Dirty-Paper Coding for the Gaussian Multiaccess Channel With Conferencing

Shraga I. Bross, *Senior Member, IEEE*, Amos Lapidoth, *Fellow, IEEE*, and Michèle Wigger, *Member, IEEE*

***Abstract*—We derive the capacity region of the two-user dirty-paper Gaussian multiaccess channel (MAC) with conferencing encoders. In this MAC, prior to each transmission block, the transmitters can hold a conference in which they can communicate with each other over error-free bit pipes of given capacities. The received signal suffers not only from additive Gaussian noise but also from additive interference, which is known noncausally to the transmitters but not to the receiver. The additive interference is modeled as Gaussian or uniform over a sphere. We show that the interference can be perfectly mitigated, i.e., that the capacity region without interference can also be achieved in its presence. This holds irrespective of whether the transmitters learn the interference before or after the conference. It follows as a corollary that also for the MAC with degraded message sets, the interference can be perfectly mitigated if it is known noncausally to the transmitters. To derive our results, we generalize Costa's single-user writing-on-dirty-paper achievability result to channels with dependent interference and not-necessarily Gaussian noise.**

***Index Terms*—Capacity, coding over spheres, conferencing encoders, interference, multiaccess channel, writing on dirty paper.**

I. Introduction

We consider a multiaccess channel (MAC), where two transmitters wish to communicate with a common receiver. Prior to each transmission block, the transmitters can hold a *conference*, i.e., they can communicate with each other over noise-free bit pipes of given capacities. Special cases of this setting are the classical MAC where the transmitters are completely unaware of each other's message—corresponding to bit pipes of zero capacity; the fully cooperative MAC—corresponding to bit pipes of infinite capacity; and the MAC with degraded message sets, where one of the encoders is fully cognizant of the message the other encoder wishes to send—corresponding to the pipe from the cognizant transmitter to the noncognizant transmitter being of zero capacity and the other pipe being of infinite capacity.

Our setting is known as *MAC with conferencing encoders*. It was introduced by Willems [24], who also derived its capacity region for discrete memoryless channels. The capacity region of the Gaussian MAC with conferencing encoders was presented in [1].

In this paper, we consider the Gaussian MAC with conferencing encoders where the received signal is also corrupted by independent additive interference that is known noncausally to both transmitters but not to the receiver. The interference is assumed to be uniformly distributed over a sphere. For this scenario, we derive a result in the spirit of Costa [3], i.e., we show that, although the receiver is not informed of the interference, the capacity region of the network with interference equals the capacity region of the network without interference. This result holds irrespective of whether the transmitters learn the interference before or after the conference. Moreover, it also holds when—instead of being uniformly distributed over a sphere—the interference is memoryless and Gaussian. Therefore, our result recovers as a special case Gel'fand and Pinsker's writing-on-dirty-paper result for the Gaussian MAC (without conferencing) [7]. The result also yields a writing-on-dirty-paper result for the MAC with degraded message sets where one transmitter knows both messages and the other only one.

Our achievability results utilize the Cohen and Lapidoth single-user writing-on-dirty-paper scheme [2]. As in their work, our achievability results do not depend on the Gaussianity of the noise but only on its ergodicity. Thus, our achievability results hold for any stationary and ergodic (not-necessarily Gaussian) noise of second moment N.

To describe our coding scheme, it is useful to think about Transmitter 1's message as comprising two parts: Transmitter 1's private message and its common message. Likewise for Transmitter 2. We refer to the pair comprising Transmitter 1's common message and Transmitter 2's common message as "the common message."

Our scheme consists of two stages. In the first stage, the transmitters use the bit pipes as in Willems's scheme [24] to exchanges messages. During this stage, Transmitter 1 reveals its common message to Transmitter 2, and Transmitter 2 reveals its common message to Transmitter 1. At the end of this stage, the common message is thus known to both transmitters; Transmitter 1's private message is known only to Transmitter 1, and Transmitter 2's private message is known only to Transmitter 2. Since in this stage, the transmitters do not use their knowledge of the interference sequence, our achievability results hold irrespective of whether the transmitters learn the interference before or after the conference.

In the second stage, the transmitters send the common message and the private messages over the MAC to the receiver. The transmission at this stage is as follows. To convey the common message, the transmitters agree, prior to communication, on an encoding scheme *à la* Cohen and Lapidoth but properly scaled

Manuscript received January 14, 2011; revised September 01, 2011; accepted March 20, 2012. Date of publication June 12, 2012; date of current version August 14, 2012. A. Lapidoth and M. Wigger were supported in part by the Swiss National Science Foundation under Grant 200021-111863/1. This paper was presented in part at the 2008 IEEE International Symposium on Information Theory.

S. I. Bross is with the Department of Electrical Engineering, Bar-Ilan University, Ramat Gan 52900, Israel (e-mail: brosss@biu.ac.il).

A. Lapidoth is with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland (e-mail: lapidoth@isi.ee.ethz.ch).

M. Wigger was with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland. She is now with the Communications and Electronics Department, Telecom ParisTech, 75013 Paris, France (e-mail: michele.wigger@telecom-paristech.fr).

Communicated by E. Erkip, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2012.2202210

to account for the power that each of them allocates to the transmission of the common message. To convey their private messages, they choose independent Cohen and Lapidoth encoding rules. Each transmitter encodes its private message and encodes the common message using the above schemes and then transmits a linear combination of the schemes' outputs. Since the transmitters use the same rule to encode the common message (except for some scaling), the channel combines the codewords corresponding to the common message coherently.

The receiver performs successive decoding and stripping: first it decodes the common message, and then it decodes the private messages. The order in which it decodes the private messages depends on the rate-pair. In each of the successive decoding steps, the receiver uses a nearest neighbor decoding rule *à la* Cohen and Lapidoth to decode its desired message and treats the effect of the messages it has not yet decoded as "additional noise." This "additional noise" is, alas, not Gaussian. Trickier still is that the way the as-of-yet-undecoded messages affect the received signal depends on the interference, so the "additional noise" and the interference are not independent. To analyze our scheme, we thus need to generalize the Cohen and Lapidoth achievability result to the case where the noise can depend (in a controlled way) on the interference.

The single-user dirty-paper channel with dependent noise and interference has been previously studied in the case where the interference and the noise are jointly Gaussian [12] and of arbitrary correlation. But for our analysis, we must study the case where the noise is not Gaussian. Our analysis thus treats channels with non-Gaussian noise, but where the interference and the noise are nearly orthogonal with very high probability (for sufficiently large blocklengths). Thus, neither of the two results implies the other. In fact, the tools used to derive the two results could be combined to treat more general channels.

Related multiaccess setups with conferencing encoders have recently been studied in [23] and [13].[1] In these works, it is assumed that the output is corrupted by *two* interference sequences, each of which is known to a different transmitter. These results show that when the two transmitters know different parts of the interference, then during the conferencing phase they should also exchange information about the interference. Unlike our scenario, in these scenarios, the capacity region depends on whether the transmitters learn the interference before or after the conferencing. Multiaccess setups without conferencing where both transmitters know only parts of the interference were also studied in [9], [10], [14]–[16], [19]–[21], and [27].

The rest of this paper is organized as follows. In Section II, we introduce some notation. In Section III, we describe the dirty-paper MAC with conferencing encoders in detail; we present the capacity region of this channel; and we prove (also based on the scheme in Section VI) our capacity result. In Section IV, we describe the generalized single-user dirty-paper channel and present our results pertaining to this channel. In Section V, we recall the single-user dirty-paper scheme from [2] and extend its analysis. Finally, in Section VI, we describe and analyze a capacity-achieving scheme for the dirty-paper MAC with conferencing encoders.

[1]The work in [13] assumes that also the receiver is cognizant of the interferences.

II. Notation

Random variables are denoted by capital letters and their realizations by lowercase letters. Vectors are denoted by bold letters: random vectors by uppercase bold letters, and deterministic vectors by lowercase bold letters. Sets and events are denoted by calligraphic letters. The notation A^n stands for the n-tuple $(A_1, \ldots, A_n)$. The transpose of a vector $\mathbf{a}$ is denoted by $\mathbf{a}^{\mathsf{T}}$; its Euclidean norm by $\|\mathbf{a}\|$; and the Euclidean inner product of two vectors $\mathbf{a}$ and $\mathbf{b}$ by $\langle \mathbf{a}, \mathbf{b} \rangle$. The set of real numbers is denoted by $\mathbb{R}$ and its n-fold Cartesian product by $\mathbb{R}^n$; the set of positive integers is denoted by $\mathbb{N}$. For a real number $a \in [0, 1]$, we use $\bar{a}$ to denote $\bar{a} \stackrel{\triangle}{=} 1 - a$. Throughout the paper, $\log(\cdot)$ denotes the logarithm to the base 2.

An *n-sphere of radius $r > 0$ centered at $\xi \in \mathbb{R}^n$* is the set of all vectors $\mathbf{x} \in \mathbb{R}^n$ satisfying

$$\|\mathbf{x} - \xi\| = r.$$

When the center of the sphere ξ is the origin $\mathbf{0}$, we call it a *centered* sphere, and when the radius of the sphere is 1, we call it a *unit* sphere.

For every vector μ on the centered unit n-sphere, the *spherical cap of half-angle θ centered at μ* is the set of all vectors $\mathbf{x}$ on the centered unit n-sphere satisfying

$$\langle \mathbf{x}, \mu \rangle \geq \cos \theta.$$

The surface area of such a spherical cap does not depend on the vector μ but only on the dimension n and the angle θ. We denote it by $C_n(\theta)$.

We say that a random n-vector is uniformly distributed over an n-sphere, if it is drawn according to a uniform probability measure over the surface of this sphere.

III. Dirty-Paper MAC With Conferencing Encoders: Setting and Results

A. Setting

Two transmitters 1 and 2 wish to convey their messages M_1 and M_2 to a common receiver. The messages M_1 and M_2 are independent and uniformly distributed over the finite sets $\mathcal{M}_1$ and $\mathcal{M}_2$. The communication takes place over a Gaussian MAC with a single additive interference that is known noncausally to both transmitters but not to the receiver. Thus, the time-t channel output Y_t corresponding to the time-t channel inputs $x_{1,t}, x_{2,t} \in \mathbb{R}$ is

$$Y_t = x_{1,t} + x_{2,t} + S_t + Z_t \tag{1}$$

where $\{Z_t\}$ is the Gaussian noise corrupting the channel, and $\{S_t\}$ is the interference. The noise sequence $\{Z_t\}$, the interference sequence $\{S_t\}$, and the messages M_1 and M_2 are independent. The noise $\{Z_t\}$ is a sequence of independent and identically distributed (IID) zero-mean variance-$N > 0$ Gaussian random variables. Denoting the blocklength by n, we can describe the distribution of the interference as follows: we stack

n interference symbols $S_1, \ldots, S_n$ into a column-vector $\mathbf{S} = (S_1, \ldots, S_n)^{\mathsf{T}}$, and we assume that this vector $\mathbf{S}$ is uniformly distributed over a centered n-sphere of radius $\sqrt{nQ} > 0$. We refer to Q as the *interference variance*. As we shall see, our results continue to hold in the more realistic case where $\{S_t\}$ are IID zero-mean, variance-Q Gaussian random variables.

Both transmitters learn the interference noncausally. Thus, prior to each block of n channel uses, they learn the sequence $S^n \triangleq (S_1, \ldots, S_n)$ of the next n interference symbols.

The two transmitters are allowed to cooperate in the following way. Prior to each block of n channel uses, the transmitters can hold a *conference*, i.e., they can exchange information over k uses of two pipes: a pipe from Transmitter 1 to Transmitter 2 and a pipe from Transmitter 2 to Transmitter 1. The pipes are assumed to be

1) perfect in the sense that any input symbol to a pipe is available immediately and error-free at the output of the pipe; and
2) of throughputs C_{12} and C_{21}, in the sense that when the κ inputs to the pipe from Transmitter 1 to Transmitter 2 take values in the sets $\mathcal{V}_{1,1}, \ldots, \mathcal{V}_{1,\kappa}$ and the κ inputs to the pipe from Transmitter 2 to Transmitter 1 take values in the sets $\mathcal{V}_{2,1}, \ldots, \mathcal{V}_{2,\kappa}$, then

$$\sum_{k=1}^{\kappa} \log |\mathcal{V}_{1,k}| \leq nC_{12} \tag{2}$$

and

$$\sum_{k=1}^{\kappa} \log |\mathcal{V}_{2,k}| \leq nC_{21}. \tag{3}$$

The communication over the pipes is assumed to be held in a conferencing way, that is, the kth inputs $V_{1,k} \in \mathcal{V}_{1,k}$ and $V_{2,k} \in \mathcal{V}_{2,k}$ can depend on the respective messages, the interference sequence S^n, and the past observed pipe-outputs

$$V_{1,k} = f_{1,k}\left(M_1, S^n, V_{2,1}, \ldots, V_{2,k-1}\right)$$
$$V_{2,k} = f_{2,k}\left(M_2, S^n, V_{1,1}, \ldots, V_{1,k-1}\right)$$

for some given sequences of encoding functions $\{f_{1,k}\}_{k=1}^{\kappa}$ and $\{f_{2,k}\}_{k=1}^{\kappa}$ where

$$f_{1,k} : \mathcal{M}_1 \times \mathbb{R}^n \times \mathcal{V}_{2,1} \times \cdots \times \mathcal{V}_{2,k-1} \longrightarrow \mathcal{V}_{1,k} \tag{4}$$
$$f_{2,k} : \mathcal{M}_2 \times \mathbb{R}^n \times \mathcal{V}_{1,1} \times \cdots \times \mathcal{V}_{1,k-1} \longrightarrow \mathcal{V}_{2,k}. \tag{5}$$

We define an (n, C_{12}, C_{21})-*conference* to be the collection of an integer number κ, two sets of input alphabets $\{\mathcal{V}_{1,1}, \ldots, \mathcal{V}_{1,\kappa}\}$ and $\{\mathcal{V}_{2,1}, \ldots, \mathcal{V}_{2,\kappa}\}$, and two sets of encoding functions $\{f_{1,1}, \ldots, f_{1,\kappa}\}$ and $\{f_{2,1}, \ldots, f_{2,\kappa}\}$ as in (4) and (5), where $n, C_{12}, C_{21}, \kappa$, and the sets $\{\mathcal{V}_{1,1}, \ldots, \mathcal{V}_{1,\kappa}\}$ and $\{\mathcal{V}_{2,1}, \ldots, \mathcal{V}_{2,\kappa}\}$ satisfy (2) and (3).

After the conference, Transmitter 1 is cognizant of its message M_1, the symbols $V_2^{\kappa} = (V_{2,1}, \ldots, V_{2,\kappa})$, and the interference sequence S^n. Similarly, Transmitter 2 is cognizant of its message M_2, the symbols $V_1^{\kappa} = (V_{1,1}, \ldots, V_{1,\kappa})$, and the interference sequence S^n. The channel input sequences $X_1^n = (X_{1,1}, \ldots, X_{1,n})$ and $X_2^n = (X_{2,1}, \ldots, X_{2,n})$ can be described as

$$X_1^n = \varphi_1^{(n)}(M_1, V_2^{\kappa}, S^n)$$
$$X_2^n = \varphi_2^{(n)}(M_2, V_1^{\kappa}, S^n)$$

where

$$\varphi_1^{(n)} : \quad \mathcal{M}_1 \times \mathcal{V}_{2,1} \times \cdots \times \mathcal{V}_{2,\kappa} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \tag{6}$$
$$\varphi_2^{(n)} : \quad \mathcal{M}_2 \times \mathcal{V}_{1,1} \times \cdots \times \mathcal{V}_{1,\kappa} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n. \tag{7}$$

We only allow encoding functions $\varphi_1^{(n)}$ and $\varphi_2^{(n)}$ that with probability 1 satisfy

$$\frac{1}{n}\|\varphi_1^{(n)}(M_1, V_2^{\kappa}, S^n)\|^2 \leq P_1 \tag{8}$$
$$\frac{1}{n}\|\varphi_2^{(n)}(M_2, V_1^{\kappa}, S^n)\|^2 \leq P_2. \tag{9}$$

The decoder applies a decoding function

$$\phi^{(n)} : \mathbb{R}^n \to \mathcal{M}_1 \times \mathcal{M}_2 \tag{10}$$

to produce the message estimates $\hat{M}_1$ and $\hat{M}_2$ based on its observed output sequence $Y^n = (Y_1, \ldots, Y_n)^{\mathsf{T}}$

$$(\hat{M}_1, \hat{M}_2) = \phi^{(n)}(Y^n).$$

An error occurs in the transmission whenever $(M_1, M_2) \neq (\hat{M}_1, \hat{M}_2)$.

This leads us to the definition of achievable rate pairs and capacity region. A blocklength n, powers (P_1, P_2) code of rate pair $\left(\frac{1}{n}\log|\mathcal{M}_1|, \frac{1}{n}\log|\mathcal{M}_2|\right)$ is a triple $\left(\varphi_1^{(n)}, \varphi_2^{(n)}, \phi^{(n)}\right)$, where $\varphi_1^{(n)}$ and $\varphi_2^{(n)}$ are of the form (6) and (7) and satisfy the power constraints (8) and (9), and where $\phi^{(n)}$ is of the form (10). We say that a rate pair (R_1, R_2) is *achievable* if for every $\delta > 0$ and every sufficiently large blocklength n, there exists an (n, C_{12}, C_{21})-conference and a blocklength n, powers (P_1, P_2) code of rates exceeding $(R_1 - \delta, R_2 - \delta)$ such that the average probability of decoding error tends to 0 as n tends to infinity, i.e.,

$$\lim_{n\to\infty} \Pr\left[(M_1, M_2) \neq (\hat{M}_1, \hat{M}_2)\right] = 0. \tag{11}$$

The *capacity region* is defined as the set of all achievable rate pairs and is denoted by $C_{\mathrm{Conf,WDP}}(P_1, P_2, N, C_{12}, C_{21})$, or by $C_{\mathrm{Conf,WDP}}$ for short.

Our setting includes various classical communication scenarios as special cases.

1) When $Q = 0$, the setup coincides with the Gaussian MAC with conferencing encoders without interference. The capacity region in this special case is denoted by C_{Conf} and was derived in [1].
2) When $C_{12} = C_{21} = \infty$, the setting is equivalent—in terms of achievable rates—to a *fully cooperative* dirty-paper MAC where both transmitters are cognizant of both messages M_1 and M_2 but cannot hold a conference. The equivalence can be seen as follows. Every scheme designed for the fully cooperative MAC exhibits the same performance on the MAC with conferencing encoders and

$C_{12} = C_{21} = \infty$ if in this latter setting, prior to transmission over the MAC, the two encoders exchange their messages over the pipes of infinite capacities. On the other hand, every scheme designed for the MAC with conferencing encoders exhibits the same performance on the fully cooperative MAC if in the latter setting the two transmitters simulate the conference by letting each transmitter compute its corresponding pipe outputs based on M_1 and M_2.

3a) When $C_{12} = C_{21} = 0$, the setting is equivalent to the dirty-paper MAC without conferencing. The capacity region of this special case is denoted by $C_{\mathrm{MAC,WDP}}$ and (in the case of a Gaussian interference) was derived by Gel'fand and Pinsker [7].

3b) When $C_{12} = C_{21} = 0$ and $Q = 0$, the setting is equivalent to the classical Gaussian MAC without conferencing and without interference. Its capacity is denoted by C_{MAC} and was reported in [5] and [26].

4a) When $C_{12} = 0$ and $C_{21} = \infty$, the setting is equivalent—again in terms of achievable rates—to a dirty-paper MAC with *degraded message sets* where Transmitter 1 is cognizant of both messages M_1 and M_2 and Transmitter 2 only of Message M_2 and where the transmitters cannot conference. The equivalence follows by similar arguments as the equivalence in case 2).

4b) When $C_{12} = 0, C_{21} = \infty$, and $Q = 0$, then the setting is equivalent to a Gaussian MAC with degraded message sets without conferencing and without interference.

B. Results

Our first result determines the capacity region of the dirty-paper MAC with conferencing encoders. The capacity region is stated after the following definition and a theorem from [1].

Definition III.1: For all parameters $P_1, P_2, N > 0$ and $C_{12}, C_{21} \geq 0$, define the region $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(P_1, P_2, N, C_{12}, C_{21})$ (or $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}$ for short) as

$$\begin{aligned}
&\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(P_1, P_2, N, C_{12}, C_{21}) \\
&\triangleq \bigcup_{0 \leq \beta_1, \beta_2 \leq 1} \Bigg\{ (R_1, R_2) : \\
&\qquad R_1 \leq \frac{1}{2} \log\left(1 + \frac{\beta_1 P_1}{N}\right) + C_{12} \\
&\qquad R_2 \leq \frac{1}{2} \log\left(1 + \frac{\beta_2 P_2}{N}\right) + C_{21} \\
&\quad R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \frac{\beta_1 P_1 + \beta_2 P_2}{N}\right) + C_{12} + C_{21} \\
&\quad R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\beta}_1 \bar{\beta}_2 P_1 P_2}}{N}\right) \Bigg\}
\end{aligned} \tag{12}$$

where $\bar{\beta}_1 = 1 - \beta_1$ and $\bar{\beta}_2 = 1 - \beta_2$.

Theorem III.2 ([1, Th. 1]): The capacity region C_{Conf} of the Gaussian MAC with conferencing encoders without interference is $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}$

$$\begin{aligned}
&C_{\mathrm{Conf}}(P_1, P_2, N, C_{12}, C_{21}) \\
&\quad = \mathcal{R}_{\mathrm{Conf},\mathcal{G}}(P_1, P_2, N, C_{12}, C_{21}).
\end{aligned}$$

The main result of this paper is that—irrespective of the interference variance Q—if the interference is known noncausally to both transmitters, then the capacity is the same as if there were no interference.

Theorem III.3: The capacity region $C_{\mathrm{Conf,WDP}}$ of the dirty-paper MAC with conferencing encoders equals that of the Gaussian MAC with conferencing encoders without interference

$$\begin{aligned}
&C_{\mathrm{Conf}}(P_1, P_2, N, C_{12}, C_{21}) \\
&\quad = C_{\mathrm{Conf,WDP}}(P_1, P_2, N, C_{12}, C_{21}, Q).
\end{aligned}$$

Proof: See Section III-C. ■

The scheme that we propose in Section VI (see also Lemmas III.10 and III.11) for achieving $C_{\mathrm{Conf,WDP}}$ does not require that the interference sequence be known before the conference begins; it suffices that it be known thereafter. Consequently:

Remark III.4: The region $C_{\mathrm{Conf,WDP}}$ is achievable even if the transmitters learn the interference only after the conferencing phase, so Theorem III.3 remains valid also in this setup.

To simplify the analysis, we have assumed that the interference sequence $\mathbf{S}$ is uniformly distributed over the centered sphere of radius $\sqrt{nQ}$, where n is the blocklength. But our results are also applicable to the case where the components of $\mathbf{S}$ are IID Gaussians of mean zero and variance Q:

Remark III.5: The result of Theorem III.3 also holds when the interference sequence $S_1, \ldots, S_n$ is IID Gaussian.

Proof: The proof of the converse to Theorem III.3 (see Section III-C) does not depend on the interference's law, so the only modification required is in the proof of the direct part. The modification (inspired by [2]) is that—having observed the IID Gaussian interference $\mathbf{S}$—the transmitters would produce their sequences (as in Section VI) as though the interference were $\sqrt{nQ}\mathbf{S}/\|\mathbf{S}\|$ but that one of the transmitters, say Transmitter 1, would then add $\sqrt{nQ}\mathbf{S}/\|\mathbf{S}\| - \mathbf{S}$ to its sequence. This will cause the channel to appear as though the interference were $\sqrt{nQ}\mathbf{S}/\|\mathbf{S}\|$ and hence uniformly distributed over the centered sphere of radius $\sqrt{nQ}$. Of course, adding $\sqrt{nQ}\mathbf{S}/\|\mathbf{S}\| - \mathbf{S}$ to its sequence might increase the transmitted power, but the additional power is negligible because for $S_1, \ldots, S_n$ IID zero-mean variance-Q Gaussians

$$\lim_{n\to\infty} \frac{1}{n} \mathsf{E}\left[\left\|\sqrt{nQ}\mathbf{S}/\|\mathbf{S}\| - \mathbf{S}\right\|^2\right] = 0.$$

■

Specializing Theorem III.3 to the case $C_{12} = C_{21} = 0$ without conferencing and combining the result with Remark III.5, recovers Gel'fand and Pinsker's dirty-paper result for the Gaussian MAC [7]. Similarly, specializing Theorem III.3 to the case $C_{12} = 0$ and $C_{21} = \infty$ results in an analogous result for the MAC with degraded message sets.

