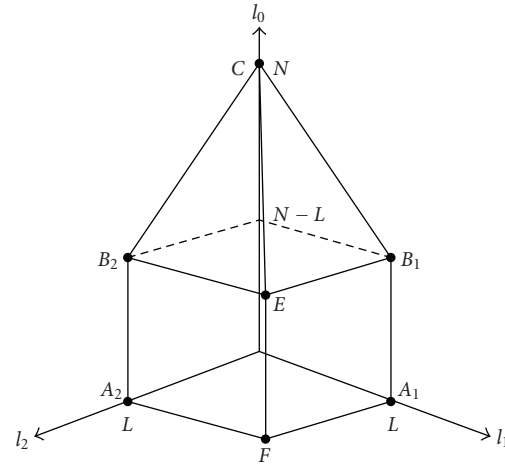

Figure 6: s.d.o.f. region (l_0, l_1, l_2) over $N + L$ dimensions of $K = 2$-user frequency-selective BCC.

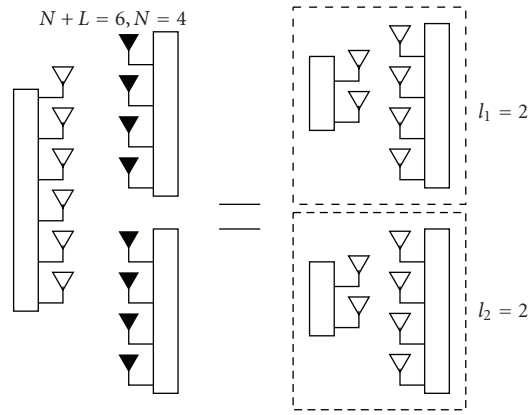

Figure 7: Equivalent MIMO interpretation for the two-user frequency-selective BCC with two confidential messages.

Remark 13. Comparing Theorems 10, 12 as well as Figures 4, 6 for $K = 2$, it clearly appears that the s.d.o.f. of $K + 1$-user BCC with K confidential messages is dominated by the s.d.o.f. of K-user BCC with K confidential messages. In other words, the s.d.o.f. region critically depends on the assumption on the eavesdropper(s) to whom each confidential message must be kept secret.

Remark 14. When only two confidential messages are transmitted in the two-user frequency-selective BCC, the set of the s.d.o.f. has the equivalent MIMO interpretation [36]. More specifically, the frequency-selective BCC (40) is equivalent to the MIMO-BCC where the transmitter with $N+L$ dimensions (antennas) sends two confidential messages to two receivers with N antennas. The secrecy constraint consumes N dimensions for each MIMO link and lets the number of effective transmit antennas be L for each user. The resulting channel is a two parallel $L \times N$ point-to-point MIMO channel without eavesdropper. Notice that the same parallel MIMO links can be obtained by applying the block

diagonalization on the MIMO-BC without secrecy constraint [36]. In other words, the secrecy constraint in the BCC with inner eavesdroppers is equivalent to the orthogonal constraint in the classical MIMO-BC. Figure 7 shows the example with $N = 4$, $L = 2$, and $K = 2$ confidential messages.

5. Numerical Examples

In order to examine the performance of the proposed Vandermonde precoding, this section provides some numerical results in different settings.

5.1. Secrecy Rate versus SNR. We evaluate the achievable secrecy rate R_1^{vdm} in (32) when the transmitter sends only a confidential message to receiver 1 (without a common message) in the presence of receiver 2 (eavesdropper) over the frequency-selective BCC studied in Section 3.

5.1.1. MISO Wiretap Channel. For the sake of comparison (albeit unrealistic), we consider the special case of the frequency-selective wiretap channel when receiver 1 has a scalar observation and the eavesdropper has N observations. This is equivalent to the MISO wiretap channel with the receiver 1 channel $\mathbf{h} \in \mathbb{C}^{1\times(N+L)}$ and the eavesdropper channel $\mathcal{T}(\mathbf{g}) \in \mathbb{C}^{N\times(N+L)}$. Without loss of generality, we assume that the observation at receiver 1 is the first row of $\mathcal{T}(\mathbf{h})$. We consider that all entries of $\mathbf{h}$, $\mathbf{g}$ are i.i.d. $\sim \mathcal{N}_{\mathcal{C}}(0, 1/(L+1))$ and average the secrecy rate over a large number of randomly generated channels with $N = 64$, $L = 16$. In Figure 8, we compare the optimal beamforming strategy [10, 13, 14] and the Vandermonde precoding as a function of SNR P. Since only one stream is sent to receiver 1, the s.d.o.f. is $1/(N+L)$. In fact, the MISO secrecy capacity in the high SNR regime is given by

$$\frac{1}{N+L}\log\left(1+(N+L)P\max_{\boldsymbol{\phi}:\mathcal{T}(\mathbf{g})\boldsymbol{\phi}=0}\left|\mathbf{h}\boldsymbol{\phi}\right|^2\right), \tag{57}$$

where $\boldsymbol{\phi} \in \mathbb{C}^{(N+L)\times 1}$ is the beamforming vector. The Vandermonde precoding achieves

$$\frac{1}{N+L}\log\left(1+(N+L)P\max_{i=1,\ldots,L}\left|\mathbf{h}\mathbf{v}_{1,i}\right|^2\right), \tag{58}$$

where $\mathbf{v}_{1,i}$ denotes the ith column of $\mathbf{V}_1 \in \mathbb{C}^{(N+L)\times L}$ orthogonal to $\mathcal{T}(\mathbf{g})$. Clearly, there exists a constant gap between (57) and (58) due to the suboptimal choice of the beamforming vector.

5.1.2. MIMO Wiretap Channel. We consider the frequency-selective wiretap channel with $N = 64$, $L = 16$. Although there exists a closed-form expression under a power-covariance constraint [15], the secrecy capacity under a total power constraint in (11) is still difficult to compute (especially for a large dimension of N and L) because it requires a search over all possible power covariances constraints. Therefore, in Figure 9, we compare the averaged secrecy rate achieved by the generalized SVD scheme [5] and the Vandermonde precoding. We assume that all entries of $\mathbf{h}, \mathbf{g}$ are i.i.d. $\sim \mathcal{N}_{\mathcal{C}}(0, 1/(L+1))$. For the Vandermonde precoding, we show the achievable rate with waterfilling power allocation (32) and equal power allocation (36) by allocating $p = (N+L)P/L$ to L streams. As observed, these two suboptimal schemes achieve the same s.d.o.f. of $L/(N+L) = 1/5$ although the generalized SVD incurs a substantial power loss. The result agrees well with Theorem 8. We remark also that the optimal waterfilling power allocation yields a negligible gain.

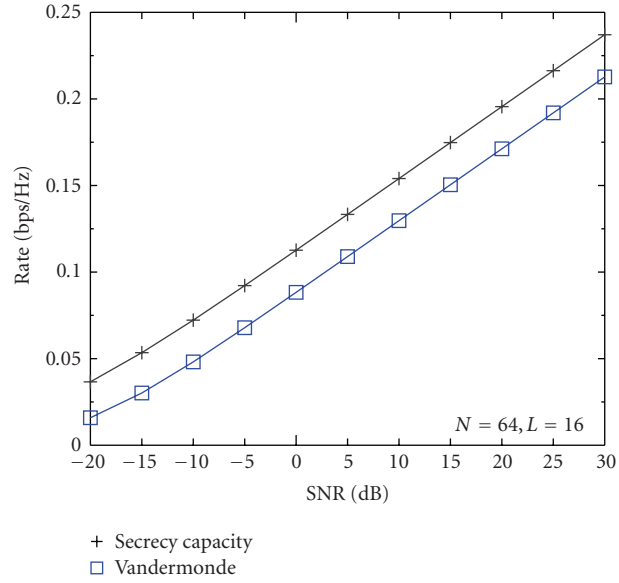

FIGURE 8: Achievable secrecy rate with one observation at receiver 1 and $N = 64$, $L = 16$ (MISO wiretap channel).

