The Capacity Region of a Product of Two Unmatched Physically Degraded Gaussian Broadcast Channels With Three Individual Messages and a Common Message

Ramy H. Gohary, *Member, IEEE*, and Timothy N. Davidson, *Member, IEEE*

***Abstract*—This paper considers a Gaussian broadcast channel with two unmatched degraded components, three individual messages, and a common message that is intended for all three receivers. It is shown that for this channel, superposition coding with Gaussian signalling is sufficient to achieve every point in the capacity region.**

***Index Terms*—Broadcast channels with physically degraded components, entropy power inequality, geometric programming, Karush–Kuhn–Tucker (KKT) conditions, relaxation, superposition coding (SPC).**

I. Introduction

In a broadcast channel (BC), a single transmitter sends messages to multiple receivers [1]. These messages may be common to all receivers or particular to an individual receiver or a subset of receivers. The vector containing the rates of these messages is said to be achievable if each receiver is able to reliably decode its intended messages. The closure of all such vectors is usually referred to as the capacity region [2].

A special class of BCs is the one in which the received signals form a Markov chain. In this case, the received signals are said to be physically degraded versions of each other, and the degradation level of each signal is given by its order in the Markov chain. For the class of physically degraded channels, superposition coding (SPC) [3] is known to attain every point on the boundary of the capacity region in the general unrestricted case [4], and in the case of Gaussian channels with a power constraint [5].

Although degraded channels are useful in modelling single-input single-output BC systems, many practical systems give rise to nondegraded channels, including those that employ multicarrier transmission [6], and the class of multiple-input multiple-output (MIMO) systems [7], [8]. In those channels, the received signals do not form a Markov chain, and the coding scheme developed in [3] cannot be applied directly to achieve every point in their capacity regions [7].

Manuscript received August 21, 2009; revised July 22, 2011; accepted May 31, 2012. Date of publication September 17, 2012; date of current version December 19, 2012. The material in this paper was presented in part at the 2010 IEEE International Symposium on Information Theory.

R. H. Gohary was with the Communications Research Centre, Ottawa, ON K1P 6M1, Canada. He is now with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: gohary@sce.carleton.ca).

T. N. Davidson is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: davidson@mcmaster.ca).

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

Digital Object Identifier 10.1109/TIT.2012.2219172

Most of the studies on nondegraded BCs have focused on the case in which only independent individual messages are sent to the receivers (see, e.g., [7]–[16]). For example, the sum capacity for the case in which individual messages are broadcast over Gaussian MIMO channels was studied in [15] and [16] and was shown in [8], [12], and [13] to be achievable by dirty paper coding (DPC) [17] with Gaussian signalling. Later, it was shown in [7] that DPC with Gaussian signalling is sufficient to attain every point in the achievable rate region. That is, DPC with Gaussian signalling is sufficient for achieving every point in the capacity region of the Gaussian MIMO BC with individual messages.

In contrast to the case of individual messages only, there has been less progress in characterizing the capacity region of general nondegraded BCs when common or partially common messages are to be transmitted along with individual messages. However, some partial results are available. For instance, for the case in which common messages may be transmitted over general nondegraded BCs, characterizations of achievable inner bounds were obtained in [18]–[20] and [21]. An inner bound that includes the bounds in [20] and [21] was developed in [22]. In [23], the bounds in [21] and [22] were carefully analyzed and their equivalence was established. In a complementary fashion, characterizations of outer bounds were obtained in [24] and [25]. Another outer bound and a review of previously known ones are presented in [26]. Further results on inner and outer bounds of the capacity region of the general BC are reported in [27]. In addition to inner and outer bounds, in the presence of two receivers, a common message intended for both receivers and an individual message intended for each receiver, characterizations of the capacity region were provided in [28] for BCs with two unmatched parallel physically degraded components and in [29] for deterministic BCs.

For a BC with three receivers, a common message and one individual message, a single-letter characterization of the capacity region was provided in [30] and this region was shown to be strictly larger than the one conjectured in [31]. For general BCs in which common, partially common and individual messages are intended for the receivers, fundamental constraints on the geometry of the capacity region were provided in [32].

In this paper, we consider a different class of BCs with three receivers. In contrast to [30], in which there is only one individual message, in the class considered herein, an individual message is sent to each of the three receivers, in addition to the common message. The channel is assumed to be Gaussian and memoryless with two unmatched physically degraded components. It will be shown that for the degradation orders considered in this paper, SPC with Gaussian signalling is sufficient to attain any point on the boundary of the capacity region. Note that, in the scenario that we consider, the individual messages are not constrained to be confidential; cf. [33] and [34].

The methodology for obtaining this result involves four stages. First, we characterize a set of rate vectors that can be achieved by SPC with Gaussian signalling. Second, we extract insight from the structure of the signalling scheme that enables this set of rate vectors to be achieved and use this insight to provide an ostensibly relaxed characterization of that set. Using the Karush–Kuhn–Tucker (KKT) optimality conditions, this relaxation is shown to be tight. Third, we use information-theoretic analysis to obtain bounds on any achievable rate vector. Finally, by combining the tight relaxation and the information-theoretic bounds, we establish the desired converse, i.e., that every achievable rate vector can be attained by SPC with Gaussian signalling.

This paper is organized as follows. In Section II, the system model considered in this paper is described along with basic definitions. In Section III, a set of rates that can be achieved by SPC with Gaussian signalling is characterized. In Section IV, each rate vector in that set is associated with the vectors of power partitions that enable it to be achieved. The resulting set of partition-rate vectors is then expressed as the intersection of two regions. Points on the boundaries of these regions are then expressed as solution of optimization problems, and relaxations of those problems are considered. A key result in our development concerns the tightness of the relaxations. This result is established in Sections V-C and V-D by examining the relationships between the solutions of the KKT systems (see Sections V-A and V-B). In Section V-E, the relationships between the relaxed optimization problems and the original region described in Section IV are established. In Section VI, information-theoretic bounds on all achievable rates are obtained and in Section VII, the entropy power inequality is applied to those bounds. The resulting inequalities along with the results pertaining to the tightness of the relaxations of the regions described in Section IV are then used in Section VII to establish the main result of this paper, i.e., the optimality of SPC with Gaussian signalling. Section VIII concludes this paper. For clarity of exposition, most of the proofs are placed in the appendix.

Notation

This paper uses conventional notation throughout. Vectors and matrices of deterministic variables are denoted by boldface symbols and their scalar entries are denoted by regular weight symbols. Subscripts of these symbols are used to refer to receivers or subchannels and superscripts are used to refer to degradation levels.

II. System Model

We consider a discrete-time BC system in which a transmitter sends messages to three receivers over two parallel unmatched Gaussian memoryless physically degraded subchannels. Using a common model [2], [10] for physically degraded channels, the considered BC system can be represented by Fig. 1. The transmitter sends an individual message M_k of rate R_k bits per channel use to receiver k, $k = 1,2,3$, and sends a common message of rate R_0 to all three receivers. The messages $M_k \in \{1,2,\ldots,2^{nR_k}\}$, $k = 0,1,2,3$, are transmitted over n channel uses of this system. Hence, in Fig. 1, X_i is used to denote the length-n vector of the transmitted signal on the ith subchannel. Its power is denoted by P_i, $i = 1,2$.

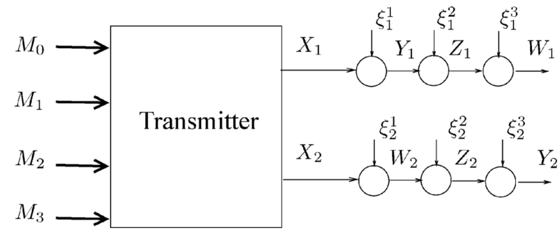

Fig. 1. BC with two unmatched Gaussian memoryless physically degraded components and three receivers. The signal observed by receivers 1, 2, and 3 on the ith subchannel is represented by Y_i, Z_i, and W_i, respectively.

The receivers in the considered system observe X_1 and X_2 in zero-mean additive white Gaussian noise. The variance of the noise at degradation level j on subchannel i is denoted by N_i^j, where $N_i^j < N_i^{j+1}$, $i,j = 1,2$. In Fig. 1, the length-n signal observed by receivers 1, 2, and 3 on the ith subchannel, is represented by Y_i, Z_i, and W_i, respectively, where $Y_1 = X_1 + \xi_1^1, Z_1 = Y_1 + \xi_1^2, W_1 = Z_1 + \xi_1^3, W_2 = X_2 + \xi_2^1, Z_2 = W_2 + \xi_2^2$, and $Y_2 = Z_2 + \xi_2^3$, and ξ_i^j represents the additional noise at degradation level j on subchannel i, which is of variance $N_i^j - N_i^{j-1}$, for $i = 1,2$, $j = 2,3$; cf. [2, p. 428].

For decoding, receiver k, $k = 1,2,3$, maps a length-n block of the signal received on each subchannel to the set of its intended messages. In particular, using d_k to denote the decoder of receiver k, for receiver 1, $d_1 : (Y_1, Y_2) \mapsto (\hat{M}_0, \hat{M}_1)$, where $\hat{M}_\ell \in \{1,\ldots,2^{nR_\ell}\}$, $\ell = 0,1$. An error event for receiver 1 occurs if $(\hat{M}_0, \hat{M}_1) \neq (M_0, M_1)$ and its average probability is denoted by $P_{e_1}^n$. The error events for receivers 2 and 3 are defined similarly and their average error probabilities are denoted by $P_{e_2}^n$ and $P_{e_3}^n$, respectively. A rate vector $\boldsymbol{R} \triangleq (R_0, R_1, R_2, R_3)$ is said to be achievable if for every $\epsilon > 0$, there exists a sequence of codes (indexed by n) such that for all sufficiently large n the probability of error $P_e^n < \epsilon$, where $P_e^n = \max\{P_{e_1}^n, P_{e_2}^n, P_{e_3}^n\}$.

The main result of this paper is that for the system in Fig. 1 with given positive[1] subchannel transmission powers P_1 and P_2, the set of rate vectors that are achievable using SPC with Gaussian signalling is the set of all achievable rates; i.e., the capacity region. Since the capacity region for the system in Fig. 1 with a total power constraint $P_1 + P_2 \leq \bar{P}$ is simply the convex hull of the union of the capacity regions for each power allocation that satisfies the constraint, our result extends directly to that case.

III. A Set of Rate Vectors Achieved by SPC With Gaussian Signalling

In this section, we will characterize a set of rate vectors that are achievable using SPC with Gaussian signalling over the BC in Fig. 1. The bounds on the partial sums of the rates achieved by these schemes are parametrized by the transmission power on

[1]Note that the case in which either P_1 or P_2 is zero corresponds to a physically degraded channel case for which the capacity region has already been fully characterized [2].

each subchannel, P_1 and P_2, and by a set of power partitions, which specify the fraction of power used to transmit each of the superimposed components of the signal transmitted on each subchannel (see, e.g., [28]). The power partition corresponding to degradation level j on subchannel i will be denoted by θ_i^j, $j = 1,2,3$, $i = 1,2$. Since power partitions on each subchannel lie in the unit simplex in $\mathbb{R}^3$ only two of them on each subchannel have to be specified. These partitions are collected in the vector $\boldsymbol{\theta} \triangleq [\theta_1^1, \theta_1^2, \theta_2^1, \theta_2^2] \in \mathcal{S}$, where

$$\mathcal{S} \triangleq \{\boldsymbol{\theta} | \theta_i^\ell \geq 0,\ \theta_i^1 + \theta_i^2 \leq 1,\ i, \ell = 1,2\}. \tag{1}$$

The following functions of $\boldsymbol{\theta}$ will simplify our descriptions:

$$f_0(\boldsymbol{\theta}) \triangleq \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big) \tag{2a}$$

$$f_{01}(\boldsymbol{\theta}) \triangleq \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big) \tag{2b}$$

$$f_{012}(\boldsymbol{\theta}) \triangleq f_{01}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) \tag{2c}$$

$$f_{0123}(\boldsymbol{\theta}) \triangleq f_{012}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big) \tag{2d}$$

$$g_0(\boldsymbol{\theta}) \triangleq \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}\Big) \tag{2e}$$

$$g_{02}(\boldsymbol{\theta}) \triangleq \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + \theta_2^1 P_2}\Big) \tag{2f}$$

$$g_{012}(\boldsymbol{\theta}) \triangleq g_{02}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big) \tag{2g}$$

$$g_{023}(\boldsymbol{\theta}) \triangleq g_{02}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big) \tag{2h}$$

$$g_{0123}(\boldsymbol{\theta}) \triangleq g_{012}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big) = g_{023}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big) \tag{2i}$$

$$h_0(\boldsymbol{\theta}) \triangleq \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1 + (\theta_2^1 + \theta_2^2)P_2}\Big) \tag{2j}$$

$$h_{03}(\boldsymbol{\theta}) \triangleq \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1}\Big) \tag{2k}$$

$$h_{023}(\boldsymbol{\theta}) \triangleq h_{03}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}{N_1^2 + \theta_1^1 P_1}\Big) \tag{2l}$$

$$h_{0123}(\boldsymbol{\theta}) \triangleq h_{023}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big). \tag{2m}$$

A. Initial Characterization

With the transmit powers P_1 and P_2 fixed, we consider a class of coding schemes based on superposition principles [1], [3]. In this class, 1) the common message is encoded jointly over the subchannels [28] and is the first message decoded at each receiver; and 2) the components of the individual messages that are transmitted on each subchannel are encoded separately. The position of each receiver in the degradation order on each subchannel enables the receiver to cancel the interference induced by signalling to more degraded receivers on that subchannel. The residual signal-to-noise ratio observed by the receivers after using SPC with a power partition vector $\boldsymbol{\theta} \in \mathcal{S}$ and interference cancellation places the following constraints on the rates that can be achieved by this scheme:

$$R_0 \leq \min\{f_0(\boldsymbol{\theta}), g_0(\boldsymbol{\theta}), h_0(\boldsymbol{\theta})\} \tag{3a}$$

$$R_0 + R_1 \leq f_{01}(\boldsymbol{\theta}) \tag{3b}$$

$$R_0 + R_2 \leq g_{02}(\boldsymbol{\theta}) \tag{3c}$$

$$R_0 + R_3 \leq h_{03}(\boldsymbol{\theta}) \tag{3d}$$

$$R_0 + R_1 + R_2 \leq \min\{f_{012}(\boldsymbol{\theta}), g_{012}(\boldsymbol{\theta})\} \tag{3e}$$

$$R_0 + R_2 + R_3 \leq \min\{g_{023}(\boldsymbol{\theta}), h_{023}(\boldsymbol{\theta})\} \tag{3f}$$

$$R_0 + R_1 + R_2 + R_3 \leq \min\{f_{0123}(\boldsymbol{\theta}), g_{0123}(\boldsymbol{\theta}), h_{0123}(\boldsymbol{\theta})\} \tag{3g}$$

with $R_i \geq 0$. In fact, the bounds in (3) are tight, in the sense that rate vectors that attain these bounds can be achieved using SPC with Gaussian signalling. Before specific signalling schemes that attain the bounds are described, we point out that in the derivation of the constraints in (3), we have used the fact that for the BC shown in Fig. 1, the constraints

$$R_0 + R_1 + R_3 \leq f_{01}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^1 + (\theta_2^2 + \theta_2^1)P_2}{N_2^1}\Big) \tag{4a}$$

$$R_0 + R_1 + R_3 \leq h_{03}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_1^1 + (\theta_1^2 + \theta_1^1)P_1}{N_1^1}\Big) \tag{4b}$$

are redundant. To show that (4a) is redundant, observe that

$$\begin{aligned} f_{01}(\boldsymbol{\theta}) &+ \frac{1}{2}\log\Big(\frac{N_2^1 + (\theta_2^2 + \theta_2^1)P_2}{N_2^1}\Big) \\ &= f_{01}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^1 + (\theta_2^1 + \theta_2^2)P_2}{N_2^1 + \theta_2^1 P_2}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big) \\ &> f_{01}(\boldsymbol{\theta}) + \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big) && (5) \\ &= f_{0123}(\boldsymbol{\theta}) && (6) \end{aligned}$$

where in (5) we have used the fact that $\log\Big(\frac{N_2^1 + (\theta_2^1 + \theta_2^2)P_2}{N_2^1 + \theta_2^1 P_2}\Big)$ is monotonically decreasing in N_2^1, and that $N_2^1 < N_2^2$. On the other hand, from (3g)

$$R_0 + R_1 + R_2 + R_3 \leq f_{0123}(\boldsymbol{\theta}). \tag{7}$$

Now, for any $R_2 > 0$, the left-hand side of (4a) is strictly less than the left-hand side of the inequality in (7) and the right-hand side of (4a) is strictly greater than the right-hand side of the inequality in (7). Hence, we conclude that the inequality in (7) is strictly tighter than the inequality in (4a), and hence the redundancy of (4a). A similar argument involving the constraint $\sum_{k=0}^{3} R_k \leq h_{0123}(\boldsymbol{\theta})$ can be used to show that (4b) is also redundant.

The bounds in (3) are tight in the sense that for an arbitrary vector of power partitions $\boldsymbol{\theta} \in \mathcal{S}$ a rate vector that attains the bounds in (3) can be achieved using one of the three SPC with Gaussian signalling modes described below. In all three modes, the common message M_0 is encoded jointly over the subchannels using a conventional Gaussian codebook containing 2^{nR_0} codewords of length $2n$. The first n entries of the chosen codeword are scaled by $\sqrt{(1-\theta_1^1-\theta_1^2)P_1}$ and are transmitted over the first subchannel, while the other n entries are scaled by $\sqrt{(1-\theta_2^1-\theta_2^2)P_2}$ and transmitted over the second subchannel. The transmitter selects the signalling mode based on which of the arguments of the minimization operator on the right-hand side of (3a) constrains the rate of the common message. In all three modes, the first action undertaken by each receiver is to decode the common message by jointly processing the signals that it observes on both subchannels.

1) *Mode 1:* The transmitter operates in this mode if, for the given $\boldsymbol{\theta}$, $f_0(\boldsymbol{\theta}) \leq \min\{g_0(\boldsymbol{\theta}), h_0(\boldsymbol{\theta})\}$. In this case, a rate vector that attains a subset of the bounds in (3) is $R_0 = f_0(\boldsymbol{\theta}), R_1 = f_{01}(\boldsymbol{\theta}) - R_0, R_2 = f_{012}(\boldsymbol{\theta}) - R_0 - R_1$, and $R_3 = f_{0123}(\boldsymbol{\theta}) - R_0 - R_1 - R_2$. To achieve this rate vector, the system operates as follows: The codeword corresponding to the individual message for receiver 1 is superimposed on the component of the codeword corresponding to the common message carried on subchannel 1, and the codewords corresponding to the individual messages for receivers 2 and 3 are superimposed on the component of the codeword corresponding to the common message carried on subchannel 2. More specifically, for $k = 1, 2, 3$, the messages M_k are encoded using separate conventional Gaussian codebooks of 2^{nR_k} codewords of length n, and the selected codewords are scaled by $\sqrt{(\theta_1^1+\theta_1^2)P_1}$, $\sqrt{\theta_2^2 P_2}$ and $\sqrt{\theta_2^1 P_2}$, respectively. After decoding the common message, receiver 1 decodes M_1 from its residual signal on subchannel 1 and receiver 2 decodes M_2 from its residual signal on subchannel 2. After decoding the common message, receiver 3 decodes M_2 from its residual signal on subchannel 2 and subsequently decodes M_3 from the remaining residual signal on subchannel 2.
2) *Mode 2:* The transmitter operates in this mode if, for the given $\boldsymbol{\theta}$, $g_0(\boldsymbol{\theta}) \leq \min\{f_0(\boldsymbol{\theta}), h_0(\boldsymbol{\theta})\}$. In this case, a rate vector that attains a subset of the bounds in (3) is $R_0 = g_0(\boldsymbol{\theta})$, $R_2 = g_{02}(\boldsymbol{\theta}) - R_0$, $R_1 = g_{012}(\boldsymbol{\theta}) - g_{02}(\boldsymbol{\theta}) = g_{0123}(\boldsymbol{\theta}) - g_{023}(\boldsymbol{\theta})$, and $R_3 = g_{023}(\boldsymbol{\theta}) - g_{02}(\boldsymbol{\theta}) = g_{0123}(\boldsymbol{\theta}) - g_{012}(\boldsymbol{\theta})$. To achieve this rate vector, the individual message for Receiver 2 is split into two submessages, $M_{2,1}$ and $M_{2,2}$ of rates $R_{2,i} = \frac{1}{2}\log\left(1 + \frac{\theta_i^2 P_i}{N_i^2+\theta_i^1 P_i}\right)$, respectively. Each submessage is encoded separately using its own conventional Gaussian codebook. The codewords for these submessages are then superimposed on the component of the codeword corresponding to the common message that is transmitted on the respective subchannel, with the codeword for $M_{2,1}$ being scaled by $\sqrt{\theta_1^2 P_1}$ and that for $M_{2,2}$ being scaled by $\sqrt{\theta_2^2 P_2}$. The individual messages for receivers 1 and 3 are encoded using separate conventional Gaussian codebooks. The codeword corresponding to M_1 is scaled by $\sqrt{\theta_1^1 P_1}$ and superimposed on the appropriately scaled codeword corresponding to $M_{2,1}$ and the first component of the codeword corresponding to M_0. The codeword corresponding to M_3 is scaled by $\sqrt{\theta_2^1 P_2}$ and superimposed on the appropriately scaled codeword corresponding to $M_{2,2}$ and the second component of the codeword corresponding to M_0. After decoding M_0, receiver 2 decodes $M_{2,1}$ from its residual signal on subchannel 1 and $M_{2,2}$ from its residual signal on subchannel 2. After decoding M_0, receiver 1 decodes $M_{2,1}$ and subsequently M_1 using its observations on subchannel 1. After decoding M_0, receiver 3 decodes $M_{2,2}$ and subsequently M_3 using its observations on subchannel 2.
3) *Mode 3:* The transmitter operates in this mode if, for the given $\boldsymbol{\theta}$, $h_0(\boldsymbol{\theta}) \leq \min\{f_0(\boldsymbol{\theta}), g_0(\boldsymbol{\theta})\}$. In this case, a rate vector that attains a subset of the bounds in (3) is $R_0 = h_0(\boldsymbol{\theta}), R_3 = h_{03}(\boldsymbol{\theta}) - R_0, R_2 = h_{023}(\boldsymbol{\theta}) - R_0 - R_3$, and $R_1 = h_{0123}(\boldsymbol{\theta}) - R_0 - R_2 - R_3$. This mode is symmetric with mode 1 in the sense that M_3 is carried on subchannel 2, and M_1 and M_2 are carried on subchannel 1.

Since the bounds in (3) can be attained using the three modes described above, those bounds characterize the rates that are achievable using the considered class of SPC with Gaussian signalling schemes with power partitions $\boldsymbol{\theta}$ and subchannel power allocation P_1 and P_2. Although the bounds in (3) are arranged in a natural way, the signalling modes suggest an alternate arrangement, which will be shown to facilitate the development of the main result. To simplify the expressions, we define

$$\Phi_{f_0}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq f_0(\boldsymbol{\theta}) - R_0 \tag{8a}$$
$$\Phi_{f_1}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq f_{01}(\boldsymbol{\theta}) - (R_0 + R_1) \tag{8b}$$
$$\Phi_{f_2}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq f_{012}(\boldsymbol{\theta}) - (R_0 + R_1 + R_2) \tag{8c}$$
$$\Phi_{f_3}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq f_{0123}(\boldsymbol{\theta}) - (R_0 + R_1 + R_2 + R_3) \tag{8d}$$
$$\Phi_{g_0}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq g_0(\boldsymbol{\theta}) - R_0 \tag{8e}$$
$$\Phi_{g_1}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq g_{02}(\boldsymbol{\theta}) - (R_0 + R_2) \tag{8f}$$
$$\Phi_{g_2}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq g_{012}(\boldsymbol{\theta}) - (R_0 + R_1 + R_2) \tag{8g}$$
$$\Phi_{g_3}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq g_{0123}(\boldsymbol{\theta}) - (R_0 + R_1 + R_2 + R_3) \tag{8h}$$
$$\Phi_{g_4}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq g_{023}(\boldsymbol{\theta}) - (R_0 + R_2 + R_3) \tag{8i}$$
$$\Phi_{h_0}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq h_0(\boldsymbol{\theta}) - R_0 \tag{8j}$$
$$\Phi_{h_1}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq h_{03}(\boldsymbol{\theta}) - (R_0 + R_3) \tag{8k}$$
$$\Phi_{h_2}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq h_{023}(\boldsymbol{\theta}) - (R_0 + R_2 + R_3) \tag{8l}$$
$$\Phi_{h_3}(\boldsymbol{\theta}, \boldsymbol{R}) \triangleq h_{0123}(\boldsymbol{\theta}) - (R_0 + R_1 + R_2 + R_3). \tag{8m}$$

Using these definitions, we can state that for given values of P_1 and P_2, a rate vector $\boldsymbol{R}$ can be achieved using SPC with Gaussian signalling with power partitions $\boldsymbol{\theta}$ if the following inequalities are satisfied:

$$
\begin{aligned}
&\Phi_{f_i}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0, \quad i=0,\dots,3 &\text{(9a)}\\
&\Phi_{g_i}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0, \quad i=0,1,3 &\text{(9b)}\\
&\Phi_{g_2}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0 &\text{(9c)}\\
&\Phi_{g_4}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0 &\text{(9d)}\\
&\Phi_{h_i}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0, \quad i=0,\dots,3 &\text{(9e)}\\
&\boldsymbol{R} \geq 0 &\text{(9f)}\\
&\boldsymbol{\theta} \in \mathcal{S} &\text{(9g)}
\end{aligned}
$$

where the inequality in (9f) is to be interpreted element-wise and $\mathcal{S}$ was defined in (1). In this statement, the constraints that hold with equality in mode 2 are treated differently from those that hold with equality in modes 1 and 3. This will facilitate subsequent analysis.

