HAL
open science

Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel

Nizar Khalfet, Samir M. Perlaza

To cite this version:

Nizar Khalfet, Samir M. Perlaza. Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel. IEEE Journal on Selected Areas in Communications, 2019, 37 (1), pp.156-170. 10.1109/jsac.2018.2872365 . hal-01874019v2

HAL Id: hal-01874019

https://hal.science/hal-01874019v2

Submitted on 3 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel

Nizar Khalfet and Samir M. Perlaza

***Abstract*—In this paper, the fundamental limits of simultaneous information and energy transmission (SIET) in the two-user Gaussian interference channel (G-IC) with and without perfect channel-output feedback are approximated by two regions in each case, i.e., an achievable region and a converse region. When the energy transmission rate is normalized by the maximum energy rate, the approximation is within a constat gap. In the proof of achievability, the key idea is the use of power-splitting between two signal components: an information-carrying component and a no-information component. The construction of the former is based on random coding arguments, whereas the latter consists in a deterministic sequence known by all transmitters and receivers. The proof of the converse is obtained via cut-set bounds, genie-aided channel models, Fano's inequality and some concentration inequalities considering that channel inputs might have a positive mean. Finally, the energy transmission enhancement due to feedback is quantified and it is shown that feedback can at most double the energy transmission rate at high signal to noise ratios.**

***Index Terms*—Feedback, Gaussian interference channel, simultaneous information and energy transmission, RF harvesting, information-energy capacity region.**

I. Introduction

Battery dependency is a critical issue when communications systems are deployed in hard-to-reach locations, e.g., remote geographical areas, concrete structures, human bodies, or disaster/war zones. In this case, the lifetime of the electronic devices or even the whole communications system is determined by the battery life. An effective remedy is using energy harvesting technologies. Specifically, energy can be harvested from different ambient sources such as light, vibrations, heat, chemical reactions, physiological processes, or the radio frequency (RF) signals produced by other communications systems. This observation rises the idea of simultaneous information and energy transmission (SIET) via RF.

The idea of wireless energy transmission traces back to Tesla in the 20-th century [34]. However, only recently the idea of SIET has been formalized. For instance, the point-to-point channel has been studied from the perspective of SIET in the asymptotic block-length regime in [19], [37], [38] and [39]. In the non-asymptotic block-length regime, very little is known as pointed out in [28] and [32]. The Gaussian multiple access channel with one non-colocated energy harvester (EH) has been studied in [1]. The special case of a colocated EH was studied in [17]. The Gaussian interference channel (G-IC) has been studied in the context of SIET in [7], [8] and [26]. A closely related result is the analysis of energy cooperation at the transmitters [31]. Nonetheless, none of these works tackles the fundamental limits of SIET in the G-IC. Other multi-user channels are studied in [2], [12], [18], [25], [29], [35] and references therein.

This paper focuses on the case of the two-user G-IC under the following assumptions: (a) A non-colocated energy harvester (EH) recollects energy from the signals sent by the transmitters; (b) a perfect channel-output feedback link might be available from each receiver to the corresponding transmitter. The main results in the case with and without feedback are twofold. First, a set of achievable information and energy transmission rates is presented. That is, a subset of the information-energy capacity region [1] is characterized. Second, a set that contains the information-energy capacity region is introduced. These two sets are shown to be a constant gap approximation to the information-energy capacity region when the energy rate is normalized by the maximum energy rate. Using these results, it is finally shown that perfect channel-output feedback can at most double the energy rate in the G-IC.

This work builds upon existing results whose foci are exclusively on information transmission, in particular [14], [20], [30] and [33]. For instance, in the proof of achievability with and without feedback, the novelty lyes upon the use of power splitting to superimpose two signal components: One that is designed to transmit information; and another that is designed to transmit energy. Note that the former naturally carries both information and energy, whereas the latter exclusively carries energy. From this perspective, the information-carrying component is built up along the same lines of the achievability scheme presented in [21] and [33] in the case with and without feedback, respectively. The no-information carrying component is built along the lines of the scheme presented in [1] using the idea of common randomness. More specifically, both the transmitters and the receivers know in advance the realization of a given random variable. This realization is used by the transmitters to correlate their transmitted signals in order to benefit the transmission of energy. At the receivers, this knowledge is used for cancelling the interference of the no-information carrying component. The proof of the converse, in the case with and without feedback, follows along

Nizar Khalfet and Samir M. Perlaza are with the CITI Laboratory, a joint laboratory between the Institut National de Recherche en Informatique et en Automatique (INRIA), the Université de Lyon and the Institut National de Sciences Apliquées (INSA) de Lyon. 6 Av. des Arts 69621 Villeurbanne, France. ({nizar.khalfet, samir.perlaza}@inria.fr)

Samir M. Perlaza is also with the Department of Electrical Engineering at Princeton University, Princeton, NJ 08544 USA.

Parts of this work were presented in the International Zurich Seminar on Information and Communication, Zurich, Switzerland, Feb., 2018 [23]; and in the 55th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, US, Oct., 2017 [22].

This research was supported in part by the European Commission under Marie Skłodowska-Curie Individual Fellowship No. 659316 and and in part by the Agence Nationale de la Recherche under Grant ANR-15-NMED-0009-03.

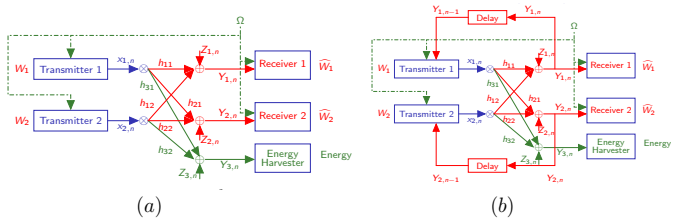

Fig. 1. Two-user Gaussian interference channels with a non-colocated energy harvester at channel use n. (a) Case without feedback; and (b) Case with perfect channel output feedback.

the lines of [14] and [33], respectively. The main difference relies on the fact that no constraint is imposed on the first moment of codewords of a given feasible coding scheme. Note for instance that a codeword with a strictly positive mean carries an amount of energy equal to its mean, i.e., direct current (DC) component. Finally, new inner and outer bounds on the energy transmission rate are obtained using Markov's concentration inequality [5].

The proofs are presented only for the case with feedback and only the intuitions behind the proofs are discussed for the case without feedback. A complete presentation of the proofs is presented in [24].

II. System Model

Consider a two-user G-IC with a non-colocated energy harvester (EH) with and without point-to-point perfect channel-output feedback (PF) from each receiver to its corresponding transmitter. These two scenarios are depicted in Figure 1(a) and Figure 1(b), respectively. Note that there is no feedback from the EH to any of the transmitters. Within this context, transmitter i, with $i \in \{1,2\}$, aims to simultaneously execute two tasks: (a) information transmission to its intended receiver; and (b) energy transmission to the EH. Section II-A and Section II-B describe these tasks independently. Later, Section II-C describes the simultaneous implementation of both tasks.

A. Information Transmission Task

From the information transmission standpoint, the goal of transmitter i, with $i \in \{1,2\}$, is to convey a message index $W_i \in \mathcal{W}_i = \{1,2,\ldots,\lfloor 2^{NR_i}\rfloor\}$ to receiver i using N channel input symbols $X_{i,1}, X_{i,2}, \ldots, X_{i,N}$. That is, information is transmitted at rate $R_i > 0$ bits per channel use. The channel coefficient from transmitter k to receiver i, with $k \in \{1,2\}$, is denoted by $h_{i,k} \in \mathbb{R}_+$, where $\mathbb{R}_+$ denotes the positive reals. At receiver i, during channel use n, input symbol $X_{i,n}$ is observed at receiver i subject to the interference produced by the symbol $X_{j,n}$ sent by transmitter j, with $j \in \{1,2\} \setminus \{i\}$, and a real additive Gaussian noise $Z_{i,n}$ with zero mean and variance σ_i^2. Hence, the channel output at receiver i during channel use n, denoted by $Y_{i,n}$, is:

$$Y_{i,n} = h_{i,i}X_{i,n} + h_{i,j}X_{j,n} + Z_{i,n}. \tag{1}$$

In the case without feedback, at each channel use n, the symbol $X_{i,n}$ sent by transmitter i depends upon the message index W_i and a randomly generated index $\Omega \in \mathbb{N}$. Let $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \to \mathbb{R}$ be the encoding function at channel use n, such that for all $n \in \{1,2,\ldots,N\}$, the following holds:

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega). \tag{2}$$

In the case with feedback, the symbol $X_{i,n}$ sent by transmitter i depends upon the indices W_i and Ω, but also upon all previous channel-outputs $Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n-d}$, with $d \in \mathbb{N}$ the feedback delay. In the following, it is assumed that d is equal to one channel use, without any loss of generality. Thus, the first channel input symbol $X_{i,1}$ depends only on the message index W_i and Ω. More specifically, $f_{i,1}^{(N)} : \mathcal{W}_i \times \mathbb{N} \to \mathbb{R}$. Alternatively, for all $n \in \{2,3,\ldots,N\}$, the encoding functions are $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \times \mathbb{R}^{n-1} \to \mathbb{R}$. Essentially,

$$X_{i,1} = f_{i,1}^{(N)}(W_i, \Omega), \tag{3a}$$

and for all $n > 1$,

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega, Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n-1}). \tag{3b}$$

In both cases, with and without feedback, the random index Ω is assumed to be known by all transmitters and receivers. Moreover, channel input symbols $X_{i,1}, X_{i,2}, \ldots, X_{i,N}$ are subject to an average power constraint of the form

$$\frac{1}{N}\sum_{n=1}^{N} \mathbb{E}_{X_{i,n}}\left[X_{i,n}^2\right] \leq P_i, \tag{4}$$

where P_i denotes the average transmit power of transmitter i in energy units per channel use. The decoder of receiver i observes the channel outputs $Y_{i,1}, Y_{i,2}, \ldots, Y_{i,N}$ and uses a decoding function $\phi_i^{(N)} : \mathbb{N} \times \mathbb{R}^N \to \mathcal{W}_i$, to get an estimate of the message indices:

$$\widehat{W_i} = \phi_i^{(N)}\left(\Omega, Y_{i,1}, Y_{i,2}, \ldots, Y_{i,N}\right), \tag{5}$$

where $\widehat{W_i}$ is an estimate of the message index W_i. The decoding error probability of a codebook of block-length N, denoted by $P_{\text{DE}}^{(N)}$, is given by

$$P_{\text{DE}}^{(N)} = \max\left[\Pr\left[\widehat{W_1} \neq W_1\right], \Pr\left[\widehat{W_2} \neq W_2\right]\right]. \tag{6}$$

The signal to noise ratio (SNR) at receiver i is denoted by

$$\text{SNR}_i = \frac{|h_{i,i}|^2 P_i}{\sigma_i^2}. \tag{7a}$$

The interference to noise ratio (INR) at receiver i is denoted by

$$\text{INR}_i = \frac{|h_{i,j}|^2 P_j}{\sigma_i^2}, \text{ with } j \neq i. \tag{7b}$$

B. Energy Transmission Task

Let $h_{3,i} \in \mathbb{R}_+$ be the channel coefficient from transmitter i to the EH. The symbols sent by the transmitters during channel use n are observed by the EH subject to an additive Gaussian noise $Z_{3,n}$ with zero mean and variance σ_3^2. More specifically, the channel output at the EH during channel use n, denoted by $Y_{3,n}$, is:

$$Y_{3,n} = h_{3,1}X_{1,n} + h_{3,2}X_{2,n} + Z_{3,n}. \tag{8}$$

From the energy transmission standpoint, the goal of both transmitters is to jointly guarantee an average energy rate at the EH.

Let $B^{(N)} : \mathbb{R}^N \to \mathbb{R}_+$ be a function that determines the energy obtained from the channel outputs $Y_{3,1}, Y_{3,2}, \ldots, Y_{3,N}$. In the following, this function is chosen to be the average energy rate (in energy-units per channel use) at the end of N channel uses. That is,

$$B^{(N)}(Y_{3,1}, Y_{3,2}, \ldots, Y_{3,N}) \triangleq \frac{1}{N}\sum_{n=1}^{N} Y_{3,n}^2, \tag{9}$$

which implies that the energy carried by a given channel output $Y_{3,t}$, with $t \in \{1, 2, \ldots, N\}$, is $Y_{3,t}^2$. This assumption is very optimistic given the dependency of the delivered DC power on higher order statistics of the channel input distribution [36], [37]. Nonetheless, from the fundamental limits point of view, any more realistic model would induce fundamental limits that are more pessimistic than the results presented in the following sections.

