Design of 3-Way Relay Channels for Throughput and Energy Efficiency

Bho Matthiesen, Student Member, IEEE, Alessio Zappone, Member, IEEE, Eduard A. Jorswieck, Senior Member, IEEE,

Abstract

Throughput and energy efficiency in 3-way relay channels are studied in this paper. Unlike previous contributions, we consider a circular message exchange. First, an outer bound and achievable sum rate expressions for different relaying protocols are derived for 3-way relay channels. The sum capacity is characterized for certain SNR regimes. Next, leveraging the derived achievable sum rate expressions, cooperative and competitive maximization of the energy efficiency are considered. For the cooperative case, both low-complexity and globally optimal algorithms for joint power allocation at the users and at the relay are designed so as to maximize the system global energy efficiency. For the competitive case, a game theoretic approach is taken, and it is shown that the best response dynamics is guaranteed to converge to a Nash equilibrium. A power consumption model for mmWave board-to-board communications is developed, and numerical results are provided to corroborate and provide insight on the theoretical findings.

Index Terms

Multi-way networks, relay systems, energy efficiency, green communications, resource allocation, fractional programming, monotonic optimization, game theory, 5G networks, mmWave communications, power control.

I. INTRODUCTION

Relays are fundamental building blocks of wireless networks. Deploying relays in areas affected by significant shadowing such as tunnels or the inside of buildings, or in areas that are far away

Part of this work has been published in [1].

The work of Bho Matthiesen and Eduard Jorswieck is supported by the German Research Foundation (DFG) in the Collaborative Research Center 912 “Highly Adaptive Energy-Efficient Computing.” The work of Alessio Zappone has been funded by the German Research Foundation (DFG) project CEMRIN, under grant ZA 747/1-2.

The authors are with the Chair for Communications Theory, Communications Laboratory, Technische Universität Dresden, Dresden, Germany (e-mail: bho.matthiesen@tu-dresden.de, alessio.zappone@tu-dresden.de, eduard.jorswieck@tu-dresden.de).
from the transmitter, allows to extend cell-coverage and increase the network’s reliability and throughput. Thus, the study of relay channels is essential in understanding the capacity limits of modern networks and for the development of novel communication schemes. One recently-proposed channel model for relay networks is the multi-way relay channel (MWRC), which models clustered communication over a relay, where the terminals in each cluster exchange information among each other with the help of a relay. Such a model applies to many communication architectures like the communication of several ground stations over a satellite, or wireless board-to-board communication in highly adaptive computing [2] where multiple chips exchange data with the help of another chip acting as relay. The MWRC was first introduced in [3], where all users in the cluster send a message and are interested in decoding the messages of all other users in the cluster. In [4] the common-rate capacity of the additive white Gaussian noise (AWGN) MWRC with full message exchange is derived and it is shown that for three and more users this capacity is achieved by decode-and-forward (DF) for signal-to-noise ratios (SNRs) below 0 dB and compute-and-forward otherwise. The same authors present in [5] the capacity region of the finite field MWRC. In [6] a constant gap approximation of the capacity region of the Gaussian 3-user MWRC with full message exchange is derived. In contrast to most other works, the authors consider private messages instead of common messages, i.e., each user transmits distinct messages to the other users instead of a common one.