5.2. The Maximum Sum Rate Point (R_0, R_1) versus SNR. We consider the frequency-selective BCC with one confidential message to receiver 1 and one common message to two receivers. In particular, we characterize the maximum sum rate-tuple corresponding to $\gamma_0 = \gamma_1$ on the boundary of the achievable rate region $\mathcal{R}_s$. Figure 10 shows the averaged maximum sum rate-tuple (R_0, R_1) of the Vandermonde precoding both with optimal input covariance computed by the greedy algorithm and with equal power allocation. We remark that there is essentially no loss with the equal power allocation.

5.3. Two-User Secrecy Rate Region in the Frequency-Selective BCC. We consider the two-user frequency-selective BCC where the transmitter sends two confidential messages (no common message) of Section 4.2. For the sake of comparison (albeit unrealistic), we consider the special case of one observation $N = 1$ at each receiver. Notice that the two-user frequency-selective BCC is equivalent to the two-user MISO BCC with $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^{1\times(L+1)}$ whose secrecy capacity region is achieved by the S-DPC scheme [18]. The proposed

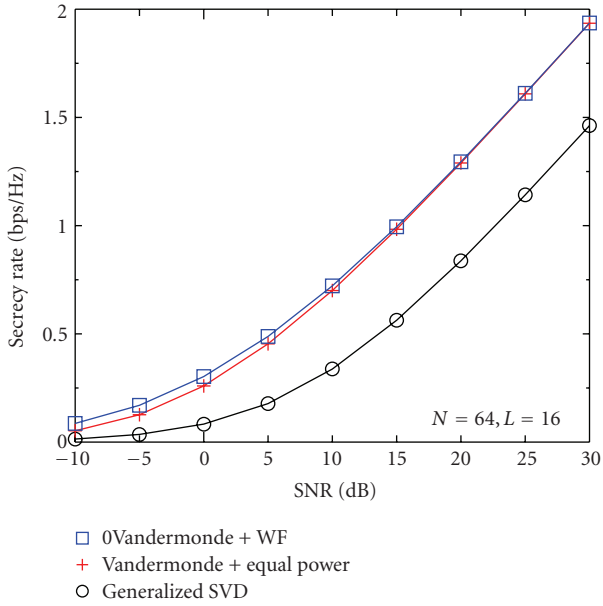

Figure 9: Achievable secrecy rate with $N = 64$, $L = 16$ (MIMO wiretap channel).

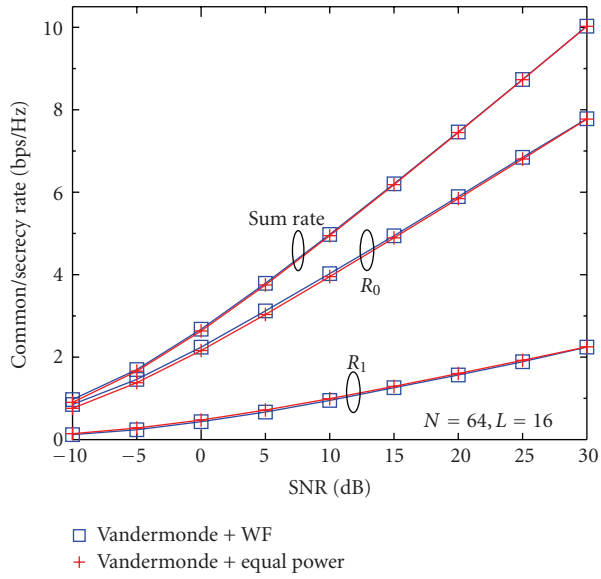

Figure 10: Achievable secrecy/common rates $N = 64$, $L = 16$ in the frequency-selective BCC.

Vandermonde precoding achieves the secrecy rate region given by all possible rate-tuples (R_1, R_2)

$$R_1 \le \frac{1}{L+1}\log\left(1 + p_1 \max_{i=1,\dots,L} \left|\mathbf{h}_1\mathbf{v}_{1,i}\right|^2\right),$$
$$R_2 \le \frac{1}{L+1}\log\left(1 + p_2 \max_{i=1,\dots,L} \left|\mathbf{h}_2\mathbf{v}_{2,i}\right|^2\right) \quad (59)$$

satisfying $p_1 + p_2 = (L+1)P$ where $\mathbf{v}_{1,i}, \mathbf{v}_{2,i}$ denotes the ith column of $\mathbf{V}_1 \in \mathbb{C}^{(N+L)\times L}$ orthogonal to $\mathbf{h}_2$, $\mathbf{V}_2 \in \mathbb{C}^{(N+L)\times L}$ orthogonal to $\mathbf{h}_1$, respectively. Figure 11 compares the averaged secrecy rate region of the Vandermonde precoding,

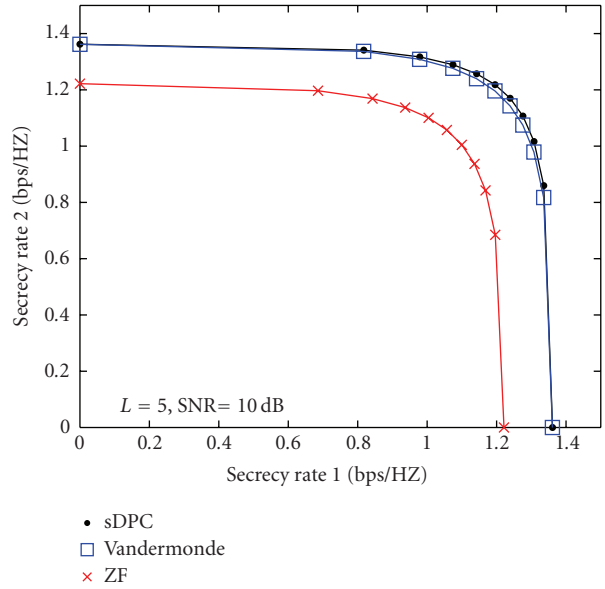

Figure 11: Achievable secrecy rate region $N = 1$, $L = 5$ (MISO-BCC).

zero-forcing beamforming, and the optimal S-DPC scheme for $L = 5$ where all entries of $\mathbf{h}_1, \mathbf{h}_2$ are i.i.d. $\sim \mathcal{N}_{\mathbb{C}}(0, 1/(L+1))$. As observed, the Vandermonde precoding achieves the near-optimal rate region. As the number of paths L increases, the gap with respect to the S-DPC becomes smaller since the Vandermonde precoding tends to choose the optimal beamformer matched to the channels.

6. Conclusions

We considered the secured communication over the frequency-selective channel by focusing on the frequency-selective BCC. In the case of a block transmission of N symbols followed by a guard interval of L symbols discarded at both receivers, the frequency-selective channel can be modeled as an $N \times (N+L)$ Toeplitz matrix. For this special type of MIMO channels, we proposed a practical yet order-optimal Vandermonde precoding which enables to send $l \le L$ streams of the confidential messages and $N - l$ streams of the common messages simultaneously over a block of $N + L$ dimensions. The key idea here consists of exploiting the frequency dimension to "hide" confidential information in the zeros of the channel seen by the unintended receiver similarly to the spatial beamforming. We also provided some application of the Vandermonde precoding in the multiuser secured communication scenarios and proved the optimality of the proposed scheme in terms of the achievable s.d.o.f. region.