The SNR of transmitter i at the EH is denoted by

$$\mathrm{SNR}_{3i} = \frac{|h_{3,i}|^2 P_i}{\sigma_3^2}. \tag{10}$$

Note that the maximum average energy rate, denoted by $B_{\max}$, is:

$$B_{\max} = \sigma_3^2 \left(1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32} + 2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}\right), \tag{11}$$

which can be achieved in the asymptotic block-length regime when both channel inputs exhibit a correlation coefficient equal to one. Hence, given an energy rate $B \in [0, B_{\max}]$, the energy shortage probability, denoted by $P_{\mathrm{ES}}^{(N)}(B)$, is:

$$P_{\mathrm{ES}}^{(N)} \triangleq \Pr\left[B^{(N)}(\boldsymbol{Y}_3) < B\right]. \tag{12}$$

C. Simultaneous Information and Energy Transmission

The system is said to operate at the information-energy rate triplet $(R_1, R_2, B) \in \mathbb{R}_+^3$ when both transmitter-receiver pairs use a transmit-receive configuration such that: (i) reliable communication at information rates R_1 and R_2 is ensured; and (ii) reliable energy transmission at energy rate B is ensured. A formal definition is given below.

Definition 1 (Achievable Rates): The triplet $(R_1, R_2, B) \in \mathbb{R}_+^3$ is achievable if for all $i \in \{1, 2\}$, there exists a sequence of encoding functions $f_{i,1}^{(N)}, f_{i,2}^{(N)}, \ldots, f_{i,N}^{(N)}$ and two decoding functions $\phi_1^{(N)}$ and $\phi_2^{(N)}$ such that both the average decoding error probability $P_{\mathrm{DE}}^{(N)}$ and the energy-shortage probability $P_{\mathrm{ES}}^{(N)}$ tend to zero as the block-length N tends to infinity. That is,

$$\limsup_{N\to\infty} P_{\mathrm{DE}}^{(N)} = 0 \text{ and} \tag{13a}$$

$$\limsup_{N\to\infty} P_{\mathrm{ES}}^{(N)} = 0. \tag{13b}$$

Using Definition 1, the fundamental limits of simultaneous information and energy transmission in the Gaussian interference channel can be described by the information-energy capacity region [2], defined as follows.

Definition 2 (Information-Energy Capacity Region): The information-energy capacity region, denoted by $\mathcal{E}^{\mathrm{F}}$ in the case with feedback and $\mathcal{E}$ in the case without feedback, corresponds to the closure of all achievable information-energy rate triplets (R_1, R_2, B).

D. A Note on the Main Results

The main results of this paper consist in descriptions of the information-energy capacity regions with feedback $\mathcal{E}^{\mathrm{F}}$ and without feedback $\mathcal{E}$. Such a description is presented in the form of an *approximation* in the sense of the definition hereunder.

Definition 3 (Approximation of a Set): Let $n \in \mathbb{N}$ be fixed. A set $\mathcal{X} \subset \mathbb{R}_+^n$ is approximated by the sets $\underline{\mathcal{X}}$ and $\overline{\mathcal{X}}$ if $\underline{\mathcal{X}} \subseteq \mathcal{X} \subseteq \overline{\mathcal{X}}$ and $\forall \boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \overline{\mathcal{X}}$ then $\Big((x_1 - \xi_1)^+, (x_2 - \xi_2)^+, \ldots, (x_n - \xi_n)^+\Big) \in \underline{\mathcal{X}}$, for some $(\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}_+^n$.

In Definition (3), the operator $\max[\cdot, 0]$ is denoted by $(\cdot)^+$.

III. Main Results: Case without Channel-Output Feedback

The information-energy capacity region $\mathcal{E}$ is approximated by the regions $\underline{\mathcal{E}} \subset \mathbb{R}_+^3$, which represents an information-energy achievable region (Theorem 1); and $\overline{\mathcal{E}} \subset \mathbb{R}_+^3$, which represents an information-energy converse region (Theorem 2). Regions $\underline{\mathcal{E}}$ and $\overline{\mathcal{E}}$ satisfy $\underline{\mathcal{E}} \subseteq \mathcal{E} \subseteq \overline{\mathcal{E}}$ and approximate the information-energy region $\mathcal{E}$ to within a given gap (Theorem 3), in the sense of Definition 3.

A. An Achievable Region

The following theorem introduces an achievable information-energy region.

Theorem 1: The information-energy capacity region $\mathcal{E}$ contains the set $\underline{\mathcal{E}} \subseteq \mathbb{R}_+^3$ of all rate tuples (R_1, R_2, B) that satisfy:

$$R_1 \le \frac{1}{2}\log\left(1 + \frac{(1-\lambda_{1e})\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{14a}$$

$$R_2 \le \frac{1}{2}\log\left(1 + \frac{(1-\lambda_{2e})\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14b}$$

$$R_1 + R_2 \le \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(1 + \frac{\lambda_{2p}\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14c}$$

$$R_1 + R_2 \le \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right) + \frac{1}{2}\log\left(1 + \frac{\lambda_{1p}\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{14d}$$

$$R_1 + R_2 \le \frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14e}$$

$$2R_1 + R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)$$
$$+\frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right)$$
$$+\frac{1}{2}\log\left(1+\frac{\lambda_{1p}\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right) \tag{14f}$$

$$R_1 + 2R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right)$$
$$+\frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)$$
$$+\frac{1}{2}\log\left(1+\frac{\lambda_{2p}\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14g}$$

$$B \leq \sigma_3^2\Big(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}$$
$$+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}\sqrt{\lambda_{1e}\lambda_{2e}}\Big), \tag{14h}$$

for some $(\lambda_{ip}, \lambda_{ie}) \in [0,1]^2$ such that $\lambda_{ip} + \lambda_{ie} \leq 1$, for all $i \in \{1,2\}$.

The proof of Theorem 1 is presented in [24]. Essentially, the achievability scheme used to obtain the region $\underline{\mathcal{E}}$ described in Theorem 1 is built upon random coding arguments using four key ingredients: (a) superposition coding [9]; (b) rate-splitting [20]; (c) common randomness [4], [27]; and (d) power-spliting [1].

The codebook of transmitter i, with $i \in \{1,2\}$, is generated by super-imposing three different code layers. The first code layer is a sub-codebook generated for the exclusive purpose of energy transmission. Note that this code layer can be chosen to be the same for both transmitters. The key point is to ensure that codewords in the first code layer of transmitter 1 and 2 exhibit a correlation factor equal to one. For each codeword in the first layer, a new sub-codebook is generated. This set of sub-codebooks is referred to as the second code layer and it is designed to broadcast information to both receivers. However, even if it is not the primary goal, these codewords naturally carry energy to the EH, as well. Finally, for each codeword in the second layer, a new sub-codebook is generated. This set of sub-codebooks is referred to as the third layer of the codebook and it is designed for the exclusive purpose of transmitting information to receiver i. Nonetheless, as for the codewords in the first and second layer, these codewords also carry energy to the EH.

In a nutshell, codewords from all layers of the codebook are capable of carrying energy to the EH but only those in the second and third layer carry both information and energy. The size of the first layer of the codebook determines the number of different codewords that can be used to transmit energy to the EH. However, the size of this layer does not have any impact on the information or energy rate of the transmitters. Alternatively, the size of the second and third layer determine the information rate of the corresponding transmitter. The exact size of each of these layers lies upon a decoding error probability analysis that is presented in [24].

Rate splitting is the ingredient that allows the convenient exploitation of the codebooks with the form described above. Note that at the beginning of each transmission, transmitter i possesses two indices to transmit: common random index Ω and message index W_i. The message index W_i is divided into two subindices: $W_{i,C}$ and $W_{i,P}$. The index Ω is used to choose a codeword in the first layer and the indices $W_{i,C}$ and $W_{i,P}$ are used to choose a codeword in the second and third layer, respectively. This justifies the name of the technique as the information rate of transmitter i is *split* into two streams: *common* and *private*. Note that the second layers contain codewords that are decoded at both receivers (common messages) whereas the third layers contain codewords that are decoded only at the intended receiver (private messages). Intuitively, the codewords from the second layer of the code of transmitter i can be decoded at receiver j, with $j \in \{1,2\}\backslash\{i\}$, which allows some interference cancellation. On the other hand, the codewords from the third layer of transmitter i are treated as interference at receiver j. The interference produced by the codewords from the first layer on both transmitters can be fully eliminated, as by assumption, the index Ω is known by all transmitters and receivers.

Finally, to prove the existence of at least one code that achieves the rates described by Theorem 1, it suffices to average the information and energy rates that are achievable by all possible codebooks that can be generated using the structure described above. If the average of such rates satisfies the inequalities in Theorem 1, then for each rate tuple in $\underline{\mathcal{E}}$, there exists at least one code that achieves such a rate tuple. Assume for instance that the codewords of the first, second and third layers of transmitter i are N-length sequences of realizations of the following three independent random variables respectively: $V \sim \mathcal{N}(0,1)$; $U_i \sim \mathcal{N}(0,\lambda_{ic})$; and $S_i \sim \mathcal{N}(0,\lambda_{ip})$, where $\lambda_{ic} + \lambda_{ip} + \lambda_{ie} \leqslant 1$. Let also the channel input of transmitter i, during any given channel use be:

$$X_i = \sqrt{P_i}S_i + \sqrt{P_i}U_i + \sqrt{\lambda_{ie}P_i}V. \tag{15}$$

At channel use n and given any possible codebook with the structure described above, the n-th channel input of transmitter i is a weighted sum of the n-th symbols of the corresponding codewords in the three layers of such codebook. The weighting is referred to as *power splitting* to highlight that a fraction λ_{ie} of the total average power P_i is used to transmit a codeword whose role is to exclusively transmit energy to the EH. The information-carrying component, which is the sum of the codewords from the second and third layers of the codebook, is transmitted using an average power $\lambda_{ic} + \lambda_{ip} \leqslant 1 - \lambda_{ie}$.

The role of the first layer of the codebook becomes clearer after the following remarks.

Remark 1**:** *When* $\lambda_{1e} = \lambda_{2e} = 1$*, the left-hand sides of inequalities* (14a)-(14g) *become zero, whereas the left-hand side of inequality* (14h) *is maximized. That is, a zero information rate is achieved at the same time that the highest energy rate* $B_{\max}$ *in* (11) *is achieved. This is essentially because the transmitted codewords belong to the first layers of the codebooks of both transmitters. Note also that the choice is made such that the correlation coefficient of both channel inputs is one.*

Remark 2: *When $\lambda_{1e} = \lambda_{2e} = 0$, the codewords of the first layers of the code are not transmitted. From this perspective, both channel input signal are independent of each other and thus, the energy rate is at most $\sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}\right)$ energy units per channel use.*

Note that Remark 1 and Remark 2 highlight the fact that the no-information component is needed to transmit energy beyond the energy rate $\sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}\right)$. Thanks to this no-information component, the signals of both transmitters can be correlated, which results into higher energy rates than those achieved by independent signals.

Remark 3: *A consequence of Remark 2 is that for all rate tuples $(R_1, R_2, B) \in \underline{\mathcal{E}}$ or $(R_1, R_2, B) \in \overline{\mathcal{E}}$, with*

$$B \leqslant \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}\right),$$

it follows that the rate pairs (R_1, R_2) form respectively the inner region or the outer region of the information capacity region described in [20]. Alternatively, for all rate tuples $(R_1, R_2, B) \in \underline{\mathcal{E}}$ or $(R_1, R_2, B) \in \overline{\mathcal{E}}$, with $B > \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}\right)$, it follows that the rate pairs (R_1, R_2) form a proper set of the inner or outer region of the information capacity region described in [20], respectively. This observation implies that a trade-off between energy and information rates is observed when $B > \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}\right)$. This is compliant with previous observations in other multi-user channels, e.g., the multiple access channel [1].