Corollary III.6 (Degraded Message Sets): The capacity region of the dirty-paper MAC with degraded message sets is the same as if the interference were not present.

Our proof that the region $C_{\mathrm{Conf,WDP}}$ is achievable (see Lemmas III.10 and III.11 and the coding scheme in Section VI) does not rely on the Gaussianity of the noise sequence $\{Z_t\}$. It suffices that it be stationary and ergodic and of second moment N.

Remark III.7: The achievability results in Theorem III.3 and Corollary III.6 hold for arbitrary stationary and ergodic noise processes of second moment N. In particular, Gel'fand and Pinsker's writing-on-dirty-paper result for the MAC holds also for such general noise processes.

Remark III.8: Our results in Theorem III.3 and Remarks III.4 and III.5 extend also to the two-user dirty-paper MAC with common and private messages (without conferencing) [8], [18]. In particular, we can recover the result in [8] that for this setup with common and private messages, and in the Gaussian case the capacity region with interference is the same as without.

C. On the Proof of Theorem III.3

The converse

$$C_{\mathrm{Conf,WDP}} \subseteq C_{\mathrm{Conf}} \tag{13}$$

follows because C_{Conf} contains the capacity region of the channel with interference even when the interference is known also to the receiver. Indeed, in this case, the receiver can subtract the interference from the channel output, thus reducing the channel to one without interference and with $\mathbf{S}$ known to all parties. In this new setup, the interference sequence $\mathbf{S}$ is independent of the messages and of the channel law, and it therefore only plays the role of a common randomness at the transmitters and the receiver, and common randomness does not increase the capacity of the Gaussian MAC with conferencing encoders.

The direct part

$$C_{\mathrm{Conf,WDP}} \supseteq C_{\mathrm{Conf}} \tag{14}$$

follows from Lemmas III.10 and III.11 ahead. Before stating these lemmas, we define a region $\mathcal{R}_{\mathrm{Ach}}(P_1, P_2, N, C_{12}, C_{21})$ through rate constraints similar to those defining $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(P_1, P_2, N, C_{12}, C_{21})$ but with two additional constraints, so

$$\mathcal{R}_{\mathrm{Ach}}(P_1, P_2, N, C_{12}, C_{21}) \subseteq \mathcal{R}_{\mathrm{Conf},\mathcal{G}}(P_1, P_2, N, C_{12}, C_{21}).$$

Nevertheless, as we shall see, this set is rich enough in the sense that its convex hull contains $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(P_1, P_2, N, C_{12}, C_{21})$, i.e., the set of all rate pairs that are achievable in the absence of interference (see Theorem III.2).

Definition III.9: Given $P_1, P_2, N > 0$ and $C_{12}, C_{21} \geq 0$, define $\mathcal{R}_{\mathrm{Ach}}(P_1, P_2, N, C_{12}, C_{21})$ (or $\mathcal{R}_{\mathrm{Ach}}$ for short) as in (15), shown at the bottom of the page, where recall that $\bar{\beta}_1 = 1 - \beta_1$ and $\bar{\beta}_2 = 1 - \beta_2$.

Lemma III.10: The convex hull of $\mathcal{R}_{\mathrm{Ach}}$ equals C_{Conf}

$$\begin{aligned}&\mathrm{conv}\left(\mathcal{R}_{\mathrm{Ach}}(P_1, P_2, N, C_{12}, C_{21})\right)\\ &\quad = C_{\mathrm{Conf}}(P_1, P_2, N, C_{12}, C_{21}).\end{aligned}$$

Proof: In Appendix A, we show that $\mathrm{conv}\left(\mathcal{R}_{\mathrm{Ach}}\right) = \mathcal{R}_{\mathrm{Conf},\mathcal{G}}$, and hence by Theorem III.2 also $\mathrm{conv}\left(\mathcal{R}_{\mathrm{Ach}}\right) = C_{\mathrm{Conf}}$. ■

Lemma III.11: Irrespective of the interference variance Q, the region $\mathrm{conv}\left(\mathcal{R}_{\mathrm{Ach}}\right)$ is achievable for the dirty-paper MAC with conferencing encoders:

$$\begin{aligned}&C_{\mathrm{Conf,WDP}}(P_1, P_2, N, C_{12}, C_{21}, Q)\\ &\quad \supseteq \mathrm{conv}\left(\mathcal{R}_{\mathrm{Ach}}(P_1, P_2, N, C_{12}, C_{21})\right).\end{aligned}$$

$$\begin{aligned}&\mathcal{R}_{\mathrm{Ach}}(P_1, P_2, N, C_{12}, C_{21})\\ &\quad \triangleq \bigcup_{0\leq\beta_1,\beta_2\leq 1} \Bigg\{(R_1, R_2) : R_1 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1}{N}\right) + C_{12} && \text{(15a)}\\ &\qquad R_1 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1}{N}\right) + \frac{1}{2}\log\left(1 + \frac{(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2})^2}{\beta_1 P_1 + \beta_2 P_2 + N}\right) && \text{(15b)}\\ &\qquad R_2 \leq \frac{1}{2}\log\left(1 + \frac{\beta_2 P_2}{N}\right) + C_{21} && \text{(15c)}\\ &\qquad R_2 \leq \frac{1}{2}\log\left(1 + \frac{\beta_2 P_2}{N}\right) + \frac{1}{2}\log\left(1 + \frac{(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2})^2}{\beta_1 P_1 + \beta_2 P_2 + N}\right) && \text{(15d)}\\ &\qquad R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1 + \beta_2 P_2}{N}\right) + C_{12} + C_{21} && \text{(15e)}\\ &\qquad R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \Bigg\} && \text{(15f)}\end{aligned}$$

Proof: A scheme achieving all rates in the region $\mathcal{R}_{\mathrm{Ach}}$ is described and analyzed in Section VI. ■

IV. Generalized Single-User Dirty-Paper Channel: Setting and Result

A. Setting

In the following two sections, we consider a single-user channel where a transmitter wishes to send a message M, which is uniformly distributed over the finite set $\mathcal{M}$. The communication takes place over an additive-interference additive-noise channel whose time-t output Y_t corresponding to its time-t input x_t is

$$Y_t = x_t + S_t + Z_t$$

where $\{S_t\}$ denotes the random interference, and $\{Z_t\}$ denotes the random noise. The interference is known noncausally to the transmitter. Thus, for a given blocklength n, the transmitter learns the entire interference sequence $S_1, \ldots, S_n$ before the transmission begins. To state our assumptions, we stack the interference symbols in a column vector

$$\mathbf{S} = (S_1, \ldots, S_n)^{\mathsf{T}}$$

and the noise symbols in a column vector

$$\mathbf{Z} = (Z_1, \ldots, Z_n)^{\mathsf{T}}.$$

We assume that the pair $(\mathbf{S}, \mathbf{Z})$ is independent of the message M, and that—as in the previous section—$\mathbf{S}$ is uniformly distributed over the centered n-sphere of radius $\sqrt{nQ}$. Our setup differs from the classical dirty-paper channel [3] in two important ways: the noise $\{Z_t\}$ need not be Gaussian, and it may depend on the interference.

The channel inputs are subject to a block-power constraint

$$\frac{1}{n}\sum_{t=1}^{n} X_t^2 \le P \quad \text{with probability one.}$$

The rate of transmission R is defined as $\frac{1}{n}\log(|\mathcal{M}|)$. A rate $R > 0$ is said to be achievable if for every $\delta, \epsilon > 0$ and for all sufficiently large blocklengths n (depending on δ and ϵ), there exist encoding and decoding rules such that a message of rate exceeding $R - \delta$ can be sent with probability of error smaller than ϵ.

B. Results

Theorem IV.1: A rate $R > 0$ is achievable for the generalized single-user dirty-paper channel with parameters P and Q, and noise $\mathbf{Z}$ whenever there exists some $N > 0$ such that for every sufficiently small $\delta > 0$ there exists some $\epsilon^* \in (0, N)$ (depending on δ) that meets the following three conditions:

1) ϵ^* is small enough so that

$$\frac{P + \alpha Q}{\sqrt{P + \alpha^2 Q}\sqrt{P + N + Q + (3 - 2\alpha)\epsilon^*}} \ge \sqrt{1 - 2^{-2(R-\delta)}\frac{P}{P + \alpha^2 Q}} \tag{16}$$

where $\alpha \stackrel{\triangle}{=} \frac{P}{P+N}$;

2) the empirical second moment of the noise satisfies

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\|\mathbf{Z}\|^2 - N\right| \le \epsilon^*\right] = 1; \tag{17}$$

3) and for every $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left[-\epsilon \le \frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \le \epsilon^*\right] = 1. \tag{18}$$

Proof: Based on the scheme in Section V and on its analysis in Lemma V.1 and Remark V.3. For details, see Appendix D. ■

As a corollary, we obtain the following "Gaussian-is-the-worst-noise"-result similar to the results in [11, Th. 1] and [2, Theorem 2.3].

Corollary IV.2: If for some $N > 0$ and every $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\|\mathbf{Z}\|^2 - N\right| \le \epsilon\right] = 1 \tag{19}$$

and

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle\right| \le \epsilon\right] = 1 \tag{20}$$

then the rate

$$\frac{1}{2}\log\left(1 + \frac{P}{N}\right)$$

is achievable.

Proof: The corollary's hypothesis implies that Conditions (17) and (18) in Theorem IV.1 are satisfied for all $\epsilon, \epsilon^* > 0$. Consequently, it suffices to show that for $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$ and every sufficiently small $\delta > 0$ we can find some $\epsilon^* \in (0, N)$ so that Condition (16) is satisfied. For a fixed $\delta > 0$ the existence of such an ϵ^* follows from the following observations.

1) The right-hand side of (16) does not depend on ϵ^* and is continuous and decreasing in δ.
2) The left-hand side of (16) does not depend on δ and is continuous in ϵ^*.
3) For $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$, $\delta = 0$, and $\epsilon^* = 0$ both sides of (16) coincide. ■

Remark IV.3: Every stationary and ergodic process $\{Z_t\}$ of second moment N that is independent of the interference $\{S_t\}$ satisfies the corollary's hypothesis. Thus, Corollary IV.2 recovers the writing-on-dirty-paper result by Cohen and Lapidoth [2] for channels with stationary and ergodic noise that is independent of the interference.

Using the same argument, we used to prove Remark III.5 we obtain

Remark IV.4: Theorem IV.1 applies also when the interference sequence $S_1, \ldots, S_n$ is IID Gaussian.

V. Generalized Single-User Dirty-Paper Channel: Coding Scheme and Analysis

In this section, we consider the generalized single-user dirty-paper channel of Section IV. Its parameters are

$$P, Q, \text{ and the law of } \mathbf{Z} \tag{21}$$

where nP is the maximal allowed energy of each codeword, where the state vector $\mathbf{S}$ is drawn uniformly at random over the centered n-sphere of radius $\sqrt{nQ}$, and where $\mathbf{Z}$ denotes the noise vector.

We shall briefly recall the scheme in [2] and extend the scheme's decoding rule. The scheme in [2] is based on coding over spheres and on nearest neighbor decoding. This is different from the scheme by Gel'fand and Pinsker in [6] which uses IID random codebooks and a joint-typicality decoder. Notice however that Gel'fand and Pinsker's scheme [6] does not directly apply to Gaussian dirty-paper setups (single-user or multiuser) because it is based on strong typicality. Replacing strong typicality with weak typicality in the description and the analysis of the Gel'fand and Pinsker scheme does not resolve this problem, because the conditional typicality lemma ([6, Lemma 2]) does not hold under weak typicality and because—even when averaged over codebooks and the random interference—the channel input sequence produced by this modified scheme is not Gaussian. Analyzing the weak-typicality decoder for Gaussian dirty-paper channels requires additional geometric arguments.

The approach based on coding over spheres and on nearest neighbor decoding considered here has simpler geometric arguments in the analysis and allows for more general results than the Gaussian codebooks and joint-typicality decoding approach of [6]. In fact, it allows us to generalize the analysis in [2] to the case where the noise $\mathbf{Z}$ and interference $\mathbf{S}$ need not be independent.

The extended decoding rule and the generalized analysis presented in this section will be useful when we apply this scheme as a building block in our scheme for the dirty-paper MAC with conferencing encoders in Section VI ahead.

For a fixed blocklength n, the scheme has the parameters

$$\tilde{N}, \tilde{P}, \tilde{Q}, \tilde{R}, \text{ and } \tilde{R}' \tag{22}$$

where $\tilde{Q}$ is nonnegative and $\tilde{N}$, $\tilde{P}$, $\tilde{R}$, and $\tilde{R}'$ are positive. Given these parameters, we define

$$\tilde{\alpha} \stackrel{\triangle}{=} \frac{\tilde{P}}{\tilde{P}+\tilde{N}}.$$

As we shall see, the rate of transmission of our scheme is $n^{-1}\log\lfloor 2^{n\tilde{R}}\rfloor$, which approaches $\tilde{R}$ as the blocklength n tends to infinity.

A. Codebook Generation

The codebook $\mathcal{C}$ consists of $\lfloor 2^{n\tilde{R}}\rfloor$ bins, each containing $\lfloor 2^{n\tilde{R}'}\rfloor$ codewords. The kth codeword in the mth bin is denoted $\mathbf{V}_{m,k}$. It is chosen at random independently of the other codewords according to the uniform distribution over the centered n-sphere of radius $\sqrt{n(\tilde{P}+\tilde{\alpha}^2\tilde{Q})}$. After the codebook is generated, it is revealed to the encoder and decoder.

B. Encoding

In order to convey Message $m \in \{1, \ldots, \lfloor 2^{n\tilde{R}}\rfloor\}$ when the interference is $\mathbf{s}$, the encoder chooses the codeword $\mathbf{v}_{m,k^*}$ in Bin m that has largest inner product with the interference $\mathbf{s}$, so

$$\langle \mathbf{v}_{m,k^*}, \mathbf{s}\rangle \geq \langle \mathbf{v}_{m,k}, \mathbf{s}\rangle, \qquad k \in \left\{1, \ldots, \lfloor 2^{n\tilde{R}'}\rfloor\right\}.$$

(Notice that $\mathbf{v}_{m,k^*}$ is uniquely defined with probability one.) The encoder then produces the channel inputs

$$\mathbf{x} = \mathbf{v}_{m,k^*} - \tilde{\alpha}\mathbf{s}$$

where $\mathbf{x} \stackrel{\triangle}{=} (x_1, \ldots, x_n)^{\mathsf{T}}$ denotes the n channel inputs stacked in an n-dimensional column-vector. We refer to $\mathbf{v}_{m,k^*}$ as the dirty-paper *codeword* and to $\mathbf{v}_{m,k^*} - \tilde{\alpha}\mathbf{s}$ as the dirty-paper *sequence*. Both depend on the message and the interference.

C. Decoding of [2]

The receiver stacks its observed channel outputs $y_1, \ldots, y_n$ into an n-dimensional column-vector $\mathbf{y} \stackrel{\triangle}{=} (y_1, \ldots, y_n)^{\mathsf{T}}$. It then applies nearest neighbor decoding, i.e., it looks for the codeword $\mathbf{v}_{\hat{m},\hat{k}}$ in the codebook $\mathcal{C}$ that is closest (in Euclidean distance) to $\mathbf{y}$. Since all the codewords in $\mathcal{C}$ are of equal Euclidean norm, this is equivalent to looking for the codeword $\mathbf{v}_{\hat{m},\hat{k}}$ that for all $m \in \{1, \ldots, \lfloor 2^{n\tilde{R}}\rfloor\}$ and all $k \in \{1, \ldots, \lfloor 2^{n\tilde{R}'}\rfloor\}$ satisfies

$$\left\langle \mathbf{v}_{\hat{m},\hat{k}}, \mathbf{y}\right\rangle \geq \langle \mathbf{v}_{m,k}, \mathbf{y}\rangle.$$

(The codeword $\mathbf{v}_{\hat{m},\hat{k}}$ is uniquely defined with probability one.) The receiver then produces the bin $\hat{m}$ of the closest codeword $\mathbf{v}_{\hat{m},\hat{k}}$ as its guess of the transmitted message.

D. Extended Decoding

We extend the decoding rule of [2] and require that in addition to $\hat{m}$, the decoder also produces $\hat{k}$.

E. Analysis

We next analyze the transmitted power and the probability of error of the extended decoding. We say that a decoding error occurred if $(M, K^*) \neq (\hat{M}, \hat{K})$. Thus, we say that an error has occurred not only when the wrong message is declared, but also when a wrong index $\hat{K}$ is declared.

Lemma V.1: Consider the performance of our scheme with parameters (22) over the generalized dirty-paper channel of parameters P and Q and with noise vector $\mathbf{Z}$. Assume that $\tilde{Q} = Q$, that $\tilde{P} \leq P$, that

$$\tilde{R}' > \frac{1}{2}\log\left(1 + \frac{\tilde{\alpha}^2 Q}{\tilde{P}}\right) \tag{23}$$

$$\tilde{R} + \tilde{R}' < \frac{1}{2}\log\left(1 + \frac{\tilde{P}}{\tilde{N}} + \frac{Q\tilde{P}}{\tilde{N}(\tilde{P}+\tilde{N})}\right) \tag{24}$$

and that for some ϵ^* in the open interval $(0, \tilde{N})$ the following three conditions are satisfied:

i) ϵ^* is sufficiently small so that[2]

$$\frac{\tilde{P}+\tilde{\alpha}Q}{\sqrt{\tilde{P}+\tilde{\alpha}^2 Q}\sqrt{\tilde{P}+\tilde{N}+Q+(3-2\tilde{\alpha})\epsilon^*}} > \sqrt{1-2^{-2(\tilde{R}+\tilde{R}')}} \tag{25}$$

ii) the empirical noise variance satisfies

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\|\mathbf{Z}\|^2 - \tilde{N}\right| \le \epsilon^*\right] = 1 \tag{26}$$

iii) and for every $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left[-\epsilon \le \frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \le \epsilon^*\right] = 1. \tag{27}$$

Then, the probability of error of the extended scheme satisfies

$$\lim_{n\to\infty} \Pr\left[(M, K^*) \ne (\hat{M}, \hat{K})\right] = 0 \tag{28}$$

and the probability that the average[3] block power exceeds the constraint P tends to 0 as n tends to infinity.

Proof: See Appendix C. ■

Corollary V.2: Consider a generalized dirty-paper channel with parameters P and Q, and where the noise $\{Z_t\}$ is independent of the interference $\{S_t\}$ and is a stationary and ergodic sequence of variance N. If in our scheme, we choose parameters $\tilde{N}, \tilde{P}, \tilde{Q}, \tilde{R}, \tilde{R}'$ satisfying Conditions (23) and (24) and such that $\tilde{N} = N$, $\tilde{Q} = Q$, and $\tilde{P} \le P$, then the probability that the extended decoder errs and the probability that the produced input sequence violates the average block power constraint P both tend to 0 as the blocklength n tends to infinity.

Proof: If $\{Z_t\}$ is an ergodic noise sequence of variance N independent of $\{S_t\}$, then Assumptions (26) and (27) are satisfied for all $\epsilon^* > 0$. Moreover, for every choice of $\tilde{P}, \tilde{R}, \tilde{R}' > 0$ satisfying (24) there is a choice of $\epsilon^* > 0$ satisfying (25) (see footnote 2). ■

Remark V.3: The extended decoder errs whenever the original decoder errs, i.e., whenever $M \ne \hat{M}$. Thus, if the assumptions in Lemma V.1 hold, then also the original scheme has probability of error tending to 0 as the blocklength n tends to infinity.

VI. Dirty-Paper MAC With Conferencing Encoders: Coding Scheme and Analysis

In this section, we present a coding scheme that achieves the region $\mathcal{R}_{\mathrm{Ach}}$ over the dirty-paper MAC with conferencing encoders and, by time-sharing, its convex hull.

[2]There always exists an $\epsilon^* > 0$ satisfying (25). This follows because of the continuity of the left-hand side and the right-hand side of (25); because the right-hand side of (25) is monotonically increasing in the sum $(\tilde{R}+\tilde{R}')$; and because for $(\tilde{R}+\tilde{R}') = \frac{1}{2}\log\left(1+\frac{\tilde{P}}{\tilde{N}}+\frac{Q\tilde{P}}{\tilde{N}(\tilde{P}+\tilde{N})}\right)$ and $\epsilon^* = 0$ the left-hand side of (25) equals its right-hand side.

[3]A scheme of asymptotically equal probability of error, but that satisfies the block-power constraint with probability one, is obtained if the transmitter sends the all-zero sequence whenever the produced dirty-paper sequence violates the power constraint.

Encoding takes place in two stages. In the first stage, as proposed by Willems [24], the transmitters use the conference to create a common message: Transmitter 1 splits Message M_1 into a private part $M_{1,p}$ of rate $R_{1,p}$ and a common part $M_{1,c}$ of rate $R_{1,c}$, where $R_{1,p}$ and $R_{1,c}$ sum to R_1 and likewise Transmitter 2. The rates of the common parts are chosen to satisfy

$$R_{1,c} \le C_{12} \tag{29}$$

$$R_{2,c} \le C_{21} \tag{30}$$

so the transmitters can use the conference to exchange the common parts of their messages. After the conference, each transmitter is cognizant of its private message and of the common-message pair $(M_{1,c}, M_{2,c})$. The transmitters do not use their knowledge of the interference sequence during the conference.