We conclude this paper by noticing that there exists a simple approach to establish secured communications. More specifically, perfect secrecy can be built in two separated blocks: (1) a precoding that cancels the channel seen by the eavesdropper to fulfill the equivocation requirement, (2) the powerful off-the-shelf encoding techniques to achieve the secrecy rate. Since the practical implementation of secrecy

encoding techniques (double binning) remains a formidable challenge, such design is of great interest for the future secrecy systems.

Appendices

A. Proof of Lemma 2

In this appendix, we consider the rank of $\mathcal{T}(\mathbf{h})\mathbf{V}_1$ where $\mathbf{V}_1$ satisfies the orthogonality $\mathcal{T}(\mathbf{g})\mathbf{V}_1 = 0$. By letting $\mathbf{v}_{1,i}$ denote the ith column of $\mathbf{V}_1$ we have $\mathbf{V}_1 = [\mathbf{v}_{1,1}, \ldots, \mathbf{v}_{1,L}]$ for the case of $l = L$. We define the matrix $\mathbf{G}$ orthogonal to $\mathbf{V}_1$ by appending $L - l$ rows $\mathbf{v}_{1,l+1}^H, \ldots, \mathbf{v}_{1,L}^H$ to $\mathcal{T}(\mathbf{g})$

$$\mathbf{G} = \begin{bmatrix} \mathcal{T}(\mathbf{g}) \\ \mathbf{v}_{1,l+1}^H \\ \vdots \\ \mathbf{v}_{1,L}^H \end{bmatrix}. \tag{A.1}$$

Notice that all $N + L - l$ rows are linearly independent. By definition of $\mathbf{V}_1$, it is not difficult to see that $\mathbf{G}$ and $\mathbf{V}_1^H$ form a complete set of basis for an $N+L$-dimensional linear space. Indeed for $l = L$ the matrix $\mathbf{G}$ reduces to $\mathcal{T}(\mathbf{g})$, while $l < L$, a subset of a projection matrix onto the null space of $\mathcal{T}(\mathbf{g})$ is appended to $\mathcal{T}(\mathbf{g})$. Hence $\mathcal{T}(\mathbf{h})$ can be expressed as

$$\mathcal{T}(\mathbf{h}) = \mathbf{H}_G + \mathbf{H}_V = \mathbf{A}\mathbf{G} + \mathbf{B}\mathbf{V}_1^H, \tag{A.2}$$

where $\mathbf{H}_G$ is the projection of $\mathcal{T}(\mathbf{h})$ onto the row vectors of $\mathbf{G}$ with an $N \times (N + L - l)$ coefficient matrix $\mathbf{A}$, $\mathbf{H}_V$ is the projection of $\mathcal{T}(\mathbf{h})$ onto the row vectors of $\mathbf{V}_1$ with an $N \times l$ coefficient matrix $\mathbf{B}$

$$\begin{aligned} \operatorname{rank}(\mathcal{T}(\mathbf{h})\mathbf{V}_1) &= \operatorname{rank}\left(\left(\mathbf{A}\mathbf{G} + \mathbf{B}\mathbf{V}_1^H\right)\mathbf{V}_1\right) \\ &\overset{(a)}{=} \operatorname{rank}(\mathbf{B}) \\ &\overset{(b)}{=} \operatorname{rank}\left(\mathbf{B}\mathbf{V}_1^H\mathbf{V}_1\mathbf{B}^H\right) \\ &= \operatorname{rank}(\mathbf{H}_V) \\ &\overset{(c)}{=} l, \end{aligned} \tag{A.3}$$

where (a) follows from the orthogonality $\mathbf{G}\mathbf{V}_1 = 0$ and $\mathbf{V}_1^H\mathbf{V}_1 = \mathbf{I}_l$, (b) follows from $\operatorname{rank}(\mathbf{B}\mathbf{B}^H) = \operatorname{rank}(\mathbf{B})$. The equality (c) is obtained as follows. We notice

$$\begin{aligned} \operatorname{rank}\begin{bmatrix} \mathbf{H}_V \\ \mathbf{G} \end{bmatrix} &\overset{(d)}{=} \operatorname{rank}\begin{bmatrix} \mathbf{H}_V + \mathbf{H}_G \\ \mathbf{G} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} \mathcal{T}(\mathbf{h}) \\ \mathbf{G} \end{bmatrix} \\ &\overset{(e)}{=} \min(N + L, 2N + L - l) = N + L, \end{aligned} \tag{A.4}$$

where in (d) adding $\mathbf{H}_G$ does not change the rank, (e) follows because any set of $N + L$ rows taken from $\mathcal{T}(\mathbf{h}), \mathbf{G}$ is linearly independent (from the assumption that $\mathbf{h}$, $\mathbf{g}$ are linearly independent). Since $\mathbf{H}_V$ is orthogonal to $\mathbf{G}$, (A.4) yields

$$\operatorname{rank}(\mathbf{H}_V) = N + L - (N + L - l) = l \tag{A.5}$$

which establishes (c).

B. Proof of Theorem 5

We consider the following three cases given in Lemma 4.

Case 1. Supposing $R_{01} < R_{02}$, we consider the objective function f_1 in (25). The objective is concave only when $\gamma_1 \geq \gamma_0$. Nevertheless, we consider the KKT conditions which are necessary for the optimality. It can be easily shown that the KKT conditions are given by (29) where $\mathbf{\Psi}_i \succeq 0$ is the Lagrangian dual matrix associated to the positive semidefiniteness constraint of $\mathbf{S}_i$ for $i = 0, 1$ and $\mu \geq 0$ is the Lagrangian dual variable associated to the total power constraint. It clearly appears that for $\gamma_1 \geq \gamma_0$ the objective is concave in $\mathbf{S}_0$, $\mathbf{S}_1$ and the problem at hand is convex. In this case, any convex optimization algorithm, the gradient-based algorithm [37] for example, can be applied to find the optimal solution while the algorithm converges to a local optimal solution for $\gamma_1 < \gamma_0$.

Case 2. Supposing $R_{02} < R_{01}$, we consider the objective function f_2 in (26). Since the problem is convex (f_2 is concave and the constraint is linear in $\mathbf{S}_0, \mathbf{S}_1$), the KKT conditions are necessary and sufficient for optimality. We form the Lagrangian and obtain the following KKT conditions:

$$\begin{aligned} \gamma_0\mathbf{G}_0^H\left(\mathbf{I}_N + \mathbf{G}_0\mathbf{S}_0\mathbf{G}_0^H\right)^{-1}\mathbf{G}_0 + \mathbf{\Psi}_0 &= \mu\mathbf{I}_{N+L-l}, \\ \gamma_1\mathbf{H}_1^H\left(\mathbf{I}_N + \mathbf{H}_1\mathbf{S}_1\mathbf{H}_1^H\right)^{-1}\mathbf{H}_1 + \mathbf{\Psi}_0 &= \mu\mathbf{I}_l, \\ \operatorname{tr}(\mathbf{S}_0) + \operatorname{tr}(\mathbf{S}_1) &= \overline{P}, \\ \operatorname{tr}(\mathbf{\Psi}_i\mathbf{S}_i) = 0, \quad i &= 0, 1, \end{aligned} \tag{B.1}$$

where $\mathbf{\Psi}_i \succeq 0$ is the Lagrangian dual matrix associated to the positive semidefiniteness constraint of $\mathbf{S}_i$ for $i = 0, 1$ and $\mu \geq 0$ is the Lagrangian dual variable associated to the total power constraint. By creating N parallel channels via SVD on $\mathbf{G}_0$, $\mathbf{H}_1$ in (28), we readily obtain the solution (30).