Remark 4: *Note that the first layer of the code does not contribute to the information rate. Hence, there is no constraint on reducing the size of the first layer to one codeword. That is, the assumption of common randomness can be soften to the knowledge of a sufficiently large codeword whose purpose is exclusively transmitting energy to the EH, e.g., a pseudo-random sequence.*

In Section V, Remark 1 - Remark 3 are highlighted in particular numerical examples.

B. A Converse Region

The following Theorem introduces an information-energy converse region.

Theorem 2: The information-energy capacity region $\mathcal{E}$ is contained into the set $\overline{\mathcal{E}} \in \mathbb{R}_+^3$, which contains all rate tuples (R_1, R_2, B) that satisfy:

$$R_1 \leq \frac{1}{2}\log(1+\beta_1\mathrm{SNR}_1), \tag{16a}$$

$$R_2 \leq \frac{1}{2}\log(1+\beta_2\mathrm{SNR}_2), \tag{16b}$$

$$R_1+R_2 \leq \frac{1}{2}\log(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1) + \frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2}{1+\beta_2\mathrm{INR}_1}\right), \tag{16c}$$

$$R_1+R_2 \leq \frac{1}{2}\log(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2) + \frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1}{1+\beta_1\mathrm{INR}_2}\right), \tag{16d}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_1\mathrm{INR}_2}\right) + \frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_2\mathrm{INR}_1}\right), \tag{16e}$$

$$2R_1+R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1}{1+\beta_1\mathrm{INR}_2}\right) + \frac{1}{2}\log(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1) + \frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_2\mathrm{INR}_1}\right), \tag{16f}$$

$$R_1+2R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2}{1+\beta_2\mathrm{INR}_1}\right) + \frac{1}{2}\log(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2) + \frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_1\mathrm{INR}_2}\right), \tag{16g}$$

$$B \leq \sigma_3^2\Big(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}\sqrt{(1-\beta_1)(1-\beta_2)}\Big), \tag{16h}$$

for some $(\beta_1, \beta_2) \in [0,1]^2$.

The proof of Theorem 2 is presented in [24]. From the information transmission perspective, the proof of the upper bounds on the information rates is identical to the proof presented in [14]. That is, (16a) and (16b) are simple cut-set bounds. The bounds (16c) - (16g) are obtained considering genie-aided channels and Fano's inequality [15]. For completness, the proof of the upper-bounds (16a)-(16g) is presented in [24]. The upper-bound on the energy transmission rate (16h) is identical in the cases with and without feedback. Thus, the reader is referred to Appendix B.

C. An Approximation to the Information Energy Capacity Region

Using the inner region $\underline{\mathcal{E}}$ and the outer region $\overline{\mathcal{E}}$, described respectively by Theorem 1 and Theorem 2, the information-energy capacity region $\mathcal{E}$ can be approximated in the sense of Definition 3. The following theorem presents this result.

Theorem 3 (Approximation of $\mathcal{E}$): Let $\underline{\mathcal{E}} \subset \mathbb{R}_+^3$ and $\overline{\mathcal{E}} \subset \mathbb{R}_+^3$ be the sets of tuples (R_1, R_2, B) described by Theorem 1 and Theorem 2, respectively. Then,

$$\underline{\mathcal{E}} \subset \mathcal{E} \subset \overline{\mathcal{E}}, \tag{17}$$

and for all $(R_1, R_2, B) \in \overline{\mathcal{E}}$ it follows that $\Big((R_1 - 1/2)^+, (R_2 - 1/2)^+, \Big(B - \frac{B_{\max}}{2}\Big)^+\Big) \in \underline{\mathcal{E}}$.

The proof of Theorem 3 is presented in [24]. It is essentially algebraic and thus, no further comment is made about this proof. Note that the approximation in Theorem 3 is not an approximation within a constant gap. This is because the gap in the energy component is at most $\frac{B_{\max}}{2}$ energy units per channel use, with $B_{\max}$ in (11). Thus, it depends on σ_3^2, SNR_{31} and SNR_{32}. A constant gap approximation is obtained only when considering the set of

tuples formed by the information rates R_1 and R_2 in bits per channel use and the normalized rate $\frac{B}{B_{\max}}$. That is, the set $\mathcal{E}' = \left\{\left(R_1, R_2, \frac{B}{B_{\max}}\right) : (R_1, R_2, B) \in \mathcal{E}\right\}$ is approximated to within $\frac{1}{2}$ units by the sets

$$\underline{\mathcal{E}}' = \left\{\left(R_1, R_2, \frac{B}{B_{\max}}\right) : (R_1, R_2, B) \in \underline{\mathcal{E}}\right\} \text{ and} \tag{18}$$

$$\overline{\mathcal{E}}' = \left\{\left(R_1, R_2, \frac{B}{B_{\max}}\right) : (R_1, R_2, B) \in \overline{\mathcal{E}}\right\}. \tag{19}$$

That is, $\underline{\mathcal{E}}' \subset \mathcal{E}' \subset \overline{\mathcal{E}}'$, and for all $(R_1, R_2, b) \in \overline{\mathcal{E}}'$, it follows that $\left((R_1 - 1/2)^+, (R_2 - 1/2)^+, (b - 1/2)^+\right) \in \underline{\mathcal{E}}'$.

IV. Main Results: Case with Perfect Channel-Output Feedback

The information-energy capacity region $\mathcal{E}^{\mathrm{F}}$ is approximated by the regions $\underline{\mathcal{E}}^{\mathrm{F}} \subset \mathbb{R}_+^3$, which represents an information-energy achievable region (Theorem 4); and $\overline{\mathcal{E}}^{\mathrm{F}} \subset \mathbb{R}_+^3$, which represents an information-energy converse region (Theorem 5). Regions $\underline{\mathcal{E}}^{\mathrm{F}}$ and $\overline{\mathcal{E}}^{\mathrm{F}}$ satisfy $\underline{\mathcal{E}}^{\mathrm{F}} \subseteq \mathcal{E}^{\mathrm{F}} \subseteq \overline{\mathcal{E}}^{\mathrm{F}}$ and approximate the information-energy region $\mathcal{E}$ to within a given gap (Definition 3).

A. An Achievable Region

The following theorem introduces an achievable information-energy region.

Theorem 4: The information-energy capacity region $\mathcal{E}^{\mathrm{F}}$ contains the set $\underline{\mathcal{E}}^{\mathrm{F}} \subseteq \mathbb{R}_+^3$ of all rate tuples (R_1, R_2, B) that satisfy:

$$R_1 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1+2\rho\sqrt{\mathrm{SNR}_1\mathrm{INR}_1}}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{20a}$$

$$R_1 \leq \frac{1}{2}\log\left(\frac{1+(1-(\rho+\lambda_{1e}))\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right) + \frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+\lambda_{2p}\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{20b}$$

$$R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2+2\rho\sqrt{\mathrm{SNR}_2\mathrm{INR}_2}}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{20c}$$

$$R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-(\rho+\lambda_{2e}))\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_1+\lambda_{1p}\mathrm{INR}_1}{1+\lambda_{1p}\mathrm{INR}_1}\right), \tag{20d}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+\lambda_{2p}\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2+2\rho\sqrt{\mathrm{SNR}_2\mathrm{INR}_2}}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{20e}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_1+\lambda_{1p}\mathrm{INR}_1}{1+\lambda_{1p}\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1+2\rho\sqrt{\mathrm{SNR}_1\mathrm{INR}_1}}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{20f}$$

$$B \leq \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}(\rho + \sqrt{\lambda_{1e}\lambda_{2e}})\right), \tag{20g}$$

for some $(\rho, \lambda_{ip}, \lambda_{ie}) \in [0,1]^3$ such that $\rho + \lambda_{ip} + \lambda_{ie} \leq 1$, for all $i \in \{1,2\}$.

The proof of Theorem 4 is presented in Appendix A and it is based on random coding arguments using rate-splitting [6], [20]; block Markov superposition coding [3], [11]; backward decoding [40], [41]; and power splitting [1].

The codebook of transmitter i, with $i \in \{1,2\}$, is generated by super-imposing four different sub-codebooks. This contrasts with the three-layer codebook used in the case without feedback. However, both codebooks share profound similarities. The first layer in the case with and without feedback are identical and play the same role. The second layer of the codebook with feedback is obtained by the union of the second layers of both transmitters in the case without feedback. The third and fourth layers of the codebook with feedback are identical to the second and third layers of the codebook without feedback. The roles of these two layers are identical in the case with and without feedback.

The convenient exploitation of this four-layer codebook is possible thanks to a rate splitting argument similar to the one used in the case without feedback. The rate of transmitter i is split into a common and a private component with message indices $W_{ic}^{(t)}$ and $W_{ip}^{(t)}$. The super-index t is used to denote the block, e.g., sequences of N channel uses. Assume that T blocks are transmitted. At the beginning of block t, each transmitter possesses five indices: the random index $\Omega^{(t)}$; the common and private message indices $W_{ic}^{(t)}$ and $W_{ip}^{(t)}$; and the common messages $W_{1c}^{(t-1)}$ and $W_{2c}^{(t-1)}$. Transmitter i obtains the message index $W_{jc}^{(t-1)}$ of transmitter j, with $j \in \{1,2\} \setminus \{i\}$, via feedback at the end of block $t-1$. For the first block $t=1$, the previous common message indices are chosen arbitrarily as $W_{1c}^{(0)} = W_{2c}^{(0)} = 1$ and are assumed to be known by all transmitters and receivers. Similarly, the last common message indices are chosen arbitrarily as $W_{1c}^{(T)} = W_{2c}^{(T)} = 1$ and are also assumed to be known by all transmitters and receivers. Under this assumption, the random index $\Omega^{(t)}$ is used to choose a codeword from the first layer; the pair $(W_{1c}^{(t-1)}, W_{2c}^{(t-1)})$ are jointly used to choose a common codeword from the second layer. Note that the second layer of transmitter 1 is identical to the second layer of transmitter 2 by construction of the code. Moreover, thanks to feedback both transmitters are able to choose the same codeword from their second layers at each block t. The message indices $W_{ic}^{(t)}$ and $W_{ip}^{(t)}$ are used at transmitter i to choose codewords from the third and fourth layers respectively.

The channel input of transmitter i at channel use n is, as in the case without feedback, a weighted sum of the n-th symbols of the corresponding codewords in each of the four layers. A

power splitting argument is also used as in the case without feedback.

At the receivers backward decoding is used. More specifically, given that the last common indices $W_{1c}^{(T)}$ and $W_{2c}^{(T)}$ are known at the transmitters, receiver i is capable of decoding $W_{1c}^{(T-1)}$ and $W_{2c}^{(T-1)}$ and $W_{ip}^{(T)}$ at the first decoding stage by using joint-typicality arguments. At the second decoding stage, $W_{1c}^{(T-1)}$ and $W_{2c}^{(T-1)}$ are used to decode $W_{1c}^{(T-2)}$, $W_{2c}^{(T-2)}$ and $W_{ip}^{(T-1)}$. The decoding goes on until decoding stage T at which only $W_{ip}^{(1)}$ is decoded as $W_{1c}^{(0)}$ and $W_{2c}^{(0)}$ are both known.

Note that Remark 1 - Remark 3 also hold for the case with feedback taking into account the differences on the structure of the codes with and without feedback. The role of the second layer of the codebook with feedback becomes clearer after the following remark.

Remark 5: The second layer of the codebooks of both transmitters are identical and thus, given the common message indices $W_{1c}^{(t-1)}$ and $W_{2c}^{(t-1)}$ at the beginning of block t, both transmitters are able to choose the same codeword to generate their corresponding channel inputs. This creates a correlation between the channel input symbols $X_{1,n}$ and $X_{2,n}$, for all $n \in \{1, 2, \ldots, N\}$, which is advantageous to increase the information transmission sum-rate and the energy transmission rate. This implies that feedback is beneficial for both information and energy transmission.

