Limited Rate Feedback in a MIMO Wiretap Channel With a Cooperative Jammer

Xinjie Yang and A. Lee Swindlehurst, Fellow, IEEE

Abstract—We study strategies for enhanced secrecy using cooperative jamming in secure communication systems with limited rate feedback. A Gaussian multiple-input multiple-output (MIMO) wiretap channel with a jamming helper is considered. The transmitter and helper both require channel state information (CSI), which is quantized at the receiver and fed back through two sum-rate-limited feedback channels. The quantization errors result in reduced beamforming gain from the transmitter, as well as interference leakage from the helper. First, under the assumption that the eavesdropper’s CSI is completely unknown, we derive a lower bound on the average main channel rate and find the feedback bit allocation that maximizes the jamming power under a constraint on the bound. For the case where statistical CSI for the eavesdropper’s channel is available, we derive a lower bound on the average secrecy rate, and we optimize the bound to find a suitable bit allocation and the transmit powers allocated to the transmitter and helper. For the case where the transmitter and helper have the same number of antennas, we obtain a closed-form solution for the optimal bit allocation. Simulations verify the theoretical analysis and demonstrate the significant performance gain that results with intelligent feedback bit allocation and power control.

Index Terms—Multiple-input multiple-output (MIMO), limited rate feedback, physical layer secrecy, cooperative jamming, random vector quantization (RVQ).

I. INTRODUCTION

The use of multiple antennas to enhance wireless security via beamforming has recently been a subject of significant interest. Depending on the availability of relevant channel state information (CSI), beamforming techniques can be used to steer information away from eavesdroppers, direct jamming signals directly at them, or fill the spatial modes orthogonal to those of the desired receiver with artificial noise [1]–[5]. Artificial interference can originate at the information source, or it can be produced by cooperating jammers present in the network [6]–[13]. In either case, multiple antennas can be used to mitigate the effect of the jamming at the legitimate receiver, provided that accurate CSI is available at both the source of the interference and the receiver.

Assuming perfect CSI at the transmitter is unrealistic in practice due to channel estimation errors in time division duplex (TDD) systems or limited rate feedback and delay in frequency division duplex (FDD) systems. The design and impact of limited rate feedback in FDD systems without secrecy considerations has been studied by a number of researchers; see for example [14]–[18]. The basic approach of these limited rate feedback methods is that, instead of full CSI, only a limited number of feedback bits representing the quantized CSI are fed back to the transmitter from the receiver. Based on the feedback, the transmitter selects the precoding matrix from a pre-designed codebook, and adapts the transmitted signals to the current eigenstructure of the channel to achieve a certain acceptable performance loss.

There has been relatively little work on the effects of limited rate feedback schemes on secrecy at the physical layer. In [19], the authors derive the optimal power allocation between the desired signal and artificial noise to maximize the secrecy rate for a given transmission power and number of feedback bits under quantized channel feedback. Moreover, they derive a scaling law between feedback bits and transmission power to maintain a constant secrecy rate loss compared to the perfect CSI case. The secrecy performance analysis of a codebook-based beamforming transmission with limited feedback is addressed in [20]. It provides an upper bound on the secrecy outage probability as a function of the amount of feedback, and demonstrates that under limited feedback, artificial-noise-aided beamforming does not exhibit any significant advantage over codebook beamforming. In [21], a multiple-input single-output (MISO) channel scenario with cooperative jamming is considered. It investigates the impact of quantized channel state information on the secrecy rate, and an adaptive bit allocation strategy is proposed to optimally divide feedback bits between the transmitter and helper channels.

In this paper, we consider the problem of limited rate feedback design for a wiretap channel with a cooperative jammer, where both the data transmitter and jamming helper require CSI feedback from the receiver. This problem is particularly interesting when the bandwidth available for feedback is limited, and the total number of feedback bits must be properly allocated between the transmitter and helper. The goal is to balance the need to achieve a strong signal from the data transmitter against the need to maximize the impact of the jamming at the eavesdropper and to minimize its impact at the receiver. Unlike [21], we assume that both the legitimate receiver and the eavesdropper possess multiple antennas, and that the transmitter could be sending multiple data streams to the receiver. We assume that both the transmitter and the cooperative jammer employ independent random vector quantization (RVQ) codebooks whose dimensions are to be optimized.
We consider two cases with respect to the eavesdropper CSI. In the first, we assume that no CSI is available for the eavesdropper, and in the second, we assume that only statistical CSI is available. For the first case, we derive a lower bound on the achievable rate for the primary channel, and following the general approach of [2], we find the feedback bit allocation that allows for maximum jamming from the helper while still maintaining the lower bound on the rate at a minimum quality-of-service level. No secrecy guarantee is possible since the eavesdropper’s CSI is completely unknown, but this approach maximizes the amount of interference available to mask the desired signal from whatever eavesdroppers are present. Next, we consider the second case, in which the statistical CSI for the eavesdropper is available. We derive a closed-form expression for the optimal bit allocation for the transmitter and jammer when they have an equal number of antennas. We then derive a lower bound for the average secrecy rate and optimize it over the power allocated to the transmitter and jammer. For both sets of CSI assumptions, simulations demonstrate a significant gain in performance when the feedback bit and power allocations are chosen according to the proposed algorithms.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and the necessary assumptions. Section III provides preliminary background information on beamforming strategies, RVQ codebooks and the calculation of average secrecy rate. In Section IV, we analyze the impact of limited rate feedback and discuss the bit allocation and power control algorithms for the two eavesdropper CSI scenarios discussed above. In Section V, we present simulation results to validate our algorithms.

Throughout the paper we use lowercase boldface letters to denote vectors and uppercase bold letters to denote matrices. The space of $m \times n$ complex matrices is denoted by $\mathbb{C}^{m \times n}$. The Hermitian transpose is represented by $(\cdot)^H$, the expectation operator $\mathbb{E}[\cdot]$, the matrix trace $\text{tr}(\cdot)$, the zero matrix $\mathbf{0}$ and the $d \times d$ identity matrix $\mathbf{I}_d$. $\{x\}^+$ denotes max $\{x,0\}$. We use $\mathbf{A}\succeq\mathbf{B}$ to denote that $\mathbf{A}$ is positive semidefinite, and write $\mathbf{A}\succeq\mathbf{0}$ if $\mathbf{A}-\mathbf{B}\succeq\mathbf{0}$. If $\mathbf{U}$ is $m \times n$ and satisfies $\mathbf{U}^H\mathbf{U} = \mathbf{I}_n$, then $\mathbf{U}^\perp$ denotes a matrix whose $m - n$ columns are orthogonal vectors that satisfy $\mathbf{U}^H\mathbf{U}^\perp = \mathbf{0}$.