Having characterized this set of achievable rate vectors, it might seem natural to proceed to a direct proof of the converse; that is, a proof that all achievable rates can be characterized using (9). However, such a direct proof has proven to be elusive. Instead, in the following sections, we will develop an alternative characterization of rate vectors that can be achieved by SPC with Gaussian signalling, before proceeding with the converse in Sections VI and VII.

IV. SPC Region

The development of an alternative characterization of the rate vectors in (9) will be facilitated by pairing each rate vector with the power partition vectors that enable that rate vector to be achieved. In particular, define the SPC region corresponding to subchannel power allocations P_1 and P_2 to be the region of all (partition-rate vectors $(\boldsymbol{\theta},\boldsymbol{R})$ such that the rate vector $\boldsymbol{R}$ is achievable using SPC with the power partitions specified by $\boldsymbol{\theta}$. That is

$$\text{SPC region} \triangleq \big\{(\boldsymbol{\theta},\boldsymbol{R}) \,\big|\, \text{constraints in (9) satisfied}\big\}. \quad (10)$$

Let us also define two other rate-partition regions

$$
\begin{aligned}
\mathrm{RSPC}_1 &\triangleq \big\{(\boldsymbol{\theta},\boldsymbol{R}) \,\big|\, \text{constraints in } (9)\backslash(9\text{d}) \text{ satisfied}\big\} \quad (11\text{a})\\
\mathrm{RSPC}_2 &\triangleq \big\{(\boldsymbol{\theta},\boldsymbol{R}) \,\big|\, \text{constraints in } (9)\backslash(9\text{c}) \text{ satisfied}\big\} \quad (11\text{b})
\end{aligned}
$$

where $\backslash$ denotes the removal of a constraint. The intersection of RSPC_1 and SPC_2 is the set of all $(\boldsymbol{\theta},\boldsymbol{R})$ pairs with $\boldsymbol{\theta} \in \mathcal{S}$ such that the constraints in (9)\(9d) are satisfied and the constraints in (9)\(9c) are satisfied. Therefore

$$\mathrm{RSPC}_1 \bigcap \mathrm{RSPC}_2 = \text{SPC region}. \quad (12)$$

A. Boundary Points

In the subsequent development, partition-rate vectors that lie on the boundary of the SPC region will be of particular interest, as will those that lie on the boundaries of RSPC_1 or RSPC_2. More specifically, we will be interested in elements of these partition-rate regions that maximize a weighted sum of the rates of the individual messages for a given rate of the common message. In order to construct an initial characterization of these points, observe from (3a) that any achievable common information rate R_0 lies in $[0, R_{0,\max}]$, where [28], [35]

$$
\begin{aligned}
R_{0,\max} \triangleq \min\Big\{&\frac{1}{2}\log\Big(\frac{N_1^1+P_1}{N_1^1}\Big)+\frac{1}{2}\log\Big(\frac{N_2^3+P_2}{N_2^3}\Big),\\
&\frac{1}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2}\Big)+\frac{1}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2}\Big),\\
&\frac{1}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3}\Big)+\frac{1}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)\Big\}.
\end{aligned} \quad (13)
$$

The arguments of the minimization in (13) are the maximum rates than can be communicated to receivers 1, 2, and 3, respectively. Using that expression, the points on the boundary of RSPC_1 that maximize a weighted sum of the rates of the individual messages can be written as

$$\bigcup_{R_0\in[0,R_{0,\max}]}\ \bigcup_{\substack{w_k\geq 0\\ \sum_k w_k=1}} \Big\{(\boldsymbol{\theta},\boldsymbol{R}) \,\Big|\, (\boldsymbol{\theta},\boldsymbol{R}) = \arg\max_{\boldsymbol{\theta},\boldsymbol{R}} \sum_{k=1}^{3} w_k R_k \text{ subject to } (9)\backslash(9\text{d})\Big\} \quad (14)$$

and those for RSPC_2 can be written as

$$\bigcup_{R_0\in[0,R_{0,\max}]}\ \bigcup_{\substack{w_k\geq 0\\ \sum_k w_k=1}} \Big\{(\boldsymbol{\theta},\boldsymbol{R}) \,\Big|\, (\boldsymbol{\theta},\boldsymbol{R}) = \arg\max_{\boldsymbol{\theta},\boldsymbol{R}} \sum_{k=1}^{3} w_k R_k \text{ subject to } (9)\backslash(9\text{c})\Big\}. \quad (15)$$

For future reference, we point out that if a $(\boldsymbol{\theta},\boldsymbol{R})$ pair in (14) also satisfies (9d), then this pair lies on the boundary of the SPC region. Similarly, if a $(\boldsymbol{\theta},\boldsymbol{R})$ pair in (15) also satisfies (9c), then this pair lies on the boundary of the SPC region.

Unfortunately, neither the optimization problem in (14) nor that in (15) is convex. To make progress in their analysis, we will consider relaxations of these problems. These relaxation are inspired by the various roles played by the power partitions in each of the three modes that enable the set of rates on the boundary of the SPC region to be achieved; cf. Section III-A. First, consider the problem in (14), which is for RSPC_1. The relaxation is constructed by replacing $\boldsymbol{\theta}$ in (14) by three vectors $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$, and $\boldsymbol{\alpha}''$, and employing $\boldsymbol{\alpha}$ in the constraints involving Φ_{f_i}, $\boldsymbol{\alpha}'$ in the constraints involving Φ_{g_i}, and $\boldsymbol{\alpha}''$ in the constraints involving Φ_{h_i}, $i=0,\dots,3$. The resulting optimization problem is

$$
\begin{aligned}
\max_{\boldsymbol{\alpha},\boldsymbol{\alpha}',\boldsymbol{\alpha}'',\{R_k\}_{k=1}^3}\quad & \sum_{k=1}^{3} w_k R_k, &\text{(16a)}\\
\text{subject to}\quad & \Phi_{f_i}(\boldsymbol{\alpha},\boldsymbol{R}) \geq 0, \quad i=0,\dots,3 &\text{(16b)}\\
& \Phi_{g_i}(\boldsymbol{\alpha}',\boldsymbol{R}) \geq 0, \quad i=0,\dots,3 &\text{(16c)}\\
& \Phi_{h_i}(\boldsymbol{\alpha}'',\boldsymbol{R}) \geq 0, \quad i=0,\dots,3 &\text{(16d)}\\
& \boldsymbol{R} \geq 0, &\text{(16e)}\\
& \boldsymbol{\alpha},\boldsymbol{\alpha}',\boldsymbol{\alpha}'' \in \mathcal{S}. &\text{(16f)}
\end{aligned}
$$

Although it is not convex, this problem can be transformed into a convex optimization problem, cf. Appendix A, and we will exploit this fact in the following sections.

The optimization problem in (16) is a relaxation of that in (14) because (16) can be made equivalent to the problem in (14) by adding the constraint $\boldsymbol{\alpha} = \boldsymbol{\alpha}' = \boldsymbol{\alpha}''$. Hence, for a given set of weights, the weighted sum-rate generated by (16) is greater than or equal to that generated by (14). This suggests that the rates generated by (16) might lie outside RSPC_1. However, in the next section, we will show that, for all weights $\{w_k\}_{k=1}^3$, the optimal rate vector in (16) can be paired with one of the optimal vectors $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}''$ to construct a partition-rate vector that lies on the boundary of RSPC_1, i.e., lies in the set in (14).

Now, consider the case of RSPC_2 and the set of boundary points in (15). An analogous discussion can be used to show that, for all weights $\{w_k\}_{k=1}^3$, the rate vector generated by the corresponding relaxation of the problem in (15) can be paired with one of the optimal vectors $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}''$ to construct a partition-rate vector that lies on the boundary of RSPC_2, i.e., lies in the set in (15).

V. Original and Relaxed Boundaries of RSPC_1 and RSPC_2

In this section, we will determine relationships between the boundary points generated by the optimization problem in (14) and its relaxed counterpart in (16). By exploiting the symmetry between receivers 1 and 3 in Fig. 1, an analogous analysis can used to establish the relationship between the boundary points generated by the optimization problem in (15) and its relaxed counterpart. We will allude to the key results for that case in Sections V-D and V-E.

To determine the relationship between the solutions to the optimization problems in (14) and (16), which correspond to RSPC_1, we will construct an explicit characterization of solutions of the KKT optimality conditions, cf. [36], for the optimization problem in (14) that also solve the KKT system for the relaxed version of that problem, (16). This characterization is of considerable interest because it is shown in Appendix B that, for regular feasible points, the KKT conditions for the relaxed problem are both necessary and sufficient for optimality, whereas for the problem in (14), the KKT conditions are only necessary for optimality [36]. Therefore, the explicit characterization of solutions to both KKT systems characterizes the optimal solutions of the problem in (14), i.e., points that are on the boundary of RSPC_1.

A. KKT Conditions of the Optimization Problems Corresponding to the Original and Relaxed Boundaries of RSPC_1

One way to analyze the KKT conditions for (14) and (16) is to partition the set of feasible rates and power partitions into nonoverlapping regions, as in [36, Example 3.3.1]. In each region, particular entries of the vector of rates and power partitions are set to zero and the remaining entries are assumed to be strictly positive. Since the functions $\Phi_{f_i}(\boldsymbol{\theta}, \boldsymbol{R})$, $\Phi_{g_i}(\boldsymbol{\theta}, \boldsymbol{R})$, $\Phi_{h_i}(\boldsymbol{\theta}, \boldsymbol{R})$ in (8) are all continuous, we will focus on the case in which all the rates and power partitions are strictly positive. Continuity of these functions implies that infinitesimal changes in the power partitions result in infinitesimal changes in the data rates. Indeed, continuity implies that the cases in which some rates and partitions vanish are limiting cases of the case that we consider as those rates and partitions approach zero.

In the forthcoming analysis, we will show that for all distinct weights $\{w_k\}_{k=1}^3$, the rate vector, $\boldsymbol{R}$, and the power partition, $\boldsymbol{\theta}$, generated by (14) are identical to the rate vector, $\boldsymbol{R}$, and one of $\boldsymbol{\alpha}, \boldsymbol{\alpha}'$, or $\boldsymbol{\alpha}''$ generated by (16). To do so, we will begin with the region in which the rate vectors and partitions generated by (14) and (16) satisfy $\boldsymbol{R} > 0$ and $\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$, where $\mathrm{Int}(\mathcal{S})$ denotes the interior of $\mathcal{S}$, which is given by

$$\mathrm{Int}(\mathcal{S}) \triangleq \{\boldsymbol{\theta} | \theta_i^\ell > 0,\ \theta_i^1 + \theta_i^2 < 1,\ i, \ell = 1, 2\}.$$

For some weight orderings, it will be seen that this assumption does not incur loss of generality, simplifies the analysis and enables insight to be drawn from the KKT systems corresponding to (14) and (16). However, for other weight orderings, it will be observed that particular rates and partitions must be set to zero for the corresponding KKT systems to be solved.

First, consider the problem in (14). Let L_0 denote the Lagrangian corresponding to (14) when $\boldsymbol{R} > 0$ and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$. Using the definitions in (8), we have

$$L_0 = \sum_{k=1}^{3} w_k R_k + \sum_{i=0}^{3} \beta_i \Phi_{f_i}(\boldsymbol{\theta}, \boldsymbol{R}) + \sum_{i=0}^{3} \beta_{i+4} \Phi_{g_i}(\boldsymbol{\theta}, \boldsymbol{R}) + \sum_{i=0}^{3} \beta_{i+8} \Phi_{h_i}(\boldsymbol{\theta}, \boldsymbol{R}).$$

Using this Lagrangian, the KKT conditions for (14) can be written as

$$\frac{\partial L_0}{\partial R_k} = 0, \qquad k = 1, 2, 3 \tag{17a}$$
$$\frac{\partial L_0}{\partial \theta_i^\ell} = 0, \qquad i, \ell = 1, 2 \tag{17b}$$
$$\Phi_{f_i}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0, \quad \beta_i \Phi_{f_i}(\boldsymbol{\theta}, \boldsymbol{R}) = 0,\ i = 0, \ldots, 3 \tag{17c}$$
$$\Phi_{g_i}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0, \quad \beta_{i+4} \Phi_{g_i}(\boldsymbol{\theta}, \boldsymbol{R}) = 0,\ i = 0, \ldots, 3 \tag{17d}$$
$$\Phi_{h_i}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0, \quad \beta_{i+8} \Phi_{h_i}(\boldsymbol{\theta}, \boldsymbol{R}) = 0,\ i = 0, \ldots, 3 \tag{17e}$$
$$\boldsymbol{\beta} \geq 0 \tag{17f}$$

where $\boldsymbol{\beta} \triangleq (\beta_0, \ldots, \beta_{11})$ is the vector of nonnegative Lagrange multipliers.

Now, let L denote the Lagrangian of the problem in (16) when $\boldsymbol{R} > 0$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$. Then

$$L = \sum_{k=1}^{3} w_k R_k + \sum_{i=0}^{3} \lambda_i \Phi_{f_i}(\boldsymbol{\alpha}, \boldsymbol{R}) + \sum_{i=0}^{3} \lambda_{i+4} \Phi_{g_i}(\boldsymbol{\alpha}', \boldsymbol{R}) + \sum_{i=0}^{3} \lambda_{i+8} \Phi_{h_i}(\boldsymbol{\alpha}'', \boldsymbol{R}).$$

Using this Lagrangian, the KKT conditions for (16) are

$$\frac{\partial L}{\partial R_k} = 0, \qquad k = 1, 2, 3 \tag{18a}$$
$$\frac{\partial L}{\partial \alpha_i^\ell} = 0, \qquad i, \ell = 1, 2 \tag{18b}$$
$$\frac{\partial L}{\partial \alpha_i'^\ell} = 0, \qquad i, \ell = 1, 2 \tag{18c}$$
$$\frac{\partial L}{\partial \alpha_i''^\ell} = 0, \qquad i, \ell = 1, 2 \tag{18d}$$
$$\Phi_{f_i}(\boldsymbol{\alpha}, \boldsymbol{R}) \geq 0,\ \lambda_i \Phi_{f_i}(\boldsymbol{\alpha}, \boldsymbol{R}) = 0,\ i = 0, \ldots, 3 \tag{18e}$$
$$\Phi_{g_i}(\boldsymbol{\alpha}', \boldsymbol{R}) \geq 0,\ \lambda_{i+4} \Phi_{g_i}(\boldsymbol{\alpha}', \boldsymbol{R}) = 0,\ i = 0, \ldots, 3 \tag{18f}$$
$$\Phi_{h_i}(\boldsymbol{\alpha}'', \boldsymbol{R}) \geq 0,\ \lambda_{i+8} \Phi_{h_i}(\boldsymbol{\alpha}'', \boldsymbol{R}) = 0,\ i = 0, \ldots, 3 \tag{18g}$$
$$\boldsymbol{\lambda} \geq 0 \tag{18h}$$

where $\boldsymbol{\lambda} \triangleq (\lambda_0, \cdots, \lambda_{11})$ is the vector of nonnegative Lagrange multipliers. Because of the (partial) decoupling of constraints, it is significantly easier to gain insight into the system of equations in (18) than it is to draw insight into the KKT system for (14); cf. (17). In particular, some results concerning the nature of the Lagrange multipliers are collected in Appendix C. Those results will be employed in the proofs of the results in the following sections.

B. Common Solutions for the KKT Conditions of the Optimization Problems Corresponding to the Original and Relaxed Boundaries of RSPC_1

In this section, we provide an explicit characterization of solutions of the KKT system in (17) that also solve the KKT system in (18). There are six ways in which the weights w_1, w_2, and w_3 can be ordered and we will construct the common solutions for each case separately. We will restrict the discussion to strict weight orderings; cases involving equal weights can be analyzed similarly. The following theorem characterizes the common solutions in the cases of three of the possible weight orderings. The other cases will be considered in Theorems 2 and 3, and Corollaries 1 and 2. A key step in the proof of these results is the identification of which Lagrange multipliers must be strictly positive.

Theorem 1: Consider a given R_0 satisfying (13) and a weight vector with $w_1 > w_2 > w_3$, $w_2 > w_1 > w_3$, or $w_3 > w_2 > w_1$. For every solution of the optimization problem in (14) that satisfies $\boldsymbol{R} > 0$ and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$, the solution of (16) satisfies $\boldsymbol{R} > 0$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$. Furthermore, for every solution of the KKT system in (18) that satisfies $\boldsymbol{R} > 0$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$, the solution of the KKT system in (17) satisfies $\boldsymbol{R} > 0$ and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$. In both cases, (17) and (18) yield identical rate vectors, $\boldsymbol{R}$, identical Lagrange multipliers, $\boldsymbol{\beta} = \boldsymbol{\lambda}$, and when

a) $w_1 > w_2 > w_3, \boldsymbol{\theta} = \boldsymbol{\alpha}$

$$R_0 + R_1 = f_{01}(\boldsymbol{\theta}),$$
$$R_0 + R_1 + R_2 = f_{012}(\boldsymbol{\theta}) \quad \text{and}$$
$$R_0 + R_1 + R_2 + R_3 = f_{0123}(\boldsymbol{\theta});$$

b) $w_2 > w_1 > w_3, \boldsymbol{\theta} = \boldsymbol{\alpha}'$,

$$R_0 + R_2 = g_{02}(\boldsymbol{\theta}), \tag{19a}$$
$$R_0 + R_1 + R_2 = g_{012}(\boldsymbol{\theta}), \quad \text{and} \tag{19b}$$
$$R_0 + R_1 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta}); \tag{19c}$$

c) $w_3 > w_2 > w_1, \boldsymbol{\theta} = \boldsymbol{\alpha}''$,

$$R_0 + R_3 = h_{03}(\boldsymbol{\theta}),$$
$$R_0 + R_2 + R_3 = h_{023}(\boldsymbol{\theta}) \quad \text{and}$$
$$R_0 + R_1 + R_2 + R_3 = h_{0123}(\boldsymbol{\theta}).$$

Proof: Part (a) is proved in Appendix D, part (b) in Appendix F, and part (c) in Appendix G. ■

The results in the cases of $w_1 > w_3 > w_2$, $w_3 > w_1 > w_2$ and $w_2 > w_3 > w_1$ have a somewhat different structure, as we show in Theorems 2 and 3, and Corollaries 1 and 2 below.

Theorem 2: Given R_0 satisfying (13) and a weight vector with either $w_1 > w_3 > w_2$ or $w_3 > w_1 > w_2$, any locally optimal solution of (14) must have $R_2 = 0$. Furthermore, for these weight orderings the optimal solution of (16) must have $R_2 = 0$.

Proof: See Appendix H. ■

Corollary 1: Consider a given R_0 satisfying (13) and a weight vector with either $w_1 > w_3 > w_2$ or $w_3 > w_1 > w_2$. For every solution of the optimization problem in (14) that satisfies $\theta_i^1 \in (0, 1)$, $i = 1, 2$ and $R_k > 0$, $k = 1, 3$ the solution of (16) satisfies $\alpha_i^1, \alpha_i'^1, \alpha_i''^1 \in (0, 1)$ $i = 1, 2$, and $R_k > 0$, $k = 1, 3$. Furthermore, for every solution of the KKT system in (18), that satisfies $\alpha_i^1, \alpha_i'^1, \alpha_i''^1 \in (0, 1)$, $i = 1, 2$, and $R_k > 0$, $k = 1, 3$ the solution of the KKT system in (17) satisfies $\theta_i^1 \in (0, 1)$, $i = 1, 2$ and $R_k > 0$, $k = 1, 3$. In both cases, (17) and (18) yield identical rate vectors, $\boldsymbol{R}$, identical Lagrange multipliers, $\boldsymbol{\beta} = \boldsymbol{\lambda}$, and when

a) $w_1 > w_3 > w_2, \boldsymbol{\theta} = \boldsymbol{\alpha}$, $\theta_2^2 = \alpha_2^2 = 0$,

$$R_0 + R_1 = f_{01}(\boldsymbol{\theta}) \quad \text{and} \quad R_0 + R_1 + R_3 = f_{0123}(\boldsymbol{\theta}); \tag{20}$$

b) $w_3 > w_1 > w_2, \boldsymbol{\theta} = \boldsymbol{\alpha}''$, $\theta_1^2 = \alpha_1''^2 = 0$,

$$R_0 + R_3 = h_{03}(\boldsymbol{\theta}) \quad \text{and} \quad R_0 + R_1 + R_3 = h_{0123}(\boldsymbol{\theta}).$$

Proof: See Appendix I. ■

Theorem 3: Given R_0 satisfying (13) and a weight vector with $w_2 > w_3 > w_1$, the optimal solution of (14) must have $R_1 = 0$. Furthermore, for this weight ordering, the optimal solution of (16) must have $R_1 = 0$.

Proof: See Appendix J. ■

Corollary 2: Consider a given R_0 satisfying (13) and a weight vector with $w_2 > w_3 > w_1$. For every solution of the optimization problem in (14) that satisfies $R_k > 0$, $k = 2, 3$

and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$, the solution of (16) satisfies $R_k > 0$, $k = 2, 3$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$. Furthermore, for every solution of the KKT system in (18) that satisfies $R_k > 0$, $k = 2, 3$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$, the solution of the KKT system in (17) satisfies $R_k > 0$, $k = 2, 3$ and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$. In both cases, (17) and (18) yield identical rates vectors, $\boldsymbol{R}$, identical Lagrange multipliers, $\boldsymbol{\beta} = \boldsymbol{\lambda}$, power partitions $\boldsymbol{\theta} = \boldsymbol{\alpha}'$, and rates that satisfy $R_0 + R_2 = g_{02}(\boldsymbol{\theta})$ and $R_0 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta})$.

Proof: See Appendix K. ■

An observation regarding Theorems 1–3 and Corollaries 1 and 2 is that for any given weight ordering, each rate vector $\boldsymbol{R}$ on the relaxed boundary of RSPC_1 can be achieved by using only one of $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$, or $\boldsymbol{\alpha}''$ as the power partition vector. Since these rate vectors are identical to the rates generated by the original problem in (14), it can be seen that to attain each point that maximizes a weighted sum of the rates of the individual messages, it suffices to operate in one of the three modes described in Section III.

C. Tightness of the Relaxed Characterization of $RSPC_1$

A consequence of the above analysis is that, for any strict weight ordering, the solution of the relaxed problem in (16) can be used to construct a partition-rate vector that lies on the boundary of RSPC_1, i.e., lies in the set in (14). For future reference, we state this result formally, and outline the proof.

Theorem 4: For any $R_0 \in [0, R_{0,\max}]$ and any strict weight ordering, pairing the optimal rates R_k, $k = 1, 2, 3$, generated by (16) with one of the optimal $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}''$ yields a partition-rate vector that lies in the set in (14).

Proof: Recall that the solutions provided in Theorems 1–3 and Corollaries 1 and 2 satisfy both the KKT conditions corresponding to the problem in (14) and those corresponding to the problem in (16). Since it has been shown that the KKT conditions corresponding to (16) are both necessary and sufficient for optimality (see Appendix B), the provided solutions are sufficient for the optimality of the weighted sum-rate generated in (16). Now, (16) is a relaxation of (14), and hence, the weighted sum-rate generated by (16) is an upper bound on that generated by the problem in (14). Since the solutions provided in Theorems 1–3 and Corollaries 1 and 2 yield identical rates for both (14) and (16), we conclude that these solutions yield the maximum weighted sum-rate in (14). The matching of $\boldsymbol{\theta}$ with one of $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}''$ in Theorems 1–3 and Corollaries 1 and 2 completes the proof. ■

D. Tightness of Relaxed Characterization for $RSPC_2$

In the above sections, we have focused on RSPC_1. Analogous results can be derived for RSPC_2 by following a similar procedure and exploiting the symmetry between receivers 1 and 3 in Fig. 1.

Although we will not formally prove those results, we will briefly state the result corresponding to Theorem 4. In particular, we have

Theorem 5: For any $R_0 \in [0, R_{0,\max}]$ and any strict weight ordering, pairing the optimal rates R_k, $k = 1, 2, 3$, generated by the relaxed problem corresponding to (15) with one of the optimal $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}''$ yields a partition-rate vector that lies in the set in (15). □

E. Where Does the Boundary of the SPC Region Coincide With the Boundaries of $RSPC_1$ and $RSPC_2$?

In the previous section, it was shown that, for all weight orderings, the partition-rate vectors in RSPC_1 can be obtained from the relaxed optimization problem in (16). By analogy, using the symmetry between receivers 1 and 3, it can be argued that, for all weight orderings, the partition-rate vectors in RSPC_2 can be obtained by a similar procedure from the corresponding relaxed optimization problem. Before proceeding with the proof of the converse, in this section, we will argue that

1. the partition-rate vectors generated by the relaxed optimization problem in (16) with any strict weight ordering except $w_2 > w_3 > w_1$ lie in the SPC region.