The additional correlation highlighted in Remark 5 is captured by the term ρ in (20). Note that the left hand side of inequalities (20e) and (20g) (information transmission sum-rate) are monotonically increasing with ρ and so is the left hand side of (20g) (energy transmission rate). The benefits of feedback in SIET are studied in Section V.

B. A Converse Region

The following theorem describes a converse region denoted by $\overline{\mathcal{E}}^{\mathrm{F}}$.

Theorem 5: The information-energy capacity region $\mathcal{E}^{\mathrm{F}}$ is contained into the set $\overline{\mathcal{E}}^{\mathrm{F}} \in \mathbb{R}_+^3$ of all rate tuples (R_1, R_2, B) that satisfy:

$$R_1 \leq \frac{1}{2}\log\left(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+2\rho\sqrt{\beta_1\mathrm{SNR}_1\beta_2\mathrm{INR}_1}\right), \tag{21a}$$

$$R_1 \leq \frac{1}{2}\log\left(1+\frac{\beta_1(1-\rho^2)\mathrm{SNR}_1}{1+\beta_1(1-\rho^2)\mathrm{INR}_2}\right) + \frac{1}{2}\log\left(1+\beta_1(1-\rho^2)\mathrm{INR}_2\right), \tag{21b}$$

$$R_2 \leq \frac{1}{2}\log\left(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+2\rho\sqrt{\beta_2\mathrm{SNR}_2\beta_1\mathrm{INR}_2}\right), \tag{21c}$$

$$R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_2(1-\rho^2)\mathrm{SNR}_2}{1+\beta_2(1-\rho^2)\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(1+\beta_2(1-\rho^2)\mathrm{INR}_1\right), \tag{21d}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_1(1-\rho^2)\mathrm{SNR}_1}{1+\beta_1(1-\rho^2)\mathrm{INR}_2}\right) + \frac{1}{2}\log\left(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+2\rho\sqrt{\beta_2\mathrm{SNR}_2\beta_1\mathrm{INR}_2}\right),$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_2(1-\rho^2)\mathrm{SNR}_2}{1+\beta_2(1-\rho^2)\mathrm{INR}_1}\right) + \frac{1}{2}\log\left(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+2\rho\sqrt{\beta_1\mathrm{SNR}_1\beta_2\mathrm{INR}_1}\right), \tag{21e}$$

$$B \leq \sigma_3^2\Big(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}} (\rho\sqrt{\beta_1\beta_2}+\sqrt{(1-\beta_1)(1-\beta_2)})\Big), \tag{21f}$$

for some $(\beta_1, \beta_2, \rho) \in [0,1]^3$.

The proof of Theorem 5 is presented in Appendix B. The intuitions behind this proof are not different from those discussed in the case without feedback. Probably, the most important step on this proof is that codebooks are not assumed to be formed by codewords with zero mean. That is, the codeword considered in this proof might have a non-zero mean, as energy can also be carried in this way. The upper bounds on the information rates heavily rely on cut-set bounds [13], Fano's inequality [15] and genie aided models. The upper-bound on the energy transmission rate is an immediate consequence of Markov's inequality [16].

C. An Approximation to the Information-Energy Capacity Region

Using the inner region $\underline{\mathcal{E}}^{\mathrm{F}}$ and the outer region $\overline{\mathcal{E}}^{\mathrm{F}}$, described respectively by Theorem 4 and Theorem 5, the information-energy capacity region $\mathcal{E}^{\mathrm{F}}$ can be approximated in the sense of Definition 3.

Theorem 6 (Approximation of $\mathcal{E}^{\mathrm{F}}$): Let $\underline{\mathcal{E}}^{\mathrm{F}} \subset \mathbb{R}_+^3$ and $\overline{\mathcal{E}}^{\mathrm{F}} \subset \mathbb{R}_+^3$ be the sets of tuples (R_1, R_2, B) described by Theorem 4 and Theorem 5, respectively. Then, $\underline{\mathcal{E}}^{\mathrm{F}} \subset \mathcal{E}^{\mathrm{F}} \subset \overline{\mathcal{E}}^{\mathrm{F}}$, and for all $(R_1, R_2, B) \in \overline{\mathcal{E}}^{\mathrm{F}}$ it follows that $\left((R_1-1)^+, (R_2-1)^+, \left(B-\frac{B_{\max}}{2}\right)^+\right) \in \underline{\mathcal{E}}^{\mathrm{F}}$.

The proof of Theorem 6 is presented in [24]. Note that a constant gap approximation can be obtained by normalizing the energy transmission rate as suggested in the case without feedback.

D. Maximum Energy Rate Improvement with Feedback

Consider the following sets of energy rates: $\underline{\mathcal{B}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \underline{\mathcal{E}}\}$, $\overline{\mathcal{B}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \overline{\mathcal{E}}\}$, $\underline{\mathcal{B}}_{\mathrm{F}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \underline{\mathcal{E}}_{\mathrm{F}}\}$, and $\overline{\mathcal{B}}_{\mathrm{F}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \overline{\mathcal{E}}_{\mathrm{F}}\}$. The maximum improvement that can be achieved on the energy rate due to feedback can be shown to be at most a factor of two. The following proposition shows this by providing upper bounds on the ratios $\frac{\max \underline{\mathcal{B}}_{\mathrm{F}}}{\max \underline{\mathcal{B}}}$ and $\frac{\max \overline{\mathcal{B}}_{\mathrm{F}}}{\max \overline{\mathcal{B}}}$.

Proposition 1 (Rate improvement with Feedback): The energy rate achievable in the two-user G-IC with perfect

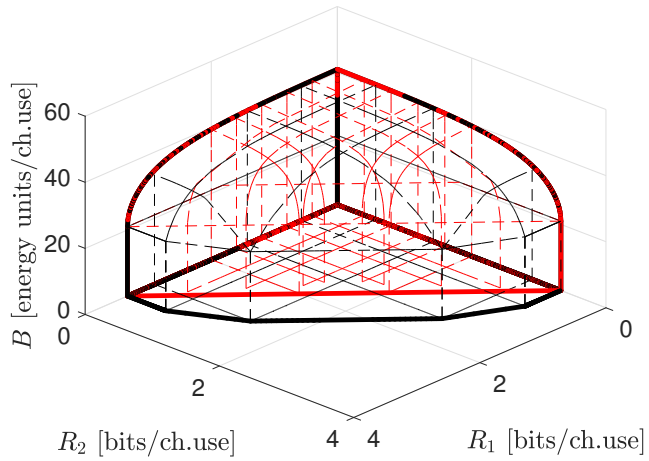

Fig. 2. 3-D superposition of $\underline{\mathcal{E}}$ and $\overline{\mathcal{E}}$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

channel-output feedback can be twice the energy rate achievable in the two-user G-IC without feedback. That is,

$$1 < \frac{\max \underline{\mathcal{B}}_{\mathrm{F}}}{\max \underline{\mathcal{B}}} \leqslant 2. \tag{22}$$

Any improvement beyond a factor of two is not feasible. That is,

$$1 < \frac{\max \overline{\mathcal{B}}_{\mathrm{F}}}{\max \overline{\mathcal{B}}} \leqslant 2. \tag{23}$$

The proof of Proposition 1 is presented in [24]. The main conclusion from Proposition 1 is that channel-output feedback can at most double the energy rate in the G-IC. Note that a similar observation is made in the case of the Gaussian multiple access chanel [1].

A tighter upper bound on the fractions $\frac{\max \underline{\mathcal{B}}_{\mathrm{F}}}{\max \underline{\mathcal{B}}}$ and $\frac{\max \overline{\mathcal{B}}_{\mathrm{F}}}{\max \overline{\mathcal{B}}}$ can be obtained by considering the exact values of σ_3^2, SNR_{31} and SNR_{32}. In the next section, some numerical examples are presented.

V. Numerical Analysis

Consider the two-user G-ICs with and without channel-output feedback depicted in Figure 1(a) and Figure 1(b) with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$. The corresponding achievable region $\underline{\mathcal{E}}$ and converse region $\overline{\mathcal{E}}$ are shown in Figure 2. In the case with feedback, the corresponding achievable region $\underline{\mathcal{E}}^{\mathrm{F}}$ and converse region $\overline{\mathcal{E}}^{\mathrm{F}}$ are shown in Figure 3. Note the strict inclusions $\underline{\mathcal{E}} \subset \overline{\mathcal{E}}$ and $\underline{\mathcal{E}}^{\mathrm{F}} \subset \overline{\mathcal{E}}^{\mathrm{F}}$ (Definition 3). Note also that for all $B \leqslant 21$ energy units, the set of triplets $(R_1, R_2, B) \in \underline{\mathcal{E}}^{\mathrm{F}}$ and the set of triplets $(R_1, R_2, B) \in \overline{\mathcal{E}}^{\mathrm{F}}$ are prisms whose bases correspond to the inner and outer regions approximating the information capacity region. For all $B > 21$, the trade-off between information transmission rates and the energy transmission rate becomes evident as both regions $\underline{\mathcal{E}}^{\mathrm{F}}$ and $\overline{\mathcal{E}}^{\mathrm{F}}$ monotonically shrink when B increases (Remark 3). The same observation can be made for the case

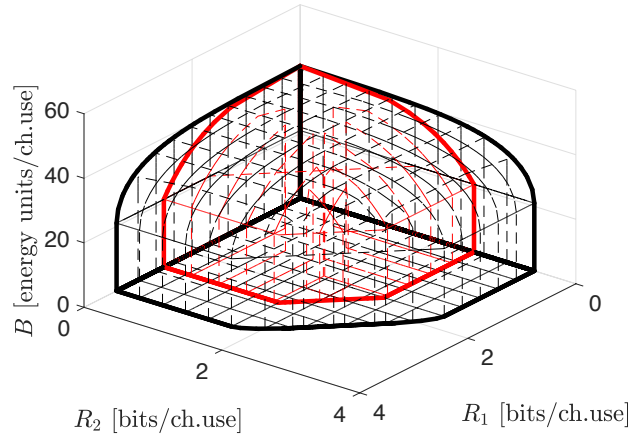

Fig. 3. 3-D superposition of $\underline{\mathcal{E}}^{\mathrm{F}}$ and $\overline{\mathcal{E}}^{\mathrm{F}}$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB and $\sigma_3^2 = 1$.

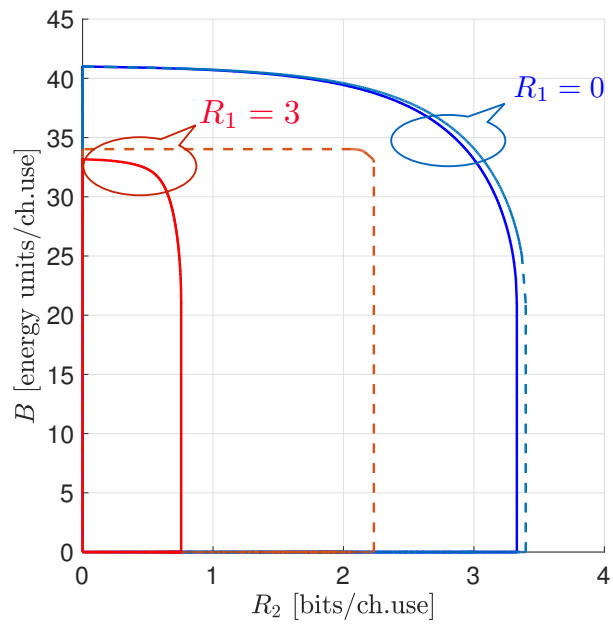

Fig. 4. Convex hull of pairs (R_2, B) that are in the sets $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \overline{\mathcal{E}}\}$ (solid line) and $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \overline{\mathcal{E}}^{\mathrm{F}}\}$ (dashed line), with $r_1 \in \{0, 3\}$. Parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

without feedback. Figure 4 shows the pairs (R_2, B) that are in the sets $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \overline{\mathcal{E}}\}$ (solid line) and $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \overline{\mathcal{E}}^{\mathrm{F}}\}$ (dashed line), with $r_1 = 0$ and $r_1 = 3$. Note that thanks to feedback, the information rate R_2 can be increased by one bit per channel use while keeping both the information rate R_1 and the energy rate B invariant.