II. SIGNAL MODELING ASSUMPTIONS

We depict the assumed scenario in Fig. 1, which features a transmitter (Alice) with $N_a$ antennas, a legitimate receiver (Bob) with $N_b$ antennas, a jamming helper (Hugo) with $N_h$ antennas, and a passive eavesdropper (Eve) with $N_e$ antennas. Eve may be an abstraction of multiple colluding eavesdroppers with a total of $N_e$ antennas, and Hugo is present to provide artificial interference to degrade the channel of any eavesdropper that may be present. The channels from Alice and Hugo to Bob are denoted as $\mathbf{H}_{ba} \in \mathbb{C}^{N_b \times N_a}$ and $\mathbf{H}_{bh} \in \mathbb{C}^{N_b \times N_h}$, and those to Eve are represented as $\mathbf{H}_{ea} \in \mathbb{C}^{N_e \times N_a}$ and $\mathbf{H}_{eh} \in \mathbb{C}^{N_e \times N_h}$. All channels are assumed to experience independent block fading, and the elements of these channel matrices are assumed to be independent and identically distributed (i.i.d.) and have a circularly symmetric complex Gaussian distribution with zero mean and unit variance. It is assumed that Bob has perfect CSI for $\mathbf{H}_{bh}$ and $\mathbf{H}_{ba}$. Alice and Hugo do not know these channels, but can obtain quantized information from Bob through two error-free, zero-delay feedback channels. Bob quantizes the feedback information by selecting the closest codewords from two codebooks respectively containing $2^{\beta_b}$ and $2^{\beta_h}$ entries. The $\mathbf{B}_a$ and $\mathbf{B}_h$ index bits corresponding to the closest codewords are separately fed back to Alice and Hugo. With the limited total rate constraint on the feedback channel, Bob is only able to feed back a fixed total number of bits $B$, such that

$$B_a + B_h = B.$$  \hspace{1cm} (1)

Let $l = \max\{N_a,N_h\}$ and $m = \min\{N_a,N_h\}$. We assume that Alice transmits a $d$-dimensional data stream $\mathbf{s}$, where $1 \leq d \leq m$, and that Hugo transmits a $p$-dimensional jamming signal $\mathbf{v}$. Alice and Hugo employ precoders $\mathbf{W}_a \in \mathbb{C}^{N_a \times d}$ and $\mathbf{W}_h \in \mathbb{C}^{N_h \times p}$, respectively, and Bob uses the beamforming matrix $\mathbf{W}_b \in \mathbb{C}^{N_b \times d}$ to recover the signal of interest. With these assumptions, the signals received by Bob and Eve are:

$$\tilde{\mathbf{y}}_b = \mathbf{W}_b^H\mathbf{y}_b = \mathbf{W}_b^H\mathbf{H}_{ba}\mathbf{W}_a\mathbf{s} + \mathbf{W}_b^H\mathbf{H}_{bh}\mathbf{W}_h\mathbf{v} + \tilde{\mathbf{n}}_b,$$  \hspace{1cm} (2)

$$\mathbf{y}_e = \mathbf{H}_{ea}\mathbf{W}_a\mathbf{s} + \mathbf{H}_{eh}\mathbf{W}_h\mathbf{v} + \mathbf{n}_e,$$  \hspace{1cm} (3)

where $\tilde{\mathbf{n}}_b = \mathbf{W}_b^H\mathbf{n}_b$. As discussed below, the beamforming matrix $\mathbf{W}_b$ is chosen such that $\mathbf{W}_b^H\mathbf{W}_b = \mathbf{I}_d$. The components of $\tilde{\mathbf{n}}_b$ and $\mathbf{n}_e$ are i.i.d. zero-mean complex Gaussian noise with variance $\sigma_b^2$ and $\sigma_e^2$ respectively. Without loss of generality, it is assumed that $\mathbf{n}_b$ has been normalized so that $\sigma_b^2 = 1$. We respectively define $P_s$ and $P_h$ as the power allocated to the information and jamming signals, which obey the following power constraints:

$$\mathbf{Q}_s = \mathbb{E}[\mathbf{ss}^H]\quad \text{tr}(\mathbf{W}_a\mathbf{Q}_s\mathbf{W}_b^H) = P_s \leq P_a$$ \hspace{1cm} (4)

$$\mathbf{Q}_v = \mathbb{E}[\mathbf{vv}^H]\quad \text{tr}(\mathbf{W}_h\mathbf{Q}_v\mathbf{W}_h^H) = P_h \leq P_h.$$ \hspace{1cm} (5)

III. ALGORITHMS AND METRICS

A. Beamforming Design

Based on his knowledge of $\mathbf{H}_{ba}$ and $\mathbf{H}_{bh}$, Bob determines appropriate precoders $\mathbf{W}_a$ and $\mathbf{W}_h$ and feeds them back to Alice and Hugo, respectively. This is advantageous in a scenario where the feedback bandwidth is limited, since the dimensionality of
the quantization problem is reduced compared with feedback of the quantized channel matrices [14].

Consider the jamming interference term \( \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_b \mathbf{v} \) received by Bob in (2). In the ideal case with perfect CSI at both Hugo and Bob, this interference could be eliminated completely, for example by choosing zero-forcing beamforming. Even in the limited feedback scenario, if \( d + p \leq N_b \), Bob has a sufficient number of antennas to cancel the interference. Note that Bob knows the true channel \( \mathbf{H}_{bh} \) as well as the quantized precoder \( \mathbf{W}_b \), since Bob fed this information back to Hugo. In this case, the beaformer \( \mathbf{W}_b \) can be chosen to be orthogonal to \( \mathbf{H}_{bh} \mathbf{W}_b \), and thus would completely eliminate the contribution from Hugo; consequently, Hugo could transmit with full power and have no impact on Bob, except to limit the number of data streams available for spatial multiplexing. A bit allocation strategy is unnecessary in this case. Specifically, given an arbitrary fixed precoder \( \mathbf{W}_b \), Bob could always design a zero-forcing receive beaformer \( \mathbf{W}_b \) that would completely eliminate the interference from Hugo. However, if \( N_b < d + p \leq N_b \), Bob has insufficient degrees of freedom for canceling the interference. Therefore, decoupling the links requires \( \mathbf{W}_b \) to be chosen orthogonal to \( \mathbf{W}_b^H \mathbf{H}_{bh} \). In general, this requires that \( N_b > d \), and hence that \( p \), the dimension of the jamming signal, is set to \( p = N_b - d \). Since Hugo uses a quantized precoder fed back from Bob, \( \mathbf{W}_b \) is no longer orthogonal to \( \mathbf{W}_b^H \mathbf{H}_{bh} \) and causes interference leakage that Bob is not able to filter out. We focus on this case throughout the paper.

Given that the instantaneous information about Eve’s CSI is unavailable, a natural approach for the system model is to choose a precoder that provides Bob with a strong signal from Alice, and a receive beaformer that is focused on the information signal. Consequently, the beaformers \( \mathbf{W}_a \) and \( \mathbf{W}_b \) are chosen from the principle right and left singular vectors of \( \mathbf{H}_{ba} \). Without access to Eve’s instantaneous CSI, a reasonable approach is for Hugo to spread the jamming power uniformly across the \( N_b - d \) jamming dimensions. We also assume a relatively high SNR scenario where errors due to the limited feedback dominate, and hence Alice also uniformly distributes her power across the \( d \) signal dimensions.