In the second stage, the transmitters communicate the private messages $M_{1,p}$ and $M_{2,p}$ and the common-message pair $(M_{1,c}, M_{2,c})$ over the MAC to the receiver. In this stage, we adopt an approach different from Willems's. In our approach the transmitters use time-sharing between two schemes. The first scheme achieves arbitrary small probability of error (for sufficiently large blocklengths) whenever the rate tuple $(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p})$ lies in the region

$$\begin{aligned}
&\mathcal{R}^{(4)}_{\mathrm{Ach},1}(P_1, P_2, N)\\
&\triangleq \bigcup_{0\le\beta_1,\beta_2\le 1} \Bigg\{(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p}):\\
&\qquad R_{1,p} \le \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{\beta_2 P_2 + N}\right)\\
&\qquad R_{2,p} \le \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{N}\right)\\
&\quad R_{1,c}+R_{2,c} \le \frac{1}{2}\log\left(1+\frac{\bar{\beta}_1 P_1 + \bar{\beta}_2 P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{\beta_1 P_1 + \beta_2 P_2 + N}\right)\Bigg\}
\end{aligned}$$

and the second scheme achieves arbitrary small probability of error whenever the rate tuple $(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p})$ lies in the region

$$\begin{aligned}
&\mathcal{R}^{(4)}_{\mathrm{Ach},2}(P_1, P_2, N)\\
&\triangleq \bigcup_{0\le\beta_1,\beta_2\le 1} \Bigg\{(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p}):\\
&\qquad R_{1,p} \le \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right)\\
&\qquad R_{2,p} \le \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{\beta_1 P_1 + N}\right)\\
&\quad R_{1,c}+R_{2,c} \le \frac{1}{2}\log\left(1+\frac{\bar{\beta}_1 P_1 + \bar{\beta}_2 P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{\beta_1 P_1 + \beta_2 P_2 + N}\right)\Bigg\}.
\end{aligned}$$

As we next show, by time-sharing between the two schemes, we can achieve arbitrary small probability of error in the second stage whenever the rate tuple $(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p})$ lies in

$$\begin{aligned}
&\mathcal{R}^{(4)}_{\mathrm{Ach}}(P_1, P_2, N)\\
&\triangleq \bigcup_{0\le\beta_1,\beta_2\le 1} \Bigg\{(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p}):
\end{aligned}$$

$$R_{1,p} \le \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right)$$
$$R_{2,p} \le \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{N}\right)$$
$$R_{1,p}+R_{2,p} \le \frac{1}{2}\log\left(1+\frac{\beta_1 P_1+\beta_2 P_2}{N}\right)$$
$$R_{1,c}+R_{2,c} \le \frac{1}{2}\log\left(1+\frac{\bar\beta_1 P_1+\bar\beta_2 P_2+2\sqrt{\bar\beta_1\bar\beta_2 P_1 P_2}}{\beta_1 P_1+\beta_2 P_2+N}\right)\Bigg\}.$$

To see this, we note that for every pair (β_1,β_2) the convex-hull of the union of the polytope in the definition of $\mathcal{R}^{(4)}_{\mathrm{Ach},1}(P_1,P_2,N)$ corresponding to this pair and the polytope corresponding to it in the definition of $\mathcal{R}^{(4)}_{\mathrm{Ach},2}(P_1,P_2,N)$ is the polytope corresponding to (β_1,β_2) in the definition of $\mathcal{R}^{(4)}_{\mathrm{Ach}}(P_1,P_2,N)$. (This is reminiscent of the fact that for fixed input distributions, the convex-hull of the union of the rectangle $\{(R_1,R_2) : R_1 \le I(X_1;Y)\ R_2 \le I(X_2;Y|X_1)\}$ corresponding to successive decoding for the classical MAC in one order and the rectangle $\{(R_1,R_2) : R_1 \le I(X_1;Y|X_2)\ R_2 \le I(X_2;Y)\}$ corresponding to successive decoding in the other is the pentagon $\{(R_1,R_2) : R_1 \le I(X_1;Y|X_2)\ R_2 \le I(X_2;Y|X_1),\ R_1+R_2 \le I(X_1,X_2;Y)\}$.)

We conclude that our technique can achieve arbitrary small probability of error over both stages, if the rate tuple $(R_{1,c},R_{2,c},R_{1,p},R_{2,p})$ simultaneously satisfies the rate constraints imposed in these two stages, i.e., if it lies in the region

$$\mathcal{R}^{(4)}_{\mathrm{Ach}} \cap \left\{(R_{1,c},R_{2,c},R_{1,p},R_{2,p}) : \begin{array}{rcl} R_{1,c} & \le & C_{12} \\ R_{2,c} & \le & C_{21} \end{array}\right\}. \tag{31}$$

Projecting the region in (31) onto the 2-D plane $(R_{1,c}+R_{1,p}, R_{2,c}+R_{2,p})$ (e.g., by means of the Fourier-Motzkin Elimination) results in $\mathcal{R}_{\mathrm{Ach}}$. That is, $\mathcal{R}_{\mathrm{Ach}}$ equals the image of the region in (31) under the mapping $(R_{1,c},R_{2,c},R_{1,p},R_{2,p}) \mapsto (R_{1,c}+R_{1,p}, R_{2,c}+R_{2,p})$. This establishes that the presented two-staged coding technique achieves all rate pairs (R_1,R_2) in $\mathcal{R}_{\mathrm{Ach}}$.

It remains to describe the two coding schemes that can be used in the second stage to achieve $\mathcal{R}^{(4)}_{\mathrm{Ach},1}$ and $\mathcal{R}^{(4)}_{\mathrm{Ach},2}$. We shall focus on the scheme that achieves $\mathcal{R}^{(4)}_{\mathrm{Ach},1}$. The other scheme is analogous but with reversed roles for the two transmitters and reversed decoding order of the private messages.

A. Scheme Achieving $\mathcal{R}^{(4)}_{\mathrm{Ach},1}$ in the Second Stage

Recall that after the conference, Transmitter 1 is cognizant of its private part $M_{1,p}$ and of the common parts $M_{1,c}$ and $M_{2,c}$, whereas Transmitter 2 is cognizant of its private part $M_{2,p}$ and of the common parts $M_{1,c}$ and $M_{2,c}$. To simplify notation, define the common message $M_0 \triangleq (M_{1,c},M_{2,c})$ of rate $R_0 = R_{1,c}+R_{2,c}$.

We begin with a sketch of the coding scheme. It is based on separately encoding messages M_0, $M_{1,p}$, and $M_{2,p}$ with a single-user dirty-paper code as described in Section V. Each transmitter then adds the dirty-paper sequence produced for its private message and a scaled version of the dirty-paper sequence produced for the common message and transmits the result. The encoders use the same dirty-paper code to encode the common message, so the channel adds these sequences coherently. The receiver applies successive decoding and stripping to decode the three messages $M_0, M_{1,p}, M_{2,p}$, where each decoding step is performed using nearest neighbor decoding. The outputs observed in the three decoding steps correspond to outputs of generalized single-user dirty-paper channels as considered in Section IV, but with different parameters.

We now describe the scheme in detail. We fix a blocklength n. Our scheme has parameters $\beta_1,\beta_2 \in [0,1]$, $\tilde P_1 \in (0,P_1)$, $\tilde P_2 \in (0,P_2)$, and $\delta_0,\delta_1,\delta_2 > 0$.

1) Preliminary Definitions: Define for each $\nu \in \{1,2\}$

$$P_{\nu,p} \triangleq \beta_\nu \tilde P_\nu \tag{32}$$

and

$$P_0 \triangleq \left(\sqrt{(1-\beta_1)\tilde P_1}+\sqrt{(1-\beta_2)\tilde P_2}\right)^2. \tag{33}$$

We interpret $P_{\nu,p}$ as the power that Transmitter ν allocates to its private message $M_{\nu,p}$, and P_0 as the received power dedicated to the common message M_0. Further, define

$$N_0 \triangleq P_{1,p}+P_{2,p}+N \tag{34}$$
$$N_1 \triangleq P_{2,p}+N \tag{35}$$
$$N_2 \triangleq N. \tag{36}$$

We will see that N_0 is the noise-variance that the receiver experiences when decoding M_0; N_1 is the noise-variance it experiences when decoding $M_{1,p}$ (if M_0 was decoded correctly); and N_2 is the noise-variance it experiences when decoding $M_{2,p}$ (if M_0 and $M_{1,p}$ were decoded correctly).

Define the n-dimensional column-vectors

$$\mathbf{S}_0 \triangleq \mathbf{S} \tag{37}$$
$$\mathbf{S}_1 \triangleq (1-\alpha_0)\mathbf{S} \tag{38}$$
$$\mathbf{S}_2 \triangleq (1-\alpha_1)(1-\alpha_0)\mathbf{S}. \tag{39}$$

We will see that these vectors correspond to the interference experienced in the different decoding steps. The vectors $\mathbf{S}_0, \mathbf{S}_1$, and $\mathbf{S}_2$ are of normalized powers $Q_0 \triangleq Q$, $Q_1 \triangleq (1-\alpha_0)^2 Q$, and $Q_2 \triangleq (1-\alpha_1)^2(1-\alpha_0)^2 Q$, where $\alpha_0 \triangleq \frac{P_0}{P_0+N_0}$ and, for $\nu \in \{1,2\}$, $\alpha_\nu \triangleq \frac{P_{\nu,p}}{P_{\nu,p}+N_\nu}$.

Finally, define the rates $R_0, R_{1,p}, R_{2,p}, R_0', R_1', R_2' > 0$ as

$$R_0' \triangleq \frac{1}{2}\log\left(1+\frac{\alpha_0^2 Q_0}{P_0}\right)+\delta_0 \tag{40}$$
$$R_1' \triangleq \frac{1}{2}\log\left(1+\frac{\alpha_1^2 Q_1}{P_{1,p}}\right)+\delta_1 \tag{41}$$
$$R_2' \triangleq \frac{1}{2}\log\left(1+\frac{\alpha_2^2 Q_2}{P_{2,p}}\right)+\delta_2 \tag{42}$$

and

$$R_0 \triangleq \frac{1}{2}\log\left(1+\frac{P_0}{N_0}\right)-2\delta_0 \tag{43}$$

$$R_{1,p} \stackrel{\triangle}{=} \frac{1}{2}\log\left(1+\frac{P_{1,p}}{N_1}\right) - 2\delta_1 \tag{44}$$

$$R_{2,p} \stackrel{\triangle}{=} \frac{1}{2}\log\left(1+\frac{P_{2,p}}{N_2}\right) - 2\delta_2. \tag{45}$$

Notice that by (40)–(45)

$$R_0' + R_0 = \frac{1}{2}\log\left(1+\frac{P_0}{N_0}+\frac{Q_0 P_0}{N_0(P_0+N_0)}\right) - \delta_0 \tag{46}$$

$$R_1' + R_{1,p} = \frac{1}{2}\log\left(1+\frac{P_{1,p}}{N_1}+\frac{Q_1 P_{1,p}}{N_1(P_{1,p}+N_1)}\right) - \delta_1 \tag{47}$$

$$R_2' + R_{2,p} = \frac{1}{2}\log\left(1+\frac{P_{2,p}}{N_2}+\frac{Q_2 P_{2,p}}{N_2(P_{2,p}+N_2)}\right) - \delta_2. \tag{48}$$

Remark VI.1: As we will see, our scheme is of rates $\frac{1}{2}\log\lfloor 2^{nR_0}\rfloor$, $\frac{1}{2}\log\lfloor 2^{nR_{1,p}}\rfloor$, and $\frac{1}{2}\log\lfloor 2^{nR_{2,p}}\rfloor$. Thus, when the blocklength n tends to infinity, the rates of our scheme tend to $(R_0, R_{1,p}, R_{2,p})$.

Remark VI.2: For every quadruple $(R_{1,c}, R_{2,c}, R_{1,p}, R_{2,p})$ in the interior of $\mathcal{R}^{(4)}_{\text{Ach},1}$, one can find $\beta_1, \beta_2 \in [0,1]$ and $0 < \tilde{P}_1 < P_1,\ 0 < \tilde{P}_2 < P_2$ so that the rate defined through the right-hand side of (43) and $\delta_0 = 0$ exceeds the sum of the common rates $R_{1,c} + R_{2,c}$ and so that the rates defined through the right-hand sides of (44) and (45) and $\delta_1 = \delta_2 = 0$ exceed the private rates $R_{1,p}$ and $R_{2,p}$. By continuity arguments, the same holds also for all sufficiently small $\delta_0, \delta_1, \delta_2 > 0$.

2) Codebook Generation: Independently construct codebooks $\mathcal{C}_0, \mathcal{C}_1$, and $\mathcal{C}_2$ as in Section V-A. Choose the parameters $\tilde{N}, \tilde{P}, \tilde{Q}, \tilde{R}, \tilde{R}'$ as the tuple N_0, P_0, Q_0, R_0, R_0' when constructing $\mathcal{C}_0$; choose them as $N_1, P_{1,p}, Q_1, R_{1,p}, R_1'$ when constructing $\mathcal{C}_1$; and choose them as $N_2, P_{2,p}, Q_2, R_{2,p}, R_2'$ when constructing $\mathcal{C}_2$.

Denote the codewords in Bin $m_0 \in \{1, \ldots, \lfloor 2^{nR_0}\rfloor\}$ of codebook $\mathcal{C}_0$ by $\{\mathbf{V}_{0,m_0,k_0}\}_{k_0=1}^{\lfloor 2^{nR_0'}\rfloor}$; the codewords in Bin $m_1 \in \{1, \ldots, \lfloor 2^{nR_{1,p}}\rfloor\}$ of codebook $\mathcal{C}_1$ by $\{\mathbf{V}_{1,m_1,k_1}\}_{k_1=1}^{\lfloor 2^{nR_1'}\rfloor}$; and the codewords in Bin $m_2 \in \{1, \ldots, \lfloor 2^{nR_{2,p}}\rfloor\}$ of codebook $\mathcal{C}_2$ by $\{\mathbf{V}_{2,m_2,k_2}\}_{k_2=1}^{\lfloor 2^{nR_2'}\rfloor}$.

3) Encoding: The transmitters encode the common message M_0 with the single-user dirty-paper encoding of Section V-B using the codebook $\mathcal{C}_0$, the parameter α_0, and assuming that the interference is $\mathbf{S}_0$. This produces the dirty-paper sequence

$$\mathbf{X}_0' = \mathbf{V}_{0,M_0,K_0^*} - \alpha_0 \mathbf{S}_0.$$

Transmitter ν also encodes its private message $M_{\nu,p}$ with the dirty-paper encoding of Section V-B. It now uses the codebook $\mathcal{C}_\nu$, the parameter α_ν, and assumes that the interference is $\mathbf{S}_\nu$. This produces the dirty-paper sequence

$$\mathbf{X}_\nu' = \mathbf{V}_{\nu,M_{\nu,p},K_\nu^*} - \alpha_\nu \mathbf{S}_\nu, \quad \nu \in \{1,2\}.$$

Transmitter ν then transmits a linear combination of $\mathbf{X}_0'$ and $\mathbf{X}_\nu'$. To describe Transmitter ν's signal, we stack its n input symbols into an n-dimensional column-vector $\mathbf{X}_\nu \stackrel{\triangle}{=} (X_{\nu,1}, \ldots, X_{\nu,n})^{\mathsf{T}}$. The inputs are then described by

$$\mathbf{X}_\nu = \lambda_\nu \mathbf{X}_0' + \mathbf{X}_\nu', \quad \nu \in \{1,2\} \tag{49}$$

where

$$\lambda_\nu \stackrel{\triangle}{=} \sqrt{\frac{(1-\beta_\nu)\tilde{P}_\nu}{P_0}}, \quad \nu \in \{1,2\}.$$

Notice that $\lambda_1 + \lambda_2 = 1$, and the channel coherently combines the transmissions of the common message M_0, so the n-dimensional column-vector of output symbols $\mathbf{Y} \stackrel{\triangle}{=} (Y_1, \ldots, Y_n)^{\mathsf{T}}$ is

$$\mathbf{Y} \stackrel{\triangle}{=} \mathbf{X}_1 + \mathbf{X}_2 + \mathbf{S} + \mathbf{Z} = \mathbf{X}_0' + \mathbf{X}_1' + \mathbf{X}_2' + \mathbf{S} + \mathbf{Z},$$

where $\mathbf{Z}$ denotes the n-dimensional column vector of noise symbols $\mathbf{Z} = (Z_1, \ldots, Z_n)^{\mathsf{T}}$.

4) Decoding: The receiver stacks its observed outputs symbols into an n-dimensional column-vector $\mathbf{Y} \stackrel{\triangle}{=} (Y_1, \ldots, Y_n)^{\mathsf{T}}$. It first decodes message M_0 based on $\mathbf{Y}$ using the extended decoding in Section V-D with codebook $\mathcal{C}_0$. Thus, in this first decoding step, the receiver treats the dirty-paper sequences $\mathbf{X}_1'$ and $\mathbf{X}_2'$ produced to encode the private messages as additional noise. The extended decoding produces the pair $(\hat{M}_0, \hat{K}_0)$ and the receiver declares the common message $\hat{M}_0$. The receiver then subtracts the decoded dirty-paper codeword $\mathbf{V}_{0,\hat{M}_0,\hat{K}_0}$ in $\mathcal{C}_0$ from $\mathbf{Y}$ to obtain

$$\hat{\mathbf{Y}}_1 \stackrel{\triangle}{=} \mathbf{Y} - \mathbf{V}_{0,\hat{M}_0,\hat{K}_0}.$$

If the first decoding stage was successful, then $\hat{\mathbf{Y}}_1 = \mathbf{X}_1' + \mathbf{X}_2' + (1-\alpha_0)\mathbf{S} + \mathbf{Z}$. And since $\mathbf{X}_0 = (\mathbf{V}_{0,\hat{M}_0,\hat{K}_0} - \alpha_0 \mathbf{S})$, by subtracting $\mathbf{V}_{0,\hat{M}_0,\hat{K}_0}$ from $\mathbf{Y}$ the interference is reduced by $\alpha_0 \mathbf{S}$. Notice that the receiver cannot subtract the dirty-paper sequence $\mathbf{X}_0'$ that was used to encode the common message M_0, because it is not cognizant of the interference $\mathbf{S}$.

Based on $\hat{\mathbf{Y}}_1$ the receiver decodes $M_{1,p}$ again by means of the extended decoding in Section V-D but now using the codebook $\mathcal{C}_1$. Thus, in this second decoding step, the receiver treats the dirty-paper sequence $\mathbf{X}_2'$ for the private message $M_{2,p}$ as additional noise. The extended decoding produces the pair $(\hat{M}_{1,p}, \hat{K}_1)$, and the receiver declares that Transmitter 1 sent the private message $\hat{M}_{1,p}$. It then subtracts the decoded codeword $\mathbf{V}_{1,\hat{M}_{1,p},\hat{K}_1}$ in $\mathcal{C}_1$ from $\hat{\mathbf{Y}}_1$ to form

$$\hat{\mathbf{Y}}_2 \stackrel{\triangle}{=} \hat{\mathbf{Y}}_1 - \mathbf{V}_{1,\hat{M}_{1,p},\hat{K}_1}.$$

If the first two decoding stages were successful, then $\hat{\mathbf{Y}}_2 = \mathbf{X}_2' + (1-\alpha_0)(1-\alpha_1)\mathbf{S} + \mathbf{Z}$.

Based on $\hat{\mathbf{Y}}_2$ the receiver finally decodes $M_{2,p}$ by means of the extended decoding now using the codebook $\mathcal{C}_2$. This decoding produces the pair $(\hat{M}_{2,p}, \hat{K}_2)$, and the receiver declares that Transmitter 2 sent the private message $\hat{M}_{2,p}$. It discards $\hat{K}_2$.

5) Error Analysis: The decoder errs whenever

$$(M_0, M_{1,p}, M_{2,p}) \neq (\hat{M}_0, \hat{M}_{1,p}, \hat{M}_{2,p}).$$

We analyze the average error probability of the scheme (averaged over all possible codebooks, messages, interferences, and noise sequences). To this end, we use a genie-aided argument as

in [17] and [25]. As there, we introduce a genie-aided decoder, but our approach differs in that we define additional error events for the genie-aided decoder.

Genie-Aided Decoder: The genie-aided decoder consists of three independent decoders. The first decoder is fed the sequence of channel outputs $\mathbf{Y}$

$$\mathbf{Y} = \mathbf{X}_0' + \mathbf{S}_0 + \mathbf{Z}_0 \tag{50}$$

where $\mathbf{S}_0$ is defined in (37), and $\mathbf{Z}_0 \triangleq \mathbf{X}_1' + \mathbf{X}_2' + \mathbf{Z}$. It applies the extended decoding in Section V-D to $\mathbf{Y}$ using the codebook $\mathcal{C}_0$, thus producing the pair $(\hat{M}_0^{\mathrm{G}}, \hat{K}_0^{\mathrm{G}})$.

The second decoder is fed by a genie the sequence $\mathbf{Y}_1 \triangleq \mathbf{Y} - \mathbf{V}_{0,M_0,K_0^*}$ (but not the original output sequence $\mathbf{Y}$), where M_0 is the correct common message and K_0^* is the index that was produced by the encoder when it encoded it. Thus

$$\mathbf{Y}_1 = \mathbf{X}_1' + \mathbf{S}_1 + \mathbf{Z}_1 \tag{51}$$

where $\mathbf{S}_1$ is defined in (38), and $\mathbf{Z}_1 \triangleq \mathbf{X}_2' + \mathbf{Z}$. It applies the extended decoding to $\mathbf{Y}_1$ with the codebook $\mathcal{C}_1$ thus producing the pair $(\hat{M}_{1,p}^{\mathrm{G}}, \hat{K}_1^{\mathrm{G}})$.

The third decoder is fed by a genie the sequence $\mathbf{Y}_2 \triangleq \mathbf{Y} - \mathbf{V}_{0,M_0,K_0^*} - \mathbf{V}_{1,M_{1,p},K_1^*}$ (but not the original output sequence $\mathbf{Y}$ or the sequence $\mathbf{Y}_1$), where $M_{1,p}$ is the correct private message of Transmitter 1, and K_1^* is the index that was produced when it was encoded. Thus

$$\mathbf{Y}_2 = \mathbf{X}_2' + \mathbf{S}_2 + \mathbf{Z}_2 \tag{52}$$

where $\mathbf{S}_2$ is defined in (39), and $\mathbf{Z}_2 \triangleq \mathbf{Z}$. Based on $\mathbf{Y}_2$ it applies the extended decoding using the codebook $\mathcal{C}_2$, thus producing the pair $(\hat{M}_{2,p}^{\mathrm{G}}, \hat{K}_2^{\mathrm{G}})$.

The error event for the genie-aided decoder structure is the event that at least one of the three single decoders errs, i.e., that

$$\begin{aligned}&(M_0, K_0^*, M_{1,p}, K_1^*, M_{2,p}, K_2^*)\\ &\quad \neq (\hat{M}_0^{\mathrm{G}}, \hat{K}_0^{\mathrm{G}}, \hat{M}_{1,p}^{\mathrm{G}}, \hat{K}_1^{\mathrm{G}}, \hat{M}_{2,p}^{\mathrm{G}}, \hat{K}_2^{\mathrm{G}}).\end{aligned}$$

Thus, the genie-aided decoder errs not only when it produces wrong messages, but also wrong indices.

The original decoder guesses the messages $(M_0, M_{1,p}, M_{2,p})$ correctly whenever the genie-aided decoder guesses the sextuple $(M_0, K_0^*, M_{1,p}, K_1^*, M_{2,p}, K_2^*)$ correctly. Thus, the probability of error of the original decoder cannot be larger than the probability of error of the genie-aided decoder.

We show that for every choice of $\delta_1 > 0$ and $\delta_2 > 0$ sufficiently small (depending on the other parameters), the probability of error of the genie-aided decoder tends to 0 as the blocklength n tends to infinity. Since the genie-aided decoder structure errs whenever one of the three single decoders errs, it suffices to show that under these conditions, the probability of error for each of the three single decoders tends to 0.

We first analyze the probability of error of the third genie-aided decoder which guesses the pair $(M_{2,p}, K_2^*)$. To this end, we notice that the outputs in (52) are the result of applying the dirty-paper coding of Section V with parameters $N_2, P_{2,p}, Q_2, R_{2,p}, R_2'$ over a dirty-paper channel with interference $\mathbf{S}_2 = (1-\alpha_0)(1-\alpha_1)\mathbf{S}$, which is uniformly distributed over an n-sphere of radius $\sqrt{nQ_2}$, and with *independent* memoryless Gaussian noise of variance N_2. Since the decoder uses the extended decoding of Section V, it follows from (42) and (48) and Corollary V.2 that

$$\lim_{n\to\infty} \Pr\left[(M_{2,p}, K_2^*) \neq (\hat{M}_{2,p}^{\mathrm{G}}, \hat{K}_2^{\mathrm{G}})\right] = 0. \tag{53}$$

We next analyze the probability of error of the first genie-aided decoder which tries to guess (M_0, K_0^*). We again notice that the outputs (50) are the result of applying the dirty-paper coding in Section V with parameters N_0, P_0, Q_0, R_0, R_0' over a generalized dirty-paper channel with interference $\mathbf{S}_0 = \mathbf{S}$, which is uniformly distributed over an n-sphere of radius $\sqrt{nQ_0}$, and with noise $\mathbf{Z}_0 = (\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1) + (\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2) + \mathbf{Z}$. Here, the noise $\mathbf{Z}_0$ depends on the interference $\mathbf{S}_0$ (because the dirty-paper sequences $(\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1)$ and $(\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2)$ depend on $\mathbf{S}$). Thus, even though the receiver applies the extended dirty-paper coding of Section V, Corollary V.2 is not sufficient to prove that the probability of error tends to 0. Instead, we need the more general Lemma V.1 combined with the following Lemma VI.3.