In a complementary fashion, using the symmetry between receivers 1 and 3, the argument that we provide can be applied to show that

2. the partition-rate vectors generated by the relaxed optimization problem corresponding to RSPC_2 with any strict weight ordering except $w_2 > w_1 > w_3$ lie in the SPC region.

To prove the first statement, we consider each weight ordering separately.

1) When $w_1 > w_2 > w_3$, for any given $R_0 \in [0, R_{0\,\max}]$, the entries of the optimal partition-rate vector for RSPC_1 are obtained by using (51) in Appendix D to substitute $\beta_i = 0$, $i = 0, 4, \ldots, 11$, in the KKT conditions in (17). The power partitions θ_2^1 and θ_2^2 are obtained by solving (17b) and the rates are obtained by substituting these partitions in (52). To show that the resulting partition-rate vector lies in the SPC region, we need to show that this vector lies in RSPC_2; cf. (12). One way to do so is to repeat the procedure used in RSPC_1 for RSPC_2 with the same weight ordering and verify that the same partition-rate vector is generated by the corresponding optimization problem; i.e., the problem in (15). However, by invoking the symmetry between receivers 1 and 3, this tedious task can be avoided. In particular, this symmetry implies that the weight ordering $w_1 > w_2 > w_3$ from the perspective of RSPC_2 corresponds to the weight ordering $w_3 > w_2 > w_1$ from the perspective of RSPC_1. Now, by considering this weight ordering, swapping R_3 with R_1, swapping $\beta_i, i = 1, 2, 3$, with $\beta_j, j = 9, 10, 11$, respectively, and replacing the functions $h_{03}(\boldsymbol{\theta})$, $h_{023}(\boldsymbol{\theta})$ and $h_{0123}(\boldsymbol{\theta})$ with $f_{01}(\boldsymbol{\theta})$, $f_{012}(\boldsymbol{\theta})$ and $f_{0123}(\boldsymbol{\theta})$, respectively, it can be seen that the partition-rate vector generated by (15) when $w_1 > w_2 > w_3$ is the same as that generated by (14) for the same weights. Hence, the partition-rate vector generated by (14) when $w_1 > w_2 > w_3$ lies in RSPC_2, and therefore also in the SPC region.

2) When $w_3 > w_2 > w_1$, the above argument applies *mutatis mutandis* to show that the optimal partition-rate vector generated by (14) for the RSPC_1 lies in RSPC_2, and hence in the SPC region.

3) When $w_1 > w_3 > w_2$, for any given $R_0 \in [0, R_{0\max}]$, the optimal partition-rate vector generated by (14) for RSPC_1 must have $\theta_2^2 = 0$ and $R_2 = 0$. The rates R_1 and R_3 are given by (20), and the optimal θ_2^1 is obtained by solving the KKT conditions in (17) with the constraints $\Phi_{f_2}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$, $\Phi_{h_2}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$ and $\Phi_{g_1}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$ removed, the corresponding Lagrange multipliers, β_i, $i = 2, 5, 10$, removed, and with $\beta_i = 0$ for $i = 0, 4, 6, \ldots, 9, 11$. To show that the resulting partition-rate vector lies in the SPC region, we need to show that this vector lies in RSPC_2; cf. (12). As in the case of the above weight orderings, we exploit the symmetry between receivers 1 and 3. In particular, the weight ordering $w_1 > w_3 > w_2$ from the perspective of RSPC_2 corresponds to the weight ordering $w_3 > w_1 > w_2$ from the perspective of RSPC_1. Now, by considering this weight ordering, swapping R_3 with R_1, swapping β_i, $i = 1, 3$, with β_j, $j = 9, 11$, respectively, and replacing the functions $h_{03}(\boldsymbol{\theta})$ and $h_{0123}(\boldsymbol{\theta})$ with $f_{01}(\boldsymbol{\theta})$ and $f_{0123}(\boldsymbol{\theta})$, respectively, it can be seen that the partition-rate vector generated by (15) when $w_1 > w_3 > w_2$ is the same as that generated by (14) for the same weights. Hence, the partition-rate vector generated by (14) when $w_1 > w_3 > w_2$ lies in RSPC_2, and therefore also in the SPC region.
4) When $w_3 > w_1 > w_2$, the above argument applies *mutatis mutandis* to show that the optimal partition-rate vector generated by (14) for the RSPC_1 lies in RSPC_2, and hence in the SPC region.
5) When $w_2 > w_1 > w_3$, for any given $R_0 \in [0, R_{0\max}]$, the entries of the optimal partition-rate vector for the RSPC_1 are obtained by using (63) in Appendix F to substitute $\beta_i = 0$, $i = 0, \ldots, 4, 8, \ldots, 11$, in the KKT conditions in (17). The power partitions are obtained by solving (17b) and the rates are obtained by substituting these partitions in (64). To show that the resulting partition-rate vector lies in the SPC region, we note that the fact that the rates are determined by the equalities in (19) implies that $R_0+R_2+R_3 = g_{023}(\boldsymbol{\theta})$. Hence, the partition-rate vector generated by (14) when $w_2 > w_1 > w_3$ lies in RSPC_2, and therefore also in the SPC region. Notice that the symmetry between receivers 1 and 3 implies that the partition-rate vector generated by (15) when $w_2 > w_3 > w_1$ lies in RSPC_1, and therefore also in the SPC region.
6) When $w_2 > w_3 > w_1$, it is shown in Appendix K that there is no solution of the system of equations in (17) that lies in the feasible set of (15), i.e., no solution that lies in RSPC_2. The symmetry between receivers 1 and 3 implies that, for this weight ordering, partition-rate vectors that lie on the boundary of the SPC region can be obtained by solving the KKT conditions corresponding to (15) in a way analogous to that used for RSPC_1 when $w_2 > w_1 > w_3$.

We summarize the above points in the following remark.

Remark 1:

1) For all weight orderings other than $w_2 > w_3 > w_1$, solving (16) yields partition-rate vectors that lie on the boundary of the SPC region. When $w_2 > w_3 > w_1$, solving (16) yields partition-rate vectors that satisfy the constraints in (9)\(9d) and violate the constraint in (9d), i.e., for this weight ordering $R_0 + R_2 + R_3 > g_{023}(\boldsymbol{\theta})$. For this weight ordering, partition-rate vectors that lie on the boundary of the SPC region can be obtained by solving (15) or, equivalently, its corresponding relaxation.
2) For all weight orderings other than $w_2 > w_1 > w_3$, solving the counterpart of (16) for RSPC_2 yields partition-rate vectors that lie on the boundary of the SPC region. When $w_2 > w_1 > w_3$, solving the counterpart of (16) for RSPC_2 yields partition-rate vectors that satisfy the constraints in (9)\(9c) and violate the constraint in (9c), i.e., for this weight ordering $R_0 + R_1 + R_2 > g_{012}(\boldsymbol{\theta})$. For this weight ordering, partition-rate vectors that lie on the boundary of the SPC region can be obtained by solving (14) or, equivalently, its corresponding relaxation. □

Our goal now is to show that for any achievable rate vector $\boldsymbol{R}$, there exists a power partition $\boldsymbol{\theta} \in \mathcal{S}$ such that $(\boldsymbol{\theta}, \boldsymbol{R}) \in \mathrm{RSPC}_i$, $i = 1, 2$. Toward that end, in the next section, we will provide information-theoretic bounds on achievable rates. In Section VII, these bounds will be used together with Theorems 1–3, Corollaries 1 and 2 and Remark 1 to establish the main result of this paper.

VI. Information-Theoretic Bounds on Achievable Rates

In this section, we provide information-theoretic bounds on the achievable rates. These bounds will be used in Section VII to show that RSPC_1 and RSPC_2 contain all achievable rate vectors.

To obtain the desired information-theoretic bounds on achievable rate vectors, let ϵ_i, $i = 1, 2, 3$, be a small positive number and let

$$\begin{array}{ll} \mathcal{U}_1^3 = [M_0, M_3, Z_2, Y_2], & \mathcal{V}_1^3 = [M_0, Z_2], \\ \mathcal{U}_2^3 = [M_0, M_1, Z_1, W_1], & \mathcal{V}_2^3 = [M_0, Z_1], \\ \mathcal{U}_1^2 = [\mathcal{U}_1^3, M_2], & \mathcal{V}_1^2 = [\mathcal{V}_1^3, M_2], \\ \mathcal{X}_1^2 = [\mathcal{V}_1^3, M_2, M_3, W_2], & \\ \mathcal{U}_2^2 = [\mathcal{U}_2^3, M_2], & \mathcal{V}_2^2 = [\mathcal{V}_2^3, M_2, M_1, Y_1], \\ \mathcal{X}_2^2 = [\mathcal{V}_2^3, M_2]. & \end{array} \tag{21}$$

In Appendix L, Fano's inequality is used to show that

$$nR_0 \leq I(\mathcal{U}_1^3; Y_1) + I(\mathcal{U}_2^3; Y_2) + n\epsilon_1 \tag{22a}$$
$$n(R_0 + R_1) \leq I(\mathcal{U}_2^3; Y_2) + I(X_1; Y_1) + n\epsilon_1 \tag{22b}$$
$$n(R_0 + R_1 + R_2) \leq I(\mathcal{U}_2^3; Y_2) + I(\mathcal{U}_2^2; Z_2|\mathcal{U}_2^3) + I(X_1; Y_1) + n\epsilon_1 + n\epsilon_2 \tag{22c}$$
$$n(R_0 + R_1 + R_2 + R_3) \leq I(\mathcal{U}_2^3; Y_2) + I(\mathcal{U}_2^2; Z_2|\mathcal{U}_2^3) + I(X_2; W_2|\mathcal{U}_2^2) + I(X_1; Y_1) + n\epsilon_1 + n\epsilon_2 + n\epsilon_3 \tag{22d}$$
$$nR_0 \leq I(\mathcal{U}_1^3; W_1) + I(\mathcal{U}_2^3; W_2) + n\epsilon_3 \tag{22e}$$
$$n(R_0 + R_3) \leq I(\mathcal{U}_1^3; W_1) + I(X_2; W_2) + n\epsilon_3 \tag{22f}$$
$$n(R_0 + R_2 + R_3) \leq I(\mathcal{U}_1^3; W_1) + I(\mathcal{U}_1^2; Z_1|\mathcal{U}_1^3) + I(X_2; W_2) + n\epsilon_2 + n\epsilon_3 \tag{22g}$$
$$n(R_0 + R_1 + R_2 + R_3) \leq I(\mathcal{U}_1^3; W_1) + I(\mathcal{U}_1^2; Z_1|\mathcal{U}_1^3) + I(X_1; Y_1|\mathcal{U}_1^2) + I(X_2; W_2) + n\epsilon_1 + n\epsilon_2 + n\epsilon_3 \tag{22h}$$

where X_i is the length-n vector transmitted on subchannel i; cf. Fig. 1.

In Appendix M, it is shown that

$$nR_0 \leq I(\mathcal{V}_1^3; Z_1) + I(\mathcal{V}_2^3; Z_2) + n\epsilon_2 \tag{23a}$$
$$n(R_0 + R_2) \leq I(\mathcal{V}_1^2; Z_1) + I(\mathcal{V}_2^2; Z_2) + n\epsilon_2 \tag{23b}$$
$$n(R_0 + R_1 + R_2) \leq I(X_1; Y_1|\mathcal{V}_1^2) + I(\mathcal{V}_1^2; Z_1) + I(\mathcal{V}_2^2; Z_2) + n\epsilon_1 + n\epsilon_2 \tag{23c}$$
$$n(R_0 + R_1 + R_2 + R_3) \leq I(X_1; Y_1|\mathcal{V}_1^2) + I(X_2; W_2|\mathcal{V}_2^2) + I(\mathcal{V}_1^2; Z_1) + I(\mathcal{V}_2^2; Z_2) + n\epsilon_1 + n\epsilon_2 + n\epsilon_3. \tag{23d}$$

In Appendix M, it is also shown that

$$nR_0 \leq I(\mathcal{V}_1^3; Z_1) + I(\mathcal{V}_2^3; Z_2) + n\epsilon_2 \tag{24a}$$
$$n(R_0 + R_2) \leq I(\mathcal{X}_1^2; Z_1) + I(\mathcal{X}_2^2; Z_2) + n\epsilon_2 \tag{24b}$$
$$n(R_0 + R_2 + R_3) \leq I(X_2; W_2|\mathcal{X}_2^2) + I(\mathcal{X}_1^2; Z_1) + I(\mathcal{X}_2^2; Z_2) + n\epsilon_1 + n\epsilon_2 \tag{24c}$$
$$n(R_0 + R_1 + R_2 + R_3) \leq I(X_1; Y_1|\mathcal{X}_1^2) + I(X_2; W_2|\mathcal{X}_2^2) + I(\mathcal{X}_1^2; Z_1) + I(\mathcal{X}_2^2; Z_2) + n\epsilon_1 + n\epsilon_2 + n\epsilon_3. \tag{24d}$$

The inequalities in (22a)–(22h) and (23) will be used to show that RSPC_1 is an outer bound on the capacity region, whereas for RSPC_2, we will use (22a)–(22h) and (24).

VII. Capacity Region of the BC in Fig. 1

In this section, we will use the information-theoretic bounds derived in Section VI to show that for every achievable rate vector $\boldsymbol{R}$, there exists a power partition $\boldsymbol{\theta}$ such that the partition-rate vector $(\boldsymbol{\theta}, \boldsymbol{R}) \in \mathrm{RSPC}_i$, $i = 1, 2$. We will then use this result together with Theorems 1–3, Corollaries 1 and 2 and Remark 1 to show that the SPC region contains the set of all achievable rates.

For RSPC_1, we have

Theorem 6: For every achievable rate vector, $\boldsymbol{R}$, there exists a power partition $\boldsymbol{\theta}$ such that the partition-rate vector $(\boldsymbol{\theta}, \boldsymbol{R})$ lies in RSPC_1.

Proof: In Appendix N, the entropy power inequality and the fact that conditioning reduces entropy are used to show that for every achievable rate vector, $\boldsymbol{R}$, there exist $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathcal{S}$ such that the inequalities in (16b)–(16f) are satisfied, i.e., every achievable rate vector lies in the feasible set for (16). From the reformulation of (16) in (40) in Appendix A, it can be seen that the set of rate vectors $\{\boldsymbol{R}\}$ that are feasible for (16) is convex. Hence, it suffices to consider the achievable rate vectors that lie on the boundary of this set. This boundary can be generated by solving (16) for all possible weights $\{w_k\}_{k=1}^3$ that belong to the unit simplex $\{(w_1, w_2, w_3) | \sum_{k=1}^3 w_k = 1, w_k \geq 0, k = 1, 2, 3\}$. Now, using Theorem 4 for every rate vector $\boldsymbol{R}$ that lies on the boundary of the feasible set of (16), there exists a partition-rate vector $(\boldsymbol{\theta}, \boldsymbol{R})$ that lies on the boundary of RSPC_1, whence the statement of the theorem. ■

For RSPC_2 we have

Theorem 7: For every achievable rate vector, $\boldsymbol{R}$, there exists a power partition, $\boldsymbol{\theta}$, such that the partition-rate vector $(\boldsymbol{\theta}, \boldsymbol{R})$ lies in RSPC_2.

Proof: The proof of this theorem parallels that of Theorem 6 but with $\boldsymbol{\alpha}'$ defined as in Appendix O. ■

We are now ready to present the main result of the paper.

Theorem 8: The capacity region of the BC in Fig. 1 with given subchannel transmission powers P_1 and P_2 is the closure of the region of rates achieved by SPC with Gaussian signalling.

Proof:

1) Achievability: By construction, all rate vectors satisfying (9) can be achieved using SPC with Gaussian signalling. In particular, given a power partition vector, $\boldsymbol{\theta} \in \mathcal{S}$, rate vectors $\boldsymbol{R}$ that lie on the boundary of the region in (9) can be achieved by operating in one of the three modes described in Section III-A.
2) Converse: From Theorem 6, we have that for every achievable rate vector, $\boldsymbol{R}$, there exists a $\boldsymbol{\theta}$ such that $(\boldsymbol{\theta}, \boldsymbol{R})$ lies in RSPC_1. Since the set of rates in RSPC_1 is convex (see (40) in Appendix A), we have that for every achievable rate vector, $\boldsymbol{R}$, there exist weights w_1, w_2, and w_3 satisfying $\sum_{k=1}^3 w_k = 1$, $w_k \geq 0$, $k = 1, 2, 3$, such that $\boldsymbol{R} \leq \boldsymbol{R}^\star$, and

$$(\boldsymbol{\theta}^\star, \boldsymbol{R}^\star) = \arg\max \sum_{k=1}^{3} w_k R_k \text{ subject to } (9)\backslash(9\text{d}) \quad \text{and} \quad R_0 \in [0, R_{0,\max}]. \tag{25}$$

In the forthcoming proof, we will consider strict weight orderings of $\{w_k\}_{k=1}^3$. The convexity of the set of rates in RSPC_1 and RSPC_2 implies that restricting the proof to these orderings suffices to prove the theorem for every point arbitrarily close to the boundary of the set of achievable rates. Hence, as before, to avoid redundancy, the proof for cases involving equal weights is omitted.

Consider an arbitrary achievable rate vector, $\boldsymbol{R}$. If there exists a weight triple that satisfies one of the following five strict weight orderings $w_1 > w_2 > w_3$, $w_2 > w_1 > w_3$, $w_3 > w_2 > w_1$, $w_1 > w_3 > w_2$, or $w_3 > w_1 > w_2$, and generates $\boldsymbol{R}^\star \geq \boldsymbol{R}$, then, from Remark 1 we have that $(\boldsymbol{\theta}^\star, \boldsymbol{R}^\star) \in \text{SPC region}$, where the SPC region was defined in (10). In that case, $\boldsymbol{R}$ is achievable by SPC with Gaussian signalling.

Now, consider the case in which the only weights that yield $\boldsymbol{R}^\star$ such that $\boldsymbol{R}^\star \geq \boldsymbol{R}$ are those that satisfy $w_2 > w_3 > w_1$. This case can be partially resolved by using Theorem 7 and the convexity of the set of rate vectors in RSPC_2, which can be verified by applying the technique in Appendix A to the relaxed version of (15). In particular, using this theorem along with the convexity result, we have that, for the considered achievable rate $\boldsymbol{R}$, there exist weights w_1, w_2, and w_3 satisfying $\sum_{k=1}^3 w_k = 1, w_k \geq 0, k = 1, 2, 3$ such that $\boldsymbol{R} \leq \boldsymbol{R}^{\star\star}$, and

$$(\boldsymbol{\theta}^{\star\star}, \boldsymbol{R}^{\star\star}) = \arg\max \sum_{k=1}^{3} w_k R_k \text{ subject to } (9)\backslash(9\text{c}) \quad \text{and} \quad R_0 \in [0, R_{0,\max}]. \tag{26}$$

If there exists such a weight triple that satisfies one of the following five weight orderings $w_1 > w_2 > w_3$, $w_2 > w_3 > w_1$, $w_3 > w_2 > w_1$, $w_1 > w_3 > w_2$, or $w_3 > w_1 > w_2$, then, from Remark 1, we have that $(\boldsymbol{\theta}^{\star\star}, \boldsymbol{R}^{\star\star}) \in$

SPC region. In that case, $\boldsymbol{R}$ is achievable by SPC with Gaussian signalling.

The only case that remains to be considered is the one in which the only weights that generate an $\boldsymbol{R}^\star$ satisfying $\boldsymbol{R} \leq \boldsymbol{R}^\star$ satisfy $w_2 > w_3 > w_1$ and the only weights that generate an $\boldsymbol{R}^{\star\star}$ satisfying $\boldsymbol{R} \leq \boldsymbol{R}^{\star\star}$ satisfy $w_2 > w_1 > w_3$. In the case of RSPC_1 with $w_2 > w_3 > w_1$, Theorem 3 indicates that $R_1^\star = 0$ and hence $R_1 = 0$. The corresponding result for RSPC_2 with $w_2 > w_1 > w_3$ shows that $R_3^{\star\star} = 0$ and hence that $R_3 = 0$. Therefore, the achievable rate vectors that cannot yet be shown to be achievable by SPC with Gaussian signalling have $R_1 = R_3 = 0$. In this case, it is immediate from (14), (15), and (10) that both RSPC_1 and RSPC_2 collapse to the SPC region; when $R_1 = R_3 = 0$, the constraints $\Phi_{g_i}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$, $i = 2, 3, 4$, become redundant. This implies that in this last case $\boldsymbol{R}$ is also achievable by SPC with Gaussian signalling, which completes the proof of the theorem. ∎

As mentioned in Section II, since the capacity region for the system in Fig. 1 with a total power constraint $P_1 + P_2 \leq \bar{P}$ is the convex hull of the union of the capacity regions for each power allocation that satisfies the constraint, Theorem 8 implies that, for the system in Fig. 1 with a total power constraint, the set of rate vectors achieved by SPC with Gaussian signalling is the capacity region.

VIII. Conclusion

This paper considered the class of BCs depicted in Fig. 1, wherein each receiver receives an individual message along with a common message that is intended to all receivers. It was shown that, for this scenario, every achievable rate vector can be attained by SPC with Gaussian signalling. Our approach to establishing this result is based on an ostensibly relaxed characterization of the region of rates that can be achieved by SPC with Gaussian signalling and on showing that this relaxation is tight. Although the focus of this paper has been restricted to the scenario depicted in Fig. 1, we suspect that the same methodology can be applied to systems with some other degradation orders and possibly with more receivers.