More specifically, define the singular value decomposition of \( \mathbf{H}_{ba} \) and \( \mathbf{W}_b^H \mathbf{H}_{bh} \) as follows:

\[
\text{svd} (\mathbf{H}_{ba} ) = \mathbf{U}_a \mathbf{A} \mathbf{V}_a^H = \begin{bmatrix} \mathbf{U}_{a1} \mathbf{V}_{a1}^T \\ \mathbf{U}_{a2} \mathbf{V}_{a2}^T \end{bmatrix}
\]

\[
\text{svd} (\mathbf{W}_b^H \mathbf{H}_{bh} ) = \mathbf{U}_b \mathbf{A}_b \mathbf{V}_b^H = \mathbf{V}_b \mathbf{A}_b \mathbf{V}_b^T
\]

where \( \mathbf{U}_{a1}, \mathbf{V}_{a1}, \mathbf{A}_b \) and \( \mathbf{V}_b \) denote the first \( d \) columns of \( \mathbf{U}_a, \mathbf{V}_a, \mathbf{A}_b \) and \( \mathbf{V}_b \). \( \mathbf{U}_{a2}, \mathbf{V}_{a2} \) and \( \mathbf{V}_b \) contain the remaining columns of \( \mathbf{U}_a, \mathbf{V}_a \) and \( \mathbf{V}_b \). \( \mathbf{A}_b \) denotes the upper left \( d \times d \) diagonal submatrix of \( \mathbf{A}_b \), and \( \mathbf{A}_b \) is the lower right diagonal submatrix. In the ideal case without quantization errors, \( \mathbf{W}_a = \mathbf{V}_{a1} \) and \( \mathbf{W}_b = \mathbf{U}_{a1} \). Furthermore, the zero-forcing precoding matrix for Hugo is \( \mathbf{W}_b = \mathbf{V}_{b1} \). Note that the ideal precoders can be uniquely defined by their orthogonal complements, since \( \mathbf{V}_{a1} = \mathbf{V}_{a2}^T \) and \( \mathbf{V}_{b1} = \mathbf{V}_{b2}^T \). Thus, the feedback from Bob must be either the precoder or its orthogonal complement, whichever is of smaller dimension and requires fewer bits to encode:

\[
\mathbf{W}_a = \begin{cases} \mathbf{V}_{a1} & \text{if } d \leq N_a - d \\ \mathbf{V}_{a2} & \text{if } d > N_a - d \end{cases}
\]

\[
\mathbf{W}_b = \begin{cases} \mathbf{V}_{b1} & \text{if } d < N_b - d \\ \mathbf{V}_{b2} & \text{if } d \geq N_b - d, \end{cases}
\]

where the hat indicates a quantized version of the matrix. The size of the codewords for \( \mathbf{W}_a \) and \( \mathbf{W}_b \) are \( N_a \times M_a \) and \( N_b \times M_b \), respectively, where \( M_a = \min \{ d, N_a - d \} \) and \( M_b = \min \{ d, N_b - d \} \). Once \( d \) is determined, Alice and Hugo use the indices fed back by Bob to determine which codebooks are to be used for the precoders.

B. Random Quantization Codebooks

The precoders \( \mathbf{W}_a \) and \( \mathbf{W}_b \) are drawn from the random quantization codebooks \( \mathcal{C}_a \) and \( \mathcal{C}_b \), respectively, which are known to Alice and Hugo beforehand. Assuming the codebooks are indexed by \( B_a \) and \( B_b \) feedback bits, respectively, the codebooks for Alice and Hugo contain \( 2^{B_a} \) and \( 2^{B_b} \) entries. Each of the codewords \( \mathbf{c}_a \in \mathcal{C}_a \) and \( \mathbf{c}_b \in \mathcal{C}_b \) are generated independently and isotropically over the \( N_a \times M_a \) and \( N_b \times M_b \) Grassmann manifold, and are assumed to be semi-unitary.

The choice of codebook used depends on whether the precoder or its orthogonal complement is fed back; in particular, we use the notation

\[
\mathcal{C}_a \triangleq \begin{cases} \mathcal{C}_{a1} & \text{if } d \leq N_a - d \\ \mathcal{C}_{a2} & \text{if } d > N_a - d \end{cases}
\]

\[
\mathcal{C}_b \triangleq \begin{cases} \mathcal{C}_{b1} & \text{if } d < N_b - d \\ \mathcal{C}_{b2} & \text{if } d \geq N_b - d, \end{cases}
\]

The precoder assignments in (6) and (7) are assumed to satisfy the following rule:

\[
\hat{\mathbf{V}}_{ij} = \arg \min_{\mathcal{C}_{ij} \in \mathcal{C}_{ij}} d^2 (\mathbf{V}_{ij}, \mathbf{C}_{ij}) \quad i = a, b; \ j = 1, 2,
\]

where \( d (\mathbf{V}_{ij}, \mathbf{C}_{ij}) \) is the chordal distance between \( \mathbf{V}_{ij} \) and \( \mathbf{C}_{ij} \), and is given by [22]

\[
d^2 (\mathbf{V}_{ij}, \mathbf{C}_{ij}) = M_i - \text{tr} (\mathbf{V}_{ij}^H \mathbf{C}_{ij} \mathbf{C}_{ij}^H \mathbf{V}_{ij}) \quad i = a, b; \ j = 1, 2.
\]

It is shown in [22] that the average distortion \( D_i \) associated with the given codebooks \( \mathcal{C}_{ij} \) is bounded above by

\[
D_i = \mathbb{E} \left[ d^2 (\hat{\mathbf{V}}_{ij}, \mathbf{V}_{ij}) \right] \leq G_i 2^{-\frac{\mu_i}{T_i}} \quad i = a, b; \ j = 1, 2,
\]

where \( T_i = d (N_i - d) \) and

\[
G_i = \left( \frac{\mu_i}{T_i} \right) \left[ \sum_{k=1}^{M_i} \frac{\Gamma (N_i - k + 1)}{\Gamma (M_i - k + 1)} \right]^{-\frac{\mu_i}{T_i}},
\]

and \( \Gamma (\cdot) \) represents the gamma function.

Since the quantized precoding matrices \( \mathbf{W}_a \) and \( \mathbf{W}_b \) are both semi-unitary, the power constraints in (4) and (5) become

\[
\text{tr} (\mathbf{Q}_a) = P_a \leq P_a \quad \text{tr} (\mathbf{Q}_b) = P_b \leq P_b.
\]
C. Average Secrecy Rate

For the case where the statistics of the eavesdropper channel are available, the metric of interest is the average achievable secrecy rate of the system. Assume s is Gaussian. Using (2) and (3), we define the average achievable secrecy rate as in [19] and [21]:

\[
\bar{R}_{sec} = \left[ I \left( S; Y_b | H_{ba} \right) - I \left( S; Y_e | H_{ea} \right) \right] + \log_2 \left( \frac{P}{N_h - d} I_{N_h} \right).
\]

\[
\bar{R}_{sec} = \left\{ \begin{array}{ll}
\log_2 \left[ \left| K_b + W_b H_{ba} W_a Q_s W_a^H H_{ba}^H W_b \right| \right] \\
\log_2 \left[ \left| K_e + H_{eb} W_a Q_e W_a^H H_{eb}^H \right| \right] + \log_2 \left( \frac{P}{N_h - d} I_{N_h} \right)
\end{array} \right.
\]

where

\[
K_b = W_b^H W_b + W_b^H H_{bb} W_b Q_s W_b^H H_{bb}^H W_b
\]

\[
K_e = \sigma_v^2 I_{N_e} + H_{eb} W_a Q_e W_a^H H_{eb}^H
\]

Since we assume a uniform power allocation for both Alice and Hugo, we have

\[
Q_s = \frac{P}{d} I_d \quad Q_e = \frac{P}{d} I_{N_h - d}.
\]

Applying the receive beamforming matrix \( W_b \), the average secrecy rate \( \bar{R}_{sec} \) can be interpreted as in (15), equation (15)-(17) as shown at the bottom of this page, for a given number of data streams \( d \).