Lemma VI.3: For a fixed choice of $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, and $\delta_0 > 0$, if the parameters δ_1 and δ_2 are chosen to satisfy

$$0 < \delta_1 < \delta_{1,0}^*(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) \tag{54}$$

$$0 < \delta_2 < \delta_{2,0}^*(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) \tag{55}$$

for some specific positive $\delta_{1,0}^*(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$ and $\delta_{2,0}^*(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$, then there exists an ϵ_0^* in the open interval $(0, N_0)$ so that the following three conditions are satisfied:

1) ϵ_0^* is sufficiently small so that

$$\begin{aligned}&\frac{P_0 + \alpha_0 Q_0}{\sqrt{P_0 + \alpha_0^2 Q_0}\sqrt{P_0 + N_0 + Q_0 + (3 - 2\alpha_0)\epsilon_0^*}}\\ &\quad > \sqrt{1 - 2^{-2(R_0 + R_0')}}\end{aligned} \tag{56}$$

2) the empirical noise variance satisfies

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\|\mathbf{Z}_0\|^2 - N_0\right| \leq \epsilon_0^*\right] = 1 \tag{57}$$

3) and for every $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left[-\epsilon \leq \frac{1}{n}\langle \mathbf{Z}_0, \mathbf{S}_0\rangle \leq \epsilon_0^*\right] = 1. \tag{58}$$

Proof: See Appendix E-B. ■

Combining this lemma with Lemma V.1, equalities (40) and (46), and Assumption $\delta_0 > 0$, we conclude that if δ_1 and δ_2 are sufficiently small to satisfy (54) and (55), then

$$\lim_{n\to\infty} \Pr\left[(M_0, K_0^*) \neq (\hat{M}_0^{\mathrm{G}}, \hat{K}_0^{\mathrm{G}})\right] = 0. \tag{59}$$

We finally analyze the probability of error of the second genie-aided decoder which guesses the pair $(M_{1,p}, K^*_{1,p})$. We notice that the outputs (51) are the result of applying the dirty-paper coding in Section V with parameters N_1, $P_{1,p}$, Q_1, $R_{1,p}$, R'_1 over a generalized dirty-paper channel with interference $\mathbf{S}_1 = (1-\alpha_0)\mathbf{S}$, which is uniformly distributed over an n-sphere of radius $\sqrt{nQ_1}$, and with noise $\mathbf{Z}_1 = (\mathbf{V}_{2,M_{2,p},K^*_2} - \alpha_2\mathbf{S}_2) + \mathbf{Z}$. Since the interference $\mathbf{S}_1$ and the noise $\mathbf{Z}_1$ are dependent, we need to apply Lemma V.1, now combined with the following Lemma VI.4.

Lemma VI.4: For a fixed choice of $\beta_1, \beta_2 > 0$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, and $\delta_0, \delta_1 > 0$, if the parameter δ_2 is chosen to satisfy

$$0 < \delta_2 < \delta^*_{2,1}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1) \tag{60}$$

for some specific positive $\delta^*_{2,1}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)$, then there exists some ϵ^*_1 in the open interval $(0, N_1)$ so that the following three conditions are satisfied:

1) ϵ^*_1 is sufficiently small so that

$$\frac{P_{1,p} + \alpha_1 Q_1}{\sqrt{P_{1,p} + \alpha_1^2 Q_1}\sqrt{P_{1,p} + N_1 + Q_1 + (3-2\alpha_1)\epsilon^*_1}} > \sqrt{1 - 2^{-2(R_1 + R'_1)}} \tag{61}$$

2) the empirical noise variance satisfies

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\|\mathbf{Z}_1\|^2 - N_1\right| \le \epsilon^*_1\right] = 1 \tag{62}$$

3) and for every $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left[-\epsilon \le \frac{1}{n}\langle \mathbf{Z}_1, \mathbf{S}_1\rangle \le \epsilon^*_1\right] = 1. \tag{63}$$

Proof: Similar to the proof of Lemma VI.3, and therefore omitted. ■

Combining this lemma with Lemma V.1, equalities (41) and (47), and Assumption $\delta_1 > 0$ establishes that when δ_2 is sufficiently small so as to satisfy (60), then

$$\lim_{n\to\infty} \Pr\left[(M_{1,p}, K^*_1) \ne (\hat{M}^{\mathrm{G}}_{1,p}, \hat{K}^{\mathrm{G}}_1)\right] = 0. \tag{64}$$

Combining (53), (59), and (64), we conclude the following.

Conclusion VI.5: For given parameters $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, and $\delta_0 > 0$, if δ_1 and δ_2 are chosen so as to satisfy

$$0 < \delta_1 < \delta^*_{1,0}(\delta_0, \beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{65}$$

$$0 < \delta_2 < \min\{\delta^*_{2,0}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0), \delta^*_{2,1}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)\} \tag{66}$$

for some specific positive $\delta^*_{1,0}(\delta_0, \beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$, $\delta^*_{2,0}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$, and $\delta^*_{2,1}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)$, then the probability of error of the genie-aided decoder tends to 0 as the blocklength n tends to infinity. This also implies that under these conditions the probability of error of our original decoder tends to 0 as n tends to infinity.

Power Constraints: By Lemma VI.6 ahead, the probability that our scheme satisfies the average block-power constraints P_1 and P_2 tends to 1 as the blocklength tends to infinity, if $\delta_0, \delta_1, \delta_2 > 0$ are chosen sufficiently small.[4]

Lemma VI.6: For given parameters $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$ and $\tilde{P}_2 \in (0, P_2)$, if our choice of $\delta_0, \delta_1, \delta_2$ satisfies

$$0 < \delta_0 < \delta^*_{0,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{67}$$

$$0 < \delta_1 < \delta^*_{1,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{68}$$

$$0 < \delta_2 < \delta^*_{2,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{69}$$

for some specific positive $\delta^*_{0,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$, $\delta^*_{1,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$, $\delta^*_{2,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$, then

$$\lim_{n\to\infty} \Pr\left[\frac{1}{n}\|\mathbf{X}_\nu\|^2 \le P_\nu\right] = 1, \qquad \nu \in \{1,2\}.$$

Proof: See Appendix E-C. ■

The achievability of $\mathcal{R}^{(4)}_{\mathrm{Ach},1}$ during the second stage follows now immediately by combining Remarks VI.1 and VI.2 with Conclusion VI.5 and Lemma VI.6.

Appendix A
Proof of Lemma III.10

By Theorem III.2, the capacity region C_{Conf} equals $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}$. Therefore, to prove the lemma, we can equivalently show

$$\mathrm{conv}\,(\mathcal{R}_{\mathrm{Ach}}) = \mathcal{R}_{\mathrm{Conf},\mathcal{G}}$$

i.e., the two inclusions

$$\mathrm{conv}\,(\mathcal{R}_{\mathrm{Ach}}) \subseteq \mathcal{R}_{\mathrm{Conf},\mathcal{G}} \tag{70}$$

and

$$\mathrm{conv}\,(\mathcal{R}_{\mathrm{Ach}}) \supseteq \mathcal{R}_{\mathrm{Conf},\mathcal{G}}. \tag{71}$$

Before proving the inclusions, we define two sets of regions and restate the regions $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}$ and $\mathcal{R}_{\mathrm{Ach}}$ in terms of these sets. For every pair $(\beta_1, \beta_2) \in [0,1] \times [0,1]$, define the region

$$\mathcal{R}_{\mathrm{Conf}}(\beta_1, \beta_2) \triangleq \Bigg\{(R_1, R_2):$$

$$R_1 \le \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1}{N}\right) + C_{12} \tag{72a}$$

$$R_2 \le \frac{1}{2}\log\left(1 + \frac{\beta_2 P_2}{N}\right) + C_{21} \tag{72b}$$

$$R_1 + R_2 \le \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1 + \beta_2 P_2}{N}\right) + C_{12} + C_{21} \tag{72c}$$

[4]The previously described scheme is easily changed to a scheme that satisfies the power constraints with probability 1 for every blocklength n, if the transmitters simply send the all-zero sequence whenever the inputs computed in (49) violate the power constraints.

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right)\Bigg\} \tag{72d}$$

and the region $\mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2)$ as shown in (73a)–(73f) at the bottom of the page. The regions $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}$ and $\mathcal{R}_{\mathrm{Ach}}$ can be expressed as

$$\mathcal{R}_{\mathrm{Conf},\mathcal{G}} = \bigcup_{0\leq\beta_1,\beta_2\leq 1} \mathcal{R}_{\mathrm{Conf},\mathcal{G}}(\beta_1, \beta_2)$$

and

$$\mathcal{R}_{\mathrm{Ach}} = \bigcup_{0\leq\beta_1,\beta_2\leq 1} \mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2).$$

We first prove Inclusion (70). Since the expression for the region $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(\beta_1, \beta_2)$ in (72) differs from the expression for $\mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2)$ in (73) only in that the latter has the additional two constraints (73b) and (73d), for every pair (β_1, β_2) the region $\mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2)$ is included in the region $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(\beta_1, \beta_2)$. By the convexity of $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}$, this concludes the proof of (70).

We next prove (71). To this end, we show that for every pair $(\beta_1, \beta_2) \in [0,1]\times[0,1]$ the region $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(\beta_1, \beta_2)$ is included in the region $\mathcal{R}_{\mathrm{Ach}}$. The proof is trivial for all pairs $(\beta_1, \beta_2) \in [0,1]\times[0,1]$ that satisfy

$$\frac{1}{2}\log\left(1 + \frac{\left(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1 + \beta_2 P_2 + N}\right) \geq \max\{C_{12}, C_{21}\} \tag{74}$$

because in this case, the rate constraints (73b) and (73d) do not actively constrain the region $\mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2)$, and thus the regions $\mathcal{R}_{\mathrm{Conf}}(\beta_1, \beta_2)$ and $\mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2)$ coincide.

The proof is more tedious for pairs $(\beta_1, \beta_2) \in [0,1]\times[0,1]$ that violate (74). We fix such a pair (β_1, β_2). By the convex-hull operator on the right-hand side of (71), it suffices to show that both *dominant* corner points of the polytope $\mathcal{R}_{\mathrm{Conf},\mathcal{G}}(\beta_1, \beta_2)$ are included in the region $\mathcal{R}_{\mathrm{Ach}}$. A corner point of a region is called *dominant* if it is of maximum sum-rate in this region [17].

In the following, we present for each dominant corner point of $\mathcal{R}_{\mathrm{Conf}}(\beta_1, \beta_2)$ a pair $(\beta_1', \beta_2') \in [0,1]\times[0,1]$ such that the chosen dominant corner point lies in $\mathcal{R}_{\mathrm{Ach}}(\beta_1', \beta_2')$.

Before presenting these choices, we characterize the two dominant corner points. To this end, we notice that since our chosen pair (β_1, β_2) violates (74) and since the sum of two nonnegative numbers cannot be smaller than the maximum of these numbers:

$$\frac{1}{2}\log\left(1 + \frac{\left(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1 + \beta_2 P_2 + N}\right) \leq C_{12} + C_{21}. \tag{75}$$

This implies that in the region $\mathcal{R}_{\mathrm{Conf}}(\beta_1, \beta_2)$, the sum-rate $R_1 + R_2$ is bounded by Constraint (72d) rather than Constraint (72c). Moreover, since for arbitrary $P_1, P_2, N > 0$ and $\beta_1, \beta_2 \in [0,1]$ the sum of the right-hand sides of (72a) and (72b) is larger than the right-hand side of (72c) both dominant corner points in $\mathcal{R}_{\mathrm{Conf}}(\beta_1, \beta_2)$ are of sum-rate

$$\frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1, P_2}}{N}\right).$$

Thus, we conclude that the two dominant corner points $\left(R_1^{(1)}, R_2^{(1)}\right)$ and $\left(R_1^{(2)}, R_2^{(2)}\right)$ in $\mathcal{R}_{\mathrm{Conf}}(\beta_1, \beta_2)$ are given by (76) and (77) at the bottom of the next page.

We only prove that the first dominant corner point $\left(R_1^{(1)}, R_2^{(1)}\right)$ lies in $\mathcal{R}_{\mathrm{Ach}}(\beta_1', \beta_2')$ for some $(\beta_1', \beta_2') \in [0,1]\times[0,1]$. By symmetry the same conclusion also holds for the second dominant corner point.

Our choice of the pair (β_1', β_2') depends on whether

$$\frac{1}{2}\log\left(1 + \frac{\beta_1 P_1}{N}\right) + C_{12} \leq \frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \tag{78}$$

or

$$\mathcal{R}_{\mathrm{Ach}}(\beta_1, \beta_2) \triangleq \Bigg\{(R_1, R_2) : R_1 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1}{N}\right) + C_{12} \tag{73a}$$

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1}{N}\right) + \frac{1}{2}\log\left(1 + \frac{(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2})^2}{\beta_1 P_1 + \beta_2 P_2 + N}\right) \tag{73b}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{\beta_2 P_2}{N}\right) + C_{21} \tag{73c}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{\beta_2 P_2}{N}\right) + \frac{1}{2}\log\left(1 + \frac{(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2})^2}{\beta_1 P_1 + \beta_2 P_2 + N}\right) \tag{73d}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1 + \beta_2 P_2}{N}\right) + C_{12} + C_{21} \tag{73e}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right)\Bigg\} \tag{73f}$$

$$\frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right)+C_{12} > \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \quad (79)$$

and on whether

$$C_{12} \le \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1+\beta_2 P_2+N}\right) \quad (80)$$

or

$$C_{12} > \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1+\beta_2 P_2+N}\right). \quad (81)$$

If (79) holds, the first dominant corner point is given by

$$R_1^{(1)} = \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \quad (82)$$
$$R_2^{(1)} = 0$$

and we choose $\beta_1' = \beta_1$ and $\beta_2' = 0$. The corner point $\left(R_1^{(1)}, R_2^{(1)}\right)$ lies in $\mathcal{R}_{\mathrm{Ach}}(\beta_1', \beta_2')$ because:

1) The rate $R_1^{(1)}$ satisfies Constraints (73a) and (73b) when in these constraints the pair (β_1, β_2) is replaced by the pair (β_1', β_2'). In fact, (73b) follows by

$$\begin{aligned} R_1^{(1)} &= \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \\ &\le \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1 P_1 P_2}}{N}\right) \\ &= \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1' P_1 P_2}}{N}\right) \\ &= \frac{1}{2}\log\left(1+\frac{\beta_1' P_1}{N}\right) \\ &\quad + \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1' P_1}+\sqrt{\bar{\beta}_2' P_2}\right)^2}{\beta_1' P_1+\beta_2' P_2+N}\right). \end{aligned} \quad (83)$$

   Constraint (73a) follows by (82) and (79) and by the choice $\beta_1' = \beta_1$.
2) Since the rate $R_2^{(1)} = 0$, it trivially satisfies Constraints (73c) and (73d) when in these constraints (β_1, β_2) is replaced by (β_1', β_2').
3) The sum-rate $R_1^{(1)} + R_2^{(1)}$ satisfies both (73e) and (73f) when β_1, β_2 are replaced by β_1', β_2'. This holds because the single-rate $R_1^{(1)}$ satisfies the more stringent constraints (73a) and (73b), and because $R_1^{(1)} + R_2^{(1)} = R_1^{(1)}$.

We now consider the case where (78) holds. In this case, the first dominant corner point is given by

$$R_1^{(1)} = \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right)+C_{12} \quad (84a)$$
$$R_2^{(1)} = \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right)-C_{12}. \quad (84b)$$

If both (78) and (80) hold, then we choose $\beta_1' = \beta_1$ and $\beta_2' = \beta_2$. The corner point $\left(R_1^{(1)}, R_2^{(1)}\right)$ as defined in (84) lies in $\mathcal{R}_{\mathrm{Ach}}(\beta_1', \beta_2')$ because:

1) The rate $R_1^{(1)}$ satisfies (73a) and (73b) when in these constraints (β_1, β_2) is replaced by (β_1', β_2'): Constraint (73a) holds by (84a) and the choice $\beta_1' = \beta_1$; and Constraint (73b) by (73a), by assumption (80), and by $\beta_1' = \beta_1$ and $\beta_2' = \beta_2$.
2) The sum-rate $R_1^{(1)} + R_2^{(1)}$ satisfies (73f) and (73e) when in these constraints (β_1, β_2) is replaced by (β_1', β_2'): Con-

$$R_1^{(1)} = \min\left\{\frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right)+C_{12}, \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right)\right\} \quad (76a)$$

$$R_2^{(1)} = \max\left\{0, \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) - C_{12}\right\} \quad (76b)$$

$$R_1^{(2)} = \max\left\{0, \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) - \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{N}\right) - C_{21}\right\} \quad (77a)$$

$$R_2^{(2)} = \min\left\{\frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{N}\right)+C_{21}, \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right)\right\} \quad (77b)$$

$$\frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1+\beta_2 P_2+N}\right) < C_{12} \le \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) \quad (85)$$

straint (73f) follows immediately by (84) and our choice $\beta_1' = \beta_1$ and $\beta_2' = \beta_2$; and Constraint (73e) follows by (73f), by (75), and because $\beta_1' = \beta_1$ and $\beta_2' = \beta_2$.

3) The rate $R_2^{(1)}$ satisfies (73c) and (73d) when (β_1, β_2) are replaced by (β_1', β_2') because

$$\begin{aligned} R_2^{(1)} &= \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \\ &\quad - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) - C_{12} \\ &= \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1+\beta_2 P_2+N}\right) \\ &\quad + \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{\beta_1 P_1+N}\right) - C_{12} \\ &= \frac{1}{2}\log\left(1+\frac{\beta_2' P_2}{\beta_1' P_1+N}\right) - C_{12} \\ &\quad + \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1' P_1}+\sqrt{\bar{\beta}_2' P_2}\right)^2}{\beta_1' P_1+\beta_2' P_2+N}\right) \\ &\le \frac{1}{2}\log\left(1+\frac{\beta_2' P_2}{N}\right) \\ &\quad + \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1' P_1}+\sqrt{\bar{\beta}_2' P_2}\right)^2}{\beta_1' P_1+\beta_2' P_2+N}\right) \end{aligned}$$

and because (75) implies

$$\begin{aligned} R_2^{(1)} &= \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \\ &\quad - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) - C_{12} \\ &\le \frac{1}{2}\log\left(1+\frac{\beta_1 P_1+\beta_2 P_2}{N}\right) \\ &\quad - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) + C_{21} \\ &= \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{N+\beta_1 P_1}\right) + C_{21} \\ &\le \frac{1}{2}\log\left(1+\frac{\beta_2 P_2}{N}\right) + C_{21} \\ &= \frac{1}{2}\log\left(1+\frac{\beta_2' P_2}{N}\right) + C_{21}. \end{aligned}$$

Before presenting our choice of (β_1', β_2') when (78) and (81) hold, we recall that in this case inequalities (85) at the bottom of the previous page hold. Moreover, since

$$\begin{aligned} &\frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{P_2}\right)^2}{\beta_1 P_1+N}\right) \\ &\quad = \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1 P_1 P_2}}{N}\right) \\ &\qquad - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) \end{aligned}$$

and since the right-hand side of (85) is strictly decreasing in $\beta_2 \in [0,1]$

$$\begin{aligned} &\frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2 P_2}\right)^2}{\beta_1 P_1+\beta_2 P_2+N}\right) \\ &\quad < C_{12} \le \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{P_2}\right)^2}{\beta_1 P_1+N}\right). \end{aligned} \tag{86}$$

By continuity of the expressions in (86), there exists a $0 \le \beta_2^* < \beta_2$ such that

$$C_{12} = \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2^* P_2}\right)^2}{\beta_1 P_1+\beta_2^* P_2+N}\right). \tag{87}$$

We can now present our choice of (β_1', β_2') when (78) and (81) hold. In this case we choose $\beta_1' = \beta_1$ and $\beta_2' = \beta_2^*$. The rate pair $\left(R_1^{(1)}, R_2^{(1)}\right)$ as defined in (84) lies in $\mathcal{R}_{\mathrm{Ach}}(\beta_1', \beta_2')$ because:

1) The rate $R_1^{(1)}$ satisfies (73a) and (73b) if (β_1, β_2) is replaced by (β_1', β_2'): Constraint (73a) holds trivially by (84a) and by $\beta_1' = \beta_1$; and Constraint (73b) holds by (73a), by (87), and because $\beta_1' = \beta_1$ and $\beta_2' = \beta_2^*$.
2) The sum-rate $R_1^{(1)} + R_2^{(1)}$ satisfies (73e) and (73f) if (β_1, β_2) is replaced by (β_1', β_2'). In fact, Constraint (73f) holds because

$$\begin{aligned} R_1^{(1)} + R_2^{(1)} &= \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right) \\ &< \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1'\bar{\beta}_2' P_1 P_2}}{N}\right) \end{aligned} \tag{88}$$

where the inequality follows because $\beta_1' = \beta_1$ and $\beta_2' < \beta_2$; and Constraint (73e) follows by (88), by (87), by $\beta_1' = \beta_1$ and $\beta_2' = \beta_2^*$, and because $C_{21} \ge 0$.

3) The rate $R_2^{(1)}$ satisfies both (73c) and (73d) if (β_1, β_2) is replaced by (β_1', β_2') because (87) implies

$$R_2^{(1)} = \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2 P_1 P_2}}{N}\right)$$

$$
\begin{aligned}
&- \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) - C_{12} \\
&\le \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2^* P_1 P_2}}{N}\right) \\
&\quad - \frac{1}{2}\log\left(1+\frac{\beta_1 P_1}{N}\right) - C_{12} \\
&= \frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}_1\bar{\beta}_2^* P_1 P_2}}{N}\right) \\
&\quad - \frac{1}{2}\log\left(1+\frac{P_1\beta_1}{N}\right) \\
&\quad - \frac{1}{2}\log\left(1+\frac{\left(\sqrt{\bar{\beta}_1 P_1}+\sqrt{\bar{\beta}_2^* P_2}\right)^2}{\beta_1 P_1+\beta_2 P_2+N}\right) \\
&= \frac{1}{2}\log\left(1+\frac{\beta_2^* P_2}{\beta_1 P_1+N}\right) \\
&\le \frac{1}{2}\log\left(1+\frac{\beta_2^* P_2}{N}\right) \\
&= \frac{1}{2}\log\left(1+\frac{\beta_2' P_2}{N}\right).
\end{aligned}
$$

APPENDIX B
ON VECTORS UNIFORMLY DISTRIBUTED OVER n-SPHERES

We present some auxiliary results on vectors that are independently and uniformly drawn over centered unit n-spheres. Recall that for $n \in \mathbb{N}$ and $\theta \in [0,\pi)$ we denote by $C_n(\theta)$ the surface area of a spherical cap of half-angle θ on a unit n-sphere (see Section II).

The proofs of the following auxiliary lemmas are based on results in [22] and omitted.

Lemma B.1: Let $\boldsymbol{\Psi}$ be uniformly distributed over the centered unit n-sphere and μ a deterministic unit-length vector in $\mathbb{R}^n$. Then

$$\Pr\left[\langle \boldsymbol{\Psi}, \mu\rangle \ge \tau\right] = \frac{C_n(\arccos(\tau))}{C_n(\pi)}, \quad 0 \le \tau \le 1. \tag{89}$$

Lemma B.2: For $0 \le \tau < 1$

$$\lim_{n\to\infty}\frac{1}{n}\log\left(\frac{C_n(\arccos(\tau))}{C_n(\pi)}\right) = \frac{1}{2}\log(1-\tau^2). \tag{90}$$

Lemma B.3: Let $f:\mathbb{R}\to(0,1]$ be so that

$$-\eta_1 \stackrel{\triangle}{=} \lim_{n\to\infty}\frac{1}{n}\log f(n) \tag{91}$$

exists and $\eta_1 > 0$. Then

$$\lim_{n\to\infty}(1-f(n))^{2^{n\eta_2}} = \begin{cases} 1, & \text{if } \eta_1 > \eta_2 \\ 0, & \text{if } \eta_1 < \eta_2. \end{cases} \tag{92}$$

Lemma B.4: For $\theta \in (0,\pi/2)$

$$\lim_{n\to\infty}\frac{C_n(\theta)}{C_n(\pi)} = 0 \tag{93}$$

whereas for $\theta \in (\pi/2,\pi)$

$$\lim_{n\to\infty}\frac{C_n(\theta)}{C_n(\pi)} = 1. \tag{94}$$

We consider independent random vectors $\{\boldsymbol{\Psi}_\ell\}_{\ell=1}^{\lfloor 2^{n\lambda}\rfloor}$, $\{\tilde{\boldsymbol{\Psi}}_{\tilde{\ell}}\}_{\tilde{\ell}=1}^{\lfloor 2^{n\tilde{\lambda}}\rfloor}$, and $\mathbf{U}$ that are uniformly distributed over centered unit n-spheres and study properties of the vector in $\{\boldsymbol{\Psi}_\ell\}$ and the vector in $\{\tilde{\boldsymbol{\Psi}}_{\tilde{\ell}}\}$ that are closest to $\mathbf{U}$.

