On Cooperation and Interference in the Weak Interference Regime

Daniel Zahavi, *Student Member, IEEE*, and Ron Dabora, *Senior Member, IEEE*

***Abstract*— Handling interference is one of the main challenges in the design of wireless networks. In this paper, we study the application of cooperation for interference management in the weak interference (WI) regime, focusing on the Z-interference channel with a causal relay (Z-ICR), in which the channel coefficients are subject to ergodic phase fading, all transmission powers are finite, and the relay is full-duplex. The phase fading model represents many practical communications systems in which the transmission path impairments mainly affect the phase of the signal, such as non-coherent wireless communications and fiber optic channels. In order to provide a comprehensive understanding of the benefits of cooperation in the WI regime, we characterize, for the first time, two major performance measures for the ergodic phase fading Z-ICR in the WI regime: the sum-rate capacity and the maximal generalized degrees-of-freedom (GDoF). In the capacity analysis, we obtain conditions on the channel coefficients, subject to which the sum-rate capacity of the ergodic phase fading Z-ICR is achieved by treating interference as noise at each receiver, and explicitly state the corresponding sum-rate capacity. In the GDoF analysis, we derive conditions on the exponents of the magnitudes of the channel coefficients, under which treating interference as noise achieves the maximal GDoF, which is explicitly characterized as well. It is shown that under certain conditions on the channel coefficients, *relaying strictly increases* both the sum-rate capacity and the maximal GDoF of the ergodic phase fading Z-interference channel in the WI regime. Our results demonstrate *for the first time* the gains from relaying in the presence of interference, *when interference is weak and the relay power is finite*, both in increasing the sum-rate capacity and in increasing the maximal GDoF, compared with the channel without a relay.**

***Index Terms*— Interference channel with a relay, weak interference, Z interference channel, channel capacity, generalized degrees-of-freedom.**

I. INTRODUCTION

THE interference channel (IC) [1] models communications scenarios in which two source-destination pairs communicate over a shared medium. The capacity region of the IC is generally unknown, but capacity characterizations exist for some special scenarios. For example, the capacity region of the IC with additive white Gaussian noise (AWGN), for the scenario in which the interference between the communicating pairs is very strong, was characterized in [2], and the capacity region for the case of strong interference (SI) was characterized in [3]. In both works it was shown that in order to achieve capacity, each receiver should *decode* both the interfering message as well as the desired message. Additional performance measures commonly used for characterizing the performance of ICs are the degrees-of-freedom (DoF) and the generalized DoF (GDoF). The GDoF for the IC was first analyzed in [4], where it was also shown that in the very strong interference regime, the maximal GDoF of the Gaussian IC is achieved by letting each receiver decode both the interfering message as well as the intended message. It thus follows that when interference is sufficiently strong, jointly decoding both messages at each receiver is the optimal strategy from both the sum-rate and the GDoF perspectives.

The weak interference (WI) regime is the opposite regime to the SI regime. In this regime, since the interference is weak, then decoding the interfering message cannot be done without constraining the rates of the desired information at each receiver. In [4] it was shown that when interference is sufficiently weak, treating interference as noise at the receivers achieves the maximal GDoF of the Gaussian IC in the WI regime; In [5]–[7] it was shown that this strategy is also sum-rate optimal in the WI regime for finite SNRs. As treating interference as noise is implemented via a low complexity, simple, point-to-point (PtP) decoding strategy, there is a strong motivation for identifying additional scenarios in which treating interference as noise at the receivers carries optimality.

In this work, we study the impact of cooperation on the communications performance in the WI regime by considering the IC with an additional relay node (ICR). The objective of the relay node in the general ICR is to simultaneously assist communications from both sources to their corresponding destinations [8], [9]. The optimal transmission strategy for the relay node in this channel is not known in general. One of the main difficulties in the design of transmission schemes is that when the relay assists one pair, it may degrade the performance of the other pair. In [10], the authors derived an achievable rate region for Gaussian ICRs by using the rate splitting technique (see, e.g., [11]) at the sources, and by employing the decode-and-forward (DF) strategy at the relay. Additional inner bounds and outer bounds on the capacity region of the ICR were derived in [12] and [13]. The capacity region of ergodic fading ICRs in the strong interference (SI) regime was studied in [14], for both Rayleigh fading and phase fading scenarios. In [14] it was shown that when relay reception is good and the interference is strong, then,

Manuscript received September 8, 2015; revised November 5, 2016; accepted January 9, 2017. Date of publication February 13, 2017; date of current version May 18, 2017. This work was supported by the Israeli science foundation under Grant 396/11. This work was presented in part at the 2013 IEEE International Symposium on Information Theory (ISIT) and in part at the 2014 IEEE ISIT.

The authors are with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Be'er-Sheva 8410501, Israel (e-mail: sdzahavi@gmail.com; ron@ee.bgu.ac.il).

Communicated by S.-Y. Chung, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2017.2668388

similarly to the IC, the optimal strategy at each receiver is to jointly decode both the desired message and the interfering message, while the optimal strategy at the relay node is to employ the DF scheme. The sum-rate capacity of the Gaussian IC with a *potent* relay in the WI regime was characterized in [15], in which it was shown that in such a scenario, compress-and-forward (CF) at the relay together with treating interference as noise at the destinations is sum-rate optimal. The sum-rate capacity of the ICR in the WI regime when all nodes have finite powers *remains unknown to date*. The ergodic sum-rate capacity of interference networks without relays, subject to phase fading, was studied in [16], and explicit sum-capacity expressions based on ergodic interference alignment (which requires channel state information (CSI) at the transmitters) were derived for networks with a finite number of users. The work [16] also derived an asymptotic sum-rate capacity expression when the number of users increases to infinity. ICs with time-varying/frequency-selective channel coefficients, in which global CSI is available at all nodes, and in addition, the magnitudes of all links have the same exponential scaling as a function of the signal-to-noise ratio (SNR), were studied in [17]. Under these conditions, [17] showed that adding a relay *does not increase the DoF region*, and that the achievable DoF for each pair in the ICR is upper bounded by 1. On the other hand, it was shown in [18] that relaying can increase the *GDoF* for symmetric Gaussian ICRs. This follows since differently from the DoF analysis, in GDoF analysis the magnitudes of different links may have different SNR scaling exponents. In [18], several GDoF upper bounds were derived for Gaussian ICRs by using the cut-set theorem and the genie-aided approach, for the case in which the source-destination, source-relay, and relay-destination links scale differently as a function of the SNR. Additionally, [18] showed that in the WI regime, *when the source-relay links are weaker than the interfering links* in the sense that their SNR scaling exponent is smaller, then the Han-Kobayashi (HK) scheme [11] achieves the maximal GDoF. The complementing scenario, i.e., *GDoF analysis when the interfering links are weaker than the source-relay links,* was considered in [19]. The GDoF analysis in [19] was based on deriving upper bounds on the sum-rate capacity of the *linear deterministic* ICR. Lastly, we note that the GDoF of the Gaussian IC with a broadcasting relay, in which the relay-destination links are *noiseless*, finite-capacity links, which are *orthogonal* to the other links in the channel, was studied in [20]. From the GDoF characterization, [20] concludes that in the WI regime, each bit per channel use transmitted by the relay can improve the sum-rate capacity by 2 bits per channel use.

To date, there has been no work that characterized the sum-rate capacity and the maximal GDoF of ergodic phase fading ICs with a causal relay in the WI regime, for scenarios in which the power of the relay is finite. In this work, we partially fill this gap by considering a special case of the ergodic phase fading ICR, in which one of the interfering links is missing, e.g., as a result of shadowing in the channel. Furthermore, we consider the scenario in which the relay node receives transmissions from only *one of the two sources*, but is received at both destinations. We refer to this channel configuration as *Z-interference channel with a relay* (Z-ICR).

Main Contributions

In this paper, we characterize for the first time the sum-rate capacity (i.e., finite-SNR performance) and the maximal GDoF (i.e., asymptotically high SNR performance) of the ergodic phase fading Z-ICR in the WI regime, when the relay is causal, has a finite transmission power, and operates in full-duplex mode. Performance gain from cooperation in the WI regime is demonstrated in both the sum-rate capacity and the maximal GDoF. In contrast to [15], which showed the optimality of CF for memoryless ICRs with AWGN and time-invariant link coefficients, we study the *ergodic phase fading* (also referred to as fast phase fading) scenario and demonstrate the optimality of DF. Throughout this paper it is assumed that the nodes have *causal* CSI only on their *incoming links* (Rx-CSI); no transmitter CSI (Tx-CSI) is assumed. The links are all subject to i.i.d. phase fading (see, e.g., [16, Sec. II] and [22, Sec. VII]) which can be applied to modeling many practical scenarios. One such example is non-coherent wireless communication [23], in which phase fading occurs due to the lack of perfect frequency synchronization between the oscillators at the transmitter and at the receiver. Phase fading channel models also apply to systems which use dithering to decorrelate signals, as well as to optic fiber channels [23]. In this work, it is assumed that the relay receives transmissions from only one of the sources, while relay transmissions are received at both destinations. Thus, differently from previous works, *the relay cannot forward desired information to one of the destinations*.

Our main contributions are summarized as follows:

- We derive an upper bound on the achievable sum-rate of the ergodic phase fading Z-ICR by using the genie-aided approach. The upper bound requires a novel design of the genie signals as well as the introduction of novel tools for proving that the bound is maximized by mutually independent, i.i.d., complex Normal channel inputs.
- We derive a lower bound on the achievable sum-rate of the ergodic phase fading Z-ICR by using DF at the relay, and by treating interference as noise at each receiver. We also state conditions on the *magnitudes of the channel coefficients* under which the sum-rate of our lower bound coincides with the sum-rate upper bound. This results in *the characterization of the sum-rate capacity of the ergodic phase fading Z-ICR in the WI regime. This is the first time capacity is characterized for a cooperative interference network in the WI regime, when all powers are finite.*
- We derive two upper bounds on the achievable GDoF of the ergodic phase fading Z-ICR, as well, as a lower bound on the achievable GDoF.

  Note that while capacity analysis is common for fading scenarios, GDoF analysis was previously applied only to *time-invariant* AWGN channels. This follows since, when the channel coefficients vary with time (e.g., a fading channel), then the GDoF generally becomes a random variable. For the ergodic phase fading model, however, as the squared magnitude of each channel coefficient is a *constant*, then GDoF analysis is relevant despite the random temporal nature of the channel coefficients. To the best of our knowledge, this is the *first time* that GDoF analysis is carried out for a fading scenario.

- We identify conditions on the *scaling of the links' magnitudes* (i.e., SNR exponents) under which our GDoF lower bound coincides with the GDoF upper bound. *This characterizes the maximal GDoF of the phase fading Z-ICR in the WI regime*.

Our results show that when certain conditions on the channel coefficients are satisfied, then adding a relay to the ergodic phase fading Z-IC *strictly increases* both the sum-rate capacity and the maximal GDoF of the channel in the WI regime. We note that the sum-rate capacity analysis in this paper has two major differences from the work of [15]: First, we consider a fading scenario while [15] considered the time-invariant AWGN case, and second, we assume that the power of the relay is *finite* while [15] considered a *potent* relay.

In the GDoF analysis, similarly to [4], [9], and [18]–[20], we consider a general setup in which the different links scale differently as a function of the SNR, which facilitates characterizing the impact of the relative link strengthes on the SNR scaling of the sum-rate. The GDoF analysis in this paper has several fundamental differences from the works [18]–[20]: First, note that [18]–[20] studied the common time-invariant Gaussian channel while we consider an ergodic fading channel; Second, unlike [18]–[20], the channel configuration studied in this work is *not symmetric* and the relay cannot forward desired information to one of the destinations. We further note that, unlike [18]–[20], GDoF optimality in the present work is achieved only in a non-symmetric scenario in which the link from the relay to one receiver scales differently than the link from the relay to the other receiver; We also emphasize that while in our work we consider a non-orthogonal scenario, the work [20] considered noiseless, orthogonal relay-destination links, and thus, relay transmissions in [20] do not interfere with the reception of the desired signal at each receiver. It therefore follows that the GDoF of the ergodic phase fading Z-ICR, studied in the present work, cannot be derived as a special case of GDoF results for Gaussian ICRs derived in [18]–[20].

The rest of this paper is organized as follows: In Section II, we define the system model and describe the notation used throughout this paper. In section III, we characterize the sum-rate capacity of the ergodic phase fading Z-ICR in the WI regime, and in section IV, we characterize the maximal GDoF of this channel in the WI regime. Finally, concluding remarks are provided in Section V.

II. Notation and System Model

We denote random variables (RVs) with upper-case letters, e.g., X, Y, and their realizations with lower-case letters, e.g., x, y. We denote the probability density function (p.d.f.) of a continuous RV X with $f_X(x)$. Double-stroke letters are used for denoting matrices, e.g., $\mathbb{A}$, $\mathbb{h}$, with the exception that $\mathbb{E}\{X\}$ denotes the stochastic expectation of X. The element at the k'th row and l'th column of the matrix $\mathbb{A}$ is denoted with $\left[\mathbb{A}\right]_{k,l}$. Bold-face letters, e.g., $\mathbf{X}$, denote column vectors, the i'th element of a vector $\mathbf{X}$, $i > 0$, is denoted with X_i, and X^j denotes the vector $(X_1, X_2, \ldots, X_j)^T$. Given a complex number x, we denote the real and the imaginary parts of x with $\mathfrak{Re}\{x\}$ and $\mathfrak{Im}\{x\}$, respectively. x^* denotes the conjugate of x, $\mathbf{X}^T$ denotes the transpose of $\mathbf{X}$, $\mathbb{A}^H$ denotes the Hermitian transpose of $\mathbb{A}$, $|\mathbb{A}|$ denotes the determinant of $\mathbb{A}$, and $\mathbb{I}_n$ denotes the $n \times n$ identity matrix. For a complex vector X^n, we define an associated real vector by stacking its real and imaginary parts: $\bar{X}^{2n} = \left(\mathfrak{Re}\{X^n\}^T, \mathfrak{Im}\{X^n\}^T\right)^T$. $\mathfrak{R}$ and $\mathfrak{C}$ denote the sets of real and of complex numbers, respectively. Given two $n \times n$ Hermitian matrices, $\mathbb{A}, \mathbb{B}$, we write $\mathbb{B} \preceq \mathbb{A}$ if $\mathbb{A} - \mathbb{B}$ is positive semidefinite (p.s.d.) and $\mathbb{B} \prec \mathbb{A}$ if $\mathbb{A} - \mathbb{B}$ is positive definite (p.d.). $\mathcal{A}_\epsilon^{(n)}(X, Y)$ denotes the set of weakly jointly typical sequences with respect to $f_{X,Y}(x, y)$. We denote the Gaussian distribution with mean μ and variance σ^2 with $\mathcal{N}(\mu, \sigma^2)$, and similarly, we denote the circularly symmetric complex Normal distribution with variance σ^2 with $\mathcal{CN}(0, \sigma^2)$. For a complex random vector, the covariance matrix and the pseudo-covariance matrix are defined as in [25, Sec. II]. Given an RV X with $\mathbb{E}\{X\} = 0$, X_G (i.e., adding a subscript "G" to the RV) denotes an RV which is distributed according to a circularly symmetric, complex Normal distribution with the same variance as the indicated RV, i.e., $X_G \sim \mathcal{CN}\big(0, \mathrm{var}\{X\}\big)$; similarly, subscript "$\bar{\mathrm{G}}$" is used to denote an RV which is distributed according to a complex Normal distribution with the same variance as the indicated RV, where the mean is explicitly specified. We emphasize that RVs with subscript "$\bar{\mathrm{G}}$" *are not necessarily circularly symmetric*. We denote $f(\mathsf{SNR}) \doteq \mathsf{SNR}^c$ if $\lim_{\mathsf{SNR}\to\infty} \frac{\log f(\mathsf{SNR})}{\log \mathsf{SNR}} = c$, and given $f(\mathsf{SNR}) \doteq \mathsf{SNR}^c$ and $g(\mathsf{SNR}) \doteq \mathsf{SNR}^d$, we write $f(\mathsf{SNR}) \dot{\leq} g(\mathsf{SNR})$ if $c \leq d$. Lastly, we note that all logarithms are of base 2.

The Z-ICR consists of two transmitters, Tx_1, Tx_2, two receivers, Rx_1, Rx_2 and a full-duplex relay node. Tx_k sends messages to Rx_k, $k \in \{1, 2\}$. The relay node receives only the signal transmitted from Tx_1 but is received at both destinations simultaneously. The signal received at Rx_1 is a combination of the transmissions of Tx_1 and of the relay along with interference from Tx_2, while the signal received at Rx_2 is a combination of the transmissions of Tx_2 and of the relay *without* interference from Tx_1. This channel model is depicted in Fig. 1. The received signals at Rx_1, Rx_2 and the relay at time i are denoted with $Y_{1,i}$, $Y_{2,i}$, and $Y_{3,i}$, respectively; the channel inputs from Tx_1, Tx_2 and the relay at time i are denoted with $X_{1,i}$, $X_{2,i}$ and $X_{3,i}$, respectively. Finally, $H_{lk,i}$ denotes the channel coefficient for the link with input $X_{l,i}$ and output $Y_{k,i}$ at time instance i. The relationship between the channel inputs and its outputs can be written as:

$$Y_{1,i} = H_{11,i}X_{1,i} + H_{21,i}X_{2,i} + H_{31,i}X_{3,i} + Z_{1,i} \quad (1a)$$

$$Y_{2,i} = H_{22,i}X_{2,i} + H_{32,i}X_{3,i} + Z_{2,i} \quad (1b)$$

$$Y_{3,i} = H_{13,i}X_{1,i} + Z_{3,i}, \quad (1c)$$

$i = 1, 2, \ldots, n$, where Z_1, Z_2 and Z_3 are mutually independent RVs, each independent and identically distributed (i.i.d.) over time according to $\mathcal{CN}(0, 1)$, and all noises are

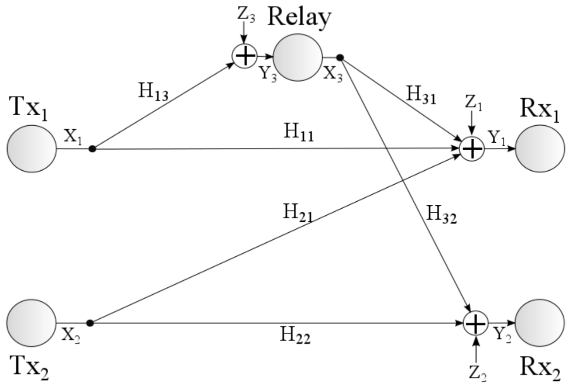


Fig. 1. The ergodic phase fading Z-ICR. The relay node receives transmissions only from Tx_1, but is received at both destinations simultaneously.

independent of the channel inputs and of the channel coefficients. The channel input signals are subject to per-symbol average power constraints: $P_{k,i} \triangleq \mathbb{E}\big\{|X_{k,i}|^2\big\} \leq 1$, $k \in \{1, 2, 3\}$. The receivers and the relay node have instantaneous *causal* Rx-CSI on their incoming links, but the transmitters and the relay do not have Tx-CSI on their outgoing links. Under the ergodic phase fading model, the channel coefficients are given by $H_{lk,i} = \sqrt{\mathsf{SNR}_{lk}} e^{j\Theta_{lk,i}}$, where $\mathsf{SNR}_{lk} \in \mathfrak{R}_+$ is a non-negative constant which corresponds to the signal-to-noise ratio for the link $H_{lk,i}$, and $\Theta_{lk,i}$ is an RV uniformly distributed over $[0, 2\pi)$, i.i.d. in time, independent of the other $\Theta_{lk,i}$'s, and independent of the additive noises $Z_{k,i}$ as well as of the transmitted signals, $X_{k,i}$, $k \in \{1, 2, 3\}$. The independence of the $\Theta_{lk,i}$'s implies that the coefficients $H_{lk,i}$'s are also mutually independent, and in addition they are independent in time, and independent of the other parameters of the scenario.

The channel coefficients causally available at Rx_1 are represented by $\tilde{H}_1 = \big(H_{11}, H_{21}, H_{31}\big)^T \in \mathfrak{C}^3 \triangleq \tilde{\mathfrak{H}}_1$, at Rx_2 they are represented by $\tilde{H}_2 = \big(H_{22}, H_{32}\big)^T \in \mathfrak{C}^2 \triangleq \tilde{\mathfrak{H}}_2$, and at the relay they are represented by $\tilde{H}_3 = H_{13} \in \mathfrak{C} \triangleq \tilde{\mathfrak{H}}_3$. Let $\underline{\tilde{H}} = (\tilde{H}_1^T, \tilde{H}_2^T, \tilde{H}_3)^T \in \mathfrak{C}^6$ be the vector of all channel coefficients, and let $\underline{\mathsf{SNR}} \triangleq \big(\mathsf{SNR}_{11}, \mathsf{SNR}_{21}, \mathsf{SNR}_{31}, \mathsf{SNR}_{22}, \mathsf{SNR}_{32}, \mathsf{SNR}_{13}\big)$. We now state several definitions:

Definition 1: An (R_1, R_2, n) code for the Z-ICR consists of two message sets, $\mathcal{M}_k \triangleq \big\{1, 2, \ldots, 2^{nR_k}\big\}$, $k = 1, 2$, two encoders at the sources, e_1, e_2, employing deterministic mappings; $e_k : \mathcal{M}_k \mapsto \mathfrak{C}^n, k \in \{1, 2\}$, and two decoders at the destinations, g_1, g_2; $g_k : \tilde{\mathfrak{H}}_k^n \times \mathfrak{C}^n \mapsto \mathcal{M}_k, k = 1, 2$. Since the relay receives transmissions only from Tx_1, the transmitted signal at the relay at time i is generated via a set of n functions $\{t_i(\cdot)\}_{i=1}^n$, such that $x_{3,i} = t_i\big(y_3^{i-1}, h_{13}^{i-1}\big) \in \mathfrak{C}, i = 1, 2, \ldots, n$.

Comment 1: Note that since the messages at the transmitters are independent and there is no feedback, then the signals transmitted from Tx_1 and from Tx_2 are necessarily independent as well. Additionally, since the relay receives transmissions only from Tx_1, then its transmitted signal is independent of the signal transmitted from Tx_2. Combining both observations we can write $f_{\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = f_{\mathbf{X}_1,\mathbf{X}_3}(\mathbf{x}_1, \mathbf{x}_3) \cdot f_{\mathbf{X}_2}(\mathbf{x}_2)$. We denote the correlation coefficient between channel inputs X_1 and X_3 at time index i with υ_i: $\upsilon_i \triangleq \frac{\mathbb{E}\{X_{1,i} X_{3,i}^*\}}{\sqrt{P_{1,i} P_{3,i}}}, 0 \leq |\upsilon_i| \leq 1$.

Definition 2: The average probability of error on an (R_1, R_2, n) code is defined as $\mathrm{P}_e^{(n)} \triangleq \Pr\Big(g_1(\tilde{H}_1^n, Y_1^n) \neq M_1 \text{ or } g_2(\tilde{H}_2^n, Y_2^n) \neq M_2\Big)$, where each message is selected independently and uniformly from its message set.

Definition 3: A rate pair (R_1, R_2) is called achievable if, for any $\epsilon > 0$ and $\delta > 0$ there exists some blocklength $n_0(\epsilon, \delta)$, such that for every $n > n_0(\epsilon, \delta)$ there exists an $(R_1 - \delta, R_2 - \delta, n)$ code with $\mathrm{P}_e^{(n)} < \epsilon$.

Definition 4: The capacity region is defined as the convex hull of all achievable rate pairs.

The objective of this work is to characterize two performance measures for the ergodic phase fading Z-ICR in the WI regime: The sum-rate capacity, which characterizes the performance at finite SNRs, and the maximal GDoF, which characterizes the performance at asymptotically high SNRs. As these cases are fundamentally different in nature, the WI regime is defined for each performance measure in accordance with the relevant notions in the literature. In the following we briefly overview the WI conditions for each performance measure, leaving the detailed discussions to the relevant sections.

- In Section III the sum-rate capacity is characterized for the WI regime, defined in Eqns. (8). Generally speaking, WI in this case occurs when the SNRs of the interfering links, SNR_{32} and SNR_{21}, are sufficiently low compared to the SNRs for carrying the desired information. This is in accordance with the acceptable notion of the WI regime for sum-rate capacity analysis at finite SNRs, see, e.g., [5]–[7].
- In Section IV the maximal GDoF is characterized for the WI regime, defined in Eqn. (35a). Generally speaking, WI in this case occurs when the exponents of the SNRs of the interfering links are sufficiently smaller than the exponents of the SNRs of the information paths. In section IV, in the statement of Thm. 3, this condition is expressed as $\lambda = \alpha \leq \frac{1}{2}$, where α and λ denote the exponential scalings of the Tx_2-Rx_1 link and of the relay-Rx_2 link, respectively. This definition is in accordance with the acceptable notion of the WI regime for DoF and GDoF analysis, see, e.g., [4], [18]. We note that in [4] and [18] the WI regime is characterized by $\alpha \leq 1$, while GDoF optimality for the communications scheme described in the current paper requires a stricter notion of WI, characterized by $\lambda = \alpha \leq \frac{1}{2}$. However, note that *in [4], treating interference as noise is GDoF optimal only for* $\alpha \leq \frac{1}{2}$, which is in agreement with our result.

III. Finite SNR Analysis: The Sum-Rate Capacity in the WI Regime

A. Preliminaries

We begin with the presentation of several lemmas used in the derivation of the sum-rate capacity. We note that while some of the following lemmas appeared in previous works for *real* variables, in the following we extend these lemmas to complex variables. Accordingly, in the appendices we include explicit proofs only for those lemmas whose proofs do not follow directly from the original proofs for real RVs.

Lemma 1: Let $\mathbf{Z}_1$ and $\mathbf{Z}_2$ be a pair of n-dimensional, circularly symmetric, complex Normal random vectors, and let $\mathbf{X}$ be an n-dimensional complex random vector whose p.d.f. is denoted by $f_{\mathbf{X}}(\mathbf{x})$. Consider the following optimization problem:

$$\begin{aligned}&\max_{f_{\mathbf{X}}(\mathbf{x})} \; h(\mathbf{X}+\mathbf{Z}_1)-h(\mathbf{X}+\mathbf{Z}_2) \\ &\text{subject to:}\quad \mathrm{tr}\big(\mathrm{cov}(\mathbf{X})\big)\le nP.\end{aligned} \tag{2}$$

Then, a circularly symmetric, complex Normal random vector $\mathbf{X}_G^{\text{Opt}} \sim \mathcal{CN}(\mathbf{0}, \mathbb{C}_X^{\text{Opt}})$ is an optimal solution to the optimization problem in (2). Additionally, if $\mathbf{Z}_1$ and $\mathbf{Z}_2$ have i.i.d. entries, i.e., $\mathbf{Z}_k \sim \mathcal{CN}(\mathbf{0}, \gamma_k \mathbb{I}_n)$, $\gamma_k \in \Re^+$, $k \in \{1,2\}$, and if it holds that $\gamma_1 \le \gamma_2$, then the optimal solution is distributed according to $\mathbf{X}_G^{\text{Opt}} \sim \mathcal{CN}\big(\mathbf{0}, P\cdot\mathbb{I}_n\big)$.

Proof: The proof is based on [5, Corollary 2] and [24, Th. 1], and is therefore not included in the manuscript. A detailed proof can be found in [34]. ■

Lemma 2: Let $\mathbf{Z}$ and $\mathbf{W}$ be a pair of n-dimensional, zero-mean, jointly circularly symmetric complex Normal random vectors with i.i.d. entries, s.t. their joint distribution can be written as

$$f_{\mathbf{W},\mathbf{Z}}(\mathbf{w},\mathbf{z}) = \prod_{i=1}^{n} f_{W,Z}(w_i, z_i). \tag{3}$$

Denote the cross-covariance matrix between Z_i and W_i with

$$\mathrm{cov}(Z_i, W_i) = \begin{bmatrix} \sigma_1^2 & \tilde{\upsilon}_{12} \\ \tilde{\upsilon}_{12}^* & \sigma_2^2 \end{bmatrix}, \quad i \in \{1,2,\dots,n\},$$

where $\sigma_1^2 > 0$, and $\sigma_2^2 > 0$. Let $\mathbf{V}$ be an n-dimensional, zero-mean, circularly symmetric complex Normal random vector with i.i.d. entries, whose covariance matrix is given by $\mathbb{E}\{\mathbf{V}\mathbf{V}^H\} = \Big(\sigma_1^2 - \frac{|\tilde{\upsilon}_{12}|^2}{\sigma_2^2}\Big)\cdot\mathbb{I}_n$. If $\mathbf{X}$ is independent of $(\mathbf{Z},\mathbf{W},\mathbf{V})$, then

$$h(\mathbf{X}+\mathbf{Z}|\mathbf{W}) = h(\mathbf{X}+\mathbf{V}).$$

Proof: The proof follows similar steps as in the proof of [5, Lemma 3] and is therefore not included in the manuscript. A detailed proof can be found in [34]. ■

Lemma 3: Let Z and W be a pair of possibly correlated, zero-mean, jointly circularly symmetric complex Normal RVs, and let $\mathbf{H}_Y$ and $\mathbf{H}_S$ be two $n\times 1$ complex random vectors. Additionally, let $\mathbf{X}$ be an $n\times 1$ complex random vector, and let Y and S be noisy observations of $\mathbf{X}$, s.t.

$$Y = \mathbf{H}_Y^T\mathbf{X} + Z \tag{4a}$$
$$S = \mathbf{H}_S^T\mathbf{X} + W. \tag{4b}$$

Consider the sequence of random vectors $\mathbf{X}^n = (\mathbf{X}_1^T, \mathbf{X}_2^T \dots, \mathbf{X}_n^T)^T$ and let $\mathbb{Q}_{\mathbf{X}_i}$ denote the covariance matrix of the $n\times 1$ vector $\mathbf{X}_i$. Furthermore, let Y^n and S^n be the corresponding observations when the noise sequences (Z^n, W^n) are i.i.d. in the sense of (3). Define $\mathbf{H} = (\mathbf{H}_Y^T, \mathbf{H}_S^T)^T$, and let $\mathbf{H}^n = (\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_n^T)^T$ be an i.i.d. sequence of random vectors, in which each $2n\times 1$ vector element is distributed according to the distribution of the $2n\times 1$ random vector $\mathbf{H}$. Then, we can bound

$$h(Y^n|S^n, \mathbf{H}^n) \le n\cdot h(Y_G|S_G, \mathbf{H}), \tag{5}$$

where Y_G and S_G denote the RVs Y and S defined in (4), obtained with $\mathbf{X}$ replaced with $\mathbf{X}_G \sim \mathcal{CN}(\mathbf{0}, \frac{1}{n}\sum_{i=1}^n \mathbb{Q}_{\mathbf{X}_i})$.