Besides throughput, another key performance metric in modern and future 5G wireless networks is energy efficiency (EE), which is steadily gaining momentum due to green and sustainable growth concerns. The need for EE is even stronger for battery-powered terminals, in order to extend their lifetimes. From a mathematical standpoint, one well-established definition of the EE of a communication system is the ratio between the system capacity or achievable rate and the total consumed power [7], [8]. With this definition, the EE is measured in bit/Joule (or bit/Hz/Joule when the rate is given in bit/s/Hz), thus naturally representing the efficiency with which the available energy is used to transmit information. Previous results on EE in relay systems mainly focus on one-way amplify-and-forward (AF) or DF schemes and do not consider the MWRC. In [9] the optimal placement of relays in cellular networks is investigated and is seen to provide power-saving gains. In [10] instead, cooperative approaches aimed at weighted sum-power minimization with fairness and rate constraints are devised. In [11] a pricing-based approach is employed to come up with energy-saving power control algorithms. There, the EE
is defined as the difference of the achievable rate and the transmit power scaled by a price. Instead, [12] considers the bit/Joule definition of EE and devises energy-efficient power control algorithms in interference networks. A cooperative approach is considered in [13], where a multiple-input multiple-output (MIMO) AF relay-assisted system is considered and the source and relay precoding matrices are allocated so as to maximize the global EE, for different channel state information (CSI) assumptions [14]. Power control in one-way and two-way relay channels with AF is studied in [15].

In this paper a 3-way relay channel is considered and both throughput and energy efficiency are analyzed and optimized. In contrast to most other works on MWRCs, we focus on a partial message exchange where each message is only intended for one receiver and, also, not every user sends a message to each other user. This might grant higher achievable rates provided one can deal with interference at the receivers which makes the analysis more involved.

In such a scenario, we make the following contributions:

1) We derive achievable sum rate expressions for the AF, DF, noisy network coding (NNC) relaying protocols, and an outer bound on the sum capacity, showing that the sum capacity is achieved by DF in the low SNR regime, and providing finite gap results for completely symmetric channels.

2) Based on the derived achievable sum rate expressions, cooperative power control algorithms to maximize the system global energy efficiency (GEE) are provided. If DF is used at the relay, the global optimum of the system GEE is found in polynomial time by means of fractional programming theory. For all other relaying schemes, leveraging fractional programming and monotonic optimization theory, two algorithms are provided to compute both a local solution with low complexity and the global optimum with higher complexity. Numerical evidence suggests that both algorithms actually achieve the global optimum.

3) Energy-efficient power control is also investigated in a competitive scenario formulating the problem as a non-cooperative game. The existence of a Nash equilibrium (NE) is shown and a best response dynamics (BRD)-based algorithm is provided which is guaranteed to always converge to a NE point.

4) The performance of the proposed algorithms and of the different relaying schemes are thoroughly compared by numerical simulations. A realistic power consumption model for a mmWave board-to-board communication system operating at 200 GHz is developed. It is
shown that, although not achieving the highest sum rate, AF relaying is more energy-efficient than all other schemes due to its low hardware complexity.

We define the function $C(x) = \log_2(1 + x)$ for $x \geq 0$. $f'(x)$ and $f''(x)$ denote the first and second derivative of $f$ as functions of $x$, respectively. $\log(\cdot)$ denotes natural logarithms.

II. SYSTEM MODEL

We consider the symmetric 3-user single-input single-output (SISO) MWRC with circular (i.e., partial) message exchange, AWGN, no direct user-to-user links, and full-duplex transmission. The users are denoted as node 1 to 3 and the relay is node $R$. We define the set of all users as $\mathcal{K} = \{1, 2, 3\}$ and the set of all nodes as $\mathcal{K}_R = \mathcal{K} \cup \{R\}$.

The 3-user MWRC consists of an uplink channel $Y_R = \sum_{k \in \mathcal{K}} X_k + Z_R$, and downlink channels $Y_k = X_k + Z_k$, $k \in \mathcal{K}$, where $X_k$ and $Y_k$ are the complex valued channel input and output at node $k \in \mathcal{K}_R$, respectively, and $Z_k$ is zero mean circularly symmetric complex Gaussian noise with power $N_R$ at the relay and $N_S$ at all other nodes.

All noise variables are mutually independent and the channel inputs are independent and identically distributed (i.i.d.) over time. All channel inputs have zero mean and an average power constraint $E|X_k|^2 \leq P_R$ and $E|X_k|^2 \leq P_S$, for $k \in \mathcal{K}$.