Appendix A
Convex Transformation of (16)

In this section, we will transform the relaxed problem in (16) into a convex form. In particular, we will show that this problem can be a cast as a geometric program. As mentioned in Section II, in the analysis the power allocations, P_1 and P_2, are assumed to be fixed. However, the methodology used in this appendix can be extended to the case in which the powers are not fixed *a priori*. In order to perform the required transformation, we use the following change of variables:

$$t_k = e^{2R_k},\ k = 0, \ldots, 3, \tag{27}$$
$$Q_i^j = \alpha_i^j P_i, \quad Q_i'^j = \alpha_i'^j P_i, \quad Q_i''^j = \alpha_i''^j P_i.$$

Now, using the monotonicity of the log function, the optimization problem in (16) can be cast as

$$\max \quad t_1^{w_1} t_2^{w_2} t_3^{w_3}, \tag{28a}$$

subject to

$$t_0(N_1^1 + Q_1^1 + Q_1^2)(N_2^3 + Q_2^1 + Q_2^2)(P_1 + N_1^1)^{-1} \times (P_2 + N_2^3)^{-1} \leq 1 \tag{28b}$$
$$N_1^1 t_0 t_1 (N_2^3 + Q_2^1 + Q_2^2)(P_1 + N_1^1)^{-1} \times (P_2 + N_2^3)^{-1} \leq 1 \tag{28c}$$
$$N_1^1 t_0 t_1 t_2 (N_2^3 + Q_2^1 + Q_2^2)(N_2^2 + Q_2^1)(P_1 + N_1^1)^{-1} \times (P_2 + N_2^3)^{-1}(N_2^2 + Q_2^1 + Q_2^2)^{-1} \leq 1 \tag{28d}$$
$$N_2^1 N_1^1 t_0 t_1 t_2 t_3 (N_2^3 + Q_2^1 + Q_2^2)(N_2^2 + Q_2^1)(P_1 + N_1^1)^{-1} \times (P_2 + N_2^3)^{-1}(N_2^1 + Q_2^1)^{-1}(N_2^2 + Q_2^1 + Q_2^2)^{-1} \leq 1 \tag{28e}$$
$$t_0(N_1^2 + Q_1'^1 + Q_1'^2)(P_1 + N_1^2)^{-1}(N_2^2 + Q_2'^1 + Q_2^{2\prime}) \times (P_2 + N_2^2)^{-1} \leq 1 \tag{28f}$$
$$t_0 t_2 (N_1^2 + Q_1'^1)(N_2^2 + Q_2'^1)(P_1 + \Delta_1^2)^{-1} \times (P_2 + N_2^2)^{-1} \leq 1 \tag{28g}$$
$$N_1^1 t_0 t_1 t_2 (N_1^2 + Q_1'^1)(P_1 + \Delta_1^2)^{-1}(P_2 + \Delta_2^2)^{-1} \times (N_2^2 + Q_2'^1)(N_1^1 + Q_1'^1)^{-1} \leq 1 \tag{28h}$$
$$N_2^1 N_1^1 t_0 t_1 t_2 t_3 (N_1^2 + Q_1'^1)(N_2^2 + Q_2'^1)(N_1^1 + Q_1'^1)^{-1} \times (N_2^1 + Q_2'^1)^{-1}(P_2 + N_2^2)^{-1}(P_1 + N_1^2)^{-1} \leq 1 \tag{28i}$$
$$t_0(N_1^3 + Q_1''^1 + Q_1''^2)(P_1 + N_1^3)^{-1}(N_2^1 + Q_2''^1 + Q_2''^2) \times (P_2 + N_2^1)^{-1} \leq 1 \tag{28j}$$
$$N_2^1 t_0 t_3 (N_1^3 + Q_1''^1 + Q_1''^2)(P_1 + N_1^3)^{-1} \times (P_2 + N_2^1)^{-1} \leq 1 \tag{28k}$$
$$N_2^1 t_0 t_2 t_3 (N_1^3 + Q_1''^2 + Q_1''^1)(N_1^2 + Q_1''^1)(P_1 + N_1^3)^{-1} \times (P_2 + N_2^1)^{-1}(N_1^2 + Q_1''^1 + Q_1''^2)^{-1} \leq 1 \tag{28l}$$
$$N_2^1 N_1^1 t_0 t_1 t_2 t_3 (N_1^3 + Q_1''^1 + Q_1''^2)(N_1^2 + Q_1''^1) \times (P_1 + N_1^3)^{-1}(P_2 + N_2^1)^{-1}(N_1^1 + Q_1''^1)^{-1} \times (N_1^2 + Q_1''^1 + Q_1''^2)^{-1} \leq 1 \tag{28m}$$
$$\sum_{\ell=1}^{2} Q_i^\ell \leq P_i, \qquad i = 1, 2 \tag{28n}$$
$$\sum_{\ell=1}^{2} Q_i'^\ell \leq P_i, \qquad i = 1, 2 \tag{28o}$$
$$\sum_{\ell=1}^{2} Q_i''^\ell \leq P_i, \qquad i = 1, 2 \tag{28p}$$
$$Q_i^\ell \geq 0, \quad Q_i'^\ell \geq 0, \quad Q_i''^\ell \geq 0, \quad i = 1, 2,\ \ell = 1, 2, 3. \tag{28q}$$

Let

$$\begin{aligned}
T_1^1 &= Q_1^1 + N_1^1/2, & T_1^2 &= Q_1^2 + N_1^1/2,\\
T_2^1 &= Q_2^1 + N_2^1, & T_2^2 &= Q_2^1 + Q_2^2 + N_2^2,\\
T_1'^1 &= Q_1'^1 + N_1^2, & T_1'^2 &= Q_1'^2 + N_1^2 - N_1^1,\\
T_2'^1 &= Q_2'^1 + N_2^2, & T_2'^2 &= Q_2'^2 + N_2^2 - N_2^1,\\
T_1''^1 &= Q_1''^1 + N_1^1, & T_1''^2 &= Q_1''^1 + Q_1''^2 + N_1^2,\\
T_2''^1 &= Q_2''^1 + N_2^1/2, \quad \text{and} & T_2''^2 &= Q_2''^2 + N_2^1/2.
\end{aligned}$$

Hence

$$Q_1^1 = T_1^1 - N_1^1/2, \quad Q_1^2 = T_1^2 - N_1^1/2 \tag{29a}$$
$$Q_2^1 = T_2^1 - N_2^1, \quad Q_2^2 = T_2^2 - T_2^1 + N_2^1 - N_2^2 \tag{29b}$$
$$Q_1^{\prime 1} = T_1^{\prime 1} - N_1^2, \quad Q_1^{\prime 2} = T_1^{\prime 2} - (N_1^2 - N_1^1) \tag{29c}$$
$$Q_2^{\prime 1} = T_2^{\prime 1} - N_2^2, \quad Q_2^{\prime 2} = T_2^{\prime 2} - (N_2^2 - N_2^1) \tag{29d}$$
$$Q_1^{\prime\prime 1} = T_1^{\prime\prime 1} - N_1^1, \quad Q_1^{\prime\prime 2} = T_1^{\prime\prime 2} - T_1^{\prime\prime 1} + N_1^1 - N_1^2 \tag{29e}$$
$$Q_2^{\prime\prime 1} = T_2^{\prime\prime 1} - N_2^1/2, \quad Q_2^{\prime\prime 2} = T_2^{\prime\prime 2} - N_2^1/2. \tag{29f}$$

Using these new variables, we can rewrite (28n) as

$$\sum_{\ell=1}^{2} T_1^\ell \le P_1 + N_1^1, \quad T_2^2 \le P_2 + N_2^2. \tag{30}$$

The constraints in (28o) can be rewritten as

$$T_i^{\prime 1} + T_i^{\prime 2} \le P_i + 2N_i^2 - N_i^1, \quad i = 1, 2 \tag{31}$$

and the constraints in (28p) can be rewritten as

$$T_1^{\prime\prime 2} \le P_1 + N_1^2, \quad \sum_{\ell=1}^{2} T_2^{\prime\prime \ell} \le P_2 + N_2^1. \tag{32}$$

We now consider the conditions in (28q). First, we note that by replacing the equalities in (28n)–(28p) by the inequalities in (30)–(32), Q_i^3, $Q_i^{\prime 3}$, and $Q_i^{\prime\prime 3}$ are eliminated from the formulation. (These variables do not appear in any other constraint.) For the first set of constraints in (28q), we have

$$\begin{array}{lcl} T_1^1 \ge N_1^1/2, & & T_1^2 \ge N_1^1/2, \\ T_2^1 \ge N_2^1, & \text{and} & T_2^2 \ge T_2^1 + N_2^2 - N_2^1. \end{array}$$

For the second set, we have

$$T_i^{\prime 1} \ge N_i^2, \quad \text{and} \quad T_i^{\prime 2} \ge N_i^2 - N_i^1, \quad i = 1, 2. \tag{33}$$

For the last set, we have

$$T_1^{\prime\prime 1} \ge N_1^1, \quad T_1^{\prime\prime 2} \ge T_1^{\prime\prime 1} + N_1^2 - N_1^1, \quad \text{and} \quad T_2^{\prime\prime \ell} \ge N_2^1/2. \tag{34}$$

Before showing how the remaining constraints can be cast as a geometric program, we recall that the degradedness condition $N_i^{\ell+1} > N_i^\ell$, for $i, \ell = 1, 2$. Hence, one can see that all the transformed constraints in (30)–(34) are in the form of posynomials that can be readily incorporated in a (convex) geometric program. Using the transformation in (29), we can write the constraints in (28b)–(28e) as

$$t_0(T_1^1 + T_1^2)(N_2^3 - N_2^2 + T_2^2)(P_1 + N_1^1)^{-1}(P_2 + N_2^3)^{-1} \le 1 \tag{35a}$$
$$N_1^1 t_0 t_1 (N_2^3 - N_2^2 + T_2^2)(P_1 + N_1^1)^{-1}(P_2 + N_2^3)^{-1} \le 1 \tag{35b}$$
$$N_1^1 t_0 t_1 t_2 (N_2^3 - N_2^2 + T_2^2)(N_2^2 - N_2^1 + T_2^1)(P_1 + N_1^1)^{-1} \times (P_2 + N_2^3)^{-1} \le T_2^2 \tag{35c}$$
$$N_2^1 N_1^1 t_0 t_1 t_2 t_3 (N_2^3 - N_2^2 + T_2^2)(N_2^2 - N_2^1 + T_2^1) \times (P_1 + N_1^1)^{-1}(P_2 + N_2^3)^{-1} \le T_2^1 T_2^2. \tag{35d}$$

Note that because $N_2^2 - N_2^1 > 0$, all the constraints in (35a)–(35d) are in the standard posynomial form. Consider now the constraints in (28f)–(28i). Using the transformations in (29), these constraints can be written as

$$t_0(T_1^{\prime 1} + T_1^{\prime 2})(P_1 + N_1^2)^{-1}(T_2^{\prime 1} + T_2^{2\prime})(P_2 + N_2^2)^{-1} \le 1 \tag{36a}$$
$$t_0 t_2 (N_1^2 - N_1^1 + T_1^{\prime 1})(N_2^2 - N_2^1 + T_2^{\prime 1})(P_1 + N_1^2)^{-1} \times (P_2 + N_2^2)^{-1} \le 1 \tag{36b}$$
$$N_1^1 t_0 t_1 t_2 (N_1^2 - N_1^1 + T_1^{\prime 1})(P_1 + N_1^2)^{-1}(P_2 + N_2^2)^{-1} \times (N_2^2 - N_2^1 + T_2^{\prime 1}) \le T_1^{\prime 1} \tag{36c}$$
$$N_2^1 N_1^1 t_0 t_1 t_2 t_3 (N_1^2 - N_1^1 + T_1^{\prime 1})(N_2^2 - N_2^1 + T_2^{\prime 1}) \times (P_1 + N_1^2)^{-1}(P_2 + N_2^2)^{-1} \le T_1^{\prime 1} T_2^{\prime 1}. \tag{36d}$$

One can also see that (36a)–(36d) are in the form of posynomial constraints. Finally, we express the constraints in (28j)–(28m) as

$$t_0(N_1^3 - N_1^2 + T_1^{\prime\prime 2})(T_2^{\prime\prime 1} + T_2^{\prime\prime 2})(P_1 + N_1^3)^{-1} \times (P_2 + N_2^1)^{-1} \le 1 \tag{37a}$$
$$N_2^1 t_0 t_3 (N_1^3 - N_1^2 + T_1^{\prime\prime 2})(P_1 + N_1^3)^{-1}(P_2 + N_2^1)^{-1} \le 1 \tag{37b}$$
$$N_2^1 t_0 t_2 t_3 (N_1^3 - N_1^2 + T_1^{\prime\prime 2})(N_1^2 - N_1^1 + T_1^{\prime\prime 1})(P_1 + N_1^3)^{-1} \times (P_2 + N_2^1)^{-1} \le T_1^{\prime\prime 2} \tag{37c}$$
$$N_2^1 N_1^1 t_0 t_1 t_2 t_3 (N_1^3 - N_1^2 + T_1^{\prime\prime 2})(N_1^2 - N_1^1 + T_1^{\prime\prime 1}) \times (P_1 + N_1^3)^{-1}(P_2 + N_2^1)^{-1} \le T_1^{\prime\prime 1} T_1^{\prime\prime 2}. \tag{37d}$$

Seeing as (37a)–(37d) are in the form of posynomial constraints, we can now write (28) as

$$\max \quad t_1^{w_1} t_2^{w_2} t_3^{w_3} \tag{38a}$$
$$\text{subject to} \quad (35\text{a})\text{–}(35\text{d}), \tag{38b}$$
$$(36\text{a})\text{–}(36\text{d}), \tag{38c}$$
$$(37\text{a})\text{–}(37\text{d}), \tag{38d}$$
$$(30)\text{–}(34). \tag{38e}$$

Since the objective is in the form of a monomial and all the constraints are in the form of posynomials, the problem in (38) is readily seen to be a geometric program.

We now transform this geometric program into a convex form. In order to do that, we take the logarithm of the objective and the constraints in (38), and we use the transformations

$$X_i^\ell = \log(T_i^\ell), \quad X_i'^\ell = \log(T_i'^\ell), \text{ and } X_i''^\ell = \log(T_i''^\ell), \quad i = 1, 2, \quad \ell = 1, 2. \tag{39}$$

We will also use (27) to write $\log(t_k) = 2R_k,\ k = 0, \ldots, 3$. Using these transformations, the problem in (38) can be written as

$$\max \sum_{k=1}^{3} w_k R_k \tag{40a}$$

subject to

$$2R_0 \le \log\Big(\frac{N_1^1 + P_1}{e^{X_1^1} + e^{X_1^2}}\Big) + \log\Big(\frac{N_2^3 + P_2}{N_2^3 - N_2^2 + e^{X_2^2}}\Big) \tag{40b}$$

$$2R_0 + 2R_1 \le \log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \log\Big(\frac{N_2^3 + P_2}{N_2^3 - N_2^2 + e^{X_2^2}}\Big) \tag{40c}$$

$$2R_0 + 2R_1 + 2R_2 \le \log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \log\Big(\frac{N_2^3 + P_2}{N_2^3 - N_2^2 + e^{X_2^2}}\Big) + \log\Big(\frac{e^{X_2^2}}{N_2^2 - N_2^1 + e^{X_2^1}}\Big) \tag{40d}$$

$$2R_0 + 2R_1 + 2R_2 + 2R_3 \le \log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \log\Big(\frac{N_2^3 + P_2}{N_2^3 - N_2^2 + e^{X_2^2}}\Big) + \log\Big(\frac{e^{X_2^2}}{N_2^2 - N_2^1 + e^{X_2^1}}\Big) + \log\Big(\frac{e^{X_2^1}}{N_2^1}\Big) \tag{40e}$$

$$2R_0 \le \log\Big(\frac{N_1^2 + P_1}{e^{X_1'^1} + e^{X_1'^2}}\Big) + \log\Big(\frac{N_2^2 + P_2}{e^{X_2'^1} + e^{X_2'^2}}\Big) \tag{40f}$$

$$2R_0 + 2R_2 \le \log\Big(\frac{N_1^2 + P_1}{N_1^2 - N_1^1 + e^{X_1'^1}}\Big) + \log\Big(\frac{N_2^2 + P_2}{N_2^2 - N_2^1 + e^{X_2'^1}}\Big) \tag{40g}$$

$$2R_0 + 2R_1 + 2R_2 \le \log\Big(\frac{N_1^2 + P_1}{N_1^2 - N_1^1 + e^{X_1'^1}}\Big) + \log\Big(\frac{N_2^2 + P_2}{N_2^2 - N_2^1 + e^{X_2'^1}}\Big) + \log\Big(\frac{e^{X_1'^1}}{N_1^1}\Big) \tag{40h}$$

$$2R_0 + 2R_1 + 2R_2 + 2R_3 \le \log\Big(\frac{N_1^2 + P_1}{N_1^2 - N_1^1 + e^{X_1'^1}}\Big) + \log\Big(\frac{N_2^2 + P_2}{N_2^2 - N_2^1 + e^{X_2'^1}}\Big) + \log\Big(\frac{e^{X_1'^1}}{N_1^1}\Big) + \log\Big(\frac{e^{X_2'^1}}{N_2^1}\Big) \tag{40i}$$

$$2R_0 \le \log\Big(\frac{N_1^3 + P_1}{N_1^3 - N_1^2 + e^{X_1''^1}}\Big) + \log\Big(\frac{N_2^1 + P_2}{e^{X_2''^1} + e^{X_2''^2}}\Big) \tag{40j}$$

$$2R_0 + 2R_3 \le \log\Big(\frac{N_1^3 + P_1}{N_1^3 - N_1^2 + e^{X_1''^1}}\Big) + \log\Big(\frac{N_2^1 + P_2}{N_2^1}\Big) \tag{40k}$$

$$2R_0 + 2R_2 + 2R_3 \le \log\Big(\frac{N_1^3 + P_1}{N_1^3 - N_1^2 + e^{X_1''^1}}\Big) + \log\Big(\frac{N_2^1 + P_2}{N_2^1}\Big) + \log\Big(\frac{e^{X_1''^2}}{N_1^2 - N_1^1 + e^{X_1''^1}}\Big) \tag{40l}$$

$$2R_0 + 2R_1 + 2R_2 + 2R_3 \le \log\Big(\frac{N_1^3 + P_1}{N_1^3 - N_1^2 + e^{X_1''^1}}\Big) + \log\Big(\frac{N_2^1 + P_2}{N_2^1}\Big) + \log\Big(\frac{e^{X_1''^2}}{N_1^2 - N_1^1 + e^{X_1''^1}}\Big) + \log\Big(\frac{e^{X_1'^1}}{N_1^1}\Big) \tag{40m}$$

$$\sum_{\ell=1}^{2} e^{X_1^\ell} \le P_1 + N_1^1, \qquad e^{X_2^2} \le P_2 + N_2^2 \tag{40n}$$

$$e^{X_i'^1} + e^{X_i'^2} \le P_i + 2N_i^2 - N_i^1, \quad i = 1, 2 \tag{40o}$$

$$e^{X_1''^2} \le P_1 + N_1^2, \qquad \sum_{\ell=1}^{2} e^{X_2''^\ell} \le P_2 + N_2^1 \tag{40p}$$

$$e^{X_1^1} \ge N_1^1/2,\ e^{X_1^2} \ge N_1^1/2,\ e^{X_2^1} \ge N_2^1 \tag{40q}$$

$$e^{X_2^2} \ge e^{X_2^1} + N_2^2 - N_2^1 \tag{40r}$$

$$e^{X_i'^1} \ge N_i^2,\ e^{X_i'^2} \ge N_i^2 - N_i^1,\ i = 1, 2 \tag{40s}$$

$$e^{X_1''^1} \ge N_1^1,\ e^{X_1''^2} \ge e^{X_1''^1} + N_1^2 - N_1^1,\ e^{X_2''^\ell} \ge N_2^1/2, \quad \ell = 1, 2. \tag{40t}$$

This problem is identical to (16), but with the power partitions parametrized by the exponential function.

Appendix B
Sufficiency of the KKT Conditions for the Relaxed Optimization Problem Corresponding to RSPC_1

First, we note that for any P_1 and P_2 greater than zero, the problem in (16) is strictly feasible. From Section V-B, it is seen that for each weight ordering, the active constraints at the provided solutions are linearly independent. (For each weight ordering, each constraint that is active at the provided solution involves a distinct partial sum of $\{R_k\}_{k=1}^3$.) Hence, using [36, Proposition 3.3.1], it is seen that the KKT conditions are necessary for optimality. We now show that these conditions are also sufficient. In order to do that, we use [36, Proposition 5.1.5]. Let $L(\mathcal{A}, \boldsymbol{\gamma})$ denote the Lagrangian function at the vector of primal variables, $\mathcal{A}$, and the Lagrange multipliers, $\boldsymbol{\gamma}$. Then, from [36, Proposition 5.1.5] it is seen that it is sufficient to show that, for any vector $\boldsymbol{\gamma} \ge 0$, if the vector $\mathcal{A}^* \triangleq (\boldsymbol{R^*}, \boldsymbol{\alpha}^*, \boldsymbol{\alpha}'^*, \boldsymbol{\alpha}''^*)$ satisfies $\nabla_{\mathcal{A}} L(\mathcal{A}, \boldsymbol{\gamma})|_{\mathcal{A}=\mathcal{A}^*} = 0$, then it maximizes $L(\mathcal{A}, \boldsymbol{\gamma})$ for all feasible vectors $\mathcal{A}$. In order to show this, recall that in Appendix A we showed that (16) can be transformed into the convex form in (40). Let L_c be the Lagrangian function that corresponds to this convex problem, and let $\mathcal{B}$ be the vector of transformed variables in (39). Now

$$\nabla_{\mathcal{A}} L(\mathcal{A}, \boldsymbol{\gamma}) = \boldsymbol{J} \nabla_{\mathcal{B}} L_c(\mathcal{B}, \boldsymbol{\gamma}) \tag{41}$$

where $\boldsymbol{J}$ is the Jacobian matrix of the transformation in (39), i.e., the ijth entry of $\boldsymbol{J}$ is given by $\frac{\partial \mathcal{B}_i}{\partial \mathcal{A}_j}$. First, notice that this transformation is continuous, one-to-one and invertible. Now, one can easily check that

$$\boldsymbol{J} = \boldsymbol{I}_3 \oplus \boldsymbol{J}_1 \oplus \boldsymbol{J}_2 \oplus \boldsymbol{J}_3 \tag{42}$$

where $\oplus$ denotes the direct sum [37, p. 24], $\boldsymbol{I}_3$ denotes the 3×3 identity matrix, and

$$\boldsymbol{J}_1 = \begin{bmatrix} \frac{P_1}{N_1^1/2+P_1\alpha_1^1} & 0 \\ 0 & \frac{P_1}{N_1^1/2+P_1\alpha_1^2} \end{bmatrix} \oplus \begin{bmatrix} \frac{P_2}{N_2^1+P_2\alpha_2^1} & 0 \\ \frac{P_2}{N_2^1+P_2\alpha_2^1+P_2\alpha_2^2} & \frac{P_2}{N_2^1+P_2\alpha_2^1+P_2\alpha_2^2} \end{bmatrix}$$

$$\boldsymbol{J}_2 = \begin{bmatrix} \frac{P_1}{N_1^2+P_1\alpha_1'^1} & 0 \\ 0 & \frac{P_1}{N_1^2-N_1^1+P_1\alpha_1'^2} \end{bmatrix} \oplus \begin{bmatrix} \frac{P_2}{N_2^2+P_2\alpha_2'^1} & 0 \\ 0 & \frac{P_2}{N_2^2-N_2^1+P_2\alpha_2'^2} \end{bmatrix}$$

$$\boldsymbol{J}_3 = \begin{bmatrix} \frac{P_1}{N_1^1+P_1\alpha_1''^1} & 0 \\ \frac{P_1}{N_1^2+P_1\alpha_1''^1+P_1\alpha_1''^2} & \frac{P_1}{N_1^2+P_1\alpha_1''^1+P_1\alpha_1''^2} \end{bmatrix} \oplus \begin{bmatrix} \frac{P_2}{N_2^1/2+P_2\alpha_2''^1} & 0 \\ 0 & \frac{P_2}{N_2^1/2+P_2\alpha_2''^2} \end{bmatrix}.$$

It is clear from (42) that for any $\mathcal{A} \geq 0$, the matrix $\boldsymbol{J}$ is nonsingular. Together with (41), this implies that $\nabla_{\mathcal{A}} L(\mathcal{A}, \boldsymbol{\gamma}) = 0$ if and only if $\nabla L_c(\mathcal{B}, \boldsymbol{\gamma}) = 0$. The convexity of the problem in (28) implies that $\nabla_{\mathcal{B}} L_c(\mathcal{B}, \boldsymbol{\gamma}) = 0$ only at the global maximum of $L_c(\mathcal{B}, \boldsymbol{\gamma})$. Hence, from the continuity and the one-to-one correspondence of the transformation in (39), it can be seen that $\nabla_{\mathcal{A}} L(\mathcal{A}, \boldsymbol{\gamma})$ equals zero only at the global maximum of $L(\mathcal{A}, \boldsymbol{\gamma})$, for any given vector $\boldsymbol{\gamma} \geq 0$, and hence for the optimal Lagrange multipliers $\boldsymbol{\gamma}^*$.

APPENDIX C
ANALYSIS OF THE LAGRANGE MULTIPLIERS OF THE RELAXED PROBLEM CORRESPONDING TO RSPC_1

In this appendix, we collect some results regarding the Lagrange multipliers in the KKT conditions in (18) for the relaxed problem in (16).

Lemma 1: Any solution of the KKT system in (18) must satisfy $\lambda_0 = \lambda_4 = \lambda_8 = 0$.

Proof: In order to find λ_0, we use the fact that $f_{01}(\boldsymbol{\alpha})$, $f_{012}(\boldsymbol{\alpha})$ and $f_{0123}(\boldsymbol{\alpha})$ are not functions of α_1^ℓ, and hence from (18b), we have

$$\frac{\partial L}{\partial \alpha_1^\ell} = \frac{-\lambda_0 P_1}{N_1^1 + (\alpha_1^1 + \alpha_1^2)P_1} = 0.$$

Therefore, for any $P_1 > 0$, $\lambda_0 = 0$. For λ_4, we apply the observation that $g_{02}(\boldsymbol{\alpha}')$, $g_{012}(\boldsymbol{\alpha}')$ and $g_{0123}(\boldsymbol{\alpha}')$ are not functions of $\alpha_i'^2$ to (18c) to write

$$\frac{\partial L}{\partial \alpha_1'^2} = \frac{-\lambda_4 P_1}{N_1^2 + (\alpha_1'^2 + \alpha_1'^1)P_1} = 0$$

which yields $\lambda_4 = 0$ for any $P_1 > 0$. Similarly, by differentiating the Lagrangian with respect to $\alpha_2''^\ell$, and using (18d), one can show that for any $P_2 > 0$, $\lambda_8 = 0$. ■

Lemma 2: For any solution of the KKT system in (18), either $\lambda_1 = \lambda_2 = \lambda_3 = 0$ or $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_3 > 0$.

Proof: In order to draw some insight into the relationship between these multipliers, we use (18b) and the definitions in Section III to write

$$\frac{\partial L}{\partial \alpha_2^1} = \frac{-(\lambda_1+\lambda_2+\lambda_3)P_2}{N_2^3 + (\alpha_2^1+\alpha_2^2)P_2} + \frac{-(\lambda_2+\lambda_3)P_2\alpha_2^2}{(N_2^2 + (\alpha_2^1+\alpha_2^2)P_2)(N_2^2+\alpha_2^1 P_2)} + \frac{\lambda_3 P_2}{N_2^1+\alpha_2^1 P_2} = 0 \quad (43)$$

$$\frac{\partial L}{\partial \alpha_2^2} = \frac{-(\lambda_1+\lambda_2+\lambda_3)P_2}{N_2^3 + (\alpha_2^1+\alpha_2^2)P_2} + \frac{(\lambda_2+\lambda_3)P_2}{N_2^2 + (\alpha_2^1+\alpha_2^2)P_2} = 0. \quad (44)$$

Since $P_2 > 0$ and λ_1, λ_2, and λ_3 are nonnegative, (43) and (44) are satisfied if and only if $\lambda_1 = \lambda_2 = \lambda_3 = 0$ or

$$\lambda_3 > 0 \qquad \text{and} \qquad \lambda_1 > 0. \quad (45)$$

Furthermore, substituting from (44) into (43), we have

$$\frac{\lambda_2+\lambda_3}{N_2^2+\alpha_2^1 P_2} = \frac{\lambda_3}{N_2^1+\alpha_2^1 P_2}.$$

Since $N_2^2 > N_2^1$, this implies that

$$\lambda_2 > 0 \quad (46)$$

unless $\lambda_1 = \lambda_2 = \lambda_3 = 0$. ■

Lemma 3: For any solution of the KKT system in (18), either $\lambda_5 = \lambda_6 = \lambda_7 = 0$ or $\lambda_5 > 0$ and $\lambda_7 > 0$.