Proof: The proof follows similar steps as in the proof of [6, Lemma 1] and is therefore not included in the manuscript. A detailed proof can be found in [34]. ■

Lemma 4: Let X_1, X_2, Z_1 and Z_2 be zero mean, circularly symmetric complex Normal RVs s.t. (X_1, X_2) is independent of (Z_1, Z_2), and in addition, $\mathbb{E}\{\Re\mathfrak{e}\{Z_1\}\Im\mathfrak{m}\{Z_2\}\} = -\mathbb{E}\{\Im\mathfrak{m}\{Z_1\}\Re\mathfrak{e}\{Z_2\}\}$ and $\mathbb{E}\{\Re\mathfrak{e}\{Z_1\}\Re\mathfrak{e}\{Z_2\}\} = \mathbb{E}\{\Im\mathfrak{m}\{Z_1\}\Im\mathfrak{m}\{Z_2\}\}$. Let c_1 and c_2 be a pair of complex constants, and let Y_1 and Y_2 be defined via

$$Y_1 = c_1\cdot X_1 + c_2\cdot X_2 + Z_1$$
$$Y_2 = c_1\cdot X_1 + c_2\cdot X_2 + Z_2.$$

Then, $I(X_1, X_2; Y_1|Y_2) = 0$ if and only if $\mathbb{E}\{Z_1 Z_2^*\} = \mathbb{E}\{|Z_2|^2\}$.

Proof: The proof is provided in Appendix A. ■

Lemma 5: Let $Z^{n+m} \triangleq \Big(\big(Z_1^n\big)^T, \big(Z_2^m\big)^T\Big)^T$, where Z_1^n and Z_2^m are two mutually independent, circularly symmetric complex Normal random vectors of lengths n and m, respectively, each with independent entries distributed according to $Z_{1,i} \sim \mathcal{CN}(0, a_{1,i}\gamma_1)$, $i \in \{1,2,\dots,n\}$ and $Z_{2,i} \sim \mathcal{CN}(0, a_{2,i}\gamma_2)$, $i \in \{1,2,\dots,m\}$, where γ_1, γ_2, and $a_{k,i}$, $k\in\{1,2\}$ are positive, real, and finite constants. Let $X^{n+m} \triangleq \Big(\big(X_1^n\big)^T, \big(X_2^m\big)^T\Big)^T$ where X_1^n and X_2^m are two complex random vectors of lengths n and m, respectively, with finite covariance matrices, and further let X^{n+m} be mutually independent of Z^{n+m}. Then, we have the following limit:

$$\lim_{\gamma_2\to\infty} I\big(X^{n+m}; X^{n+m}+Z^{n+m}\big) = I\big(X_1^n; X_1^n + Z_1^n\big).$$

Proof: The proof is provided in Appendix B. ■

Lemma 6: Let $\tilde{Z}_1^n$ and $\tilde{Z}_2^n$ be a pair of possibly correlated, n-dimensional circularly symmetric complex Normal random vectors, each with independent entries, i.e., $\tilde{Z}_k^n \sim \mathcal{CN}(0, \tilde{\mathbb{D}}_k^z)$, $k\in\{1,2\}$, where $\tilde{\mathbb{D}}_k^z$, $k\in\{1,2\}$ are two $n\times n$ diagonal matrices with real and positive entries on their main diagonals. Let $\tilde{\mathbb{V}}_1$ and $\tilde{\mathbb{V}}_2$ be two $2n\times n$ deterministic complex matrices, s.t. $\tilde{\mathbb{V}}_k^H\tilde{\mathbb{V}}_k = \tilde{\mathbb{D}}_k^{-1}$, where $\tilde{\mathbb{D}}_k$, $k\in\{1,2\}$ are two $n\times n$ diagonal matrices with real and positive entries on their main diagonals. Let X^{2n} be a $2n\times 1$ complex random vector with distribution $f_{X^{2n}}(x^{2n})$, independent of $(\tilde{Z}_1^n, \tilde{Z}_2^n)$, and let $\bar{X}^{4n}$ be the stacking of the real and imaginary parts of X^{2n}, i.e., $\bar{X}^{4n} \triangleq \Big(\big(\Re\mathfrak{e}\{X^{2n}\}\big)^T, \big(\Im\mathfrak{m}\{X^{2n}\}\big)^T\Big)^T$. Consider the following optimization problem:

$$\max_{f(x^{2n}):\mathrm{cov}(\bar{X}^{4n})\preceq\mathbb{S}} h(\tilde{\mathbb{V}}_1^H\cdot X^{2n} + \tilde{Z}_1^n) - h(\tilde{\mathbb{V}}_2^H\cdot X^{2n} + \tilde{Z}_2^n). \tag{6}$$

Then, a *zero-mean* complex Normal random vector, X_G^{2n}, is an optimal solution for (6).

Proof: The proof is provided in Appendix C. ■

B. Sum-Rate Capacity in the WI Regime

Let $\mathcal{C}(\underline{\text{SNR}})$ denote the capacity region of the Z-ICR, for a given $\underline{\text{SNR}}$. The sum-rate capacity of the ergodic phase fading

Z-ICR in the WI regime is characterized in the following theorem:

Theorem 1: Consider the ergodic phase fading Z-ICR with only Rx-CSI, defined in Section II. If $\underline{\mathsf{SNR}}$ *satisfies*

$$\frac{\mathsf{SNR}_{11} + \mathsf{SNR}_{31}}{1 + \mathsf{SNR}_{21}} \leq \mathsf{SNR}_{13}, \tag{7}$$

and if there exist two real scalars β_1 *and* β_2 *which satisfy* $0 \leq \beta_1, \beta_2 \leq 1$*, and*

$$\mathsf{SNR}_{32}(1 + \mathsf{SNR}_{21})^2 \leq \beta_1 \Big(\mathsf{SNR}_{31}(1 - \beta_2) - 2\mathsf{SNR}_{32}\mathsf{SNR}_{11} \Big) \tag{8a}$$

$$\mathsf{SNR}_{21}(1 + \mathsf{SNR}_{32})^2 \leq \beta_2 \cdot \mathsf{SNR}_{22}(1 - \beta_1), \tag{8b}$$

then, the sum-rate capacity of the channel is given by

$$\sup_{(R_1, R_2) \in \mathcal{C}(\underline{\mathsf{SNR}})} (R_1 + R_2) = \log\left(1 + \frac{\mathsf{SNR}_{11} + \mathsf{SNR}_{31}}{1 + \mathsf{SNR}_{21}}\right) + \log\left(1 + \frac{\mathsf{SNR}_{22}}{1 + \mathsf{SNR}_{32}}\right), \tag{9}$$

and it is achieved by $X_k \sim \mathcal{CN}(0, 1), k \in \{1, 2, 3\}$*, mutually independent.*

Comment 2: Observe that the conditions in (8) are satisfied if SNR_{32} and SNR_{21} are small compared to SNR_{31} and SNR_{22}, respectively. As SNR_{32} and SNR_{21} correspond to the strengths of the interfering links, conditions (8) correspond to the WI regime. To make this point more explicit, note that $\mathsf{SNR}_{32}(1 + \mathsf{SNR}_{21})^2 \geq \mathsf{SNR}_{32}$, hence (8a) implies that $\mathsf{SNR}_{32} \leq \beta_1 \cdot \mathsf{SNR}_{31}(1 - \beta_2)$, and similarly (8b) implies that $\mathsf{SNR}_{21} \leq \beta_2 \cdot \mathsf{SNR}_{22}(1 - \beta_1)$.

Comment 3: Note that condition (7) corresponds to good reception at the relay, in the sense that decoding the message sent by Tx_1 at the relay does not constrain the information rate from Tx_1 to Rx_1. This condition facilitates the sum-rate optimality of DF, as the constraints on the achievable rates are now only due to the rate constraints for reliable decoding at the destinations.

Comment 4: Note that in the ergodic phase fading case, the magnitudes of the channel coefficients are constants while the phases of the channel coefficients vary i.i.d. over time and are mutually independent across the fading links. Thus, in the ergodic phase fading model, the channel coefficients induce randomly varying phases upon the components of the received signal arriving at each receiver after traveling across the different links. Intuitively, having mutually independent and uniformly distributed i.i.d. phases does not allow achieving non-zero correlation between the components of the received signal, and consequently implies that there is no loss of optimality in transmitting uncorrelated codewords. In particular, if the optimal input distribution is complex Normal, then the absence of correlation between the codebooks implies that the optimal codebooks are generated independently of each other. Indeed, in the derivations in the manuscript, it is rigorously proved that the optimal channel inputs for the ergodic phase fading Z-ICR are generated according to mutually independent complex Normal random variables. The optimality of mutually independent channel inputs is one of the fundamental advantages of the communications scheme we use in this manuscript, since it means that there is no need for coordinated transmission in order to optimally benefit from the relay. As will be clarified later, this fact greatly simplifies both the achievability scheme as well as the practical incorporation of cooperative transmission in interference networks. In contrast, for the no-fading case (commonly referred to as the AWGN channel) both the magnitudes and the phases of the channel coefficients are constants. Consequently, in the no-fading channel the correlation between the channel inputs is maintained at the received signal components, and hence, the optimal codebooks may be correlated. This fact greatly complicates the optimal achievability scheme as well as makes the derivations for the upper bounds significantly more complicated.

Proof: The proof of Thm. 1 consists of the following three steps:

1) We derive an upper bound on the sum-rate of the ergodic phase fading Z-ICR by letting each receiver observe an appropriate genie signal. In particular, we show that the upper bound is maximized by mutually independent, zero-mean circularly symmetric complex Normal channel inputs, i.i.d. in time.
2) We characterize an achievable rate region for the Z-ICR by using codebooks generated according to a mutually independent circularly symmetric complex Normal distribution, i.i.d. in time, and by employing the DF scheme at the relay, together with treating the interfering signal as noise at each receiver.
3) Combining the conditions for the upper bound and for the lower bound we obtain the conditions for characterizing the sum-rate capacity of the ergodic phase fading Z-ICR in the WI regime, and explicitly state the corresponding expressions.

In the following subsections, we provide the detailed derivations for the above steps: Step 1 is carried out in Section III-C, Step 2 is carried out in Section III-D, and finally, Step 3 is detailed in Section III-E.

C. Step 1: An Upper Bound on the Sum-Rate Capacity

The upper bound on the sum-rate capacity of the ergodic phase fading Z-ICR is summarized in the following theorem:

Theorem 2: Consider the phase fading Z-ICR with only Rx-CSI, defined in Section II. If there are two real scalars β_1 *and* β_2 *which satisfy* $0 \leq \beta_1, \beta_2 \leq 1$*, and*

$$\mathsf{SNR}_{32}(1 + \mathsf{SNR}_{21})^2 \leq \beta_1 \Big(\mathsf{SNR}_{31}(1 - \beta_2) - 2\mathsf{SNR}_{32}\mathsf{SNR}_{11} \Big) \tag{10a}$$

$$\mathsf{SNR}_{21}(1 + \mathsf{SNR}_{32})^2 \leq \beta_2 \cdot \mathsf{SNR}_{22}(1 - \beta_1), \tag{10b}$$

then, the sum-rate capacity is upper bounded by

$$\sup_{(R_1, R_2) \in \mathcal{C}(\underline{\mathsf{SNR}})} (R_1 + R_2) \leq I(X_1, X_3; Y_1 | \tilde{H}_1) + I(X_2; Y_2 | \tilde{H}_2), \tag{11}$$

where the mutual information expressions are evaluated with mutually independent, zero mean, circularly symmetric complex Normal channel inputs, distributed according to $X_k \sim \mathcal{CN}(0,1)$.

Proof: We use a genie to provide additional information to the receivers. Let W_1^n and W_2^n be two arbitrarily correlated, circularly symmetric, complex Normal random vectors, each with i.i.d. elements distributed $\mathcal{CN}(0,1)$, such that $f_{W_1^n,W_2^n}(w_1^n,w_2^n) = \prod_{i=1}^n f_{W_1,W_2}(w_{1,i},w_{2,i})$. In addition, (W_1^n,W_2^n) are independent of (X_1^n,X_2^n,X_3^n). For $i\in\{1,2,\ldots,n\}$, and we further let $W_{k,i}$ be correlated with $Z_{k,i}$, $k=1,2$:

$$\begin{aligned}\mathrm{cov}(W_{k,i},Z_{k,i}) &= \mathbb{E}\left\{\begin{bmatrix}W_{k,i}\\ Z_{k,i}\end{bmatrix}\begin{bmatrix}W_{k,i}^* & Z_{k,i}^*\end{bmatrix}\right\}\\ &= \begin{bmatrix}1 & \tilde{\upsilon}_k\\ \tilde{\upsilon}_k^* & 1\end{bmatrix},\quad k\in\{1,2\},\end{aligned}$$

and further assume $\mathbb{E}\{\mathfrak{Re}\{W_{k,i}\}\mathfrak{Im}\{Z_{k,i}\}\} = -\mathbb{E}\{\mathfrak{Im}\{W_{k,i}\}\mathfrak{Re}\{Z_{k,i}\}\}$ and $\mathbb{E}\{\mathfrak{Re}\{W_{k,i}\}\mathfrak{Re}\{Z_{k,i}\}\} = \mathbb{E}\{\mathfrak{Im}\{W_{k,i}\}\mathfrak{Im}\{Z_{k,i}\}\}$. Note that since $\mathrm{var}(W_k)=\mathrm{var}(Z_k)=1, k=1,2$, then $|\tilde{\upsilon}_k| \le \sqrt{\mathrm{var}(W_k)\cdot\mathrm{var}(Z_k)} = 1$. Define the signals

$$S_{1,i} \triangleq H_{11,i}X_{1,i} + H_{31,i}X_{3,i} + \eta_1 W_{1,i} \qquad (12a)$$

$$S_{2,i} \triangleq H_{22,i}X_{2,i} + \eta_2 W_{2,i}, \qquad (12b)$$

$i\in\{1,2,\ldots,n\}$, where η_1 and η_2 are two complex-valued constants determined by the genie. Assume that at time i, the genie provides the signals $S_{1,i}$ and $S_{2,i}$ to Rx_1 and Rx_2, respectively. For an achievable rate pair (R_1,R_2), let $(\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3)$ be random vectors of length n representing the statistics of the achievable codebook, and define

$$q_{k,i} \triangleq \mathrm{cov}\big(X_{k,i}\big) \equiv \mathrm{var}(X_{k,i}),\quad k\in\{1,2,3\}, \qquad (13a)$$

$$\mathbb{Q}_{\mathbf{X}}^{(i)} \triangleq \mathrm{cov}(X_{1,i},X_{2,i},X_{3,i}),\qquad i\in\{1,2,\ldots,n\}, \quad (13b)$$

and

$$P_k = \frac{1}{n}\sum_{i=1}^n q_{k,i},\quad k\in\{1,2,3\}.$$

It is emphasized that at this point, properness of $(\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3)$ is not assumed. Next, recall that in Comment 1 we concluded that $\mathbf{X}_2$ is independent of $(\mathbf{X}_1,\mathbf{X}_3)$, while $\mathbf{X}_1$ and $\mathbf{X}_3$ may be statistically dependent. It follows that the n-letter input distribution for $(\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3)$ must satisfy $f_{\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) = f_{\mathbf{X}_1,\mathbf{X}_3}(\mathbf{x}_1,\mathbf{x}_3)\cdot f_{\mathbf{X}_2}(\mathbf{x}_2)$. From this observation, we conclude that for $i\in\{1,2,\ldots,n\}$

$$\mathbb{Q}_{\mathbf{X}}^{(i)} = \begin{bmatrix} q_{1,i} & 0 & \upsilon_i\sqrt{q_{1,i}q_{3,i}}\\ 0 & q_{2,i} & 0\\ \upsilon_i^*\sqrt{q_{1,i}q_{3,i}} & 0 & q_{3,i}\end{bmatrix}, \qquad (14a)$$

$$\begin{aligned}\mathbb{Q}_G &\triangleq \frac{1}{n}\sum_{i=1}^n \mathbb{Q}_{\mathbf{X}}^{(i)}\\ &\equiv \begin{bmatrix} P_1 & 0 & \upsilon\sqrt{P_1P_3}\\ 0 & P_2 & 0\\ \upsilon^*\sqrt{P_1P_3} & 0 & P_3\end{bmatrix},\end{aligned} \qquad (14b)$$

where $|\upsilon_i| \le 1$ in (14a). Note that by the Cauchy-Schwartz inequality [32]

$$\begin{aligned}\left|\frac{1}{n}\sum_{i=1}^n \upsilon_i\sqrt{q_{1,i}q_{3,i}}\right| &\le \sqrt{\frac{1}{n}\sum_{i=1}^n\left|\upsilon_i\sqrt{q_{1,i}}\right|^2}\sqrt{\frac{1}{n}\sum_{i=1}^n\left|\sqrt{q_{3,i}}\right|^2}\\ &\le \sqrt{\frac{1}{n}\sum_{i=1}^n q_{1,i}}\sqrt{\frac{1}{n}\sum_{i=1}^n q_{3,i}}\\ &= \sqrt{P_1P_3},\end{aligned}$$

thus $|\upsilon| \le 1$ in (14b). Lastly, define the random vector $(X_{1G},X_{2G},X_{3G})^T \sim \mathcal{CN}(\mathbf{0},\mathbb{Q}_G)$.

Let M_1 be the transmitted message at Tx_1, $\hat{M}_1$ denote the decoded message at Rx_1, and $P_{e,1}^{(n)}$, denote the probability of error in decoding M_1 at Rx_1. The rate R_1 can be upper bounded as follows:

$$\begin{aligned} nR_1 &= H(M_1)\\ &= H(M_1|Y_1^n,\tilde{H}_1^n) + I(M_1;Y_1^n,\tilde{H}_1^n)\\ &\overset{(a)}{\le} 1 + P_{e,1}^{(n)}nR_1 + I(X_1^n;Y_1^n,\tilde{H}_1^n)\\ &\overset{(b)}{=} 1 + P_{e,1}^{(n)}nR_1 + I(X_1^n;Y_1^n|\tilde{H}_1^n)\\ &\overset{(c)}{\le} 1 + P_{e,1}^{(n)}nR_1 + I(X_1^n,X_3^n;Y_1^n,S_1^n|\tilde{H}_1^n).\end{aligned}$$

Here, (a) follows from Fano's inequality [31, Th. 2.10.1], and from the data processing inequality [31, Th. 2.8.1], since $M_1 - X_1^n - (Y_1^n,\tilde{H}_1^n)$ form a Markov chain; (b) follows since the transmitted symbols X_1^n are independent of the channel coefficients $\tilde{H}_1^n$; and (c) follows from the chain rule of mutual information and since mutual information is non-negative. Next, define $n\epsilon_{1n} \triangleq 1 + P_{e,1}^{(n)}nR_1$, and observe that since (R_1,R_2) is achievable, then $P_{e,1}^{(n)} \to 0$ for $n\to\infty$, and therefore $\epsilon_{1n}\to 0$ as $n\to\infty$. Hence, we obtain

$$\begin{aligned} &n(R_1-\epsilon_{1n})\\ &\quad\le I(X_1^n,X_3^n;Y_1^n,S_1^n|\tilde{H}_1^n)\\ &\quad= I(X_1^n,X_3^n;S_1^n|\tilde{H}_1^n) + I(X_1^n,X_3^n;Y_1^n|S_1^n,\tilde{H}_1^n)\\ &\quad= h(S_1^n|\tilde{H}_1^n) - h(S_1^n|X_1^n,X_3^n,\tilde{H}_1^n)\\ &\qquad + h(Y_1^n|S_1^n,\tilde{H}_1^n) - h(Y_1^n|X_1^n,X_3^n,S_1^n,\tilde{H}_1^n)\\ &\quad\overset{(a)}{\le} h(S_1^n|\tilde{H}_1^n) - h(S_1^n|X_1^n,X_3^n,\tilde{H}_1^n)\\ &\qquad + n\cdot h(Y_{1G}|S_{1G},\tilde{H}_1) - h(Y_1^n|X_1^n,X_3^n,S_1^n,\tilde{H}_1^n)\\ &\quad\overset{(b)}{=} h(S_1^n|\tilde{H}_1^n) - n\cdot h(S_{1G}|X_{1G},X_{3G},\tilde{H}_1)\\ &\qquad + n\cdot h(Y_{1G}|S_{1G},\tilde{H}_1) - h(Y_1^n|X_1^n,X_3^n,S_1^n,\tilde{H}_1^n),\end{aligned} \qquad (15)$$

where $(X_{1G},X_{2G},X_{3G})^T \sim \mathcal{CN}(\mathbf{0},\mathbb{Q}_G)$ with $\mathbb{Q}_G$ defined in (14b), and S_{1G} is obtained from (12a) by replacing X_1 and X_3 with X_{1G} and X_{3G}, respectively. In the above transitions, (a) follows from Lemma 3 using the assignment $\mathbf{X} = (X_1,X_2,X_3)^T$, $\mathbf{H}_Y = (H_{11},H_{21},H_{31})^T$, $\mathbf{H}_S = (H_{11},0,H_{31})^T$, and with $\mathbb{Q}_{\mathbf{X}_i}$ defined in (14a) and (b) follows from the transitions detailed below:

$$\begin{aligned} h\big(S_1^n|X_1^n, X_3^n, \tilde{H}_1^n\big) &= h\big(\eta_1 \cdot (W_1^n)|X_1^n, X_3^n, \tilde{H}_1^n\big) \\ &\stackrel{(c)}{=} h\big(\eta_1 \cdot (W_1^n)\big) \\ &\stackrel{(d)}{=} n \cdot h\big(\eta_1 W_1\big) \\ &\stackrel{(e)}{=} n \cdot h\big(\eta_1 W_1 | X_{1G}, X_{3G}, \tilde{H}_1\big) \\ &= n \cdot h\big(S_{1G}|X_{1G}, X_{3G}, \tilde{H}_1\big), \end{aligned}$$

where step (c) follows since W_1^n is independent of $(X_1^n, X_3^n, \tilde{H}_1^n)$; step (d) follows since W_1^n has i.i.d. entries, and step (e) is valid for any joint distribution on (X_{1G}, X_{3G}) independent of W_1, and thus, in this step we let (X_{1G}, X_{3G}) be distributed according to the joint distribution $\mathcal{CN}(\mathbf{0}, \mathbb{Q}_{G13})$, where

$$\mathbb{Q}_{G13} \triangleq \begin{bmatrix} P_1 & \upsilon\sqrt{P_1 P_3} \\ \upsilon^*\sqrt{P_1 P_3} & P_3 \end{bmatrix}.$$

Applying similar steps and identical arguments for R_2, we obtain the upper bound:

$$\begin{aligned} n(R_2 - \epsilon_{2n}) \leq\, & h(S_2^n|\tilde{H}_2^n) - n \cdot h(S_{2G}|X_{2G}, \tilde{H}_2) \\ & + n \cdot h(Y_{2G}|S_{2G}, \tilde{H}_2) \\ & - h(Y_2^n|X_2^n, S_2^n, \tilde{H}_2^n). \end{aligned} \tag{16}$$

Since the maximizing complex Normal distributions in (15) and (16) are identical and equal to $(X_{1G}, X_{2G}, X_{3G})^T \sim \mathcal{CN}(\mathbf{0}, \mathbb{Q}_G)$, then (15) and (16) can be combined into a single bound on the sum-rate:

$$\begin{aligned} & n(R_1 + R_2 - \epsilon_{1n} - \epsilon_{2n}) \\ & \leq n \cdot \Big(h(Y_{1G}|S_{1G}, \tilde{H}_1) - h(S_{1G}|X_{1G}, X_{3G}, \tilde{H}_1) \\ & \qquad + h(Y_{2G}|S_{2G}, \tilde{H}_2) - h(S_{2G}|X_{2G}, \tilde{H}_2) \Big) \\ & \quad + \Big(h(S_1^n|\tilde{H}_1^n) - h(Y_1^n|X_1^n, X_3^n, S_1^n, \tilde{H}_1^n) \\ & \qquad + h(S_2^n|\tilde{H}_2^n) - h(Y_2^n|X_2^n, S_2^n, \tilde{H}_2^n) \Big). \end{aligned} \tag{17}$$

In the following proposition, we identify the maximizing distribution for the first brackets in the right-hand side of (17):

Proposition 1: The expression in the first brackets in the right-hand side of (17) *is maximized with mutually independent circularly symmetric complex Normal channel inputs distributed according to* $X_k \sim \mathcal{CN}(0, 1), k \in \{1, 2, 3\}$.

Proof: The proof is provided in Appendix D. ■

Next, we show that the expression in the second brackets in the right-hand side of (17) is also maximized by mutually independent, and i.i.d. in time channel inputs, distributed according to $X_{k,i} \sim \mathcal{CN}(0, 1), k \in \{1, 2, 3\}, i \in \{1, 2, \ldots, n\}$. Assume that there exists a pair of complex scalars $\tilde{\upsilon}_1$ and η_2 s.t. $|\tilde{\upsilon}_1| \leq 1$ and $\mathsf{SNR}_{21}|\eta_2|^2 \leq \mathsf{SNR}_{22}(1 - |\tilde{\upsilon}_1|^2)$, and let V_1^n be an n-dimensional random vector with i.i.d. elements distributed according to $V_{1,i} \sim \mathcal{CN}(0, 1 - |\tilde{\upsilon}_1|^2), i \in \{1, 2, \ldots, n\}$. Additionally, let $\mathbb{H}_{h_{lk}}^{(n)}$ be an $n \times n$ diagonal matrix s.t. $\big[\mathbb{H}_{h_{lk}}^{(n)}\big]_{i,i} = h_{lk,i}$. Then

$$\begin{aligned} & h(S_2^n|\tilde{H}_2^n) - h(Y_1^n|X_1^n, X_3^n, S_1^n, \tilde{H}_1^n) \\ &= \mathbb{E}_{\tilde{H}_1^n, \tilde{H}_2^n}\Big\{ h\big(\mathbb{H}_{h_{22}}^{(n)} X_2^n + \eta_2 W_2^n | \tilde{H}_2^n = \tilde{h}_2^n\big) \\ & \qquad - h\big(\mathbb{H}_{h_{21}}^{(n)} X_2^n + Z_1^n | W_1^n, \tilde{H}_1^n = \tilde{h}_1^n\big)\Big\} \\ &\stackrel{(a)}{=} \mathbb{E}_{\tilde{H}_1^n, \tilde{H}_2^n}\Big\{ h\big(\mathbb{H}_{h_{22}}^{(n)} X_2^n + \eta_2 W_2^n | \tilde{H}_2^n = \tilde{h}_2^n\big) \\ & \qquad - h\big(\mathbb{H}_{h_{21}}^{(n)} X_2^n + V_1^n | \tilde{H}_1^n = \tilde{h}_1^n\big)\Big\} \\ &\stackrel{(b)}{\leq} \mathbb{E}_{\tilde{H}_1, \tilde{H}_2}\Big\{ n \cdot h\big(h_{22} X_{2G} + \eta_2 W_2 | \tilde{H}_2 = \tilde{h}_2\big) \\ & \qquad - n \cdot h\big(h_{21} X_{2G} + V_1 | \tilde{H}_1 = \tilde{h}_1\big)\Big\} \\ &\stackrel{(c)}{\leq} n \cdot \Big(h(H_{22} X_{2G} + \eta_2 W_2 | \tilde{H}_2) \\ & \qquad - h(H_{21} X_{2G} + V_1 | \tilde{H}_1) \Big)\Big|_{\substack{P_1 = P_2 = P_3 = 1 \\ \upsilon = 0}}, \end{aligned} \tag{18}$$

where (a) follows from Lemma 2. For step (b) we first use [25, eq. (13)][1] and obtain that since $\tilde{h}_1^n$ and $\tilde{h}_2^n$ are given, then we can write

$$\begin{aligned} & h\big(\mathbb{H}_{h_{22}}^{(n)} X_2^n + \eta_2 W_2^n | \tilde{H}_2^n = \tilde{h}_2^n\big) \\ &= h\Big(X_2^n + \Big(\mathbb{H}_{h_{22}}^{(n)}\Big)^{-1} \eta_2 W_2^n \big| \tilde{H}_2^n = \tilde{h}_2^n\Big) + \log\Big(\big(\mathsf{SNR}_{22}\big)^n\Big) \\ & h\big(\mathbb{H}_{h_{21}}^{(n)} X_2^n + V_1^n | \tilde{H}_1^n = \tilde{h}_1^n\big) \\ &= h\Big(X_2^n + \Big(\mathbb{H}_{h_{21}}^{(n)}\Big)^{-1} V_1^n \big| \tilde{H}_1^n = \tilde{h}_1^n\Big) + \log\Big(\big(\mathsf{SNR}_{21}\big)^n\Big). \end{aligned}$$

Next, we note that since the magnitudes of channel coefficients for each link are equal, then the noise vectors $\Big(\mathbb{H}_{h_{22}}^{(n)}\Big)^{-1} \eta_2 W_2^n$ and $\Big(\mathbb{H}_{h_{21}}^{(n)}\Big)^{-1} V_1^n$ each has i.i.d. elements. Step (b) now follows from Lemma 1 which states that if $\mathsf{SNR}_{21}|\eta_2|^2 \leq \mathsf{SNR}_{22} \cdot (1 - |\tilde{\upsilon}_1|^2)$, then, subject to the trace constraint $\mathrm{tr}\big(\mathrm{cov}(X_2^n)\big) \leq \sum_{i=1}^n q_{2,i} \equiv nP_2$, we have that $h\Big(X_2^n + \Big(\mathbb{H}_{h_{22}}^{(n)}\Big)^{-1} \eta_2 W_2^n \big| \tilde{H}_2^n = \tilde{h}_2^n\Big) - h\Big(X_2^n + \Big(\mathbb{H}_{h_{21}}^{(n)}\Big)^{-1} V_1^n \big| \tilde{H}_1^n = \tilde{h}_1^n\Big)$ is maximized by X_2^n distributed according to a circularly symmetric complex Normal distribution, with i.i.d. elements, each distributed according to $X_{2G} \sim \mathcal{CN}(0, P_2)$. To prove Step (c) recall that $0 \leq P_1, P_2, P_3 \leq 1$; Step (c) then follows since $\Big(h(H_{22} X_{2G} + \eta_2 W_2 | \tilde{H}_2) - h(H_{21} X_{2G} + V_1 | \tilde{H}_1)\Big)$ does not depend on (υ, P_1, P_3) and thus, we can set $\upsilon = 0, P_1 = P_3 = 1$. Additionally, if $\mathsf{SNR}_{21}|\eta_2|^2 \leq \mathsf{SNR}_{22}(1 - |\tilde{\upsilon}_1|^2)$, then the derivative of the expression in step (b) with respect to P_2 is non-negative,[2] and thus, this expression is a non-decreasing function of P_2, from which we conclude that it is maximized with $P_2 = 1$.

[1] For a complex random vector $\mathbf{X}$ and any complex matrix $\mathbb{A}$ it holds that $h\big(\mathbb{A} \cdot \mathbf{X}\big) = h\big(\mathbf{X}\big) + 2 \cdot \log\big|\det(\mathbb{A})\big|$.