Case 3. For $0 < \theta < 1$, we consider the objective function f_3 in (27). In the following we focus on $\gamma_0 > 0$. Notice that if $\gamma_0 = 0$ we have $R_{01} = R_{02} = 0$ which yields the corner point $(0, R_1^{\text{vdm}})$ where R_1^{vdm} denotes the secrecy rate characterized in (32). The KKT conditions, necessary for the optimality, are given by (31) where $\mathbf{\Psi}_i \succeq 0$ is the Lagrangian dual matrix associated to the positive semidefiniteness constraints for $i = 0, 1$ and $\mu \geq 0$ is the Lagrangian dual variable associated to the total power constraint. The gradient-based algorithm [37] can be applied to find the solution satisfying these KKT conditions. Although this algorithm yields the optimal and unique solution for $\gamma_1 \geq \gamma_0\theta$, the algorithm converges to a local optimal solution for $\gamma_1 < \gamma_0\theta$.

C. Proof of Lemma 9

In the following, we provide the encoding/decoding scheme to achieve a vertex point within R corresponding to a specific encoding order π. Our proof builds on the successive Gel'fand-Pinsker coding [38] and random binning for ensuring the perfect secrecy. The overall region R is obtained

by taking the union over all possible $K!$ encoding orders followed by the convex hull operation. We extensively use the notation $A_\varepsilon^{(n)}(P_{X,Y})$ to denote a set of jointly typical sequences x, y of length n with respect to the distribution $P(x, y)$. We let $\varepsilon > 0$ arbitrary small for a large n.

(a) Codebook Generation. Fix $P(u), P(v_1 \mid u), \ldots, P(v_K \mid u)$ and $P(x \mid u, v_1, \ldots, v_K)$. We define for $k = 1, \ldots, K$

$$\begin{aligned} L_{\pi(k)} &\triangleq I(V_{\pi(k)}; Z \mid V_{\pi(1)}, \ldots, V_{\pi(k-1)}, U) - \varepsilon \\ M_{\pi(k)} &\triangleq I(V_{\pi(k)}; V_{\pi(1)}, \ldots, V_{\pi(k-1)} \mid U) + \varepsilon \end{aligned} \tag{C.1}$$

and we let $M_{\pi(1)} = 0$. the joint distribution factors as The stochastic encoder randomly generates

(i) i.i.d. codewords $u(w_0)$ according $P(\mathrm{u}^{\mathrm{n}}) = \prod_{i=1}^{n} P(u_i)$ where $w_0 \in \{1, \ldots, 2^{n(R_0 - \varepsilon)}\}$.

(ii) For user $\pi(1)$, $2^{nI(V_{\pi(1)}; Y_{\pi(1)} \mid U)} = 2^{n(R_{\pi(1)} + L_{\pi(1)})}$ i.i.d. codewords $\mathrm{v}_{\pi(1)}(w_{\pi(1)}, j_{\pi(1)})$ with $P(\mathrm{v}_{\pi(1)}^n) = \prod_{i=1}^{n} P(v_{\pi(1)})$, where the indices are given by

$$w_{\pi(1)} \in \left\{1, \ldots, 2^{nR_{\pi(1)}}\right\}, \qquad j_{\pi(1)} \in \left\{1, \ldots, 2^{nL_{\pi(1)}}\right\}. \tag{C.2}$$

(iii) For user $\pi(k)$, $2^{nI(V_{\pi(k)}; Y_{\pi(k)} \mid U)} = 2^{n(R_{\pi(k)} + L_{\pi(k)} + M_{\pi(k)})}$ i.i.d. codewords $\mathrm{v}_{\pi(k)}(w_{\pi(k)}, j_{\pi(k)}, i_{\pi(k)})$ with $P(\mathrm{v}_{\pi(k)}^n) = \prod_{i=1}^{n} P(v_{\pi(k)})$, where the indices are given by

$$\begin{aligned} w_{\pi(k)} \in \left\{1, \ldots, 2^{nR_{\pi(k)}}\right\}, j_{\pi(1)} \in \left\{1, \ldots, 2^{nL_{\pi(k)}}\right\}, \\ i_{\pi(k)} \in \left\{1, \ldots, 2^{nL_{\pi(k)}}\right\}. \end{aligned} \tag{C.3}$$

(b) Encoding. To send the messages $w_0, w_1, \ldots, w_k$, we first choose randomly the index w_0 and the corresponding codeword $u(w_0)$. Given the common message $u(w_0)$, we choose randomly the codeword $\mathrm{v}_{\pi(1)}$ within the bin $w_{\pi(1)}$, that is, the index $j_{\pi(1)}$, such that $(\mathrm{u}, \mathrm{v}_{\pi(1)}) \in A\varepsilon^n(P_{U,V_{\pi(1)}})$. Then successively choose the codeword $\mathrm{v}_{\pi(k)}$, that is, the indices $j_{\pi(k)}, i_{\pi(k)}$, such that

$$(\mathrm{u}, \mathrm{v}_{\pi(1)}, \ldots, \mathrm{v}_{\pi(k)}) \in A_\varepsilon^n\left(P_{U,V_{\pi(1)}}, \ldots, \mathrm{V}_{\pi(k)}\right). \tag{C.4}$$

If there are more than one such sequence, it randomly selects one. Finally the encoder selects according to $P(x \mid v_1, \ldots, v_K)$.

(c) Decoding. The received signals at the K legitimate receivers are $\mathrm{y}_1^n, \ldots, \mathrm{y}_{\mathrm{k}}^{\mathrm{n}}$, the outputs of the channels $P(y_k^n \mid x^n) = \prod_{i=1}^{n} P(y_k^n \mid x^n)$ for any k. Receiver k chooses $w_0^{(k)}, w_k$ so that

$$\left(\mathrm{u}\left(w_0^{(k)}\right), \mathrm{v}_{\mathrm{k}}(w_k, j_k), \mathrm{y}_{\mathrm{k}}\right) \in A_\varepsilon^{(n)}(P_{U,V_k,Y_k}) \tag{C.5}$$

if such pair $w_0^{(k)}, w_k$ exists and unique. Otherwise it declares an error.

(d) Error Probability Analysis. Without loss of generality, we assume that the message set is $w_0 = w_1 = \ldots = w_k = 1$. We remark that an error is declared if one or more of the following events occur.

(i) Encoding fails

$$\begin{aligned} E_1 \triangleq \{&(\mathrm{u}(1), \mathrm{v}_{\pi(1)}(1, j_{\pi(1)}), \ldots, \mathrm{v}_{\pi(\mathrm{k})}(1, j_{\pi(K)}, i_{\pi(K)})) \\ &\notin A_\varepsilon^{(n)}(P_{U,V_1,\ldots,V_K})\Big\}. \end{aligned} \tag{C.6}$$

From the construction of the codebook above, we have $P(E_1) \leq \varepsilon$.