Figure 5 shows the set of pairs (R_1, R_2) that are in the sets $\{(R_1, R_2) \in \mathbb{R}_+^2 : (R_1, R_2, b) \in \overline{\mathcal{E}}\}$ (solid line) and $\{(R_1, R_2) \in \mathbb{R}_+^2 : (R_1, R_2, b) \in \overline{\mathcal{E}}^{\mathrm{F}}\}$ (dashed line), with $b = 21$ and $b = 35$. Note that thanks to feedback, both the information rates R_1 and R_2 can be increased more than half a bit per channel use while keeping the energy rate B constant.

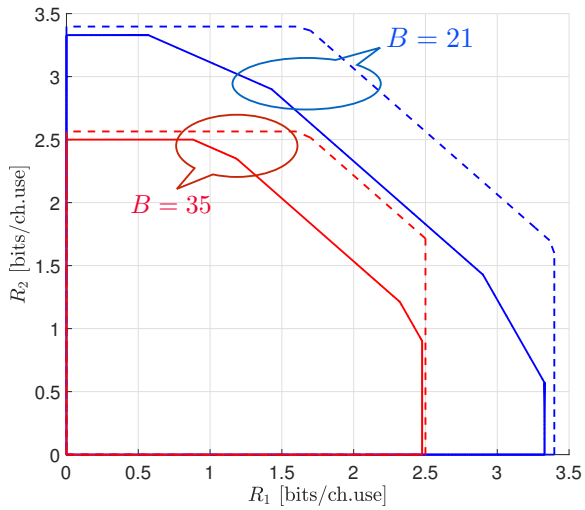

Fig. 5. Convex hull of pairs (R_1, R_2) that are in the sets $\{(R_1, R_2) \in \mathbb{R}_+^2 : (R_1, R_2, b) \in \overline{\mathcal{E}}\}$ (solid line) and $\{(R_1, R_2) \in \mathbb{R}_+^2 : (R_1, R_2, b) \in \overline{\mathcal{E}}^{\mathrm{F}}\}$ (dashed line), with $b \in \{21, 35\}$. Parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

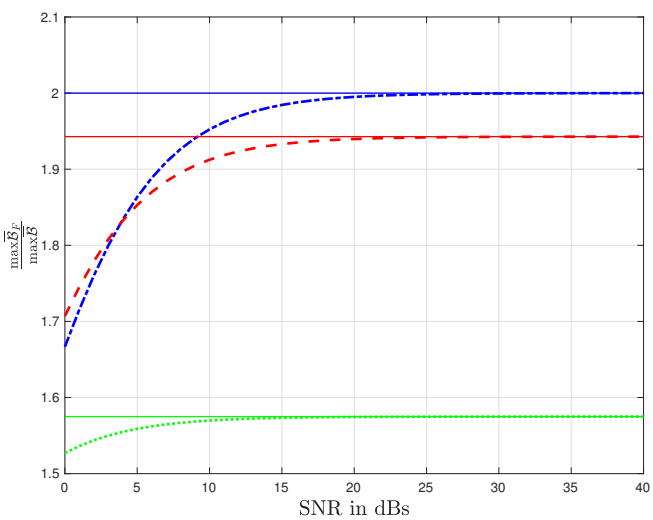

Fig. 6. The ratio $\frac{\max \overline{\mathcal{B}}_{\mathrm{F}}}{\max \overline{\mathcal{B}}}$ for different ratios of SNR at the EH, i.e., $\mathrm{SNR}_{31} = \mathrm{SNR}_{32} = \mathrm{SNR}$; $\frac{\mathrm{SNR}_{31}}{2} = \mathrm{SNR}_{32} = \mathrm{SNR}$; and $\frac{\mathrm{SNR}_{31}}{10} = \mathrm{SNR}_{32} = \mathrm{SNR}$, respectively. Parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

Figure 6 shows the ratio $\frac{\max \overline{\mathcal{B}}_{\mathrm{F}}}{\max \overline{\mathcal{B}}}$ for different ratios of SNR at the EH, i.e., $\mathrm{SNR}_{31} = \mathrm{SNR}_{32} = \mathrm{SNR}$; $\frac{\mathrm{SNR}_{31}}{2} = \mathrm{SNR}_{32} = \mathrm{SNR}$; and $\frac{\mathrm{SNR}_{31}}{10} = \mathrm{SNR}_{32} = \mathrm{SNR}$, respectively. Note that the upper bound in Proposition 1 is tight in the case of the symmetric case.

VI. Conclusions

In this paper, the information-energy capacity regions of the two-user Gaussian interference channel with and without perfect channel output feedback have been approximated by two regions, i.e., an achievable region and a converse region. When the energy transmission rate is normalized by the maximum energy rate, the approximation of these information-energy capacity regions is within a constant gap. In the proof of achievability, the key idea is the use of power-splitting between two signal components: an information-carrying component and a no-information component. Random coding arguments are used for the case of the information-carrying component, whereas a deterministic sequence known by all transmitters and receivers is used for the no-information component. The proof of the converse of the information rates follows along the same lines of the case in which only information is transmitted. The difference stems from lifting the constraints on the mean of the channel input signals. The proof of converse of the energy rate uses Markov's concentration inequality.

The results presented in this paper are a first step in the study of the fundamental limits of simultaneous information and energy transmission, nonetheless many questions are left open. On the one hand, there exits sufficient evidence that the use of multiple antennas at either the transmitters or the receivers enhances the energy rate [42]. However, very little is known from the perspectives of fundamental limits. On the other hand, an interesting question is about the degradation of the energy rates due to noisy feedback or rate-limited feedback. Similarly, another interesting question is about the benefits of other topologies of feedback, i.e., feedback from the receivers to both transmitters.

Appendix A
Proof of Theorem 4

The proof of Theorem 4 is divided into two parts. The first part consists of the proof of (20a)-(20g); and the second part consists of the proof of (20g).

A. *Proof of* (20a)-(20g)

Codebook Generation: Fix a strictly positive joint probability distribution:

$$P_{VUU_1U_2S_1S_2}(v, u, u_1, u_2, s_1, s_2) = P_V(v)P_{U|V}(u|v) \\ P_{U_1|UV}(u_1|u, v)P_{U_2|U,V}(u_2|u, v)P_{S_1|UU_1V}(x_1|u, u_1, v) \\ P_{S_2|UU_2V}(s_2|u, u_2, v), \tag{24}$$

for all $(v, u, u_1, u_2, s_1, s_2) \in (\mathcal{X}_1 \cap \mathcal{X}_2)^2 \times (\mathcal{X}_1 \cap \mathcal{X}_2)^2$. Let $R_E, R_{1,C}$, $R_{2,C}$, $R_{1,P}$ and $R_{2,P}$ be non-negative real numbers. For transmitter 1, generate 2^{NR_E} i.i.d N-length codewords $\boldsymbol{v}(\omega) = (v_1(\omega), \ldots, v_N(\omega))$ according to

$$P_{\boldsymbol{V}}(\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_V(v_m(\omega)), \tag{25}$$

with $\omega \in \{1, 2, \ldots, 2^{NR_E}\}$. For each codeword $\boldsymbol{v}(\omega)$, generate $2^{N(R_{1,C}+R_{2,C})}$ i.i.d. N-length codewords $\boldsymbol{u}(\omega, s, r) = (u_1(\omega, s, r), \ldots, u_N(\omega, s, r))$ according to

$$P_{\boldsymbol{U}|\boldsymbol{V}}(\boldsymbol{u}(\omega, s, r)|\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{U|V}(u_m(\omega, s, r)|v_m(\omega)), \tag{26}$$

with $s \in \{1, \ldots, 2^{NR_{1,C}}\}$ and $r \in \{1, \ldots, 2^{NR_{2,C}}\}$. For transmitter 1, for each codeword $\boldsymbol{u}(\omega, s, r)$, generate $2^{NR_{1,C}}$ i.i.d.

N-length codewords $\boldsymbol{u}_1(\omega,s,r,k) = (u_{1,1}(\omega,s,r,k),\ldots, u_{1,N}(\omega,s,r,k))$ according to

$$P_{\boldsymbol{U_1}|\boldsymbol{UV}}\Big(\boldsymbol{u}_1(\omega,s,r,k)|\boldsymbol{u}(s,r),\boldsymbol{v}(\omega)\Big) = \prod_{m=1}^{N} P_{U_1|UV}\Big(u_{1,m}(\omega,s,r,k)|u_m(\omega,s,r),v_m(\omega)\Big), \quad (27)$$

with $k \in \{1,\ldots,2^{NR_{1,C}}\}$. For each tuple of codewords $(\boldsymbol{v}(\omega),\boldsymbol{u}(\omega,s,r),\boldsymbol{u}_1(\omega,s,r,k))$, generate $2^{NR_{1,P}}$ i.i.d. N-length codewords $\boldsymbol{s}_1(\omega,s,r,k,l) = (s_{1,1}(\omega,s,r,k,l),\ldots,s_{1,N}(\omega,s,r,k,l))$ according to

$$P_{\boldsymbol{S_1}|\boldsymbol{U_1UV}}\Big(\boldsymbol{s}_1(\omega,s,r,k,l)|\boldsymbol{u}_1(\omega,s,r,k),\boldsymbol{u}(\omega,s,r),\boldsymbol{v}(\omega)\Big) = \prod_{m=1}^{N} P_{\boldsymbol{S}_1|\boldsymbol{U}_1\boldsymbol{UV}}\Big(s_{1,m}(\omega,s,r,k,l)|u_{1,m}(\omega,s,r,k),u_m(\omega,s,r), v_m(\omega)\Big), \quad (28)$$

with $l \in \{1,\ldots,2^{NR_{1,P}}\}$.

For encoder 2, for each codeword $\boldsymbol{u}(\omega,s,r)$, generate $2^{NR_{2,C}}$ i.i.d. N-length codewords $\boldsymbol{u}_2(\omega,s,r,q) = (u_{2,1}(\omega,s,r,q),\ldots,u_{2,1}(\omega,s,r,q))$ according to

$$P_{\boldsymbol{U_2}|\boldsymbol{UV}}(\boldsymbol{u}_2(\omega,s,r,q)|\boldsymbol{u}(\omega,s,r),\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{U_2|UV}\Big(u_{2,m}(\omega,s,r,q)|u_m(\omega,s,r), v_m(\omega)\Big), \quad (29)$$

with $q \in \{1,\ldots,2^{NR_{2,C}}\}$. For each tuple of codewords $(\boldsymbol{v}(\omega),\boldsymbol{u}(\omega,s,r),\boldsymbol{u}_2(\omega,s,r,q))$, generate $2^{NR_{2,P}}$ i.i.d. N-length codewords $\boldsymbol{s}_2(\omega,s,r,q,z) = (s_{2,1}(\omega,s,r,q,z),\ldots,s_{2,N}(\omega,s,r,q,z))$ according to

$$P_{\boldsymbol{S}_2|\boldsymbol{U}_2\boldsymbol{UV}}(\boldsymbol{s}_2(\omega,s,r,q,z)|\boldsymbol{u}_2(\omega,s,r,q),\boldsymbol{u}(\omega,s,r),\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{S_2|U_2UV}(s_{2,m}\Big(\omega,s,r,q,z)|u_{2,m}(\omega,s,r,q),u_m(\omega,s,r), v_m(\omega)\Big), \quad (30)$$

with $z \in \{1,\ldots,2^{NR_{2,P}}\}$.