[2] The derivative is

$$\begin{aligned} & \frac{\partial}{\partial P_2} \log\left(\frac{\mathsf{SNR}_{22} P_2 + |\eta_2|^2}{\mathsf{SNR}_{21} P_2 + (1 - |\tilde{\upsilon}_2|^2)} \right) \\ &= \frac{\mathsf{SNR}_{21} P_2 + (1 - |\tilde{\upsilon}_2|^2)}{\mathsf{SNR}_{22} P_2 + |\eta_2|^2} \cdot \frac{\mathsf{SNR}_{22}\big(\mathsf{SNR}_{21} P_2 + (1 - |\tilde{\upsilon}_2|^2)\big) - \mathsf{SNR}_{21}\big(\mathsf{SNR}_{22} P_2 + |\eta_2|^2\big)}{\big(\mathsf{SNR}_{21} P_2 + (1 - |\tilde{\upsilon}_2|^2)\big)^2} \\ &= \frac{\mathsf{SNR}_{22}(1 - |\tilde{\upsilon}_2|^2) - \mathsf{SNR}_{21}|\eta_2|^2}{\big(\mathsf{SNR}_{22} P_2 + |\eta_2|^2\big)\big(\mathsf{SNR}_{21} P_2 + (1 - |\tilde{\upsilon}_2|^2)\big)}. \end{aligned}$$

Next, assume that there exists a pair of complex scalars $\tilde{\upsilon}_2$ and η_1 s.t.

$$|\tilde{\upsilon}_2| \leq 1 \tag{19a}$$

and

$$\mathsf{SNR}_{32}|\eta_1|^2 \leq \mathsf{SNR}_{31}\big(1-|\tilde{\upsilon}_2|^2\big) - 2\mathsf{SNR}_{32}\mathsf{SNR}_{11}, \tag{19b}$$

and let V_2^n be an n-dimensional random vector with i.i.d. elements, each distributed according to $V_{2,i} \sim \mathcal{CN}(0, 1-|\tilde{\upsilon}_2|^2)$, $i \in \{1,2,\ldots,n\}$. It now follows that

$$\begin{aligned}
&h(S_1^n|\tilde{H}_1^n) - h(Y_2^n|X_2^n, S_2^n, \tilde{H}_2^n)\\
&\overset{(a)}{=} \mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{h\big(\mathbb{H}_{h_{11}}^{(n)}X_1^n + \mathbb{H}_{h_{31}}^{(n)}X_3^n + \eta_1 W_1^n|\tilde{H}_1^n = \tilde{h}_1^n\big)\\
&\qquad - h\big(\mathbb{H}_{h_{32}}^{(n)}X_3^n + Z_2^n|W_2^n, \tilde{H}_2^n = \tilde{h}_2^n\big)\Big\}\\
&\overset{(b)}{=} \mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{h\big(\mathbb{H}_{h_{11}}^{(n)}X_1^n + \mathbb{H}_{h_{31}}^{(n)}X_3^n + \eta_1 W_1^n|\tilde{H}_1^n = \tilde{h}_1^n\big)\\
&\qquad - h\big(\mathbb{H}_{h_{32}}^{(n)}X_3^n + V_2^n|\tilde{H}_2^n = \tilde{h}_2^n\big)\Big\}\\
&\overset{(c)}{\leq} \mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{h\big(\mathbb{H}_{h_{11}}^{(n)}X_{1\bar{G}}^n + \mathbb{H}_{h_{31}}^{(n)}X_{3\bar{G}}^n + \eta_1 W_1^n|\tilde{H}_1^n = \tilde{h}_1^n\big)\\
&\qquad - h\big(\mathbb{H}_{h_{32}}^{(n)}X_{3\bar{G}}^n + V_2^n|\tilde{H}_2^n = \tilde{h}_2^n\big)\Big\}\\
&\overset{(d)}{\leq} n\cdot\Bigg(h(H_{11}X_{1G} + H_{31}X_{3G} + \eta_1 W_1|\tilde{H}_1)\\
&\qquad - h(H_{32}X_{3G} + V_2|\tilde{H}_2)\Bigg)\Bigg|_{\substack{P_1=P_2=P_3=1\\ \upsilon=0}}.
\end{aligned} \tag{20}$$

In the above transitions (a) follows from the fact that the relay receives transmissions only from Tx$_1$, which makes (X_1^n, X_3^n) necessarily independent of X_2^n, and (b) follows from Lemma 2. For (c) we apply Lemma 6 by first setting

$$X^{2n} = \big((X_1^n)^T, (X_3^n)^T\big)^T, \tag{21a}$$

$$\tilde{\mathbb{V}}_1^H = \big(\mathbb{H}_{h_{11}}^{(n)}, \mathbb{H}_{h_{31}}^{(n)}\big), \tag{21b}$$

$$\tilde{\mathbb{V}}_2^H = \big(\mathbb{O}_{n\times n}, \mathbb{H}_{h_{32}}^{(n)}\big), \tilde{Z}_1^n \triangleq \eta_1 W_1^n, \tilde{Z}_2^n \triangleq V_2^n, \tag{21c}$$

where $\mathbb{O}_{n\times n}$ is an $n\times n$ matrix in which all entries are equal to zero. To determine the matrix $\mathbb{S}$ for the application of Lemma 6 we consider the random vectors $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ *corresponding to the achievable code*, and for $k \in \{1,3\}$, we let $\mathbf{X}_{kR} \triangleq \mathfrak{Re}\{\mathbf{X}_k\}$ and $\mathbf{X}_{kI} \triangleq \mathfrak{Im}\{\mathbf{X}_k\}$ be two $n\times 1$ real random vectors, and $\bar{\mathbf{X}}_k \triangleq (\mathbf{X}_{kR}^T, \mathbf{X}_{kI}^T)^T$ be an $2n \times 1$ real random vector, where $k \in \{1,3\}$. Lastly, we define the random vectors $\bar{\mathbf{X}}_R \triangleq \big(\mathbf{X}_{1R}^T, \mathbf{X}_{3R}^T\big)^T$, $\bar{\mathbf{X}}_I \triangleq \big(\mathbf{X}_{1I}^T, \mathbf{X}_{3I}^T\big)^T$, and $\bar{\mathbf{X}} \triangleq \big((\bar{\mathbf{X}}_R)^T, (\bar{\mathbf{X}}_I)^T\big)^T$, all corresponding to the achievable code. The matrix $\mathbb{S}$ for the application of Lemma 6 is determined via

$$\begin{aligned}
\mathbb{S} \triangleq \text{cov}(\bar{\mathbf{X}}) &= \begin{bmatrix} \mathbb{Q}_{RR} & \mathbb{Q}_{RI} \\ \mathbb{Q}_{RI}^T & \mathbb{Q}_{II}\end{bmatrix},\\
\mathbb{Q}_{RR} &= \text{cov}\big(\bar{\mathbf{X}}_R\big),\\
\mathbb{Q}_{II} &= \text{cov}\big(\bar{\mathbf{X}}_I\big)\\
\mathbb{Q}_{RI} &= \text{cov}\big(\bar{\mathbf{X}}_R, \bar{\mathbf{X}}_I\big).
\end{aligned}$$

Lastly we note that $\tilde{\mathbb{V}}_1^H \cdot \tilde{\mathbb{V}}_1 = (\mathsf{SNR}_{11} + \mathsf{SNR}_{31})\cdot \mathbb{I}_n$ and $\tilde{\mathbb{V}}_2^H \cdot \tilde{\mathbb{V}}_2 = \mathsf{SNR}_{32}\cdot\mathbb{I}_n$, which satisfies the conditions of Lemma 6. It thus follows that $\Big(\big(X_{1\bar{G}}^n\big)^T, \big(X_{3\bar{G}}^n\big)^T\Big)^T = X_{\bar{G}}^{2n}$ is a $2n\times 1$ zero mean *complex* Normal random vector whose covariance matrix satisfies $\text{cov}\Big(\big(\mathfrak{Re}\{X_{\bar{G}}^{2n}\}\big)^T, \big(\mathfrak{Im}\{X_{\bar{G}}^{2n}\}\big)^T\Big) \preceq \mathbb{S}$. Consequently, we have that the maximizing $\big(X_{1\bar{G}}, X_{3\bar{G}}\big)$, obtained from the optimal $X_{\bar{G}}^{2n}$, satisfy for $k=1,3$:

$$\mathbb{E}\big\{|X_{k\bar{G},i}|^2\big\} \leq \mathbb{E}\big\{|X_{k,i}|^2\big\} = 1 \tag{22a}$$

$$\begin{aligned}
\text{tr}\big\{\text{cov}(X_{k\bar{G}}^n)\big\} &= \text{tr}\Big\{\text{cov}\big(\mathfrak{Re}\big\{X_{k\bar{G}}^n\big\}\big)\Big\} + \text{tr}\Big\{\text{cov}\big(\mathfrak{Im}\big\{X_{k\bar{G}}^n\big\}\big)\Big\}\\
&\leq \text{tr}\big\{\text{cov}(X_{kR})\big\} + \text{tr}\big\{\text{cov}(X_{kI})\big\}\\
&= \sum_{i=1}^n q_{k,i}\\
&\leq n,
\end{aligned} \tag{22b}$$

where $q_{k,i}$ denotes the variance of *complex* symbol X_k at time index $i \in \{1,2,\ldots,n\}$, $k \in \{1,3\}$, in the achievable code. Lastly, Step (d) is proved in Appendix E using relationships (22).

Plugging (18) and (20) into the second line of (17), we conclude that if it is possible to choose $\tilde{\upsilon}_1, \tilde{\upsilon}_2, \eta_1$ and η_2 s.t.

$$\mathsf{SNR}_{32}|\eta_1|^2 \leq \mathsf{SNR}_{31}\big(1-|\tilde{\upsilon}_2|^2\big) - 2\mathsf{SNR}_{32}\mathsf{SNR}_{11} \tag{23a}$$

$$\mathsf{SNR}_{21}|\eta_2|^2 \leq \mathsf{SNR}_{22}\big(1-|\tilde{\upsilon}_1|^2\big), \tag{23b}$$

then the sum-rate is upper bounded by

$$\begin{aligned}
&n(R_1 + R_2 - \epsilon_{1n} - \epsilon_{2n})\\
&\leq n\cdot\Bigg(h(Y_{1G}|S_{1G}, \tilde{H}_1) - h(S_{1G}|X_{1G}, X_{3G}, \tilde{H}_1)\\
&\qquad + h(Y_{2G}|S_{2G}, \tilde{H}_2) - h(S_{2G}|X_{2G}, \tilde{H}_2)\Bigg)\\
&\quad + n\cdot\Bigg(h(H_{22}X_{2G} + \eta_2 W_2|\tilde{H}_2) - h(H_{21}X_{2G} + V_1|\tilde{H}_1)\\
&\qquad + h(H_{11}X_{1G} + H_{31}X_{3G} + \eta_1 W_1|\tilde{H}_1)\\
&\qquad - h(H_{32}X_{3G} + V_2|\tilde{H}_2)\Bigg)\\
&\overset{(a)}{=} n\cdot\Bigg(h(Y_{1G}|S_{1G}, \tilde{H}_1) - h(S_{1G}|X_{1G}, X_{3G}, \tilde{H}_1)\\
&\qquad + h(Y_{2G}|S_{2G}, \tilde{H}_2) - h(S_{2G}|X_{2G}, \tilde{H}_2)\\
&\qquad + h(S_{2G}|\tilde{H}_2)\\
&\qquad - h(H_{11}X_{1G} + H_{21}X_{2G} + H_{31}X_{3G} + Z_1|X_{1G}, X_{3G}, W_1, \tilde{H}_1)\\
&\qquad + h(S_{1G}|\tilde{H}_1)\\
&\qquad - h(H_{22}X_{2G} + H_{32}X_{3G} + Z_2|X_{2G}, W_2, \tilde{H}_2)\Bigg)\\
&= n\cdot\Bigg(h(Y_{1G}|S_{1G}, \tilde{H}_1) + I(X_{1G}, X_{3G}; S_{1G}|\tilde{H}_1)\\
&\qquad + h(Y_{2G}|S_{2G}, \tilde{H}_2) + I(X_{2G}; S_{2G}|\tilde{H}_2)\\
&\qquad - h(Y_{1G}|X_{1G}, X_{3G}, S_{1G}, \tilde{H}_1)\\
&\qquad - h(Y_{2G}|X_{2G}, S_{2G}, \tilde{H}_2)\Bigg)\\
&= n\cdot\Big(I(X_{1G}, X_{3G}; Y_{1G}, S_{1G}|\tilde{H}_1) + I(X_{2G}; Y_{2G}, S_{2G}|\tilde{H}_2)\Big),
\end{aligned}$$

s.t. all the expressions are evaluated with circularly symmetric complex Normal channel inputs $(X_{1G}, X_{2G}, X_{3G})^T \sim$

$\mathcal{CN}(\mathbf{0}, \mathbb{Q}_G^{\text{Opt}})$ where

$$\mathbb{Q}_G^{\text{Opt}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In the above transitions, (a) follows from Lemma 2. It thus follows that for n large enough, the sum-rate capacity is upper-bounded by:

$$\sup_{(R_1,R_2)\in\mathcal{C}(\underline{\mathsf{SNR}})} (R_1 + R_2) \le I(X_{1G}, X_{3G}; Y_{1G}, S_{1G}|\tilde{H}_1) + I(X_{2G}; Y_{2G}, S_{2G}|\tilde{H}_2), \quad (24)$$

where $(X_{1G}, X_{2G}, X_{3G})^T \sim \mathcal{CN}(\mathbf{0}, \mathbb{Q}_G^{\text{Opt}})$. To complete the proof of Theorem 2, note that

$$\begin{aligned} I(X_{1G}, X_{3G}; Y_{1G}, S_{1G}|\tilde{H}_1) &= I(X_{1G}, X_{3G}; Y_{1G}|\tilde{H}_1) \\ &\quad + I(X_{1G}, X_{3G}; S_{1G}|Y_{1G}, \tilde{H}_1) \\ I(X_{2G}; Y_{2G}, S_{2G}|\tilde{H}_2) &= I(X_{2G}; Y_{2G}|\tilde{H}_2) \\ &\quad + I(X_{2G}; S_{2G}|Y_{2G}, \tilde{H}_2). \end{aligned}$$

Hence, if we find conditions under which

$$I(X_{1G}, X_{3G}; S_{1G}|Y_{1G}, \tilde{H}_1) = 0 \quad (25a)$$
$$I(X_{2G}; S_{2G}|Y_{2G}, \tilde{H}_2) = 0, \quad (25b)$$

then, when these conditions are satisfied, an upper bound on the sum-rate capacity is given by

$$\sup_{(R_1,R_2)\in\mathcal{C}(\underline{\mathsf{SNR}})} (R_1 + R_2) \le I(X_{1G}, X_{3G}; Y_{1G}|\tilde{H}_1) + I(X_{2G}; Y_{2G}|\tilde{H}_2), \quad (26)$$

where $(X_{1G}, X_{2G}, X_{3G})^T \sim \mathcal{CN}(\mathbf{0}, \mathbb{Q}_G^{\text{Opt}})$. To that aim, note that from (25a) we obtain

$$\begin{aligned} &I(X_{1G}, X_{3G}; S_{1G}|Y_{1G}, \tilde{H}_1) = 0 \\ &\Leftrightarrow \mathbb{E}_{\tilde{H}_1}\Big\{I(X_{1G}, X_{3G}; h_{11}X_{1G} + h_{31}X_{3G} + \eta_1 W_1|h_{11}X_{1G} \\ &\qquad + h_{31}X_{3G} + h_{21}X_{2G} + Z_1, \tilde{H}_1 = \tilde{h}_1)\Big\} = 0. \end{aligned}$$

From Lemma 4 we conclude that this is satisfied if for all values of h_{lk} it holds that

$$\begin{aligned} \mathbb{E}\big\{(\eta_1 W_1)(h_{21}X_{2G} + Z_1)^*\big\} &= \mathbb{E}\big\{|h_{21}X_{2G} + Z_1|^2\big\} \\ \Leftrightarrow \eta_1 \tilde{v}_1 &= 1 + \mathsf{SNR}_{21}. \end{aligned} \quad (27)$$

Applying the same arguments to (25b), we obtain

$$\begin{aligned} &I(X_{2G}; S_{2G}|Y_{2G}, \tilde{H}_2) = 0 \\ &\Leftrightarrow \mathbb{E}_{\tilde{H}_2}\Big\{I(X_{2G}; h_{22}X_{2G} + \eta_2 W_2|h_{22}X_{2G} \\ &\qquad + h_{32}X_{3G} + Z_2, \tilde{H}_2 = \tilde{h}_2)\Big\} = 0 \\ &\overset{(a)}{\Leftrightarrow} \mathbb{E}\big\{(\eta_2 W_2)(h_{32}X_{3G} + Z_2)^*\big\} = \mathbb{E}\big\{|h_{32}X_{3G} + Z_2|^2\big\} \\ &\Leftrightarrow \eta_2 \tilde{v}_2 = 1 + \mathsf{SNR}_{32}, \end{aligned} \quad (28)$$

where (a) follows again from Lemma 4.
Combining (27) and (28) with (23), we conclude that (26) constitutes an upper-bound on the sum-rate capacity if it is possible to construct a genie signal with parameters $\tilde{v}_1$, $\tilde{v}_2$, η_1 and η_2 s.t.

$$\begin{aligned} \mathsf{SNR}_{32}|\eta_1|^2 &\le \mathsf{SNR}_{31}\big(1 - |\tilde{v}_2|^2\big) - 2\mathsf{SNR}_{32}\mathsf{SNR}_{11} \\ \mathsf{SNR}_{21}|\eta_2|^2 &\le \mathsf{SNR}_{22}\big(1 - |\tilde{v}_1|^2\big) \\ \eta_1 \tilde{v}_1 &= 1 + \mathsf{SNR}_{21} \\ \eta_2 \tilde{v}_2 &= 1 + \mathsf{SNR}_{32}. \end{aligned}$$

We note that this can be done if

$$\mathsf{SNR}_{32}(1 + \mathsf{SNR}_{21})^2 \le |\tilde{v}_1|^2\Big(\mathsf{SNR}_{31}\big(1 - |\tilde{v}_2|^2\big) - 2\mathsf{SNR}_{32}\mathsf{SNR}_{11}\Big) \quad (29a)$$

$$\mathsf{SNR}_{21}(1 + \mathsf{SNR}_{32})^2 \le |\tilde{v}_2|^2\Big(\mathsf{SNR}_{22}\big(1 - |\tilde{v}_1|^2\big)\Big). \quad (29b)$$

In conclusion, if we can find two complex scalars $\tilde{v}_1$ and $\tilde{v}_2$ s.t. $0 \le |\tilde{v}_1|, |\tilde{v}_2| \le 1$, for which (29) is satisfied, then an upper-bound on the sum-rate capacity is given by

$$\sup_{(R_1,R_2)\in\mathcal{C}(\underline{\mathsf{SNR}})} (R_1 + R_2) \le \Big\{I(X_{1G}, X_{3G}; Y_{1G}|\tilde{H}_1) + I(X_{2G}; Y_{2G}|\tilde{H}_2)\Big\}, \quad (30)$$

where $X_{kG} \sim \mathcal{CN}(0, 1), k \in \{1, 2, 3\}$, mutually independent. The proof of Theorem 2 is completed by identifying $\beta_1 \equiv |\tilde{v}_1|^2$ and $\beta_2 \equiv |\tilde{v}_2|^2$. ■

D. Step 2: An Achievable Rate Region

We next characterize an achievable rate region for the ergodic phase fading Z-ICR. This region is stated in the following proposition:

Proposition 2: Consider the ergodic phase fading Z-ICR with only Rx-CSI, defined in Section II. Let the channel inputs be generated i.i.d. in time according to $X_k \sim \mathcal{CN}(0, 1)$, $k \in \{1, 2, 3\}$, *mutually independent. If it holds that*

$$I(X_1, X_3; Y_1|\tilde{H}_1) \le I(X_1; Y_3|X_3, \tilde{H}_3), \quad (31)$$

then an achievable rate region for the Z-ICR is given by all the non-negative rate pairs (R_1, R_2) *satisfying*

$$R_1 \le I(X_1, X_3; Y_1|\tilde{H}_1) \quad (32a)$$
$$R_2 \le I(X_2; Y_2|\tilde{H}_2). \quad (32b)$$

Proof: The achievability is based on the DF strategy at the relay. Fix the blocklength n and the input distribution $f_{X_1,X_2,X_3}(x_1, x_2, x_3) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot f_{X_3}(x_3)$, with $X_k \sim \mathcal{CN}(0, 1), k = 1, 2, 3$. We employ a transmission scheme in which $B - 1$ messages are transmitted using nB channel symbols:

a) Code Construction: For each message $m_k \in \mathcal{M}_k$, $k \in \{1, 2\}$ select a codeword $\mathbf{x}_k(m_k)$ according to the p.d.f. $f_{\mathbf{X}_k}\big(\mathbf{x}_k(m_k)\big) = \prod_{i=1}^n f_{X_k}\big(x_{k,i}(m_k)\big)$. For each $\tilde{m}_1 \in \mathcal{M}_1$ select a codeword $\mathbf{x}_3(\tilde{m}_1)$ according to the p.d.f. $f_{\mathbf{X}_3}\big(\mathbf{x}_3(\tilde{m}_1)\big) = \prod_{i=1}^n f_{X_3}\big(x_{3,i}(\tilde{m}_1)\big)$.

b) Encoding at Block b: At block b, Tx_k transmits the message $m_{k,b}$ via the codeword $\mathbf{x}_k(m_{k,b}), k \in \{1,2\}$. Let $\hat{m}_{1,b-1}$ denote the decoded message at the relay at block $b-1$. At block b, the relay transmits the codeword $\mathbf{x}_3(\hat{m}_{1,b-1})$. At block $b=1$ the relay transmits the codeword $\mathbf{x}_3(1)$, and at block $b=B$, Tx_1 and Tx_2 transmit the codewords $\mathbf{x}_1(1)$ and $\mathbf{x}_2(1)$, respectively.

c) Decoding at the Relay: The decoding process at the relay is similar to the one used for deriving the capacity result in [22, Sec. VII-D]. For decoding $m_{1,b}$, the decoder at the relay looks for a unique, $m_1 \in \mathcal{M}_1$ that satisfies:

$$\Big(\mathbf{x}_1(m_1), \mathbf{x}_3(\hat{m}_{1,b-1}), \mathbf{y}_3(b), \tilde{\mathbf{h}}_3(b)\Big) \in \mathcal{A}_\epsilon^{(n)}(X_1, X_3, Y_3, \tilde{H}_3).$$

From [22, eq. (15)] it directly follows that the relay can decode reliably if n is large enough, as long as

$$R_1 \leq I(X_1; Y_3, \tilde{H}_3|X_3) \stackrel{(a)}{=} I(X_1; Y_3|X_3, \tilde{H}_3),$$

where (a) holds since the channel coefficients are independent of the transmitted symbols.

d) Decoding at Rx_1: Rx_1 uses a backward block decoding scheme as in [22, Appendix A] while treating the signal from Tx_2 as additive noise (recall that the codebooks are generated independently). Assume the relay has correctly decoded all the messages $\{m_{1,b}\}_{b=1}^{B-1}$. Assuming that Rx_1 has correctly decoded $m_{1,b+1}$, then, in order to decode $m_{1,b}$, Rx_1 generates the sets:

$$\begin{aligned}\mathcal{E}_{0,b} \triangleq \Big\{\hat{m}_1 \in \mathcal{M}_1 : \big(&\mathbf{x}_1(m_{1,b+1}), \mathbf{x}_3(\hat{m}_1), \mathbf{y}_1(b+1),\\ &\tilde{\mathbf{h}}_1(b+1)\big) \in \mathcal{A}_\epsilon^{(n)}(X_1, X_3, Y_1, \tilde{H}_1)\Big\},\\ \mathcal{E}_{1,b} \triangleq \Big\{\hat{m}_1 \in \mathcal{M}_1 : \big(&\mathbf{x}_1(\hat{m}_1), \mathbf{y}_1(b),\\ &\tilde{\mathbf{h}}_1(b)\big) \in \mathcal{A}_\epsilon^{(n)}(X_1, Y_1, \tilde{H}_1)\Big\}.\end{aligned}$$

Rx_1 then decodes $m_{1,b}$ by finding a unique $m_1 \in \mathcal{E}_{0,b} \cap \mathcal{E}_{1,b}$. Note that since the codewords are independent of each other, error events associated with $\mathcal{E}_{0,b}$ are independent of error events associated with $\mathcal{E}_{1,b}$. Thus, by using standard joint-typicality arguments [31, Th. 7.6.1], it follows that decoding can be done reliably by taking n large enough as long as

$$\begin{aligned}R_1 &\leq I(X_1; Y_1, \tilde{H}_1) + I(X_3; Y_1, \tilde{H}_1|X_1)\\ &= I(X_1, X_3; Y_1, \tilde{H}_1)\\ &= I(X_1, X_3; Y_1|\tilde{H}_1).\end{aligned}$$

e) Decoding at Rx_2: Rx_2 treats the signal from the relay as additive noise. This can be done since the codebooks are generated independently. The decoder at Rx_2 is therefore the decoder for PtP channels: At block b the decoder looks for a unique message $m_2 \in \mathcal{M}_2$ that satisfies

$$\Big(\mathbf{x}_2(m_2), \mathbf{y}_2(b), \tilde{\mathbf{h}}_2(b)\Big) \in \mathcal{A}_\epsilon^{(n)}(X_2, Y_2, \tilde{H}_2).$$

It thus follows from [31, Th. 9.1.1] that Rx_2 can reliably decode $m_{2,b}$ if n is large enough, as long as

$$R_2 \leq I(X_2; Y_2, \tilde{H}_2) = I(X_2; Y_2|\tilde{H}_2).$$

Finally, we observe that if (31) is satisfied, i.e., if $I(X_1, X_3; Y_1|\tilde{H}_1) \leq I(X_1; Y_3|X_3, \tilde{H}_3)$, then the decoder at the relay can reliably decode the signal from Tx_1 whenever Rx_1 can. Consequently, we conclude that any rate pair inside the region specified in (32) is achievable. ■

E. Step 3: The Sum-Rate Capacity in the WI Regime

Note that when the conditions in (10) and (31) hold (corresponding to conditions (8) and (7) in Thm. 1, respectively), then the *upper bound* on the sum-rate in (11) coincides with the *achievable* sum-rate obtained from (32), where both sum-rate expressions are evaluated with mutually independent channel inputs, distributed according to $X_k \sim \mathcal{CN}(0,1)$, $k \in \{1,2,3\}$. This results in a characterization of the *sum-rate capacity* for the ergodic phase fading Z-ICR. The proof of Theorem 1 is completed by observing that for mutually independent channel inputs distributed according to $X_k \sim \mathcal{CN}(0,1), k \in \{1,2,3\}$, the mutual information expressions in (11) and (31) are explicitly written as

$$\begin{aligned}I(X_1, X_3; Y_1|\tilde{H}_1) &= \log\left(1 + \frac{\mathsf{SNR}_{11} + \mathsf{SNR}_{31}}{1 + \mathsf{SNR}_{21}}\right)\\ I(X_2; Y_2|\tilde{H}_2) &= \log\left(1 + \frac{\mathsf{SNR}_{22}}{1 + \mathsf{SNR}_{32}}\right)\\ I(X_1; Y_3|X_3, \tilde{H}_3) &= \log(1 + \mathsf{SNR}_{13}),\end{aligned}$$

from which the explicit expressions (7) and (9) are obtained. ■

F. Comments

Comment 5: Consider the scenario in which the relay is off, referred to as the Z-IC [21, Th. 2]. This scenario can be obtained from the Z-ICR by letting $\mathsf{SNR}_{31} = \mathsf{SNR}_{32} = 0$. In this case, since decoding at the relay does not constrain the rates, from Theorem 1 we conclude that if $\mathsf{SNR}_{21} \leq \mathsf{SNR}_{22}$, then the sum-rate capacity of the ergodic phase fading Z-IC is given by

$$\begin{aligned}\sup_{\substack{(R_1,R_2)\in\mathcal{C}(\underline{\mathsf{SNR}}),\\ \text{s.t. } \mathsf{SNR}_{31}=\mathsf{SNR}_{32}=0}} (R_1 + R_2) &= \log\left(1 + \frac{\mathsf{SNR}_{11}}{1 + \mathsf{SNR}_{21}}\right)\\ &\quad + \log\Big(1 + \mathsf{SNR}_{22}\Big)\\ &\triangleq C_{\text{sum}}^{\text{PF-Z-IC}}(\underline{\mathsf{SNR}}),\end{aligned} \tag{33}$$

which is similar to the sum-rate capacity expression for the AWGN Z-IC in the WI regime characterized in [21, Th. 2] (although in the current work the channel is subject to ergodic phase fading).

Comment 6: An interesting question that arises is whether adding a relay node to the Z-IC increases the sum-rate in the WI regime, when the interfering signal is treated as noise at each receiver. In Fig. 2 we show that the answer to this question is positive.

Fig. 2 depicts the sum-rate of (32) with and without a relay (as discussed in Comment 5, turning off the relay is achieved by setting $\mathsf{SNR}_{31} = \mathsf{SNR}_{32} = 0$ in Eqns. (8) and (9)), for Z-ICR scenarios in which condition (31), or equivalently (7), is satisfied. We consider a symmetric setting by letting $\mathsf{SNR}_{11} = \mathsf{SNR}_{22} = \mathsf{SNR}_{31} = \mathsf{SNR}_{\text{d}}$, and

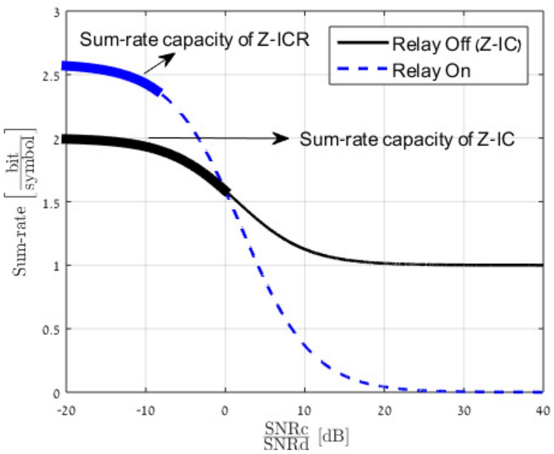


Fig. 2. The sum-rates of Proposition 2 and of (33) for the scenario in which $\mathsf{SNR}_{11} = \mathsf{SNR}_{22} = \mathsf{SNR}_{31} \triangleq \mathsf{SNR}_{\rm d} = 1$, and $\mathsf{SNR}_{21} = \mathsf{SNR}_{32} \triangleq \mathsf{SNR}_{\rm c}$.

$\mathsf{SNR}_{21} = \mathsf{SNR}_{32} = \mathsf{SNR}_{\rm c}$. Thus, $\mathsf{SNR}_{\rm c}$ and $\mathsf{SNR}_{\rm d}$ denote the strengths of the interfering links and of the links carrying desired information, respectively, and the relative strength of the interference is given as $\frac{\mathsf{SNR}_{\rm c}}{\mathsf{SNR}_{\rm d}}$. It can be seen from the figure that when the interference is sufficiently weak, the relay increases the sum-rate, which follows as the rate increase for Tx_1-Rx_1 is greater than the rate decrease for Tx_2-Rx_2.

At interference levels, $\frac{\mathsf{SNR}_{\rm c}}{\mathsf{SNR}_{\rm d}}$, which correspond to the thick lines in each plot, treating the interfering signal as noise is sum-rate optimal, and the resulting achievable sum-rates from (32) and (33) correspond to the sum-rate capacities for the Z-ICR and for the Z-IC, respectively. Thus, it is evident from Fig. 2 that in some scenarios, adding a relay node and employing the communications scheme described in the proof of Proposition 2, *strictly increases* the sum-rate capacity of the ergodic phase fading Z-IC in the WI regime, $C_{\rm sum}^{\text{PF-Z-IC}}(\underline{\mathsf{SNR}})$. In particular, we observe that for the symmetric scenario of Fig. 2, adding a relay strictly increases the sum-rate *capacity* as long as $\frac{\mathsf{SNR}_{\rm c}}{\mathsf{SNR}_{\rm d}} < 0$ [dB], and for sufficiently weak interference, e.g., $\frac{\mathsf{SNR}_{\rm c}}{\mathsf{SNR}_{\rm d}} < -10$ [dB], DF at the relay achieves the sum-rate capacity of the Z-ICR.