We consider a circular message exchange as illustrated in Fig. 1 where user $q(k)$ wants message $m_k$ with $q = [2, 3, 1]$. We also define $l(k)$ as the index of the interfering (i.e., unwanted) message at user $q(k)$ as $l = [3, 1, 2]$.

A $(2^{nR_1}, 2^{nR_2}, 2^{nR_3}, n)$ code for the 3-user MWRC consists of three message sets $\mathcal{M}_k = [1 : 2^{nR_k}]$, one for each user $k \in \mathcal{K}$, three encoders, where encoder $k \in \mathcal{K}$ assigns a symbol $x_{ki}(m_k, y_{ki}^{i-1})$ to each message $m_k \in \mathcal{M}_k$ and received sequence $y_{ki}^{i-1}$ for $i \in [1 : n]$, a relay encoder that assigns a symbol $x_{0i}(y_{0i}^{i-1})$ to every past received sequence $y_{0i}^{i-1}$ for $i \in [1 : n]$, and
three decoders, where decoder $q(k) \in \mathcal{K}$ assigns an estimate $\hat{m}_k \in \mathcal{M}_k$ or an error message $e$ to each pair $(m_{q(k)}, y_{q(k)})$.

We assume that the message triple $(M_1, M_2, M_3)$ is uniformly distributed over $\mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3$. The average probability of error is defined as $P_e^{(n)} = \Pr\{\hat{M}_k \neq M_k \text{ for some } k \in \mathcal{K}\}$.

A rate triple $(R_1, R_2, R_3)$ is said to be achievable if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, 2^{nR_3}, n)$ codes such that $\lim_{n \to \infty} P_e^{(n)} = 0$. The capacity region of the 3-user MWRC is the closure of the set of achievable rates. The sum rate is defined as $R_\Sigma = \max \{R_1 + R_2 + R_3 : (R_1, R_2, R_3) \in \mathcal{R}\}$, where $\mathcal{R}$ is an achievable rate region. Whenever $\mathcal{R}$ is the capacity region, we call $R_\Sigma$ the sum capacity $C_\Sigma$.

III. Bounds on the Sum Capacity

We start our treatment of the symmetric 3-user MWRC by deriving an upper bound on the sum capacity and then continue with several inner bounds.

A. Outer Bound

This outer bound consists of the cut set bound in the uplink and a downlink bound [16] that takes the side information at the receivers into account.

Lemma 1: The sum capacity of the symmetric 3-user MWRC is upper bounded as

$$C_\Sigma \leq \min\left\{\frac{3}{2} C\left(\frac{P_R}{N_S}\right), 3 C\left(\frac{P_S}{N_R}\right)\right\}.$$ (1)

Proof: Assume the relay knows the messages $M_1$, $M_2$, and $M_3$ a priori. Then we are looking at a broadcast channel (BC) with receiver message side information.

From [16, Theorem 4] we have that every achievable rate tuple $(R_1, R_2, R_3)$ must satisfy $R_i + R_j \leq C\left(\frac{P_R}{N_S}\right)$, for $i \neq j$ and $i, j \in \mathcal{K}$. With the simplex algorithm (see, e.g., [17]) we can show that $R_\Sigma \leq \frac{3}{2} C\left(\frac{P_R}{N_S}\right)$.

Applying the cut set bound [18] for the cuts $\mathcal{S} = \{k\}, k \in \mathcal{K}$, results in $R_k \leq C\left(\frac{P_e}{N_R}\right)$, for $k \in \mathcal{K}$, and, thus, $R_\Sigma \leq 3 C\left(\frac{P_e}{N_R}\right)$. $\blacksquare$

B. Amplify-and-Forward

We first consider AF relaying where the relay scales the observed signal by a positive constant and broadcasts it back to the users. The transmitted symbol at the relay is $X_R = \alpha Y_R$, where
α is a normalization factor chosen such that the transmit power constraint at the relay is met, i.e., \( \alpha = \sqrt{P_R' / (\sum_{k \in K} P_k' + N_R)} \), where \( P_k' \), \( k \in K_R \), is the actual transmit power of node \( k \) satisfying the average power constraints. The receiver first removes its self-interference from the received signal and then decodes for its desired message while treating the remaining interference as noise. We split the transmission into three equal length blocks and switch off user \( i \) in time slot \( i \). This reduces interference and allows for higher transmission powers in the other two time slots while still meeting the average power constraint.