The following lemma, discussed in [24] and [25], provides an asymptotic approximation to the average mutual information of a MIMO channel for a large number of antennas.

**Lemma 1:** Given an \( r \times t \) matrix \( \mathbf{H} \) composed of independent, circular complex, zero-mean, unit variance Gaussian elements, then for asymptotically large \( r, t \) we have

\[
\mathbb{E} \left[ \log_2 \left| I + \frac{P}{t} \mathbf{H} \mathbf{H}^H \right| \right] = t F \left( \frac{r}{t}, P \right),
\]

where

\[
F \left( \beta, \rho \right) = \log_2 \left( 1 + \rho \left( \sqrt{\beta} + 1 \right)^2 \right)
\]

\[
+ \left( \beta + 1 \right) \log_2 \left( \frac{1 + \sqrt{1 - a}}{2} \right)
\]

\[
- \left( \log_2 e \right) \sqrt{\beta} \frac{1 - \sqrt{1 - a}}{1 + \sqrt{1 - a}}
\]

\[
+ \left( \beta - 1 \right) \log_2 \left( \frac{1 + \gamma}{\gamma + \sqrt{1 - a}} \right)
\]

\[
a = \frac{4 \rho \sqrt{\beta}}{1 + \rho \left( \sqrt{\beta} + 1 \right)^2} \quad \gamma = \frac{\sqrt{\beta} - 1}{\sqrt{\beta} + 1}
\]

Although (18) was originally derived using the central limit theorem under an asymptotic assumption on the number of antennas, the approximation works quite well even for a small number of antennas [25].

IV. LIMITED RATE FEEDBACK ANALYSIS

A. Unknown Eavesdropper CSI

To begin our analysis, we investigate the scenario where Eve’s channel state information is completely unknown. A formal secrecy metric is impossible to define without any knowledge of Eve, so a reasonable alternative is to maximize the amount of jamming broadcast by Hugo subject to a certain acceptable rate for Bob, i.e., \( \bar{R}_b (d) \geq \bar{R}_e \) taking the effects of the limited feedback quantization error into account. Although the secrecy rate of such a scheme cannot be quantified, this approach aims at making the unintended reception of the signal as difficult as possible [26]. Direct evaluation of (17), at the bottom of the page, in terms of the parameters of interest is difficult, so instead we focus on optimizing a lower bound on (17) derived below.

**Lemma 2:** Under the random quantization codebook model, we have the following properties:

\[
\mathbb{E} \left[ V_{ij}^H \hat{V}_{ij} V_{ij}^H \hat{V}_{ij} \right] = \mathbb{E} \left[ \hat{V}_{ij} V_{ij} \hat{V}_{ij} V_{ij} \right] = \left( 1 - \frac{D_i}{M_i} \right) I_{M_i},
\]

**Proof:** See Appendix A.

\[
\bar{R}_{sec} (d) = \left\{ \begin{array}{ll}
\bar{R}_b (d) - \mathbb{E}_{H_{ba}, H_{ea}, W_a, W_b} \left[ \log_2 \left[ \frac{\sigma_v^2 I_{N_e} + \frac{P}{N_h - d} H_{eb} W_a Q_e W_a^H H_{eb}^H}{\sigma_v^2 I_{N_e} + \frac{P}{N_h - d} H_{eb} W_a Q_e W_a^H H_{eb}^H} \right] \right] + \log_2 \left( \frac{P}{N_h - d} I_{N_h} \right) \end{array} \right.
\]

where

\[
\bar{R}_b (d) \triangleq \mathbb{E}_{H_{ba}, H_{ea}, W_a, W_b} \left[ \log_2 \left[ \frac{W_b^H W_b + \frac{P}{N_h - d} W_b^H H_{bb} W_b H_{bb}^H W_b}{W_b^H W_b + \frac{P}{N_h - d} H_{bb} W_b H_{bb}^H W_b} \right] + \log_2 \left( \frac{P}{N_h - d} I_{N_h} \right) \right]
\]

\[
= \mathbb{E}_{H_{ba}, H_{ea}, W_a, W_b} \left[ \log_2 \left[ I_d + \frac{P}{N_h - d} U_{a1} V_{a1} H_{h1}^H W_h W_{h1}^H U_{h1} + \frac{P}{d} A_{a1} V_{a1} W_a W_a^H V_{a1} A_{a1}^H \right] \right]
\]

\[
+ \mathbb{E}_{H_{ba}, H_{ea}, W_a, W_b} \left[ \log_2 \left[ I_d + \frac{P}{N_h - d} U_{a1} V_{1h1} H_{h1}^H W_h W_{h1}^H U_{h1} + \frac{P}{d} A_{a1} V_{a1} W_a W_a^H V_{a1} A_{a1}^H \right] \right].
\]
Theorem 1: The average rate of the main channel using random quantization codebooks of size $B_a$ and $B_h$ can be approximately bounded below by

$$\mathcal{R}_h(d) > \mathcal{R}_{h,LB}(d)$$

$$\triangleq \mathcal{R}(d) - M_h \log_2 \left( 1 + \frac{P_i N_h G_h 2^{-\frac{\beta}{M_h}}}{(N_h - d) M_h} \right) - M_h \log_2 \frac{M_h}{M_h - G_h 2^{-\frac{\beta}{M_h}}},$$

(20)

where $\mathcal{R}(d) = dF \left( \frac{t}{\pi}, \frac{m}{\pi} P_s \right)$ is the ideal rate achieved with perfect CSI.

Proof: See Appendix B.

The use of limited rate feedback produces the two negative terms in (20). The first term is due to interference leakage from Hugo, and the second is due to mismatch with the desired beamforming for Alice. Together they constitute the throughput loss caused by quantization errors. It is clear from (20) that $\mathcal{R}_h(d) = \mathcal{R}(d)$ and the rate loss is 0 if the feedback rate is infinite, i.e., if $B_a, B_h \to \infty$.