Comment 7: Note that for the set of channel coefficients satisfying (7) and (8), the sum-rate capacity stated in (9) is an upper bound on the sum-rate capacity of the ergodic phase fading ICR (*with both interfering links active*) in the weak interference regime, when the relay node receives transmissions only from Tx_1 (as is the case in Theorem 1). If, in addition, the relay node receives the transmissions of Tx_2, then a new coding strategy must be developed for the WI regime in order to facilitate simultaneous enhancement of the desired signal at both destinations. Finding the optimal scheme and the corresponding sum-rate capacity is currently an open issue that requires further research.

Comment 8: Fig. 3 shows the region of relay locations in the 2D-plane in which DF at the relay achieves the sum-rate capacity of the ergodic phase fading Z-ICR in the WI regime. This figure was obtained using a channel model in which the attenuation $\sqrt{\mathsf{SNR}_{ij}}$ is linked to the distance from node i to node j, d_{ij}, via $\mathsf{SNR}_{ij} = \frac{1}{d_{ij}^4}$. This attenuation model corresponds to the two-ray propagation model. Note that since the signal from the relay is desired at Rx_1 and is treated as noise at Rx_2, then for the WI conditions to hold, the relay should be closer to Rx_1 (to strengthen the desired signal at Rx_1) and farther away from Rx_2 (to decrease the interference at Rx_2). However, the relay should remain relatively close to Tx_1 to allow reliable decoding of the messages from Tx_1 at the relay.

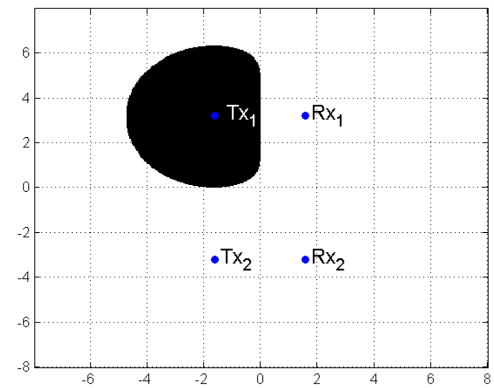


Fig. 3. The geographical position of the relay in a 2D-plane where the conditions of Theorem 1 are satisfied.

IV. Asymptotic SNR Analysis: The Maximal GDoF in the WI Regime

In this section, we characterize the maximal GDoF of the ergodic phase fading Z-ICR in the WI regime. Since GDoF analysis characterizes the performance in the asymptotically high SNR regime, i.e., $\mathsf{SNR}_{lk} \to \infty$ for all links, then, in order to analyze the effect of different link conditions, we consider a scenario in which the magnitudes of the channel coefficients scale differently as a function of the SNR. Letting α, β, γ and λ be four non-negative real numbers, in this section we consider a model in which

$$\mathsf{SNR}_{11} = \mathsf{SNR}, \quad \mathsf{SNR}_{22} = \mathsf{SNR}, \tag{34a}$$
$$\mathsf{SNR}_{21} = \mathsf{SNR}^{\alpha}, \quad \mathsf{SNR}_{32} = \mathsf{SNR}^{\lambda} \tag{34b}$$
$$\mathsf{SNR}_{13} = \mathsf{SNR}^{\gamma}, \quad \mathsf{SNR}_{31} = \mathsf{SNR}^{\beta}. \tag{34c}$$

Observe that the direct links scale as SNR, the interfering links from Tx_2 to Rx_1, and from the relay to Rx_2 scale as SNR^{α} and SNR^{λ}, respectively, and the links on the cooperation path from Tx_1 to the relay, and from the relay to Rx_1 scale as SNR^{γ} and SNR^{β}, respectively. Let $\mathcal{C}(\mathsf{SNR})$ denote the capacity region of the Z-ICR for a given value of the parameter SNR, and define $C_{\rm sum}(\mathsf{SNR}) \triangleq \max_{(R_1,R_2)\in\mathcal{C}(\mathsf{SNR})}\big(R_1 + R_2\big)$. Then, the GDoF is defined as (see also [18, Definition 1]):

$$\text{GDoF} \triangleq \lim_{\mathsf{SNR}\to\infty} \frac{C_{\rm sum}(\mathsf{SNR})}{\log\left(\mathsf{SNR}\right)}.$$

In the following theorem, we characterize the maximal GDoF of the ergodic phase fading Z-ICR in the WI regime:

Theorem 3: Consider the ergodic phase fading Z-ICR with only Rx-CSI, defined in Section II. If the interference is symmetric and weak in the sense of

$$\lambda = \alpha \leq \frac{1}{2}, \tag{35a}$$

and it also holds that

$$1 + 2\alpha < \beta \leq \gamma + \alpha, \tag{35b}$$

then the maximal GDoF of the channel is

$$\text{GDoF}_{\max} = 1 + \beta - 2\alpha, \tag{36}$$

and it is achieved with mutually independent, zero mean complex Normal channel inputs with positive powers satisfying $0 < P_k \leq 1,\ k \in \{1, 2, 3\}$.

Comment 9: In the following we intuitively explain conditions (35) in Thm. 3. Note that (35a) corresponds to the weak interference regime in the sense of [4] and [18], namely, that the interfering links are *exponentially weaker* than the direct links in the sense that $\lim_{\mathsf{SNR}\to\infty} \frac{\mathsf{SNR}^{\alpha}}{\mathsf{SNR}} = 0$. We note that while the results of [4] and [18] for the symmetric scenario hold as long as the scaling exponent of the interfering links satisfies $\alpha \leq 1$, the GDoF optimality result of Thm. 3 requires $\lambda = \alpha \leq \frac{1}{2}$, i.e., Thm. 3 requires a smaller exponential scaling of the interference strength, compared to the minimal exponential scaling of the interference required for WI in [4] and [18]. It follows that the WI regime for the GDoF result of Thm. 3 corresponds to a subset of the WI regime applicable for the results of [4] and [18]. Yet, we note that in [4], GDoF optimality of *treating interference as noise* was shown to hold only for $\alpha \leq \frac{1}{2}$, which is in agreement with our characterization (see [4, Sec. V-B]). Next, consider (35b): Observe that (35b) can be written as $1 + \alpha < \beta - \alpha \leq \gamma$, which is equivalent to the inequality $\mathsf{SNR}^{1+\alpha} < \frac{\mathsf{SNR}^{\beta}}{\mathsf{SNR}^{\alpha}} \leq \mathsf{SNR}^{\gamma}$. Note that $\frac{\mathsf{SNR}^{\beta}}{\mathsf{SNR}^{\alpha}} \leq \mathsf{SNR}^{\gamma}$ implies that the relay reception is good enough such that the SNR on the incoming link at the relay, (i.e., SNR^{γ}) is higher than the SNR on the link from the relay to Rx_1, when interference is treated as additive noise at Rx_1 (i.e., $\frac{\mathsf{SNR}^{\beta}}{\mathsf{SNR}^{\alpha}}$). The inequality $\mathsf{SNR}^{1+\alpha} < \frac{\mathsf{SNR}^{\beta}}{\mathsf{SNR}^{\alpha}}$ implies that interference should be weak enough s.t. the SNR on the link from the relay to Rx_1, achieved by treating interference as additive noise at Rx_1 (i.e., $\frac{\mathsf{SNR}^{\beta}}{\mathsf{SNR}^{\alpha}}$), will be higher than the SNR of the direct link from Tx_1 to Rx_1 augmented by the interference at Rx_1, (i.e., $\mathsf{SNR}^{1+\alpha}$). Hence, the second inequality represents an additional weak interference condition.

Proof: The proof of Thm. 3 consists of the following steps:

1) We derive an upper bound on the GDoF of the ergodic phase fading Z-ICR by combining two bounds: A bound derived using a genie, and a bound obtained by following the derivations of the cut-set bound theorem [31, Th. 15.10.1].
2) We derive a lower bound on the GDoF by considering the communications scheme used in Section III-D.
3) We derive conditions on the SNR exponents of the channel coefficients under which our lower bound coincides with the upper bound, thereby, characterizing the maximal GDoF of the ergodic phase fading Z-ICR in the weak interference regime, subject to these conditions.

In the following subsections, we provide a detailed proof for the above steps. Specifically, Steps 1 is carried out in Subsection IV-A, Step 2 is carried out in Subsection IV-B, and finally, Step 3 is detailed in Subsection IV-C.

A. An Upper Bound on the Achievable GDoF

An upper bound on the achievable GDoF of the Z-ICR is stated in the following theorem:

Theorem 4: Consider the ergodic phase fading Z-ICR with only Rx-CSI, stated in Section II. If $\beta > 2\lambda + 1$, *then an upper bound on the achievable GDoF is given by*

$$\begin{aligned}\text{GDoF}^{+} = \min\Big\{ &\max\big\{2, 1 + \min\{\beta, \gamma\}\big\},\\ &\max\{\alpha + \lambda, \beta, 1 + \beta - \alpha - \lambda\}\Big\}.\end{aligned} \tag{38}$$

Proof: The upper bound is obtained as a combination of two bounds: The first bound is derived using a genie, and the second bound is derived by following the derivation of the cut-set theorem [31, Th. 15.10.1].

1) An Upper Bound Using a Genie: Consider the following genie signals:

$$\begin{aligned}S_{1,i} &= H_{32,i}X_{3,i} + Z_{2,i}\\ S_{2,i} &= H_{21,i}X_{2,i} + Z_{1,i},\end{aligned}$$

$i \in \{1, 2, \ldots, n\}$. Suppose that a genie provides $\{S_{1,i}\}_{i=1}^{n}$ to Rx_1 and $\{S_{2,i}\}_{i=1}^{n}$ to Rx_2, i.e., the genie provides to Rx_2 an interference-free, noisy version of its desired signal as it is received at Rx_1, and to Rx_1 it provides a noisy version of the relay signal component observed at Rx_2. Let M_k denote the message transmitted from Tx_k, and let $\hat{M}_k$ denote the decoded message at Rx_k. Additionally, let $P_{e,k}^{(n)}$ denote the probability of error in the estimation of M_k at Rx_k and define $n\epsilon_{kn} \triangleq 1 + P_{e,k}^{(n)} nR_k,\ k \in \{1, 2\}$. Then, for an achievable rate pair (R_1, R_2), we obtain:

$$\begin{aligned}nR_1 &= H(M_1)\\ &= H(M_1) - H(M_1|Y_1^n, \underline{\tilde{H}}^n) + H(M_1|Y_1^n, \underline{\tilde{H}}^n)\\ &\stackrel{(a)}{\leq} I(M_1; Y_1^n, \underline{\tilde{H}}^n) + n\epsilon_{1n}\\ &\stackrel{(b)}{\leq} I(X_1^n; Y_1^n, \underline{\tilde{H}}^n) + n\epsilon_{1n}\\ &\stackrel{(c)}{=} I(X_1^n; Y_1^n|\underline{\tilde{H}}^n) + n\epsilon_{1n}\\ &\stackrel{(d)}{\leq} I(X_1^n; Y_1^n|\underline{\tilde{H}}^n) + I(X_3^n; Y_1^n|\underline{\tilde{H}}^n, X_1^n)\\ &\quad + I(X_1^n, X_3^n; S_1^n|\underline{\tilde{H}}^n, Y_1^n) + n\epsilon_{1n}\\ &= h(S_1^n|\underline{\tilde{H}}^n) - h(S_1^n|X_1^n, X_3^n, \underline{\tilde{H}}^n) + h(Y_1^n|S_1^n, \underline{\tilde{H}}^n)\\ &\quad - h(Y_1^n|X_1^n, X_3^n, S_1^n, \underline{\tilde{H}}^n) + n\epsilon_{1n}\\ &\stackrel{(e)}{=} h(S_1^n|\underline{\tilde{H}}^n) - h(Z_2^n)\\ &\quad + h(Y_1^n|S_1^n, \underline{\tilde{H}}^n) - h(S_2^n|\underline{\tilde{H}}^n) + n\epsilon_{1n},\end{aligned} \tag{39}$$

where (a) follows from Fano's inequality [31, Th. 2.10.1], (b) follows from the data processing inequality [31, Th.

2.8.1] as $M_1 - X_1^n - (Y_1^n, \tilde{H}_1^n)$ forms a Markov chain, (c) follows since channel inputs X_1^n are independent of the channel coefficients $\tilde{H}_1^n$, (d) follows since mutual information is non-negative, and (e) follows since

$$\begin{aligned} h(Y_1^n|X_1^n, X_3^n, S_1^n, \underline{\tilde{H}}^n) &= h(Y_1^n|X_1^n, X_3^n, Z_2^n, \underline{\tilde{H}}^n) \\ &= h\big(\{H_{21,i}X_{2,i} + Z_{1,i}\}_{i=1}^n|X_1^n, X_3^n, Z_2^n, \underline{\tilde{H}}^n\big) \\ &\stackrel{(f)}{=} h\big(\{H_{21,i}X_{2,i} + Z_{1,i}\}_{i=1}^n|\underline{\tilde{H}}^n\big) \\ &\equiv h(S_2^n|\underline{\tilde{H}}^n), \end{aligned}$$

where (f) follows since X_1^n and X_3^n are independent of X_2^n, which follows since the message sets at the sources are mutually independent, and since the relay receives transmissions only from Tx_1. Similarly, for R_2 we have

$$\begin{aligned} nR_2 \leq\ & h(S_2^n|\underline{\tilde{H}}^n) - h(Z_1^n) + h(Y_2^n|S_2^n, \underline{\tilde{H}}^n) \\ & - h(S_1^n|\underline{\tilde{H}}^n) + n\epsilon_{2n}. \end{aligned} \qquad (40)$$

Let $R_{\text{sum}} = R_1 + R_2$, denote $\upsilon_i \triangleq \frac{\mathbb{E}\{X_{1,i}X_{3,i}^*\}}{\sqrt{P_{1,i}P_{3,i}}}$, where $|\upsilon_i| \leq 1$, and define $\theta_i \triangleq \arg\{h_{11,i}h_{31,i}^*\upsilon_i\}$. Then, by combining (39) and (40) we obtain the bound in (42), as shown at the bottom of this page, where, in deriding (42), step (a) follows from [8, Lemma 2], which states that given the set of channel coefficients at time i, $\underline{\tilde{h}}_i$, then $h(Y_{k,i}|S_{k,i}, \underline{\tilde{H}}_i = \underline{\tilde{h}}_i)$ is maximized with $Y_{k,i}$ and $S_{k,i}$, $k \in \{1, 2\}$ distributed according to the zero-mean, jointly circularly symmetric complex Normal distribution with the covariance matrix $\text{cov}(Y_{kG,i}, S_{kG,i}|\underline{\tilde{h}}_i) = \text{cov}(Y_{k,i}, S_{k,i}|\underline{\tilde{h}}_i)$, $k \in \{1, 2\}$. Note that in this step and also in the subsequent equality, the maximizing υ_i, $P_{1,i}$, $P_{2,i}$, and $P_{3,i}$ are generally functions of $\underline{\tilde{h}}_i$. Step (b) follows since $0 \leq |\upsilon_i| \leq 1$ and $-1 \leq \cos(\theta_i) \leq 1$ and since the logarithm function is a monotonically increasing function of its argument, (c) follows since both sums of logarithmic functions in (41), as shown at the bottom of this page, are maximized by $P_{1,i} = P_{2,i} = P_{3,i} = 1$, $i \in \{1, 2, \ldots, n\}$. To see this point we consider each of the sums separately:

1) Begin by considering the first logarithmic term in (41): We now show that the expression (43), as shown at the

$$\begin{aligned} & n\big(R_{\text{sum}} - \epsilon_{1n} - \epsilon_{2n}\big) \\ & \leq h(Y_1^n|S_1^n, \underline{\tilde{H}}^n) - h(Z_1^n) + h(Y_2^n|S_2^n, \underline{\tilde{H}}^n) - h(Z_2^n) \\ & \stackrel{(a)}{\leq} \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Big\{ \max_{\substack{0\leq|\upsilon_i|\leq 1 \\ P_{k,i}\leq 1, k\in\{1,2,3\}}} h(Y_{1G,i}|S_{1G,i}, \underline{\tilde{H}}_i = \underline{\tilde{h}}_i)\Big\} - n\cdot h(Z_1) \\ & \qquad + \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Big\{ \max_{\substack{0\leq|\upsilon_i|\leq 1 \\ P_{k,i}\leq 1, k\in\{1,2,3\}}} h(Y_{2G,i}|S_{2G,i}, \underline{\tilde{H}}_i = \underline{\tilde{h}}_i)\Big\} - n\cdot h(Z_2) \\ & = \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Bigg\{ \max_{\substack{0\leq|\upsilon_i|\leq 1 \\ P_{k,i}\leq 1, k\in\{1,2,3\}}} \log\Bigg(1 + P_{2,i}|h_{21,i}|^2 + \frac{P_{1,i}|h_{11,i}|^2 + P_{3,i}|h_{31,i}|^2 + P_{1,i}P_{3,i}|h_{11,i}|^2|h_{32,i}|^2(1-|\upsilon_i|^2)}{1 + P_{3,i}|h_{32,i}|^2} \\ & \qquad + \frac{2|h_{11,i}||h_{31,i}^*|\sqrt{P_{1,i}P_{3,i}}|\upsilon_i|\cos(\theta_i)}{1 + P_{3,i}|h_{32,i}|^2}\Bigg)\Bigg\} \\ & \qquad + \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Bigg\{ \max_{P_{k,i}\leq 1, k\in\{1,2,3\}} \log\Bigg(1 + P_{3,i}|h_{32,i}|^2 + \frac{P_{2,i}|h_{22,i}|^2}{1 + P_{2,i}|h_{21,i}|^2}\Bigg)\Bigg\} \\ & \stackrel{(b)}{\leq} \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Bigg\{ \max_{P_{k,i}\leq 1, k\in\{1,2,3\}} \log\Bigg(1 + P_{2,i}|h_{21,i}|^2 + \frac{P_{1,i}|h_{11,i}|^2 + P_{3,i}|h_{31,i}|^2 + P_{1,i}P_{3,i}|h_{11,i}|^2|h_{32,i}|^2}{1 + P_{3,i}|h_{32,i}|^2} \\ & \qquad + \frac{2|h_{11,i}||h_{31,i}^*|\sqrt{P_{1,i}P_{3,i}}}{1 + P_{3,i}|h_{32,i}|^2}\Bigg)\Bigg\} \\ & \qquad + \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Bigg\{ \max_{P_{k,i}\leq 1, k\in\{1,2,3\}} \log\Bigg(1 + P_{3,i}|h_{32,i}|^2 + \frac{P_{2,i}|h_{22,i}|^2}{1 + P_{2,i}|h_{21,i}|^2}\Bigg)\Bigg\} \qquad (41) \\ & \stackrel{(c)}{\leq} \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Bigg\{ \log\Bigg(1 + |h_{21,i}|^2 + \frac{|h_{11,i}|^2 + |h_{31,i}|^2 + |h_{11,i}|^2|h_{32,i}|^2 + 2|h_{11,i}||h_{31,i}^*|}{1 + |h_{32,i}|^2}\Bigg)\Bigg\} \\ & \qquad + \sum_{i=1}^n \mathbb{E}_{\underline{\tilde{H}}_i}\Bigg\{ \log\Bigg(1 + |h_{32,i}|^2 + \frac{|h_{22,i}|^2}{1 + |h_{21,i}|^2}\Bigg)\Bigg\} \\ & \stackrel{(d)}{=} n\log\Bigg(1 + \text{SNR}_{21} + \frac{\text{SNR}_{11} + \text{SNR}_{31} + \text{SNR}_{11}\text{SNR}_{32} + 2\sqrt{\text{SNR}_{11}\text{SNR}_{31}}}{1 + \text{SNR}_{32}}\Bigg) + n\log\Bigg(1 + \text{SNR}_{32} + \frac{\text{SNR}_{22}}{1 + \text{SNR}_{21}}\Bigg) \end{aligned} \qquad (42)$$

$$\frac{P_{1,i}|h_{11,i}|^2 + P_{3,i}|h_{31,i}|^2 + P_{1,i}P_{3,i}|h_{11,i}|^2|h_{32,i}|^2 + 2|h_{11,i}||h^*_{31,i}|\sqrt{P_{1,i}P_{3,i}}}{1 + P_{3,i}|h_{32,i}|^2} \tag{43}$$

$$\frac{\partial}{\partial P_{3,i}}\left\{\frac{P_{1,i}|h_{11,i}|^2 + P_{3,i}|h_{31,i}|^2 + P_{1,i}P_{3,i}|h_{11,i}|^2|h_{32,i}|^2 + 2|h_{11,i}||h^*_{31,i}|\sqrt{P_{1,i}P_{3,i}}}{1 + P_{3,i}|h_{32,i}|^2}\right\} = \frac{|h_{31,i}|^2 + |h_{11,i}||h^*_{31,i}|\frac{\sqrt{P_{1,i}}}{\sqrt{P_{3,i}}} - |h_{11,i}||h^*_{31,i}||h_{32,i}|^2\sqrt{P_{1,i}}\sqrt{P_{3,i}}}{\left(1 + P_{3,i}|h_{32,i}|^2\right)^2} \tag{44}$$

$$\begin{aligned} R_{\text{sum}} &\leq \log\left(1 + \mathsf{SNR}_{21} + \frac{\mathsf{SNR}_{11} + \mathsf{SNR}_{31} + \mathsf{SNR}_{11}\mathsf{SNR}_{32} + 2\sqrt{\mathsf{SNR}_{11}\mathsf{SNR}_{31}}}{1 + \mathsf{SNR}_{32}}\right) \\ &\quad + \log\left(1 + \mathsf{SNR}_{32} + \frac{\mathsf{SNR}_{22}}{1 + \mathsf{SNR}_{21}}\right) \\ &\overset{(a)}{\doteq} \log\left(\mathsf{SNR}^{\max\{\alpha,1,\beta-\lambda\}}\right) + \log\left(\mathsf{SNR}^{\max\{\lambda,1-\alpha\}}\right) \end{aligned} \tag{45}$$

top of the page which appears in the first summation in (41), increases monotonically with respect to both $P_{1,i}$ and $P_{3,i}$. To that aim we note that from inspecting the expression (43) it is evident that it increases monotonically with respect to $P_{1,i}$, for any $P_{3,i} \geq 0$. Next, for any fixed $0 \leq P_{1,i} \leq 1$, we differentiate (43) with respect to $P_{3,i}$ and obtain (44), as shown at the top of this page. From (44) we note that as $0 \leq P_{1,i}, P_{3,i} \leq 1$, then the derivative is positive if $|h_{31,i}|^2 > |h_{11,i}||h^*_{31,i}||h_{32,i}|^2$, or equivalently, if $\mathsf{SNR}_{31} > \mathsf{SNR}_{32}\sqrt{\mathsf{SNR}_{11}\mathsf{SNR}_{31}}$, which is satisfied if $\beta > 2\lambda + 1$. We conclude that (43) increases monotonically with respect to $0 \leq P_{1,i}, P_{3,i} \leq 1$. This conclusion, combined with the facts that the expression in the logarithm in the first summation in (41) is monotone increasing in $P_{2,i}$, and that the logarithm function itself is monotone increasing, leads to the conclusion that if $\beta > 2\lambda+1$, then the first logarithmic expression in (41) monotonically increases with respect to $P_{1,i}$, $P_{2,i}$ and $P_{3,i}$, hence it is maximized by setting $P_{1,i} = P_{2,i} = P_{3,i} = 1$.

2) Now consider the second term in (41): The function $\frac{a_1 x}{1+b_1 x}$ monotonically increases with respect to x as long as $a_1, b_1 > 0$ and thus, letting $a_1 = |h_{22,i}|^2$ and $b_1 = |h_{21,i}|^2$, we conclude that $\frac{P_{2,i}|h_{22,i}|^2}{1+P_{2,i}|h_{21,i}|^2}$ increases with respect to $P_{2,i}$. It also immediately follows that

$$\log\left(1 + P_{3,i}|h_{32,i}|^2 + \frac{P_{2,i}|h_{22,i}|^2}{1 + P_{2,i}|h_{21,i}|^2}\right)\Bigg|_{P_{2,i}=1} = \log\left(1 + P_{3,i}|h_{32,i}|^2 + \frac{|h_{22,i}|^2}{1 + |h_{21,i}|^2}\right)$$

is maximized by $P_{3,i} = 1$.

We conclude that if $\beta > 2\lambda + 1$ then (41) is maximized when all nodes transmit at their maximum available power: $P_{1,i} = P_{2,i} = P_{3,i} = 1, i \in \{1, 2, \ldots, n\}$. Finally, step (d) follows since in the ergodic phase fading model, the magnitudes of the channel coefficients are constants and do not depend on the time index, and therefore the expectation can be omitted. Observe that as (R_1, R_2) is achievable, then for $k \in \{1, 2\}$, $P^{(n)}_{e,k} \to 0$ as $n \to \infty$, and hence, $\epsilon_{kn} \to 0$ as $n \to \infty$. We therefore conclude that the sum rate is asymptotically bounded by (45), as shown at the top of this page. In the derivation leading to (45), step (a) follows since $\max\{1, \beta\} \geq \frac{1+\beta}{2}$. We note that if $\beta > 2\lambda + 1$, then $\beta - \lambda > 1 + \lambda \geq 1$, hence, $\max\{\alpha, 1, \beta - \lambda\} = \max\{\alpha, \beta - \lambda\}$, and

$$\begin{aligned} &\max\{\alpha, \beta - \lambda\} + \max\{\lambda, 1 - \alpha\} \\ &\quad = \max\{\alpha + \lambda, \beta, 1, 1 + \beta - \alpha - \lambda\}. \end{aligned}$$

Therefore, if $\beta > 2\lambda + 1$ then the genie-aided GDoF upper bound is given by

$$\mathrm{GDoF}^+_1 = \max\{\alpha + \lambda, \beta, 1 + \beta - \alpha - \lambda\}. \tag{46}$$

■

2) An Upper Bound Based on the Cut-Set Theorem: We derive three rate bounds following along the lines of the proof of the cut-set theorem [31, Th. 15.10.1]. First, we derive an upper bounds on R_1 by considering the cut $\mathcal{S} = \{\mathrm{Tx}_1, \mathrm{Relay}, \mathrm{Rx}_2\}$, $\mathcal{S}^c = \{\mathrm{Tx}_2, \mathrm{Rx}_1\}$, i.e., allowing full cooperation between Tx_1 and the Relay. For this cut we obtain

$$\begin{aligned} nR_1 &\overset{(a)}{\leq} I(M_1; Y^n_1, \underline{\tilde{H}}^n | M_2) + n\epsilon_{1n} \\ &= \sum_{i=1}^{n}\Big[h(Y_{1,i}, \underline{\tilde{H}}_i | Y^{i-1}_1, \underline{\tilde{H}}^{i-1}, M_2) \\ &\qquad - h(Y_{1,i}, \underline{\tilde{H}}_i | Y^{i-1}_1, \underline{\tilde{H}}^{i-1}, M_1, M_2)\Big] + n\epsilon_{1n} \\ &\overset{(b)}{=} \sum_{i=1}^{n}\Big[h(\underline{\tilde{H}}_i) + h(Y_{1,i} | Y^{i-1}_1, \underline{\tilde{H}}^i, M_2) \\ &\qquad - h(\underline{\tilde{H}}_i) - h(Y_{1,i} | Y^{i-1}_1, \underline{\tilde{H}}^i, M_1, M_2)\Big] + n\epsilon_{1n} \\ &\overset{(c)}{\leq} \sum_{i=1}^{n}\Big[h(Y_{1,i} | Y^{i-1}_1, \underline{\tilde{H}}^i, M_2, X_{2,i}) \\ &\qquad - h(Y_{1,i} | Y^{i-1}_1, \underline{\tilde{H}}^i, M_1, M_2, X_{1,i}, X_{2,i}, X_{3,i})\Big] + n\epsilon_{1n} \end{aligned}$$

$$\overset{(d)}{\leq} \sum_{i=1}^{n}\Big[h(Y_{1,i}|X_{2,i},\underline{\tilde{H}}_i) - h(Y_{1,i}|X_{1,i},X_{2,i},X_{3,i},\underline{\tilde{H}}_i)\Big] + n\epsilon_{1n}$$
$$\overset{(e)}{\leq} \sum_{i=1}^{n}\mathbb{E}_{\underline{\tilde{H}}_i}\Big\{\max_{\substack{0\leq|v_i|\leq 1\\ P_{k,i}\leq 1,\\ k\in\{1,2,3\}}} I(X_{1G,i},X_{3G,i};Y_{1G,i}|X_{2G,i},\underline{\tilde{H}}_i=\underline{\tilde{h}}_i)\Big\} + n\epsilon_{1n}$$
$$\overset{(f)}{\leq} \sum_{i=1}^{n}\mathbb{E}_{\underline{\tilde{H}}_i}\Big\{\log\Big(1+|h_{11,i}|^2+|h_{31,i}|^2\Big)\Big\} + n\epsilon_{1n}$$
$$= n\cdot\log\Big(1+\mathsf{SNR}_{11}+\mathsf{SNR}_{31}\Big)+n\epsilon_{1n}, \qquad (47)$$

where (a) follows from Fano's inequality [31, Th. 2.10.1] and since the messages from Tx_1 and Tx_2 are drawn independently, (b) follows since channel coefficients are i.i.d. in time and are independent of the channel inputs, of the noise, and of the messages sent by the sources, (c) follows since $X_{2,i}$ is a deterministic function of M_2 and since adding conditioning decreases the differential entropy, (d) follows since adding conditioning can only decrease the differential entropy, and since the channel outputs at time i depend only on the channel inputs and the channel coefficients at time i. To prove step (e) first note that

$$h(Y_{1,i}|X_{2,i},\underline{\tilde{H}}_i) - h(Y_{1,i}|X_{1,i},X_{2,i},X_{3,i},\underline{\tilde{H}}_i) = h(Y_{1,i}|X_{2,i},\underline{\tilde{H}}_i) - h(Z_{1,i}).$$

Step (e) then follows from [8, Lemma 2] which states that the conditional entropy is maximized by jointly circularly symmetric complex normal channel inputs with covariance matrix $\text{cov}(X_{2,i},Y_{1,i})$. Note that we can write

$$(X_{2,i},Y_{1,i}) = \begin{bmatrix} 0 & 1 & 0 & 0\\ H_{11,i} & H_{21,i} & H_{31,i} & 1\end{bmatrix}\begin{bmatrix}X_{1,i}\\ X_{2,i}\\ X_{3,i}\\ Z_{1,i}\end{bmatrix}.$$

As the pair $(X_{2,i},Y_{1,i})$ is a linear transformation of a random vector, and as in addition, $(X_{2,i},Y_{1,i})$ is distributed according to a zero mean, jointly complex Gaussian distribution, we conclude that the joint distribution of $(X_{2G,i},Y_{1G,i})$ with the covariance matrix $\text{cov}(X_{2,i},Y_{1,i})$ is obtained by letting $(X_{1,i},X_{2,i},X_{3,i},Z_{1,i})$ be a jointly complex Gaussian random vector, which, in turn is obtained when $(X_{1,i},X_{2,i},X_{3,i})$ is a jointly complex Normal vector with covariance matrix $\text{cov}(X_{1,i},X_{2,i},X_{3,i})$. Finally, step (f) follows from [8, eq. (A.10)].