**Lemma 2:** In the 3-user MWRC, the sum rate

\[
R_{\text{AF-IAN}} = C \left( \frac{3P_S P_R}{N_R P_R + 3P_S N_S + N_S N_R} \right)
\]

is achievable with AF relaying and treating interference as noise (IAN) at the receivers.

**Proof:** The received signal of user \( q(i) \) is \( Y_{q(i)} = \sum_{k \in K} \alpha X_k + \alpha Z_R + Z_{q(i)} \). Since \( X_{q(i)} \) is known at receiver \( q(i) \), it can be removed from the received signal. Thus, \( \hat{Y}_{q(i)} = Y_{q(i)} - \alpha X_{q(i)} = \alpha X_i + \alpha X_{l(i)} + \alpha Z_R + Z_{q(i)} \). Treating the interfering symbol \( X_{l(i)} \) as additional noise, the maximum achievable rate \( R_i \) is \( C \left( \frac{\alpha^2 P'_l}{\alpha^2 P'_l + \alpha^2 N_R + N_S} \right) \). Plugging in the expression for \( \alpha \) we have

\[
R_i \leq C \left( \frac{P'_R P'_l}{P'_R P'_l + P'_R N_R + (\sum_{k \in K} P_k' + N_R) N_S} \right),
\]

for all \( i \in K \).

It is easily shown that the right-hand side (RHS) of (3) is monotonically increasing in \( P'_R \). Thus, \( P'_R = P_r \) is optimal.

Next, we split the transmission into three equal length blocks, and switch off user \( i \) in time slot \( i \), i.e., \( P'_i = 0 \). Then, the achievable sum rate in time slot \( i \) is

\[
\frac{1}{3} \left( \frac{P'_R P'_{q(i)}}{P_R N_R + (P'_{q(i)} + P'_{l(i)} + N_R) N_S} \right) + \frac{1}{3} \left( \frac{P'_R P'_{l(i)}}{P_R P'_{q(i)} + P_R N_R + (P'_{q(i)} + P'_{l(i)} + N_R) N_S} \right).
\]

Since each user transmits only in two out of three time slots the average power constraint is met as long as \( \sum_{j \in K \setminus \{i\}} P'^j_i \leq 3P_S \). Because (4) is monotonically increasing in \( P'_{q(i)} \) and \( P'_{l(i)} \), this constraint should be met with equality.

\(^1\)Notation: \( P'_k \) is the transmit power used by user \( k \) in time slot \( i \).
We set $P_{1}^{\prime} = 3PS - p$, $P_{3}^{\prime} = p$, $P_{2}^{3} = 3PS - p$, $P_{2}^{l} = p$, $P_{3}^{l} = 3PS - p$, $P_{3}^{2} = p$, for some $p \in [0,3PS]$ in (4) and get

$$\frac{1}{3} C \left( \frac{3PSR}{NRPR + NS(3PS + NR)} \right)$$

as the achievable sum rate in time slot $i$. Summing over all time slots results in (2).

**Proposition 1:** The transmission scheme in Lemma 2 achieves strictly higher sum rates than using maximum transmit power at all nodes for AF-IAN.