(ii) Decoding step 1 fails; there does not exist a jointly typical sequence for some k, that is,

$$E_2^k \triangleq \left\{\left(\mathrm{u}(1), \mathrm{v}_{\mathrm{k}}(1, j_k, i_k), \mathrm{y}_{\mathrm{k}}\right) \notin A_\varepsilon^{(n)}(P_{U,V_k,Y_k})\right\}. \tag{C.7}$$

From joint typicality [39] we have $P(E_2^k) \leq \varepsilon$ for any k.

(iii) Decoding step 2 fails; there exits other sequences satisfying the joint typicality for some k

$$\begin{aligned} E_3^k \triangleq \Big\{&\forall\left(w_0^{(k)}, w_k\right) \neq (1,1), \left(\mathrm{u}(w_0), \mathrm{v}_{\mathrm{k}}(w_k, j_k, i_k), \mathrm{y}_{\mathrm{k}}\right) \\ &\in A_\varepsilon^{(n)}(P_{U,V_k,Y_k})\Big\}. \end{aligned} \tag{C.8}$$

It can be shown that we have $P(E_3^k) \leq \varepsilon$ if

$$R_k + L_k + M_k \leq I(V_k; Y_k \mid U) \tag{C.9}$$

for any k. Hence, the error probability $P_e^{(n)} = P(E_1 \bigcup (\cup_k E_{2k}) \bigcup (\cup_k E_{3k})) \leq \varepsilon$ if the rate-tuple satisfies (44).

(e) Equivocation Calculation. To prove the equivocation requirement

$$\sum_{k \in \mathcal{K}} R_k - \frac{1}{n} H(W_{\mathcal{K}} \mid Z^n) \leq \frac{\varepsilon}{n}, \ \mathcal{K} \subseteq \{1, \ldots, K\}, \tag{C.10}$$

where we denote $W_{\mathcal{K}} = \{W_k, k \in \mathcal{K}\}$, we remark that it is sufficient to verify the above inequality for $\mathcal{K} = \{1, \ldots, K\}$ due to [24, Lemma 1]. Hence, we check whether the the sum

rate secrecy constraint is satisfied by the proposed encoding strategy.

$$\begin{aligned}
H(W_1,\ldots,W_K \mid Z^n) &\overset{(a)}{\geq} H(W_1,\ldots,W_K \mid Z^n,U^n)\\
&= H(W_1,\ldots,W_K,Z^n \mid U^n) - H(Z^n \mid U^n)\\
&= H(W_1,\ldots,W_K,V_1^n,\ldots,V_K^n,Z^n \mid U^n)\\
&\quad - H(V_1^n,\ldots,V_K^n \mid W_1,\ldots,W_K,Z^n,U^n) - H(Z^n \mid U^n)\\
&= H(W_1,\ldots,W_K,V_1^n,\ldots,V_K^n \mid U^n)\\
&\quad + H(Z^n \mid W_1,\ldots,W_K,V_1^n,\ldots,V_K^n,U^n)\\
&\quad - H(V_1^n,\ldots,V_K^n \mid W_1,\ldots,W_K,Z^n,U^n) - H(Z^n \mid U^n)\\
&\overset{(b)}{\geq} H(W_1,\ldots,W_K,V_1^n,\ldots,V_K^n \mid U^n)\\
&\quad + H(Z^nW_1,\ldots,W_K,V_1^n,\ldots,V_K^n,U^n) - n\varepsilon - H(Z^n)\\
&\overset{(c)}{=} H(W_1,\ldots,W_K,V_1^n,\ldots,V_K^n \mid U^n)\\
&\quad + H(Z^n \mid V_1^n,\ldots,V_K^n,U^n) - n\varepsilon - H(Z^n \mid U^n)\\
&\overset{(d)}{\geq} H(V_1^n,\ldots,V_K^n \mid U^n) + H(Z^n \mid V_1^n,\ldots,V_K^n,U^n)\\
&\quad - n\varepsilon - H(Z^n \mid U^n)\\
&= H(V_1^n,\ldots,V_K^nU^n) - I(V_1^n,\ldots,V_K^n;Z^n \mid U^n) - n\varepsilon\\
&\overset{(e)}{=} \sum_{k=1}^{K} H\left(V_k^n \mid U^n\right) - \sum_{j=2}^{K} I\left(V_{\pi(j)}^n;V_{\pi(1)}^n,\ldots,V_{\pi(j-1)}^nU^n\right)\\
&\quad - I(V_1^n,\ldots,V_K^n;Z^n \mid U^n) - n\varepsilon\\
&\overset{(f)}{\geq} \sum_{k=1}^{K} I\left(V_k^n;Y_k^n \mid U^n\right) - \sum_{j=2}^{K} I\left(V_{\pi(j)}^n;V_{\pi(1)}^n,\ldots,V_{\pi(j-1)}^n \mid U^n\right)\\
&\quad - I(V_1^n,\ldots,V_K^n;Z^n \mid U^n) - n\varepsilon\\
&\overset{(g)}{\geq} n\sum_{k=1}^{K} R_k - n\varepsilon,
\end{aligned} \tag{C.11}$$

where (a) follows because the conditioning decrease the entropy, (b) follows from Fano's inequality [39] stating that for a sufficiently large n we have

$$\begin{aligned}
&H(V_1^n,\ldots,V_K^n \mid W_1,\ldots,W_K,Z^n,U^n)\\
&\quad \leq 1 + nP_{e,\text{eav}}^{(n)} \sum_{k=1}^{K} \left(L_{\pi(k)} + M_{\pi(k)}\right) \leq n\varepsilon,
\end{aligned} \tag{C.12}$$

where $P_{e,\text{eav}}^{(n)}$ denotes the eavesdropper's error probability when decoding $V_1^n,\ldots,V_K^n$ with the knowledge on the message indices $w_1,\ldots,w_K$. We have that $P_{e,\text{eav}}^{(n)} \to 0$ as $n \to \infty$ if $\sum_{k=1}^{K}(L_{\pi(k)} + M_{\pi(k)}) \leq I(V_1^n,\ldots,V_K^n;Z \mid U^n) + \sum_{j=2}^{K} I(V_{\pi(j)}^n;V_{\pi(1)}^n,\ldots,V_{\pi(j-1)}^n \mid U^n)$. (c) follows from the Markov chain $W_1,\ldots,W_K \to V_1^n,\ldots,V_K^n \to Z_1^n,\ldots,Z_K^n$, (d) follows by ignoring a nonnegative term $H(W_1,\ldots,W_K \mid V_1^n,\ldots,V_K^n,U)$, (e) follows because $H(V_1^n,\ldots,V_K^n \mid U^n) = \sum_{k=1}^{K} H(V_k^n \mid U^n) - \sum_{j=2}^{K} I(V_{\pi(j)}^n;V_{\pi(1)}^n,\ldots,V_{\pi(j-1)}^n \mid U^n)$ for any permutation π over the set $\{1,\ldots,K\}$, (f) follows because $H(V_k^n \mid U^n) \geq I(V_k^n;Y_k^n \mid U^n)$ for any k, finally (g) follows because the successive encoder yields the sum rate given by

$$\begin{aligned}
&\sum_{k=1}^{K} I\left(V_k^n;Y_k^n \mid U\right) - \sum_{j=2}^{K} I\left(V_{\pi(j)}^n;V_{\pi(1)}^n,\ldots,V_{\pi(j-1)}^n \mid U\right)\\
&\qquad - I(V_1^n,\ldots,V_K^n;Z^nU).
\end{aligned} \tag{C.13}$$

This establishes the achievability.

D. Proof of Theorem 10

The achievability follows by extending Theorem 8 to the case of K confidential messages. First we remark that as a straightforward extension of Lemma 2 the following lemma holds.