The lower bound is tight when the number of feedback bits is sufficiently large and properly allocated, so we use it as an approximation to the average link rate. Since the codebook size is fixed for all the channel realizations, our optimization problem is to find a bit allocation strategy for Alice and Hugo that maximizes the jamming power subject to the constraint that the lower bound on the average rate in (20) is above the target value $R_t$. In addition to optimal values for $B_a$ and $B_h$, there is a best choice for $d$ that maximizes the jamming power. In practical applications, the possible number of values for $d$ is limited ($d \leq m$), so our approach is to repeat the optimization over $B_a$ and $B_h$ for each possible $d$, and then choose the value for $d$ whose optimal bit allocation provides the largest jamming power. Since Alice is unable to exploit rate or power adaptation, she always transmits with full power in this scenario, i.e., $P_s = P_h$. Consequently, we choose $B_a$ and $B_h$ to maximize the jamming power:

$$P^*_i(d) = \max_{B_a, B_h} P_i(B_a, B_h)$$

s.t. $\mathcal{R}_{h,LB}(d) = R_t$

$$B_a + B_h = B$$

$$B_a, B_h \in \mathbb{Z}^+$$

$$0 \leq P_i \leq P_h,$$

where $\mathbb{Z}^+$ is the set of non-negative integers.

In general, the required optimization is an integer programming problem. However, if the integer constraint is relaxed, the following closed-form expression for the optimal solution can

1) $2G_h < M_h (1 - r_0^{-1})$ and $P_h \geq \left( \frac{N_h - d}{N_h G_h} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}}$

$$\begin{align*}
&B^*_h(d) = B \\
P^*_i(d) = \left( \frac{N_h - d}{N_h G_h} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}} & \quad \text{if } 0 < R_t < \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h}
\end{align*}$$

(22a)

$$\begin{align*}
&B^*_h(d) = B - T_h \log_2 \frac{2G_h}{M_h} r_0^{-1} 2^{-\frac{\beta}{r_0}}
\end{align*}$$

$$\left\{ \begin{array}{ll} 
& \text{if } \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} \leq R_t \leq \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}} \\
&B^*_h(d) = 0 \end{array} \right.$$}

$$\begin{align*}
P^*_i(d) = \left( \frac{N_h - d}{N_h G_h} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}} & \quad \text{if } \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}} < R_t < \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}}
\end{align*}$$

2) $M_h (1 - r_0^{-1}) \leq 2G_h < M_h 2^{-\frac{\beta}{r_0}} (1 - r_0^{-1})$ and $P_h \geq \left( \frac{N_h - d}{4N_h G_h^2} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}}$

$$\begin{align*}
&B^*_h(d) = B - T_h \log_2 \frac{2G_h}{M_h} r_0^{-1} 2^{-\frac{\beta}{r_0}}
\end{align*}$$

$$\left\{ \begin{array}{ll} 
& \text{if } 0 < R_t \leq \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}} \\
&B^*_h(d) = 0 \end{array} \right.$$}

$$\begin{align*}
P^*_i(d) = \left( \frac{N_h - d}{N_h G_h} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}} & \quad \text{if } \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}} < R_t \leq \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}}
\end{align*}$$

(22b)

3) $M_h 2^{-\frac{\beta}{r_0}} (1 - r_0^{-1}) \leq 2G_h < 2M_h 2^{-\frac{\beta}{r_0}} (1 - r_0^{-1})$ and $P_h \geq \left( \frac{N_h - d}{N_h G_h} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}}$

$$\begin{align*}
&B^*_h(d) = 0 \\
P^*_i(d) = \left( \frac{N_h - d}{N_h G_h} \right) \left( \frac{M_h - G_h}{M_h} \right)^2 r_0^{-1} 2^{-\frac{\beta}{r_0}} & \quad \text{if } 0 < R_t \leq \mathcal{R}(d) - M_h \log_2 \frac{M_h}{M_h - 2G_h} 2^{-\frac{\beta}{r_0}}
\end{align*}$$

(22c)
be obtained when Alice and Hugo have the same number of antennas.

Theorem 2: When \( N_s = N_h \), the optimal solution to (21) for \( B_h \) without the integer constraint, and the corresponding maximum jamming power is shown in (22a)–(22c), shown at the bottom of the previous page, for a given number of data streams \( d \), where \( r = \frac{N_h G_h}{2} \) and \( r_0 = \frac{N_s}{2} \).

Proof: By setting the lower bound in (20) equal to \( R_t \) and replacing \( B_a \) by \( B - B_h \), the jamming power can be expressed as

\[
P_i(B_h) = \frac{(N_h - d) M_h}{N_h G_h 2^{\frac{\beta - \beta_h}{r}}} \left( \frac{M_h - G_h 2^{\frac{\beta - \beta_h}{r}}}{M_h} r - 1 \right).
\]

(23)

After relaxing the integer constraints, the optimization problem in (21) can be rewritten in standard form as

\[
P_i^*(d) = \max_{B_h} P_i(B_h)
\]

s.t. \( B - B_h \geq 0 \)
\( B_h \geq 0 \)
\( P_i(B_h) \geq 0 \)
\( P_h - P_i(B_h) \geq 0 \).

We solve this problem in general, but, due to space limitations, only present the results for \( P_h \) sufficiently large. For each case, we present the threshold. The closed-form solution is given in (22a)–(22c).

The optimal \( d^* \) is taken to be the one that leads to the maximum jamming power \( P_i^*(d) \). For the actual feedback link, we search the integer values above and below \( B_h^*(d) \) in (23) to determine the integer bit allocation \( B_h^*(d^*) \) and \( B_a^*(d^*) \) and the actual jamming power \( P_i^*(d^*) \).

If \( R_t > R(d) - M_h \log_2 \frac{M_h}{M_h - G_h 2^{\frac{\beta_h}{r}}} \), the target rate \( R_t \) cannot be achieved and the link is assumed to be in outage.

B. Statistical Eavesdropper CSI

In the last section, we developed a feedback bit allocation strategy for situations with no eavesdropper CSI that provided the maximum possible jamming power to disrupt potential eavesdroppers, subject to the constraint that a lower bound on the average rate of the desired link is above a target level \( R_t \). However, the secrecy rate of such a scheme cannot in general be guaranteed; a well-endowed eavesdropper in the right location could end up with a better quality signal, and secrecy would be lost.

Here we assume that statistical information about Eve’s channel is available. In particular, we assume isotropic distributions for \( H_{sa} \) and \( H_{sh} \), and we investigate maximizing the average secrecy rate in (15) by adjusting both the feedback bit allocation and the amount of transmission power at Alice and Hugo:

\[
\max_{B_s, B_h, P_s, P_h} R_{s\sec}(d)
\]

s.t. \( B_s + B_h = B \)
\( B_s, B_h \in \mathbb{Z}^+ \)
\( 0 \leq P_s \leq P_a \)
\( 0 \leq P_t \leq P_h \).

Our approach first optimizes the secrecy rate over \( B_s, B_h \) for fixed \( P_s, P_h \), and then addresses the power allocation.

Theorem 3: When \( N_s = N_h \), the optimal solution for \( B_h \) without the integer constraint is shown in (25a)–(25c), shown at the bottom of this page, for a given number of data streams \( d \) and fixed jamming power \( P_t \).