Next, by using the cut $\mathcal{S}=\{\text{Tx}_1,\text{Rx}_2\}$, $\mathcal{S}^c=\{\text{Relay},\text{Tx}_2,\text{Rx}_1\}$, i.e., by allowing full cooperation between Rx_1 and the relay, we obtain an additional upper bound on R_1. This bound is expressed as:

$$nR_1 \leq \sum_{i=1}^{n}\mathbb{E}_{\underline{\tilde{H}}_i}\Big\{\max_{\substack{0\leq|v_i|\leq 1\\ P_{k,i}\leq 1,k\in\{1,2,3\}}} I(X_{1G,i};Y_{1G,i},Y_{3G,i}|X_{2G,i},X_{3G,i},\underline{\tilde{H}}_i=\underline{\tilde{h}}_i)\Big\}+n\epsilon_{1n}$$
$$\overset{(a)}{\leq} \sum_{i=1}^{n}\mathbb{E}_{\underline{\tilde{H}}_i}\Big\{\log\Big(1+|h_{11,i}|^2+|h_{13,i}|^2\Big)\Big\}+n\epsilon_{1n}$$
$$= n\cdot\log\Big(1+\mathsf{SNR}_{11}+\mathsf{SNR}_{13}\Big)+n\epsilon_{1n}, \qquad (48)$$

where (a) follows from [8, eq. (A.5)].

Lastly, we use the cut $\mathcal{S}=\{\text{Tx}_2,\text{Rx}_1\}$, $\mathcal{S}^c=\{\text{Relay},\text{Tx}_1,\text{Rx}_2\}$ to obtain an upper bound on R_2:

$$nR_2 \leq \sum_{i=1}^{n}\mathbb{E}_{\underline{\tilde{H}}_i}\Big\{\max_{\substack{0\leq|v_i|\leq 1\\ P_{k,i}\leq 1,k\in\{1,2,3\}}} I(X_{2G,i};Y_{2G,i}|X_{3G,i},\underline{\tilde{H}}_i=\underline{\tilde{h}}_i)\Big\}+n\epsilon_{2n}$$
$$= n\cdot\log\Big(1+\mathsf{SNR}_{22}\Big)+n\epsilon_{2n}. \qquad (49)$$

Since for $n\to\infty$ we have $\epsilon_{kn}\to 0, k\in\{1,2\}$, then by combining (47)-(49) we obtain

$$R_{\text{sum}} \leq \min\Big\{\log\Big(1+\mathsf{SNR}_{11}+\mathsf{SNR}_{31}\Big), \log\Big(1+\mathsf{SNR}_{11}+\mathsf{SNR}_{13}\Big)\Big\} + \log\Big(1+\mathsf{SNR}_{22}\Big)$$
$$\doteq \min\Big\{\log\Big(\mathsf{SNR}^{\max\{1,\beta\}}\Big), \log\Big(\mathsf{SNR}^{\max\{1,\gamma\}}\Big)\Big\} + \log\Big(\mathsf{SNR}\Big).$$

Thus, the cut-set based GDoF upper bound is given by:

$$\text{GDoF}_2^{+} = 1+\min\big\{\max\{1,\beta\},\max\{1,\gamma\}\big\} = \max\big\{2, 1+\min\{\beta,\gamma\}\big\}. \qquad (50)$$

Note that (50) holds for any relationship between α,β,γ and λ. We conclude that an upper bound on the GDoF of the Z-ICR is given by the minimum of (46) and (50), which coincides with (38). ■

B. A Lower Bound on the Achievable GDoF

A lower bound on the achievable GDoF of the ergodic phase fading Z-ICR is stated in the following proposition:

Proposition 3: Consider the ergodic phase fading Z-ICR defined in Section II. The GDoF of this channel is lower bounded by

$$\text{GDoF}^{-} = \min\Big\{\gamma, \max\big\{(1-\alpha)^+,(\beta-\alpha)^+\big\}\Big\}+(1-\lambda)^+. \qquad (51)$$

Proof: We use a communications scheme similar to the communications scheme of Section III-D: The transmitters use mutually independent codebooks generated according to the i.i.d. (in time) complex Normal distributions: $X_{k,i}\sim\mathcal{CN}(0,P_k), k\in\{1,2,3\}, i\in\{1,2,\ldots,n\}$, $0<P_k\leq 1$, $k\in\{1,2,3\}$. Encoding is based on the DF scheme at the relay, and for decoding we use a backward decoding scheme at Rx_1, and a PtP decoding rule at Rx_2, where both receivers treat the additive interference as noise. Repeating the analysis in the proof of Prop. 2 it follows that this coding scheme results in the following achievable rate region for the Z-ICR:

$$R_1 \leq \min\big\{I(X_1,X_3;Y_1|\tilde{H}_1), I(X_1;Y_3|X_3,\tilde{H}_3)\big\} \qquad (52a)$$
$$R_2 \leq I(X_2;Y_2|\tilde{H}_2). \qquad (52b)$$

Explicitly evaluating the mutual information expressions in (52) with the Gaussian p.d.f. for (X_1,X_2,X_3) specified above, we arrive at

$$I(X_1, X_3; Y_1|\tilde{H}_1) = \mathbb{E}_{\tilde{H}_1}\left\{\log\left(1 + \frac{P_1|h_{11}|^2 + P_3|h_{31}|^2}{1 + P_2|h_{21}|^2}\right)\right\}$$
$$\doteq \log\left(\mathsf{SNR}^{(1-\alpha)^+} + \mathsf{SNR}^{(\beta-\alpha)^+}\right), \quad (53)$$
$$I(X_1; Y_3|X_3, \tilde{H}_3) = \mathbb{E}_{\tilde{H}_3}\left\{\log(1 + P_1|h_{13}|^2)\right\}$$
$$= \log\left(1 + P_1 \cdot \mathsf{SNR}^{\gamma}\right)$$
$$\doteq \log\left(\mathsf{SNR}^{\gamma}\right) \quad (54)$$
$$I(X_2; Y_2|\tilde{H}_2) = \mathbb{E}_{\tilde{H}_2}\left\{\log\left(1 + \frac{P_2|h_{22}|^2}{1 + P_3|h_{32}|^2}\right)\right\}$$
$$= \log\left(1 + \frac{P_2 \cdot \mathsf{SNR}}{1 + P_3 \cdot \mathsf{SNR}^{\lambda}}\right)$$
$$\doteq \log\left(\mathsf{SNR}^{(1-\lambda)^+}\right). \quad (55)$$

Combining (53)-(55), we obtain the achievable GDoF stated in (51). ■

C. The Optimality of Treating Interference as Noise

In this section, we derive conditions on the SNR exponents of the channel coefficients under which the lower bound in (51) coincides with the upper bound in (38), thereby characterizing the maximal GDoF for the ergodic phase fading Z-ICR in the WI regime. Consider (51) and note that if $\alpha \leq \frac{1}{2}$ and $1 + 2\alpha < \beta \leq \gamma + \alpha$, then it follows that $\gamma \geq 1$, $\beta \geq 1$, $\max\left\{(1-\alpha)^+, (\beta-\alpha)^+\right\} = \beta - \alpha$, and $\min\{\gamma, \beta - \alpha\} = \beta - \alpha$. If it additionally holds that $\lambda = \alpha$ then (51) results in $\text{GDoF}^- = 1 + \beta - 2\alpha$.

Next, consider (38) and note that if $\lambda = \alpha$, $1 + 2\alpha < \beta$, and $\alpha \leq \frac{1}{2}$, then

$$\max\{\alpha + \lambda, \beta, 1 + \beta - \alpha - \lambda\} = \max\{2\alpha, \beta, 1 + \beta - 2\alpha\}$$
$$= \max\{\beta, 1 + \beta - 2\alpha\}$$
$$= 1 + \beta - 2\alpha,$$

and thus, (38) specializes to

$$\text{GDoF}^+ = \min\left\{\max\left\{2, 1 + \min\{\beta, \gamma\}\right\}, 1 + \beta - 2\alpha\right\}.$$

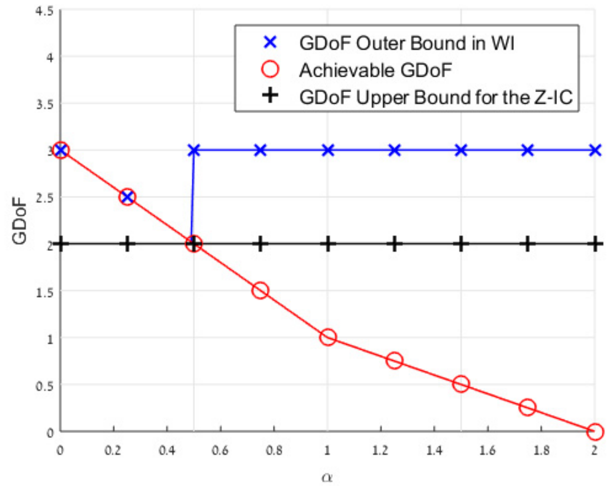
Next, from (35b) we conclude that $\gamma \geq 1$ and $\beta > 1+2\alpha \geq 1$, and hence, $\max\left\{2, 1+\min\{\beta, \gamma\}\right\} = 1+\min\{\beta, \gamma\}$. Lastly, the condition $\beta \leq \gamma + \alpha$ implies that $\beta \leq \min\{\beta, \gamma\} + 2\alpha$, i.e., $1 + \beta - 2\alpha \leq 1 + \min\{\beta, \gamma\}$, and thus, it follows that $\text{GDoF}^+ = 1 + \beta - 2\alpha$. We conclude that if (35) is satisfied, then (38) coincides with (51) and both are equal to $1 + \beta - 2\alpha$, thereby characterizing the maximal GDoF for the ergodic phase fading Z-ICR subject to (35). The maximizing input distribution follows directly from the input distribution used in the proof of the lower bound in Prop. 3, namely $X_k \sim \mathcal{CN}(0, P_k)$, $0 < P_k \leq 1$, $k \in \{1, 2, 3\}$. ■

D. Discussion

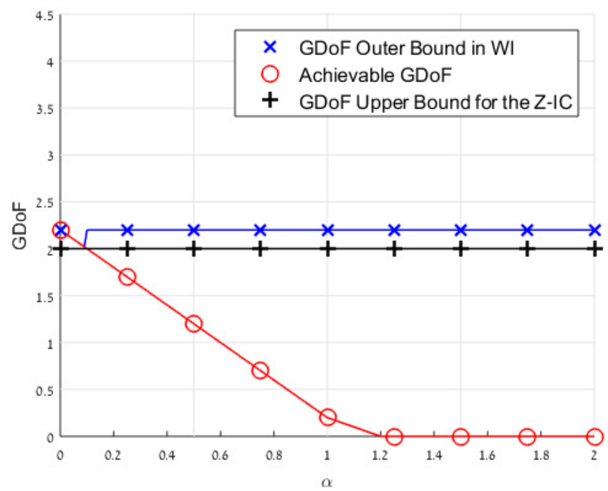
Comment 10: Consider the ergodic phase fading Z-IC: An upper bound on the achievable sum-rate for this channel is obtained by applying cut-set theorem [31, Th. 15.10.1]:

$$R_{\text{sum}}^{\text{PF Z-IC}}(\mathsf{SNR}) \leq \max_{\left\{\substack{f(x_1)f(x_2): \\ \mathbb{E}\left\{|X_k|^2\right\} \leq 1,\ k \in \{1,2\}}\right\}} \left\{I(X_1; Y_1|X_2, \underline{\tilde{H}}) + I(X_2; Y_2|\underline{\tilde{H}})\right\}$$
$$\doteq \log(\mathsf{SNR}) + \log(\mathsf{SNR}). \quad (56)$$

It follows that the GDoF for this channel is upper bounded by 2. Comparing the GDoF upper bound of the phase fading Z-IC with the lower bound on the GDoF of the phase fading Z-ICR stated in (51), we note that if $1 \leq \beta \leq \gamma + \alpha$, then for the ergodic phase fading Z-ICR we have $\text{GDoF}^-_{\text{Z-ICR}} = (\beta - \alpha)^+ + (1 - \lambda)^+$. Hence, when $2 < (\beta - \alpha)^+ + (1 - \lambda)^+$ the relay node *strictly increases* the GDoF of the ergodic phase fading Z-IC *even in scenarios in which the relay receives transmissions only from one of the transmitters*, and the interference is treated as noise at both destinations. In Fig. 4, the GDoF upper bound and the GDoF lower bound



(a)



(b)

Fig. 4. The upper bound on the GDoF of (38), and the achievable GDoF of (51) for the phase fading Z-ICR, together with the GDoF upper bound for the Z-IC given in (56). (a) $\beta = \gamma = 2$. (b) $\beta = \gamma = 1.2$.

for the Z-ICR, as well as the upper bound on the GDoF of the Z-IC, are plotted vs. α for two sets of (β, γ): $\beta = \gamma = 2$ and $\beta = \gamma = 1.2$, subject to ergodic phase fading. Observe that the GDoF for the Z-ICR is *strictly* greater than that of the Z-IC for $\beta = \gamma = 2$, when $\alpha < 0.5$ and for $\beta = \gamma = 1.2$, when $\alpha < 0.1$. Fig. 4 also clearly demonstrates the GDoF optimality of treating interference as noise in the WI regime.

Comment 11: Note that from Theorems 1 and 3 we conclude that mutually independent channel inputs achieve both the sum-rate capacity and the maximal GDoF of the ergodic phase fading Z-ICR in the weak interference regime. Hence, using the communications scheme described in Section III-D, there is no need for coordinating the codebooks of Tx_1 and of the relay to achieve optimality in both perspectives (capacity and GDoF). This observation suggests that when adding a relay to the interference network considered in this manuscript, the transmission scheme at the sources should remain unchanged, and that only the receivers should be modified to take advantage of the relay transmissions when decoding the messages from the sources, in order to improve performance. This conclusion substantially simplifies adding relay nodes to existing wireless communications networks, and provides a strong support for user cooperation for interference management in the weak interference regime.

Comment 12: From the derivation of the achievable GDoF in Section IV, it directly follows that the maximal GDoF of the ergodic phase fading Z-ICR can be achieved with channel inputs generated according to mutually independent i.i.d Gaussians *with any arbitrary non-zero power*, and it is not necessary to use the maximal power $P_k = 1$, $k = 1, 2, 3$, for generating the channel inputs. Note, however, that the technical derivation of the GDoF upper bound does require $P_k = 1$, $k = 1, 2, 3$, because we first upper bound the rate at *any* SNR and then take $\mathsf{SNR} \to \infty$. Yet, the achievability scheme can obtain the maximal GDoF when the nodes transmit with any finite positive powers, as long as the conditions of Thm. 3 are satisfied.

V. Conclusions

In this paper, we studied the two major performance measures of the ergodic phase fading Z-ICR: The sum-rate capacity and the GDoF. We focused on scenarios in which the interference is weak and the relay receives transmissions only from Tx_1. We first characterized the sum-rate capacity of the ergodic phase fading Z-ICR in the WI regime. This is the first capacity result for the Z-ICR in the WI regime in which the relay power is finite. Next, we explained why GDoF analysis is relevant for this channel model although the fading process is ergodic, and then characterized the maximal GDoF for this channel in the weak interference regime. To the best of our knowledge, this is the first time that GDoF analysis is carried out for a fading scenario. For both performance measures, optimal performance was achieved by treating the interfering signal as additive noise at the destination receivers, in combination with using the DF strategy at the relay.

Our results show that adding a relay to the Z-IC enhances both its sum-capacity and GDoF compared to communications without a relay. Combined with our previous results on fading ICRs in the SI regime [8], [9], [14], we conclude that there is a very strong motivation for employing relay nodes for interference management in both the WI regime as well as in the strong interference regime. Additionally, the fact that the optimal channel inputs are mutually independent both in the strong interference regime and in the weak interference regime, further motivates incorporating relay nodes into existing wireless networks. The results in this paper constitute a starting point for studying the combination of cooperation and interference in the WI regime.

Appendix A
Proof of Lemma 4

The proof is based on the proof of [6, Lemma 8]. Recall that a circularly symmetric, complex Normal random vector can be represented as a random vector with double the length, whose components are real, jointly Gaussian RVs. It follows that, as [6, Lemma 7] is stated for real Gaussian random vectors,[3] it holds also for circularly symmetric complex Normal random vectors. Letting $\mathbf{X} \triangleq (X_1, X_2)^T$ then from [6, Lemma 7] we obtain the following equivalence for $(\mathbf{X}, Y_1, Y_2)$:

$$I(\mathbf{X}; Y_1|Y_2) = 0 \Leftrightarrow \mathbb{E}\{Y_1|\mathbf{X}, Y_2\} = \mathbb{E}\{Y_1|Y_2\}. \qquad \text{(A.1)}$$

Next, let $Y_{kR} = \mathfrak{Re}\{Y_k\}$, $Y_{kI} = \mathfrak{Im}\{Y_k\}$, $Z_{kR} = \mathfrak{Re}\{Z_k\}$, $Z_{kI} = \mathfrak{Im}\{Z_k\}$, $D = c_1 \cdot X_1 + c_2 \cdot X_2$, $D_R = \mathfrak{Re}\{D\}$, and $D_I = \mathfrak{Im}\{D\}$. Additionally, denote $\bar{Y}_k = (Y_{kR}, Y_{kI})^T$, $\bar{Z}_k = (Z_{kR}, Z_{kI})^T$, $k = 1, 2$, $\bar{D} = (D_R, D_I)^T$, and $\mathbb{D} = \begin{bmatrix} D_R^2 & D_R D_I \\ D_I D_R & D_I^2 \end{bmatrix}$. Finally, define $\mathbf{X}_R = \mathfrak{Re}\{\mathbf{X}\}$, $\mathbf{X}_I = \mathfrak{Im}\{\mathbf{X}\}$, $\bar{\mathbf{X}} = \left(\mathbf{X}_R^T, \mathbf{X}_I^T\right)^T$, and $\mathbb{Z} = \frac{1}{\text{var}(Z_2)} \begin{bmatrix} 2\mathbb{E}\{Z_{1R}Z_{2R}\} & 2\mathbb{E}\{Z_{1R}Z_{2I}\} \\ -2\mathbb{E}\{Z_{1R}Z_{2I}\} & 2\mathbb{E}\{Z_{1R}Z_{2R}\} \end{bmatrix}$. With these definitions, for the MMSE estimate of Y_1 based on $(\mathbf{X}, Y_2)$ we write

$$\begin{aligned}
\mathbb{E}\{Y_1|\mathbf{X}, Y_2\} &= \mathbb{E}\{\bar{Y}_1|\bar{\mathbf{X}}, \bar{Y}_2\} \\
&= \mathbb{E}\{\bar{Y}_1|\bar{\mathbf{X}}, \bar{Z}_2\} \\
&\overset{(a)}{=} \bar{D} + \mathbb{E}\{\bar{Z}_1|\bar{Z}_2\} \\
&= \bar{D} + \mathbb{E}\{\bar{Z}_1 \cdot \bar{Z}_2^T\} \left(\mathbb{E}\{\bar{Z}_2 \cdot \bar{Z}_2^T\}\right)^{-1} \bar{Z}_2 \\
&= \bar{D} + \mathbb{E}\left\{\begin{bmatrix} Z_{1R}Z_{2R} & Z_{1R}Z_{2I} \\ Z_{1I}Z_{2R} & Z_{1I}Z_{2I} \end{bmatrix}\right\} \\
&\qquad \cdot \left(\mathbb{E}\left\{\begin{bmatrix} Z_{2R}^2 & Z_{2R}Z_{2I} \\ Z_{2I}Z_{2R} & Z_{2I}^2 \end{bmatrix}\right\}\right)^{-1} \bar{Z}_2 \\
&\overset{(b)}{=} \bar{D} + \mathbb{E}\left\{\begin{bmatrix} Z_{1R}Z_{2R} & Z_{1R}Z_{2I} \\ -Z_{1R}Z_{2I} & Z_{1R}Z_{2R} \end{bmatrix}\right\} \\
&\qquad \cdot \left(\begin{bmatrix} \frac{1}{2}\text{var}(Z_2) & 0 \\ 0 & \frac{1}{2}\text{var}(Z_2) \end{bmatrix}\right)^{-1} \bar{Z}_2 \\
&= \bar{D} + \frac{1}{\text{var}(Z_2)} \begin{bmatrix} 2\mathbb{E}\{Z_{1R}Z_{2R}\} & 2\mathbb{E}\{Z_{1R}Z_{2I}\} \\ -2\mathbb{E}\{Z_{1R}Z_{2I}\} & 2\mathbb{E}\{Z_{1R}Z_{2R}\} \end{bmatrix} \bar{Z}_2 \\
&= \bar{D} + \mathbb{Z} \cdot \bar{Z}_2 \\
&= \bar{D}\,(\mathbb{I}_2 - \mathbb{Z}) + \mathbb{Z} \cdot \bar{Y}_2. \qquad \text{(A.2)}
\end{aligned}$$

[3]Reference [6, Lemma 7] states that for the real Gaussian random vectors $\mathbf{X}$, Y, and S, the following three statements are equivalent: (1) $I(\mathbf{X}; S|Y) = 0$, (2) $\mathbf{X} - Y - S$ form a Markov chain, and (3) $\hat{S}(\mathbf{X}, Y)$, the MMSE estimate of S given $(\mathbf{X}, Y)$, is equal to $\hat{S}(Y)$ the MMSE estimate of S given Y.

Here, (a) follows from [30, Th. 23.7.4] by using the real vector representation for the complex RVs, and (b) follows from the conditions on the cross-correlations in the statement of the lemma.

Computing $E\{Y_1|Y_2\}$ explicitly we obtain

$$\begin{aligned}
&\mathbb{E}\{Y_1|Y_2\}\\
&\quad= \mathbb{E}\{\bar{Y}_1|\bar{Y}_2\}\\
&\quad= \mathbb{E}\left\{\left(\bar{D}+\bar{Z}_1\right)\cdot\left(\bar{D}+\bar{Z}_2\right)^T\right\}\\
&\qquad\qquad\cdot\left(\mathbb{E}\left\{\left(\bar{D}+\bar{Z}_2\right)\cdot\left(\bar{D}+\bar{Z}_2\right)^T\right\}\right)^{-1}\bar{Y}_2\\
&\quad= \mathbb{E}\left\{\bar{D}\cdot\bar{D}^T+\bar{Z}_1\cdot\bar{Z}_2^T\right\}\\
&\qquad\qquad\cdot\left(\mathbb{E}\left\{\bar{D}\cdot\bar{D}^T+\bar{Z}_2\cdot\bar{Z}_2^T\right\}\right)^{-1}\bar{Y}_2\\
&\quad=\left(\frac{1}{\text{var}(Z_2)}2\mathbb{E}\{\mathbb{D}\}+\mathbb{Z}\right)\cdot\left(\frac{1}{\text{var}(Z_2)}2\mathbb{E}\{\mathbb{D}\}+\mathbb{I}_2\right)^{-1}\bar{Y}_2.
\end{aligned}\tag{A.3}$$

Comparing (A.2) and (A.3) we conclude that $\mathbb{E}\{Y_1|\mathbf{X},Y_2\} = \mathbb{E}\{Y_1|Y_2\}$, if and only if $\mathbb{Z}=\mathbb{I}_2$, namely

$$\mathbb{E}\{Z_{1R}Z_{2R}\} = \mathbb{E}\{Z_{1I}Z_{2I}\} = \frac{1}{2}\text{var}(Z_2)$$
$$\mathbb{E}\{Z_{1R}Z_{2I}\} = \mathbb{E}\{Z_{2R}Z_{1I}\} = 0.$$

Note that $\mathbb{E}\{|Z_2|^2\} = \text{var}(Z_2)$, and that $\mathbb{E}\{Z_1\cdot Z_2^*\} = \mathbb{E}\big\{(Z_{1R}Z_{2R}+Z_{1I}Z_{2I})+j\cdot(-Z_{1R}Z_{2I}+Z_{2R}Z_{1I})\big\}$. Thus, subject to the conditions of the lemma $\mathbb{Z}=\mathbb{I}_2$ is equivalent to $\mathbb{E}\{|Z_2|^2\}=\mathbb{E}\{Z_1\cdot Z_2^*\}$. ∎

Appendix B
Proof of Lemma 5

We follow the same approach as the proof of [24, Lemma 13]. Let $\mathbf{X}_1 = X_1^n$, $\mathbf{X}_2 = X_2^m$, $\mathbf{Z}_1 = Z_1^n$, $\mathbf{Z}_2 = Z_2^m$, $\mathbf{X} = (\mathbf{X}_1^T,\mathbf{X}_2^T)^T$, and $\mathbf{Z} = (\mathbf{Z}_1^T,\mathbf{Z}_2^T)^T$. By the chain rule of mutual information we have

$$\begin{aligned}I(\mathbf{X};\mathbf{X}+\mathbf{Z}) = I(\mathbf{X}_1;\mathbf{X}_1+\mathbf{Z}_1)+I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1)\\ +I(\mathbf{X}_2;\mathbf{X}+\mathbf{Z}|\mathbf{X}_1).\end{aligned}$$

In the following, we will show that

$$\lim_{\gamma_2\to\infty} I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1) = \lim_{\gamma_2\to\infty} I(\mathbf{X}_2;\mathbf{X}+\mathbf{Z}|\mathbf{X}_1) = 0,\tag{B.1}$$

which will complete the proof. Starting with $I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1)$, we note that

$$\begin{aligned}
&I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1)\\
&\quad= h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1)-h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1,\mathbf{Z}_1)\\
&\quad\overset{(a)}{\le} h(\mathbf{X}_2+\mathbf{Z}_2)-h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1,\mathbf{Z}_1)\\
&\quad\overset{(b)}{=} h(\mathbf{X}_2+\mathbf{Z}_2)-h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1)\\
&\quad= I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2)\\
&\quad\overset{(c)}{\le} I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2),
\end{aligned}$$

where (a) follows since conditioning does not increase entropy [31, Th. 2.6.5], (b) follows since $\mathbf{Z}_1$ is independent of $(\mathbf{Z}_2,\mathbf{X}_1,\mathbf{X}_2)$, and (c) follows since $I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2)\le I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2)$, and

$$\begin{aligned}
&I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2)\\
&\quad= I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2)+h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_2)-h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1,\mathbf{X}_2)\\
&\quad= I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2).
\end{aligned}\tag{B.2}$$

Next, note that

$$\begin{aligned}
&I(\mathbf{X}_2;\mathbf{X}+\mathbf{Z}|\mathbf{X}_1)\\
&\quad= I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1)+h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{X}_1,\mathbf{X}_2+\mathbf{Z}_2)\\
&\qquad -h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_2+\mathbf{Z}_2)\\
&\quad= h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1)-h(\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1,\mathbf{X}_2)\\
&\quad\le h(\mathbf{X}_2+\mathbf{Z}_2)-h(\mathbf{Z}_2)\\
&\quad= I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2).
\end{aligned}\tag{B.3}$$

In summary, from (B.2) and (B.3) we obtain that

$$I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1)\le I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2)$$
$$I(\mathbf{X}_2;\mathbf{X}+\mathbf{Z}|\mathbf{X}_1)\le I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2).$$

Finally, let $\mathbb{D}_2$ denote a diagonal matrix with $\{a_{2,i}\}_{i=1}^m$, s.t. $|a_{2,i}|<\infty, 1\le i\le m$, on its diagonal, where $\{a_{2,i}\}_{i=1}^m$ are defined in the statement of the lemma. The proof is then completed by noting that for a given input covariance matrix $\mathbb{K}_{X_2}$, a circularly symmetric complex Normal $\mathbf{X}_2\sim\mathcal{CN}(\mathbf{0},\mathbb{K}_{X_2})$ maximizes $I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2)$ [25, Th. 2], hence

$$\begin{aligned}
\lim_{\gamma_2\to\infty} I(\mathbf{X}_2;\mathbf{X}_2+\mathbf{Z}_2) &\le \lim_{\gamma_2\to\infty}\Big(\log\big(|\gamma_2\cdot\mathbb{D}_2+\mathbb{K}_{X_2}|\big)-\log\big(|\gamma_2\cdot\mathbb{D}_2|\big)\Big)\\
&= \lim_{\gamma_2\to\infty}\log(|\mathbb{I}+\gamma_2^{-1}\cdot\mathbb{D}_2^{-1}\cdot\mathbb{K}_{X_2}|)\\
&= 0.
\end{aligned}$$

It thus follows that

$$\lim_{\gamma_2\to\infty} I(\mathbf{X}_1;\mathbf{X}_2+\mathbf{Z}_2|\mathbf{X}_1+\mathbf{Z}_1) = \lim_{\gamma_2\to\infty} I(\mathbf{X}_2;\mathbf{X}+\mathbf{Z}|\mathbf{X}_1) = 0,$$

which results in the desired equality (B.1). ∎

Appendix C
Proof of Lemma 6

We follow steps similar to those used in the proof of [24, Corollary 6], generalizing the derivations to hold for *complex vectors*. Using the notation of [25, Sec. III.A], define the $4n\times 2n$ real matrix $\mathbb{V}_{kk}$ as

$$\mathbb{V}_{kk}\triangleq\begin{bmatrix}\mathfrak{Re}\{\tilde{\mathbb{V}}_k\} & -\mathfrak{Im}\{\tilde{\mathbb{V}}_k\}\\ \mathfrak{Im}\{\tilde{\mathbb{V}}_k\} & \mathfrak{Re}\{\tilde{\mathbb{V}}_k\}\end{bmatrix}\quad k\in\{1,2\}.$$

Next, let $\mathbf{X}=X^{2n}$, and define the $4n\times 1$ real-valued vector $\bar{\mathbf{X}}\triangleq(\mathbf{X}_R^T,\mathbf{X}_I^T)^T$, where $\mathbf{X}_R$ and $\mathbf{X}_I$ are the real and the imaginary parts of $\mathbf{X}$, respectively. Note that similarly to [25, Sec. III-A], by using this notation we have $\mathbb{V}_{kk}^T\bar{\mathbf{X}} = \left(\mathfrak{Re}\{\tilde{\mathbb{V}}_k^H\mathbf{X}\}^T,\mathfrak{Im}\{\tilde{\mathbb{V}}_k^H\mathbf{X}\}^T\right)^T$; also note that this notation preserves the orthogonality among the columns of $\mathbb{V}_{kk}$ s.t. $\mathbb{V}_{kk}^T\mathbb{V}_{kk} = \mathbb{D}_{kk}^{-1} = \begin{bmatrix}\tilde{\mathbb{D}}_k^{-1} & \mathbb{O}_{n\times n}\\ \mathbb{O}_{n\times n} & \tilde{\mathbb{D}}_k^{-1}\end{bmatrix}, k\in\{1,2\}$, where $\mathbb{D}_{kk}^{-1}\in\mathfrak{R}^{2n\times 2n}$ and $\mathbb{O}_{m\times k}$ is the all-zero $m\times k$ marix. Therefore, we can find two $4n\times 2n$ matrices $\mathbb{V}_{12}$ and $\mathbb{V}_{21}$ s.t.