Lemma D.15. *For $\sum_{k=1}^{K} l_k \leq L$, there exists a matrix $[\mathbf{V}_1,\ldots,\mathbf{V}_K]$ with $\sum_{k=1}^{K} l_k$ orthonormal columns with size $N+L$ satisfying*

$$\mathcal{T}(\mathbf{g})\mathbf{V}_k = 0_{N\times l_k}, \quad k = 1,\ldots,K, \tag{D.1}$$

$$\begin{aligned}
\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\left(\sum_{j\in\mathcal{K}} \mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right) &= \sum_{j\in\mathcal{K}} l_j,\\
\forall\,\mathcal{K} &\subseteq \{1,\ldots,K\},
\end{aligned} \tag{D.2}$$

where l_k denotes the number of columns of $\mathbf{V}_k$

A sketch of proof is given in Appendix E.

We let $\mathbf{V}_0$ be unitary matrix with $N + L - \sum_{k=1}^{K} l_k$ orthonormal columns in the null space of $[\mathbf{V}_1,\ldots,\mathbf{V}_K]$ such that $\mathbf{V}_0^H[\mathbf{V}_1,\ldots,\mathbf{V}_K] = 0$. In other words, the Vandermonde precoder $\mathbf{V} = [\mathbf{V}_0,\ldots,\mathbf{V}_K]$ is a squared unitary matrix satisfying $\mathbf{V}\mathbf{V}^H = \mathbf{I}_{N+L}$. Based on the Vandermonde precoder $\mathbf{V}$, we construct the transmit vector $\mathbf{x}$ as

$$\mathbf{x} = \sum_{k=0}^{K} \mathbf{V}_k\mathbf{u}_k, \tag{D.3}$$

where $\mathbf{u}_0,\mathbf{u}_1,\ldots,\mathbf{u}_K$ are mutually independent Gaussian vectors with zero mean and covariance $\mathbf{S}_0,\mathbf{S}_1,\ldots,\mathbf{S}_K$ satisfying $\sum_{i=0}^{K} \operatorname{tr}(\mathbf{S}_i) \leq (N+L)P$. From the orthogonality properties (D.1), the received signals become

$$\begin{aligned}
\mathbf{y}_k &= \mathcal{T}(\mathbf{h}_k)\mathbf{V}_0\mathbf{u}_0 + \mathcal{T}(\mathbf{h}_k)\mathbf{V}_k\mathbf{u}_k\\
&\quad + \mathcal{T}(\mathbf{h}_k)\sum_{j\neq k} \mathbf{V}_j\mathbf{u}_j + \mathbf{n}_k, \quad k = 1,\ldots,K,\\
\mathbf{z} &= \mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{u}_0 + \boldsymbol{\nu},
\end{aligned} \tag{D.4}$$

where receiver k observes the common message, the intended confidential message, and the interference from other users, while receiver $K+1$ observes only the common message. By letting $U = \mathbf{V}_0\mathbf{u}_0$, $V_k = U + \mathbf{V}_k\mathbf{u}_k$ for $k = 1,\ldots,K$, $X = U + \sum_{k=1}^{K} V_k$ and considering the equal power allocation to all $N+L$ streams, we readily obtain

$$I(U;Y_k) = \frac{1}{N+L} \times \log \frac{\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=0}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right|}{\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=1}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right|}, \quad \forall k, \tag{D.5}$$

$$I(U;Z) = \frac{1}{N+L}\log\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{g})^H\right|, \tag{D.6}$$

$$I(V_k;Y_k \mid U) = \frac{1}{N+L} \times \log \frac{\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=1}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right|}{\left|\mathbf{I}_N + P\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=1,j\neq k}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right|}, \quad \forall k, \tag{D.7}$$

$$I(V_{\mathcal{K}};Z \mid U) = 0, \quad \forall \mathcal{K} \subseteq \{1,\ldots,K\}, \tag{D.8}$$

and we also have $H(V_{\mathcal{K}} \mid U) = \sum_{k\in\mathcal{K}} H(V_k \mid U)$ from the independency between $V_1,\ldots,V_K$ conditioned on U. Plugging this together with (D.7) and (D.8) into (44), we have

$$R_k \le I(V_k;Y_k \mid U), \quad k = 1,\ldots,K, \qquad \sum_{k\in\mathcal{K}} R_k \le \sum_{k\in\mathcal{K}} I(V_k;Y_k \mid U). \tag{D.9}$$

In order to find the d.o.f. region, we notice

$$\begin{aligned}\operatorname{rank}\left(\mathcal{T}(\mathbf{g})\mathbf{V}_0\mathbf{V}_0^H\mathcal{T}(\mathbf{g})^H\right) &= \operatorname{rank}(\mathcal{T}(\mathbf{g})\mathbf{V}_0)\\ &\overset{(a)}{=} \operatorname{rank}(\mathcal{T}(\mathbf{g})[\mathbf{V}_0\mathbf{V}_1,\ldots,\mathbf{V}_K])\\ &\overset{(b)}{=} \operatorname{rank}(\mathcal{T}(\mathbf{g})) = N,\end{aligned} \tag{D.10}$$

$$\begin{aligned}\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=0}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right) &= \operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\mathbf{V}\mathbf{V}^H\mathcal{T}(\mathbf{h}_k)^H\right)\\ &\overset{(b)}{=} \operatorname{rank}(\mathcal{T}(\mathbf{h}_k)) = N,\end{aligned} \tag{D.11}$$

$$\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=1}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right) \overset{(c)}{=} \sum_{j=1}^{K} l_j, \tag{D.12}$$

$$\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\left(\sum_{j=1,j\neq k}^{K}\mathbf{V}_j\mathbf{V}_j^H\right)\mathcal{T}(\mathbf{h}_k)^H\right) \overset{(c)}{=} \sum_{j=1,j\neq k}^{K} l_j, \tag{D.13}$$

where (a) follows from orthogonality between $\mathcal{T}(\mathbf{g})$ and $\mathbf{V}_k$ for $k \ge 1$, (b) follows from the fact that $\mathbf{V} = [\mathbf{V}_0 \cdots \mathbf{V}_K]$ is unitary satisfying $\mathbf{V}\mathbf{V}^H = \mathbf{I}$, and (c) follows from Lemma D.15. From (D.11) and (D.12), we readily obtain $r_0 \le (N - \sum_{k=1}^{K} l_k)/(N+L)$, which is dominated by (D.10). Combining (D.12) and (D.13), we obtain $r_k \le l_k/(N+L)$ for $k = 1,\ldots,K$. This completes the achievability.

The converse follows by a natural extension of Theorem 8 to the $K+1$-user BCC. To obtain the constraint (48), we consider that the first K receivers perfectly cooperate to decode the K confidential messages and one common message. By treating these K receivers as a *virtual* receiver with KN antennas, we immediately obtain the bound (48) corresponding to the s.d.o.f. of the MIMO wiretap channel with the virtual receiver channel $[\mathcal{T}(\mathbf{h}_1)^T,\ldots,\mathcal{T}(\mathbf{h}_K)^T]^T$ and the eavesdropper channel $\mathcal{T}(\mathbf{g})$. The bound (49) is obtained by noticing that the total number of streams that receiver k can decode is limited by the d.o.f. of $\mathcal{T}(\mathbf{h}_k)$, that is, N. Namely, we have the following K inequalities:

$$l_0 + l_k \le N, \quad k = 1,\ldots,K \tag{D.14}$$

which yields $l_0 \le N - \max_k l_k$. Further by letting $l_k = L$ for any $k \in \{1,\ldots,K\}$ and and $l_j = 0$ for any $j \neq k$, we obtain $l_0 \le N - L$. Adding the last inequality and (48), we obtain (49). This establishes the converse.