Proof: Using (18c), we have

$$\frac{\partial L}{\partial \alpha_1'^1} = \frac{-(\lambda_5+\lambda_6+\lambda_7)P_1}{N_1^2+\alpha_1'^1 P_1} + \frac{(\lambda_6+\lambda_7)P_1}{N_1^1+\alpha_1'^1 P_1} = 0 \quad (47)$$

$$\frac{\partial L}{\partial \alpha_2'^1} = \frac{-(\lambda_5+\lambda_6+\lambda_7)P_2}{N_2^2+\alpha_2'^1 P_2} + \frac{\lambda_7 P_2}{N_2^1+\alpha_2'^1 P_2} = 0. \quad (48)$$

Since P_1 and P_2 are strictly greater than zero, from (47) and (48) and the nonnegativity of λ_5, λ_6, and λ_7, we have either $\lambda_5 = \lambda_6 = \lambda_7 = 0$, or

$$\lambda_5 > 0 \qquad \text{and} \qquad \lambda_7 > 0.$$

■

Lemma 4: For any solution of the KKT system in (18), either $\lambda_9 = \lambda_{10} = \lambda_{11} = 0$ or $\lambda_9 > 0$, λ_{10} and $\lambda_{11} > 0$.

Proof: Using the definitions in Section II, we differentiate the Lagrangian with respect to $\alpha_1''^1$ and $\alpha_1''^2$. Substituting into (18d), we conclude that either $\lambda_9 = \lambda_{10} = \lambda_{11} = 0$, or

$$\lambda_9 > 0, \qquad \lambda_{10} > 0, \qquad \text{and} \qquad \lambda_{11} > 0. \quad (49)$$

■

APPENDIX D
PROOF OF THEOREM 1—ITEM (A)

To prove this theorem for $w_1 > w_2 > w_3$, we will consider the KKT system of the original problem corresponding to $\boldsymbol{R} > 0$ and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$, which is given in (17), and the KKT system of the relaxed problem corresponding to $\boldsymbol{R} > 0$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$, which is given in (18).

A) Solution of the KKT System of the Original Problem Corresponding to RSPC_1 *for* $w_1 > w_2 > w_3$*:* For $w_1 > w_2 > w_3$, (17a) yields

$$\beta_1 > \beta_5 + \beta_{10} \geq 0 \quad \text{and} \quad \beta_2 + \beta_5 + \beta_6 > \beta_9 \geq 0. \tag{50}$$

From the first inequality in (50), it is seen that $\beta_1 > 0$. Using this fact in (17b) with $i = 2$ and $\ell = 2$, it can be seen that $\beta_2 + \beta_3 > 0$. For the moment, we will assume that $\beta_2 > 0$ and $\beta_3 > 0$ and we will show later that this assumption is without loss of generality.

In Appendix E, we show that one solution of the KKT conditions can be obtained by setting

$$\beta_i = 0, \; i = 0, 4, \ldots, 11. \tag{51}$$

Using (51), we have $w_1 = \beta_1 + \beta_2 + \beta_3$, $w_2 = \beta_2 + \beta_3$, and $w_3 = \beta_3$. The complementarity slackness conditions for this choice of Lagrange multipliers yield

$$R_0 + R_1 = f_{01}(\boldsymbol{\theta}), \tag{52a}$$
$$R_0 + R_1 + R_2 = f_{012}(\boldsymbol{\theta}), \quad \text{and} \tag{52b}$$
$$R_0 + R_1 + R_2 + R_3 = f_{0123}(\boldsymbol{\theta}). \tag{52c}$$

We now show that because $\beta_2 + \beta_3 > 0$, we can assume that $\beta_2 > 0$ and $\beta_3 > 0$, without loss of generality. Toward that end, we observe that:

1) If $\beta_2 > 0$ and $\beta_3 = 0$, then R_2 is determined by the second equality in (52). Now, in this case, it may not be immediately clear that R_3 is determined by the last equality in (52). Substituting from (17b) with $i = 2$ and $\ell = 2$, into (17b) with $i = 2$ and $\ell = 1$, we have that $\beta_7 > 0$, which implies that $R_0 + R_1 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta})$. Now, suppose that R_3 is not determined by the last equality in (52). In that case, we would have $R_3 < \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big)$. Since $R_0 + R_1 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta})$, it would follow that $R_0 + R_1 + R_2 > g_{012}(\boldsymbol{\theta})$, which contradicts the left inequality in (17d) with $i = 2$. Hence, in this case, R_3 must be determined by the equalities in (52).
2) If $\beta_2 = 0$ and $\beta_3 > 0$, then $R_2 + R_3$ is determined by the third equality in (52). Furthermore, from the last two terms of the left-hand side of the equality in (17b) with $i = 2$ and $\ell = 1$ we have that $\beta_5 + \beta_6 > 0$. We will use contradiction to show that $\beta_5 = \beta_6 = 0$.

   Suppose that $\beta_5 > 0$. In that case, $R_0 + R_2 = g_{02}(\boldsymbol{\theta})$. From the left inequality in (17c) with $i = 2$, we have $R_2 \leq \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big)$, because $\beta_1 > 0$ leads to R_1 being determined by the first equality in (52). Now

$$\begin{aligned}
R_0 &\geq g_{02}(\boldsymbol{\theta}) - \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) \\
&= \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + \theta_2^1 P_2}\Big) \\
&\qquad - \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) \\
&= \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}\Big) \\
&> g_0(\boldsymbol{\theta})
\end{aligned}$$

   which contradicts the left inequality in (17d) with $i = 0$. Hence, we conclude that $\beta_5 = 0$.

   Now, assume that $\beta_6 > 0$ and consider the last two terms of the left-hand side of the equality in (17b) with $i = 1$ and $\ell = 1$. Since $\beta_5 = 0$, it is seen that $\beta_{10} > 0$. In that case, the complementarity slackness condition in (17e) with $i = 2$ implies that $R_0 + R_2 + R_3 = h_{023}(\boldsymbol{\theta})$, but since $\beta_1 > 0$ and $\beta_3 > 0$, we have $R_2 + R_3 = \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big)$. Hence, in this case, we have

$$\begin{aligned}
R_0 &= \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1}\Big) \\
&\quad + \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}{N_1^2 + \theta_1^1 P_1}\Big) \\
&\quad - \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) - \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big) \\
&= \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1 + \theta_2^1 P_2}\Big) \\
&\quad + \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}{N_1^2 + \theta_1^1 P_1}\Big) \\
&\quad - \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big) \\
&> \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) \\
&\quad + \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1 + \theta_2^1 P_2}\Big) \\
&\quad - \frac{1}{2}\log\Big(\frac{N_2^1 + (\theta_2^1 + \theta_2^2)P_2}{N_2^1 + \theta_2^1 P_2}\Big) \\
&= \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}{N_1^2 + \theta_1^1 P_1}\Big) \\
&\quad + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1 + (\theta_2^1 + \theta_2^2)P_2}\Big) \\
&> h_0(\boldsymbol{\theta})
\end{aligned}$$

   which violates the left inequality in (17e) with $i = 0$. Hence, we conclude that $\beta_6 = 0$, $\beta_2 > 0$, $\beta_3 > 0$, and that the rates R_1, R_2 and R_3 are determined by (52).

B) Solution of the KKT System of the Relaxed Problem Corresponding to RSPC_1 *for* $w_1 > w_2 > w_3$*:* Using (18a) with $k = 1$ and $k = 2$ yields

$$\lambda_1 > \lambda_5 + \lambda_{10} \geq 0 \tag{53}$$

and using (18a) with $k = 2$ and $k = 3$ yields $\lambda_2 + \lambda_5 + \lambda_6 > \lambda_9 \geq 0$. From (53), we have $\lambda_1 > 0$. Hence, from Lemma 2, we have

$$\lambda_2 > 0 \qquad \text{and} \qquad \lambda_3 > 0. \tag{54}$$

Using (18e), we have

$$R_0 + R_1 = f_{01}(\boldsymbol{\alpha}), \tag{55a}$$
$$R_0 + R_1 + R_2 = f_{012}(\boldsymbol{\alpha}) \quad \text{and} \tag{55b}$$
$$R_0 + R_1 + R_2 + R_3 = f_{0123}(\boldsymbol{\alpha}). \tag{55c}$$

Now, we set

$$\lambda_i = 0, \qquad i = 0, 4, \ldots, 11. \tag{56}$$

Notice that the setting of $\{\lambda_i\}$ in (56) is the only possible one. In particular, λ_i, $i = 0, 4, 8$, must be zero by Lemma 1.

Furthermore, Lemmas 3 and 4 imply that any other setting of $\{\lambda_i\}_{i=5,6,7,9,10,11}$ would yield a number of linearly independent equations that exceeds the number of unknowns.

Using (55) it can be seen that the setting in (56) solves the KKT system of equations in (18), for any $\boldsymbol{\alpha}'$, and $\boldsymbol{\alpha}''$ that satisfy (18f) and (18g).

C) Identifying Solutions for the KKT Systems of the Original and Relaxed Optimization Problems Corresponding to RSPC_1 *for* $w_1 > w_2 > w_3$*:* For $w_1 > w_2 > w_3$, we now compare the solution of the KKT system of the relaxed problem in the previous section with that of the KKT system of the original problem (14) in Appendix D-A. In particular, let $\boldsymbol{\lambda} = \boldsymbol{\beta}$. Since a solution of (17) exists with $\beta_i = 0,\ i = 0, 4, \ldots, 11$, it is seen that, for this solution, (17b) with $i = 2$ and $\ell = 1$ becomes identical to (43) and (44), respectively, when $\boldsymbol{\alpha} = \boldsymbol{\theta}$.

To complete the proof of the first statement of the theorem when $w_1 > w_2 > w_3$, we note that if the optimal solution of (14) in this case satisfies $(\boldsymbol{\theta}, \boldsymbol{R}) > 0$, then $\boldsymbol{\alpha}'$ and $\boldsymbol{\alpha}''$ can be arbitrarily chosen so that (18f) and (18g) are satisfied. In particular, since $(\boldsymbol{\theta}, \boldsymbol{R})$ belongs to the feasible set of (14), then setting $\boldsymbol{\alpha} = \boldsymbol{\theta}$ and setting $\boldsymbol{\alpha}' = \boldsymbol{\alpha}'' = \boldsymbol{\alpha}$ is guaranteed to satisfy (18f) and (18g). To complete the proof of the second statement of the theorem, we note from the previous section, that for this weight ordering, the solution of the KKT system in (18) when $(\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'', \boldsymbol{R}) > 0$ must have the entries of the vector $\boldsymbol{R}$ determined by (55), which implies that there exist $\boldsymbol{\alpha}'$ and $\boldsymbol{\alpha}''$ that satisfy (18f) and (18g). Finally, it is seen that with $\boldsymbol{\alpha}$ identified with $\boldsymbol{\theta}$, the rates generated by (55) are identical to those generated by (52), which completes the proof of item (a) of Theorem 1.

Appendix E
Proving That for $w_1 > w_2 > w_3$, $\beta_i = 0,\ i = 0, 4, \ldots, 11$

When $w_1 > w_2 > w_3$, we have $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 > 0$; see Appendix D-A. First, let us assume that $\beta_{10} + \beta_{11} > 0$. We will use contradiction to show that $\beta_{10} = \beta_{11} = 0$.

For this weight ordering, $\beta_1 > 0$, $\beta_2 > 0$, and $\beta_3 > 0$, and hence, the rates are determined by (17c). Therefore, from the second equality in (17c) with $i = 1$, we have

$$R_1 = \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^1 + \alpha_2^2)P_2}\Big) - R_0. \tag{57}$$

Let us assume that $\beta_{10} > 0$ and $\beta_{11} > 0$. In that case, using (17e) with $i = 2$ and $i = 3$, we have

$$R_1 = \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big). \tag{58}$$

Substituting from (58) into (57) yields

$$\begin{aligned} R_0 &= \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big) + \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) \\ &\quad - \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big) \\ &> \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big) \\ &\quad + \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1 + (\theta_1^1 + \theta_1^2)P_1}\Big) = f_0(\boldsymbol{\theta}) \end{aligned}$$

which contradicts the first constraint in (17c) with $i = 0$. Hence, it is seen that the case of $\beta_{10} > 0$ and $\beta_{11} > 0$ can be eliminated.

A similar argument can be used to eliminate the possibility that either β_{10} or β_{11} is greater than zero. If $\beta_{10} = 0$ and $\beta_{11} > 0$, then from (17b) with $i = 1$ and $\ell = 1$, $\beta_5 > 0$. Using the complementarity slackness condition associated with β_5 and the fact that R_2 is determined by the equality in (17c) with $i = 2$, one can show that R_0 violates the left constraint in (17d) with $i = 0$. Hence, the possibility of $\beta_{10} = 0$ and $\beta_{11} > 0$ can also be eliminated.

Now, let us consider the possibility that $\beta_{10} > 0$ and $\beta_{11} = 0$. In this case, $R_1 \le \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big)$. Using this expression in the first equality in (52) yields an R_0 that violates the first constraint in (17c) with $i = 0$. Hence, the possibility that $\beta_{10} > 0$ and $\beta_{11} = 0$ can be eliminated.

In the argument above, we have shown that $\beta_{10} + \beta_{11} = 0$. Since the last term on the left-hand side of (17b) with $i = 2$ and $\ell = 1$ is the only nonnegative term, we conclude that $\beta_0 = \beta_4 = \beta_8 = \beta_9 = 0$. Substituting $\beta_i = 0$ for $i = 0, 4, 8, 9, 10, 11$ into (17b) with $i = 1$ and $\ell = 1$ we have,

$$\frac{\beta_6 + \beta_7}{N_1^1 + \theta_1^1 P_1} = \frac{\beta_5 + \beta_6 + \beta_7}{N_1^2 + \theta_1^1 P_1}. \tag{59}$$

Now, we will use contradiction to show that $\beta_5 = 0$. To do that, we note that if $\beta_5 > 0$ then from the last equality in (17d) with $i = 1$

$$R_0 + R_2 = \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + \theta_2^1 P_2}\Big). \tag{60}$$

However, because $\beta_2 > 0$, we have

$$R_2 = \frac{1}{2}\log\Big(\frac{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}{N_2^2 + \theta_2^1 P_2}\Big). \tag{61}$$

Substituting from (61) into (60) and simplifying yields

$$\begin{aligned} R_0 &= \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + (\theta_2^1 + \theta_2^2)P_2}\Big) \\ &> g_0(\boldsymbol{\theta}) \end{aligned}$$

which violates (17d) with $i = 0$ for any $\theta_2^2 > 0$. Hence, we must have $\beta_5 = 0$ which, using (59) and the fact that $N_1^2 > N_1^1$, leads to $\beta_6 = \beta_7 = 0$. We have thus shown that when $w_1 > w_2 > w_3$, $\beta_i = 0,\ i = 0, 4, \ldots, 11$, as desired.

Appendix F
Proof of Theorem 1—Item (b)

A) Solution of the KKT System of the Original Problem Corresponding to RSPC_1 *for* $w_2 > w_1 > w_3$*:* For $w_2 > w_1 > w_3$, the equalities in (17a) yield

$$\beta_5 + \beta_{10} > \beta_1 \ge 0 \quad \text{and} \quad \beta_2 + \beta_5 + \beta_6 > \beta_9 \ge 0. \tag{62}$$

In Appendix E, it was shown that if $\beta_1 > 0$ then $\beta_5 = \beta_{10} = 0$, which contradicts the first inequality in (62). Hence, it is seen that in this case $\beta_1 = 0$. A similar argument can be used to show that $\beta_9 = 0$. We will show below that a solution of the KKT conditions in this case can be obtained by setting $\beta_2 = \beta_3 = 0$.

For $\beta_2 + \beta_3 = 0$, we have from (17b) with $i = 2$ and $\ell = 2$ that $\beta_0 = \beta_4 = \beta_8 = 0$. Applying this fact and the fact that $\beta_9 = 0$ to (17b) with $i = 1$ and $\ell = 2$ yields $\beta_{10} = \beta_{11} = 0$. Using

$$\beta_i = 0,\ i = 0, \ldots, 4, 8, \ldots, 11 \tag{63}$$

in (17a) and (17b) with the current weight ordering yields $\beta_5 > 0$, $\beta_6 > 0$ and $\beta_7 > 0$, from which we have

$$R_0 + R_2 = g_{02}(\boldsymbol{\theta}), \tag{64a}$$
$$R_0 + R_1 + R_2 = g_{012}(\boldsymbol{\theta}) \quad \text{and} \tag{64b}$$
$$R_0 + R_1 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta}). \tag{64c}$$

It remains to show that the rates generated in (64) are feasible, i.e., that setting $\beta_2 = \beta_3 = 0$ yields a solution of the KKT system of equations. Suppose that $\beta_2 + \beta_3 > 0$, then from (17b) with $i = 2$ and $\ell = 2$ we have that at least one $\{\beta_i\}_{i=0,4,8}$ is greater than zero. Using this fact in (17b) with $i = 1$ and $\ell = 2$ yields $\beta_{10} + \beta_{11} > 0$. Now, suppose that $\beta_5 + \beta_6 + \beta_7 = 0$. (This assumption results in fewer constraints being active.) In this case, we have $\beta_2 > 0, \beta_3 > 0, \beta_{10} > 0$ and $\beta_{11} > 0$. The complementarity slackness now implies that at least one of the constraints on R_0 is active, and for each rate R_k, $k = 1, 2, 3$, one can find two active constraints. Equating these expressions, it is seen that in this case, we have four equations in the four unknowns, $\{\theta_i^\ell\}_{i,\ell=1,2}$. Solving for these unknowns and substituting into (17b), we have four equations in five unknowns, $\beta_2, \beta_3, \beta_{10}, \beta_{11}$ and one of $\{\beta_i\}_{i=0,4,8}$, in addition to the three equations in (17a). That is, in total we have seven linearly independent linear equations in five unknowns. Since these equations cannot be consistent we conclude that one must have $\beta_2 + \beta_3 = 0$ for the problem to be feasible.

B) Solution of the KKT System of the Relaxed Problem Corresponding to RSPC_1 *for* $w_2 > w_1 > w_3$*:* For this weight ordering, the KKT conditions of the relaxed problem yield

$$\lambda_5 + \lambda_{10} > \lambda_1 \geq 0 \quad \text{and} \quad \lambda_2 + \lambda_5 + \lambda_6 > \lambda_9 \geq 0. \tag{65}$$

Now, set $\lambda_i = 0,\ i = 0, \ldots, 4, 9, 10, 11$. Notice that this setting of $\{\lambda_i\}$ is the only possible one. In particular, λ_i, $i = 0, 4, 8$, must be zero by Lemma 1 in Appendix C. Furthermore, Lemmas 2 and 4 imply that any other setting of $\{\lambda_i\}_{i=5,6,7,9,10,11}$ would yield a number of linearly independent equations that exceeds the number of unknowns. Using this setting in (18a), along with Lemma 3 and the fact that $w_2 > w_1 > w_3$, we have that

$$R_0 + R_2 = g_{02}(\boldsymbol{\alpha}'),$$
$$R_0 + R_1 + R_2 = g_{012}(\boldsymbol{\alpha}') \quad \text{and}$$
$$R_0 + R_1 + R_2 + R_3 = g_{0123}(\boldsymbol{\alpha}').$$

Now, $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}''$ can be arbitrarily chosen so that (18e) and (18g) are satisfied, respectively.

C) Identifying Solutions of the KKT Systems of the Original and Relaxed Optimization Problems Corresponding to RSPC_1 *for* $w_2 > w_1 > w_3$*:* For $w_2 > w_1 > w_3$, the solution of the KKT system of the relaxed problem obtained in the previous section is now compared with that of the KKT system of the original problem in (14). Since a solution of (17) exists with $\beta_i = 0,\ i = 0, \ldots, 4, 8, \ldots, 11$, it is seen that, for this solution, (17b) with $i = 2$ and $\ell = 1$, and (17b) with $i = 2$ and $\ell = 2$ become identical to (43) and (44), respectively, when $\boldsymbol{\alpha}' = \boldsymbol{\theta}$. Now, we choose $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}''$ arbitrarily so that (18e) and (18g) are satisfied. Similar to the case considered in Appendix D, it can be seen that such $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}''$ exist when the solution of (14) satisfies $\boldsymbol{R} > 0$ and $\boldsymbol{\theta} \in \mathrm{Int}(\mathcal{S})$, or conversely, when the solution of the system in (18) satisfies $\boldsymbol{R} > 0$ and $\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$. Finally, it is seen that with $\boldsymbol{\alpha}'$ identified with $\boldsymbol{\theta}$, the rates generated by (55) are identical to those generated by (52), which completes the proof of item (b) of Theorem 1.

Appendix G
Proof of Theorem 1—Item (c)

A) Solution of the KKT System of the Original Problem Corresponding to RSPC_1 *for* $w_3 > w_2 > w_1$*:* The ordering $w_3 > w_2 > w_1$ (17a) implies that

$$\beta_9 > 0 \qquad \text{and} \qquad \beta_5 + \beta_{10} > 0. \tag{66}$$

Using $\beta_9 > 0$ in (17b) with $i = 1$ and $\ell = 2$ implies that $\beta_{10} + \beta_{11} > 0$. We will show that in this case $\beta_{10} > 0$ and $\beta_{11} > 0$. Hence, the rates in this case are given by

$$R_0 + R_3 = h_{03}(\boldsymbol{\theta}), \tag{67a}$$
$$R_0 + R_2 + R_3 = h_{023}(\boldsymbol{\theta}) \quad \text{and} \tag{67b}$$
$$R_0 + R_1 + R_2 + R_3 = h_{0123}(\boldsymbol{\theta}). \tag{67c}$$

Toward that end, we will show that both the case $\beta_{10} = 0$ and $\beta_{11} > 0$ and the case of $\beta_{10} > 0$ and $\beta_{11} = 0$ yield rates that violate the KKT conditions. First, we consider the case of $\beta_{10} = 0$ and $\beta_{11} > 0$. In this case, (66) implies that $\beta_5 > 0$. Thus

$$R_0 + R_2 = \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \theta_1^1 P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_1^2 + \theta_2^1 P_2}\Big). \tag{68}$$

Using (17e) with $i = 2$ implies that

$$R_2 \leq \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2) P_1}{N_1^2 + \theta_1^1 P_1}\Big). \tag{69}$$

Substituting from (69) in (68) implies that

$$R_0 \geq \frac{1}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + (\theta_1^1 + \theta_1^2) P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^2 + P_2}{N_1^2 + \theta_2^1 P_2}\Big)$$
$$> g_0(\boldsymbol{\theta})$$

which violates the condition in (17d) with $i = 0$ for any $\theta_2^2 > 0$. Hence, we conclude that β_{10} must be greater than zero.

Now we consider the case of $\beta_{10} > 0$ and $\beta_{11} = 0$. In this case

$$R_2 = \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2) P_1}{N_1^2 + \theta_1^1 P_1}\Big). \tag{70}$$

Substituting $\beta_{11} = 0$ in (17b) with $i = 1$, $\ell = 1$ and $\ell = 2$ and simplifying, it can be shown that

$$\beta_6 + \beta_7 > 0. \tag{71}$$

We will use contradiction to show that $\beta_6 = \beta_7 = 0$, which contradicts (71).

Suppose that $\beta_6 > 0$. In this case

$$R_0+R_1+R_2 = \frac{1}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+\theta_1^1 P_1}\Big)+\frac{1}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\theta_2^1 P_2}\Big) + \frac{1}{2}\log\Big(\frac{N_1^1+\theta_1^1 P_1}{N_1^1}\Big). \quad (72)$$

Since $\beta_{11} = 0$ and $\beta_{10} > 0$, we know that

$$R_1 \leq \frac{1}{2}\log\Big(\frac{N_1^1+\theta_1^1 P_1}{N_1^1}\Big). \quad (73)$$

Substituting from (73), (70) into (72) yields

$$R_0 \geq \frac{1}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+(\theta_1^1+\theta_1^2)P_1}\Big)+\frac{1}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\theta_2^1 P_2}\Big) > g_0(\boldsymbol{\theta}),$$

which violates the condition in (17d) with $i = 0$ for any $\theta_2^2 > 0$. Hence, we conclude that if $\beta_{10} > 0$ and $\beta_{11} = 0$, β_6 must be equal to zero.