**Proof:** Using maximum transmit power in (3), the achievable sum rate with AF-IAN is

$$3 C \left( \frac{PRPS}{PRPS + PRNS + 3PSNS + NRSN} \right).$$

This rate is always less than (2) if

$$(3PSPR + NRRP + 3SNS + NSNR) - (2SPPR + NRRP + 3SNS + NSNR)(NRRP + 3SNS + NSNR) \geq 0$$

Expanding the left-hand side (LHS), we have $P_{3}^{R}P_{3}^{3}R(2NSNR + 2NRRP + 3PRPS + 6PSNS) \geq 0$, which is easily seen to hold.

**Remark 1:** Since the time division scheme allows for higher transmit powers, this approach also outperforms AF-IAN schemes where one user is turned off during the whole transmission.

**Remark 2:** Since (5) is not dependent on $p$, Lemma 2 includes TDMA as a special case (with $p = 0$ or $p = 3PS$).

While AF-IAN offers low relay and decoder complexity, simultaneous non-unique decoding (SND) [19] may be employed at the receivers to achieve higher transmission rates. In SND, the interfering message is jointly decoded with the desired message instead of treating it as additional noise.

**Lemma 3:** In the 3-user MWRC, the sum rate

$$R_{\Sigma}^{AF-SND} = \frac{3}{2} C \left( \frac{2PSPR}{NRPR + 3PSNS + NSNR} \right)$$

is achievable with AF relaying and SND at the receivers.

**Proof:** Using the approach from [20] and standard techniques from network information theory [19], we can show that a rate tuples $(R_{1}, R_{2}, R_{3})$ is achievable in the symmetric MWRC
with AF relaying and SND at the receivers if all rates are non-negative and satisfy
\[
\sum_{i \in S} R_i \leq C \left( \frac{P_R \sum_{i \in S} P'_i}{N_R P_R + N_S \left( N_R + \sum_{j \in K} P'_j \right)} \right),
\]
for all \( S \in \{ \{k\}, \{k, l(k)\} : k \in K \} \). Setting \( P'_k = P_S, k \in K \), and applying the simplex algorithm [17], it is straightforward to show (7).

C. Decode-and-Forward

In DF relaying, the relay completely decodes the messages of each user and then broadcasts them back to all users. The achievable rate region is the intersection of the capacity regions of the 3-user multiple-access channel (MAC) and the BC with receiver side information and partial decoding at the receivers.

**Lemma 4:** In the 3-user MWRC, the sum rate
\[
R^D_{\Sigma} = \min \left\{ \frac{3}{2} C \left( \frac{P_R}{N_S} \right), \ C \left( \frac{3P_S}{N_R} \right) \right\}
\]
is achievable with DF relaying.

**Proof:** The achievable rate region is given in [3, Proposition 2].

Using the simplex algorithm (see, e.g., [17]) we can show that
\[
R^D_{\Sigma} = \begin{cases} 
\frac{3}{2} D & \text{if } D < A \text{ and } C > \frac{3}{2} D, \text{ or } A < D < B \text{ and } C > \frac{3}{2} D \\
C & \text{if } D < A \text{ and } C < \frac{3}{2} D, \text{ or } A < D < B \text{ and } C < \frac{3}{2} D, \text{ or } B < D
\end{cases}
\]
where \( A = C \left( \frac{P_S}{N_R} \right), B = C \left( \frac{2P_s}{N_R} \right), C = C \left( \frac{3P_S}{N_R} \right), \text{ and } D = C \left( \frac{P_R}{N_S} \right). \)

We observe that \( B < D \Leftrightarrow 2P_S < P_R \). Thus, it holds that \( (1 + 3 \frac{P_S}{N_R})^2 < \left( 1 + \frac{3P_R}{2N_S} \right)^2 < (1 + \frac{P_R}{N_S})^3 \) which is equivalent to \( C < \frac{3}{2} D \). Hence, with \( B < D \Rightarrow C < \frac{3}{2} D \) it can be seen from (9) that \( R^D_{\Sigma} = \min \left\{ \frac{3}{2} D, C \right\}. \)

**Remark 3:** The result from [3] implements a full message exchange. However, from the proof of Lemma 1 it can be seen that in the symmetric case the relaxed decoding requirements due to the partial message exchange considered here can not result in higher rates for DF.
D. Noisy Network Coding

NNC [21] generalizes the compress-and-forward (CF) coding scheme for the relay channel [22] and network coding for graphical networks [23] to discrete memoryless networks (DMNs).