$\mathbb{V}_1 \triangleq (\mathbb{V}_{11}, \mathbb{V}_{12})$ and $\mathbb{V}_2 \triangleq (\mathbb{V}_{21}, \mathbb{V}_{22})$ are two $4n \times 4n$ matrices with orthogonal columns and hence, they can be written as $\mathbb{V}_k = \mathbb{U}_k \cdot \mathbb{C}_k$, where $\mathbb{U}_k \triangleq (\mathbb{U}_{k1}, \mathbb{U}_{k2}), k \in \{1, 2\}$ is a $4n \times 4n$ matrix with orthonormal columns, $\mathbb{U}_{lk}, l, k \in \{1, 2\}$ are four $4n \times 2n$ matrices with orthogonal columns, and $\mathbb{C}_k, k \in \{1, 2\}$ is a $4n \times 4n$ diagonal matrix whose elements are given by:

$$\mathbb{C}_k \triangleq \begin{bmatrix} \sqrt{\sum_{j=1}^{4n} \left|[\mathbb{V}_k]_{j,1}\right|^2} & 0 & \dots & 0 \\ 0 & \sqrt{\sum_{j=1}^{4n} \left|[\mathbb{V}_k]_{j,2}\right|^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\sum_{j=1}^{4n} \left|[\mathbb{V}_k]_{j,4n}\right|^2} \end{bmatrix}. \tag{C.1}$$

Let $\mathbb{C}_{kk}, k \in \{1, 2\}$ be two $2n \times 2n$ diagonal matrices whose elements are given by $\left[\mathbb{C}_{11}\right]_{i,i} = \left[\mathbb{C}_1\right]_{i,i}$, and $\left[\mathbb{C}_{22}\right]_{i,i} = \left[\mathbb{C}_2\right]_{2n+i,2n+i}, i \in \{1, 2, 3, \dots, 2n\}$. With these assignments we can represent $\mathbb{C}_k = \begin{bmatrix} \mathbb{C}_{1k} & \mathbb{O}_{2n\times 2n} \\ \mathbb{O}_{2n\times 2n} & \mathbb{C}_{2k} \end{bmatrix}$, where $\mathbb{C}_{12}$ and $\mathbb{C}_{21}$ are defined to satisfy (C.1). Using the above definitions we can write $\mathbb{V}_{kk}^T = \mathbb{C}_{kk}^T \cdot \mathbb{U}_{kk}^T, k \in \{1, 2\}$. Recall the definition of $\mathbb{D}_{kk}^{-1} \in \mathfrak{R}_{++}^{2n\times 2n}$:

$$\begin{aligned} \mathbb{D}_{kk}^{-1} &= \mathbb{V}_{kk}^T \mathbb{V}_{kk} \\ &= \mathbb{C}_{kk}^T \cdot \mathbb{U}_{kk}^T \cdot \mathbb{U}_{kk} \cdot \mathbb{C}_{kk} \\ &= \mathbb{C}_{kk}^T \mathbb{C}_{kk} \\ &= \mathbb{C}_{kk} \mathbb{C}_{kk} \\ &= \mathbb{C}_{kk}^2, \qquad k \in \{1, 2\}. \end{aligned}$$

Thus, from [27, Proposition 8.1.2 and Lemma 8.2.1] it follows that we can write $\mathbb{D}_{kk}^{-\frac{1}{2}} = \mathbb{C}_{kk} = \mathbb{C}_{kk}^T, k \in \{1, 2\}$. Using these definitions, we obtain

$$\mathbb{V}_{kk}^T = \mathbb{C}_{kk}^T \cdot \mathbb{U}_{kk}^T = \mathbb{D}_{kk}^{-\frac{1}{2}} \cdot \mathbb{U}_{kk}^T, \quad k \in \{1, 2\}. \tag{C.2}$$

Next, let $\mathbb{D}_{12}, \mathbb{D}_{21}, \bar{\mathbb{D}}_{11}$ and $\bar{\mathbb{D}}_{22}$ be four arbitrary $2n \times 2n$ dimensional diagonal matrices with real positive elements and define the following $4n \times 4n$ matrices:

$$\mathbb{D}_1 \triangleq \begin{bmatrix} \mathbb{D}_{11} & \mathbb{O}_{2n\times 2n} \\ \mathbb{O}_{2n\times 2n} & \mathbb{D}_{12} \end{bmatrix}, \tag{C.3a}$$

$$\mathbb{D}_2 \triangleq \begin{bmatrix} \mathbb{D}_{21} & \mathbb{O}_{2n\times 2n} \\ \mathbb{O}_{2n\times 2n} & \mathbb{D}_{22} \end{bmatrix}, \tag{C.3b}$$

$$\bar{\mathbb{D}}_1 \triangleq \begin{bmatrix} \bar{\mathbb{D}}_{11} & \mathbb{O}_{2n\times 2n} \\ \mathbb{O}_{2n\times 2n} & \mathbb{D}_{12} \end{bmatrix}, \tag{C.3c}$$

$$\bar{\mathbb{D}}_2 \triangleq \begin{bmatrix} \mathbb{D}_{21} & \mathbb{O}_{2n\times 2n} \\ \mathbb{O}_{2n\times 2n} & \bar{\mathbb{D}}_{22} \end{bmatrix}. \tag{C.3d}$$

Here, $\mathbb{D}_{lk}$ is a diagonal matrix with real positive elements defined by the vector $\boldsymbol{d}_{lk} = (d_{lk,1}, d_{lk,2}, \dots, d_{lk,2n}), lk \in \{12, 21, 11, 22\}$ on its diagonal. Similarly, $\bar{\mathbb{D}}_{kk}$ is a diagonal matrix with real positive elements defined by the vector $\bar{\boldsymbol{d}}_{kk} = (\bar{d}_{kk,1}, \bar{d}_{kk,2}, \dots, \bar{d}_{kk,2n}), k \in \{1, 2\}$ on its diagonal.

For $k \in \{1, 2\}$ set $\mathbb{K}_{\bar{Z}_k} \triangleq \mathbb{U}_k \cdot (\bar{\mathbb{D}}_k \cdot \mathbb{D}_k) \cdot \mathbb{U}_k^T$, let $\bar{\mathbf{Z}}_k$ be a $4n \times 1$ real-valued random vector distributed according to $\mathcal{N}\big(0, \mathbb{K}_{\bar{Z}_k}\big)$, and consider the following optimization problem:

$$\begin{aligned} &\max_{f(\bar{\mathbf{x}})} \; h(\bar{\mathbf{X}} + \bar{\mathbf{Z}}_1) - h(\bar{\mathbf{X}} + \bar{\mathbf{Z}}_2) \\ &\text{subject to:} \quad \mathrm{cov}(\bar{\mathbf{X}}) \preceq \mathbb{S}, \end{aligned} \tag{C.4}$$

where maximization is carried out over all $4n \times 1$ real vectors $\bar{\mathbf{X}}$. Denote $\mathrm{cov}(\bar{\mathbf{X}}) \equiv \mathbb{K}_{\bar{X}}$. From [24, Th. 1] it directly follows that a Gaussian random vector $\bar{\mathbf{X}}$ is an optimal solution to this problem. Furthermore, from [8, Lemma 1], we obtain that for any pair of complex random vectors $\bar{\mathbf{X}}$ and $\bar{\mathbf{X}}^{\mathrm{zm}} \triangleq \bar{\mathbf{X}} - \mathbb{E}\{\bar{\mathbf{X}}\}$, it holds that $h(\bar{\mathbf{X}}) = h(\bar{\mathbf{X}}^{\mathrm{zm}})$. Thus, we conclude that a *zero-mean* Gaussian random vector is the optimal solution to (C.4). Hence,

$$\begin{aligned} &\max_{f(\bar{\mathbf{x}}):\ \mathrm{cov}(\bar{\mathbf{X}})\preceq\mathbb{S}} h(\bar{\mathbf{X}} + \bar{\mathbf{Z}}_1) - h(\bar{\mathbf{X}} + \bar{\mathbf{Z}}_2) \\ &= \max_{\mathbb{K}_{\bar{X}}:\ 0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\frac{1}{2}\log\left(|\mathbb{K}_{\bar{X}} + \mathbb{K}_{\bar{Z}_1}|\right) \\ &\qquad -\frac{1}{2}\log\left(|\mathbb{K}_{\bar{X}} + \mathbb{K}_{\bar{Z}_2}|\right)\Big\}. \end{aligned} \tag{C.5}$$

Adding $h(\bar{\mathbf{Z}}_2) - h(\bar{\mathbf{Z}}_1)$ to both sides of (C.5) we obtain

$$\begin{aligned} &\max_{f(\bar{\mathbf{x}}):\ \mathrm{cov}(\bar{\mathbf{X}})\preceq\mathbb{S}} I(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_1) - I(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_2) \\ &= \max_{\mathbb{K}_{\bar{X}}:\ 0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\frac{1}{2}\log\left(|\mathbb{I}_{4n} + \mathbb{K}_{\bar{Z}_1}^{-1}\mathbb{K}_{\bar{X}}|\right) \\ &\qquad -\frac{1}{2}\log\left(|\mathbb{I}_{4n} + \mathbb{K}_{\bar{Z}_2}^{-1}\mathbb{K}_{\bar{X}}|\right)\Big\}. \end{aligned} \tag{C.6}$$

Next, define $\mathbb{U}_k^T \bar{\mathbf{Z}}_k \triangleq \hat{\mathbf{Z}}_k \equiv (\hat{\mathbf{Z}}_{k1}^T, \hat{\mathbf{Z}}_{k2}^T)^T$, where $\hat{\mathbf{Z}}_{kl}, (k, l) \in \{1, 2\} \times \{1, 2\}$ are four $2n \times 1$ vectors. Note that $\mathbb{U}_k$ is an orthogonal matrix, i.e., $\mathbb{U}_k^T \mathbb{U}_k = \mathbb{I}_{4n}$, and thus the covariance matrix of $\hat{\mathbf{Z}}_k$ is given by

$$\begin{aligned} \mathbb{K}_{\hat{Z}_k} &= \mathbb{U}_k^T \cdot \mathbb{K}_{\bar{Z}_k} \cdot \mathbb{U}_k \\ &= \mathbb{U}_k^T \cdot \mathbb{U}_k \cdot (\bar{\mathbb{D}}_k \cdot \mathbb{D}_k) \cdot \mathbb{U}_k^T \cdot \mathbb{U}_k \\ &= (\bar{\mathbb{D}}_k \cdot \mathbb{D}_k), \qquad k \in \{1, 2\}. \end{aligned} \tag{C.7}$$

Since $\hat{\mathbf{Z}}_k$, is a linear transformation of a Gaussian vector $\bar{\mathbf{Z}}_k$, it is a Gaussian random vector. Hence, it follows from (C.7) that $\hat{\mathbf{Z}}_{k1}$ and $\hat{\mathbf{Z}}_{k2}$ are mutually independent for $k \in \{1, 2\}$. Next, note that for any real random vector $\mathbf{X}$ and any real matrix $\mathbb{A}$ it holds that (see [25, eq. (13)]):

$$h\big(\mathbb{A} \cdot \mathbf{X}\big) = h\big(\mathbf{X}\big) + \log\big|\det(\mathbb{A})\big|. \tag{C.8}$$

Hence, since any orthogonal matrix is invertible, then we can write

$$I(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_1) = I(\mathbb{U}_1^T \bar{\mathbf{X}}; \mathbb{U}_1^T \bar{\mathbf{X}} + \mathbb{U}_1^T \bar{\mathbf{Z}}_1) \tag{C.9a}$$

$$I(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_2) = I(\mathbb{U}_2^T \bar{\mathbf{X}}; \mathbb{U}_2^T \bar{\mathbf{X}} + \mathbb{U}_2^T \bar{\mathbf{Z}}_2). \tag{C.9b}$$

Using Lemma 5, it follows that

$$\begin{aligned} \lim_{\boldsymbol{d}_{12}\to\infty} I(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_1) &= \lim_{\boldsymbol{d}_{12}\to\infty} I(\mathbb{U}_1^T \bar{\mathbf{X}}; \mathbb{U}_1^T \bar{\mathbf{X}} + \mathbb{U}_1^T \bar{\mathbf{Z}}_1) \\ &= I(\mathbb{U}_{11}^T \bar{\mathbf{X}}; \mathbb{U}_{11}^T \bar{\mathbf{X}} + \hat{\mathbf{Z}}_{11}) \\ \lim_{\boldsymbol{d}_{21}\to\infty} I(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_2) &= \lim_{\boldsymbol{d}_{21}\to\infty} I(\mathbb{U}_2^T \bar{\mathbf{X}}; \mathbb{U}_2^T \bar{\mathbf{X}} + \mathbb{U}_2^T \bar{\mathbf{Z}}_2) \\ &= I(\mathbb{U}_{22}^T \bar{\mathbf{X}}; \mathbb{U}_{22}^T \bar{\mathbf{X}} + \hat{\mathbf{Z}}_{22}), \end{aligned}$$

where we use $\boldsymbol{d}_{lk} \to \infty$ to imply that all the elements of $\boldsymbol{d}_{lk}$ go to infinity, i.e., $d_{lk,1}, d_{lk,2}, \ldots, d_{lk,2n} \to \infty$. Thus,

$$\lim_{\substack{\boldsymbol{d}_{12}\to\infty\\ \boldsymbol{d}_{21}\to\infty}} \left(I(\bar{\mathbf{X}};\bar{\mathbf{X}}+\bar{\mathbf{Z}}_1) - I(\bar{\mathbf{X}};\bar{\mathbf{X}}+\bar{\mathbf{Z}}_2)\right) = I(\mathbb{U}_{11}^T\bar{\mathbf{X}};\mathbb{U}_{11}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{11}) - I(\mathbb{U}_{22}^T\bar{\mathbf{X}};\mathbb{U}_{22}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{22}). \quad (C.10)$$

Note that for an $n \times m$ matrix $\mathbb{A}$ and an $m \times n$ matrix $\mathbb{B}$, Sylvester's determinant theorem [29, p. 271] states that $|\mathbb{I}_n + \mathbb{A}\mathbb{B}| = |\mathbb{I}_m + \mathbb{B}\mathbb{A}|$, and that given two square matrices $\mathbb{A}$ and $\mathbb{B}$, it holds that $(\mathbb{A}\cdot\mathbb{B})^{-1} = \mathbb{B}^{-1}\mathbb{A}^{-1}$, thus, due to the continuity of $\log(\mathbb{I}+\mathbb{A})$ over the semidefinite $\mathbb{A}$ we obtain from (C.6) (see, e.g., [24, eq. (164)])

$$\begin{aligned}
&\lim_{\substack{\boldsymbol{d}_{12}\to\infty\\ \boldsymbol{d}_{21}\to\infty}} \Big(\log\left(\left|\mathbb{I}_{4n} + \mathbb{K}_{\bar{Z}_1}^{-1}\mathbb{K}_{\bar{X}}\right|\right) - \log\left(\left|\mathbb{I}_{4n} + \mathbb{K}_{\bar{Z}_2}^{-1}\mathbb{K}_{\bar{X}}\right|\right)\Big)\\
&= \lim_{\substack{\boldsymbol{d}_{12}\to\infty\\ \boldsymbol{d}_{21}\to\infty}} \Big\{\log\left(\left|\mathbb{I}_{4n} + \mathbb{U}_1\cdot(\bar{\mathbb{D}}_1\cdot\mathbb{D}_1)^{-1}\cdot\mathbb{U}_1^T\cdot\mathbb{K}_{\bar{X}}\right|\right)\\
&\qquad - \log\left(\left|\mathbb{I}_{4n} + \mathbb{U}_2\cdot(\bar{\mathbb{D}}_2\cdot\mathbb{D}_2)^{-1}\cdot\mathbb{U}_2^T\cdot\mathbb{K}_{\bar{X}}\right|\right)\Big\}\\
&= \log\left(\left|\mathbb{I}_{4n} + \mathbb{U}_1\cdot\begin{bmatrix}(\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11})^{-1} & \mathbb{O}_{2n\times 2n}\\ \mathbb{O}_{2n\times 2n} & \mathbb{O}_{2n\times 2n}\end{bmatrix}\cdot\mathbb{U}_1^T\cdot\mathbb{K}_{\bar{X}}\right|\right)\\
&\quad - \log\left(\left|\mathbb{I}_{4n} + \mathbb{U}_2\cdot\begin{bmatrix}\mathbb{O}_{2n\times 2n} & \mathbb{O}_{2n\times 2n}\\ \mathbb{O}_{2n\times 2n} & (\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11})^{-1}\end{bmatrix}\cdot\mathbb{U}_2^T\cdot\mathbb{K}_{\bar{X}}\right|\right)\\
&= \log\left(\left|\mathbb{I}_{4n} + \mathbb{U}_{11}(\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11})^{-1}\mathbb{U}_{11}^T\mathbb{K}_{\bar{X}}\right|\right)\\
&\quad - \log\left(\left|\mathbb{I}_{4n} + \mathbb{U}_{22}(\bar{\mathbb{D}}_{22}\cdot\mathbb{D}_{22})^{-1}\mathbb{U}_{22}^T\mathbb{K}_{\bar{X}}\right|\right)\\
&= \log\left(\left|\mathbb{I}_{2n} + (\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11})^{-1}\mathbb{U}_{11}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{11}\right|\right)\\
&\quad - \log\left(\left|\mathbb{I}_{2n} + (\bar{\mathbb{D}}_{22}\cdot\mathbb{D}_{22})^{-1}\mathbb{U}_{22}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{22}\right|\right). \qquad (C.11)
\end{aligned}$$

Moreover, the convergence of (C.11) is uniform in $\mathbb{K}_{\bar{X}}$, because the continuity of $\log(|\mathbb{I}+\mathbb{A}|)$ over $\mathbb{A}$ is uniform, and $\mathbb{U}_k^T\cdot\mathbb{K}_{\bar{X}}\cdot\mathbb{U}_k, k\in\{1,2\}$ is bounded for $0 \preceq \mathbb{K}_{\bar{X}} \preceq \mathbb{S}$, in the sense that $0 \preceq \mathbb{U}_k^T\cdot\mathbb{K}_{\bar{X}}\cdot\mathbb{U}_k \preceq \mathbb{U}_k^T\cdot\mathbb{S}\cdot\mathbb{U}_k$. We thus have[4] (see e.g. [24, eq. (165)])

$$\begin{aligned}
&\lim_{\substack{\boldsymbol{d}_{12}\to\infty\\ \boldsymbol{d}_{21}\to\infty}} \Big\{\max_{\mathbb{K}_{\bar{X}}:0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\log\left(\left|\mathbb{I}_{4n} + \mathbb{K}_{\bar{Z}_1}^{-1}\mathbb{K}_{\bar{X}}\right|\right)\\
&\qquad - \log\left(\left|\mathbb{I}_{4n} + \mathbb{K}_{\bar{Z}_2}^{-1}\mathbb{K}_{\bar{X}}\right|\right)\Big\}\Big\}\\
&= \max_{\mathbb{K}_{\bar{X}}:0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\log\left(\left|\mathbb{I}_{2n} + (\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11})^{-1}\mathbb{U}_{11}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{11}\right|\right)\\
&\qquad - \log\left(\left|\mathbb{I}_{2n} + (\bar{\mathbb{D}}_{22}\cdot\mathbb{D}_{22})^{-1}\mathbb{U}_{22}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{22}\right|\right)\Big\}. \qquad (C.12)
\end{aligned}$$

[4] For uniform convergence $\lim_{y\to 0} f_y(x) = f(x)$ (see `http://www2.math.umd.edu/~czaja/chap1.pdf`): For any $\epsilon < 0$, $\exists\delta > 0$ s.t. $|f_y(x) - f(x)| < \epsilon$, $\forall y < \delta$, $\forall x \in \mathfrak{R}$. Hence, $\exists\delta > 0$ s.t. $\max_{x\in\mathfrak{R}} |f_y(x) - f(x)| \le \epsilon$, $\forall y < \delta$. Thus, $\exists\delta > 0$, s.t. $\forall y < \delta$, $|\max_{x\in\mathfrak{R}} f_y(x) - \max_{x\in\mathfrak{R}} f(x)| \le \max_{x\in\mathfrak{R}} |f_y(x) - f(x)| \le \epsilon$. Thus, $\lim_{y\to 0}\max_{x\in\mathfrak{R}} f_y(x) = \max_{x\in\mathfrak{R}} f(x)$.

Now, using (C.6), (C.10) and (C.12) we obtain

$$\begin{aligned}
&\max_{f(\bar{\mathbf{x}}):\,\mathrm{cov}(\bar{\mathbf{X}})\preceq\mathbb{S}} I(\mathbb{U}_{11}^T\bar{\mathbf{X}};\mathbb{U}_{11}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{11}) - I(\mathbb{U}_{22}^T\bar{\mathbf{X}};\mathbb{U}_{22}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{22})\\
&= \max_{\mathbb{K}_{\bar{X}}:\;0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\frac{1}{2}\log\left(\left|\mathbb{I}_{2n} + (\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11})^{-1}\mathbb{U}_{11}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{11}\right|\right)\\
&\qquad -\frac{1}{2}\log\left(\left|\mathbb{I}_{2n} + (\bar{\mathbb{D}}_{22}\cdot\mathbb{D}_{22})^{-1}\mathbb{U}_{22}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{22}\right|\right)\Big\},
\end{aligned}$$

or equivalently,

$$\begin{aligned}
&\max_{f(\bar{\mathbf{x}}):\;\mathrm{cov}(\bar{\mathbf{X}})\preceq\mathbb{S}} h(\mathbb{U}_{11}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{11}) - h(\mathbb{U}_{22}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{22})\\
&= \max_{\mathbb{K}_{\bar{X}}:\;0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\frac{1}{2}\log\left((\pi e)^n\left|\bar{\mathbb{D}}_{11}\cdot\mathbb{D}_{11} + \mathbb{U}_{11}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{11}\right|\right)\\
&\qquad -\frac{1}{2}\log\left((\pi e)^n\left|\bar{\mathbb{D}}_{22}\cdot\mathbb{D}_{22} + \mathbb{U}_{22}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{22}\right|\right)\Big\}.
\end{aligned}$$

Recall that $\mathbb{D}_{kk}, k = 1, 2$ is a diagonal matrix with real and positive entries. Thus, from [27, Proposition 8.1.2 and Lemma 8.2.1] we have that $\mathbb{D}_{kk}^{-1}, k\in\{1,2\}$ can be written as $\mathbb{D}_{kk}^{-1} = \mathbb{A}^2$, where $\mathbb{A}$ is a p.d. matrix defined as $\mathbb{A} \triangleq \mathbb{D}_{kk}^{-\frac{1}{2}}$. Using (C.8) once more, we obtain for $k = 1, 2$,

$$h(\mathbb{U}_{kk}^T\bar{\mathbf{X}}+\hat{\mathbf{Z}}_{kk}) = h(\mathbb{D}_{kk}^{-\frac{1}{2}}\mathbb{U}_{kk}^T\bar{\mathbf{X}} + \mathbb{D}_{kk}^{-\frac{1}{2}}\hat{\mathbf{Z}}_{kk}) - \log(|\mathbb{D}_{kk}^{-\frac{1}{2}}|).$$

Since both $\mathbb{D}_{kk}$ and $\bar{\mathbb{D}}_{kk}$ are diagonal matrices then,

$$\mathbb{D}_{kk}^{-\frac{1}{2}}\cdot\bar{\mathbb{D}}_{kk}\cdot\mathbb{D}_{kk}\cdot\mathbb{D}_{kk}^{-\frac{1}{2}} = \bar{\mathbb{D}}_{kk}\cdot\mathbb{D}_{kk}^{-\frac{1}{2}}\cdot\mathbb{D}_{kk}\cdot\mathbb{D}_{kk}^{-\frac{1}{2}} = \bar{\mathbb{D}}_{kk},$$

and hence, we obtain that

$$\begin{aligned}
&\max_{f(\bar{\mathbf{x}}):\;\mathrm{cov}(\bar{\mathbf{X}})\preceq\mathbb{S}} \Big\{h(\mathbb{D}_{11}^{-\frac{1}{2}}\mathbb{U}_{11}^T\bar{\mathbf{X}} + \mathbb{D}_{11}^{-\frac{1}{2}}\hat{\mathbf{Z}}_{11})\\
&\qquad - h(\mathbb{D}_{22}^{-\frac{1}{2}}\mathbb{U}_{22}^T\bar{\mathbf{X}} + \mathbb{D}_{22}^{-\frac{1}{2}}\hat{\mathbf{Z}}_{22})\Big\}\\
&= \max_{\mathbb{K}_{\bar{X}}:0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\frac{1}{2}\log\left((\pi e)^n\left|\bar{\mathbb{D}}_{11} + \mathbb{D}_{11}^{-\frac{1}{2}}\mathbb{U}_{11}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{11}\mathbb{D}_{11}^{-\frac{1}{2}}\right|\right)\\
&\qquad -\frac{1}{2}\log\left((\pi e)^n\left|\bar{\mathbb{D}}_{22} + \mathbb{D}_{22}^{-\frac{1}{2}}\mathbb{U}_{22}^T\mathbb{K}_{\bar{X}}\mathbb{U}_{22}\mathbb{D}_{22}^{-\frac{1}{2}}\right|\right)\Big\}.
\end{aligned}$$

Recall that from (C.2) we have $\mathbb{V}_{ll}^T = \mathbb{D}_{ll}^{-\frac{1}{2}}\mathbb{U}_{ll}^T$, and define $\mathbf{Z}_{ll} \triangleq \mathbb{D}_{ll}^{-\frac{1}{2}}\hat{\mathbf{Z}}_{ll}, l \in \{1,2\}$. Additionally, from (C.7) we conclude that $\hat{\mathbf{Z}}_{ll} \sim \mathcal{N}(\mathbf{0}, \bar{\mathbb{D}}_{ll}\cdot\mathbb{D}_{ll}), l\in\{1,2\}$. Hence, since $\bar{\mathbb{D}}_{ll}$ and $\mathbb{D}_{ll}$ are positive and real diagonal matrices, then $\mathbf{Z}_{ll}$ is distributed according to $\mathbf{Z}_{ll} \sim \mathcal{N}(\mathbf{0}, \bar{\mathbb{D}}_{ll})$. Thus,

$$\begin{aligned}
&\max_{f(\bar{\mathbf{x}}):\,\mathrm{cov}(\bar{\mathbf{X}})\preceq\mathbb{S}} h(\mathbb{V}_{11}^T\bar{\mathbf{X}}+\mathbf{Z}_{11}) - h(\mathbb{V}_{22}^T\bar{\mathbf{X}}+\mathbf{Z}_{22})\\
&= \max_{\mathbb{K}_{\bar{X}}:\;0\preceq\mathbb{K}_{\bar{X}}\preceq\mathbb{S}} \Big\{\frac{1}{2}\log\left((\pi e)^n\left|\bar{\mathbb{D}}_{11} + \mathbb{V}_{11}^T\mathbb{K}_{\bar{X}}\mathbb{V}_{11}\right|\right)\\
&\qquad -\frac{1}{2}\log\left((\pi e)^n\left|\bar{\mathbb{D}}_{22} + \mathbb{V}_{22}^T\mathbb{K}_{\bar{X}}\mathbb{V}_{22}\right|\right)\Big\}. \qquad (C.13)
\end{aligned}$$

The proof of Lemma 6 is completed by recalling that the elements of $\bar{\mathbb{D}}_{kk}, k \in \{1,2\}$ are chosen arbitrarily and thus, (C.13) holds for any p.d. $\bar{\mathbb{D}}_{kk}$, and in particular for $\bar{\mathbb{D}}_{kk} \triangleq \begin{bmatrix}\frac{1}{2}\tilde{\mathbb{D}}_k^z & \mathbb{O}_{n\times n}\\ \mathbb{O}_{n\times n} & \frac{1}{2}\tilde{\mathbb{D}}_k^z\end{bmatrix}$, $k\in\{1,2\}$, and by noting that a Gaussian random vector $\bar{\mathbf{X}}$ achieves the r.h.s. of (C.13) with equality, i.e., a zero-mean complex Normal $\mathbf{X}$ is an optimal solution to (6). This completes the proof of Lemma 6. ∎

Appendix D
Proof of Proposition 1

First note that from the construction of the genie signals, it follows that the entropy expressions $h(S_{1G}|X_{1G}, X_{3G}, \tilde{H}_1)$ and $h(S_{2G}|X_{2G}, \tilde{H}_2)$ do not depend on υ and on (P_1, P_2, P_3). Next, recall that $(X_{1G}, X_{2G}, X_{3G})^T \sim \mathcal{CN}(\mathbf{0}, \mathbb{Q}_G)$ and consider $h(Y_{2G}|S_{2G}, \tilde{H}_2)$: Defining $\theta_{\eta_2^*\tilde{\upsilon}_2^*} \triangleq \arg\{\eta_2^*\tilde{\upsilon}_2^*\}$ we can write:

$$\begin{aligned}
&h(Y_{2G}|S_{2G}, \tilde{H}_2) - \log(\pi e)\\
&\overset{(a)}{=} \mathbb{E}_{\tilde{H}_2}\left\{\log\left(\text{var}(Y_{2G}) - \frac{|\mathbb{E}\{Y_{2G}S_{2G}^*\}|^2}{\text{var}(S_{2G})}\right)\right\}\\
&= \mathbb{E}_{\tilde{H}_2}\left\{\log\left(1 + \mathsf{SNR}_{22}P_2 + \mathsf{SNR}_{32}P_3 - \frac{|\mathsf{SNR}_{22}P_2 + \eta_2^*\tilde{\upsilon}_2^*|^2}{\mathsf{SNR}_{22}P_2 + |\eta_2|^2}\right)\right\}\\
&= \mathbb{E}_{\tilde{H}_2}\left\{\log\left(1 + \mathsf{SNR}_{22}P_2 + \mathsf{SNR}_{32}P_3 - \frac{\mathsf{SNR}_{22}^2P_2^2 + |\eta_2^*|^2|\tilde{\upsilon}_2^*|^2 + 2|\eta_2^*||\tilde{\upsilon}_2^*|\cos(\theta_{\eta_2^*\tilde{\upsilon}_2^*})\mathsf{SNR}_{22}P_2}{\mathsf{SNR}_{22}P_2 + |\eta_2|^2}\right)\right\}\\
&\triangleq \mathbb{E}_{\tilde{H}_2}\left\{\log\left(f_1(P_2)\right)\right\},
\end{aligned}$$

where (a) follows from a direct calculation via $h(Y_{2G}|S_{2G}, \tilde{H}_2) = h(Y_{2G}, S_{2G}|\tilde{H}_2) - h(S_{2G}|\tilde{H}_2)$. Observe that $f_1(P_2)$ is independent of (υ, P_1), and that it increases with respect to P_3. Additionally, since (Y_{2G}, S_{2G}) are jointly circularly symmetric complex Normal when $\tilde{H}_2 = \tilde{h}_2$ is given, we note that

$$\begin{aligned}
&\frac{\partial f_1(P_2)}{\partial P_2}\\
&= \left[\frac{\left(\mathsf{SNR}_{22}|\eta_2|^2 - 2|\eta_2^*||\tilde{\upsilon}_2^*|\cos(\theta_{\eta_2^*\tilde{\upsilon}_2^*})\mathsf{SNR}_{22}\right)}{\left(\mathsf{SNR}_{22}P_2 + |\eta_2|^2\right)^2} \cdot\left(\mathsf{SNR}_{22}P_2 + |\eta_2|^2\right)\right]\\
&\quad - \left[\frac{\left(\mathsf{SNR}_{22}P_2|\eta_2|^2 - |\eta_2^*|^2|\tilde{\upsilon}_2^*|^2 - 2|\eta_2^*||\tilde{\upsilon}_2^*|\cos(\theta_{\eta_2^*\tilde{\upsilon}_2^*})\mathsf{SNR}_{22}P_2\right)}{\left(\mathsf{SNR}_{22}P_2 + |\eta_2|^2\right)^2}\cdot\mathsf{SNR}_{22}\right]\\
&= \frac{\mathsf{SNR}_{22}\cdot\left(|\eta_2|^4 + |\eta_2^*|^2|\tilde{\upsilon}_2^*|^2 - 2|\eta_2|^2|\eta_2^*||\tilde{\upsilon}_2^*|\cos(\theta_{\eta_2^*\tilde{\upsilon}_2^*})\right)}{\left(\mathsf{SNR}_{22}P_2 + |\eta_2|^2\right)^2}\\
&\geq \frac{\mathsf{SNR}_{22}\left(|\eta_2^*|^2 - |\eta_2^*||\tilde{\upsilon}_2^*|\right)^2}{\left(\mathsf{SNR}_{22}P_2 + |\eta_2|^2\right)^2}\\
&\geq 0.
\end{aligned}$$

We therefore conclude that $h(Y_{2G}|S_{2G}, \tilde{H}_2)$ is maximized with $P_2 = P_3 = 1$, and that setting $\upsilon = 0$ and $P_1 = 1$ does not affect the value of $h(Y_{2G}|S_{2G}, \tilde{H}_2)$.