We now consider the other possibility for (71) to be satisfied; i.e., that $\beta_7 > 0$. Using the fact that $\beta_6 = 0$ and substituting in (17b) with $i = 2$, $\ell = 1$ and $\ell = 2$ and simplifying yields

$$\beta_2 + \beta_5 > 0. \quad (74)$$

We will show that both β_2 and β_5 must be zero, which yields a contradiction.

1) If $\beta_5 > 0$, R_0+R_2 is given by (68). However, since $\beta_9 > 0$ and $\beta_{10} > 0$, R_2 is given by (70), which implies that

$$R_0 = \frac{1}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+(\theta_1^1+\theta_1^2)P_1}\Big)+\frac{1}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\theta_2^1 P_2}\Big) > g_0(\boldsymbol{\theta})$$

which violates the condition in (17d) with $i = 0$ for any $\theta_2^2 > 0$. Hence, we conclude that if $\beta_{10} > 0$, $\beta_{11} = 0$ and $\beta_7 > 0$, then $\beta_5 = 0$.

2) If $\beta_2 > 0$, the condition (17c) with $i = 2$ implies that

$$R_0+R_1+R_2 = \frac{1}{2}\log\Big(\frac{N_2^3+P_2}{N_2^3+(\theta_2^1+\theta_2^2)P_2}\Big) + \frac{1}{2}\log\Big(\frac{N_1^1+P_1}{N_1^1}\Big)+\frac{1}{2}\log\Big(\frac{N_2^2+(\theta_2^1+\theta_2^2)P_2}{N_2^2+\theta_2^1 P_2}\Big).$$

Now, because (17c) with $i = 3$ must be satisfied, we know that

$$R_3 \leq \frac{1}{2}\log\Big(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\Big). \quad (75)$$

Using the fact that $\beta_9 > 0$ implies that

$$R_0+R_3 = \frac{1}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\theta_1^1+\theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)$$

and substituting from (75) yields

$$R_0 \geq \frac{1}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\theta_1^1+\theta_1^2)P_1}\Big)+\frac{1}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1+\theta_2^1 P_2}\Big) > h_0(\boldsymbol{\theta})$$

which violates the condition in (17e) with $i = 0$ for any $\theta_2^2 > 0$. Hence, we conclude that if $\beta_{10} > 0$, $\beta_{11} = 0$ and $\beta_7 > 0$, then $\beta_2 = 0$.

Hence, if $\beta_{10} > 0, \beta_{11} = 0$ and $\beta_7 > 0$, we must have $\beta_2+\beta_5 = 0$, which violates (74). Therefore, β_7 must be equal to zero and $\beta_{10} > 0$ and $\beta_{11} > 0$, which imply (67).

B) Solution of the KKT System of the Relaxed Problem Corresponding to RSPC_1 *for* $w_3 > w_2 > w_1$*:* The analysis of the KKT system of the relaxed problem in (18) and the identification of the solution uses a technique similar to the one used for the case of $w_1 > w_2 > w_3$ and is omitted for brevity.

Appendix H
Proof of Theorem 2

We proceed by contradiction. In particular, we will show that assuming that $\boldsymbol{R} > 0$ and $\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathrm{Int}(\mathcal{S})$, will yield a contradiction. This contradiction is resolved by setting R_2 to zero.

A) Original Problem With $w_1 > w_3 > w_2$ *and* $w_3 > w_1 > w_2$*:* In this section, we show that, given R_0 satisfying (13) and a weight vector with $\min\{w_1, w_3\} > w_2$, any locally optimal solution of (14) must have $R_2 = 0$. First, assume that $R_k > 0$, for $k = 1, 2, 3$. We will show in this section that for the considered weight settings this assumption leads to a contradiction.

We begin by noting that under the assumption that $R_k > 0$, using $w_1 > w_2$ and $w_3 > w_2$ in (17a) yields

$$\beta_1 > \beta_5 + \beta_{10} \geq 0 \quad \text{and} \quad \beta_9 > \beta_2 + \beta_5 + \beta_6 \geq 0. \quad (76)$$

We will now show that β_1 and β_9 cannot be strictly greater than zero simultaneously, and we will use this to conclude that $R_2 = 0$. If $\beta_1 > 0$, we have from (17b) with $i = 2$ and $\ell = 2$ that $\beta_2 + \beta_3 > 0$. Similarly, if $\beta_9 > 0$, we have from (17b) with $i = 1$ and $\ell = 2$ that $\beta_{10} + \beta_{11} > 0$. Also, since $\beta_1 > 0$, we have that

$$R_0 + R_1 = f_{01}(\boldsymbol{\theta}) \quad (77)$$

and since $\beta_9 > 0$

$$R_0 + R_3 = h_{03}(\boldsymbol{\theta}). \quad (78)$$

2) Showing That $\beta_3 = \beta_{11} = 0$*:* Suppose that $\beta_{11} > 0$. In this case, using (78) and the equality in (17e) with $i = 3$, we have

$$R_1 + R_2 = \frac{1}{2}\log\Big(\frac{N_1^1+\theta_1^1 P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_1^2+(\theta_1^1+\theta_1^2)P_1}{N_1^2+\theta_1^1 P_1}\Big). \quad (79)$$

From (77) and (79), we have

$$\begin{aligned} R_0 &> R_0 + R_1 - (R_1 + R_2) \qquad (80)\\ &= R_0 - R_2\\ &= \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big)\\ &\quad - \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big) - \frac{1}{2}\log\Big(\frac{N_1^2 + (\theta_1^1 + \theta_1^2)P_1}{N_1^2 + \theta_1^1 P_1}\Big)\\ &> \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big)\\ &\quad - \frac{1}{2}\log\Big(\frac{N_1^1 + \theta_1^1 P_1}{N_1^1}\Big) - \frac{1}{2}\log\Big(\frac{N_1^1 + (\theta_1^1 + \theta_1^2)P_1}{N_1^1 + \theta_1^1 P_1}\Big)\\ &= \frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1 + (\theta_1^1 + \theta_1^2)P_1}\Big)\\ &\qquad + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\theta_2^1 + \theta_2^2)P_2}\Big) \qquad (81) \end{aligned}$$

which contradicts the inequality constraint in (17c) with $i = 0$. Hence, we conclude that $\beta_{11} = 0$. Notice that the contradiction here is resolved if $R_2 = 0$. Using a similar argument, we conclude that $\beta_3 = 0$. The contradiction for the latter case is also resolved if $R_2 = 0$.

3) Showing That $\beta_2 = \beta_{10} = 0$: In order to show that R_k cannot be strictly greater than zero for all $k = 1, 2, 3$, we have to show that $\beta_2 = 0$ (or alternatively that $\beta_{10} = 0$). This will contradict (17b) with $i = 2$ and $\ell = 2$ for $\beta_1 > 0$.

If $\beta_2 > 0$, then from the equality in (17c) with $i = 2$ we have $R_0 + R_1 + R_2 = f_{012}(\boldsymbol{\theta})$. Using this in the inequality constraint in (17c) with $i = 3$, we have $R_3 \le \frac{1}{2}\log\Big(\frac{N_2^1 + \alpha_2^1 P_2}{N_2^1}\Big)$. Using that in (78) yields

$$\begin{aligned} R_0 &\ge \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1}\Big)\\ &\qquad - \frac{1}{2}\log\Big(\frac{N_2^1 + \theta_2^1 P_2}{N_2^1}\Big)\\ &= \frac{1}{2}\log\Big(\frac{N_1^3 + P_1}{N_1^3 + (\theta_1^1 + \theta_1^2)P_1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^1 + P_2}{N_2^1 + \theta_2^1 P_2}\Big). \end{aligned}$$

This inequality contradicts the inequality constraint in (17e) with $i = 0$ for any $\theta_2^2 > 0$. The argument that $\beta_{10} = 0$ can be made analogously.

D) Relaxed Problem With $w_1 > w_3 > w_2$ and $w_3 > w_1 > w_2$: In this section, we show that, when $\min\{w_1, w_3\} > w_2$, the optimal solution of (16) must have $R_2 = 0$. To show this, we will assume that $R_k > 0$, for $k = 1, 2, 3$, and then proceed by contradiction. If $R_k > 0$, for $k = 1, 2, 3$, then from (18a) with $k = 1$ and $k = 2$ we have that

$$\lambda_1 > \lambda_5 + \lambda_{10} \ge 0. \qquad (82)$$

and from (18a) with $k = 2$ and $k = 3$, we have that

$$\lambda_9 > \lambda_5 + \lambda_6 + \lambda_2 \ge 0. \qquad (83)$$

Now, using (83) and (82), we have from (45), (46), (49), (18e) and (18g) that

$$\begin{aligned} R_0 + R_1 &= f_{01}(\boldsymbol{\alpha}), && (84a)\\ R_2 &= \frac{1}{2}\log\Big(\frac{N_2^2 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^2 + \alpha_2^1 P_2}\Big) \quad \text{and} && (84b)\\ R_3 &= \frac{1}{2}\log\Big(\frac{N_2^1 + \alpha_2^1 P_2}{N_2^1}\Big) && (84c) \end{aligned}$$

and

$$\begin{aligned} R_0 + R_3 &= h_{03}(\boldsymbol{\alpha}''), && (85a)\\ R_2 &= \frac{1}{2}\log\Big(\frac{N_1^2 + (\alpha_1''^1 + \alpha_1''^2)P_1}{N_1^2 + \alpha_1''^1 P_1}\Big) \quad \text{and} && (85b)\\ R_1 &= \frac{1}{2}\log\Big(\frac{N_1^1 + \alpha_1''^1 P_1}{N_1^1}\Big). && (85c) \end{aligned}$$

We will now show that $R_2 = 0$ and that $\alpha_1''^2 = \alpha_2^2 = 0$. Consider the value of the objective that corresponds to the equalities in (84), namely

$$\begin{aligned} &w_1\Big(\frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^1 + \alpha_2^2)P_2}\Big) - R_0\Big)\\ &+ w_2\frac{1}{2}\log\Big(\frac{N_2^2 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^2 + \alpha_2^1 P_2}\Big) + w_3\frac{1}{2}\log\Big(\frac{N_2^1 + \alpha_2^1 P_2}{N_2^1}\Big). \qquad (86) \end{aligned}$$

Observe that, for a given R_0, the value of the objective does not depend on the partitions α_1^1 and α_1^2. Consider now the value of the objective that corresponds to power partitions $\gamma_2^1 = \alpha_2^1 + \alpha_2^2$ and $\gamma_2^2 = 0$. In this case, the value of the objective is given by

$$\begin{aligned} &w_1\Big(\frac{1}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{1}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + \gamma_2^1 P_2}\Big) - R_0\Big)\\ &\qquad + w_3\frac{1}{2}\log\Big(\frac{N_2^1 + \gamma_2^1 P_2}{N_2^1}\Big). \qquad (87) \end{aligned}$$

Subtracting (86) from (87) yields

$$\begin{aligned} &w_3\frac{1}{2}\log\Big(\frac{N_2^1 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^1 + \alpha_2^1 P_2}\Big)\\ &\qquad - w_2\frac{1}{2}\log\Big(\frac{N_2^2 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^2 + \alpha_2^1 P_2}\Big) > 0 \end{aligned}$$

where the inequality follows from the fact that $N_2^2 > N_2^1$ and $w_2 < w_3$. Hence, it is seen that, for this weight ordering, the rates in (84) are not optimal unless $R_2 = 0$ and $\alpha_2^2 = 0$. Using a similar argument, we can show that the rates in (85) are not optimal unless $R_2 = 0$ and $\alpha_1''^2 = 0$.

APPENDIX I
PROOF OF COROLLARY 1

In Theorem 2, we have shown that for these weight orderings both the original problem in (14) and its relaxed counterpart in (16) yield $R_2 = 0$.

We now consider the optimization problems corresponding to the ones in (14) and (16) but with $R_2 = 0$. In this case, it is straightforward to see that for these problems the constraints

$\Phi_{g_1}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0, \Phi_{f_2}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0$ and $\Phi_{h_2}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0$ are redundant. Using this observation, the KKT systems corresponding to $R_2 = 0$ can be obtained from (17) and (18) by removing the constraints $\Phi_{g_1}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0, \Phi_{f_2}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0$ and $\Phi_{h_2}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0$ and their corresponding Lagrange multipliers, namely, β_i and λ_i, $i = 2, 5, 10$.

For the original problem corresponding to (14), following an analysis similar to the one used for other weight orderings, it can be seen that no solution of the KKT system exists if both θ_1^2 and θ_2^2 are greater than zero. For the case in which $w_1 > w_3$ and $R_2 = 0$, it can be shown from the analysis of the KKT system that the solution of (14) must have $\theta_2^2 = 0$. Now, assuming that $\theta_1^2 > 0$ and $\theta_2^2 = 0$ yields $\beta_6 + \beta_7 + \beta_{11} = 0$. This implies that $\beta_0 = \beta_4 = \beta_8 = \beta_9 = 0$. A solution of the KKT system in this case is found with $\beta_1 > 0$ and $\beta_3 > 0$.

For the case in which $w_3 > w_1$, a similar analysis for the original problem reveals that $\beta_9 > 0$, which can be shown to yield $\beta_{11} > 0$ and $\theta_1^2 = 0$.

For the relaxed problem corresponding to (16), the condition that $w_1 > w_3$ implies that $\lambda_1 > 0$, which, using Lemma 2, implies that $\lambda_3 > 0$ and $\alpha_2^2 = 0$. Analogously, the condition that $w_3 > w_1$ implies that $\lambda_9 > 0$, which, using Lemma 4, implies that $\lambda_{11} > 0$ and $\alpha_1^2 = 0$.

The statement of the corollary follows from noting that identifying $\boldsymbol{\lambda}$ with $\boldsymbol{\beta}$, and identifying $\boldsymbol{\alpha}$ with $\boldsymbol{\theta}$ for $w_1 > w_3$, and $\boldsymbol{\alpha}''$ with $\boldsymbol{\theta}$ for $w_3 > w_1$, the relaxed problem corresponding to (16) yields the same rates as the original problem corresponding to (14).

APPENDIX J
PROOF OF THEOREM 3

A) Original Problem With $w_2 > w_3 > w_1$: To prove this theorem, we will first assume that there exists a solution of the KKT system with $R_k > 0$, for $k = 1, 2, 3$. Then, we will show that this assumption yields an objective that is monotonically decreasing in R_1, which leads us to conclude that the optimal R_1 is equal to zero.

To analyze the KKT conditions for $w_2 > w_3 > w_1$, we invoke this ordering in (18a), which yields

$$\beta_2 + \beta_5 + \beta_6 > \beta_9 \geq 0 \qquad \text{and} \qquad \beta_9 + \beta_{10} > \beta_1 \geq 0 \quad (88)$$

from which we have

$$\beta_5 + \beta_{10} > \beta_1 \geq 0. \quad (89)$$

As argued in Appendix F, if $\beta_1 > 0$ then $\beta_5 = \beta_{10} = 0$, which contradicts the second inequality in (89). Similarly, if $\beta_9 > 0$ then $\beta_2 = \beta_5 = \beta_6 = 0$, which contradicts the first inequality in (88). Hence, we conclude that $\beta_1 = \beta_9 = 0$. Since $\beta_9 = 0$, we have from the second inequality in (88) that

$$\beta_{10} > 0. \quad (90)$$

Using the first inequality in (88) in (17b) with $i = 2$ and $\ell = 1$ yields $\beta_3 + \beta_7 > 0$. Using (90) in (17b) with $i = 1$ and $\ell = 1$ yields $\beta_6 + \beta_7 + \beta_{11} > 0$. For the number of linearly independent equations not to exceed the number of unknowns in the KKT system, we will show that $\beta_5, \beta_7 > 0$ and $\beta_6 = 0$. To do so, we will consider all possible assumptions for β_5, β_6, and β_7.

1) $[\beta_5 + \beta_6 + \beta_7 = 0]$, $[\beta_5 > 0, \beta_6 > 0, \beta_7 > 0]$, $[\beta_7 = 0, \beta_5 > 0, \beta_6 > 0]$ and $[\beta_6 + \beta_7 = 0, \beta_5 > 0]$: These assumptions can be eliminated by using an argument similar to the one in Appendix F-A to show that they result in a number of independent linear equations that exceeds the number of unknowns;
2) $[\beta_5 + \beta_6 = 0, \beta_7 > 0]$: For this assumption, the first inequality in (88) yields $\beta_2 > 0$. If $R_3 < \frac{1}{2}\log\left(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\right)$, which is possible because we have no conditions on β_3, the fact that $\beta_7 > 0$ implies that $R_0 + R_1 + R_2 > g_{012}(\boldsymbol{\theta})$, which violates the inequality constraint in (17d) with $i = 2$. Now, if $R_3 = \frac{1}{2}\log\left(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\right)$, we will have a number of independent equations that exceeds the number of unknowns. Hence, this assumption can be eliminated;
3) $[\beta_5 + \beta_7 = 0, \beta_6 > 0]$: This assumption yields $\beta_3 > 0$. If $R_3 < \frac{1}{2}\log\left(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\right)$, which is possible because under this assumption $\beta_7 = 0$, the fact that $\beta_3 > 0$ implies that $R_0 + R_1 + R_2 > f_{012}(\boldsymbol{\theta})$, which violatesthe inequality constraint in (17c) with $i = 2$. Now, if $R_3 = \frac{1}{2}\log\left(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\right)$, we will have a number of independent linear equations that exceeds the number of unknowns. Hence, this assumption can be eliminated;
4) $[\beta_5 = 0, \beta_6 > 0, \beta_7 > 0]$: For this assumption, $R_3 = \frac{1}{2}\log\left(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\right)$ and $R_0 + R_1 + R_2 = g_{012}(\boldsymbol{\theta})$. If $R_1 < \frac{1}{2}\log\left(\frac{N_1^1+\theta_1^1 P_1}{N_1^1}\right)$, we have $R_0 + R_2 > g_{02}(\boldsymbol{\theta})$, which violates the inequality constraint in (17d) with $i = 1$, and if $R_1 = \frac{1}{2}\log\left(\frac{N_1^1+\theta_1^1 P_1}{N_1^1}\right)$, the number of independent linear equations exceeds the number of unknowns. Hence, this assumption can be eliminated;
5) $[\beta_6 = 0, \beta_5 > 0, \beta_7 > 0]$: For this assumption, we have

$$R_0 + R_2 = g_{02}(\boldsymbol{\theta}) \quad (91)$$
$$R_0 + R_1 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta}). \quad (92)$$

Counting the number of linearly independent equations in this case, it can be seen that it is equal to the number of unknowns. Now, the partition-rate vector $(\boldsymbol{\theta}, \boldsymbol{R})$ generated by (14) for this ordering must satisfy $\Phi_{h_3}(\boldsymbol{\theta},\boldsymbol{R}) \geq 0$. Using the fact that $\beta_{10} > 0$, cf. (90), we can write $R_1 \leq \frac{1}{2}\log\left(\frac{N_1^1+\theta_1^1 P_1}{N_1^1}\right)$. Subtracting (91) from (92) yields $R_1 + R_3 = g_{0123}(\boldsymbol{\theta}) - g_{02}(\boldsymbol{\theta})$. Using this equality, the objective in (14) can be expressed as

$$w_2 R_2 + w_3 R_3 + w_1 R_1 = w_2\big(g_{02}(\boldsymbol{\theta}) - R_0\big) + w_3\big(g_{0123}(\boldsymbol{\theta}) - g_{02}(\boldsymbol{\theta})\big) - (w_3 - w_1)R_1.$$

Now, for any $\boldsymbol{\theta} \in \mathcal{S}$ and $R_0 \in [0, R_{0,\max}]$, because $w_3 > w_1$, it can be seen that the objective is monotonically decreasing in R_1. In particular

$$w_2 R_2 + w_3 R_3 + w_1 R_1 \leq w_2\big(g_{02}(\boldsymbol{\theta}) - R_0\big) + w_3\big(g_{0123}(\boldsymbol{\theta}) - g_{02}\big)$$

where equality holds if and only if $R_1 = 0$, which establishes the first statement of the theorem.

B) Relaxed Problem With $w_2 > w_3 > w_1$: In analogous manner to the analysis of the original problem, we begin by assuming that $R_k > 0$, $k = 1, 2, 3$. In this case, the weight ordering $w_2 > w_3 > w_1$ yields

$$\lambda_2+\lambda_5+\lambda_6 > \lambda_9 \geq 0 \qquad \text{and} \qquad \lambda_9+\lambda_{10} > \lambda_1 \geq 0. \tag{93}$$

Using the second inequality and invoking Lemma 4 in Appendix C implies that $\lambda_9, \lambda_{10}, \lambda_{11} > 0$. From the first inequality and Lemmas 2 and 3 it can be seen that either $\lambda_1, \lambda_2, \lambda_3 > 0$, or $\lambda_5, \lambda_7 > 0$. In either case, it can be verified that the number of linearly independent equations exceeds the number of unknowns. Hence, we conclude that at least one of $\{R_k\}_{k=1}^3$ must be zero. However, since (14) is strictly feasible, it is straightforward to see that the current weight ordering implies that if R_2 or R_3 are zero, then R_1 must be zero, which proves the second statement of the theorem.

Appendix K
Proof of Corollary 2

From Theorem 3, we have that for $w_2 > w_3 > w_1$, the optimal partition-rate vector, $(\boldsymbol{\theta}, \boldsymbol{R})$, must have $R_1 = 0$. Considering the optimization problems corresponding to the ones in (14) and (16) but with $R_1 = 0$, it can be seen that for these problems the constraints $\Phi_{f_1}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$, $\Phi_{g_2}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$ and $\Phi_{h_3}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$ are redundant. Using this observation, the KKT systems corresponding to $R_1 = 0$ can be obtained from (17) and (18) by removing the constraints $\Phi_{f_1}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$, $\Phi_{g_2}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$ and $\Phi_{h_3}(\boldsymbol{\theta}, \boldsymbol{R}) \geq 0$ and their corresponding Lagrange multipliers, namely, β_i and λ_i, $i = 1, 6, 11$.

For the original problem corresponding to (14), the KKT stationarity conditions yield

$$\frac{\beta_2+\beta_3-\beta_4}{N_2^2+(\theta_2^1+\theta_2^2)P_2} = \frac{\beta_0+\beta_2+\beta_3}{N_2^3+(\theta_2^1+\theta_2^2)P_2} + \frac{\beta_8}{N_2^1+(\theta_2^1+\theta_2^2)P_2} \tag{94}$$

$$\frac{\beta_2+\beta_3+\beta_5+\beta_7}{N_2^2+\theta_2^1 P_2} = \frac{\beta_3+\beta_7}{N_2^1+\theta_2^1 P_2} \tag{95}$$

$$\frac{\beta_{10}-\beta_4}{N_1^2+(\theta_1^1+\theta_1^2)P_1} = \frac{\beta_0}{N_1^1+(\theta_1^1+\theta_1^2)P_1} + \frac{\beta_8+\beta_9+\beta_{10}}{N_1^3+(\theta_1^1+\theta_1^2)P_1} \tag{96}$$

$$\frac{\beta_5+\beta_7+\beta_{10}}{N_1^2+\theta_1^1 P_1} = \frac{\beta_7}{N_1^1+\theta_1^1 P_1}. \tag{97}$$

Now, the condition that $w_3 > w_2$ implies that

$$\beta_2+\beta_5 > 0. \tag{98}$$

We will use contradiction to show that β_5 is not equal to zero. To do so, assume that $\beta_5 = 0$. Using this in (97) yields that either $\beta_7 = \beta_{10} = 0$, or $\beta_7 > 0$ and $\beta_{10} > 0$. We will show that both cases yield a contradiction, which shows that $\beta_5 > 0$.

First, consider the case of $\beta_7 = \beta_{10} = 0$. In this case, (96) yields $\beta_0 = \beta_4 = \beta_8 = \beta_9 = 0$. Using this in (94) yields $\beta_2 + \beta_3 = 0$. However, this contradicts (98) because of the assumption that $\beta_5 = 0$. Hence, we conclude that, if $\beta_5 = 0$, we must have $\beta_7 > 0$ and $\beta_{10} > 0$.

Next, we consider the case of $\beta_7 > 0$ and $\beta_{10} > 0$ with the current assumption of $\beta_5 = 0$. The fact that $\beta_2 > 0$ implies that $R_0 + R_2 = f_{012}(\boldsymbol{\theta})$ and that

$$R_3 \leq \frac{1}{2}\log\Big(\frac{N_2^1+\theta_2^1 P_2}{N_2^1}\Big). \tag{99}$$

Now, because $\beta_7 > 0$, we have

$$R_0 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta}). \tag{100}$$

Substituting from (99) into (100) yields $R_0 + R_2 \geq g_{012}(\boldsymbol{\theta}) > g_{02}(\boldsymbol{\theta})$, for any $\theta_1^1 > 0$. Hence, we conclude that when $\beta_5 = 0$, β_7 and β_{10} cannot be greater than zero. Combining this result with the result obtained for the case of $\beta_5 = \beta_7 = \beta_{10} = 0$ yields the desired contradiction.