Encoding: Let $W_i^{(t)}$ be represented by the message index $W_{i,C}^{(t)} \in \{1,2,\ldots,2^{NR_{i,C}}\}$ and the message index $W_{i,P}^{(t)} \in \{1,2,\ldots,2^{NR_{i,P}}\}$. The message index $W_{i,P}^{(t)}$ must be reliably decoded at receiver i and the message index $\Omega^{(t)}$ is known by both transmitters and receivers. The index $W_{i,C}^{(t-1)}$ must be reliably decoded by both receivers and transmitter j (via feedback). Consider Markov encoding over T blocks. At encoding step t, with $t \in \{1,2,\ldots,T\}$, transmitter 1 sends the codeword $\boldsymbol{x}_1^{(t)} = \theta_1\Big(\boldsymbol{v}(\Omega^{(t)},\boldsymbol{u}\left(\Omega^{(t)},W_{1,C}^{(t-1)},W_{2,C}^{(t-1)}\right), \boldsymbol{u}_1\left(\Omega^{(t)},W_{1,C}^{(t-1)},W_{2,C}^{(t-1)},W_{1,C}^{(t)}\right), \boldsymbol{s}_1\big(\Omega^{(t)},W_{1,C}^{(t-1)},W_{2,C}^{(t-1)}, W_{1,C}^{(t)},W_{1,P}^{(t)}\big)\Big)$, where $\theta_1 : \mathcal{X}^N \times (\mathcal{X}_1 \cup \mathcal{X}_2)^N \times \mathcal{X}_1 \times \mathcal{X}_1^N \rightarrow \mathcal{X}_1^N$ is a function that transforms the codewords $\boldsymbol{v}(\Omega^{(t)})$, $\boldsymbol{u}\Big(\Omega^{(t)},W_{1,C}^{(t-1)},W_{2,C}^{(t-1)}\Big)$, $\boldsymbol{u}_1\left(\Omega^{(t)},W_{1,C}^{(t-1)},W_{2,C}^{(t-1)},W_{1,C}^{(t)}\right)$, and $\boldsymbol{s}_1\big(\Omega^{(t)},W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)},W_{1,C}^{(t)},W_{1,P}^{(t)}\big)$ into the N-dimensional vector $\boldsymbol{x}_1^{(t)}$. The indices $W_{1,C}^{(0)} = W_{1,C}^{(T)} = s^*$ and $W_{2,C}^{(0)} = W_{2,C}^{(T)} = r^*$, and the pair $(s^*,r^*) \in \{1,2,\ldots,2^{NR_{1,C}}\} \times \{1,2,\ldots,2^{NR_{2,C}}\}$ are pre-defined and known by both receivers and transmitters. Transmitter 2 follows a similar encoding scheme.

Decoding: Both receivers decode their message indices at the end of block T in a backward decoding fashion. At each decoding step t, with $t \in \{1,2,\ldots,T\}$, receiver 1 obtains the indices $\left(\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)}\ ,\widehat{W}_{1,P}^{(T-(t-1))}\right)$ from the channel output $\boldsymbol{y}_1$.

The tuple $\left(\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},\widehat{W}_{1,P}^{(T-(t-1))}\right)$ is the unique tuple that satisfy:

$$\Big((\boldsymbol{v}(\Omega^{(t)}),\boldsymbol{u}\left(\Omega^{(t)},\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)}\right), \boldsymbol{u}_1\left(\Omega^{(t)},\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},W_{1,C}^{(T-(t-1))}\right), \boldsymbol{s}_1\left(\Omega^{(t)},\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},W_{1,C}^{(T-(t-1))},\widehat{W}_{1,P}^{(T-(t-1))}\right), \boldsymbol{u}_2\left(\Omega^{(t)},\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},W_{2,C}^{(T-(t-1))}\right),\boldsymbol{y}_1^{(T-(t-1))}\Big) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}, \quad (31)$$

where $W_{1,C}^{(T-(t-1))}$ and $W_{2,C}^{(T-(t-1))}$ are assumed to be perfectly decoded in the previous decoding step $t-1$. The set $\mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}$ represent the set of jointly typical N-length sequences of the random variables V,U,U_1,S_1,U_2, and Y_1, with $\epsilon > 0$. Finally, receiver 2 follows a similar decoding scheme.

Probability of Error Analysis: An error might occur during encoding step t at transmitter 1 if the index $\widehat{W}_{2,C}^{(t-1)}$ is not correctly decoded. Without any loss of generality, let $W_{2,C}^{(t-1)} = 1$ and $\widehat{W}_{2,C}^{(t-1)} = k$. Define the event E_k that describes the case in which there exist another message index $k \neq 1$ that satisfy:

$$\Big(\boldsymbol{v}(\Omega^{(t)}),\boldsymbol{u}\left(\Omega^{(t)},W_{1,C}^{(t-2)},W_{2,C}^{(t-2)}\right), \boldsymbol{u}_1\left(\Omega^{(t)},W_{1,C}^{(t-2)},W_{2,C}^{(t-2)},W_{1,C}^{(t-1)}\right), \boldsymbol{s}_1\big(\Omega^{(t)},W_{1,C}^{(t-2)},W_{2,C}^{(t-2)},W_{1,C}^{(t-1)},W_{1,P}^{(t-1)}\big), \boldsymbol{u}_2\left(\Omega^{(t)},W_{1,C}^{(t-2)},W_{2,C}^{(t-2)},k\right)\Big) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)},$$

with $t \in \{2,3,\ldots,T\}$ and $W_{2,C}^{(t-2)}$ is assumed to be perfectly decoded in the previous block $t-1$. Then, the probability of event E_k can be bounded as follows:

$$\Pr(E_k) \overset{(a)}{=} \Pr\Big[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{S}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}\Big] \leqslant 2^{-N(I(Y_1;U_2|X_1,V)-4\epsilon)},$$

where the probability operator $\Pr[.]$ in (a) applies with respect to a probability distribution $P_{\boldsymbol{VUU}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{Y}_1|V}P_{\boldsymbol{U}_2|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u}_2$

are independent from the output of the channel $\boldsymbol{y}_1$. The error probability becomes arbitrarily small (as N goes to infinity) if

$$R_{2c} \leq I(U_2,;Y_1|X_1,U,V). \tag{32}$$

An error might occur during the (backward) decoding step t if the indices $\widehat{W}_{1,C}^{(T-t)}$,$\widehat{W}_{2,C}^{(T-t)}$ or $\widehat{W}_{1,P}^{(T-(t-1)}$ are not decoded correctly given that the indices $W_{1,C}^{(T-(t-1))}$ and $W_{2,C}^{(T-(t-1))}$ were correctly decoded in the previous decoding step $t-1$. These errors might arise for two reasons: (i) there does not exist a tuple $\left(\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},\widehat{W}_{1,P}^{(T-(t-1))}\right)$ that satisfies (31), or (ii) there exist several tuples $\left(\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},\widehat{W}_{1,P}^{(T-(t-1))}\right)$ that simultaneously satisfy (31). From the asymptotic equipartition property [10], the probability of error due to (i) tends to zero when N grows to infinity. Consider the error due to (ii) and define the event E_{srl} that describes the case in which the codewords $\boldsymbol{v}(\Omega^{(t)}), \boldsymbol{u}\left(\Omega^{(t)},s,r\right), \boldsymbol{u}_1\left(\Omega^{(t)},s,r,W_{1,C}^{(T-(t-1))}\right)$, $\boldsymbol{s}_1\left(\Omega^{(t)},s,r,W_{1,C}^{(T-(t-1))},l\right)$, and $\boldsymbol{u}_2\left(\Omega^{(t)},s,r,W_{2,C}^{(T-(t-1))}\right)$ are jointly typical with $\boldsymbol{y}_1^{(T-(t-1))}$ during decoding step t. Assume now that the codeword to be decoded at decoding step t corresponds to the indices $(s,r,l)=(1,1,1)$. This is without loss of generality due to the symmetry of the code. Then, the probability of error due to (ii) during decoding step t, can be bounded as follows:

$$\begin{aligned}
&\Pr\left[\bigcup_{(s,r,l)\neq(1,1,1)} E_{srl}\right] \leq 2^{N(R_{1C}+R_{2C}+R_{2P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)}\\
&+2^{N(R_{1C}+R_{2C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)}\\
&+2^{N(R_{1C}+R_{1P}-I(U,X_1,U_2;Y_1V)+4\epsilon)}\\
&+2^{N(R_{1C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)}\\
&+2^{N(R_{2C}+R_{1P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)}\\
&2^{N(R_{2C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)}\\
&+2^{N(R_{1P}-I(X_1;Y_1|U,U_1,U_2,V)+4\epsilon)}.
\end{aligned} \tag{33}$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence in general, from (33), reliable decoding holds under the following conditions:

$$\begin{aligned}
R_{2C} &\leq I(U_2;Y_1|X_1,U,V), && (34a)\\
R_{1P} &\leq I(X_1;Y_1|U_1,U_2,U,V), && (34b)\\
R_{1C}+R_{2C}+R_{1P} &\leq I(U,X_1,U_2;Y_1|V), && (34c)\\
R_{1C} &\leq I(U_1;Y_2|X_2,U,V), && (34d)\\
R_{2P} &\leq I(X_2;Y_2|U_1,U_2,U,V), && (34e)\\
R_{1C}+R_{2C}+R_{2P} &\leq I(U,X_2,U_1;Y_2|V). && (34f)
\end{aligned}$$

The proof continues by applying a Fourrier-Motzkin elimination on (34), which yields,

$$\begin{aligned}
R_1 &\leq I(U,X_1,U_2;Y_1|V), && (35a)\\
R_1 &\leq I(U_1;Y_2|U,X_2,V)+I(X_1;Y_1|U_1,U_2,U,V), && (35b)\\
R_2 &\leq I(U,X_2,U_1;Y_2|V), && (35c)\\
R_2 &\leq I(U_2;Y_1|U,X_1,V)+I(X_2;Y_2|U_1,U_2,U,V), && (35d)\\
R_1+R_2 &\leq I(X_1;Y_1|U_1,U_2,U,V)+I(V,U_2,X_1;Y_1), && (35e)\\
R_1+R_2 &\leq I(X_2;Y_2|U_1,U_2,U,V)+I(V,U_2,X_1;Y_1). && (35f)
\end{aligned}$$

The proof of Theorem 4 continues as follows. Let $k\in\{0,1\}$ be fixed and consider that the distribution in (24) is the following Gaussian input distribution:

$$\begin{aligned}
&V\sim\mathcal{N}(0,1);\ U\sim\mathcal{N}(0,\rho);\ U_k\sim\mathcal{N}(0,\lambda_{kc});\\
&\text{and } S_k\sim\mathcal{N}(0,\lambda_{kp}),
\end{aligned} \tag{36}$$

where $(\rho,\lambda_{kp},\lambda_{kc},\lambda_{ke})\in[0,1]^4$ and $\rho+\lambda_{kp}+\lambda_{kc}+\lambda_{ke}\leq 1$. The input symbol is generated deterministically given the mutually independent random variables V, U, U_k, and X_{kp} as follows:

$$X_k=\sqrt{P_k}U+\sqrt{P_k}X_{kp}+\sqrt{P_k}U_k+\sqrt{\lambda_{ke}P_k}V. \tag{37}$$

By symmetry, it suffices to prove (20a), (20b) and (20e). The choice of the Gaussian input distribution in (35) yields:

$$\begin{aligned}
&I(U,X_1,U_2;Y_1|V)\\
&=\frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1+2\rho\sqrt{\mathrm{SNR}_1\mathrm{INR}_1}}{1+\lambda_{2p}\mathrm{INR}_1}\right),
\end{aligned} \tag{38a}$$

which proves (20a). With the same power setting in (36), the following holds

$$I(U_1;Y_2|U,X_2,V)=\frac{1}{2}\log\left(\frac{1+(1-(\rho+\lambda_{1e}))\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{38b}$$

and

$$I(X_1;Y_1|U,U_1,U_2,V)=\frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+\lambda_{2p}\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right). \tag{38c}$$

This proves (20b). Finally, using (38b) and (38c), yields the proof of (20e).

B. Proof of 20g

The choice of the channel input in (36) guarantees that the random variables $Y_{3,1},\ldots,Y_{3,n}$ are independently and identically distributed. For all $n\in\{1,2,\ldots,N\}$, $Y_{3,n}$ follows a zero-mean Gaussian distribution with variance $\bar{B}$ given by

$$\begin{aligned}
\bar{B}=\mathbb{E}\left[Y_{3,n}^2\right]\leq h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}\sqrt{P_1P_2}\\
(\rho+\sqrt{\lambda_{1e}\lambda_{2e}})+\sigma_3^2.
\end{aligned}$$

By the weak law of large numbers, it holds that

$$\lim_{N\to\infty}\Pr\left[B^{(N)}<\bar{B}\right]=0. \tag{39}$$

From (39), it holds that for any energy B which satisfies $0<B\leq\bar{B}$, it holds that

$$\lim_{N\to\infty}\Pr\left[B^{(N)}<B\right]=0. \tag{40}$$

This proves (20g) and completes the proof of Theorem 4.