Next, consider $h(Y_{1G}|S_{1G}, \tilde{H}_1)$, and define

$$\begin{aligned}
c_1 &\triangleq \mathsf{SNR}_{11}P_1 + \mathsf{SNR}_{31}P_3 && \text{(D.1a)}\\
c_2 &\triangleq 2\sqrt{\mathsf{SNR}_{11}P_1\mathsf{SNR}_{31}P_3} && \text{(D.1b)}\\
c_3 &\triangleq 1 + \mathsf{SNR}_{21}P_2 && \text{(D.1c)}\\
c_4 &\triangleq |\eta_1^*|^2 && \text{(D.1d)}\\
c_5 &\triangleq |\eta_1^*||\tilde{\upsilon}_1^*| && \text{(D.1e)}\\
\theta_1 &\triangleq \arg(h_{11}h_{31}^*\upsilon) && \text{(D.1f)}\\
\theta_2 &\triangleq \arg(\eta_1^*\tilde{\upsilon}_1^*). && \text{(D.1g)}
\end{aligned}$$

With these definitions we can write an explicit expression for $h(Y_{1G}|S_{1G}, \tilde{H}_1) - \log(\pi e)$ as in Eqn. (D.2), as shown at the top of the next page where step (a) in the derivation leading to (D.2) follows from an explicit analytical calculation.[5] Next, differentiating $f_2(|\upsilon|)$ with respect to $|\upsilon|$ we obtain

$$\frac{\partial f_2(|\upsilon|)}{\partial|\upsilon|} = -\frac{c_2\cdot\left(c_4^2 + c_5^2 - 2c_4c_5\cos(\theta_2)\right)}{(c_1 + c_4 - c_2|\upsilon|)^2},$$

and we note that since both c_4 and c_5 are non-negative real numbers, then $0 \leq (c_4 - c_5)^2 \leq \left(c_4^2 + c_5^2 - 2c_4c_5\cos(\theta_2)\right)$. Thus, as c_2 is positive, then the derivative of $f_2(|\upsilon|)$ with respect to $|\upsilon|$ is non-positive, i.e., $f_2(|\upsilon|)$ is a non-increasing function of $|\upsilon|$, and hence, it is maximized at $|\upsilon| = 0$. Next, setting $|\upsilon| = 0$ in $f_2(|\upsilon|)$ we obtain:

$$\begin{aligned}
f_2(|\upsilon|)\Big|_{|\upsilon|=0} &= \frac{c_3c_4 + c_1(c_3 + c_4) - c_5^2 - 2c_5c_1\cos(\theta_2)}{c_1 + c_4}\\
&\triangleq f_3(P_1, P_2, P_3).
\end{aligned}$$

Note that $f_3(P_1, P_2, P_3)$ is a monotonically increasing function of c_3 and is independent of c_2. Additionally, note that since both c_4 and c_5 are nonnegative real numbers, then

$$\frac{\partial f_3(P_1, P_2, P_3)}{\partial c_1} = \frac{c_4^2 + c_5^2 - 2c_4c_5\cos(\theta_2)}{(c_1 + c_4)^2} \geq 0.$$

We therefore conclude that $f_3(P_1, P_2, P_3)$ is a non-decreasing function of c_1. From the definition of (c_1, c_2, c_3) in (D.1), we conclude that $f_3(P_1, P_2, P_3)$ is non-decreasing with respect to P_1, P_2 and P_3. In conclusion, $h(Y_{1G}|S_{1G}, \tilde{H}_1)$ is maximized with $\upsilon = 0$ and $P_1 = P_2 = P_3 = 1$. This completes the proof of Proposition 1. ■

Appendix E
Proving the Maximizing Distribution for Equation (20) in Theorem 2 is i.i.d.

The objective of this appendix is to derive the maximum value and characterize the associated maximizing distribution, for the expectation:

$$\begin{aligned}
\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Bigg\{&h\left(\mathbb{H}_{h_{31}}^{(n)}X_{3\bar{G}}^n + \mathbb{H}_{h_{11}}^{(n)}X_{1\bar{G}}^n + \eta_1\cdot W_1^n\big|\tilde{H}_1^n = \tilde{h}_1^n\right)\\
&-h\left(\mathbb{H}_{h_{32}}^{(n)}X_{3\bar{G}}^n + V_2^n\big|\tilde{H}_2^n = \tilde{h}_2^n\right)\Bigg\}, \qquad \text{(E.1)}
\end{aligned}$$

where $W_1^n \sim \mathcal{CN}(\mathbf{0}, \mathbb{I}_n)$, and $V_2^n \sim \mathcal{CN}\left(\mathbf{0}, \left(1 - |\upsilon_2|^2\right)\mathbb{I}_n\right)$.

[5] See details in [34].

$$
\begin{aligned}
&h(Y_{1G}|S_{1G}, \tilde{H}_1) - \log(\pi e)\\
&= \mathbb{E}_{\tilde{H}_1}\left\{ \log\left(\mathrm{var}(Y_{1G}) - \frac{|\mathbb{E}\{Y_{1G}S_{1G}^*\}|^2}{\mathrm{var}(S_{1G})}\right)\right\}\\
&= \frac{1}{2\pi}\int_{\theta_1=0}^{2\pi}\left(\log\left(2\pi\cdot\left(c_3 + \frac{c_1c_4 + c_2c_4\cos(\theta_1)|v| - c_5^2 - (c_1 + c_2\cos(\theta_1)|v|)\cdot 2c_5\cos(\theta_2)}{c_1 + c_2\cos(\theta_1)|v| + c_4}\right)\right) - \log(2\pi)\right) d\theta_1\\
&\overset{(a)}{=} \log\left(2\pi\cdot\frac{c_3c_4 + c_1(c_3+c_4) - c_5^2 - c_2c_3|v| - c_2c_4|v| - 2c_5(c_1 - c_2|v|)\cos(\theta_2)}{c_1 + c_4 - c_2|v|}\right) - \log(2\pi)\\
&\triangleq \log\left(f_2(|v|)\right)
\end{aligned}
\tag{D.2}
$$

Let $\mathbb{h}_{h_{kl}}^{(n)}$ denote a realization of $\mathbb{H}_{h_{kl}}^{(n)}$, where $(k,l) \in \{(1,1),(3,1),(3,2)\}$. Next, consider $h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n + \mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n + \eta_1\cdot W_1^n \middle| \tilde{H}_1^n = \tilde{h}_1^n\right)$, and define the $2n\times 2n$ real matrices

$$\mathbb{h}_{31} \triangleq \begin{bmatrix} \mathfrak{Re}\left\{\mathbb{h}_{h_{31}}^{(n)}\right\} & -\mathfrak{Im}\left\{\mathbb{h}_{h_{31}}^{(n)}\right\} \\ \mathfrak{Im}\left\{\mathbb{h}_{h_{31}}^{(n)}\right\} & \mathfrak{Re}\left\{\mathbb{h}_{h_{31}}^{(n)}\right\}\end{bmatrix}, \tag{E.2a}$$

$$\mathbb{h}_{11} \triangleq \begin{bmatrix} \mathfrak{Re}\left\{\mathbb{h}_{h_{11}}^{(n)}\right\} & -\mathfrak{Im}\left\{\mathbb{h}_{h_{11}}^{(n)}\right\} \\ \mathfrak{Im}\left\{\mathbb{h}_{h_{11}}^{(n)}\right\} & \mathfrak{Re}\left\{\mathbb{h}_{h_{11}}^{(n)}\right\}\end{bmatrix}, \tag{E.2b}$$

$$\bar{X}_{k\bar{G}}^{2n} = \left(\left(\mathfrak{Re}\left\{X_{k\bar{G}}^n\right\}\right)^T, \left(\mathfrak{Im}\left\{X_{k\bar{G}}^n\right\}\right)^T\right)^T, \quad k\in\{1,3\}.$$

Since $X_{1\bar{G}}^n$ and $X_{3\bar{G}}^n$ are two zero-mean complex jointly Gaussian random vectors, then $\left(\bar{X}_{1\bar{G}}^{2n}, \bar{X}_{3\bar{G}}^{2n}\right)$ are zero mean real jointly Gaussian vectors. Note that since $\mathbb{h}_{h_{31}}^{(n)}$ is a diagonal matrix with $\left[\mathbb{h}_{h_{31}}^{(n)}\right]_{i,i} = h_{31,i}$, then $\mathfrak{Re}\left\{\mathbb{h}_{h_{31}}^{(n)}\right\}$ and $\mathfrak{Im}\left\{\mathbb{h}_{h_{31}}^{(n)}\right\}$ are diagonal matrices. Next, define $\theta_{31,i} \triangleq \arg\left\{h_{31,i}\right\}$, and write $h_{31,i} = \sqrt{\mathsf{SNR}}_{31}\cdot e^{j\theta_{31,i}}$. Additionally, for $k\in\{1,3\}$ we define the $n\times n$ real diagonal matrices $\mathbb{C}_{k1}$ and $\mathbb{S}_{k1}$ whose diagonal elements are given by $[\mathbb{C}_{k1}]_{i,i} = \cos\left(\theta_{k1,i}\right)$, $[\mathbb{S}_{k1}]_{i,i} = \sin\left(\theta_{k1,i}\right)$. With these definitions we write

$$\mathfrak{Re}\left\{\mathbb{h}_{h_{k1}}^{(n)}\right\} = \sqrt{\mathsf{SNR}_{k1}}\cdot\mathbb{C}_{k1}, \quad \mathfrak{Im}\left\{\mathbb{h}_{h_{k1}}^{(n)}\right\} = \sqrt{\mathsf{SNR}_{k1}}\cdot\mathbb{S}_{k1}.$$

Letting $\mathbb{U}_{31} \triangleq \frac{1}{\sqrt{\mathsf{SNR}_{31}}}\mathbb{h}_{31} = \begin{bmatrix}\mathbb{C}_{31} & -\mathbb{S}_{31}\\ \mathbb{S}_{31} & \mathbb{C}_{31}\end{bmatrix}$, we obtain $\mathbb{U}_{31}^T\mathbb{U}_{31} = \mathbb{I}_{2n}$. Hence, we conclude that $\mathbb{U}_{31}$ is a $2n\times 2n$ orthogonal matrix, and that the matrix $\mathbb{h}_{31}$ can now be written as $\mathbb{h}_{31} = \sqrt{\mathsf{SNR}_{31}}\cdot\mathbb{U}_{31}$. Similarly we define $\mathbb{U}_{11} = \begin{bmatrix}\mathbb{C}_{11} & -\mathbb{S}_{11}\\ \mathbb{S}_{11} & \mathbb{C}_{11}\end{bmatrix}$. Next, consider $\eta_1\cdot W_1^n$. Begin by writing:

$$
\begin{aligned}
\eta_1\cdot W_1^n &= \left(\mathfrak{Re}\{\eta_1\} + j\cdot\mathfrak{Im}\{\eta_1\}\right)\left(\mathfrak{Re}\left\{W_1^n\right\} + j\cdot\mathfrak{Im}\left\{W_1^n\right\}\right)\\
&= \left(\mathfrak{Re}\{\eta_1\}\mathfrak{Re}\left\{W_1^n\right\} - \mathfrak{Im}\{\eta_1\}\mathfrak{Im}\left\{W_1^n\right\}\right)\\
&\quad + j\cdot\left(\mathfrak{Re}\{\eta_1\}\mathfrak{Im}\left\{W_1^n\right\} + \mathfrak{Im}\{\eta_1\}\mathfrak{Re}\left\{W_1^n\right\}\right),
\end{aligned}
\tag{E.3}
$$

and define $\bar{W}_\eta^{2n} \triangleq \left(\left(\mathfrak{Re}\{\eta_1\cdot W_1^n\}\right)^T, \left(\mathfrak{Im}\{\eta_1\cdot W_1^n\}\right)^T\right)^T$ and $\bar{W}_1^{2n} \triangleq \left(\left(\mathfrak{Re}\left\{W_1^n\right\}\right)^T, \left(\mathfrak{Im}\left\{W_1^n\right\}\right)^T\right)^T$. Since W_1^n is an i.i.d. circularly symmetric complex Gaussian vector, where each element has a unit variance, then we conclude that[6] $\bar{W}_1^{2n}$ is a real jointly Normal vector with covariance matrix $\mathrm{cov}(\bar{W}_1^{2n}) = \frac{1}{2}\mathbb{I}_{2n}$, see [25, Lemma 5]. Next, from (E.3) it directly follows that $\mathfrak{Re}\left\{\eta_1\cdot W_1^n\right\} \sim \mathcal{N}\left(\mathbf{0},\ |\eta_1|^2\frac{1}{2}\mathbb{I}_n\right)$ and $\mathfrak{Im}\left\{\eta_1\cdot W_1^n\right\} \sim \mathcal{N}\left(\mathbf{0}, |\eta_1|^2\frac{1}{2}\mathbb{I}_n\right)$, and consequently, $\bar{W}_\eta^{2n} \sim \mathcal{N}\left(\mathbf{0}, \frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\right)$. Defining $\bar{V}_2^{2n} = \left(\left(\mathfrak{Re}\{V_2^n\}\right)^T, \left(\mathfrak{Im}\{V_2^n\}\right)^T\right)^T$, we similarly conclude $\bar{V}_2^{2n} \sim \mathcal{N}\left(\mathbf{0}, \frac{1-|v_2|^2}{2}\mathbb{I}_{2n}\right)$. Lastly, write

$$
\begin{aligned}
&\mathbb{h}_{h_{k1}}^{(n)}X_{k\bar{G}}^n\\
&\quad = \left(\mathfrak{Re}\left\{\mathbb{h}_{h_{k1}}^{(n)}\right\}\mathfrak{Re}\left\{X_{k\bar{G}}^n\right\} - \mathfrak{Im}\left\{\mathbb{h}_{h_{k1}}^{(n)}\right\}\mathfrak{Im}\left\{X_{k\bar{G}}^n\right\}\right)\\
&\qquad + j\cdot\left(\mathfrak{Re}\left\{\mathbb{h}_{h_{k1}}^{(n)}\right\}\mathfrak{Im}\left\{X_{k\bar{G}}^n\right\} + \mathfrak{Im}\left\{\mathbb{h}_{h_{k1}}^{(n)}\right\}\mathfrak{Re}\left\{X_{k\bar{G}}^n\right\}\right).
\end{aligned}
$$

From the above assignments it also follows that $\mathbb{h}_{h_{k1}}^{(n)}X_{k\bar{G}}^n$ is statistically equivalent to $\mathbb{h}_{k1}\bar{X}_{k\bar{G}}^{2n}$, $k\in\{1,3\}$, and that $\mathbb{h}_{h_{32}}^{(n)}X_{3\bar{G}}^n$ is statistically equivalent to $\mathbb{h}_{32}\bar{X}_{3\bar{G}}^{2n}$. Hence,

$$
\begin{aligned}
&\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{h\left(\mathbb{H}_{h_{31}}^{(n)}X_{3\bar{G}}^n + \mathbb{H}_{h_{11}}^{(n)}X_{1\bar{G}}^n + \eta_1\cdot W_1^n\middle|\tilde{H}_1^n = \tilde{h}_1^n\right)\right\}\\
&\quad = \mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n} + \mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n} + \bar{W}_\eta^{2n}\middle|\tilde{H}_1^n = \tilde{h}_1^n\right)\right\}.
\end{aligned}
\tag{E.4}
$$

Denote the $2n\times 2n$ covariance matrices for $\bar{X}_{3\bar{G}}^{2n}$ and $\bar{X}_{1\bar{G}}^{2n}$ by $\mathbb{K}_{\bar{X}_{3\bar{G}}}$ and $\mathbb{K}_{\bar{X}_{1\bar{G}}}$, respectively, and let $\mathbb{K}_{\bar{X}_{13\bar{G}}} \triangleq \mathbb{E}\left\{\bar{X}_{1\bar{G}}^{2n}\cdot\left(\bar{X}_{3\bar{G}}^{2n}\right)^T\right\}$. Next, as $X_{1\bar{G}}^n$ and $X_{3\bar{G}}^n$ are complex jointly Gaussian and both are independent of W_1^n, it follows that given $\tilde{H}_1^n = \tilde{h}_1^n$, then $\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n} + \mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n} + \bar{W}_\eta^{2n}$ is a jointly Gaussian real random vector, whose covariance matrix, $\mathrm{cov}\left(\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n} + \mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n} + \bar{W}_\eta^{2n}\right) \triangleq \mathbb{T}_1$, is given by:

$$
\begin{aligned}
\mathbb{T}_1 &= \mathbb{h}_{31}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{31}^T + \mathbb{h}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{h}_{11}^T\\
&\qquad + \mathbb{h}_{11}\mathbb{K}_{\bar{X}_{13\bar{G}}}\mathbb{h}_{31}^T + \mathbb{h}_{31}\mathbb{K}_{\bar{X}_{13\bar{G}}}^T\mathbb{h}_{11}^T + \frac{|\eta_1|^2}{2}\mathbb{I}_{2n},
\end{aligned}
$$

[6]For complex Normal $\mathbf{Z} = \mathbf{X} + jY$, where $\mathbf{X}$ and Y are real vectors, define $\mu = \mathbb{E}\{\mathbf{Z}\}$, $\mathbb{G} = \mathbb{E}\{(\mathbf{Z}-\mu)(\mathbf{Z}-\mu)^H\}$, and $\mathbb{J} = \mathbb{E}\{(\mathbf{Z}-\mu)(\mathbf{Z}-\mu)^T\}$. Then, $(\mathbf{X}^T, Y^T)^T$ is a real jointly Gaussian vector with covariance matrix $\mathrm{cov}(\mathbf{X}, Y) = \frac{1}{2}\mathfrak{Im}\{-\mathbb{G}+\mathbb{J}\}$. For a circularly symmetric complex Normal vector $\mathbf{Z}$ we have $\mu = 0$, $\mathbb{J} = \mathbb{O}$, and $\mathrm{cov}(\mathbf{X}) = \mathrm{cov}(Y) = \frac{1}{2}\mathfrak{Re}\{\mathbb{G}\}$. When $\mathbf{Z}$ is circularly symmetric complex Normal with i.i.d. elements, each has a unit variance, then $\mathbb{G} = \mathbb{I}_n$ is real, hence, $\mathrm{cov}(\mathbf{X}, Y) = \mathbb{O}$, and from joint Gaussianity it follows that the real and imaginary parts of $\mathbf{Z}$ are mutually independent, see `http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf`.

hence, its differential entropy is given by [25, p. 1296]:

$$h\left(\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n}+\mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n}+\bar{W}_{\eta}^{2n}\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)=\frac{1}{2}\log\det\left((2\pi)^{2n}\mathbb{T}_1\right).$$

Note that the expectation over $\tilde{H}_1^n$ in (E.4) is taken with respect to the phases of the channel coefficients, which are mutually independent over the links, i.i.d. over time, and distributed uniformly over $[0, 2\pi)$. Since $\cos(x+\pi) = -\cos(x)$ and $\sin(x+\pi) = -\sin(x)$, then replacing $\mathbb{h}_{11}$ with $-\mathbb{h}_{11}$ is equivalent to shifting the phases of all time realizations of H_{11}, i.e., all coefficients in $\mathbb{h}_{h_{11}}^{(n)}$ by π. As the complex exponential is periodic with a period of 2π, and the expectation spans a continuous intervals of 2π radians, a constant phase shift to all elements of the diagonal matrix $\mathbb{h}_{11}^{(n)}$ does not affect the expectation, see proof of [22, Th. 8], and consequently:

$$\begin{aligned}
&\mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n+\mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\right\}\\
&\quad=\mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n-\mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\right\}\\
&\quad=\mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n}-\mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n}+\bar{W}_{\eta}^{2n}\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\right\}. \quad (E.5)
\end{aligned}$$

Let the covariance matrix for $\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n}-\mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n}+\bar{W}_{\eta}^{2n}$ be denoted with $\mathbb{T}_2 \triangleq \operatorname{cov}\left(\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n}-\mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n}+\bar{W}_{\eta}^{2n}\right)$. $\mathbb{T}_2$ can be written as

$$\begin{aligned}
\mathbb{T}_2 &= \mathbb{h}_{31}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{31}^T+\mathbb{h}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{h}_{11}^T\\
&\quad-\mathbb{h}_{11}\mathbb{K}_{\bar{X}_{13\bar{G}}}\mathbb{h}_{31}^T-\mathbb{h}_{31}\mathbb{K}_{\bar{X}_{13\bar{G}}}^T\mathbb{h}_{11}^T+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}.
\end{aligned}$$

Using $\mathbb{T}_1$ and $\mathbb{T}_2$ we can write

$$\begin{aligned}
&\mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n+\mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\right\}\\
&=\frac{1}{2}\mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n+\mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\right\}\\
&\quad+\frac{1}{2}\mathbb{E}_{\tilde{H}_1^n}\left\{h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n-\mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\right\}\\
&=\frac{1}{2}\mathbb{E}_{\tilde{H}_1^n}\left\{\frac{1}{2}\log\det\left((2\pi)^{2n}\ \mathbb{T}_1\right)\right\}\\
&\quad+\frac{1}{2}\mathbb{E}_{\tilde{H}_1^n}\left\{\frac{1}{2}\log\det\left((2\pi)^{2n}\ \mathbb{T}_2\right)\right\}\\
&\overset{(a)}{\le}\mathbb{E}_{\tilde{H}_1^n}\left\{\frac{1}{2}\log\det\left(\frac{1}{2}(2\pi)^{2n}\mathbb{T}_1+\frac{1}{2}(2\pi)^{2n}\mathbb{T}_2\right)\right\}\\
&=\mathbb{E}_{\tilde{H}_1^n}\left\{\frac{1}{2}\log\det\left((2\pi)^{2n}\right)\right.\\
&\quad\left.+\frac{1}{2}\log\det\left(\mathbb{h}_{31}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{31}^T+\mathbb{h}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{h}_{11}^T+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\right)\right\},
\end{aligned}$$

where the inequality (a) follows from the concavity of the logdet function in space of p.d symmetric matrices [26, Sec. 3.1.5]. Observe that the inequality (a) is obtained with equality when $\mathbb{K}_{\bar{X}_{13\bar{G}}} = \mathbb{O}_{2n\times 2n}$. As $\bar{X}_{3\bar{G}}^{2n}$ and $\bar{X}_{1\bar{G}}^{2n}$, are jointly Gaussian, then zero cross-correlation implies $\bar{X}_{3\bar{G}}^{2n}$ and $\bar{X}_{1\bar{G}}^{2n}$ are mutually independent. Note that $\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{h\left(\mathbb{h}_{h_{32}}^{(n)}X_{3\bar{G}}^n+V_2^n\middle|\tilde{H}_2^n=\tilde{h}_2^n\right)\right\} = \mathbb{E}_{\tilde{H}_2^n}\left\{h\left(\mathbb{h}_{h_{32}}^{(n)}X_{3\bar{G}}^n+V_2^n\middle|\tilde{H}_2^n=\tilde{h}_2^n\right)\right\}$ is not affected by the correlation between $\bar{X}_{3\bar{G}}^{2n}$ and $\bar{X}_{1\bar{G}}^{2n}$, hence, we conclude that (E.1) is maximized by mutually independent $\bar{X}_{3\bar{G}}^{2n}$ and $\bar{X}_{1\bar{G}}^{2n}$, and henceforth we shall proceed with this assumption.

Next, define $\hat{\mathbb{K}} \triangleq \operatorname{cov}\left\{\mathbb{h}_{31}\bar{X}_{3\bar{G}}^{2n}+\mathbb{h}_{11}\bar{X}_{1\bar{G}}^{2n}\middle|\tilde{H}_1^n=\tilde{h}_1^n\right\} = \mathbb{h}_{31}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{31}^T+\mathbb{h}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{h}_{11}^T$. As $\mathbb{K}_{\bar{X}_{3\bar{G}}}$ and $\mathbb{K}_{\bar{X}_{1\bar{G}}}$ are both covariance matrices for real random vectors, they are both symmetric, hence, $\hat{\mathbb{K}}$ and $\operatorname{cov}\left(\mathbb{h}_{32}\bar{X}_{3\bar{G}}^{2n}\middle|\tilde{H}_2^n=\tilde{h}_2^n\right) \triangleq \mathbb{h}_{32}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{32}^T$ are both real symmetric matrices whose eigenvalues are real and non-negative. For a square and symmetric real matrix $\mathbb{B}$, we use $\operatorname{eig}(\mathbb{B})$ to denote a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbb{B}$, when the eigenvectors are orthogonal and have a unit norm, and the eigenvalues appear in ascending order in $\operatorname{eig}(\mathbb{B})$, i.e., $\operatorname{eig}(\mathbb{B})_{i,i} \le \operatorname{eig}(\mathbb{B})_{k,k}$ for $i \le k$, [33, p. 549]. In the following we will refer only to eigenvalue decompositions of this form. We next write the eigenvalue representations for $\mathbb{h}_{32}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{32}^T$ and for $\hat{\mathbb{K}}$ as $\mathbb{h}_{32}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{32}^T = \hat{\mathbb{U}}_{32\bar{G}}\hat{\mathbb{D}}_{32\bar{G}}\hat{\mathbb{U}}_{32\bar{G}}^T$, and $\hat{\mathbb{K}} = \hat{\mathbb{U}}\hat{\mathbb{D}}\hat{\mathbb{U}}^T$. As the determinant of an orthogonal matrix is 1, we can write

$$\begin{aligned}
&h\left(\mathbb{h}_{h_{31}}^{(n)}X_{3\bar{G}}^n+\mathbb{h}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\middle|\tilde{H}_1^n=\tilde{h}_1^n\right)\\
&\quad-h\left(\mathbb{h}_{h_{32}}^{(n)}X_{3\bar{G}}^n+V_2^n\middle|\tilde{H}_2^n=\tilde{h}_2^n\right)\\
&=\frac{1}{2}\log\det\left(\hat{\mathbb{K}}+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\right)\\
&\quad-\frac{1}{2}\log\det\left(\mathbb{h}_{32}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{32}^T+\frac{1-|\tilde{\upsilon}_2|^2}{2}\mathbb{I}_{2n}\right)\\
&=\frac{1}{2}\log\det\left(\hat{\mathbb{D}}+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\right)\\
&\quad-\frac{1}{2}\log\det\left(\hat{\mathbb{D}}_{32\bar{G}}+\frac{1-|\tilde{\upsilon}_2|^2}{2}\mathbb{I}_{2n}\right).
\end{aligned}$$

Next, we state the following basic fact: For a real symmetric matrix $\mathbb{A}$, and a real orthogonal matrix $\mathbb{Q}$ then $\operatorname{eig}(\mathbb{A}) = \operatorname{eig}\left(\mathbb{Q}\mathbb{A}\mathbb{Q}^T\right)$. Note that for orthogonal $\mathbb{Q}$ then $\mathbb{A}$ and $\mathbb{Q}\mathbb{A}\mathbb{Q}^T$ are called similar and the simple fact stated above is also referred to as "similarity preserves eigenvalues", see [33, pp. 508 and 549].

As $\mathbb{K}_{\bar{X}_{3\bar{G}}}$ is a real $2n\times 2n$ symmetric matrix, we can express $\mathbb{K}_{\bar{X}_{3\bar{G}}} = \mathbb{A}_{3\bar{G}}\mathbb{D}_{3\bar{G}}\mathbb{A}_{3\bar{G}}^T$, where $\mathbb{A}_{3\bar{G}}$ is an orthogonal matrix and $\mathbb{D}_{3\bar{G}}$ is a diagonal matrix, whose diagonal elements are in ascending order: $d_{3\bar{G},i} \triangleq \left[\mathbb{D}_{3\bar{G}}\right]_{i,i} \le \left[\mathbb{D}_{3\bar{G}}\right]_{k,k}$, $\forall 1 \le i \le k \le 2n$. Next, using similarity we write

$$\begin{aligned}
\hat{\mathbb{D}}_{32\bar{G}} &= \operatorname{eig}\left(\mathbb{h}_{32}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{32}^T\right)\\
&= \operatorname{eig}\left(\sqrt{\mathsf{SNR}_{32}}\mathbb{U}_{32}\mathbb{A}_{3\bar{G}}\mathbb{D}_{3\bar{G}}\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{32}^T\sqrt{\mathsf{SNR}_{32}}\right)\\
&= \mathsf{SNR}_{32}\mathbb{D}_{3\bar{G}}.
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbb{h}_{32}\mathbb{D}_{3\bar{G}}\mathbb{h}_{32}^T &= \sqrt{\mathsf{SNR}_{32}}\mathbb{U}_{32}\mathbb{D}_{3\bar{G}}\mathbb{U}_{32}^T\sqrt{\mathsf{SNR}_{32}}\\
&= \mathbb{U}_{32}\hat{\mathbb{D}}_{32\bar{G}}\mathbb{U}_{32}^T. \quad (E.6)
\end{aligned}$$

From (E.6) we obtain that

$$\frac{1}{2}\log\det\left(\mathbb{h}_{32}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{32}^T+\frac{1-|\tilde{v}_2|^2}{2}\mathbb{I}_{2n}\right) = \frac{1}{2}\log\det\left(\mathsf{SNR}_{32}\mathbb{D}_{3\bar{G}}+\frac{1-|\tilde{v}_2|^2}{2}\mathbb{I}_{2n}\right). \quad \text{(E.7)}$$

From (E.6) and (E.7) we conclude that given $\mathbb{K}_{\bar{X}_{3\bar{G}}}$, the same differential entropy can be achieved by replacing $\bar{X}_{3\bar{G}}^{2n}$ with an $\tilde{X}_{3\bar{G}}^{2n}\sim\mathcal{N}\left(\mathbf{0},\mathbb{D}_{3\bar{G}}\right)$. We also recall that from the power constraint $P_{k,i}\triangleq\mathbb{E}\{|X_{k,i}|^2\}\leq 1$, $k\in\{1,2,3\}$ we have $\mathrm{tr}\{\mathbb{D}_{3\bar{G}}\}=\mathrm{tr}\left\{\mathbb{K}_{\bar{X}_{3\bar{G}}}\right\}\leq n$.