Using $\beta_5 > 0$ in (97) yields $\beta_7 > 0$, which implies that

$$R_0 + R_2 = g_{02}(\boldsymbol{\theta}) \qquad \text{and} \qquad R_0 + R_2 + R_3 = g_{0123}(\boldsymbol{\theta}).$$

We now consider the KKT system of the relaxed problem corresponding to (16). In this problem, the condition that $w_2 > w_3$ implies that $\lambda_5 > 0$, which, using Lemma 2 in Appendix C, implies that $\lambda_7 > 0$.

The statement of the corollary follows from noting that identifying $\boldsymbol{\lambda}$ with $\boldsymbol{\beta}$, $\boldsymbol{\alpha}'$ with $\boldsymbol{\theta}$ yields the same rates as the original problem corresponding to (14).

Noting that for any $\boldsymbol{\theta}$, with $\theta_1^1 > 0$, $g_{0123}(\boldsymbol{\theta}) > g_{023}(\boldsymbol{\theta})$, it is seen that the partition-rate vector generated by (14) for $w_2 > w_3 > w_1$ does not belong to the feasible set of (15), and hence does not belong to the intersection of the two regions, which is the SPC region.

Appendix L
Proof of the Information-Theoretic Bounds in Section VI

Given some small positive reals ϵ_j, $j = 1, 2, 3$, for every achievable rate, there is a sufficiently large n such that $P_{e_j}^n < \epsilon_j$. It follows from Fano's inequality that

$$H(M_0, M_1|Y_1, Y_2) \leq n\epsilon_1 \tag{101a}$$

$$H(M_0, M_2|Z_1, Z_2) \leq n\epsilon_2 \tag{101b}$$

$$H(M_0, M_3|W_1, W_2) \leq n\epsilon_3. \tag{101c}$$

In order to obtain (22a), we have from Fano's inequality that

$$\begin{aligned} nR_0 = H(M_0) &\leq I(M_0; Y_1, Y_2) + n\epsilon_1 \\ &= I(M_0; Y_2) + I(M_0; Y_1|Y_2) + n\epsilon_1 \\ &\leq I(M_0; Y_2) + I(M_0, Y_2, Z_2; Y_1) + n\epsilon_1 \\ &\leq I(M_0, M_1, Z_1, W_1; Y_2) \\ &\qquad + I(M_0, M_3, Z_2, Y_2; Y_1) + n\epsilon_1 \\ &= I(\mathcal{U}_1^3; Y_1) + I(\mathcal{U}_2^3; Y_2) + n\epsilon_1 \end{aligned} \tag{102}$$

where $\mathcal{U}_1^3$ and $\mathcal{U}_2^3$ are defined in (21). Due to the symmetry between receivers 1 and 3, (22e) can be proved in a similar manner.

We now show how to obtain the bound in (22b). Using Fano's inequality

$$\begin{aligned} n(R_0+R_1) &\le I(M_0,M_1;Y_1,Y_2)+n\epsilon_1\\ &= I(M_0,M_1;Y_2)+I(M_0,M_1;Y_1|Y_2)+n\epsilon_1\\ &\le I(M_0,M_1;Y_2)+I(M_0,M_1,Y_2;Y_1)+n\epsilon_1\\ &\le I(M_0,M_1,W_1,Z_1;Y_2)+I(X_1;Y_1)+n\epsilon_1\\ &\le I(\mathcal{U}_2^3;Y_2)+I(X_1;Y_1)+n\epsilon_1. \end{aligned} \tag{103}$$

Invoking the symmetry between receivers 1 and 3, one can prove (22f) in a similar fashion.

In order to obtain the bound in (22c), we write

$$\begin{aligned} n(R_0+R_1+R_2) &\le I(M_0,M_1;Y_1,Y_2)+I(M_2;Z_1,Z_2)\\ &\quad +n\epsilon_1+n\epsilon_2\\ &\le I(M_0,M_1;Y_1,Y_2)\\ &\quad +I(M_2;Z_1,Z_2|M_0,M_1)\\ &\quad +n\epsilon_1+n\epsilon_2\\ &\le I(M_0,M_1;Y_2)+I(M_0,M_1;Y_1|Y_2)\\ &\quad +I(M_2;Z_1,W_1|M_0,M_1)\\ &\quad +I(M_2;Z_2|M_0,M_1,Z_1,W_1)\\ &\quad +n\epsilon_1+n\epsilon_2. \end{aligned}$$

By adding and subtracting the term $I(Z_1,W_1;Y_2|M_0,M_1)$ in the above expression, we obtain

$$\begin{aligned} n(R_0+R_1+R_2) &\le I(M_0,M_1,Z_1,W_1;Y_2)\\ &\quad +I(M_0,M_1;Y_1|Y_2)\\ &\quad +I(M_2;Z_1,W_1|M_0,M_1)\\ &\quad +I(M_2;Z_2|M_0,M_1,Z_1,W_1)\\ &\quad -I(Z_1,W_1;Y_2|M_0,M_1)\\ &\quad +n\epsilon_1+n\epsilon_2. \end{aligned} \tag{104}$$

Further bounding of the right-hand side yields

$$\begin{aligned} n(R_0+R_1+R_2) &\le I(\mathcal{U}_2^3;Y_2)+I(\mathcal{U}_2^2;Z_2|\mathcal{U}_2^3)\\ &\quad +I(M_0,M_1,Y_2;Y_1)\\ &\quad +I(M_2;Z_1,W_1|M_0,M_1,Y_2)\\ &\quad +n\epsilon_1+n\epsilon_2\\ &\le I(\mathcal{U}_2^3;Y_2)+I(\mathcal{U}_2^2;Z_2|\mathcal{U}_2^3)\\ &\quad +I(M_0,M_1,Y_2;Y_1)\\ &\quad +I(M_2;Y_1|M_0,M_1,Y_2)\\ &\quad +n\epsilon_1+n\epsilon_2\\ &\le I(\mathcal{U}_2^3;Y_2)+I(\mathcal{U}_2^2;Z_2|\mathcal{U}_2^3)\\ &\quad +I(X_1;Y_1)+n\epsilon_1+n\epsilon_2. \end{aligned}$$

In a similar manner, one can prove the bound in (22g).

In order to obtain a bound on the sum rate in (22d), we have

$$\begin{aligned} &n(R_0+R_1+R_2+R_3)\\ &\le I(M_0,M_1;Y_1,Y_2)+I(M_2;Z_1,Z_2)\\ &\quad +I(M_3;W_1,W_2)+n\epsilon_1+n\epsilon_2+n\epsilon_3\\ &\le I(M_0,M_1;Y_2)+I(M_0,M_1;Y_1|Y_2)+I(M_2;W_1,Z_1)\\ &\quad +I(M_2;Z_2|W_1,Z_1)+I(M_3;W_1,Z_1)\\ &\quad +I(M_3;W_2|W_1,Z_1)+n\epsilon_1+n\epsilon_2+n\epsilon_3\\ &\le I(M_0,M_1;Y_2)+I(M_0,M_1;Y_1|Y_2)\\ &\quad +I(M_2;W_1,Z_1|M_0,M_1)\\ &\quad +I(M_2;Z_2|M_0,M_1,W_1,Z_1)\\ &\quad +I(M_3;W_1,Z_1|M_0,M_1,M_2)\\ &\quad +I(M_3;W_2|W_1,Z_1,M_0,M_1,M_2)\\ &\quad +n\epsilon_1+n\epsilon_2+n\epsilon_3\\ &= I(M_0,M_1;Y_2)+I(M_0,M_1;Y_1|Y_2)\\ &\quad +I(M_2,M_3;Z_1,W_1|M_0,M_1)\\ &\quad +I(M_2;Z_2|M_0,M_1,W_1,Z_1)\\ &\quad +I(M_3;W_2|W_1,Z_1,M_0,M_1,M_2)\\ &\quad +I(W_1,Z_1;Y_2|M_0,M_1)-I(W_1,Z_1;Y_2|M_0,M_1)\\ &\quad +n\epsilon_1+n\epsilon_2+n\epsilon_3\\ &\le I(M_0,M_1,W_1,Z_1;Y_2)+I(M_0,M_1;Y_1|Y_2)\\ &\quad +I(M_2,M_3;Z_1,W_1|M_0,M_1,Y_2)\\ &\quad +I(M_2;Z_2|M_0,M_1,W_1,Z_1)\\ &\quad +I(M_3;W_2|W_1,Z_1,M_0,M_1,M_2)+n\epsilon_1\\ &\quad +n\epsilon_2+n\epsilon_3\\ &\le I(\mathcal{U}_2^3;Y_2)+I(\mathcal{U}_2^2;Z_2|\mathcal{U}_2^3)+I(X_2;W_2|\mathcal{U}_2^2)\\ &\quad +I(X_1;Y_1)+n\epsilon_1+n\epsilon_2+n\epsilon_3. \end{aligned}$$

One can use the same technique to obtain the bound in (22h).

Appendix M
Proof of the Inequalities in (23) and (24)

Inequalities (23a) for RSPC_1 and (24a) for RSPC_2 are identical, and to prove them, observe that

$$\begin{aligned} nR_0 &\le I(M_0;Z_1,Z_2)+n\epsilon_2\\ &= I(M_0;Z_1)+I(M_0;Z_2|Z_1)+n\epsilon_2\\ &\le I(M_0,Z_2;Z_1)+I(M_0,Z_1;Z_2)+n\epsilon_2\\ &= I(\mathcal{V}_1^3;Z_1)+I(\mathcal{V}_2^3;Z_2)+n\epsilon_2. \end{aligned}$$

To prove (23b), we use Fano's inequality to write

$$\begin{aligned} n(R_0+R_2) &\le I(M_0,M_2;Z_1,Z_2)+n\epsilon_2\\ &= I(M_0,M_2;Z_1|Z_2)+I(M_0,M_2;Z_2)+n\epsilon_2\\ &\le I(M_0,M_2,Z_2;Z_1)+I(M_0,M_2,M_1,Y_1;Z_2)\\ &\qquad +n\epsilon_2\\ &= I(\mathcal{V}_1^2;Z_1)+I(\mathcal{V}_2^2;Z_2)+n\epsilon_2. \end{aligned}$$

Similarly, to prove (24b), we have

$$\begin{aligned} n(R_0+R_2) &\le I(M_0,M_2;Z_1,Z_2)+n\epsilon_2 \\ &= I(M_0,M_2;Z_1)+I(M_0,M_2;Z_2|Z_1)+n\epsilon_2 \\ &\le I(M_0,M_2,M_3,W_2;Z_1)+I(M_0,M_2,Z_1;Z_2) \\ &\qquad +n\epsilon_2 \\ &= I(\mathcal{X}_1^2;Z_1)+I(\mathcal{X}_2^2;Z_2)+n\epsilon_2. \end{aligned}$$

We now prove (23c). Using Fano's inequality

$$\begin{aligned} n(R_0+R_1+R_2) &\le I(M_0,M_2;Z_1,Z_2)+I(M_1;Y_1,Y_2) \\ &\quad +n\epsilon_1+n\epsilon_2 \\ &\le I(M_0,M_2;Z_1,Z_2) \\ &\quad +I(M_1;Y_1,Y_2|M_0,M_2) \\ &\quad +n\epsilon_1+n\epsilon_2 \qquad (105) \\ &\le I(M_0,M_2;Z_1,Z_2) \\ &\quad +I(M_1;Y_1,Z_2|M_0,M_2) \\ &\quad +n\epsilon_1+n\epsilon_2 \qquad (106) \\ &= I(M_0,M_2;Z_1|Z_2)+I(M_0,M_2;Z_2) \\ &\quad +I(M_1;Z_2|M_0,M_2) \\ &\quad +I(M_1;Y_1|Z_2,M_0,M_2) \\ &\quad +n\epsilon_1+n\epsilon_2 \\ &\le I(M_0,M_2,Z_2;Z_1)+I(M_0,M_2,M_1;Z_2) \\ &\quad +I(M_1;Y_1|Z_2,M_0,M_2)+n\epsilon_1+n\epsilon_2 \\ &\le I(M_0,M_2,Z_2;Z_1) \\ &\quad +I(M_0,M_2,M_1,Y_1;Z_2) \\ &\quad +I(M_1;Y_1|Z_2,M_0,M_2)+n\epsilon_1+n\epsilon_2 \\ &\le I(\mathcal{V}_1^2;Z_1)+I(\mathcal{V}_2^2;Z_2)+I(X_1;Y_1|\mathcal{V}_1^2) \\ &\quad +n\epsilon_1+n\epsilon_2 \qquad (107) \end{aligned}$$

where in (105) we used the independence of (M_0,M_2) and M_1, and in (106) we used the fact that Z_2 is less degraded than Y_2. Using the symmetry between the received signals (W_2,W_1) and (Y_1,Y_2), an analogous argument can be used to prove (24c).

In order to prove (23d), we use Fano's inequality to write

$$\begin{aligned} n(R_0+R_1+R_2+R_3) &\le I(M_0,M_2;Z_1,Z_2) \\ +I(M_1;Y_1,Y_2)+I(M_3;W_1,W_2)&+n\epsilon_1+n\epsilon_2+n\epsilon_3. \end{aligned} \qquad (108)$$

From (108)

$$\begin{aligned} &n(R_0+R_1+R_2+R_3) \qquad (109) \\ &\quad\le I(M_0,M_2;Z_1,Z_2)+I(M_1;Y_1,Z_2)+I(M_3;Y_1,W_2) \\ &\qquad +n\epsilon_1+n\epsilon_2+n\epsilon_3 \qquad (110) \\ &\quad\le I(M_0,M_2;Z_1,Z_2)+I(M_1;Y_1,Z_2|M_0,M_2) \\ &\qquad +I(M_3;Y_1,W_2|M_0,M_2,M_1) \\ &\qquad +n\epsilon_1+n\epsilon_2+n\epsilon_3 \qquad (111) \\ &\quad= I(M_0,M_2;Z_1,Z_2)+I(M_1;Y_1|M_0,M_2) \\ &\qquad +I(M_1;Z_2|M_0,M_2,Y_1)+I(M_3;Y_1|M_0,M_2,M_1) \\ &\qquad +I(M_3;W_2|M_0,M_2,M_1,Y_1)+n\epsilon_1+n\epsilon_2+n\epsilon_3 \\ &\quad= I(M_0,M_2;Z_1,Z_2)+I(M_1,M_3;Y_1|M_0,M_2) \\ &\qquad +I(M_1;Z_2|M_0,M_2,Y_1)+I(Z_2;Y_1|M_0,M_2) \\ &\qquad +I(M_3;W_2|M_0,M_2,M_1,Y_1) \\ &\qquad -I(Z_2;Y_1|M_0,M_2)+n\epsilon_1+n\epsilon_2+n\epsilon_3 \qquad (112) \\ &\quad\le I(M_0,M_2;Z_2)+I(M_0,M_2;Z_1|Z_2) \\ &\qquad +I(M_1,M_3;Y_1|M_0,M_2,Z_2) \\ &\qquad +I(M_3;W_2|M_0,M_2,M_1,Y_1)+I(Z_2;Y_1|M_0,M_2) \\ &\qquad +I(M_1;Z_2|M_0,M_2,Y_1)+n\epsilon_1+n\epsilon_2+n\epsilon_3 \\ &\quad= I(M_0,M_2,Y_1;Z_2)+I(M_0,M_2;Z_1|Z_2) \\ &\qquad +I(M_1,M_3;Y_1|M_0,M_2,Z_2) \\ &\qquad +I(M_3;W_2|M_0,M_2,M_1,Y_1) \\ &\qquad +I(M_1;Z_2|M_0,M_2,Y_1) \\ &\qquad +n\epsilon_1+n\epsilon_2+n\epsilon_3 \\ &\quad\le I(M_0,M_2,Y_1,M_1;Z_2)+I(M_0,M_2,Z_2;Z_1) \\ &\qquad +I(M_1,M_3;Y_1|M_0,M_2,Z_2) \\ &\qquad +I(M_3;W_2|M_0,M_2,M_1,Y_1)+n\epsilon_1+n\epsilon_2+n\epsilon_3 \\ &\quad= I(\mathcal{V}_2^2;Z_2)+I(\mathcal{V}_1^2;Z_1)+I(X_1;Y_1|\mathcal{V}_1^2) \\ &\qquad +I(X_2;W_2|\mathcal{V}_2^2)+n\epsilon_1+n\epsilon_2+n\epsilon_3 \end{aligned}$$

where in (110) we have used the fact that Y_1 is less degraded than W_1, and Z_2 is less degraded than Y_2. In (111) we used the observation that M_0 and M_2 are independent of M_1 and M_3, and that M_1 and M_3 are independent of each other. The term $I(Z_2;Y_1|M_0,M_2)$ is added and subtracted in (112) in order to introduce Z_2 in the conditioning of the second term in (112). Using the symmetry between the received signals (W_2,W_1) and (Y_1,Y_2), an analogous argument can be used to prove (24d).

Appendix N
Application to the Gaussian Channel—RSPC_1

In this section, we will take the first step towards showing that RSPC_1 contains the set of all achievable rate vectors. We will show that for every achievable rate vector there exist vectors $\boldsymbol{\alpha},\boldsymbol{\alpha}',\boldsymbol{\alpha}''$ such that the inequalities in (16b)–(16f) are satisfied. The argument will be based on invoking the inequalities in (22) and (23) in the case in which each subchannel is Gaussian. (The corresponding argument for RSPC_2 is almost identical, and will be briefly discussed in Appendix O.) We begin by observing that [2]

$$H(Y_1) \le \frac{n}{2}\log\big(2\pi e(P_1+N_1^1)\big) \qquad (113a)$$

$$H(Z_1) \le \frac{n}{2}\log\big((2\pi e(P_1+N_1^2)\big) \qquad (113b)$$

$$H(W_2) \le \frac{n}{2}\log\big((2\pi e(P_2+N_2^1)\big) \qquad (113c)$$

$$H(Z_2) \le \frac{n}{2}\log\big((2\pi e(P_2+N_2^2)\big). \qquad (113d)$$

In the following sections, we will specify the vectors $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$ and $\boldsymbol{\alpha}''$ and we will employ those vectors and the entropy power inequality to provide the desired bounds.

A) Specifying the Vectors $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$ and $\boldsymbol{\alpha}''$:

1) Specifying $\boldsymbol{\alpha}$: Since conditioning reduces entropy, we conclude that there exist two nonnegative reals α_1^1 and α_1^2 satisfying $\alpha_1^1 + \alpha_1^2 \leq 1$ such that

$$H(Y_1|\mathcal{U}_1^3) = \frac{n}{2}\log(2\pi e((\alpha_1^1 + \alpha_1^2)P_1 + N_1^1)) \quad (114)$$
$$H(Y_1|\mathcal{U}_1^2) = \frac{n}{2}\log(2\pi e(\alpha_1^1 P_1 + N_1^1)). \quad (115)$$

Similarly, there exist α_2^1 and α_2^2 satisfying $\alpha_2^1 + \alpha_2^2 \leq 1$ such that

$$H(Z_2|\mathcal{U}_2^3) = \frac{n}{2}\log(2\pi e((\alpha_2^1 + \alpha_2^2)P_2 + N_2^2)) \quad (116)$$
$$H(Z_2|\mathcal{U}_2^2) = \frac{n}{2}\log(2\pi e(\alpha_2^1 P_2 + N_2^2)). \quad (117)$$

In (115), we have used the fact that $\mathcal{U}_1^2$ contains more information about Y_1 than $\mathcal{U}_1^3$, and that $\mathcal{U}_2^2$ contains more information about Z_2 than $\mathcal{U}_2^3$. This fact is immediately apparent from the definitions in (21).

2) Specifying $\boldsymbol{\alpha}'$: Because conditioning reduces entropy, there also exist nonnegative reals $\alpha_1'^1$, $\alpha_1'^2$, $\alpha_2'^1$ and $\alpha_2'^2$ such that $\alpha_i'^1 + \alpha_i'^2 \leq 1$, $i = 1, 2$, and

$$\begin{aligned} H(Y_1|M_0, Z_2) &= H(Y_1|\mathcal{V}_1^3) \\ &= \frac{n}{2}\log\Big(2\pi e\big((\alpha_1'^1 + \alpha_1'^2)P_1 + N_1^1\big)\Big) \end{aligned} \quad (118)$$
$$\begin{aligned} H(W_2|M_0, Z_1) &= H(W_2|\mathcal{V}_2^3) \\ &= \frac{n}{2}\log\Big(2\pi e\big((\alpha_2'^1 + \alpha_2'^2)P_2 + N_2^1\big)\Big) \end{aligned} \quad (119)$$
$$\begin{aligned} H(Y_1|M_0, M_2, Z_2) &= H(Y_1|\mathcal{V}_1^2) \\ &= \frac{n}{2}\log\big(2\pi e(\alpha_1'^1 P_1 + N_1^1)\big) \end{aligned} \quad (120)$$
$$\begin{aligned} H(W_2|M_0, M_2, Y_1, M_1) &= H(W_2|\mathcal{V}_2^2) \\ &= \frac{n}{2}\log\big(2\pi e(\alpha_2'^1 P_2 + N_2^1)\big) \end{aligned} \quad (121)$$

where (121) follows from the fact that Y_1 is a less degraded version of Z_1.

3) Specifying $\boldsymbol{\alpha}''$: Using, once again, the fact that conditioning reduces entropy, one can find nonnegative reals $\alpha_1''^1$, $\alpha_1''^2$, $\alpha_2''^1$, $\alpha_2''^2$ such that $\alpha_i''^1 + \alpha_i''^2 \leq 1$, $i = 1, 2$, and

$$H(Z_1|\mathcal{U}_1^3) = \frac{n}{2}\log(2\pi e((\alpha_1''^1 + \alpha_1''^2)P_1 + N_1^2)) \quad (122)$$
$$H(Z_1|\mathcal{U}_1^2) = \frac{n}{2}\log(2\pi e(\alpha_1''^1 P_1 + N_1^2)) \quad (123)$$
$$H(W_2|\mathcal{U}_2^3) = \frac{n}{2}\log(2\pi e((\alpha_2''^1 + \alpha_2''^2)P_2 + N_2^1)) \quad (124)$$
$$H(W_2|\mathcal{U}_2^2) = \frac{n}{2}\log(2\pi e(\alpha_2''^1 P_2 + N_2^1)). \quad (125)$$

B) Applying the Entropy Power Inequality

1) Applying the Entropy Power Inequality With $\boldsymbol{\alpha}$: Using a technique similar to the one used in [5], the entropy power inequality can be shown to yield

$$H(Z_1|\mathcal{U}_1^3) \geq \frac{n}{2}\log(2\pi e((\alpha_1^1 + \alpha_1^2)P_1 + N_1^2)) \quad (126)$$
$$H(Z_1|\mathcal{U}_1^2) \geq \frac{n}{2}\log(2\pi e(\alpha_1^1 P_1 + N_1^2)) \quad (127)$$
$$H(W_1|\mathcal{U}_1^3) \geq \frac{n}{2}\log(2\pi e((\alpha_1^1 + \alpha_1^2)P_1 + N_1^3)) \quad (128)$$
$$H(W_1|\mathcal{U}_1^2) \geq \frac{n}{2}\log(2\pi e(\alpha_1^1 P_1 + N_1^3)). \quad (129)$$

Similarly

$$H(W_2|\mathcal{U}_2^3) \leq \frac{n}{2}\log(2\pi e((\alpha_2^1 + \alpha_2^2)P_2 + N_2^1)) \quad (130)$$
$$H(W_2|\mathcal{U}_2^2) \leq \frac{n}{2}\log(2\pi e(\alpha_2^1 P_2 + N_2^1)) \quad (131)$$

where in (130) and (131), we have used the entropy power inequality in the reverse direction. Using the entropy power inequality on (116) and (117), we obtain

$$H(Y_2|\mathcal{U}_2^3) \geq \frac{n}{2}\log(2\pi e((\alpha_2^1 + \alpha_2^2)P_2 + N_2^3)) \quad (132)$$
$$H(Y_2|\mathcal{U}_2^2) \geq \frac{n}{2}\log(2\pi e(\alpha_2^1 P_2 + N_2^3)). \quad (133)$$

2) Applying the Entropy Power Inequality With $\boldsymbol{\alpha}'$:

$$\begin{aligned} H(Z_1|M_0, Z_2) &= H(Z_1|\mathcal{V}_1^3) \\ &\geq \frac{n}{2}\log\Big(2\pi e\big((\alpha_1'^1 + \alpha_1'^2)P_1 + N_1^2\big)\Big) \end{aligned} \quad (134)$$
$$\begin{aligned} H(Z_2|M_0, Z_1) &= H(Z_2|\mathcal{V}_2^3) \\ &\geq \frac{n}{2}\log\Big(2\pi e\big((\alpha_2'^1 + \alpha_2'^2)P_2 + N_2^2\big)\Big) \end{aligned} \quad (135)$$
$$\begin{aligned} H(Z_1|M_0, M_2, Z_2) &= H(Z_1|\mathcal{V}_1^2) \\ &\geq \frac{n}{2}\log\big(2\pi e(\alpha_1'^1 P_1 + N_1^2)\big) \end{aligned} \quad (136)$$
$$\begin{aligned} H(Z_2|M_0, M_2, Y_1, M_1) &= H(Z_2|\mathcal{V}_2^2) \\ &\geq \frac{n}{2}\log\big(2\pi e(\alpha_2'^1 P_2 + N_2^2)\big). \end{aligned} \quad (137)$$

3) Applying the Entropy Power Inequality With $\boldsymbol{\alpha}''$:

$$H(Y_1|\mathcal{U}_1^3) \leq \frac{n}{2}\log(2\pi e((\alpha_1''^1 + \alpha_1''^2)P_1 + N_1^1)) \quad (138)$$
$$H(W_1|\mathcal{U}_1^3) \geq \frac{n}{2}\log(2\pi e((\alpha_1''^1 + \alpha_1''^2)P_1 + N_1^3)) \quad (139)$$
$$H(Y_1|\mathcal{U}_1^2) \leq \frac{n}{2}\log(2\pi e(\alpha_1''^1 P_1 + N_1^1)) \quad (140)$$
$$H(W_1|\mathcal{U}_1^2) \geq \frac{n}{2}\log(2\pi e(\alpha_1''^1 P_1 + N_1^3)) \quad (141)$$
$$H(Z_2|\mathcal{U}_2^3) \geq \frac{n}{2}\log(2\pi e((\alpha_2''^1 + \alpha_2''^2)P_2 + N_2^2)) \quad (142)$$
$$H(Y_2|\mathcal{U}_2^3) \geq \frac{n}{2}\log(2\pi e((\alpha_2''^1 + \alpha_2''^2)P_2 + N_2^3)) \quad (143)$$
$$H(Z_2|\mathcal{U}_2^2) \geq \frac{n}{2}\log(2\pi e(\alpha_2''^1 P_2 + N_2^2)) \quad (144)$$
$$H(Y_2|\mathcal{U}_2^2) \geq \frac{n}{2}\log(2\pi e(\alpha_2''^1 P_2 + N_2^3)). \quad (145)$$

Using (113)–(145), we now prove our target inequalities.