Lemma B.5: Let $\lambda \ge 0$ and $\gamma \in (0,1)$ be given. Let $n \in \mathbb{N}$, and let $\boldsymbol{\Psi}_1,\ldots,\boldsymbol{\Psi}_{\lfloor 2^{n\lambda}\rfloor}$ and $\mathbf{U}$ be independent random vectors uniformly distributed over the centered unit n-sphere. Then, if $\gamma > \sqrt{1-2^{-2\lambda}}$

$$\lim_{n\to\infty}\Pr\left[\bigcup_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}\left(\langle\boldsymbol{\Psi}_\ell,\mathbf{U}\rangle \ge \gamma\right)\right] = 0 \tag{95}$$

whereas if $\gamma < \sqrt{1-2^{-2\lambda}}$:

$$\lim_{n\to\infty}\Pr\left[\bigcup_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}\left(\langle\boldsymbol{\Psi}_\ell,\mathbf{U}\rangle \ge \gamma\right)\right] = 1. \tag{96}$$

Proof: We condition on $\mathbf{U} = \mathbf{u}$. By the independence of the vectors $\boldsymbol{\Psi}_1,\ldots,\boldsymbol{\Psi}_{\lfloor 2^{n\lambda}\rfloor}$, and $\mathbf{U}$

$$
\begin{aligned}
&\Pr\left[\bigcup_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}\left(\langle\boldsymbol{\Psi}_\ell,\mathbf{u}\rangle \ge \gamma\right)\middle|\mathbf{U}=\mathbf{u}\right] \\
&\quad = 1 - \prod_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}\left(1-\Pr\left[\langle\boldsymbol{\Psi}_\ell,\mathbf{u}\rangle \ge \gamma\right]\right).
\end{aligned} \tag{97}
$$

By the uniform distribution of $\boldsymbol{\Psi}_\ell$ and by Lemma B.1

$$\Pr\left[\langle\boldsymbol{\Psi}_\ell,\mathbf{u}\rangle \ge \gamma\right] = \frac{C_n(\arccos(\gamma))}{C_n(\pi)}, \quad \ell \in \left\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\right\} \tag{98}$$

irrespective of the unit-length vector $\mathbf{u}$. Inserting (98) into (97) and taking expectation with respect to $\mathbf{U}$, we obtain

$$
\begin{aligned}
&\Pr\left[\bigcup_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}\left(\langle\boldsymbol{\Psi}_\ell,\mathbf{U}\rangle \ge \gamma\right)\right] \\
&\quad = 1 - \left(1-\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)^{\lfloor 2^{n\lambda}\rfloor}.
\end{aligned} \tag{99}
$$

In the following, we distinguish the cases $\gamma > \sqrt{1-2^{-2\lambda}}$ and $\gamma < \sqrt{1-2^{-2\lambda}}$. We first treat the case $\gamma > \sqrt{1-2^{-2\lambda}}$, where

$$\frac{1}{2}\log\left(\frac{1}{1-\gamma^2}\right) > \lambda$$

and therefore by Lemma B.2

$$-\lim_{n\to\infty}\frac{1}{n}\log\left(\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right) > \lambda. \tag{100}$$

Notice that the probability in (99) can be upper bounded as

$$\Pr\left[\bigcup_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}(\langle\boldsymbol{\Psi}_\ell,\mathbf{U}\rangle\geq\gamma)\right] \leq 1-\left(1-\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)^{2^{n\lambda}} \tag{101}$$

where we used that $0\leq\left(1-\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)\leq 1$ and thus the mapping $x\mapsto\left(1-\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)^x$ is decreasing in $x>0$. The desired limit (95) follows then by (100), (101), and Lemma B.3.

In contrast, if $\gamma<\sqrt{1-2^{-2\lambda}}$, then

$$\frac{1}{2}\log\left(\frac{1}{1-\gamma^2}\right)<\lambda$$

and therefore by Lemma B.2:

$$-\lim_{n\to\infty}\frac{1}{n}\log\left(\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)<\lambda. \tag{102}$$

We choose $\tilde{\lambda}$ so that

$$-\lim_{n\to\infty}\frac{1}{n}\log\left(\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)<\tilde{\lambda}<\lambda \tag{103}$$

and notice that for sufficiently large n the probability in (99) can be lower bounded as

$$\Pr\left[\bigcup_{\ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}}(\langle\boldsymbol{\Psi}_\ell,\mathbf{U}\rangle\geq\gamma)\right] \geq 1-\left(1-\frac{C_n(\arccos(\gamma))}{C_n(\pi)}\right)^{2^{n\tilde{\lambda}}}. \tag{104}$$

The desired limit (96) follows then by combining (103), (104), and Lemma B.3. ■

Lemma B.6: Let $n\in\mathbb{N}$ and $\lambda,\tilde{\lambda}\geq 0$ be given, and let the vectors $\boldsymbol{\Psi}_1,\ldots,\boldsymbol{\Psi}_{\lfloor 2^{n\lambda}\rfloor}$, $\tilde{\boldsymbol{\Psi}}_1,\ldots,\tilde{\boldsymbol{\Psi}}_{\lfloor 2^{n\tilde{\lambda}}\rfloor}$, and $\mathbf{U}$ be IID random vectors uniformly distributed over the centered unit n-sphere. Let L^* be a random variable taking value in $\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\}$ and $\tilde{L}^*$ a random variable taking value in $\{1,\ldots,\lfloor 2^{n\tilde{\lambda}}\rfloor\}$ so that

$$\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle\geq\langle\boldsymbol{\Psi}_\ell,\mathbf{U}\rangle,\qquad \ell\in\{1,\ldots,\lfloor 2^{n\lambda}\rfloor\} \tag{105}$$

$$\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle\geq\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{\ell}},\mathbf{U}\right\rangle,\qquad \tilde{\ell}\in\{1,\ldots,\lfloor 2^{n\tilde{\lambda}}\rfloor\}. \tag{106}$$

For every γ, $\tilde{\gamma}\in(0,1)$ so that $\gamma<\sqrt{1-2^{-2\lambda}}$, and $\tilde{\gamma}<\sqrt{1-2^{-2\tilde{\lambda}}}$, and for every $\epsilon^*>\left(\sqrt{1-2^{-2\lambda}}-\gamma\right)\left(\sqrt{1-2^{-2\tilde{\lambda}}}-\tilde{\gamma}\right)$ and every $\epsilon>0$:

$$\lim_{n\to\infty}\Pr\left[-\epsilon\leq\left\langle\boldsymbol{\Psi}_{L^*}-\gamma\mathbf{U},\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}-\tilde{\gamma}\mathbf{U}\right\rangle\leq\epsilon^*\right]=1. \tag{107}$$

Proof: Choose $\epsilon_1^*,\epsilon_2^*>0$ so that

$$\epsilon_1^*\cdot\epsilon_2^*<\epsilon^* \tag{108}$$

$$\gamma+\epsilon_1^*>\sqrt{1-2^{-2\lambda}} \tag{109}$$

$$\tilde{\gamma}+\epsilon_2^*>\sqrt{1-2^{-2\tilde{\lambda}}}. \tag{110}$$

Since $\epsilon^*>\left(\sqrt{1-2^{-2\lambda}}-\gamma\right)\left(\sqrt{1-2^{-2\tilde{\lambda}}}-\tilde{\gamma}\right)$ it is always possible to find such $\epsilon_1^*,\epsilon_2^*>0$.

Define the vectors $\mathbf{W}\triangleq(\boldsymbol{\Psi}_{L^*}-\gamma\mathbf{U})$ and $\tilde{\mathbf{W}}\triangleq\left(\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}-\tilde{\gamma}\mathbf{U}\right)$, and decompose them into components that are orthogonal to $\mathbf{U}$ and components that are parallel to $\mathbf{U}$, i.e., decompose $\mathbf{W}$ into

$$\mathbf{W}^\perp\triangleq\boldsymbol{\Psi}_{L^*}-\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle\,\mathbf{U}$$

$$\mathbf{W}^\parallel\triangleq(\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle-\gamma)\,\mathbf{U}$$

and $\tilde{\mathbf{W}}$ into

$$\tilde{\mathbf{W}}^\perp\triangleq\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}-\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle\mathbf{U}$$

$$\tilde{\mathbf{W}}^\parallel\triangleq\left(\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle-\tilde{\gamma}\right)\mathbf{U}.$$

By the orthogonality of the components and because $\mathbf{U}$ has unit norm

$$\begin{aligned}\left\langle\mathbf{W},\tilde{\mathbf{W}}\right\rangle&=\left\langle\mathbf{W}^\perp,\tilde{\mathbf{W}}^\perp\right\rangle+\left\langle\mathbf{W}^\parallel,\tilde{\mathbf{W}}^\parallel\right\rangle\\&=\left\langle\mathbf{W}^\perp,\tilde{\mathbf{W}}^\perp\right\rangle+\left(\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle-\gamma\right)\left(\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle-\tilde{\gamma}\right).\end{aligned}$$

We notice that if the inner product $\left\langle\mathbf{W}^\perp,\tilde{\mathbf{W}}^\perp\right\rangle$ lies in the interval $[-\epsilon,(\epsilon^*-\epsilon_1^*\epsilon_2^*)]$, if the term $(\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle-\gamma)$ lies in $[0,\epsilon_1^*]$, and if $\left(\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle-\tilde{\gamma}\right)$ lies in $[0,\epsilon_2^*]$, then the inner product $\left\langle\mathbf{W},\tilde{\mathbf{W}}\right\rangle$ lies in $[-\epsilon,\epsilon^*]$. We thus obtain the bound (111) at the bottom of the page, and notice that to establish the lemma it suffices to prove the limits

$$\lim_{n\to\infty}T_1(n)=1 \tag{112}$$

$$\lim_{n\to\infty}T_2(n)=1 \tag{113}$$

where $T_1(n)$ and $T_2(n)$ are defined in (111). We first prove limit (113). To this end, notice that the inner products $\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle$ and $\left\langle\tilde{\boldsymbol{\Psi}}_{L^*},\mathbf{U}\right\rangle$ are independent, because $\{\boldsymbol{\Psi}_\ell\}$ and $\{\tilde{\boldsymbol{\Psi}}_\ell\}$ are independent of each other and of the unit-length vector $\mathbf{U}$ and they are IID uniformly distributed over the centered unit n-sphere. Hence

$$\begin{aligned}&\Pr\left[\gamma\leq\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle\leq\gamma+\epsilon_1^*,\tilde{\gamma}\leq\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle\leq\tilde{\gamma}+\epsilon_2^*\right]\\&=\Pr\left[\gamma\leq\langle\boldsymbol{\Psi}_{L^*},\mathbf{U}\rangle\leq\gamma+\epsilon_1^*\right]\\&\quad\cdot\Pr\left[\tilde{\gamma}\leq\left\langle\tilde{\boldsymbol{\Psi}}_{\tilde{L}^*},\mathbf{U}\right\rangle\leq\tilde{\gamma}+\epsilon_2^*\right].\end{aligned} \tag{114}$$

By (109), (110), and the assumptions on γ and $\tilde{\gamma}$ in the lemma

$$\gamma<\sqrt{1-2^{-2\lambda}}<\gamma+\epsilon_1^*$$

$$\tilde{\gamma}<\sqrt{1-2^{-2\tilde{\lambda}}}<\tilde{\gamma}+\epsilon_2^*.$$

Hence, by Lemma B.5, the two factors on the right-hand side of (114) both tend to 1 as n tends to infinity; this establishes the desired limit (113).

We now prove limit (112). To this end, we notice that conditional on $\langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\rangle = t_{\Psi U}$ and on $\langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\rangle = t_{\tilde{\Psi} U}$, the vectors $\mathbf{W}^{\perp}$ and $\tilde{\mathbf{W}}^{\perp}$ are independent and uniformly distributed over centered $(n-1)$-dimensional spheres of radii $\sqrt{1-t_{\Psi U}^2}$ and $\sqrt{1-t_{\tilde{\Psi} U}^2}$. Thus, by Lemma B.1, for $t_{\Psi U} \in [\gamma, \gamma+\epsilon_1^*]$ and $t_{\tilde{\Psi} U} \in [\tilde{\gamma}, \tilde{\gamma}+\epsilon_2^*]$, inequality (116) at the bottom of the page holds. where the last inequality follows because $C_{n-1}(\arccos(x))$ is monotonically decreasing in $x \in (-1,1)$. Since lower bound (116) is independent of $t_{\Psi U}$ and of $t_{\tilde{\Psi} U}$, taking expectation over $\gamma \le \langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\rangle \le \gamma+\epsilon_1^*$ and $\tilde{\gamma} \le \langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\rangle \le \tilde{\gamma}+\epsilon_2^*$ results in inequality (117) at the bottom of the page. By Lemma B.4 and because $\arccos\left(-\frac{\epsilon}{\sqrt{1-\gamma^2}\sqrt{1-\tilde{\gamma}^2}}\right) \in (\pi/2, \pi)$ whereas $\arccos\left(\frac{\epsilon^*-\epsilon_1^*\epsilon_2^*}{2\sqrt{1-\gamma^2}\sqrt{1-\tilde{\gamma}^2}}\right) \in (0, \pi/2)$, the right-hand side of (117) tends to 1 as the blocklength n tends to infinity. This establishes the desired limit (112), and thus concludes the proof. ■

APPENDIX C
PROOF OF LEMMA V.1

We prove Lemma V.1 in Appendix C-D. We first recall the setup and the notation of our single-user dirty-paper scheme in Appendix C-A. The notation and the assumptions are valid throughout this appendix. We then present some definitions in Appendix C-B and auxiliary lemmas in Appendix C-C.

A) Setup and Notation: We briefly recall the notation in the single-user dirty-paper scheme of Section V and the assumptions of Lemma V.1.

Recall that the scheme has parameters $\tilde{N} > 0$, $\tilde{P} > 0$, $\tilde{Q} \ge 0$, $\tilde{R} > 0$, and $\tilde{R}' > 0$, and that we defined $\tilde{\alpha} = \frac{\tilde{P}}{\tilde{P}+\tilde{N}}$. By the assumptions in the lemma, the parameter $\tilde{Q} = Q$ and the

$$\begin{aligned}\Pr\left[-\epsilon \le \left\langle \mathbf{W}, \tilde{\mathbf{W}}\right\rangle \le \epsilon^*\right] &\ge \Pr\left[-\epsilon \le \left\langle \mathbf{W}^{\perp}, \tilde{\mathbf{W}}^{\perp}\right\rangle \le (\epsilon^* - \epsilon_1^*\epsilon_2^*), 0 \le \left(\left\langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\right\rangle - \gamma\right) \le \epsilon_1^*, 0 \le \left(\left\langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\right\rangle - \tilde{\gamma}\right) \le \epsilon_2^*\right] \\ &= \underbrace{\Pr\left[-\epsilon \le \left\langle \mathbf{W}^{\perp}, \tilde{\mathbf{W}}^{\perp}\right\rangle \le (\epsilon^* - \epsilon_1^*\epsilon_2^*) \middle| 0 \le (\langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\rangle - \gamma) \le \epsilon_1^*, 0 \le \left(\left\langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\right\rangle - \tilde{\gamma}\right) \le \epsilon_2^*\right]}_{\triangleq T_1(n)} \\ &\quad \cdot \underbrace{\Pr\left[0 \le (\langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\rangle - \gamma) \le \epsilon_1^*, 0 \le \left(\left\langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\right\rangle - \tilde{\gamma}\right) \le \epsilon_2^*\right]}_{\triangleq T_2(n)} \end{aligned} \tag{111}$$

$$\Pr\left[-\epsilon \le \left\langle \mathbf{W}^{\perp}, \tilde{\mathbf{W}}^{\perp}\right\rangle \le (\epsilon^* - \epsilon_1^*\epsilon_2^*) \middle| \langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\rangle = t_{\Psi U}, \left\langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\right\rangle = t_{\tilde{\Psi} U}\right] \tag{115}$$

$$= \frac{C_{n-1}\left(\arccos\left(-\frac{\epsilon}{\sqrt{1-t_{\Psi U}^2}\sqrt{1-t_{\tilde{\Psi} U}^2}}\right)\right) - C_{n-1}\left(\arccos\left(\frac{\epsilon^*-\epsilon_1^*\epsilon_2^*}{\sqrt{1-t_{\Psi U}^2}\sqrt{1-t_{\tilde{\Psi} U}^2}}\right)\right)}{C_{n-1}(\pi)}$$

$$\ge \frac{C_{n-1}\left(\arccos\left(-\frac{\epsilon}{\sqrt{1-\gamma^2}\sqrt{1-\tilde{\gamma}^2}}\right)\right) - C_{n-1}\left(\arccos\left(\frac{\epsilon^*-\epsilon_1^*\epsilon_2^*}{\sqrt{1-\gamma^2}\sqrt{1-\tilde{\gamma}^2}}\right)\right)}{C_{n-1}(\pi)} \tag{116}$$

$$\Pr\left[-\epsilon \le \left\langle \mathbf{W}^{\perp}, \tilde{\mathbf{W}}^{\perp}\right\rangle \le (\epsilon^* - \epsilon_1^*\epsilon_2^*) \middle| \gamma \le \langle \boldsymbol{\Psi}_{L^*}, \mathbf{U}\rangle \le \gamma+\epsilon_1^*, \tilde{\gamma} \le \left\langle \tilde{\boldsymbol{\Psi}}_{\tilde{L}^*}, \mathbf{U}\right\rangle \le \tilde{\gamma}+\epsilon_2^*\right]$$

$$\ge \frac{C_{n-1}\left(\arccos\left(-\frac{\epsilon}{\sqrt{1-\gamma^2}\sqrt{1-\tilde{\gamma}^2}}\right)\right) - C_{n-1}\left(\arccos\left(\frac{\epsilon^*-\epsilon_1^*\epsilon_2^*}{\sqrt{1-\gamma^2}\sqrt{1-\tilde{\gamma}^2}}\right)\right)}{C_{n-1}(\pi)} \tag{117}$$

parameters $\tilde{N} > 0$, $\tilde{P} > 0$, $\tilde{R} > 0$, and $\tilde{R}' > 0$ are chosen in a way that (23) and (24) are satisfied and that there exists a number $\epsilon^* \in (0, \tilde{N})$ satisfying conditions (25)–(27). In the following, let ϵ^* be such a number.

Recall that $\mathbf{V}_{M,K^*}$ denotes the dirty-paper codeword produced by the encoder (see Section V-B), and that $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$, and $\mathbf{S}$ denote the vectors obtained when the input symbols, the output symbols, the noise symbols, and the interference symbols are stacked on top of each other. Also, recall that the inputs and outputs of our scheme can be written as $\mathbf{X} = \mathbf{V}_{M,K^*} - \tilde{\alpha}\mathbf{S}$ and $\mathbf{Y} = \mathbf{V}_{M,K^*} + (1-\tilde{\alpha})\mathbf{S} + \mathbf{Z}$.

B) Some Definitions: The following definitions and choices will be used in Sections C and D of this Appendix. Choose an $\epsilon_1 > 0$ such that

$$\frac{\tilde{P} + \tilde{\alpha}Q}{\sqrt{\tilde{P} + \tilde{\alpha}^2 Q}\sqrt{\tilde{P} + \tilde{N} + Q + (3 - 2\tilde{\alpha})\epsilon^*}} > \Upsilon(\epsilon_1) > \sqrt{1 - 2^{-2(\tilde{R}+\tilde{R}')}} \tag{118}$$

where $\Upsilon(\epsilon_1)$ is defined as

$$\Upsilon(\epsilon_1) \stackrel{\triangle}{=} \frac{\tilde{P} + \tilde{\alpha}Q}{\sqrt{\tilde{P} + \tilde{\alpha}^2 Q}\sqrt{\tilde{P} + \tilde{N} + Q}} - \epsilon_1. \tag{119}$$

By condition (25) such an ϵ_1 always exists. For brevity, in the following, we also write Υ instead of $\Upsilon(\epsilon_1)$.

Choose an $\epsilon_2 > 0$ such that

$$1 > \sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P} + \tilde{\alpha}^2 Q}} + \epsilon_2 > \sqrt{1 - 2^{-2\tilde{R}'}}. \tag{120}$$

By (120) and because R' satisfies (23)

$$\sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P} + \tilde{\alpha}^2 Q}} + \epsilon_2 > \sqrt{1 - 2^{-2\tilde{R}'}} > \sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P} + \tilde{\alpha}^2 Q}}. \tag{121}$$

Define the inner products

$$T_{VS} \stackrel{\triangle}{=} \left\langle \frac{\mathbf{V}_{M,K^*}}{\sqrt{n(\tilde{P} + \tilde{\alpha}^2 Q)}}, \frac{\mathbf{S}}{\sqrt{nQ}} \right\rangle \tag{122}$$

$$T_{VY} \stackrel{\triangle}{=} \left\langle \frac{\mathbf{V}_{M,K^*}}{\sqrt{n(\tilde{P} + \tilde{\alpha}^2 Q)}}, \frac{\mathbf{Y}}{\|\mathbf{Y}\|} \right\rangle \tag{123}$$

and the interval

$$\mathcal{I}_{VS} \stackrel{\triangle}{=} \left[\sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P} + \tilde{\alpha}^2 Q}}, \sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P} + \tilde{\alpha}^2 Q}} + \epsilon_2 \right]. \tag{124}$$

C) Auxiliary Lemmas: The following three auxiliary lemmas will be useful when proving Lemma V.1 in Section D of this appendix. The lemmas establish that for large blocklengths with very high probability, the chosen dirty-paper codeword $\mathbf{V}_{M,K^*}$ satisfies the following three properties.

1) The angle between $\mathbf{V}_{M,K^*}$ and $\mathbf{S}$ is close to $\arccos\left(\sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P} + \tilde{\alpha}^2 Q}}\right)$ (Lemma C.1).
2) The angle between $\mathbf{V}_{M,K^*}$ and $\mathbf{Z}$ is close to $\pi/2$ (Lemma C.2).
3) The angle between $\mathbf{V}_{M,K^*}$ and $\mathbf{Y}$ is smaller than $\arccos(\Upsilon(\epsilon_1))$ (Lemma C.3).

Lemma C.1: Let T_{VS} be defined as in (122) and let $\mathcal{I}_{VS}$ be defined as in (124) for an $\epsilon_2 > 0$ satisfying (120). Then,

$$\lim_{n\to\infty} \Pr\left[T_{VS} \in \mathcal{I}_{VS} | M = 1, K^* = 1\right] = 1.$$

Proof: By the symmetry of the code construction and the encoding, the probability that T_{VS} lies in the interval $\mathcal{I}_{VS}$ does not depend on the value of K^*. Thus

$$\Pr\left[T_{VS} \in \mathcal{I}_{VS} | M = 1, K^* = 1\right] = \Pr\left[T_{VS} \in \mathcal{I}_{VS} | M = 1\right] \tag{125}$$

and it suffices to show that the right-hand side of (125) tends to 1 as the blocklength n tends to infinity.