1) \( 2G_h < M_h \)

\[
B_h^*(d, P_t) = \begin{cases} 
B & \text{if } P_t > \frac{(N_s - d)M_s}{N_h (M_h - 2G_h)} \\
T_h \log_2 \left( \frac{-(P_s N_s G_s + \sqrt{(P_s N_s G_s)^2 + P_s N_s (N_s - d)M_s^2 2^{\frac{\beta_h}{r}}})}{(N_s - d)M_h} \right) & \text{if } \frac{1}{N_h (M_h - 2G_h)} \leq P_t \leq \frac{1}{N_h M_s 2^{\frac{\beta_h}{r}} - 2G_h} \\
0 & \text{if } \frac{1}{N_h M_s 2^{\frac{\beta_h}{r}} - 2G_h} \leq P_t \leq \frac{1}{N_h (M_h - 2G_h)} \\
0 & \text{if } P_t < \frac{1}{N_h (M_h - 2G_h)}.
\end{cases}
\]

(25a)

2) \( M_h \leq 2G_h < M_h 2^{\frac{\beta_h}{r}} \)

\[
B_h^*(d, P_t) = \begin{cases} 
T_h \log_2 \left( \frac{-(P_s N_s G_s + \sqrt{(P_s N_s G_s)^2 + P_s N_s (N_s - d)M_s^2 2^{\frac{\beta_h}{r}}})}{(N_s - d)M_h} \right) & \text{if } P_t \leq \frac{1}{N_h M_s 2^{\frac{\beta_h}{r}} - 2G_h} \\
0 & \text{if } \frac{1}{N_h M_s 2^{\frac{\beta_h}{r}} - 2G_h} \leq P_t \leq \frac{1}{N_h (M_h - 2G_h)} \\
0 & \text{if } P_t < \frac{1}{N_h (M_h - 2G_h)}.
\end{cases}
\]

(25b)

3) \( 2G_h \geq M_h 2^{\frac{\beta_h}{r}} \)

\[
B_h^*(d, P_t) = 0
\]

(25c)
Proof: Define \( \mathbf{H}_{e_a} \triangleq \mathbf{H}_{e_a} \mathbf{W}_a \) and \( \mathbf{H}_{e_h} \triangleq \mathbf{H}_{e_h} \mathbf{W}_h \). The average secrecy rate is given by

\[
\overline{R}_{sec} (d) = \left\{ \overline{R}_b (d) - \mathbb{E}[\mathbf{H}_{e_a} \mathbf{H}_{e_h}^H \{ R_e (d) \}] \right\}^+ ,
\]

(26)

where

\[
R_e (d) \triangleq \log_2 \left[ \frac{\sigma^2 N_s + \frac{\rho}{N_h - d} \mathbb{E}[\mathbf{H}_{e_a} \mathbf{H}_{e_h}^H \mathbf{H}_{e_h} \mathbf{H}_{e_a}^H]}{\sigma^2 N_s + \frac{\rho}{N_h - d} \mathbb{E}[\mathbf{H}_{e_a} \mathbf{H}_{e_h}^H \mathbf{H}_{e_h} \mathbf{H}_{e_a}^H]} \right] .
\]

Because each quantized precoding matrix is isotropically distributed over the Grassmann manifold, changing the number of codebook entries does not affect its distribution. In addition, the quantized precoding matrices are independent of Eve’s channel. Varying the bit allocations does not affect the probability distribution of \( \mathbf{H}_{e_a} \) and \( \mathbf{H}_{e_h} \), since it is only determined by their dimensions. Thus, the second term in (26) does not depend on \( B_a \) and \( B_h \), and hence the initial optimization problem over \( B_a \) and \( B_h \) is given by

\[
\max_{B_a, B_h} \overline{R}_b (d) \quad \text{s.t.} \quad B_a + B_h = B \quad B_a, B_h \in \mathbb{Z}^+. \]

To solve this problem, we first relax the integer constraint. Then, we use the average link rate lower bound in (20) as an approximation to the above objective function. Since \( \overline{R}_b (d) \) is independent of \( B_a \) and \( B_h \), substituting \( B - B_h \) for \( B_a \), the optimal bit allocation problem is converted to

\[
\min_{d, B_h} \left( 1 + \frac{P_s N_h G_h 2^{-\frac{B_h}{\sigma^2} \rho}}{(N_h - d) M_h} \right) \left( \frac{M_h}{M_a - G_a 2^{-\frac{W_h}{\sigma^2} \rho}} \right) \quad \text{s.t.} \quad B_h - B \leq 0 \quad -B_h \leq 0 .
\]

(27)

This leads to a standard convex optimization problem that can be solved by standard methods. As in the previous case, a closed-form solution can be obtained in (25a)–(25c) when Alice and Hugo have the same number of antennas.

With the integer bit allocation \( \hat{B}_h \) determined (\( \hat{B}_h (d, P_s) \) rounded to the nearest integer), the idea is to substitute it back into the expression for the average secrecy rate so that it can be optimized over the power allocation. This is prohibitively complicated for the general expression of (15), so instead we proceed by substituting the expressions for \( \hat{B}_h \) of Theorem 3 in the lower bound of (20). This leads to the following result.

**Theorem 4:** The average secrecy rate under the optimal bit allocation strategy can be approximately bounded below by

\[
\overline{R}_{sec} (d) \geq \overline{R}_{sec, LB} (d) \]

\[
\triangleq \left\{ dF \left( \frac{l m}{dP_s} \right) - M_h \log_2 \left( 1 + \frac{P_s N_h G_h 2^{-\frac{B_h}{\sigma^2} \rho}}{(N_h - d) M_h} \right) \right. \]

\[
- (N_h - d) \left[ F \left( \frac{N_h}{N_h - d} \frac{P_s}{\sigma^2} \right) - F \left( \frac{N_h}{N_h - d} \frac{P_s}{\sigma^2} \right) \right] - N_h \log_2 \left( 1 + \frac{P_s}{\sigma^2} \right) - M_h \log_2 \left( \frac{M_h}{M_h - G_h 2^{-\frac{W_h}{\sigma^2} \rho}} \right) \}
\]

(28)

**Proof:** See Appendix C.

We assume Bob chooses values for \( P_s, P_i \), and \( d \) to maximize the average secrecy rate lower bound given in (28):

\[
\max_{d \leq P_s \leq P_h} \overline{R}_{sec, LB} (d) .
\]

This can be done by performing a 2-dimensional line search for \( P_s \) and \( P_i \) for each candidate \( d \), and then choosing the value for \( d \) that provides the largest value for the lower bound.

**V. Simulation Results**

In this section, we verify the validity of our analytical results through Monte Carlo simulations. For a given channel realization, we use the numerical results in [27] to randomly generate the associated quantized feedback. Utilizing the statistics of random quantization codebooks, this method simulates the quantization procedure without generating an actual codebook, and reduces the computational complexity as \( B \) grows.