C) Proving the Converse of the Inequalities in (16b): For this set of inequalities, we will apply the inequalities in (113), (114)–(117) and (126)–(133) to (22a)–(22d).

1) Proving the Converse of the Inequality in (16b) Corresponding to $i = 0$: From (22a)

$$\begin{aligned} nR_0 &\le H(Y_1) - H(Y_1|\mathcal{U}_1^3) + H(Y_2) - H(Y_2|\mathcal{U}_1^3) + n\epsilon_1 \\ &\le \frac{n}{2}\log(2\pi e(P_1 + N_1^1)) \\ &\quad - \frac{n}{2}\log(2\pi e((\alpha_1^1 + \alpha_1^2)P_1 + N_1^1)) \\ &\quad + \frac{n}{2}\log(2\pi e(P_2 + N_2^3)) \\ &\quad - \frac{n}{2}\log(2\pi e((\alpha_2^1 + \alpha_2^2)P_2 + N_2^3)) + n\epsilon_1 \\ &= \frac{n}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1 + (\alpha_1^1 + \alpha_1^2)P_1}\Big) \\ &\quad + \frac{n}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^2 + \alpha_2^1)P_2}\Big) + n\epsilon_1. \end{aligned}$$

2) Proving the Converse of the Inequality in (16b) Corresponding to $i = 1$: From (22b)

$$\begin{aligned} n(R_0 + R_1) &\le H(Y_1) - H(Y_1|X_1) + H(Y_2) - H(Y_2|\mathcal{U}_2^3) \\ &\quad + n\epsilon_1 \\ &\le \frac{n}{2}\log(2\pi e(P_1 + N_1^1)) - \frac{n}{2}\log(2\pi e(N_1^1)) \\ &\quad + \frac{n}{2}\log(2\pi e(P_2 + N_2^3)) \\ &\quad - \frac{n}{2}\log(2\pi e((\alpha_2^1 + \alpha_2^2)P_2 + N_2^3)) + n\epsilon_1 \\ &= \frac{n}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) \\ &\quad + \frac{n}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^2 + \alpha_2^1)P_2}\Big) + n\epsilon_1. \end{aligned}$$

3) Proving the Converse of the Inequality in (16b) Corresponding to $i = 2$: Using (22c)

$$\begin{aligned} &n(R_0 + R_1 + R_2) \\ &\le I(\mathcal{U}_2^3; Y_2) + I(\mathcal{U}_2^2; Z_2|\mathcal{U}_2^3) + I(X_1; Y_1) + n\epsilon_1 + n\epsilon_2 \\ &\le H(Y_1) - H(Y_1|X_1) + H(Y_2) - H(Y_2|\mathcal{U}_2^3) \\ &\quad + H(Z_2|\mathcal{U}_2^3) - H(Z_2|\mathcal{U}_2^2) + n\epsilon_1 + n\epsilon_2 \\ &\le \frac{n}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{n}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^2 + \alpha_2^1)P_2}\Big) \\ &\quad + \frac{n}{2}\log(2\pi e((\alpha_2^1 + \alpha_2^2)P_2 + N_2^2)) \\ &\quad - \frac{n}{2}\log(2\pi e(\alpha_2^1 P_2 + N_2^2)) + n\epsilon_1 + n\epsilon_2 \\ &= \frac{n}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{n}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^2 + \alpha_2^1)P_2}\Big) \\ &\quad + \frac{n}{2}\log\Big(\frac{N_2^2 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^2 + \alpha_2^1 P_2}\Big) + n\epsilon_1 + n\epsilon_2. \end{aligned}$$

4) Proving the Converse of the Inequality in (16b) Corresponding to $i = 3$: To prove the converse of this inequality, we use (22d) to write

$$\begin{aligned} &n(R_0 + R_1 + R_2 + R_3) \\ &\le \frac{n}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{n}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^2 + \alpha_2^1)P_2}\Big) \\ &\quad + \frac{n}{2}\log\Big(\frac{N_2^2 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^2 + \alpha_2^1 P_2}\Big) + H(W_2|\mathcal{U}_2^2) \\ &\quad - H(W_2|X_2) + n\epsilon_1 + n\epsilon_2 + n\epsilon_3 \\ &\le \frac{n}{2}\log\Big(\frac{N_1^1 + P_1}{N_1^1}\Big) + \frac{n}{2}\log\Big(\frac{N_2^3 + P_2}{N_2^3 + (\alpha_2^2 + \alpha_2^1)P_2}\Big) \\ &\quad + \frac{n}{2}\log\Big(\frac{N_2^2 + (\alpha_2^1 + \alpha_2^2)P_2}{N_2^2 + \alpha_2^1 P_2}\Big) \\ &\quad + \frac{n}{2}\log(2\pi e(\alpha_2^1 P_2 + N_2^1)) - \frac{n}{2}\log(2\pi e N_2^1) \\ &\quad + n\epsilon_1 + n\epsilon_2 + n\epsilon_3 \end{aligned} \tag{146}$$

where in (146), we have used the upper bound on $H(W_2|\mathcal{U}_2^2)$ in (131) and the fact that $H(W_2|X_2) = \frac{n}{2}\log(2\pi e N_2^1)$.

D) Proving the Converse of the Inequalities in (16c): For this set of inequalities, we will apply the inequalities in (113), (118)–(119) and (134)–(137) to (23a)–(23d).

1) Proving the Converse of the Inequality in (16c) Corresponding to $i = 0$: Expressing (23a) in terms of the conditional entropy and using (134) and (135), it can be shown that

$$\begin{aligned} nR_0 \le \frac{n}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + (\alpha_1'^1 + \alpha_1'^2)P_1}\Big) \\ + \frac{n}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + (\alpha_2'^1 + \alpha_2'^2)P_2}\Big) + n\epsilon_2. \end{aligned}$$

2) Proving the Converse of the Inequality in (16c) Corresponding to $i = 1$: Expressing (23b) in terms of the conditional entropy and using (113), (136), and (137)

$$\begin{aligned} n(R_0 + R_2) \le \frac{n}{2}\log\Big(\frac{N_2^2 + P_2}{N_2^2 + \alpha_2'^1 P_2}\Big) + \frac{n}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \alpha_1'^1 P_1}\Big) \\ + n\epsilon_2. \end{aligned}$$

3) Proving the Converse of the Inequality in (16c) Corresponding to $i = 2$: Using (23c)

$$\begin{aligned} &n(R_0 + R_1 + R_2) \\ &\le H(Z_2) - H(Z_2|M_0, M_2, Y_1, M_1) + H(Z_1) \\ &\quad - H(Z_1|M_0, M_2, Z_2) + H(Y_1|M_0, M_2, Z_2) \\ &\quad - H(Y_1|M_1, M_3, M_0, M_2, Z_2) + n(\epsilon_1 + \epsilon_2) \\ &\le \frac{n}{2}\log\big((2\pi e(P_2 + N_2^2)\big) - \frac{n}{2}\log\big(2\pi e(\alpha_2'^1 P_2 + N_2^2\big) \\ &\quad + \frac{n}{2}\log\big((2\pi e(P_1 + N_1^2)\big) - \frac{n}{2}\log\big(2\pi e(\alpha_1'^1 P_1 + N_1^2)\big) \\ &\quad + \frac{n}{2}\log\big(2\pi e(\alpha_1'^1 P_1 + N_1^1)\big) - \frac{n}{2}\log(2\pi e N_1^1) \\ &\quad + n(\epsilon_1 + \epsilon_2) \\ &= \frac{n}{2}\log\Big(\frac{(N_2^2 + P_2}{N_2^2 + \alpha_2'^1 P_2}\Big) + \frac{n}{2}\log\Big(\frac{N_1^2 + P_1}{N_1^2 + \alpha_1'^1 P_1}\Big) \\ &\quad + \frac{n}{2}\log\Big(\frac{N_1^1 + \alpha_1'^1 P_1}{N_1^1}\Big) + n(\epsilon_1 + \epsilon_2). \end{aligned}$$

4) Proving the Converse of Inequality in (16c) Corresponding to $i = 3$: Using (23d)

$$\begin{aligned}
&n(R_0+R_1+R_2+R_3)\\
&\quad\le H(Z_2)-H(Z_2|M_0,M_2,M_1,Y_1)+H(Z_1)\\
&\qquad -H(Z_1|M_0,M_2,Z_2)+H(Y_1|M_0,M_2,Z_2)\\
&\qquad -H(Y_1|M_0,M_2,Z_2,M_1,M_3)\\
&\qquad +H(W_2|M_0,M_2,Y_1,M_1)\\
&\qquad -H(W_2|M_0,M_2,Y_1,M_1,M_3)+n(\epsilon_1+\epsilon_2+\epsilon_3)\\
&\quad\le \frac{n}{2}\log\big((2\pi e(P_2+N_2^2)\big)-\frac{n}{2}\log\big(2\pi e(\alpha_2'^1P_2+N_2^2\big)\\
&\qquad +\frac{n}{2}\log\big(2\pi e(P_1+N_1^2)\big)-\frac{n}{2}\log\big(2\pi e(\alpha_1'^1P_1+N_1^2)\big)\\
&\qquad +\frac{n}{2}\log\big(2\pi e(\alpha_1'^1P_1+N_1^1)\big)+\frac{n}{2}\log\big(2\pi e(\alpha_2'^1P_2+N_2^1)\big)\\
&\qquad -H(Y_1|M_0,M_2,Z_2,M_1,M_3)\\
&\qquad -H(W_2|M_0,M_2,Y_1,M_1,M_3)+n(\epsilon_1+\epsilon_2+\epsilon_3)\\
&\quad= \frac{n}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\alpha_2'^1P_2}\Big)+\frac{n}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+\alpha_1'^1P_1}\Big)\\
&\qquad +\frac{n}{2}\log\Big(\frac{N_1^1+\alpha_1'^1P_1}{N_1^1}\Big)+\frac{n}{2}\log\Big(\frac{N_2^1+\alpha_2'^1P_2}{N_2^1}\Big)\\
&\qquad +n(\epsilon_1+\epsilon_2+\epsilon_3)
\end{aligned}\tag{147}$$

where in (147), we have used the fact that

$$H(Y_1|M_0,M_2,Z_2,M_1,M_3)\ge\frac{n}{2}\log(2\pi eN_1^1)$$

and

$$H(W_2|M_0,M_2,Y_1,M_1,M_3)\ge\frac{n}{2}\log(2\pi eN_2^1).$$

E) Proving the Converse of the Inequalities in (16d): For this set of inequalities, we will apply (113), (122)–(125), (138)–(145) to (22e)–(22h).

1) Proving the Converse of the Inequality in (16d) Corresponding to $i = 0$:

$$\begin{aligned}
nR_0&\le H(W_1)-H(W_1|\mathcal{U}_1^3)+H(W_2)-H(W_2|\mathcal{U}_2^3)+n\epsilon_3\\
&\le\frac{n}{2}\log(2\pi e(P_1+N_1^3))\\
&\quad-\frac{n}{2}\log(2\pi e(N_1^3+(\alpha_1''^1+\alpha_1''^2)P_1))\\
&\quad+\frac{n}{2}\log(2\pi e(P_2+N_2^1))\\
&\quad-\frac{n}{2}\log(2\pi e((\alpha_2''^1+\alpha_2''^2)P_2+N_2^1))+n\epsilon_3\\
&=\frac{n}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\alpha_1''^2+\alpha_1''^1)P_1}\Big)\\
&\quad+\frac{n}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1+(\alpha_2''^2+\alpha_2''^1)P_2}\Big)+n\epsilon_3.
\end{aligned}$$

2) Proving the Converse of the Inequality in (16d) Corresponding to $i = 1$: Using (22f)

$$\begin{aligned}
n(R_0+R_3)&\le H(W_2)-H(W_2|X_2)+H(W_1)\\
&\quad-H(W_1|\mathcal{U}_1^3)+n\epsilon_3\\
&\le\frac{n}{2}\log(2\pi e(P_2+N_2^1))-\frac{n}{2}\log(2\pi eN_2^1)\\
&\quad+\frac{n}{2}\log(2\pi e(P_1+N_1^3))\\
&\quad-\frac{n}{2}\log(2\pi e((\alpha_1''^1+\alpha_1''^2)P_1+N_1^3))+n\epsilon_3\\
&=\frac{n}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)\\
&\quad+\frac{n}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\alpha_1''^1+\alpha_1''^2)P_1}\Big)+n\epsilon_3.
\end{aligned}$$

3) Proving the Converse of the Inequality in (16d) Corresponding to $i = 2$: In order to prove this inequality, (22g) is used to show that

$$\begin{aligned}
&n(R_0+R_2+R_3))\\
&\quad\le H(W_1)-H(W_1|\mathcal{U}_1^3)+H(Z_1|\mathcal{U}_1^3)-H(Z_1|\mathcal{U}_1^2)\\
&\qquad+H(W_2)-H(W_2|X_2)+n\epsilon_2+n\epsilon_3\\
&\quad\le\frac{n}{2}\log(2\pi e(P_1+N_1^3))\\
&\qquad-\frac{n}{2}\log(2\pi e((\alpha_1''^1+\alpha_1''^2)P_1+N_1^3))\\
&\qquad+\frac{n}{2}\log(2\pi e((\alpha_1''^1+\alpha_1''^2)P_1+N_1^2))\\
&\qquad-\frac{n}{2}\log(2\pi e(\alpha_1''^1P_1+N_1^2))\\
&\qquad+\frac{n}{2}\log(2\pi e(P_2+N_2^1))-\frac{n}{2}\log(2\pi eN_2^1)+n\epsilon_2+n\epsilon_3\\
&\quad\le\frac{n}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\alpha_1''^2+\alpha_1''^1)P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_1^2+(\alpha_1''^1+\alpha_1''^2)P_1}{N_1^2+\alpha_1''^1P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)+n\epsilon_2+n\epsilon_3.
\end{aligned}$$

4) Proving the Converse of the Inequality in (16d) Corresponding to $i = 3$: Using (22h)

$$\begin{aligned}
&n(R_0+R_1+R_2+R_3)\\
&\quad\le\frac{n}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\alpha_1''^2+\alpha_1''^1)P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_1^2+(\alpha_1''^1+\alpha_1''^2)P_1}{N_1^2+\alpha_1''^1P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)+H(Y_1|\mathcal{U}_1^2)-H(Y_1|X_1)+n\epsilon_1\\
&\qquad+n\epsilon_2+n\epsilon_3\\
&\quad\le\frac{n}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\alpha_1''^2+\alpha_1''^1)P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_1^2+(\alpha_1''^1+\alpha_1''^2)P_1}{N_1^2+\alpha_1''^1P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)+\frac{n}{2}\log(2\pi e(\alpha_1''^1P_1+N_1^1))\\
&\qquad-\frac{n}{2}\log(2\pi eN_1^1)+n\epsilon_1+n\epsilon_2+n\epsilon_3
\end{aligned}$$

$$
\begin{aligned}
&= \frac{n}{2}\log\Big(\frac{N_1^3+P_1}{N_1^3+(\alpha_1''^2+\alpha_1''^1)P_1}\Big)\\
&\quad+\frac{n}{2}\log\Big(\frac{N_1^2+(\alpha_1''^1+\alpha_1''^2)P_1}{N_1^2+\alpha_1''^1P_1}\Big)\\
&\quad+\frac{n}{2}\log\Big(\frac{N_2^1+P_2}{N_2^1}\Big)+\frac{n}{2}\log\Big(\frac{N_1^1+\alpha_1''^1P_1}{N_1^1}\Big)+n\epsilon_1\\
&\quad+n\epsilon_2+n\epsilon_3.
\end{aligned}
$$

APPENDIX O
APPLICATION TO THE GAUSSIAN CHANNEL—RSPC_2

The application of the entropy power inequality to show that RSPC_2 contains the set of all achievable rate vector uses essentially the same methodology as that used in Appendix N for RSPC_1, but with the partitions $\boldsymbol{\alpha}'$ chosen so as to satisfy the following equalities:

$$H(Y_1|\mathcal{V}_1^3)=\frac{n}{2}\log\Big(2\pi e\big((\alpha_1'^1+\alpha_1'^2)P_1+N_1^1\big)\Big) \tag{148}$$

$$H(W_2|\mathcal{V}_2^3)=\frac{n}{2}\log\Big(2\pi e\big((\alpha_2'^1+\alpha_2'^2)P_2+N_2^1\big)\Big) \tag{149}$$

$$H(Y_1|\mathcal{X}_1^2)=\frac{n}{2}\log\big(2\pi e(\alpha_1'^1P_1+N_1^1)\big) \tag{150}$$

$$H(W_2|\mathcal{X}_2^2)=\frac{n}{2}\log\big(2\pi e(\alpha_2'^1P_2+N_2^1)\big). \tag{151}$$

Using these partitions along with the inequalities in (24a)–(24d) yields

$$
\begin{aligned}
&n(R_0+R_2)\\
&\quad\le\frac{n}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\alpha_2'^1P_2}\Big)+\frac{n}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+\alpha_1'^1P_1}\Big)+n\epsilon_2\\
&n(R_0+R_2+R_3)\\
&\quad\le\frac{n}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\alpha_2'^1P_2}\Big)+\frac{n}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+\alpha_1'^1P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_2^1+\alpha_2'^1P_2}{N_2^1}\Big)+n(\epsilon_2+\epsilon_3)\\
&n(R_0+R_1+R_2+R_3)\\
&\quad\le\frac{n}{2}\log\Big(\frac{N_2^2+P_2}{N_2^2+\alpha_2'^1P_2}\Big)+\frac{n}{2}\log\Big(\frac{N_1^2+P_1}{N_1^2+\alpha_1'^1P_1}\Big)\\
&\qquad+\frac{n}{2}\log\Big(\frac{N_1^1+\alpha_1'^1P_1}{N_1^1}\Big)+\frac{n}{2}\log\Big(\frac{N_2^1+\alpha_2'^1P_2}{N_2^1}\Big)\\
&\qquad+n(\epsilon_1+\epsilon_2+\epsilon_3)
\end{aligned}
$$

which is the desired converse.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Dr. Zhi-Quan (Tom) Luo of the University of Minnesota for his insightful comments and suggestions during the development of this work. The first author would also like to thank Dr. Tricia Willink of the Communications Research Centre, Ottawa, Ontario, Canada, and Dr. Halim Yanikomeroglu of Carleton University, Ottawa, Ontario, Canada, for their support and encouragement during various phases of this work. The authors would also like to express their gratitude to the anonymous reviewers for their diligent reading of our manuscript and their constructive suggestions and insights.

REFERENCES

[1] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[3] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 2, pp. 197–207, Mar. 1973.
[4] R. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Peredachi Inf.*, vol. 10, pp. 185–193, Jul.–Sep. 1974.
[5] P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 279–280, Mar. 1974.
[6] A. J. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 219–240, Jan. 2001.
[7] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
[8] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.
[9] D. Hughes-Hartogs, "The capacity of a degraded spectral Gaussian broadcast channel," Ph.D. dissertation, Stanford Univ., Stanford, CA, Jul. 1975.
[10] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 374–377, May 1978.
[11] D. N. C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. Int. Symp. Inf. Theory*, Ulm, Germany, Jun. 1997, p. 27.
[12] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
[13] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 265–2668, Oct. 2003.
[14] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 361–374, Feb. 2006.
[15] S. Vishwanath, G. Kramer, S. Shamai (Shitz), S. Jafar, and A. Goldsmith, "*Capacity bounds for Gaussian vector broadcast channels*," in G. J. Foschini and S. Verdu (eds) Multiantenna Channels: Capacity, Coding and Signal Processing, AMS-DIMACS Series, Vol 62, pp. 107-122, 2003. (Papers from the DIMACS Workshop on Signal Processing for Wireless Transmission, Piscataway, NJ, October 2002).
[16] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
[17] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.
[18] T. M. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 4, pp. 399–404, Jul. 1975.
[19] E. C. van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 180–190, Mar. 1975.
[20] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.
[21] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Peredachi Inf.*, vol. 16, pp. 24–34, Jan.–Mar. 1980.
[22] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.
[23] Y. Liang, G. Kramer, and H. Poor, "Equivalence of two inner bounds on the capacity region of the broadcast channel," in *Proc. Allerton Conf. Commun, Control, Comput.*, Monticello, IL, Sep. 2008, pp. 1417–1421.
[24] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.

[25] C. Nair, An outer bound for 2-receiver discrete memoryless broadcast channels 2008 [Online]. Available: http://arxiv.org/abs/0807.3593

[26] Y. Liang, G. Kramer, and S. Shamai (Shitz), "Capacity outer bounds for broadcast channels," in *Proc. IEEE Inf. Theory Workshop*, Porto, Portugal, May 2008, pp. 2–4.

[27] A. A. Gohari, A. El Gamal, and V. Anantharam, "On an outer bound and an inner bound for the general broadcast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, Jun. 2010, pp. 540–544.

[28] A. A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Peredachi Inf.*, vol. 16, pp. 3–23, Jan.–Mar. 1980.

[29] T. S. Han, "The capacity region for the deterministic broadcast channel with a common message," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 122–125, Jan. 1981.

[30] C. Nair and A. El Gamal, "The capacity region of a class of 3-receiver broadcast channels with degraded message sets," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1706–1710.

[31] S. Borade, L. Zheng, and M. Trott, "Multilevel broadcast networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1151–1155.

[32] L. Grokop and D. N. C. Tse, "Fundamental constraints on multicast capacity regions," Sep. 2008 [Online]. Available: http://arxiv.org/abs/0809.2835v1

[33] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[34] Y.-K. Chia and A. El Gamal, "3-receiver broadcast channels with common and confidential messages," 2011 [Online]. Available: http://arxiv.org/abs/0910.1407

[35] N. Jindal and A. J. Goldsmith, "Capacity and dirty paper coding for Gaussian broadcast channels with common information," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, p. 214.

[36] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Nashua, NH: Athena Scientific, 1999.

[37] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

Ramy H. Gohary (S'02–M'06) received the B.Eng. (Hons.) degree from Assiut University, Egypt in 1996, the M.Sc. degree from Cairo University, Egypt, in 2000, and the Ph.D. degree from McMaster University, Ontario, Canada in 2006, all in electronics and communications engineering. He received the Natural Sciences and Engineering Research Council visiting fellowship award in 2007.

Dr. Gohary was a visiting fellow with the Terrestrial Wireless Systems Branch, Communications Research Centre, Canada. He is currently the project manager of the Carleton-RIM (Research In Motion) research project.

His research interests include analysis and design of MIMO wireless communication systems, applications of optimization and geometry in signal processing and communications, information-theoretic aspects of multiuser communication systems, and applications of iterative detection and decoding techniques in multiple antenna and multiuser systems.

Timothy N. Davidson (M'96) received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 1995.

He is currently a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he holds the (Tier II) Canada Research Chair in Communication Systems and is currently serving as Chair of the Department. His research interests lie in the general areas of communications, signal processing, and control.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from UWA, the 1991 Rhodes Scholarship for Western Australia, and a 2011 Best Paper Award from the IEEE Signal Processing Society. He has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II. He has also served as a Guest Co-Editor of issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. Dr. Davidson is currently serving as Vice-Chair of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking. He is a Registered Professional Engineer in the Province of Ontario.