E. Proof of Lemma D.15

We consider $\operatorname{rank}(\mathcal{T}(\mathbf{h}_k)\sum_{j\in\mathcal{K}}\mathbf{V}_j\mathbf{V}_j^H\mathcal{T}(\mathbf{h}_k)^H)$ for a subset $\mathcal{K} \subseteq \{1,\ldots,K\}$. First we let $\mathbf{v}_{c,1},\ldots,\mathbf{v}_{c,L}$ denote L orthonormal columns that form a unitary Vandermonde matrix orthogonal to $\mathcal{T}(\mathbf{g})$. For any subset $\mathcal{L} \subseteq \{1,\ldots,L\}$, we let $\mathbf{V}_{c,\mathcal{L}}$ be the unitary matrix formed by $|\mathcal{L}|$ columns corresponding to the subset $\mathcal{L}$ taken from $\mathbf{v}_{c,1},\ldots,\mathbf{v}_{c,L}$. Since a unitary matrix formed by $\{\mathbf{V}_k\}_{k\in\mathcal{K}}$ for any $\mathcal{K}$ can be expressed equivalently as $\mathbf{V}_{c,\mathcal{L}}$, we consider $\operatorname{rank}(\mathcal{T}(\mathbf{h}_k)\mathbf{V}_{c,\mathcal{L}}\mathbf{V}_{c,\mathcal{L}}^H\mathcal{T}(\mathbf{h}_k)^H)$. For a given $\mathcal{L}$, we let $\mathbf{V}_{c,\overline{\mathcal{L}}}$ denote a unitary matrix composed by $L - |\mathcal{L}|$ columns corresponding to the complementary set $\overline{\mathcal{L}}$ such that $\mathcal{L} + \overline{\mathcal{L}} = \{1,\ldots,L\}$. In order to derive the rank, we follow the

same approach as Appendix A. We define the matrix $\mathbf{G}_{\mathcal{L}} \in \mathbb{C}^{(N+L-|\mathcal{L}|)\times(N+L)}$ orthogonal to $\mathbf{V}_{c,\mathcal{L}}$ by appending $\mathbf{V}_{c,\overline{\mathcal{L}}}^H$ to $\mathcal{T}(\mathbf{g})$

$$\mathbf{G}_{\mathcal{L}} = \begin{bmatrix} \mathcal{T}(\mathbf{g}) \\ \mathbf{V}_{c,\overline{\mathcal{L}}}^H \end{bmatrix}, \tag{E.1}$$

where the $N + L - |\mathcal{L}|$ rows are linearly independent. Since $\mathbf{G}_{\mathcal{L}}$ and $\mathbf{V}_{c,\mathcal{L}}^H$ form a complete set of an $N + L$-dimensional linear space, $\mathcal{T}(\mathbf{h}_k)$ can be expressed as

$$\mathcal{T}(\mathbf{h}_k) = \mathbf{A}_{k,\mathcal{L}}\mathbf{G}_{\mathcal{L}} + \mathbf{B}_{k,\mathcal{L}}\mathbf{V}_{c,\mathcal{L}}^H, \quad k = 1,\ldots,K, \tag{E.2}$$

where $\mathbf{A}_{k,\mathcal{L}}, \mathbf{B}_{k,\mathcal{L}}$ is a coefficient matrix with dimension $N \times (N+L-|\mathcal{L}|)$, $N \times |\mathcal{L}|$, respectively. By recalling that any set of $N+L$ rows taken from $\mathcal{T}(\mathbf{h}_k), \mathcal{T}(\mathbf{g})$ is linearly independent for $k = 1,\ldots,K$ (from the assumption that $\mathbf{g}, \mathbf{h}_1,\ldots,\mathbf{h}_K$ are linearly independent), we can repeat the same argument as Appendix A and obtain

$$\begin{aligned} &\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\mathbf{V}_{c,\mathcal{L}}\mathbf{V}_{c,\mathcal{L}}^H\mathcal{T}(\mathbf{h}_k)^H\right) = |\mathcal{L}|, \\ &\qquad \forall \mathcal{L} \subseteq \{1,\ldots,L\},\ k = 1,\ldots,K \end{aligned} \tag{E.3}$$

which yields the result.

F. Proof of Theorem 12

The achievability follows by generalizing Theorem 8 for the case of two confidential messages. We remark that by symmetry Lemma 2 for one beamforming matrix $\mathbf{V}_1$ can be trivially extended to two beamforming matrices $\mathbf{V}_1$ and $\mathbf{V}_2$. Namely, we have

Lemma F.16. *For $l_1 \le L$ and $l_2 \le L$, there exists $\mathbf{V}_k$ with l_k orthnormal columns for $k = 1, 2$ satisfying*

$$\mathcal{T}(\mathbf{h}_k)\mathbf{V}_j = 0_{N\times l_j}, \quad k = 1,2,\ j \neq k, \tag{F.1}$$

$$\operatorname{rank}(\mathcal{T}(\mathbf{h}_k)\mathbf{V}_k) = l_k, \quad k = 1,2. \tag{F.2}$$

Further, we let $\mathbf{V}_0$ be a unitary matrix with $M = N + L - \operatorname{rank}([\mathbf{V}_1\mathbf{V}_2])$ orthonormal columns in the null space of $[\mathbf{V}_1\mathbf{V}_2]$ such that $\mathbf{V}_0^H[\mathbf{V}_1\mathbf{V}_2] = 0_{M\times(l_1+l_2)}$. We construct $\mathbf{x}$ by Gaussian superposition coding based on the Vandermonde precoder $\mathbf{V}_0$, $\mathbf{V}_1$, and $\mathbf{V}_2$. From (F.1), each user observes the vector of its confidential message and that of the common message, that is,

$$\begin{aligned} \mathbf{y}_1 &= \mathcal{T}(\mathbf{h}_1)(\mathbf{V}_0\mathbf{u}_0 + \mathbf{V}_1\mathbf{u}_1) + \mathbf{n}_1, \\ \mathbf{y}_2 &= \mathcal{T}(\mathbf{h}_2)(\mathbf{V}_0\mathbf{u}_0 + \mathbf{V}_2\mathbf{u}_2) + \mathbf{n}_2. \end{aligned} \tag{F.3}$$

By letting $U = \mathbf{V}_0\mathbf{u}_0$, $V_k = U + \mathbf{V}_k\mathbf{u}_k$ for $k = 1, 2$, $X = V_1 + V_2$ and considering equal power allocation to all streams with $p = (N + L)P/(M + l_1 + l_2)$, we readily obtain

$$\begin{aligned} &I(U; Y_k) \\ &= \frac{1}{N+L} \times \log \frac{\left|\mathbf{I}_N + p\mathcal{T}(\mathbf{h}_k)\left(\mathbf{V}_0\mathbf{V}_0^H + \mathbf{V}_k\mathbf{V}_k^H\right)\mathcal{T}(\mathbf{h}_k)^H\right|}{\left|\mathbf{I}_N + p\mathcal{T}(\mathbf{h}_k)\mathbf{V}_k\mathbf{V}_k^H\mathcal{T}(\mathbf{h}_k)^H\right|}, \\ &I(V_k; Y_k \mid U) = \frac{1}{N+L}\log\left|\mathbf{I}_N + p\mathcal{T}(\mathbf{h}_k)\mathbf{V}_k\mathbf{V}_k^H\mathcal{T}(\mathbf{h}_k)^H\right|, \\ &I(V_1; Y_2, V_2 \mid U) = I(V_2; Y_1, V_1 \mid U) = 0. \end{aligned} \tag{F.4}$$