We consider a case where Alice and Hugo have an equal number of antennas, i.e., \( N_a = N_h = 4 \), and we set \( N_h = 2 \). The limited feedback bandwidth \( B \) is 20 bits, and all results are based on averages obtained over 1000 independent channel realizations. While conventional commercial systems employ a coarser feedback quantization, we note that the application considered here is significantly different and much more advanced than those addressed in current systems. Very accurate CSI is required if one is attempting to null a strong interferer and obtain a reasonable secrecy rate, as we are trying to do in this case. In addition, the benefit of the proposed optimization is less apparent for small numbers of feedback bits.

Fig. 2 compares the numerical evaluation of (16) with the approximate lower bound in Theorem 1 as a function of the number of bits allocated to Hugo (\( B_h \)). The figure shows the results for the two possible values of \( d = 1 \) and \( d = 2 \), with transmit powers fixed at \( P_s = 10 \) dB, 20 dB and \( P_i = 10 \) dB. This figure verifies the accuracy of the approximate lower bound especially at the maximum rate where \( B_a \) and \( B_h \) are properly allocated; most importantly, the peaks of the two curves coincide exactly.
In the next few examples, we assume no CSI is available for the eavesdropper, and thus we focus on maximizing the power available for jamming. Fig. 3 shows the optimal bit allocation versus target rate with $P_i = 20$ dB for different $d$. In most cases, the interference leakage term in (20) dominates, and thus Hugo receives a higher allocation of bits than Alice. Note that even for very small target rates, the optimal solution sometimes still allocates a small number of bits for feedback to Alice, since the loss in beamforming gain cannot be compensated for by reduced interference (i.e., when $d = 1$, those 2 remaining bits are more valuable to Alice than they are to Hugo). Fig. 4 shows the optimal jamming power versus target rate that is achieved with $P_i = 20$ dB. When the target rate for the main channel is low, a single data stream is transmitted and the rest of the dimensions are used to interfere with the eavesdropper at higher jamming power. As the target rate increases, a single data stream can no longer meet the rate requirement in the presence of any significant jamming. Thus at this point it is better to switch to $d = 2$ and achieve higher jamming power over the two remaining spatial dimensions. The available jamming power eventually decays to zero as the constraint on the quality of the main link becomes more stringent. Figs. 5 and 6 plot the optimal jamming power as a function of the number of bits allocated to Hugo for small and large values of $R_i$ respectively. It is clear that the proper choice of $B_h$ and $d$ make a significant difference in the amount of jamming power that is available for interfering with any eavesdroppers that are present. Figs. 5 and 6 illustrate the advantage of our results over a non-optimized feedback allocation. Fig. 7 plots the average rate for the main channel (dashed line) together with the target lower bound on the rate (solid line), indicating that the constraint on the lower bound is met in all instances. Also plotted is the rate achieved in the ideal case (dash-dot lines) where there are no quantization errors for $d = 1$ (about 9 bps/Hz) and $d = 2$ (about 14 bps/Hz). The fact that the target rate approaches the ideal rate of 14 bps/Hz when
$d = 2$ is a reflection of the fact that the available jamming power is essentially zero.

For the case where statistical information about Eve’s CSI is available, we show the average secrecy rate as a function of jamming power with $N_e = 2$ in Fig. 8, both for the case with no feedback quantization error (solid lines) and with feedback quantization (dashed lines) assuming the optimal feedback allocation. In this example, the power constraints are 20 dB and 40 dB at Alice and Hugo respectively. This figure illustrates the trade-off associated with the use of a cooperative jammer in a scenario with limited feedback and hence inaccurate CSI; at a certain point the jamming hurts the desired receiver more than it confuses the eavesdropper. In Fig. 9, we add curves that represent the lower bound of the average secrecy rate in Theorem 4, which are used to identify the optimal bit allocation for each value of $d$. The peak values of the lower bound accurately coincide with the peak values of the actual secrecy rate, so use of the lower bound allows the optimal system parameters to be found in the absence of the exact rate. In this case, the peak of the lower bound for $d = 2$ exceeds that for $d = 1$, and leads to the choice of an optimal jamming power of $P^*_j = 16$ dB, indicated by the vertical red line. Fig. 10 demonstrates the accuracy of the average achievable secrecy rate expression with respect to transmit power $P^*_s$. The plot shows that the average secrecy rate achieved by using the optimal results ($P^*_t$, $d^*$ and $B^*_h$) obtained from (29) is essentially identical to the best possible average secrecy rate obtained from Monte Carlo simulations according to (24). This figure also illustrates that transmitting with full power at Alice achieves the maximum average secrecy rate.

VI. CONCLUSION

This paper has considered power and bit allocation strategies for enhanced secrecy in a limited rate feedback MIMO wiretap channel involving a cooperative jammer. We considered two cases, one where no information about the eavesdroppers is available, and one where statistical channel state information is available. With no information about the eavesdroppers, we showed how to choose the allocation of feedback bits to the transmitter and helper in order to maximize the amount of jamming power available to interfere with the eavesdroppers, subject to maintaining a lower bound on the target rate for the desired link. A closed-form solution was found for the special case where the transmitter and jammer have the same number of antennas. For the case of statistical CSI, we derived an approximate lower bound on the average secrecy rate, and again found a closed-form solution for the feedback bit allocation that maximizes this lower bound for an equal number of transmit antennas. Optimization of the transmit power in this case requires a two-dimensional numerical search. Simulation results
indicate the accuracy of the approximations used in the paper, and demonstrate how proper choice of the feedback bit allocation can dramatically enhance the security provided by the cooperative jammer.

**APPENDIX A**

**Proof of Lemma 2**

Using the definition of chordal distance in (8), the average distortion associated with the codebooks \(C_i\) is given by

\[
D_i = E \left[ d^2 \left( V_{ij}, \tilde{V}_{ij} \right) \right]
\]

\[
= E \left[ M_i - tr \left( V_{ij}^H \tilde{V}_{ij} V_{ij} \right) \right]
\]

\[
= M_i - E \left[ tr \left( \tilde{V}_{ij}^H V_{ij} \tilde{V}_{ij} V_{ij} \right) \right]
\]

\[
= M_i - E \left[ tr \left( \tilde{V}_{ij}^H V_{ij} \tilde{V}_{ij} V_{ij} \right) \right].
\]

From the analysis in [27] we know that \(E \left[ \tilde{V}_{ij}^H V_{ij} \tilde{V}_{ij} V_{ij} \right] \) and \(E \left[ \tilde{V}_{ij}^H V_{ij} \tilde{V}_{ij} V_{ij} \right] \) are multiples of the identity matrix. Therefore, we have

\[
E \left[ \tilde{V}_{ij}^H V_{ij} \tilde{V}_{ij} V_{ij} \right] = E \left[ \tilde{V}_{ij}^H V_{ij} V_{ij} \tilde{V}_{ij} \right]
\]

\[
= \frac{E \left[ tr \left( \tilde{V}_{ij}^H V_{ij} \tilde{V}_{ij} V_{ij} \right) \right]}{M_i}
\]

\[
= \left( 1 - \frac{D_i}{M_i} \right) I_{M_i}.
\]

**APPENDIX B**

**Proof of Theorem 1**

First, the average rate of the main channel \(\overline{R}_b(d)\) in (17) can be bounded below by (30)–(32).