Lemma 5: In the 3-user MWRC, the sum rate

\[ R_{\Sigma}^{NNC} = \frac{3}{2} C \left( \frac{2P_S P_R}{N_R P_R + 2P_S N_S + N_S N_R} \right) \]  

is achievable with NNC and SND.

Proof: Use [21, Theorem 2] and identify \( D_0 = \emptyset \) and \( D_k = \{ q(k) \} \) for \( k \in \mathcal{K} \). Assume \( \hat{Y}_i = Y_i + \hat{Z}_i \) with \( \hat{Z}_i \sim CN(0, Q_i) \) for \( i \in \mathcal{K}_R \), and \( Q = \emptyset \), i.e., no time-sharing is used. Then, the achievable rate region is

\[ R_k < C \left( \frac{P_S}{N_R + Q_0} \right), \quad \sum_{i \in \mathcal{K} \setminus \{ k \}} R_i < \min \left\{ C \left( \frac{2P_S}{N_R + Q_0} \right), \ C \left( \frac{P_R}{N_S} \right) - C \left( \frac{N_R}{Q_0} \right) \right\}, \]

for each \( k \in \mathcal{K} \).

Using the simplex algorithm [17], it is obtained that the sum rate is \( R_{\Sigma} = \frac{3}{2} C \left( \frac{2P_S}{N_R + Q_0} \right) \) if \( Q_0 > \frac{N_S (N_R + 2P_S)}{P_R} \), and \( R_{\Sigma} = \frac{3}{2} C \left( \frac{P_R}{N_S} \right) - C \left( \frac{N_R}{Q_0} \right) \), otherwise. Due to monotoncity in \( Q_0 \), the optimal choice for \( Q_0 \) is \( \frac{N_S (N_R + 2P_S)}{P_R} \). Evaluating \( R_{\Sigma} \) at this point results in (10).

E. Comparisons and Insights

In this subsection, we use the results from above to state some analytical results on the sum capacity and compare achievable sum rates of the considered relaying schemes.

Theorem 1: If \( \frac{P_n}{N_S} \leq \left( 1 + 3 \frac{P_S}{N_R} \right)^{2/3} - 1 \), the sum capacity of the symmetric 3-user MWRC is \( C_{\Sigma} = \frac{3}{2} C \left( \frac{P_n}{N_S} \right) \) and is achieved using the DF strategy from Lemma 4.

Proof: Equation (8) is equal to (1) if in both equations the first term of the minimum is active. Since \( C \left( \frac{3P_S}{N_R} \right) \leq 3 C \left( \frac{P_S}{N_R} \right) \), this is the case as long as \( \frac{3}{2} C \left( \frac{P_R}{N_S} \right) \leq C \left( \frac{3P_S}{N_R} \right) \). This condition is equivalent to the condition in Theorem 1.

Remark 4: For the completely symmetric channel where \( P = P_S = P_R \) and \( N = N_S = N_R \), the condition in Theorem 1 is equivalent to \( \frac{P}{N} \leq 3 + 2\sqrt{3} \approx 8.1 \) dB.