To this end, notice that the vectors $\left\{\frac{\mathbf{V}_{1,k}}{\sqrt{n(\tilde{P}+\tilde{\alpha}^2 Q)}}\right\}_{k=1}^{\lfloor 2^{n\tilde{R}'}\rfloor}$ and $\frac{\mathbf{S}}{\sqrt{nQ}}$ are IID random vectors uniformly distributed over the centered unit n-sphere, and thus, by the way the encoder chooses V_{M,K^*}, by the definition of T_{VS} and $\mathcal{I}_{VS}$, by assumption (121), and by Lemma B.5 (see Appendix B), limit (126) at the bottom of the page follows. Combined with (125), this concludes the proof. ■

Lemma C.2: For every $\tilde{\epsilon} > 0$

$$\lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\left\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\right\rangle\right| \le \tilde{\epsilon} \middle| M = 1, K^* = 1\right] = 1.$$

Proof: We only prove

$$\lim_{n\to\infty} \Pr\left[T_{VS} \in \mathcal{I}_{VS} | M = 1\right] = \lim_{n\to\infty} \left(\Pr\left[\bigcup_{k\in\{1,\ldots,\lfloor 2^{n\tilde{R}'}\rfloor\}} \left(\left\langle \frac{\mathbf{V}_{1,k}}{\sqrt{n(\tilde{P}+\tilde{\alpha}^2 Q)}}, \frac{\mathbf{S}}{\sqrt{nQ}} \right\rangle \ge \sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P}+\tilde{\alpha}^2 Q}} \right) \right] - \Pr\left[\bigcup_{k\in\{1,\ldots,\lfloor 2^{n\tilde{R}'}\rfloor\}} \left(\left\langle \frac{\mathbf{V}_{1,k}}{\sqrt{n(\tilde{P}+\tilde{\alpha}^2 Q)}}, \frac{\mathbf{S}}{\sqrt{nQ}} \right\rangle \ge \sqrt{\frac{\tilde{\alpha}^2 Q}{\tilde{P}+\tilde{\alpha}^2 Q}} + \epsilon_2 \right) \right] \right) = 1 \tag{126}$$

$$\lim_{n\to\infty} \Pr\left[\frac{1}{n}\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}\middle| M=1, K^*=1\right] = 1; \tag{127}$$

limit

$$\lim_{n\to\infty} \Pr\left[\frac{1}{n}\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \leq \tilde{\epsilon}\middle| M=1, K^*=1\right] = 1$$

follows then immediately from (127) and the symmetry of $\mathbf{S}$ and $\{\mathbf{V}_{m,k}\}$.

Fix $\tilde{\epsilon} > 0$, and choose $\epsilon_3 > 0$ depending on $\tilde{\epsilon}$ sufficiently small as will be described later. Define the inner product

$$T_{ZS} \stackrel{\triangle}{=} \left\langle \frac{\mathbf{Z}}{\|\mathbf{Z}\|}, \frac{\mathbf{S}}{\sqrt{nQ}} \right\rangle$$

and the intervals

$$\mathcal{I}_Z \stackrel{\triangle}{=} \left[\tilde{N} - \epsilon^*, \tilde{N} + \epsilon^*\right] \tag{128}$$

$$\mathcal{I}_{ZS} \stackrel{\triangle}{=} \left[-\epsilon_3, \frac{\epsilon^*}{\sqrt{\tilde{N}-\epsilon^*}\sqrt{Q}}\right] \tag{129}$$

where recall that by assumption $\tilde{N} > \epsilon^* > 0$, and thus $\mathcal{I}_Z$ and $\mathcal{I}_{ZS}$ are well defined.

In the following, we condition on the event $(M, K^*) = (1,1)$, and therefore, $\mathbf{V}_{M,K^*}$ is given by $\mathbf{V}_{1,1}$. We upper bound the probability that the inner product $\frac{1}{n}\langle \mathbf{V}_{1,1}, \mathbf{Z}\rangle$ exceeds $-\tilde{\epsilon}$ by (130) at the bottom of the page, where the first inequality follows by the law of total probability and dropping one of the terms; and the second inequality follows because for all events $\mathcal{A}$ and $\mathcal{B}$: $\Pr(\mathcal{A}\cap\mathcal{B}) \geq \Pr(\mathcal{A}) + \Pr(\mathcal{B}) - 1$. By Lemma C.1, the probability $\Pr[T_{VS} \in \mathcal{I}_{VS} | M=1, K^*=1]$ tends to 1 as n tends to infinity. Thus, to establish (127), it suffices to show the limits

$$\lim_{n\to\infty} T_3(n) = 1 \tag{131}$$

and

$$\lim_{n\to\infty} T_4(n) = 1 \tag{132}$$

where $T_3(n)$ and $T_4(n)$ are defined at the bottom of the page. We first show (131). To this end, we decompose the vectors $\mathbf{V}_{1,1}$ and $\mathbf{Z}$ into a part that is parallel to $\mathbf{S}$ and a part that is orthogonal to $\mathbf{S}$, i.e., we write $\mathbf{V}_{1,1} = \mathbf{V}_{1,1}^{\perp} + \mathbf{V}_{1,1}^{\parallel}$, where

$$\mathbf{V}_{1,1}^{\perp} \stackrel{\triangle}{=} \mathbf{V}_{1,1} - \sqrt{\frac{\tilde{P}+\tilde{\alpha}^2 Q}{Q}} T_{VS} \cdot \mathbf{S}$$

$$\mathbf{V}_{1,1}^{\parallel} \stackrel{\triangle}{=} \sqrt{\frac{\tilde{P}+\tilde{\alpha}^2 Q}{Q}} T_{VS} \cdot \mathbf{S}$$

and we write $\mathbf{Z} = \mathbf{Z}^{\parallel} + \mathbf{Z}^{\perp}$, where

$$\mathbf{Z}^{\perp} \stackrel{\triangle}{=} \mathbf{Z} - \frac{\|\mathbf{Z}\|}{\sqrt{nQ}} T_{ZS} \cdot \mathbf{S}$$

$$\mathbf{Z}^{\parallel} \stackrel{\triangle}{=} \frac{\|\mathbf{Z}\|}{\sqrt{nQ}} T_{ZS} \cdot \mathbf{S}.$$

The inner product of interest can then be expressed as the sum

$$\frac{1}{n}\langle \mathbf{V}_{1,1}, \mathbf{Z}\rangle = \frac{1}{n}\langle \mathbf{V}_{1,1}^{\perp}, \mathbf{Z}^{\perp}\rangle + \frac{1}{n}\langle \mathbf{V}_{1,1}^{\parallel}, \mathbf{Z}^{\parallel}\rangle. \tag{133}$$

In the following, we condition on the tuple $(\mathbf{S}, T_{VS}, T_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2) = (\mathbf{s}, t_{VS}, t_{ZS}, \varsigma^2)$ for some n-dimensional vector $\mathbf{s}$ of norm $\sqrt{nQ}$, some $t_{VS} \in \mathcal{I}_{VS}$, some $t_{ZS} \in \mathcal{I}_{ZS}$, and some $\varsigma \geq 0$ so that $\varsigma^2 \in \mathcal{I}_Z$. Then

$$\frac{1}{n}\|\mathbf{V}_{1,1}^{\perp}\|^2 = (\tilde{P}+\tilde{\alpha}^2 Q)(1-t_{VS}^2) \tag{134}$$

$$\begin{aligned}
&\Pr\left[\frac{1}{n}\langle \mathbf{V}_{1,1}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}\middle| M=1, K^*=1\right] \\
&\quad\geq \Pr\left[\frac{1}{n}\langle \mathbf{V}_{1,1}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}\middle| M=1, K^*=1, T_{ZS}\in\mathcal{I}_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z, T_{VS}\in\mathcal{I}_{VS}\right] \\
&\qquad\cdot \Pr\left[T_{ZS}\in\mathcal{I}_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2\in\mathcal{I}_Z, T_{VS}\in\mathcal{I}_{VS}\middle| M=1, K^*=1\right] \\
&\quad\geq \underbrace{\Pr\left[\frac{1}{n}\langle \mathbf{V}_{1,1}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}\middle| M=1, K^*=1, T_{ZS}\in\mathcal{I}_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z, T_{VS}\in\mathcal{I}_{VS}\right]}_{\stackrel{\triangle}{=}T_3(n)} \\
&\qquad\cdot \left(\underbrace{\Pr\left[T_{ZS}\in\mathcal{I}_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2\in\mathcal{I}_Z\middle| M=1, K^*=1\right]}_{\stackrel{\triangle}{=}T_4(n)} + \Pr[T_{VS}\in\mathcal{I}_{VS}|M=1,K^*=1] - 1\right)
\end{aligned} \tag{130}$$

$$\frac{1}{n}\|\mathbf{Z}^{\perp}\|^2 = \varsigma^2(1-t_{ZS}^2) \tag{135}$$

$$\frac{1}{n}\langle \mathbf{V}_{1,1}^{\|}, \mathbf{Z}^{\|}\rangle = \sqrt{\tilde{P}+\tilde{\alpha}^2 Q}\cdot t_{VS}t_{ZS}\varsigma \tag{136}$$

and the vector $\mathbf{V}_{1,1}^{\perp}$ is uniformly distributed over the centered $(n-1)$-sphere of radius $\sqrt{n(\tilde{P}+\tilde{\alpha}^2Q)(1-t_{VS}^2)}$ independent of $\mathbf{Z}^{\perp}$. We combine these observations with (133) to obtain expression (137) at the bottom of the page, where the first equality follows by (133) and (136), and the second by the uniform distribution of $\mathbf{V}_{1,1}^{\perp}$, by Lemma B.1 (see Appendix B), and by (134) and (135).

Notice that the expression on the right-hand side of (137) only depends on t_{VS}, t_{ZS}, and ς, but not on $\mathbf{s}$. Further, notice that the mapping $x \mapsto C_{n-1}(\arccos(x))$ is monotonically decreasing for $x \in [-1,1]$ and the mapping $x \mapsto \frac{x}{\sqrt{1-x^2}}$ is monotonically increasing for $x \in [-1,1]$. We can thus lower bound the right-hand side of (137) if we replace t_{VS}, t_{ZS}, and ς by the corresponding boundary points of the intervals $\mathcal{I}_{VS}$, $\mathcal{I}_{ZS}$, and $\mathcal{I}_Z$. Taking then expectation with respect to $T_{VS} \in \mathcal{I}_{VS}$, $T_{ZS} \in \mathcal{I}_{ZS}$, $\frac{1}{n}\|\mathbf{Z}\| \in \mathcal{I}_Z$, and $\mathbf{S}$ results in the bound (138) at the bottom of the page. When $\epsilon_3 > 0$ is chosen sufficiently small, the argument of the $\arccos$ is negative, and consequently, the $\arccos$ lies in the interval $(\pi/2, \pi)$. Hence by Lemma B.4 (see Appendix B), for sufficiently small $\epsilon_3 > 0$, the right-hand side of (138) tends to 1 as n tends to infinity, which establishes (131).

It remains to prove (132). To this end, notice that if the noise vector $\mathbf{Z}$ satisfies both

$$-\epsilon_3\sqrt{Q}\sqrt{(\tilde{N}-\epsilon^*)} \le \frac{1}{n}\langle \mathbf{Z},\mathbf{S}\rangle \le \epsilon^*$$

and

$$\tilde{N}-\epsilon^* \le \frac{1}{n}\|\mathbf{Z}\|^2 \le \tilde{N}+\epsilon^*$$

then $T_{ZS} \in \mathcal{I}_{ZS}$, i.e.,

$$-\epsilon_3 \le \left\langle \frac{\mathbf{Z}}{\|\mathbf{Z}\|}, \frac{\mathbf{S}}{\sqrt{nQ}} \right\rangle \le \frac{\epsilon^*}{\sqrt{\tilde{N}-\epsilon^*}\sqrt{Q}}.$$

Thus, the event that both $T_{ZS} \in \mathcal{I}_{ZS}$ and $\frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z$ are satisfied, includes the event that both $-\epsilon_3\sqrt{Q}\sqrt{(\tilde{N}-\epsilon^*)} \le \frac{1}{n}\langle \mathbf{Z},\mathbf{S}\rangle \le \epsilon^*$ and $\frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z$ are satisfied. Therefore

$$\begin{aligned}
&\Pr\left[\frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z, T_{ZS} \in \mathcal{I}_{ZS} \middle| M=1, K^*=1\right] \\
&\ge \Pr\left[\frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z, \right. \\
&\qquad \left. -\epsilon_3\sqrt{Q(\tilde{N}-\epsilon^*)} \le \frac{1}{n}\langle \mathbf{Z},\mathbf{S}\rangle \le \epsilon^* \middle| M=1, K^*=1\right] \\
&= \Pr\left[\frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z, -\epsilon_3\sqrt{Q(\tilde{N}-\epsilon^*)} \le \frac{1}{n}\langle \mathbf{Z},\mathbf{S}\rangle \le \epsilon^*\right]
\end{aligned} \tag{139}$$

where the last equality follows because by symmetry of the code construction and the encoding, the law of the pair $(\mathbf{Z},\mathbf{S})$ does not depend on the pair (M,K^*).

Since, by Assumptions (26) and (27) in Lemma V.1, for every $\epsilon_3 > 0$

$$\lim_{n\to\infty} \Pr\left[-\epsilon_3\sqrt{Q(\tilde{N}-\epsilon^*)} \le \frac{1}{n}\langle \mathbf{Z},\mathbf{S}\rangle \le \epsilon^*, \frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z\right] = 1 \tag{140}$$

$$\begin{aligned}
&\Pr\left[\frac{1}{n}\langle \mathbf{V}_{1,1},\mathbf{Z}\rangle \ge -\tilde{\epsilon} \middle| M=1, K^*=1, \mathbf{S}=\mathbf{s}, T_{ZS}=t_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2=\varsigma^2, T_{VS}=t_{VS}\right] \\
&\quad = \Pr\left[\frac{1}{n}\langle \mathbf{V}_{1,1}^{\perp},\mathbf{Z}^{\perp}\rangle \ge -\tilde{\epsilon} - \sqrt{(\tilde{P}+\tilde{\alpha}^2Q)}t_{VS}t_{ZS}\varsigma \right| M=1, K^*=1, \mathbf{S}=\mathbf{s}, \\
&\qquad\qquad T_{ZS}=t_{ZS}, \|\mathbf{Z}\|^2=\varsigma^2, T_{VS}=t_{VS}\Bigg] \\
&\quad = \frac{C_{n-1}\left(\arccos\left(-\frac{\tilde{\epsilon}}{\sqrt{(\tilde{P}+\tilde{\alpha}^2Q)(1-t_{VS}^2)(1-t_{ZS}^2)}\varsigma} - \frac{t_{VS}t_{ZS}}{\sqrt{(1-t_{VS}^2)(1-t_{ZS}^2)}}\right)\right)}{C_{n-1}(\pi)}
\end{aligned} \tag{137}$$

$$\begin{aligned}
&\Pr\left[\frac{1}{n}\langle \mathbf{V}_{1,1},\mathbf{Z}\rangle \ge -\tilde{\epsilon} \middle| M=1, K^*=1, T_{ZS}\in\mathcal{I}_{ZS}, \frac{1}{n}\|\mathbf{Z}\|^2 \in \mathcal{I}_Z, T_{VS}\in\mathcal{I}_{VS}\right] \\
&\quad \ge \frac{C_{n-1}\left(\arccos\left(-\frac{\tilde{\epsilon}}{\sqrt{\tilde{P}\left(1-\min\left\{\epsilon_3^2, \frac{(\epsilon^*)^2}{(\tilde{N}-\epsilon^*)Q}\right\}\right)(\tilde{N}+\epsilon^*)}} + \frac{\epsilon_3\left(\sqrt{\frac{\tilde{\alpha}^2Q}{\tilde{P}+\tilde{\alpha}^2Q}}+\epsilon_2\right)}{\sqrt{\left(1-\left(\sqrt{\frac{\tilde{\alpha}^2Q}{\tilde{P}+\tilde{\alpha}^2Q}}+\epsilon_2\right)^2\right)(1-\epsilon_3^2)}}\right)\right)}{C_{n-1}(\pi)}
\end{aligned} \tag{138}$$

the desired limit (132) follows. ■

Lemma C.3: Let $\Upsilon(\epsilon_1)$ be defined as in (119) for some $\epsilon_1 > 0$ so that (118) holds, and let T_{VY} be defined as in (123). Then

$$\lim_{n\to\infty} \Pr\left[T_{VY} \geq \Upsilon(\epsilon_1) | M = 1, K^* = 1\right] = 1.$$

Proof: Fix an $\tilde{\epsilon} > 0$ so that

$$\frac{\tilde{P} + \tilde{\alpha}^2 Q - \tilde{\epsilon}}{\sqrt{\tilde{P} + \tilde{\alpha}^2 Q}\sqrt{\tilde{P} + \tilde{N} + Q - 2\tilde{\epsilon} + (3 - 2\tilde{\alpha})\epsilon^*}} \geq \Upsilon(\epsilon_1). \tag{141}$$

By (118), such an $\tilde{\epsilon} > 0$ always exists.

To prove Lemma C.3, we first notice that by the law of total probability and dropping, one of the terms (142) at the bottom of the page is obtained. We can thus prove the lemma by showing that the two terms $T_5(n)$ and $T_6(n)$ in (142) tend to 1 as the blocklength n tends to infinity. We first show that

$$\lim_{n\to\infty} T_6(n) = 1. \tag{143}$$

By Lemmas C.1 and C.2, the probabilities of events $T_{VS} \in \mathcal{I}_{VS}$ and $\frac{1}{n}\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}$ tend to 1 as n tends to infinity. Further, by assumptions (26) and (27) in Lemma V.1 on the noise vector $\mathbf{Z}$ and by the symmetry of the code construction also the conditional probabilities of events $\frac{1}{n}\|\mathbf{Z}\|^2 \leq \tilde{N} + \epsilon^*$ and $\frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \leq \epsilon^*$ given $(M, K^*) = (1, 1)$ tend to 1 as n tends to infinity. Limit (143) then follows by combining these four limits.

It remains to prove limit

$$\lim_{n\to\infty} T_5(n) = 1. \tag{144}$$

We show that conditional on $(M, K^*) = (1, 1)$, on $T_{VS} \in \mathcal{I}_{VS}$, on $\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}$, on $\frac{1}{n}\|\mathbf{Z}\|^2 \leq \tilde{N} + \epsilon^*$, and on $\frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \leq \epsilon^*$, for arbitrary blocklength n

$$T_{VY} \geq \Upsilon \tag{145}$$

holds with probability 1, which establishes (144). The inner product T_{VY} is given by (146) at the bottom of the page. Notice that the right-hand side of (146) is monotonically increasing in $\frac{1}{n}\langle \mathbf{V}_{1,1}, (1 - \tilde{\alpha})\mathbf{S} + \mathbf{Z}\rangle$ and monotonically decreasing in $\frac{1}{n}\|\mathbf{Z}\|^2$ and $\frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle$. Hence, conditioned on $(M, K^*) = (1, 1)$, on $T_{VS} \in \mathcal{I}_{VS}$, on $\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}$, on $\frac{1}{n}\|\mathbf{Z}\|^2 \leq \tilde{N} + \epsilon^*$, and on $\frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \leq \epsilon^*$

$$T_{VY} \geq \frac{\tilde{P} + \tilde{\alpha} Q - \tilde{\epsilon}}{\sqrt{\tilde{P} + \tilde{\alpha}^2 Q}\sqrt{\tilde{P} + Q + \tilde{N} - 2\tilde{\epsilon} + (3 - 2\tilde{\alpha})\epsilon^*}}. \tag{147}$$

Combining (147) with (141), we conclude that conditioned on $(M, K^*) = (1, 1)$, on $T_{VS} \in \mathcal{I}_{VS}$, on $\langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}$, on $\frac{1}{n}\|\mathbf{Z}\|^2 \leq \tilde{N} + \epsilon^*$, and on $\frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \leq \epsilon^*$, for every $n \in \mathbb{N}$, inequality (145) holds with probability 1. This establishes Limit (144), and thus combined with (143) and (142) the proof of the lemma. ■

D) Proof of Lemma V.1: We first prove that the probability of error tends to 0 as the block-length n tends to infinity. By the symmetry of the code construction and the encoding, we can assume without loss of generality that $(M, K^*) = (1, 1)$. Then, the error event is given by $(1, 1) \neq (\hat{M}, \hat{K})$, which by

$$\begin{aligned}
&\Pr\left[T_{VY} \geq \Upsilon | M = 1, K^* = 1\right] \\
&\geq \underbrace{\Pr\left[T_{VY} \geq \Upsilon \middle| M = 1, K^* = 1, T_{VS} \in \mathcal{I}_{VS}, \langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}, \frac{1}{n}\|\mathbf{Z}\|^2 \leq \tilde{N} + \epsilon^*, \frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \leq \epsilon^*\right]}_{\triangleq T_5(n)} \\
&\quad \cdot \underbrace{\Pr\left[T_{VS} \in \mathcal{I}_{VS}, \langle \mathbf{V}_{M,K^*}, \mathbf{Z}\rangle \geq -\tilde{\epsilon}, \frac{1}{n}\|\mathbf{Z}\|^2 \leq N + \epsilon^*, \frac{1}{n}\langle \mathbf{Z}, \mathbf{S}\rangle \leq \epsilon^* \middle| M = 1, K^* = 1\right]}_{\triangleq T_6(n)}
\end{aligned} \tag{142}$$

$$\begin{aligned}
T_{VY} = &\frac{\tilde{P} + \tilde{\alpha}^2 Q + \frac{1}{n}\langle \mathbf{V}_{1,1}, (1 - \tilde{\alpha})\mathbf{S} + \mathbf{Z}\rangle}{\sqrt{\tilde{P} + \tilde{\alpha}^2 Q}} \\
&\frac{1}{\sqrt{\tilde{P} + (1 - 2\tilde{\alpha} + 2\tilde{\alpha}^2)Q + \frac{2}{n}\langle \mathbf{V}_{1,1}, (1 - \tilde{\alpha})\mathbf{S} + \mathbf{Z}\rangle + \frac{1}{n}\|\mathbf{Z}\|^2 + \frac{2(1-\tilde{\alpha})}{n}\langle \mathbf{Z}, \mathbf{S}\rangle}}
\end{aligned} \tag{146}$$

the nearest neighbor decoding rule in Section V-D is equivalent to the event

$$\bigcup_{\substack{m\in\{1,\ldots,\lfloor 2^{n\tilde{R}}\rfloor\}\\ k\in\{1,\ldots,\lfloor 2^{n\tilde{R}'}\rfloor\}\\ (m,k)\neq(1,1)}} \left(\langle \mathbf{V}_{m,k},\mathbf{Y}\rangle \geq \langle \mathbf{V}_{1,1},\mathbf{Y}\rangle\right).$$

Thus, Limit (148) at the bottom of the page establishes the desired claim on the probability of error. To prove (148), we first bound the probability as in inequality (149) at the bottom of the page, where $T_{VS}, T_{VY}, \mathcal{I}_{VS}$, and $\Upsilon(\epsilon_1)$ are defined in the previous Section C of this appendix. By Lemmas C.1 and C.3, the terms $T_7(n)$ and $T_8(n)$ tend to 0 as the blocklength n tends to infinity, i.e.,

$$\lim_{n\to\infty} T_7(n) = 0 \tag{150}$$

$$\lim_{n\to\infty} T_8(n) = 0. \tag{151}$$

It thus remains to prove limit

$$\lim_{n\to\infty} T_9(n) = 0. \tag{152}$$

For $m \in \{1,\ldots,\lfloor 2^{n\tilde{R}}\rfloor\}$ and $k \in \{1,\ldots,\lfloor 2^{n\tilde{R}'}\rfloor\}$, we define the unit-norm vectors

$$\boldsymbol{\Psi}_{m,k} \stackrel{\triangle}{=} \frac{\mathbf{V}_{m,k}}{\sqrt{n(\tilde{P}+\tilde{\alpha}^2 Q)}}. \tag{153}$$

Not conditioned on the pair (M,K^*), these vectors are independent and uniformly distributed over the centered unit n-sphere. That means their density[5] over the unit n-sphere is given by $\frac{1}{C_n(\pi)}$. Conditioned on $(M,K^*) = (1,1)$ however, this is not true anymore. In the following, we condition on $(M,K^*) = (1,1)$ and on $(\mathbf{S},T_{VS},T_{VY},\mathbf{Y}) = (\mathbf{s},t_{VS},t_{VY},\mathbf{y})$, for some n-dimensional vector $\mathbf{s}$ with norm $\sqrt{nQ}$, some $t_{VS} \in \mathcal{I}_{VS}$, some $t_{VY} \geq \Upsilon$, and some n-dimensional vector $\mathbf{y}$. Conditional on $(M,K^*) = (1,1)$ and $(\mathbf{S},T_{VS},T_{VY},\mathbf{Y}) = (\mathbf{s},t_{VS},t_{VY},\mathbf{y})$, the vectors $\{\boldsymbol{\Psi}_{m,k}\}$, for pairs (m,k) not equal to $(1,1)$, are still independent, but for some vectors the density over the centered unit n-sphere has changed. Whereas the vectors $\{\boldsymbol{\Psi}_{m,k}\}$, for $m \neq 1$, are still uniformly distributed over the entire centered unit n-sphere, the vectors $\{\boldsymbol{\Psi}_{1,k}\}$, for $k \neq 1$, are uniformly distributed only over the centered unit n-sphere without the spherical cap of half-angle $\arccos(t_{VS})$ centered at $\mathbf{s}$. Since the surface area of the sphere without this cap equals $C_n(\pi) - C_n(\arccos(t_{VS}))$, which for all $t_{VS} \in \mathcal{I}_{VS}$ is larger than $\frac{1}{2}C_n(\pi)$, we obtain that for all vectors $\{\boldsymbol{\Psi}_{m,k}\}$ where $(m,k) \neq (1,1)$, the conditional density on the centered unit n-sphere is upper bounded by $\frac{2}{C_n(\pi)}$.