\[
\overline{R}_b(d) \geq E_{H_{bs}, H_{ba}, W_h, W_a, w_h} \left[ \log_2 \left( I_d + \frac{P_i}{N_h - d} \Lambda_{h1}^{V_{h1}^H W_h W_h^H V_{h1}} \right) \right]
\]

\[
\geq E_{H_{bs}, H_{ba}, W_h, W_a, w_h} \left[ \log_2 \left( I_d + \frac{P_i}{N_h - d} P_i N_h \Lambda_{h1}^{V_{h1}^H W_h W_h^H V_{h1}} \right) \right]
\]

\[
\geq E_{H_{bs}, H_{ba}, W_h, W_a, w_h} \left[ \log_2 \left( I_d + P_i N_h \Lambda_{h1}^{V_{h1}^H W_h W_h^H V_{h1}} \right) \right].
\]

The first term in (33) is the average achievable rate using ideal beamforming. We define the effective channel from Alice to Bob as \(H_{ba} = W_h^H H_{ba}\). Since \(W_h\) is a submatrix of a unitary matrix, \(H_{ba}\) is also a zero-mean complex Gaussian matrix with i.i.d. elements. However, the elements of \(H_{ba}\) have a variance greater than unity due to the truncation of the \(m - d\) smallest eigenvalues. In order to apply the random matrix result stated in (18), it is necessary to normalize the effective channel \(H_{ba}\) to obtain unit variance elements. Because the exact normalization constant is difficult to obtain analytically, we scale \(H_{ba}\) by an approximate factor \(\sqrt{\frac{P_i}{N_h}}\). In the sequel, this normalization factor is absorbed into the transmit power. Since the quantity \(E_{H_{bs}} \log_2 \left[ I_d + \frac{P_i}{N_h} \Lambda_{h1} \right] \) is approximately given by \(dF\left( \frac{1}{2}, 2 \frac{P_i}{N_h} \right)\), it can be represented as \(\overline{R}_b(d)\).

The second term in (33) is due to interference leakage from Hugo, and we have

\[
-E \log_2 \left( I_d + \frac{P_i}{N_h - d} \Lambda_{h1} V_{h1}^H W_h W_h^H V_{h1} \right)
\]

\[
\geq -E \log_2 \left( I_d + \frac{P_i}{N_h - d} \Lambda_{h1} V_{h1}^H W_h W_h^H V_{h1} \right)
\]

\[
\geq -E \log_2 \left( I_d + \frac{P_i N_h}{N_h - d} \Lambda_{h1} V_{h1}^H W_h W_h^H V_{h1} \right)
\]

\[
= -E \log_2 \left( I_d + \frac{P_i N_h \Lambda_{h1} V_{h1}^H W_h W_h^H V_{h1}}{(N_h - d) M_h} \right).
\]
2) if $M_h = N_h - d$, (35) becomes
\[
\mathbb{E} \left[ \log_2 \left( 1 + \frac{P_t N_h}{N_h - d} \mathbf{V}_{a1}^H \tilde{\mathbf{V}}_{a2} \mathbf{W} \mathbf{W}^H \mathbf{V}_{a1} \right) \right] = \frac{-\sigma_h^2}{2} \approx M_a \log_2 \left( 1 - \frac{D_h}{M_a} \right).
\]

In summary, the approximate lower bound for the average rate of the main channel is given by
\[
\mathbb{E} \left[ \log_2 \left( 1 + \frac{P_t N_h}{N_h - d} \mathbf{V}_{a1}^H \tilde{\mathbf{V}}_{a2} \mathbf{W} \mathbf{W}^H \mathbf{V}_{a1} \right) \right] \geq 2 \left( \text{average secrecy rate using the optimal bit allocation strategy} \right).
\]

In summary, the approximate lower bound for the average rate of the main channel is given by
\[
\mathbb{E} \left[ \log_2 \left( 1 + \frac{P_t N_h}{N_h - d} \mathbf{V}_{a1}^H \tilde{\mathbf{V}}_{a2} \mathbf{W} \mathbf{W}^H \mathbf{V}_{a1} \right) \right] \geq 2 \left( \text{average secrecy rate using the optimal bit allocation strategy} \right).
\]

In summary, the approximate lower bound for the average rate of the main channel is given by
\[
\mathbb{E} \left[ \log_2 \left( 1 + \frac{P_t N_h}{N_h - d} \mathbf{V}_{a1}^H \tilde{\mathbf{V}}_{a2} \mathbf{W} \mathbf{W}^H \mathbf{V}_{a1} \right) \right] \geq 2 \left( \text{average secrecy rate using the optimal bit allocation strategy} \right).
\]
REFERENCES


Xinjie Yang received the B.Eng. degree in electronic information engineering from Central South University, Changsha, China, in 2007. She completed the M.S. degree in electrical and computer engineering from University of California, Irvine, USA, in 2009, and is currently a Ph.D. candidate.

She is affiliated with the Center for Pervasive Communications and Computing, and interned at Broadcom Corporation, Irvine, CA, from 2011–2012 and 2015–2016. Her research focuses on signal processing for wireless communications, with specific interests in physical layer security, limited rate feedback techniques, MIMO systems and cooperative communications.

A. Lee Swindlehurst (F’04) received the B.S., summa cum laude, and M.S. degrees in electrical engineering from Brigham Young University, Provo, Utah, in 1985 and 1986, respectively, and the Ph.D. degree in electrical engineering from Stanford University in 1991. From 1986–1990, he was employed at ESL, Inc., of Sunnyvale, CA, where he was involved in the design of algorithms and architectures for several radar and sonar signal processing systems. He was on the faculty of the Department of Electrical and Computer Engineering at Brigham Young University from 1990–2007, where he was a Full Professor and served as Department Chair from 2003–2006. During 1996–1997, he held a joint appointment as a visiting scholar at both Uppsala University, Uppsala, Sweden, and at the Royal Institute of Technology, Stockholm, Sweden. From 2006–07, he was on leave working as Vice President of Research for ArrayComm LLC in San Jose, California. He is currently the Associate Dean for Research and Graduate Studies in the Henry Samueli School of Engineering at the University of California Irvine (UCI), a Professor of the Electrical Engineering and Computer Science Department at UCI, and a Hans Fischer Senior Fellow in the Institute for Advanced Studies at the Technical University of Munich. His research interests include sensor array signal processing for radar and wireless communications, detection and estimation theory, and system identification, and he has over 275 publications in these areas.

Dr. Swindlehurst is a past Secretary of the IEEE Signal Processing Society, past Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, and past member of the Editorial Boards for the EURASIP Journal on Wireless Communications and Networking, IEEE SIGNAL PROCESSING MAGAZINE, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is a recipient of several paper awards: the 2000 IEEE W. R. G. Baker Prize Paper Award, the 2006 and 2010 IEEE Signal Processing Society’s Best Paper Awards, the 2006 IEEE Communications Society Stephen O. Rice Prize in the Field of Communications Theory, and is co-author of a paper that received the IEEE Signal Processing Society Young Author Best Paper Award in 2001.