Appendix B
Proof of Theorem 5

Fix an information-energy rate triplet (R_1, R_2, B) achievable with a given coding scheme (Definition 1). Denote by $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ the channel inputs resulting from transmitting the independent messages (W_1, Ω) and (W_2, Ω) using such coding scheme. Denote by $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ the corresponding channel outputs. Define the following random variables:

$$S_1 = h_{21}X_1 + Z_2 \text{ and} \tag{41}$$

$$S_2 = h_{12}X_2 + Z_1, \tag{42}$$

where, Z_1 and Z_2 are real Gaussian random variables independent of each other with zero means and variances σ_1^2 and σ_2^2, respectively. Using assumption (13a), Fano's inequality and following similar steps as in [30], it can be shown that the information rates R_1 and R_2 must satisfy the following inequalities

$$NR_1 \leq \sum_{n=1}^{N}[h(Y_{1,n}) - h(Z_{1,n})] + \mathrm{o}(N), \tag{43a}$$

$$NR_1 \leq \sum_{n=1}^{N}[h(Y_{2,n}|X_{2,n}) - h(Z_{2,n}) + h(Y_{1,n}|X_{2,n}, S_{1,n}) - h(Z_{1,n})] + \mathrm{o}(N), \tag{43b}$$

$$NR_2 \leq \sum_{n=1}^{N}[h(Y_{2,n}) - h(Z_{2,n})] + \mathrm{o}(N), \tag{43c}$$

$$NR_2 \leq \sum_{n=1}^{N}[h(Y_{1,n}|X_{1,n}) - h(Z_{1,n}) + h(Y_{2,n}|X_{1,n}, S_{2,n}) - h(Z_{2,n})] + \mathrm{o}(N), \tag{43d}$$

$$N(R_1+R_2) \leq \sum_{n=1}^{N}[h(Y_{1,n}|S_{1,n}, X_{2,n}) - h(Z_{1,n}) + h(Y_{2,n}) - h(Z_{2,n})] + \mathrm{o}(N), \tag{43e}$$

$$N(R_1+R_2) \leq \sum_{n=1}^{N}[h(Y_{2,n}|S_{2,n}, X_{1,n}) - h(Z_{2,n}) + h(Y_{1,n}) - h(Z_{1,n})] + \mathrm{o}(N), \tag{43f}$$

where $\frac{o(N)}{N}$ tends to zero as N tends to infinity. Using assumption (13b), for any $\eta > 0$ there exists $N_0(\eta)$ such that for any $n \geq N_0(\eta)$ it holds that

$$\Pr\left[B^{(N)} \geqslant B\right] \geqslant 1 - \eta. \tag{44}$$

Using Markov's inequality, the probability in (44) can be upper-bounded as follows:

$$B\Pr\left[B^{(N)}(\boldsymbol{Y}_3) \geqslant B\right] \leqslant \mathbb{E}[B^{(N)}]. \tag{45}$$

Combining (44) and (45) yields: $B(1-\eta) \leqslant \mathbb{E}[B^{(N)}(\boldsymbol{Y}_3)]$, which can be written as

$$(B - \delta^{(N)}) \leqslant \mathbb{E}[B^{(N)}(\boldsymbol{Y}_3)], \tag{46}$$

for some $\delta^{(N)} > 0$ (for sufficiently large N). The bounds in (43) and (46) are evaluated assuming that the channel inputs $X_{1,n}$ and $X_{2,n}$ are arbitrary correlated random variables, such that $\mu_{i,n} \triangleq \mathbb{E}[X_{i,n}]$; $\gamma_{i,n}^2 \triangleq \mathrm{Var}[X_{i,n}]$; and $\lambda_n \triangleq \mathrm{Cov}[X_{1,n}X_{2,n}]$, for all $n \in \{1, \ldots, N\}$ and for all $i \in \{1, 2\}$. The input sequence must satisfy the input power constraint which can be written, for $i \in \{1, 2\}$, as follows:

$$\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[X_{i,n}^2\right] = \left(\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\right) + \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{i,n}^2\right) \leqslant P_i. \tag{47}$$

Using this notation, the following holds:

$$h(Y_{1,n}) \leqslant \frac{1}{2}\log\left(2\pi e(h_{11}^2\gamma_{1,n}^2 + h_{12}^2\gamma_{2,n}^2 + 2h_{11}h_{12}\lambda_n + \sigma_1^2)\right), \tag{48a}$$

$$h(Y_{2,n}|X_{2,n}) \leqslant \frac{1}{2}\log\left(1 + \gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_2^2}\right) + \frac{1}{2}\log(2\pi e\sigma_2^2), \text{ and} \tag{48b}$$

$$h(Y_{1,n}|X_{2,n}, S_{1,n}) \leqslant \frac{1}{2}\log\left(1 + \frac{\gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_2^2}}\right) + \frac{1}{2}\log(2\pi e\sigma_1^2). \tag{48c}$$

Finally, the bounds in (43) can be rewritten as follows:

$$NR_1 \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(\frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2} + \frac{2h_{11}h_{12}\lambda_n}{\sigma_1^2} + 1\right) + \mathrm{o}(N), \tag{49a}$$

$$NR_1 \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \gamma_{1,n}^2(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2})\frac{h_{21}^2}{\sigma_2^2}\right) + \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_2^2}}\right) + \mathrm{o}(N), \tag{49b}$$

$$NR_2 \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2} + \frac{2h_{22}h_{21}\lambda_n}{\sigma_2^2} + 1\right) + \mathrm{o}(N), \tag{49c}$$

$$NR_2 \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \gamma_{2,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_1^2}\right) + \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\gamma_{2,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{22}^2}{\sigma_2^2}}{1 + \gamma_{2,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{12}^2}{\sigma_1^2}}\right) + \mathrm{o}(N), \tag{49d}$$

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2} + \frac{2h_{22}h_{21}\lambda_n}{\sigma_2^2} + 1\right) + \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_2^2}}\right) + \mathrm{o}(N), \tag{49e}$$

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} \frac{1}{2} \log \left(\frac{h_{11}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{12}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{2 h_{11} h_{12} \lambda_n}{\sigma_1^2} + 1 \right)$$

$$+ \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \frac{\gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{22}^2}{\sigma_2^2}}{1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{12}^2}{\sigma_1^2}} \right) + \mathrm{o}(N). \tag{49f}$$

The average received energy rate is given by

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} Y_{3,n}^2\right] = \sigma_3^2 + h_{31}^2 \left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2 + \mu_{1,n}^2)\right)$$
$$+ h_{32}^2 \left(\frac{1}{N}\sum_{n=N}^{n}(\gamma_{2,n}^2 + \mu_{2,n}^2)\right)$$
$$+ 2 h_{31} h_{32} \left(\frac{1}{N}\sum_{n=1}^{N}(\lambda_n + \mu_{1,n}\mu_{2,n})\right). \tag{50}$$

Using Cauchy-Schwarz inequality, the expected value of the energy rate in (50) can be upper-bounded as follows:

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} Y_{3,n}^2\right] \leqslant \sigma_3^2 + h_{31}^2 \left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2 + \mu_{1,n}^2)\right)$$
$$+ h_{32}^2 \left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2 + \mu_{2,n}^2)\right)$$
$$+ 2 h_{31} h_{32} \Bigg(\left|\frac{1}{N}\sum_{n=1}^{N}\lambda_n\right|$$
$$+ \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}^2\right)^{1/2} \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{2,n}^2\right)^{1/2} \Bigg). \tag{51}$$

Combining (46) and (51) yields the following upper-bound on the energy rate B:

$$B \leqslant \sigma_3^2 + h_{31}^2 \left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2 + \mu_{1,n}^2)\right) \tag{52}$$
$$+ h_{32}^2 \left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2 + \mu_{2,n}^2)\right)$$
$$+ 2 h_{31} h_{32} \Bigg(\left|\frac{1}{N}\sum_{n=1}^{N}\lambda_n\right|$$
$$+ \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}^2\right)^{1/2} \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{2,n}^2\right)^{1/2} \Bigg) + \delta_N. \tag{53}$$

In order to obtain a single-letterization of the upper-bound given by constraints (49) and (52), define also $\mu_i^2 \triangleq \frac{1}{N}\sum_{n=1}^{N}\mu_{i,n}^2$; $\gamma_i^2 \triangleq \frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2$; $\beta_i \triangleq \frac{\gamma_i^2}{P_i} = \frac{P_i - \mu_i^2}{P_i}$; and $\rho \triangleq \frac{\left(\frac{1}{N}\sum_{n=1}^{N}\lambda_n\right)}{|\gamma_1||\gamma_2|}$. With this notation, the input power constraint in (47) can be rewritten as

$$\gamma_i^2 + \mu_i^2 \leqslant P_i, \quad i \in \{1,2\}. \tag{54}$$

The following inequality is proved in [24].

$$\frac{1}{N}\sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\gamma_{i,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ii}^2}{\sigma_i^2}}{1 + \gamma_{i,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ji}^2}{\sigma_j^2}}\right) \leqslant$$
$$\frac{1}{2}\log\left(1 + \frac{\gamma_i^2\left(1-\rho^2\right)\frac{h_{ii}^2}{\sigma_i^2}}{1 + \gamma_i^2\left(1-\rho^2\right)\frac{h_{ji}^2}{\sigma_j^2}}\right). \tag{55}$$

By the concavity of the mutual information, applying Jensen's inequality and (55) in the bounds (49) yields,

$$R_1 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\gamma_1^2}{\sigma_1^2} + \frac{h_{12}^2\gamma_2^2}{\sigma_1^2} + 2\rho\sqrt{\frac{h_{11}^2 h_{12}^2 \gamma_1^2\gamma_2^2}{\sigma_1^4}}\right), \tag{56a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1 + \frac{h_{21}^2}{\sigma_2^2}\gamma_1^2\left(1-\rho^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + \frac{\gamma_1^2(1-\rho^2)\frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_1^2(1-\rho^2)\frac{h_{21}^2}{\sigma_2^2}}\right), \tag{56b}$$

$$R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\gamma_2^2}{\sigma_2^2} + \frac{h_{21}^2\sigma_1^2}{\sigma_2^2} + 2\rho\sqrt{\frac{h_{22}h_{21}\gamma_1^2\gamma_2^2}{\sigma_2^4}}\right), \tag{56c}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1 + \frac{h_{12}^2}{\sigma_1^2}\gamma_2^2\left(1-\rho^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + \frac{\gamma_2^2\left(1-\rho^2\right)\frac{h_{22}^2}{\sigma_2^2}}{1 + \gamma_2^2\left(1-\rho^2\right)\frac{h_{12}^2}{\sigma_1^2}}\right), \tag{56d}$$

$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\gamma_2^2}{\sigma_2^2} + \frac{h_{21}^2\gamma_1^2}{\sigma_2^2} + 2\rho\sqrt{\frac{h_{22}^2 h_{21}^2\gamma_1^2\gamma_2^2}{\sigma_2^4}}\right)$$
$$+ \frac{1}{2}\log\left(1 + \frac{\gamma_1^2(1-\rho^2)\frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_1^2(1-\rho^2)\frac{h_{21}^2}{\sigma_2^2}}\right), \tag{56e}$$

$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\gamma_1^2}{\sigma_1^2} + \frac{h_{12}^2\gamma_2^2}{\sigma_1^2} + 2\rho\sqrt{\frac{h_{11}^2 h_{12}^2\gamma_1^2\gamma_2^2}{\sigma_1^4}}\right)$$
$$+ \frac{1}{2}\log\left(1 + \frac{\gamma_2^2(1-\rho^2)\frac{h_{22}^2}{\sigma_2^2}}{1 + \gamma_2^2(1-\rho^2)\frac{h_{12}^2}{\sigma_1^2}}\right). \tag{56f}$$