Next, using matrix similarity we write

$$\begin{aligned}
&\mathrm{eig}\left(\hat{\mathbb{K}}+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\right)\\
&=\mathrm{eig}\left(\mathbb{h}_{31}\mathbb{K}_{\bar{X}_{3\bar{G}}}\mathbb{h}_{31}^T+\mathbb{h}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{h}_{11}^T+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\right)\\
&=\mathrm{eig}\bigg(\mathsf{SNR}_{31}\mathbb{U}_{31}\mathbb{A}_{3\bar{G}}\mathbb{D}_{3\bar{G}}\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{31}^T+\mathsf{SNR}_{11}\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\\
&\qquad+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\bigg) \qquad \text{(E.8)}\\
&=\mathrm{eig}\bigg(\mathbb{U}_{31}\mathbb{A}_{3\bar{G}}\bigg(\mathsf{SNR}_{31}\mathbb{D}_{3\bar{G}}\\
&\qquad+\mathsf{SNR}_{11}\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\mathbb{A}_{3\bar{G}}\\
&\qquad+\frac{|\eta_1|^2}{2}\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{31}^T\mathbb{I}_{2n}\mathbb{U}_{31}\mathbb{A}_{3\bar{G}}\bigg)\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{31}^T\bigg)\\
&=\mathrm{eig}\bigg(\mathsf{SNR}_{31}\mathbb{D}_{3\bar{G}}+\mathsf{SNR}_{11}\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\mathbb{A}_{3\bar{G}}\\
&\qquad+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\bigg). \qquad \text{(E.9)}
\end{aligned}$$

Let $\mathbf{a}_{3\bar{G},i}\triangleq\left[\mathbb{A}_{3\bar{G}}\right]_{(1\ldots 2n),i}$ denote that i'th column of $\mathbb{A}_{3\bar{G}}$. Applying Hadamard's inequality we obtain

$$\begin{aligned}
&\det\bigg(\mathsf{SNR}_{11}\mathbb{A}_{3\bar{G}}^T\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\mathbb{A}_{3\bar{G}}\\
&\qquad+\mathsf{SNR}_{31}\mathbb{D}_{3\bar{G}}+\frac{|\eta_1|^2}{2}\mathbb{I}_{2n}\bigg)\\
&\leq\prod_{i=1}^{2n}\bigg(\mathsf{SNR}_{11}\left(\mathbf{a}_{3\bar{G},i}\right)^T\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\mathbf{a}_{3\bar{G},i}\\
&\qquad+\mathsf{SNR}_{31}d_{3\bar{G},i}+\frac{|\eta_1|^2}{2}\bigg). \qquad \text{(E.10)}
\end{aligned}$$

Combining (E.7), (E.9), (E.10) and the fact that if matrices have the same eigenvalues their determinants are identical, we obtain the upper bound on (E.1), which is stated in (E.11), as shown at the top of the next page, where step (a) in the derivation leading to (E.11) is due to the fact that the logarithm function is a concave function and the application Jensen's inequality.

Next, we define $\mathbb{G}\triangleq\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}$, and compute the expectation $\bar{\mathbb{G}}=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\{\mathbb{G}\}=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\right\}$. To that aim, define

$$\begin{aligned}
\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RR}&=\mathbb{E}\left\{\left(\mathfrak{Re}\left\{X_{1\bar{G}}^n\right\}\left(\mathfrak{Re}\left\{X_{1\bar{G}}^n\right\}\right)^T\right)\right\},\\
\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}&=\mathbb{E}\left\{\left(\mathfrak{Re}\left\{X_{1\bar{G}}^n\right\}\left(\mathfrak{Im}\left\{X_{1\bar{G}}^n\right\}\right)^T\right)\right\},\\
\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}&=\mathbb{E}\left\{\left(\mathfrak{Im}\left\{X_{1\bar{G}}^n\right\}\left(\mathfrak{Re}\left\{X_{1\bar{G}}^n\right\}\right)^T\right)\right\},\\
\mathbb{K}_{\bar{X}_{1\bar{G}}}^{II}&=\mathbb{E}\left\{\left(\mathfrak{Im}\left\{X_{1\bar{G}}^n\right\}\left(\mathfrak{Im}\left\{X_{1\bar{G}}^n\right\}\right)^T\right)\right\},
\end{aligned}$$

and recall the definitions $\mathbb{U}_{31}=\begin{bmatrix}\mathbb{C}_{31}&-\mathbb{S}_{31}\\\mathbb{S}_{31}&\mathbb{C}_{31}\end{bmatrix}$ and $\mathbb{U}_{11}=\begin{bmatrix}\mathbb{C}_{11}&-\mathbb{S}_{11}\\\mathbb{S}_{11}&\mathbb{C}_{11}\end{bmatrix}$. Using these definitions we write the product as a block matrix: $\mathbb{G}=\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}=\begin{bmatrix}\mathbb{G}_{11}&\mathbb{G}_{12}\\\mathbb{G}_{21}&\mathbb{G}_{22}\end{bmatrix}$.

To compute $\bar{\mathbb{G}}_{12}\triangleq\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\{\mathbb{G}_{12}\}$ where averaging is carried out over the phase variables $\{\theta_{31,l},\theta_{11,l}\}_{l=1}^n$, we first write explicitly

$$\begin{aligned}
\bar{\mathbb{G}}_{12}=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{&-\mathbb{C}_{31}\mathbb{S}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\mathbb{C}_{31}\mathbb{S}_{11}-\mathbb{S}_{31}\mathbb{C}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\mathbb{S}_{31}\mathbb{C}_{11}\\
&+\mathbb{C}_{31}\mathbb{C}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\mathbb{C}_{31}\mathbb{C}_{11}+\mathbb{S}_{31}\mathbb{S}_{11}\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\mathbb{S}_{31}\mathbb{S}_{11}\Big\}.
\end{aligned}$$

Next, we note that the products of the matrices $\mathbb{C}_{lm}$ and $\mathbb{S}_{uv}$ are diagonal matrices. Thus, for $l\neq m$, $\left[\bar{\mathbb{G}}_{12}\right]_{l,m}$ correspond to averaging over phases from different times and are thus zero. It therefore remains to consider the diagonal elements of $\bar{\mathbb{G}}_{12}$:

$$\begin{aligned}
\left[\bar{\mathbb{G}}_{12}\right]_{l,l}&=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{-\cos\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\right]_{l,l}\cdot\cos\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\\
&\qquad-\sin\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\right]_{l,l}\cdot\sin\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\\
&\qquad+\cos\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\right]_{l,l}\cdot\cos\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\\
&\qquad+\sin\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\right]_{l,l}\cdot\sin\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\Big\}\\
&=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{-\cos^2\left(\theta_{31,l}\right)\sin^2\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\right]_{l,l}\\
&\qquad-\sin^2\left(\theta_{31,l}\right)\cos^2\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\right]_{l,l}\\
&\qquad+\cos^2\left(\theta_{31,l}\right)\cos^2\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\right]_{l,l}\\
&\qquad+\sin^2\left(\theta_{31,l}\right)\sin^2\left(\theta_{11,l}\right)\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\right]_{l,l}\Big\}\\
&=-\frac{1}{4}\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\right]_{l,l}-\frac{1}{4}\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{IR}\right]_{l,l}+\frac{1}{4}\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\right]_{l,l}\\
&\qquad+\frac{1}{4}\left[\mathbb{K}_{\bar{X}_{1\bar{G}}}^{RI}\right]_{l,l}.
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{h\left(\mathbb{H}_{h_{31}}^{(n)}X_{3\bar{G}}^n+\mathbb{H}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\big|\tilde{H}_1^n=\tilde{h}_1^n\right)\\
&\qquad\qquad -h\left(\mathbb{H}_{h_{32}}^{(n)}X_{3\bar{G}}^n+V_2^n\big|\tilde{H}_2^n=\tilde{h}_2^n\right)\Big\}\\
&\le \mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{\sum_{i=1}^{2n}\frac{1}{2}\left(\log\left(\mathsf{SNR}_{11}\left(\mathbf{a}_{3\bar{G},i}\right)^T\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\mathbf{a}_{3\bar{G},i}+\mathsf{SNR}_{31}d_{3\bar{G},i}+\frac{|\eta_1|^2}{2}\right)\right.\right.\\
&\qquad\qquad\left.\left.-\log\left(\mathsf{SNR}_{32}d_{3\bar{G},i}+\frac{1-|\tilde{v}_2|^2}{2}\right)\right)\right\}\\
&\overset{(a)}{\le}\sum_{i=1}^{2n}\left(\frac{1}{2}\log\left(\mathsf{SNR}_{11}\left(\mathbf{a}_{3\bar{G},i}\right)^T\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\right\}\mathbf{a}_{3\bar{G},i}+\mathsf{SNR}_{31}d_{3\bar{G},i}+\frac{|\eta_1|^2}{2}\right)\right.\\
&\qquad\qquad\left.-\frac{1}{2}\log\left(\mathsf{SNR}_{32}d_{3\bar{G},i}+\frac{1-|\tilde{v}_2|^2}{2}\right)\right)
\end{aligned}\tag{E.11}$$

Lastly, we note that $\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{IR}=\left(\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RI}\right)^T$, hence $\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{IR}\right]_{l,l}=\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RI}\right]_{l,l}$, and consequently $\left[\bar{\mathbb{G}}_{12}\right]_{l,l}=\mathbb{O}_{n\times n}$. We conclude that $\bar{\mathbb{G}}_{12}=\mathbb{O}_{n\times n}$. Similarly, we obtain $\bar{\mathbb{G}}_{21}\triangleq\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\{\mathbb{G}_{21}\}=\mathbb{O}_{n\times n}$. Next, we compute $\bar{\mathbb{G}}_{11}\triangleq\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\{\mathbb{G}_{11}\}$. Again, we note that unless a term has two identical matrices in the product, its expectation is necessarily zero:

$$\begin{aligned}
\bar{\mathbb{G}}_{11}=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{&\mathbb{C}_{31}\mathbb{C}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\mathbb{C}_{31}\mathbb{C}_{11}+\mathbb{S}_{31}\mathbb{S}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\mathbb{S}_{31}\mathbb{S}_{11}\\
&+\mathbb{C}_{31}\mathbb{S}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\mathbb{C}_{31}\mathbb{S}_{11}+\mathbb{S}_{31}\mathbb{C}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\mathbb{S}_{31}\mathbb{C}_{11}\Big\}.
\end{aligned}$$

Similarly to the computation of $\bar{\mathbb{G}}_{12}$, we note that the products of the matrices $\mathbb{C}_{lm}$ and $\mathbb{S}_{uv}$ are diagonal matrices. Thus, for $l\neq m$ $\left[\bar{\mathbb{G}}_{11}\right]_{l,m}$ correspond to averaging over phases from different times and are thus zero. Hence, it remains to consider the diagonal elements of $\bar{\mathbb{G}}_{11}$:

$$\begin{aligned}
\left[\bar{\mathbb{G}}_{11}\right]_{l,l}&=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{\cos\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\right]_{l,l}\cdot\cos\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\\
&\quad+\sin\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\right]_{l,l}\cdot\sin\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\\
&\quad+\cos\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\right]_{l,l}\cdot\cos\left(\theta_{31,l}\right)\sin\left(\theta_{11,l}\right)\\
&\quad+\sin\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\right]_{l,l}\cdot\sin\left(\theta_{31,l}\right)\cos\left(\theta_{11,l}\right)\Big\}\\
&=\frac{1}{4}\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\right]_{l,l}+\frac{1}{4}\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\right]_{l,l}+\frac{1}{4}\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\right]_{l,l}+\frac{1}{4}\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\right]_{l,l}\\
&=\frac{1}{2}\left(\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{RR}\right]_{l,l}+\left[\mathbb{K}_{\tilde{X}_{1\bar{G}}}^{II}\right]_{l,l}\right)\\
&\le\frac{1}{2},
\end{aligned}$$

where the last inequality follows from the per-symbol power constraint $P_{k,i}\triangleq\mathbb{E}\{|X_{k,i}|^2\}\le 1$, $k\in\{1,2,3\}$, and the condition on the optimal covariance matrix of Lemma 6. Following similar steps we obtain that $\bar{\mathbb{G}}_{22}$ is a diagonal matrix with non-negative elements, each smaller than $\frac{1}{2}$.

We conclude that $\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\right\}=\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{\begin{bmatrix}\mathbb{G}_{11}&\mathbb{G}_{12}\\\mathbb{G}_{21}&\mathbb{G}_{22}\end{bmatrix}\right\}=\bar{\mathbb{G}}$ is a diagonal matrix whose diagonal elements are all non-negative and less than $\frac{1}{2}$:

$$0\le\left[\bar{\mathbb{G}}\right]_{k,k}\triangleq\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{[\mathbb{G}]_{k,k}\right\}\le\frac{1}{2}.$$

Lastly, observe that as $\bar{\mathbb{G}}$ is a diagonal matrix, we can write

$$\begin{aligned}
\left(\mathbf{a}_{3\bar{G},i}\right)^T\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\left\{\mathbb{U}_{31}^T\mathbb{U}_{11}\mathbb{K}_{\tilde{X}_{1\bar{G}}}\mathbb{U}_{11}^T\mathbb{U}_{31}\right\}\mathbf{a}_{3\bar{G},i}&\equiv\left(\mathbf{a}_{3\bar{G},i}\right)^T\cdot\bar{\mathbb{G}}\cdot\mathbf{a}_{3\bar{G},i}\\
&=\sum_{k=1}^{2n}\left(\left[\mathbb{A}_{3\bar{G}}\right]_{k,i}\right)^2\bar{\mathbb{G}}_{k,k}.
\end{aligned}$$

Proceeding with the bounding, we define

$$d_{m,i}\triangleq\sum_{k=1}^{2n}\left(\left[\mathbb{A}_{3\bar{G}}\right]_{k,i}\right)^2\bar{\mathbb{G}}_{k,k}.$$

The facts that $0\le\left[\bar{\mathbb{G}}\right]_{k,k}\le\frac{1}{2}$ and that $\mathbb{A}_{3\bar{G}}$ is orthogonal, imply that

$$d_{m,i}\triangleq\sum_{k=1}^{2n}\left(\left[\mathbb{A}_{3\bar{G}}\right]_{k,i}\right)^2\bar{\mathbb{G}}_{k,k}\le\frac{1}{2}\sum_{k=1}^{2n}\left(\left[\mathbb{A}_{3\bar{G}}\right]_{k,i}\right)^2=\frac{1}{2},$$

and also that $d_{m,i}\ge 0$. Using the above definition of $d_{m,i}$ in (E.11), we can upper bound (E.1) as

$$\begin{aligned}
&\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Big\{h\left(\mathbb{H}_{h_{31}}^{(n)}X_{3\bar{G}}^n+\mathbb{H}_{h_{11}}^{(n)}X_{1\bar{G}}^n+\eta_1\cdot W_1^n\big|\tilde{H}_1^n=\tilde{h}_1^n\right)\\
&\qquad -h\left(\mathbb{H}_{h_{32}}^{(n)}X_{3\bar{G}}^n+V_2^n\big|\tilde{H}_2^n=\tilde{h}_2^n\right)\Big\}\\
&\le\sum_{i=1}^{2n}\left(\frac{1}{2}\log\left(\mathsf{SNR}_{11}d_{m,i}+\mathsf{SNR}_{31}d_{3\bar{G},i}+\frac{|\eta_1|^2}{2}\right)\right.\\
&\qquad\left.-\frac{1}{2}\log\left(\mathsf{SNR}_{32}d_{3\bar{G},i}+\frac{1-|\tilde{v}_2|^2}{2}\right)\right),
\end{aligned}$$

where $0 \leq d_{m,i} \leq \frac{1}{2}$ and $\sum_{i=1}^{2n} d_{3\bar{G},i} \equiv \text{tr}\{\mathbb{D}_{3\bar{G}}\} \leq n$, $d_{3\bar{G},i} \geq 0$.

We conclude that we can write the upper bound on (E.1) as follows:

$$\mathbb{E}_{\tilde{H}_1^n,\tilde{H}_2^n}\Bigg\{h\left(\mathbb{H}_{h_{31}}^{(n)} X_{3\bar{G}}^n + \mathbb{H}_{h_{11}}^{(n)} X_{1\bar{G}}^n + \eta_1 \cdot W_1^n \middle| \tilde{H}_1^n = \tilde{h}_1^n\right) - h\left(\mathbb{H}_{h_{32}}^{(n)} X_{3\bar{G}}^n + V_2^n \middle| \tilde{H}_2^n = \tilde{h}_2^n\right)\Bigg\}$$

$$\leq \max_{\left\{\substack{\{d_{3\bar{G},i}\}_{i=1}^{2n},\\ \{d_{m,i}\}_{i=1}^{2n}}\right\}} \sum_{i=1}^{2n} \frac{1}{2}\left(\log\left(\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \frac{|\eta_1|^2}{2}\right) - \log\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \frac{1-|\tilde{v}_2|^2}{2}\right)\right), \quad \text{(E.12)}$$

$$\text{subject to: } \sum_{i=1}^{2n} d_{3\bar{G},i} \leq n,$$

$$d_{3\bar{G},i} \geq 0, \quad 0 \leq d_{m,i} \leq \frac{1}{2}, \ i = 1, 2, \ldots, 2n.$$

Next, defining $\bar{N}_1 \triangleq \frac{|\eta_1|^2}{2}$ and $\bar{N}_2 \triangleq \frac{1-|\tilde{v}_2|^2}{2}$, we can write

$$\begin{aligned}&\log\left(\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right) - \log\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)\\ &= \log\left(\frac{\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1}{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}\right)\\ &= \frac{1}{\ln 2}\ln\left(\frac{\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1}{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}\right).\end{aligned}$$

The function $\frac{\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1}{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}$ is a linear fractional function in $(d_{m,i}, d_{3\bar{G},i})$ with positive denominator, and is thus a quasilinear function, see [26, Example 3.32]. Consequently, depending on $d_{m,i}$, it is either a monotone increasing function of $d_{3\bar{G},i}$ or a monotone decreasing function of $d_{3\bar{G},i}$. To find this threshold, we differentiate $\frac{\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1}{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}$ w.r.t $d_{3\bar{G},i}$:

$$\begin{aligned}&\frac{\partial}{\partial d_{3\bar{G},i}}\left\{\frac{\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1}{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}\right\}\\ &= \frac{\mathsf{SNR}_{31}\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)}{\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)^2} - \frac{\mathsf{SNR}_{32}\left(\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right)}{\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)^2}\\ &= \frac{\bar{N}_2 \cdot \mathsf{SNR}_{31} - \mathsf{SNR}_{32}\mathsf{SNR}_{11} d_{m,i} - \mathsf{SNR}_{32} \cdot \bar{N}_1}{\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)^2}.\end{aligned}$$

For $\frac{\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1}{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}$ to be monotone increasing in $d_{3\bar{G},i}$, the derivative must be positive. This occurs if

$$\bar{N}_2 \cdot \mathsf{SNR}_{31} - \mathsf{SNR}_{32}\mathsf{SNR}_{11} d_{m,i} - \mathsf{SNR}_{32} \cdot \bar{N}_1 > 0.$$

As the feasible $d_{m,i}$ must satisfy $d_{m,i} \leq \frac{1}{2}$, it directly follows from (19b) that indeed for all feasible $d_{m,i}$ it holds that

$$\mathsf{SNR}_{32}|\eta_1|^2 < \mathsf{SNR}_{31}\left(1 - |\tilde{v}_2|^2\right) - 2d_{m,i}\mathsf{SNR}_{32}\mathsf{SNR}_{11}. \quad \text{(E.13)}$$

Therefore, we conclude that for all feasible $d_{m,i}$, the objective increases with $d_{3\bar{G},i}$, hence, the objective is concave in the feasible region. It thus follows that there is a unique solution to the optimization problem (E.12) (see [26, Ch. 3.4]), and that this solution corresponds to the global maxima. Next, in order to maximize the upper bound (E.12), we define $\mathcal{N}_2 \triangleq \{1, 2, \ldots 2n\}$ and rewrite the problem as

$$\max_{\left\{\substack{\{d_{3\bar{G},i}\}_{i=1}^{2n},\\ \{d_{m,i}\}_{i=1}^{2n}}\right\}} \sum_{i=1}^{2n} \frac{1}{2}\left(\log\left(\mathsf{SNR}_{11} d_{m,i} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right) - \log\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)\right)$$

$$\text{subject to: } \sum_{i=1}^{2n} d_{3\bar{G},i} - n \leq 0,$$

$$-d_{3\bar{G},i} \leq 0, \ -d_{m,i} \leq 0, \ d_{m,i} - \frac{1}{2} \leq 0, \ i \in \mathcal{N}_2.$$

First, note that the objective is monotone increasing with $d_{m,i}$, hence it is maximized by letting $d_{m,i} = \frac{1}{2}, \ i \in \mathcal{N}_2$. Thus, the optimization problem can be written as

$$\max_{\{d_{3\bar{G},i}\}_{i=1}^{2n}} \sum_{i=1}^{2n} \frac{1}{2}\left(\log\left(\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right) - \log\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)\right)$$

$$\text{subject to: } \sum_{i=1}^{2n} d_{3\bar{G},i} - n \leq 0, \quad -d_{3\bar{G},i} \leq 0, \ i \in \mathcal{N}_2.$$

The Lagrangian for the above optimization function is

$$\begin{aligned}&L\left(\{d_{3\bar{G},i}, \varphi_{3,i}\}_{i=1}^{2n}, \psi_3\right)\\ &\triangleq \sum_{i=1}^{2n} \frac{1}{2}\left(\log\left(\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right) - \log\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)\right)\\ &\quad - \psi_3\left(\sum_{i=1}^{2n} d_{3\bar{G},i} - n\right) - \sum_{i=0}^{2n} \varphi_{1,i}(-d_{m,i}).\end{aligned}$$

To find the optimal solution, we first write the Karush-Kuhn-Tucker (KKT) necessary conditions for optimality [26, Ch. 5.5.3], and solve for the maximizing $\left(\{d_{3\bar{G},i}, \varphi_{3,i}\}_{i=1}^{2n}, \psi_3\right)$. As (E.13) guarantees concavity of the objective function, the optimal solution is the unique maximum. The KKT conditions for the above problem are:

$$\frac{\partial}{\partial d_{3\bar{G},i}} L\left(\left\{d_{3\bar{G},i}, \varphi_{3,i}, \psi_3\right\}_{i=1}^{2n}\right) = 0 \quad \text{(E.14a)}$$

$$\varphi_{3,i}\left(-d_{3\bar{G},i}\right) = 0, \quad \varphi_{3,i} \geq 0, \; i \in \mathcal{N}_2, \quad \text{(E.14b)}$$

$$\psi_3\left(\sum_{i=1}^{2n} d_{3\bar{G},i} - n\right) = 0, \quad \psi_3 \geq 0. \quad \text{(E.14c)}$$

Evaluating (E.14a) explicitly we obtain

$$\begin{aligned}
&\frac{\partial}{\partial d_{3\bar{G},i}} L\left(\left\{d_{3\bar{G},i}, \varphi_{3,i}\right\}_{i=1}^{2n}, \psi_3\right) \\
&= \frac{1}{\ln 2} \cdot \frac{\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2}{\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1} \\
&\quad \cdot \frac{\bar{N}_2 \cdot \mathsf{SNR}_{31} - \mathsf{SNR}_{32}\mathsf{SNR}_{11}\frac{1}{2} - \mathsf{SNR}_{32} \cdot \bar{N}_1}{\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)^2} \\
&\quad - \psi_3 + \varphi_{3,i} \\
&= \frac{1}{\ln 2} \cdot \frac{\bar{N}_2 \cdot \mathsf{SNR}_{31} - \mathsf{SNR}_{32}\mathsf{SNR}_{11}\frac{1}{2} - \mathsf{SNR}_{32} \cdot \bar{N}_1}{\left(\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right)\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)} \\
&\quad - \psi_3 + \varphi_{3,i} \\
&= 0.
\end{aligned}$$

As the objective is a monotone increasing function of $d_{3\bar{G},i}$, and the first derivative decreases as $d_{3\bar{G},i}$ increases, then $d_{3\bar{G},i} > 0$, and thus, from (E.14b) it follows that $\varphi_{3,i} = 0$.

From Eqn. (E.13), it follows that $\bar{N}_2 \cdot \mathsf{SNR}_{31} - \mathsf{SNR}_{32}\mathsf{SNR}_{11}\frac{1}{2} - \mathsf{SNR}_{32} \cdot \bar{N}_1 > 0$. With $\varphi_{3,i} = 0$ and $d_{3\bar{G},i} > 0$, $\forall i \in \mathcal{N}_2$, we obtain

$$\begin{aligned}
\psi_3 &= \frac{\bar{N}_2 \cdot \mathsf{SNR}_{31} - \mathsf{SNR}_{32}\mathsf{SNR}_{11}\frac{1}{2} - \mathsf{SNR}_{32} \cdot \bar{N}_1}{\ln 2\left(\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right)\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right)} \\
&> 0.
\end{aligned}$$

It remains to see what values of $d_{3\bar{G},i}$ can be considered. From the above equation we conclude that the denominator has to be a positive constant:

$$\left(\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31} d_{3\bar{G},i} + \bar{N}_1\right)\left(\mathsf{SNR}_{32} d_{3\bar{G},i} + \bar{N}_2\right) = C_0.$$

Hence, we obtain the quadratic equation for $d_{3\bar{G},i}$:

$$\begin{aligned}
&\mathsf{SNR}_{32} \cdot \mathsf{SNR}_{31} \cdot d_{3\bar{G},i}^2 \\
&\quad + \left(\mathsf{SNR}_{32} \cdot \bar{N}_1 + \frac{1}{2}\mathsf{SNR}_{32} \cdot \mathsf{SNR}_{11} + \mathsf{SNR}_{31} \cdot \bar{N}_2\right) d_{3\bar{G},i} \\
&\quad + \bar{N}_2\left(\frac{1}{2}\mathsf{SNR}_{11} + \bar{N}_1\right) - C_0 = 0,
\end{aligned}$$

from which we conclude that since the coefficient of $d_{3\bar{G},i}$ is positive, then the equation has at most one positive root. It thus follows that for $i \in \mathcal{N}_2$, $d_{3\bar{G},i} = d_3$ is a constant, and hence, from the sum condition $\sum_{i=1}^{2n} d_{3\bar{G},i} = n$ we conclude that $d_{3\bar{G},i} = d_3 = \frac{1}{2}, \quad i \in \mathcal{N}_2$. The maximal objective is therefore:

$$\begin{aligned}
&\sum_{i=1}^{2n} \frac{1}{2}\left(\log\left(\mathsf{SNR}_{11}\frac{1}{2} + \mathsf{SNR}_{31}\frac{1}{2} + \bar{N}_1\right)\right. \\
&\qquad\qquad \left. - \log\left(\mathsf{SNR}_{32}\frac{1}{2} + \bar{N}_2\right)\right) \\
&\quad = n\left(\log\left(\mathsf{SNR}_{11} + \mathsf{SNR}_{31} + |\eta_1|^2\right)\right. \\
&\qquad\qquad \left. - \log\left(\mathsf{SNR}_{32} + \left(1 - |\upsilon_2|^2\right)\right)\right). \quad \text{(E.15)}
\end{aligned}$$

We note that the maximum of the objective can be achieved with equality by letting $\bar{X}_{3\bar{G}}^{2n}$ and $\bar{X}_{1\bar{G}}^{2n}$ be mutually independent real Gaussian vectors with i.i.d. elements, each with zero mean and variance of $\frac{1}{2}$. With this assignment $\mathbb{K}_{\bar{X}_{1\bar{G}}} = \mathbb{K}_{\bar{X}_{3\bar{G}}} = \frac{1}{2}\mathbb{I}_{2n}$. The resulting complex variables are circularly symmetric complex Normal $X_{1\bar{G}}^n, X_{3\bar{G}}^n \sim \mathcal{CN}(\mathbf{0}, \mathbb{I}_n)$.

References

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math., Statist., Probab.*, vol. 1. 1961, pp. 611–644.

[2] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.

[3] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.

[4] 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.

[5] X. Shang, G. Kramer, and B. Chen, "A new outer bound and noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.

[6] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.

[7] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.

[8] D. Zahavi and R. Dabora, "Capacity theorems for the fading interference channel with a relay and feedback links," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5185–5213, Aug. 2012.

[9] D. Zahavi, L. Zhang, I. Maric, R. Dabora, A. J. Goldsmith, and S. Cui, "Diversity-multiplexing tradeoff for the interference channel with a relay," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 963–982, Feb. 2015.

[10] O. Sahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel," in *Proc. IEEE GLOBECOM Commun. Theory Symp.*, Washington, DC, USA, Nov. 2007, pp. 1627–1631.

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

[12] O. Sahin, E. Erkip, and O. Simeone, "Interference channel with a relay: Models, relaying strategies, bounds," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2009, pp. 90–95.

[13] I. Maric, R. Dabora, and A. Goldsmith, "Relaying in the presence of interference: Achievable rates, interference forwarding, and outer bounds," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4342–4354, Jul. 2012.

[14] R. Dabora, "The capacity region of the fading interference channel with a relay in the strong interference regime," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5172–5184, Aug. 2012.

[15] Y. Tian and A. Yener, "The Gaussian interference relay channel: Improved achievable rates and sum rate upperbounds using a potent relay," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2865–2879, May 2011.

[16] S. A. Jafar, "The ergodic capacity of phase-fading interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 7685–7694, Dec. 2011.

[17] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation, and full duplex operation," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2334–2344, May 2009.
[18] A. Chaaban and A. Sezgin, "On the generalized degrees of freedom of the Gaussian interference relay channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4432–4461, Jul. 2012.
[19] S. Gherekhloo, A. Chaaban, and A. Sezgin, "Cooperation for interference management: A GDoF perspective," *IEEE Trans. Inf. Theory*, vol. 32, no. 12, pp. 6986–7029, Dec. 2016.
[20] L. Zhou and W. Yu, "Incremental relaying for the Gaussian interference channel with a degraded broadcasting relay," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2794–2815, May 2013.
[21] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.
[22] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
[23] M. Katz and S. Shamai (Shitz), "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2257–2270, Oct. 2004.
[24] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.
[25] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[27] D. S. Bernstein, *Matrix Mathematics*. Prinston, NJ, USA: Prinston Univ. Press, 2009.
[28] D. Ouellette, "Schur complements and statistics," *Linear Algebra Appl.*, vol. 36, pp. 187–295, Mar. 1981.
[29] C. Pozrikidis, *An Introduction to Grids, Graphs, and Networks*. Oxford, U.K.: Oxford Univ. Press, 2014.
[30] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[31] T. M. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.
[32] J. M. Steele, *The Cauchy-Schwartz Master Class*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[33] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA, USA: SIAM, 2001.
[34] D. Zahavi and R. Dabora, "On cooperation and interference in the weak interference regime (full version with detailed proofs)." [Online]. Available: http://arxiv.org/abs/1701.08337

Daniel Zahavi (S'14) received the B.Sc. degree in 2009 from Technion, Israel Institute of Technology, Israel, and the M.Sc. and the Ph.D. degrees in 2012 and 2017, respectively, from Ben-Gurion University of the Negev, Israel. From 2010 to 2014 he was with the Signal Corps of Israel Defense Forces where he served as a researcher and later as a team leader in the fields of wireless communications and digital signal processing. His research interests include network information theory, and MIMO communications.

Ron Dabora (M'07–SM'14) received his BSc and MSc degrees in 1994 and 2000, respectively, from Tel-Aviv University, and his PhD degree in 2007 from Cornell University, all in Electrical Engineering. From 1994 to 2000, he worked as an engineer at the Ministry of Defense of Israel, and from 2000 to 2003, he was with the Algorithms Group at Millimetrix Broadband Networks, Israel. From 2007 to 2009, he was a postdoctoral researcher at the Department of Electrical Engineering, Stanford University. Since 2009, he is an assistant professor at the Department of Electrical and Computer Engineering, Ben-Gurion University, Israel. His research interests include network information theory, wireless communications, and power line communications. Dr. Dabora served as a TPC member in a number of international conferences including WCNC, PIMRC, and ICC. From 2012 to 2014, he served as an associate editor for the IEEE SIGNAL PROCESSING LETTERS, and he currently serves as a senior area editor for the IEEE SIGNAL PROCESSING LETTERS.