With this upper bound and with Lemma B.1 (see Appendix B) we obtain the first inequality in (154) at the bottom of the next page, independent of $\mathbf{s}$, $t_{VS} \in \mathcal{I}_{VS}$, and $\mathbf{y}$. The last inequality in (154) follows because $t_{VY} > 0$ and thus the difference $\left(1 - 2\frac{C_n(\arccos(t_{VY}))}{C_n(\pi)}\right)$ lies in the open interval $(0,1)$ and the

[5]Here, we abuse terminology. In fact it is not a density with respect to the Lebesgue measure, but only with respect to the measure induced by the uniform distribution over the centered unit n-sphere.

$$\lim_{n\to\infty} \Pr\left[\bigcup_{\substack{m\in\{1,\ldots,\lfloor 2^{n\tilde{R}}\rfloor\}\\ k\in\{1,\ldots,\lfloor 2^{n\tilde{R}'}\rfloor\}\\ (m,k)\neq(1,1)}} \left(\langle \mathbf{V}_{m,k},\mathbf{Y}\rangle \geq \langle \mathbf{V}_{1,1},\mathbf{Y}\rangle\right) \middle| M=1,K^*=1\right] = 0 \tag{148}$$

$$\begin{aligned}
&\Pr\left[\bigcup_{(m,k)\neq(1,1)} \left(\langle \mathbf{V}_{m,k},\mathbf{Y}\rangle \geq \langle \mathbf{V}_{1,1},\mathbf{Y}\rangle\right) \middle| M=1,K^*=1\right]\\
&\quad\leq \underbrace{\Pr\left[T_{VS}\notin\mathcal{I}_{VS}|M=1,K^*=1\right]}_{T_7(n)} + \underbrace{\Pr\left[T_{VY}<\Upsilon|M=1,K^*=1\right]}_{\stackrel{\triangle}{=}T_8(n)}\\
&\qquad+\underbrace{\Pr\left[\bigcup_{(m,k)\neq(1,1)} \left(\langle \mathbf{V}_{m,k},\mathbf{Y}\rangle \geq \langle \mathbf{V}_{1,1},\mathbf{Y}\rangle\right) \middle| M=1,K^*=1,T_{VY}\geq\Upsilon,T_{VS}\in\mathcal{I}_{VS}\right]}_{\stackrel{\triangle}{=}T_9(n)}
\end{aligned} \tag{149}$$

mapping $x \mapsto \left(1-2\frac{C_n(\arccos(t_{VY}))}{C_n(\pi)}\right)^x$ is decreasing in $x>0$. Since the function $x \mapsto C_n(\arccos(x))$ is monotonically decreasing in $x \in [-1,1]$, for all $t_{VY} \geq \Upsilon$ inequality (154) is further upper bounded by $1-\left(1-2\frac{C_n(\arccos(\Upsilon))}{C_n(\pi)}\right)^{2^{n(\tilde{R}+\tilde{R}')}}$, and therefore, taking expectation with respect to $T_{VY} \geq \Upsilon$, $T_{VS} \in \mathcal{I}_{VS}$, $\mathbf{S}$, and $\mathbf{Y}$ results in (155) at the bottom of the page. Notice that because

$$\Upsilon > \sqrt{1-2^{-2(\tilde{R}+\tilde{R}')}}$$

by [2, eq. (54)], and since the constant factor 2 in the logarithm does not change the limit

$$-\lim_{n\to\infty}\frac{1}{n}\log\left(2\frac{C_n(\arccos(\Upsilon))}{C_n(\pi)}\right)=\frac{1}{2}\log\left(\frac{1}{1-\Upsilon^2}\right) > (\tilde{R}+\tilde{R}'). \tag{156}$$

The desired limit (152) follows then by Lemma B.3 (see Appendix B) and by inequalities (155) and (156).

We next consider the probability that the input sequence $\mathbf{X}=\mathbf{V}_{M,K^*}-\tilde{\alpha}\mathbf{S}$ violates the average block-power constraint P. By symmetry, again this probability does not depend on the value of M and K^*, and we can assume $M=K^*=1$. Thus, we wish to prove that

$$\lim_{n\to\infty}\Pr\left[\frac{1}{n}\|\mathbf{X}\|^2 \leq P \middle| M=K^*=1\right]=1. \tag{157}$$

Notice that

$$\frac{1}{n}\|\mathbf{X}\|^2=\frac{1}{n}\|\mathbf{V}_{M,K^*}-\tilde{\alpha}\mathbf{S}\|^2 \tag{158}$$

$$=2\tilde{\alpha}^2Q+P-\frac{2\tilde{\alpha}\langle\mathbf{V}_{M,K^*},\mathbf{S}\rangle}{n} \tag{159}$$

and by Lemma C.1

$$\lim_{n\to\infty}\Pr\left[\frac{\langle\mathbf{V}_{M,K^*},\mathbf{S}\rangle}{n}\geq\tilde{\alpha}Q\middle|M=K^*=1\right]=1. \tag{160}$$

Combining (158) with (160) establishes (157) and thus concludes the proof.

Appendix D
Proof of Theorem IV.1

We show that every rate $R>0$ satisfying the conditions in Theorem IV.1 is achievable. To this end, we present for every rate below R a choice of parameters so that our scheme in Section V communicates at this rate with vanishing probability of error and with an average transmit power no larger than P. This establishes the theorem.

We fix a rate $R>0$ satisfying the conditions in Theorem IV.1. Then, we fix a sufficiently small $\delta>0$ and corresponding $N>0$ and $\epsilon^*>0$ so that assumptions 1–3 in the theorem are met. We now describe the choice of parameters for which we apply our scheme in Section V. We choose

$$\begin{aligned}\tilde{R}&\stackrel{\triangle}{=}R-2\delta\\ \tilde{Q}&\stackrel{\triangle}{=}Q\\ \tilde{N}&\stackrel{\triangle}{=}N\end{aligned} \tag{161}$$

and $\tilde{P}\in(0,P)$ sufficiently close to P such that

$$\tilde{R}\leq\frac{1}{2}\log\left(1+\frac{\tilde{P}}{\tilde{N}}\right)-\delta \tag{162}$$

and

$$\frac{\tilde{P}+\tilde{\alpha}Q}{\sqrt{\tilde{P}+\tilde{\alpha}^2Q}\sqrt{\tilde{P}+\tilde{N}+Q+(3-2\tilde{\alpha})\epsilon^*}} > \sqrt{1-2^{-2(R-\frac{3}{2}\delta)}\frac{\tilde{P}}{\tilde{\alpha}^2Q+\tilde{P}}} \tag{163}$$

where recall that $\tilde{\alpha}\stackrel{\triangle}{=}\frac{\tilde{P}}{\tilde{P}+\tilde{N}}$.

Notice that such a choice of $\tilde{P}$ always exists, as proved by the following three observations:

$$\begin{aligned}&\Pr\left[\bigcup_{(m,k)\neq(1,1)}\left(\left\langle\frac{\mathbf{V}_{m,k}}{\sqrt{n(\tilde{P}+\tilde{\alpha}^2Q)}},\frac{\mathbf{y}}{\|\mathbf{y}\|}\right\rangle\geq T_{VY}\right)\middle|M=1,K^*=1,\mathbf{S}=\mathbf{s},T_{VY}=t_{VY},T_{VS}=t_{VS},\mathbf{Y}=\mathbf{y}\right]\\ &\quad\leq 1-\left(1-2\frac{C_n(\arccos(t_{VY}))}{C_n(\pi)}\right)^{\lfloor 2^{n\tilde{R}}\rfloor\lfloor 2^{n\tilde{R}'}\rfloor-1}\\ &\quad\leq 1-\left(1-2\frac{C_n(\arccos(t_{VY}))}{C_n(\pi)}\right)^{2^{n(\tilde{R}+\tilde{R}')}}\end{aligned} \tag{154}$$

$$\begin{aligned}&\Pr\left[\bigcup_{(m,k)\neq(1,1)}\left(\left\langle\frac{\mathbf{V}_{m,k}}{\sqrt{n(\tilde{P}+\tilde{\alpha}^2Q)}},\frac{\mathbf{Y}}{\|\mathbf{Y}\|}\right\rangle\geq T_{VY}\right)\middle|M=1,K^*=1,T_{VY}\geq\Upsilon,T_{VS}\in\mathcal{I}_{VS}\right]\\ &\quad\leq 1-\left(1-2\frac{C_n(\arccos(\Upsilon))}{C_n(\pi)}\right)^{2^{n(\tilde{R}+\tilde{R}')}}\end{aligned} \tag{155}$$

1) Every rate R that satisfies condition (16) in the theorem for arbitrary small $\delta > 0$ must be smaller than $\frac{1}{2}\log\left(1+\frac{P}{N}\right)$. This holds because for $\delta = \epsilon^* = 0$ and $R = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$ the left-hand side of (16) equals its right-hand side; because the left-hand side of (16) is independent of δ and decreasing in ϵ^*; and because its right-hand side is independent of ϵ^* and increasing in δ.
2) Therefore, by (161) and by continuity, inequality (162) holds for all $\tilde{P}$ that are sufficiently close to P.
3) Also inequality (163) holds for all $\tilde{P}$ that are sufficiently close to P. This follows again by condition (16) in the theorem, by the continuity of the expressions involved in (16), and because the right-hand side of (16) is decreasing in δ.

We also choose

$$\tilde{R}' \stackrel{\triangle}{=} \frac{1}{2}\log\left(1+\frac{\tilde{\alpha}^2 Q}{\tilde{P}}\right) + \delta/2. \tag{164}$$

In the following, we show that our choice of parameters $(\tilde{N}, \tilde{P}, \tilde{Q}, \tilde{R}, \tilde{R}')$ satisfies the conditions in Lemma V.1 for the chosen ϵ^*. By Lemma V.1 and Remark V.3, this then establishes that the probability of error of our scheme and the probability of violating the average block-power constraint tend to 0 as the blocklength tends to infinity.

We first notice that conditions (23) and (24) in Lemma V.1 follow from (162) and (164), from the fact that δ is positive, and from simple algebraic manipulations. Moreover, conditions ii) and iii) in Lemma V.1 hold trivially because they coincide with conditions 2 and 3 in the theorem. Finally, condition i) in Lemma V.1 follows from (163) and because by (161) and (164)

$$\begin{aligned}\sqrt{1-2^{-2(\tilde{R}+\tilde{R}')}} &= \sqrt{1-2^{-2(R-2\delta)}\frac{\tilde{P}}{\tilde{\alpha}^2 Q+\tilde{P}}2^{-\delta}} \\ &= \sqrt{1-2^{-2(R-\frac{3}{2}\delta)}\frac{\tilde{P}}{\tilde{\alpha}^2 Q+\tilde{P}}}.\end{aligned}$$

Appendix E
Proof of Lemmas VI.3 and VI.6

Before proving Lemmas VI.3 and VI.6 in Sections B and C of this Appendix, we recall the setup and notation of Section VI, and we present some definitions and lemmas in Section A of this Appendix.

Consider the dirty-paper MAC with conferencing in Section III-B where the transmitters have powers P_1 and P_2, the interference vector $\mathbf{S}$ is uniformly distributed over a centered n-sphere of radius $\sqrt{nQ}$, the noise vector $\mathbf{Z}$ has IID zero-mean variance-N Gaussian components, and the two transmitters have access to noise-free pipes of rates C_{12} and C_{21}. We consider the scheme in Section VI. That means, we assume that in a first stage the two transmitters exchange the common parts of their messages $M_{1,c}$ and $M_{2,c}$ over the pipes, and in a second stage the transmitters and the receiver apply the scheme in Section VI-A for some choice of parameters $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, $\delta_0, \delta_1, \delta_2 > 0$.

Let, in the following, the parameters $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, and $\delta_0, \delta_1, \delta_2 > 0$ be fixed. Given these parameters, let P_0, $P_{1,p}$, $P_{2,p}$, N_0, N_1, N_2, and Q_0, Q_1, Q_2 be the powers, the noise variances, and the interference variances as defined in Section VI-A1; let $\mathbf{S}_0, \mathbf{S}_1, \mathbf{S}_2$ and $\mathbf{Z}_0, \mathbf{Z}_1, \mathbf{Z}_2$ be the interference and noise sequences experienced in the various decoding steps as defined in Sections VI-A1 and VI-A5; let $R_0', R_1', R_2', R_0, R_1, R_2$ be the rates defined through (40)–(45); let $\mathbf{V}_{0,M_0,K_0^*}$, $\mathbf{V}_{1,M_{1,p},K_1^*}$, and $\mathbf{V}_{2,M_{2,p},K_2^*}$ be the dirty-paper codewords and $\mathbf{X}_0'$, $\mathbf{X}_1'$, and $\mathbf{X}_2'$ the corresponding dirty-paper sequences as defined in Sections VI-A2 and VI-A3.

A) Some Definitions and Auxiliary Lemmas:

Definition E.1: For fixed $\epsilon', \epsilon_0', \epsilon_1', \epsilon_2' > 0$, define the following 14 events.

1) Let $\mathcal{E}_0(\epsilon_0')$ denote the event that

$$\alpha_0 Q_0 \le \frac{1}{n}\left\langle \mathbf{V}_{0,M_0,K_0^*}, \mathbf{S}_0\right\rangle \le \alpha_0 Q_0 + \epsilon_0'.$$

2) For $i \in \{1,2\}$, let $\mathcal{E}_i(\epsilon_i')$ denote the event that

$$\alpha_i Q_i \le \frac{1}{n}\left\langle \mathbf{V}_{i,M_{i,p},K_i^*}, \mathbf{S}_i\right\rangle \le \alpha_i Q_i + \epsilon_i'.$$

3) For $i \in \{1,2\}$, let $\mathcal{E}_{0i}(\epsilon_0', \epsilon_i', \epsilon')$ denote the event that

$$\begin{aligned}-\epsilon' \le \frac{1}{n}\left\langle (\mathbf{V}_{0,M_0,K_0^*} - \alpha_0\mathbf{S}_0), (\mathbf{V}_{i,M_{i,p},K_i^*} - \alpha_i\mathbf{S}_i)\right\rangle \\ \le \epsilon_0'\epsilon_i'.\end{aligned}$$

4) Let $\mathcal{E}_{12}(\epsilon_1', \epsilon_2', \epsilon')$ denote the event that

$$\begin{aligned}-\epsilon' \le \frac{1}{n}\left\langle (\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1), (\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2)\right\rangle \\ \le \epsilon_1'\epsilon_2'.\end{aligned}$$

5) Let $\mathcal{E}_{00}(\epsilon_0')$ denote the event that

$$P_0 - 2\epsilon_0' \le \frac{1}{n}\|\mathbf{V}_{0,M_0,K_0^*} - \alpha_0\mathbf{S}_0\|^2 \le P_0.$$

6) For $i \in \{1,2\}$, let $\mathcal{E}_{ii}(\epsilon_i')$ denote the event that

$$P_{i,p} - 2\epsilon_i' \le \frac{1}{n}\|\mathbf{V}_{i,M_{i,p},K_i^*} - \alpha_i\mathbf{S}_i\|^2 \le P_{i,p}.$$

7) Let $\mathcal{E}_{Z0}(\epsilon')$ denote the event that

$$\frac{1}{n}\left|\left\langle \mathbf{V}_{0,M_0,K_0^*}, \mathbf{Z}\right\rangle\right| \le \epsilon'.$$

8) For $i \in \{1,2\}$, let $\mathcal{E}_{Zi}(\epsilon')$ denote the event that

$$\frac{1}{n}\left|\left\langle \mathbf{V}_{i,M_{i,p},K_i^*}, \mathbf{Z}\right\rangle\right| \le \epsilon'.$$

9) Let $\mathcal{E}_{ZS}(\epsilon')$ denote the event that

$$\frac{1}{n}|\langle \mathbf{S}, \mathbf{Z}\rangle| \leq \epsilon'.$$

10) Let $\mathcal{E}_{Z}(\epsilon')$ denote the event that

$$N - \epsilon' \leq \frac{1}{n}\|\mathbf{Z}\|^2 \leq N + \epsilon'.$$

Lemma E.2: Given parameters $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, and $\delta_0, \delta_1, \delta_2 > 0$, there exist positive $\epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$, $\epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)$, $\epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2)$ with the following properties:

i) The probability that all 14 events in Definition E.1 simultaneously occur tends to 1 as the blocklength n tends to infinity for all $\epsilon' > 0$ and all $\epsilon'_0, \epsilon'_1, \epsilon'_2$ satisfying

$$\epsilon'_0 > \epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) \tag{165}$$
$$\epsilon'_1 > \epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1) \tag{166}$$
$$\epsilon'_2 > \epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2). \tag{167}$$

ii) For fixed $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$:
   1) $\epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) \downarrow 0$ as $\delta_0 \downarrow 0$
   2) $\epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1) \downarrow 0$ as $\delta_1 \downarrow 0$
   3) $\epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2) \downarrow 0$ as $\delta_2 \downarrow 0$.

Proof: Define $\epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$, $\epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)$, and $\epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2)$ as in (168) at the bottom of the page. By (40) and because $\delta_0 > 0$

$$\lim_{\delta_0 \downarrow 0} \epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) = 0.$$

Moreover, by definition, for all $\epsilon'_0 > \epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$

$$\sqrt{\frac{\alpha_0^2 Q_0}{P_0 + \alpha_0^2 Q_0}} + \frac{\epsilon'_0}{\sqrt{P_0 + \alpha_0^2 Q_0}\sqrt{Q_0}} > \sqrt{1 - 2^{-2R'_0}}. \tag{169}$$

Similarly, by (41) and (42) and because $\delta_1, \delta_2 > 0$, for $i \in \{1,2\}$

$$\lim_{\delta_i \downarrow 0} \epsilon_i(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_i) = 0$$

and for all $\epsilon'_i > \epsilon_i(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_i)$

$$\sqrt{\frac{\alpha_i^2 Q_1}{P_{i,p} + \alpha_i^2 Q_i}} + \frac{\epsilon'_i}{\sqrt{P_{i,p} + \alpha_i^2 Q_i}\sqrt{Q_i}} > \sqrt{1 - 2^{-2R'_i}}. \tag{170}$$

In the following, we prove that since $\epsilon'_0, \epsilon'_1, \epsilon'_2 > 0$ satisfy Inequalities (169) and (170), the probability of all 14 events in Definition E.1 tends to 1 as n tends to infinity. This then establishes the lemma.

By the symmetry of the code construction and of the encodings, we can assume without loss of generality that $M_0 = M_{1,p} = M_{2,p} = 1$. Moreover, it suffices to show that the probability of each individual event tends to 1 as n tends to infinity.

We first consider event $\mathcal{E}_0(\epsilon'_0)$ and notice that it is equivalent to event

$$\begin{aligned}\sqrt{\frac{\alpha_0^2 Q_0}{P_0 + \alpha_0^2 Q_0}} &\leq \frac{1}{n}\left\langle \frac{\mathbf{V}_{0,1,K_0^*}}{\sqrt{P_0 + \alpha_0^2 Q_0}}, \frac{\mathbf{S}_0}{\sqrt{Q_0}} \right\rangle \\ &\leq \sqrt{\frac{\alpha_0^2 Q_0}{P_0 + \alpha_0^2 Q_0}} + \frac{\epsilon'_0}{\sqrt{P_0 + \alpha_0^2 Q_0}\sqrt{Q_0}}.\end{aligned} \tag{171}$$

Since by (40) and by (169)

$$\begin{aligned}\sqrt{\frac{\alpha_0^2 Q_0}{P_0 + \alpha_0^2 Q_0}} &< \sqrt{1 - 2^{-2R'_0}} \\ &< \sqrt{\frac{\alpha_0^2 Q_0}{P_0 + \alpha_0^2 Q_0}} + \frac{\epsilon'_0}{\sqrt{P_0 + \alpha_0^2 Q_0}}\end{aligned} \tag{172}$$

the auxiliary Lemma B.5 (see Appendix B) immediately yields

$$\lim_{n \to \infty} \Pr\left[\mathcal{E}_0(\epsilon'_0)\right] = 1. \tag{173}$$

Similar limits for events $\mathcal{E}_1(\epsilon'_1)$ and $\mathcal{E}_2(\epsilon'_2)$ can be proved in the same way.

We next fix an arbitrary $\epsilon' > 0$ and consider the event $\mathcal{E}_{01}(\epsilon'_0, \epsilon'_1, \epsilon')$. This event is equivalent to event (174) at the bottom of the next page. Similar to (172), by (41), (42), and (170), we have

$$\sqrt{\frac{\alpha_1^2 Q_1}{P_{1,p} + \alpha_1^2 Q_1}} < \sqrt{1 - 2^{-2R'_1}}$$

$$\epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) \triangleq \left(\sqrt{\frac{\alpha_0^2 Q_0}{P_0 + \alpha_0^2 Q_0}} - \sqrt{1 - 2^{-2R'_0}}\right)\sqrt{P_0 + \alpha_0^2 Q_0}\sqrt{Q_0} \tag{168a}$$

$$\epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1) \triangleq \left(\sqrt{\frac{\alpha_1^2 Q_1}{P_{1,p} + \alpha_1^2 Q_1}} - \sqrt{1 - 2^{-2R'_1}}\right)\sqrt{P_{1,p} + \alpha_1^2 Q_1}\sqrt{Q_1} \tag{168b}$$

$$\epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2) \triangleq \left(\sqrt{\frac{\alpha_2^2 Q_2}{P_{2,p} + \alpha_2^2 Q_2}} - \sqrt{1 - 2^{-2R'_2}}\right)\sqrt{P_{2,p} + \alpha_2^2 Q_2}\sqrt{Q_2} \tag{168c}$$

$$< \sqrt{\frac{\alpha_1^2 Q_1}{P_{1,p}+\alpha_1^2 Q_1}} + \frac{\epsilon_1'}{\sqrt{P_{1,p}+\alpha_1^2 Q_1}}. \tag{175}$$

Combining the auxiliary Lemma B.6 (see Appendix B) with (172) and (175), we can conclude that

$$\lim_{n\to\infty} \Pr\left[\mathcal{E}_{01}(\epsilon_0',\epsilon_1',\epsilon')\right] = 1. \tag{176}$$

Similar limits for events $\mathcal{E}_{02}(\epsilon_0',\epsilon_2',\epsilon')$ and $\mathcal{E}_{12}(\epsilon_1',\epsilon_2',\epsilon')$ can be proved in the same way.