The upper-bound on the energy rate (52) satisfies :

$$B \leqslant \sigma_3^2 + h_{31}^2(\gamma_1^2 + \mu_1^2) + h_{32}^2(\gamma_2^2 + \mu_2^2)$$
$$+ 2 h_{21} h_{22}(|\rho|\,|\gamma_1||\gamma_2| + |\mu_1||\mu_2|). \tag{56g}$$

In the limit when N tends to infinity, any information-energy rate triplet $(R_1, R_2, B) \in \mathcal{E}^{\mathrm{F}}$ satisfies (56) for some γ_1, γ_2, μ_1, μ_2 according to (54) and for some $\rho \in [-1, 1]$. Let $\mathcal{R}^{\mathrm{F}}(\gamma_1, \gamma_2, \mu_1, \mu_2, \rho)$ denote the set of information-energy rate triplets satisfying (56), for some γ_1, γ_2, μ_1, μ_2 satisfying (54)

and for some $\rho \in [-1,1]$. Thus, it holds that

$$\mathcal{E}^{\mathrm{F}} \subseteq \bigcup_{\substack{0\leqslant\gamma_1^2+\mu_1^2\leqslant P_1\\ 0\leqslant\gamma_2^2+\mu_2^2\leqslant P_2\\ -1\leqslant\rho\leqslant 1}} \mathcal{R}^{\mathrm{F}}(\gamma_1^2,\gamma_2^2,\mu_1,\mu_2,\rho). \tag{57}$$

In this union, it suffices to consider $0 \leqslant \rho \leqslant 1$ because for any $-1 \leqslant \rho \leqslant 1$, $\mathcal{R}^{\mathrm{F}}(\gamma_1^2,\gamma_2^2,\mu_1^2,\mu_2^2,\rho) \subseteq \mathcal{R}^{\mathrm{F}}(\gamma_1^2,\gamma_2^2,\mu_1^2,\mu_2^2,|\rho|)$. Thus,

$$\mathcal{E}^{\mathrm{F}} \subseteq \bigcup_{\substack{0\leqslant\gamma_1^2+\mu_1^2\leqslant P_1\\ 0\leqslant\gamma_2^2+\mu_2^2\leqslant P_2\\ -1\leqslant\rho\leqslant 1}} \mathcal{R}^{\mathrm{F}}(\gamma_1^2,\gamma_2^2,\mu_1,\mu_2,\rho) \subseteq \bigcup_{\substack{\gamma_1^2+\mu_1^2=P_1\\ \gamma_2^2+\mu_2^2=P_2\\ 0\leqslant\rho\leqslant 1}} \mathcal{R}^{\mathrm{F}}(\gamma_1^2,\gamma_2^2,\mu_1,\mu_2,\rho). \tag{58}$$

Using the definition of β_i, any region $\mathcal{R}(\gamma_1,\gamma_2,\mu_1,\mu_2,\rho)$ in the union over all $(\mu_1,\mu_2,\gamma_1,\gamma_2)$ that satisfy $\gamma_1^2+\mu_1^2=P_1$, $\gamma_2^2+\mu_2^2=P_2$ and $0\leqslant\rho\leqslant 1$, can be rewritten as follows:

$$R_1 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\beta_1P_1}{\sigma_1^2}+\frac{h_{12}^2\beta_2P_2}{\sigma_1^2}+2\rho\sqrt{\frac{h_{11}^2h_{12}^2\beta_1\beta_2P_1P_2}{\sigma_1^4}}\right), \tag{59a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1+\frac{h_{21}^2}{\sigma_2^2}\beta_1P_1\left(1-\rho^2\right)\right) \tag{59b}$$

$$+\frac{1}{2}\log\left(1+\frac{\beta_1P_1(1-\rho^2)\frac{h_{11}^2}{\sigma_1^2}}{1+\beta_1P_1(1-\rho^2)\frac{h_{21}^2}{\sigma_2^2}}\right), \tag{59c}$$

$$R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\beta_2P_2}{\sigma_2^2}+\frac{h_{21}^2\beta_1P_1}{\sigma_2^2}+2\rho\sqrt{\frac{h_{22}h_{21}\beta_1\beta_2P_1P_2}{\sigma_2^4}}\right), \tag{59d}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\frac{h_{12}^2}{\sigma_1^2}\beta_2P_2\left(1-\rho^2\right)\right)$$

$$+\frac{1}{2}\log\left(1+\frac{\beta_2P_2\left(1-\rho^2\right)\frac{h_{22}^2}{\sigma_2^2}}{1+\beta_2P_2\left(1-\rho^2\right)\frac{h_{12}^2}{\sigma_1^2}}\right), \tag{59e}$$

$$R_1+R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\beta_2P_2}{\sigma_2^2}+\frac{h_{21}^2\beta_1P_1}{\sigma_2^2}+2\rho\sqrt{\frac{h_{22}^2h_{21}^2\beta_1\beta_2P_1P_2}{\sigma_2^4}}\right)$$

$$+\frac{1}{2}\log\left(1+\frac{\beta_1P_1\left(1-\rho^2\right)\frac{h_{11}^2}{\sigma_1^2}}{1+\beta_1P_1\left(1-\rho^2\right)\frac{h_{21}^2}{\sigma_2^2}}\right), \tag{59f}$$

$$R_1+R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\beta_1P_1}{\sigma_1^2}+\frac{h_{12}^2\beta_2P_2}{\sigma_1^2}+2\rho\sqrt{\frac{h_{11}^2h_{12}^2\beta_1\beta_2P_1P_2}{\sigma_1^4}}\right)$$

$$+\frac{1}{2}\log\left(1+\frac{\beta_2P_2\left(1-\rho^2\right)\frac{h_{22}^2}{\sigma_2^2}}{1+\beta_2P_2\left(1-\rho^2\right)\frac{h_{12}^2}{\sigma_1^2}}\right), \tag{59g}$$

$$B \leqslant \sigma_3^2+h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}(|\rho|\sqrt{\beta_1P_1\beta_2P_2}$$

$$+\sqrt{(1-\beta_1)(1-\beta_2)P_1P_2}), \tag{59h}$$

for some $(\beta_1,\beta_2)\in[0,1]^2$ and $\rho\in[0,1]$. Hence, using the definitions in (7) and (10), the region (59) contains all information-energy rate triplets (R_1,R_2,B) satisfying constraints (21), which completes the proof of Theorem 5.

References

[1] S. Belhadj Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback enhances simultaneous wireless information and energy transmission in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5244–5265, Aug. 2017.

[2] S. M. Belhadj Amor and S. M. Perlaza, "Fundamental limits of simultaneous energy and information transmission," in *Proc. 23rd International Symposium on Telecommunications*, Thessaloniki, Greece, May 2016.

[3] P. Bergmans, "Random coding theorems for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, Mar. 1973.

[4] R. A. Berry and D. N. C. Tse, "Shannon meets Nash on the interference channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2821–2836, May 2011.

[5] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities : A Nonasymptotic Theory of Independence*. New York, USA: Oxford University Press, Mar. 2013.

[6] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Sep. 1978.

[7] S. Chao, W.-C. Li, and T.-H. Chang, "Simultaneous information and energy transfer: A two-user MISO interference channel case," in *Proc. Signal Processing for Communications Symposium*, Anaheim, CA, USA, Dec. 2012, pp. 3862–3867.

[8] ——, "Wireless information and energy transfer in multi-antenna interference channel," *IEEE Transactions on Signal Processing*, vol. 62, no. 23, pp. 6249–6264, Dec. 2014.

[9] H.-F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On the Han-Kobayashi region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3188–3195, Jul. 2008.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley-Interscience Publication, 2005.

[11] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 27, no. 3, pp. 292–298, May 1981.

[12] Z. Ding, S. M. Perlaza, I. Esnaola, and H. V. Poor, "Power allocation strategies in energy harvesting wireless cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 846–860, Feb. 2014.

[13] P. Elias, A. Feinstein, and C. E. Shannon, "A note on the maximum flow through a network," *IEEE Trans. Inf. Theory*, vol. 2, no. 4, pp. 117–119, Dec. 1956.

[14] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[15] R. Fano, *Transmission of Information - A Statistical Theory of Communication*. Cambridge, MA: MIT Press, Mar. 1961.

[16] W. Feller, *An Introduction to Probability Theory and Its Application*, 2nd ed. New York, NY: John Wiley and Sons, 1971, vol. 2.

[17] A. M. Fouladgar; and O. Simeone, "On the transfer of information and energy in multi-user systems," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1733–1736, Nov. 2012.

[18] M. Gastpar, "Gaussian multiple-access channels under received-power constraints," in *Proc. IEEE Information Theory Workshop*, San Antonio, TX, USA, Mar. 2004, pp. 452–457.

[19] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer," in *Proc. IEEE International Symposium on Information Theory*, Austin, TX, USA, Jun. 2010, pp. 2363–2367.

[20] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, 1981.

[21] ——, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, pp. 49–60, Jan. 1981.

[22] N. Khalfet and S. M. Perlaza, "Simultaneous information and energy transmission in Gaussian interference channels with feedback," in *Proc. 55th Annual Allerton Conference on Communications, Control, and Computing*, Monticello, IL, USA, Oct. 2017.

[23] ——, "Simultaneous information and energy transmission in Gaussian interference channels," in *Proc. International Zurich Seminar on Information and Communication*, Zurich, Switzerland, Feb. 2018.

[24] ——, "Simultaneous information and energy transmission in the interference channel," INRIA Grenoble - Rhône-Alpes, Grenoble, France, Tech. Rep. 9102, Oct. 2017.

[25] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 104–110, Nov. 2014.

[26] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a two-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4210–4221, Aug. 2013.

[27] S. M. Perlaza, R. Tandon, H. V. Poor, and Z. Han, "Perfect output feedback in the two-user decentralized interference channel," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5441–5462, Oct. 2015.

[28] S. M. Perlaza, A. Tajer, and H. V. Poor, "Simultaneous energy and information transmission: A finite block-length analysis," in *Proc. 19th IEEE International Workshop on Signal Processing Advances in Wireless Communication*, Kalamata, Greece, Jun. 2018.

[29] P. Popovski, A. M. Fouladgar, and O. Simeone, "Interactive joint transfer of energy and information," *IEEE Transactions on Communications*, vol. 61, no. 5, pp. 2086–2097, May 2013.

[30] V. Quintero, S. M. Perlaza, I. Esnaola, and J.-M. Gorce, "Approximate capacity region of the two-user Gaussian interference channel with noisy channel-output feedback," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 5326 – 5358, Jul. 2018.

[31] D. Sacristan-Murga and D. Guenduez, "Joint energy and information transmission in a two-user symmetric interference channel," in *Proc. 20th European Wireless Conference*, Barcelona, Spain, 2014.

[32] K. G. Shenoy and V. Sharma, "Finite blocklength achievable rates for energy harvesting AWGN channels with infinite buffer," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, Jul. 2016.

[33] C. Suh and D. N. C. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

[34] N. Tesla, *Apparatus for transmitting electrical energy*. New York, NY: United States Patent Office, Dec. 1914, vol. US1119732 A.

[35] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy harvesting wireless communications: A review of recent advances," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[36] M. Varasteh, B. Rassouli, and B. Clerckx, "On capacity-achieving distributions over complex AWGN channels under nonlinear power constraints and their applications to SWIPT," *CoRR*, vol. abs/1712.01226, 2017. [Online]. Available: http://arxiv.org/abs/1712.01226

[37] ——, "Wireless information and power transfer over an AWGN channel: Nonlinearity and asymmetric Gaussian signaling," in *Proc. IEEE Information Theory Workshop*, Kaohsiung, Taiwan, Nov. 2017.

[38] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE International Symposium on Information Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1612–1616.

[39] ——, "On energy/information cross-layer architectures," in *Proc. IEEE International Symposium on Information Theory*, Cambridge, MA, USA, Jul. 2012, pp. 1356–1360.

[40] F. M. J. Willems, "Information theoretical results for multiple access channels," Ph.D. dissertation, Katholieke Universiteit, Department of Electrical Engineering, Leuven, Belgium, Oct. 1982.

[41] F. M. J. Willems and E. C. Van Der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 313–327, May. 1985.

[42] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1989–2001, May. 2013.