We remark

$$\begin{aligned} &\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\mathbf{V}_k\mathbf{V}_k^H\mathcal{T}(\mathbf{h}_k)^H\right) \\ &\qquad = \operatorname{rank}(\mathcal{T}(\mathbf{h}_k)\mathbf{V}_k) = l_k, \quad k = 1,2, \\ &\operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\left(\mathbf{V}_0\mathbf{V}_0^H + \mathbf{V}_k\mathbf{V}_k^H\right)\mathcal{T}(\mathbf{h}_k)^H\right) \\ &\qquad = \operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)[\mathbf{V}_0\mathbf{V}_k]\begin{bmatrix}\mathbf{V}_0^H \\ \mathbf{V}_k^H\end{bmatrix}\mathcal{T}(\mathbf{h}_k)^H\right) \\ &\qquad \overset{(a)}{=} \operatorname{rank}\left(\mathcal{T}(\mathbf{h}_k)\left[\mathbf{V}_0\mathbf{V}_k\mathbf{V}_j\right]\right) \\ &\qquad \overset{(b)}{=} \operatorname{rank}(\mathcal{T}(\mathbf{h}_k)) = N, \end{aligned} \tag{F.5}$$

where (a) follows from orthogonality between $\mathcal{T}(\mathbf{h}_k)$ and $\mathbf{V}_j$ for $j \neq k$, (b) follows because $[\mathbf{V}_0\mathbf{V}_1\mathbf{V}_2]$ or $[\mathbf{V}_0\mathbf{V}_2\mathbf{V}_1]$ spans a complete $N + L$-dimensional space. These equations yield $l_0 + l_k \le N$ for $k = 1, 2$. This establishes the achievability.

The converse follows by noticing that the constraints (55) and (56) correspond to trivial upper bounds. To obtain (55), we consider the special case when the transmitter sends only one confidential message to one of two receivers in the presence of the eavesdropper. When sending one confidential message to receiver 1, the two-user frequency-selective BCC reduces to the MIMO wiretap channel with the legitimate channel $\mathcal{T}(\mathbf{h}_1)$ and the eavesdropper channel $\mathcal{T}(\mathbf{h}_2)$, whose s.d.o.f. is upper bounded by L. The same bound holds for receiver 2 when transmitting one confidential message to receiver 2 in the presence the eavesdropper (receiver 1). The upper bounds (56) follow because the total number of streams per receiver is limited by the individual $(N + L) \times N$ MIMO link. This establishes the converse.

Acknowledgments

The work is supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM++. The work of M. Debbah is supported by Alcatel-Lucent within the Alcatel-Lucent Chair on Flexible Radio at Supelec. The authors wish to thank Yingbin Liang for helpful discussions, and the anonymous reviewers for constructive comments. The material in this paper was partially presented at IEEE 19th International

Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France, September 2008.

References

[1] A. D. Wyner, "The wiretap channel," *Bell System Technical Journal*, vol. 54, 1975.

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, 1978.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

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

[5] A. Khisti and G. Wornell, "The MIMOME channel," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing*, 2007.

[6] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '08)*, pp. 116–120, Toronto, Canada, 2008.

[7] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[8] S. Goel and R. Negi, "Secret communication in presence of colluding eavesdroppers," in *Proceedings of the 62nd IEEE Vehicular Technology Conference (VTC '05)*, vol. 3, Atlantic City, NJ, USA, 2005.

[9] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '05)*, pp. 2152–2155, 2005.

[10] A. Khisti and G. Wornell, "Secure transmission with multiple antennas: the MISOME wiretap channel," submitted to *IEEE Transactions on Information Theory*, 2007.

[11] T. Liu and S. Shamai, "A note on the secrecy capacity of the multi-antenna wiretap channel," submitted to *IEEE Transactions on Information Theory*, 2007.

[12] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '08)*, pp. 524–528, Toronto, Canada, 2008.

[13] S. Shafiee and S. Ulukus, "Achievable rates in Gaussian MISO channels with secrecy constraints," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2466–2470, 2007.

[14] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 2471–2475, Nice, France, 2007.

[15] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Communications and Networking*. In press.

[16] H. D. Ly, T. Liu, and Y. Liang, "MIMO broadcasting with common, private, and confidential messages," in *Proceedings of the International symposium on Information Theory and Its Applications (ISITA '08)*, Auckland, New Zealand, December 2008.

[17] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: secrecy rate regions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2493–2507, 2008.

[18] R. Liu and H. V. Poor, "Secrecy capacity region of a multi-antenna Gaussian broadcast channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1235–1249, 2009.

[19] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "MIMO Gaussian broadcast channels with confidential messages," in *Proceedings of the IEEE Symposium on Information Theory (ISIT '09)*, Seoul, Korea, June-July 2009.

[20] L. C. Choo and K. K. Wong, "The K-receiver broadcast channel with confidential messages," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '09)*, Seoul, Korea, 2009.

[21] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[22] L. C. Choo and K. K. Wong, "The three-receiver broadcast channel with degraded message sets and confidential messages," submitted to *IEEE Transactions on Information Theory*.

[23] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," submitted to *EURASIP Journal on Wireless Communications and Networking*.

[24] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," submitted to *IEEE Transactions on Information Theory*.

[25] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy rate region of the broadcast channel," preprint, 2008, http://arxiv.org/abs/0806.4200.

[26] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy capacity region of the Gaussian MIMO broadcast channel," preprint, 2009, http://arxiv.org/abs/0903.3261.

[27] O. O. Koyluoglu, H. El Gamal, L. Lai, and H. V. Poor, "Interference alignment for secrecy," preprint, 2008, http://arxiv.org/abs/0810.1187.

[28] ANSI/IEEE Std 802.11, Edition (R2003), 1999, http://standards.ieee.org/getieee802/download/802.11-1999.pdf.

[29] Air Interface for Fixed and Mobile Broadband Wireless Access Systems, 2005, http://standards.ieee.org/getieee802/download/802.16e-2005.pdf.

[30] http://www.3gpp.org/Highlights/LTE/LTE.htm.

[31] Ø. Ryan and M. Debbah, "Asymptotic behaviour of random vandermonde matrices with entries on the unit circle," *IEEE Transactions on Information Theory*, vol. 55, no. 7, 2009.

[32] L. S. Cardoso, M. Kobayashi, Ø. Ryan, and M. Debbah, "Vandermonde frequency division multiplexing for cognitive radio," in *Proceedings of the 9th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '08)*, pp. 421–425, Recife, Brazil, 2008.

[33] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Lagrange/Vandermonde MUI eliminating user codes forquasi-synchronous CDMA in unknown multipath," *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 2057–2073, 2000.

[34] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer, New York, NY, USA, 1994.

[35] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.

[36] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: dirty paper coding versus linear precoding," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4787–4792, 2007.

[37] H. Viswanathan, S. Venkatesan, and H. Huang, "Downlink capacity evaluation of cellular networks with known-interference cancellation," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 802–811, 2003.

[38] X. Zhang, J. Chen, SB. Wicker, and T. Berger, "Successive coding in multiuser information theory," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 2246–2254, 2007.

[39] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.