We now prove limit

$$\lim_{n\to\infty} \Pr\left[\mathcal{E}_{00}(\epsilon_0')\right] = 1. \tag{177}$$

Notice that event $\mathcal{E}_{00}(\epsilon_0')$ occurs whenever event $\mathcal{E}_0(\epsilon_0')$ occurs. In fact, whenever $\mathcal{E}_0(\epsilon_0')$ occurs, then

$$\begin{aligned}
&\frac{1}{n}\|\mathbf{V}_{0,1,K_0^*} - \alpha_0 \mathbf{S}_0\|^2 \\
&\quad = \frac{1}{n}\|\mathbf{V}_{0,1,K_0^*}\|^2 + \frac{1}{n}\alpha_0^2\|\mathbf{S}_0\|^2 - 2\alpha_0 \frac{1}{n}\left\langle \mathbf{V}_{0,1,K_0^*}, \mathbf{S}_0\right\rangle \\
&\quad \le P_0 + \alpha_0^2 Q_0 + \alpha_0^2 Q_0 - 2\alpha_0^2 Q_0 \\
&\quad = P_0
\end{aligned}$$

and similarly

$$\begin{aligned}
&\frac{1}{n}\|\mathbf{V}_{0,1,K_0^*} - \alpha_0 \mathbf{S}_0\|^2 \\
&\quad = \frac{1}{n}\|\mathbf{V}_{0,1,K_0^*}\|^2 + \frac{1}{n}\alpha_0^2\|\mathbf{S}_0\|^2 - 2\alpha_0 \frac{1}{n}\left\langle \mathbf{V}_{0,1,K_0^*}, \mathbf{S}_0\right\rangle \\
&\quad \ge P_0 - 2\alpha_0\epsilon_0' \\
&\quad \ge P_0 - 2\epsilon_0'.
\end{aligned}$$

Thus, limit (176) combined with the last two bounds establishes (177). Similar limits for events $\mathcal{E}_{11}(\epsilon_1')$ and $\mathcal{E}_{22}(\epsilon_2')$ can be proved in the same way.

Since $\mathbf{S}$, $\mathbf{V}_{0,1,K_0^*}$, $\mathbf{V}_{1,1,K_1^*}$, and $\mathbf{V}_{2,1,K_2^*}$ are uniformly distributed over centered n-spheres and independent of $\mathbf{Z}$, by [2, eq. (54) and Lemma 4.1], we can conclude that for any $\epsilon' > 0$ the probability of events $\mathcal{E}_{Z0}(\epsilon')$, $\mathcal{E}_{Z1}(\epsilon')$, $\mathcal{E}_{Z2}(\epsilon')$, and $\mathcal{E}_{ZS}(\epsilon')$ tends to 1 as n tends to infinity.

Finally, by the ergodicity of $\mathbf{Z}$ and by the weak law of large numbers, for every $\epsilon' > 0$

$$\lim_{n\to\infty} \Pr\left[\mathcal{E}_Z(\epsilon')\right] = 1$$

which concludes the proof. ■

Lemma E.3: Given parameters $\beta_1, \beta_2 \in [0,1]$, $\tilde{P}_1 \in (0, P_1)$, $\tilde{P}_2 \in (0, P_2)$, and $\delta_0 > 0$, if the parameters δ_1 and δ_2 satisfy

$$0 < \delta_1 < \delta_{1,0}^*(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0) \tag{178}$$

$$0 < \delta_2 < \delta_{2,0}^*(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0) \tag{179}$$

for some specific positive $\delta_{1,0}^*(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0)$ and $\delta_{2,0}^*(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0)$, then it is possible to find $\epsilon_1', \epsilon_2' > 0$ satisfying the following three properties:

1)

$$\max\{\epsilon_1' + \epsilon_2', \epsilon_1'\epsilon_2'\} < \frac{N_0}{2} \tag{180}$$

2) for every $\epsilon' > 0$

$$\begin{aligned}
\lim_{n\to\infty} \Pr\Big[&\mathcal{E}_1(\epsilon_1'), \mathcal{E}_2(\epsilon_2'), \mathcal{E}_{12}(\epsilon_1',\epsilon_2',\epsilon'), \mathcal{E}_{11}(\epsilon_1') \\
&\mathcal{E}_{22}(\epsilon_2'), \mathcal{E}_{Z1}(\epsilon'), \mathcal{E}_{Z2}(\epsilon'), \mathcal{E}_{ZS}(\epsilon'), \mathcal{E}_Z(\epsilon')\Big] = 1
\end{aligned} \tag{181}$$

where the events $\mathcal{E}_1(\epsilon_1'), \mathcal{E}_2(\epsilon_2'), \mathcal{E}_{12}(\epsilon_1',\epsilon_2',\epsilon')$, $\mathcal{E}_{11}(\epsilon_1')$, $\mathcal{E}_{22}(\epsilon_2')$, $\mathcal{E}_{Z1}(\epsilon'), \mathcal{E}_{Z2}(\epsilon'), \mathcal{E}_{ZS}(\epsilon')$, and $\mathcal{E}_Z(\epsilon')$ are defined in Definition E.1;

3)

$$\begin{aligned}
&\frac{1}{\sqrt{P_0+\alpha_0^2 Q_0}} \\
&\cdot \frac{P_0+\alpha_0 Q_0}{\sqrt{P_0+N_0+Q_0+(3-2\alpha_0)\max\{2(\epsilon_1'+\epsilon_2'), 2\epsilon_1'\epsilon_2'\}}} \\
&\qquad > \sqrt{1-2^{-2(R_0+R_0')}}.
\end{aligned} \tag{182}$$

Proof: We first notice that (180) holds whenever

$$0 < \epsilon_1', \epsilon_2' < \min\left\{\sqrt{\frac{N_0}{2}}, \frac{N_0}{4}\right\}. \tag{183}$$

By (46) and simple algebraic manipulations

$$\frac{P_0+\alpha_0 Q_0}{\sqrt{P_0+\alpha_0^2 Q_0}\sqrt{P_0+N_0+Q_0}} = \sqrt{1-2^{-2(R_0+R_0'+\delta_0)}}. \tag{184}$$

Therefore, since the right-hand side of (184) is increasing in δ_0, since $\delta_0 > 0$, and by continuity, there exists a positive

$$\begin{aligned}
\frac{\epsilon_1'\epsilon_1'}{\sqrt{P_0+\alpha_0^2 Q_0}\sqrt{P_{1,p}+\alpha_1^2 Q_1}} &\ge \frac{1}{n}\left\langle \left(\frac{\mathbf{V}_{1,1,K_0^*}}{\sqrt{P_0+\alpha_0^2 Q_0}} - \sqrt{\frac{\alpha_0^2 Q_0}{P_0+\alpha_0^2 Q_0}}\frac{\mathbf{S}_0}{\sqrt{Q_0}}\right), \left(\frac{\mathbf{V}_{1,1,K_1^*}}{\sqrt{P_{1,p}+\alpha_1^2 Q_1}} - \sqrt{\frac{\alpha_1^2 Q_1}{P_{1,p}+\alpha_1^2 Q_1}}\frac{\mathbf{S}_0}{\sqrt{Q_0}}\right)\right\rangle \\
&\ge -\frac{\epsilon'}{\sqrt{P_0+\alpha_0^2 Q_0}\sqrt{P_{1,p}+\alpha_1^2 Q_1}}
\end{aligned} \tag{174}$$

$\epsilon(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0) > 0$ such that inequality (182) holds for all ϵ_1', ϵ_2' that satisfy

$$0 < \epsilon_1', \epsilon_2' \le \epsilon(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0). \tag{185}$$

Finally, by Lemma E.2, there exist positive $\epsilon_1(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_1)$ and $\epsilon_2(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_2)$ that tend to 0 as δ_1 and δ_2 tend to 0, respectively, and such that Limit (181) holds for all $\epsilon_1', \epsilon_2' > 0$ that satisfy

$$\epsilon_1' > \epsilon_1(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_1) \tag{186}$$
$$\epsilon_2' > \epsilon_2(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_2). \tag{187}$$

By (183)–(187), we conclude that if

$$\max\left\{\epsilon_1(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_1), \epsilon_2(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_2)\right\} < \min\left\{\epsilon(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0), \frac{N_0}{4}, \sqrt{\frac{N_0}{2}}\right\} \tag{188}$$

then it is possible to choose $\epsilon_1', \epsilon_2' > 0$ satisfying (180)–(182). Since $\epsilon(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0) > 0$ and $N_0 > 0$, and since for fixed $\beta_1,\beta_2 \in [0,1]$, $\tilde{P}_1 \in (0,P_1)$, $\tilde{P}_2 \in (0,P_2)$, and $\delta_0 > 0$, we can make $\epsilon_1(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_1)$ and $\epsilon_2(\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_2)$ arbitrary small by choosing $\delta_1 > 0$ and $\delta_2 > 0$ sufficiently small, the lemma follows. ■

B) Proof of Lemma VI.3: By the definition of $\mathbf{Z}_0$

$$\begin{aligned}
&\frac{1}{n}\|\mathbf{Z}_0\|^2 \\
&= \frac{1}{n}\|\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1\|^2 + \frac{1}{n}\|\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2\|^2 \\
&\quad + \frac{2}{n}\left\langle\left(\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1\right), \left(\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2\right)\right\rangle \\
&\quad + \frac{1}{n}\|\mathbf{Z}\|^2 + \frac{2}{n}\left\langle\mathbf{V}_{1,M_{1,p},K_1^*}, \mathbf{Z}\right\rangle + \frac{2}{n}\left\langle\mathbf{V}_{2,M_{2,p},K_2^*}, \mathbf{Z}\right\rangle \\
&\quad - (\alpha_1(1-\alpha_0) + \alpha_2(1-\alpha_1)(1-\alpha_0))\frac{2}{n}\langle\mathbf{S}, \mathbf{Z}\rangle
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{n}\langle\mathbf{Z}_0, \mathbf{S}_0\rangle &= \frac{1}{n}\left\langle\left(\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1\right), \mathbf{S}_0\right\rangle \\
&\quad + \frac{1}{n}\left\langle\left(\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2\right), \mathbf{S}_0\right\rangle \\
&\quad + \frac{1}{n}\langle\mathbf{Z}, \mathbf{S}_0\rangle.
\end{aligned}$$

Thus, whenever events

$$\begin{aligned}
&\mathcal{E}_1(\epsilon_1'), \mathcal{E}_2(\epsilon_2'), \mathcal{E}_{12}(\epsilon_1', \epsilon_2', \epsilon'), \mathcal{E}_{11}(\epsilon_1') \\
&\mathcal{E}_{22}(\epsilon_2'), \mathcal{E}_{Z1}(\epsilon'), \mathcal{E}_{Z2}(\epsilon'), \mathcal{E}_{ZS}(\epsilon')\mathcal{E}_Z(\epsilon')
\end{aligned} \tag{189}$$

occur for some $\epsilon_1', \epsilon_2', \epsilon' > 0$, then

$$-9\epsilon' - 2(\epsilon_1' + \epsilon_2') \le \frac{1}{n}\|\mathbf{Z}_0\|^2 - N_0 \le 7\epsilon' + 2\epsilon_1'\epsilon_2' \tag{190}$$

and

$$-\epsilon' \le \frac{1}{n}\langle\mathbf{Z}_0, \mathbf{S}_0\rangle \le \epsilon_1' + \epsilon_2' + \epsilon'. \tag{191}$$

Let in the following the parameters $\delta_1, \delta_2 > 0$ be sufficiently small (as a function of $\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2,\delta_0$) to satisfy (178) and (179) in Lemma E.3. Then, by Lemma E.3, we can choose $\epsilon_1', \epsilon_2' > 0$ so as to satisfy conditions (180)–(182) in Lemma E.3. By (180) and (182), we can further choose a small $\epsilon > 0$ such that

$$\epsilon_0^* \stackrel{\triangle}{=} \max\{2(\epsilon_1' + \epsilon_2'), 2\epsilon_1'\epsilon_2'\} + \epsilon \tag{192}$$

satisfies $\epsilon_0^* < N_0$ and condition (56) in Lemma VI.3. Moreover, since (190) and (191) hold whenever the events in (189) occur, and since by our choice of $\epsilon_1', \epsilon_2' > 0$ for all $\epsilon' > 0$ the probability of these events tends to 1 a $n \to \infty$ [limit (181)], our choice of ϵ_0^* in (181) also satisfies Conditions (57) and (58) in Lemma VI.3. This concludes the proof.

C) Proof of Lemma VI.6: Notice that whenever events $\mathcal{E}_0(\epsilon_0')$, $\mathcal{E}_1(\epsilon_1')$, $\mathcal{E}_2(\epsilon_2')$, $\mathcal{E}_{01}(\epsilon_0', \epsilon_1', \epsilon')$ and $\mathcal{E}_{02}(\epsilon_0', \epsilon_2', \epsilon')$ occur for some $\epsilon_0', \epsilon_1', \epsilon_2', \epsilon' > 0$, then

$$\begin{aligned}
&\frac{1}{n}\|\mathbf{X}_1\|^2 \\
&= \frac{1}{n}\|\mathbf{X}_0'\lambda_1 + \mathbf{X}_1'\|^2 \\
&= \left(\frac{1}{n}\|\mathbf{V}_{0,M_0,K_0^*}\|^2 - \frac{2\alpha_0}{n}\left\langle\mathbf{V}_{0,M_0,K_0^*}, \mathbf{S}_0\right\rangle + \frac{\alpha_0^2}{n}\|\mathbf{S}_0\|^2\right)\lambda_1^2 \\
&\quad + \frac{1}{n}\|\mathbf{V}_{1,M_{1,p},K_1^*}\|^2 - \frac{2\alpha_1}{n}\left\langle\mathbf{V}_{1,M_{1,p},K_1^*}, \mathbf{S}_1\right\rangle + \frac{1}{n}\|\mathbf{S}_1\|^2 \\
&\quad + \frac{2}{n}\lambda_1\left\langle\left(\mathbf{V}_{0,M_0,K_0^*} - \alpha_0\mathbf{S}_0\right), \left(\mathbf{V}_{1,M_{1,p},K_1^*} - \alpha_1\mathbf{S}_1\right)\right\rangle \\
&\le P_0\frac{\tilde{P}_1 - P_{1,p}}{P_0} + P_{1,p} + 2\sqrt{\frac{\tilde{P}_1 - P_{1,p}}{P_0}}\epsilon_0'\epsilon_1' \\
&= \tilde{P}_1 + 2\sqrt{\frac{\tilde{P}_1 - P_{1,p}}{P_0}}\epsilon_0'\epsilon_1'
\end{aligned} \tag{193}$$

and

$$\begin{aligned}
&\frac{1}{n}\|\mathbf{X}_2\|^2 \\
&= \frac{1}{n}\|\mathbf{X}_0'\lambda_2 + \mathbf{X}_2'\|^2 \\
&= \frac{1}{n}\|\mathbf{V}_{0,M_0,K_0^*} - \alpha_0\mathbf{S}_0\|^2\lambda_2^2 + \frac{1}{n}\|\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2\|^2 \\
&\quad + \frac{2}{n}\lambda_2\left\langle\left(\mathbf{V}_{0,M_0,K_0^*} - \alpha_0\mathbf{S}_0\right), \left(\mathbf{V}_{2,M_{2,p},K_2^*} - \alpha_2\mathbf{S}_2\right)\right\rangle \\
&\le P_0\frac{\tilde{P}_2 - P_{2,p}}{P_0} + P_{2,p} + 2\sqrt{\frac{\tilde{P}_2 - P_{2,p}}{P_0}}\epsilon_0'\epsilon_2' \\
&= \tilde{P}_2 + 2\sqrt{\frac{\tilde{P}_2 - P_{2,p}}{P_0}}\epsilon_0'\epsilon_2'.
\end{aligned} \tag{194}$$

Thus, if we choose $\epsilon_0', \epsilon_1', \epsilon_2' > 0$ so that

$$\epsilon_0'\epsilon_1' \le (P_1 - \tilde{P}_1)\frac{\sqrt{P_0}}{2\sqrt{\tilde{P}_1 - P_{1,p}}} \tag{195}$$

$$\epsilon_0'\epsilon_2' \le (P_2 - \tilde{P}_2)\frac{\sqrt{P_0}}{2\sqrt{\tilde{P}_2 - P_{2,p}}} \tag{196}$$

then the channel input sequences $\mathbf{X}_1$ and $\mathbf{X}_2$ satisfy the block-power constraints (8) and (9).

In the following, we argue that if $\delta_0, \delta_1, \delta_2 > 0$ are sufficiently small (depending on $\beta_1,\beta_2,\tilde{P}_1,\tilde{P}_2$), then there exists a choice $\epsilon_0', \epsilon_1', \epsilon_2' > 0$ satisfying (195) and (196) and such that

for all $\epsilon' > 0$ the probabilities of events $\mathcal{E}_0(\epsilon'_0)$, $\mathcal{E}_1(\epsilon'_1)$, $\mathcal{E}_2(\epsilon'_2)$, $\mathcal{E}_{01}(\epsilon'_0, \epsilon'_1, \epsilon')$, and $\mathcal{E}_{02}(\epsilon'_0, \epsilon'_2, \epsilon')$ tend to 1 as n tends to ∞. This establishes the lemma.

Recall that by Lemma E.2, there exist $\epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$, $\epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)$, and $\epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2)$ such that for every $\epsilon' > 0$ and all $\epsilon'_0, \epsilon'_1, \epsilon'_2 > 0$ satisfying

$$\epsilon'_0 > \epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0) \tag{197}$$

$$\epsilon'_1 > \epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1) \tag{198}$$

$$\epsilon'_2 > \epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2) \tag{199}$$

the probability of events $\mathcal{E}_0(\epsilon'_0)$, $\mathcal{E}_1(\epsilon'_1)$, $\mathcal{E}_2(\epsilon'_2)$, $\mathcal{E}_{01}(\epsilon'_0, \epsilon'_1, \epsilon')$, and $\mathcal{E}_{02}(\epsilon'_0, \epsilon'_2, \epsilon')$ tends to 1 as n tends to infinity. Moreover, for fixed $\beta_1, \beta_2 \in (0, 1)$, $\tilde{P}_1 \in (0, P_1)$, and $\tilde{P}_2 \in (0, P_2)$ the bounds $\epsilon_0(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_0)$, $\epsilon_1(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_1)$, $\epsilon_2(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2, \delta_2)$ tend to 0 as $\delta_0, \delta_1, \delta_2 \downarrow 0$. Thus, there exist positive $\delta^*_{0,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$, $\delta^*_{1,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$, $\delta^*_{2,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2)$ such that whenever

$$0 < \delta_0 < \delta^*_{0,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{200}$$

$$0 < \delta_1 < \delta^*_{1,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{201}$$

$$0 < \delta_2 < \delta^*_{2,P}(\beta_1, \beta_2, \tilde{P}_1, \tilde{P}_2) \tag{202}$$

then it is possible to choose $\epsilon'_0, \epsilon'_1, \epsilon'_2 > 0$ that simultaneously satisfy (195)–(199). This concludes the proof.

ACKNOWLEDGMENT

The authors thank Stephane Tinguely and Tsachy Weissman for helpful discussions.

REFERENCES

[1] S. I. Bross, A. Lapidoth, and M. A. Wigger, "The Gaussian MAC with conferencing encoders," in *Proc. Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 2702–2706.

[2] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1639–1667, Jun. 2002.

[3] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[4] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

[5] T. M. Cover, "Some advances in broadcast channels," in *Advances in Communication Systems*, A. J. Viterbi, Ed. San Francisco, CA: Academic, 1975, vol. 4, pp. 229–260.

[6] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems Control Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[7] S. I. Gel'fand and M. S. Pinsker, "On Gaussian channels with random parameters," in *Proc. Int. Symp. Inf. Theory*, Taskent, U.S.S.R., Sep. 1984, pp. 247–250.

[8] R. Khosravi-Farsani and F. Marvasti, "Multiple access channels with cooperative encoders and channel state information," submitted to *Eur. Trans. Telecom,* Sep. 2010 [Online]. Available: http://arxiv.org/abs/1009.6008

[9] S. P. Kotagiri and J. N. Laneman, "Multiaccess channels with state known to one encoder: A case of degraded message sets," in *Proc. Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1566–1570.

[10] S. P. Kotagiri and J. N. Laneman, "Multiaccess channels with state known to some encoder and independent messages," *EURASIP J. Wireless Commun. Netw.*, Mar. 2008.

[11] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520–1529, Sep. 1996.

[12] N. Merhav and S. Shamai (Shitz), "Information rates subject to state masking," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2254–2261, Jun. 2007.

[13] H. H. Permuter, S. Shamai (Shitz), and A. Somekh-Baruch, "Message and state cooperation in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6379–6396, Oct. 2011.

[14] T. Philosof, A. Khisti, U. Erez, and R. Zamir, "Lattice strategies for the dirty multiple access channel," in *Proc. Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 386–390.

[15] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5006–5035, Aug. 2011.

[16] T. Philosof and R. Zamir, "On the loss of single-letter characterization: The dirty multiple access channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2442–2454, Jun. 2009.

[17] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 364–375, Mar. 1996.

[18] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 4, pp. 1037–1076, Sep. 1973.

[19] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdu, "Cooperative multiple-access encoding with states available at one transmitter," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4448–4460, Oct. 2008.

[20] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdu, "Cooperative multiple-access encoding with states available at one transmitter," in *Proc. Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1556–1560.

[21] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdu, "Cooperative encoding with asymmetric state information at the transmitters," in *Proc. 44th Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 2006, pp. 392–401.

[22] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[23] I.-H. Wang, "Distributed interference cancellation in multiple access channels," 2010 [Online]. Available: http://arxiv.org/abs/1011.3588

[24] F. M. J. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 441–445, May 1983.

[25] J. M. Wozencraft and I. M. Jacobs*, Principles of Communication Engineering*. New York: Wiley, 1965.

[26] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 2–10, Jan. 1974.

[27] A. Zaidi, S. P. Kotagiri, J. N. Laneman, and L. Vandendorpe, "Multiaccess channels with state known to one encoder: Another case of degraded message sets," in *Proc. Int. Symp. Inf. Theory*, Seoul, Korea, Jul. 2009, pp. 2376–2380.

Shraga I. Bross (S'89–M'92–SM'09) received the B.Sc. and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, in 1978 and 1983, and the Ph.D. degree from the University of Maryland, College Park in 1991, all in electrical engineering. During the 1991–1992 academic year, he was a postdoctoral Fellow in the ECE Department at the University of Waterloo, Canada. During 1992–1998, he was with Orckit Communications Ltd., Tel-Aviv, Israel, in the capacity of a Senior Scientist. From 1998 to 2006, he was a Senior Research Fellow at the EE Department, Technion. Since 2007, he has been a Senior Lecturer at the School of Engineering, Bar-Ilan University, Israel. His research interests are in digital communications and information theory.

Amos Lapidoth (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (*summa cum laude*, 1986), the B.Sc. degree in electrical engineering (*summa cum laude*, 1986), and the M.Sc. degree in electrical engineering (1990) all from the Technion—Israel Institute of Technology. He received the Ph.D. degree in electrical engineering from Stanford University in 1995.

In the years 1995–1999 he was an Assistant and Associate Professor at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at ETH Zurich in Switzerland. He is the author of the book *A Foundation in Digital Communication*, published by Cambridge University Press in 2009. His research interests are in digital communications and information theory.

Dr. Lapidoth served in the years 2003–2004 and 2009 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Michèle Wigger (S'05–M'09) received the M.Sc. degree in electrical engineering (with distinction) and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009, she was a Postdoctoral Researcher at the ITA Center, University of California, San Diego. Since December 2009, she has been an Assistant Professor at Telecom ParisTech, Paris, France. Her research interests are in information and communications theory; in particular in wireless networks, feedback channels, and channels with states.