Theorem 2: The sum capacity of the completely symmetric 3-user MWRC is achieved within 0.877 bit by NNC and within 1.5 bit by AF-SND.
Proof: Let $S = \frac{P_R}{N_S} = \frac{P_S}{N_R}$. The achievable sum rate for NNC and AF-SND is $C\left(\frac{2S^2}{1+(1+a)S}\right)$ with $a = 2$ for NNC and $a = 3$ for AF. Its difference to the outer bound is

$$\Delta = \frac{3}{2} C(S) - \frac{3}{2} C\left(\frac{2S^2}{1+(1+a)S}\right) = \frac{3}{2} \log\left(\frac{1 + (2 + a)S + (1 + a)S^2}{1 + (1 + a)S + 2S^2}\right).$$

From the first derivative of $\Delta$ it can be seen that this is an increasing function in $S$ for the relevant choices of $a$. Thus, its maximum is attained at $S \to \infty$ and has value $\lim_{S \to \infty} \Delta = \frac{3}{2} \log\left(\frac{1+a}{2}\right)$.

**Corollary 1:** The sum degrees of freedom (DoF) of the 3-user MWRC with circular message exchange is 1.5.

**Proof:** The sum DoF $d_\Sigma$ is defined as $\lim_{P \to \infty} \frac{C_\Sigma(P)}{\log(P)}$ where $C_\Sigma(P)$ is the sum capacity of the completely symmetric MWRC. From the proof of Theorem 2, for the outer and inner bounds we respectively have

$$d_{\Sigma}^{\text{out}} \leq \lim_{S \to \infty} 3 \log\left(\frac{1+S}{S}\right) = \frac{3}{2} \quad \text{and} \quad d_{\Sigma}^{\text{in}} \geq \lim_{S \to \infty} \frac{3 \log\left(\frac{1 + \frac{2S^2}{1+(1+a)S}}{S}\right)}{\log(S)} = \frac{3}{2}.$$

Since both bounds coincide, $d_\Sigma = \frac{3}{2}$. 

**Remark 5:** It is easily shown that DF and AF-IAN achieve a sum DoF of 1.

**Proposition 2:** It holds that $R_{\Sigma}^{\text{NNC}} \geq R_{\Sigma}^{\text{AF-SND}} \geq R_{\Sigma}^{\text{AF-IAN}}$.

**Proof:** The first inequality follows directly from comparing (7) and (10). For the second, note that $R_{\Sigma}^{\text{AF-SND}} \geq R_{\Sigma}^{\text{AF-IAN}}$ is equivalent to

$$\left(1 + \frac{2P_SP_R}{N_RP_R + 3P_SN_S + N_SN_R}\right)^3 \geq \left(1 + \frac{3P_SP_R}{N_RP_R + 3P_SN_S + N_SN_R}\right)^2.$$

By expanding both sides, it is also easily seen that this inequality holds.

**Remark 6:** DF does not fit in this order since it outperforms NNC only in certain SNR regimes. Even for the completely symmetric case, the intersection point between DF and NNC has to be determined numerically. In this case it is approximately 14.27 dB.

**IV. RESOURCE ALLOCATION FOR ENERGY EFFICIENCY MAXIMIZATION**

From a physical standpoint, the efficiency in the use of a given resource is the benefit-cost ratio, i.e. the benefit obtained from using the resource, divided by the cost associated to the use of the resource. Applying this general concept to the use of energy to transmit data over
a communication link, leads to defining the EE as the amount of data that can be reliably transmitted in a given time interval $T$, divided by the resulting energy consumption over the same time interval.

In this section, we will deal with the problem of allocating the transmit powers at the users and relay for EE maximization. Both cooperative and competitive resource allocation will be considered.

A. Cooperative resource allocation

From a system-wide perspective, the benefit and cost related to the use of energy over a time interval $T$ are identified as the amount of data that can be reliably transmitted in the whole network during $T$, and the total energy consumption in the network. This leads to defining the system GEE as [8], [24]

$$\text{GEE} = \frac{TBR_{\Sigma}}{R_{\Sigma}} = \frac{BR_{\Sigma}}{\phi P_s + \psi P_r + P_c}. \quad (11)$$

In (11), $B$ is the communication bandwidth and $R_{\Sigma}$ is the achievable sum rate in bit/s/Hz, which depends on the particular relay protocol and receive scheme. Therefore, the numerator $TBR_{\Sigma}$ is the total amount of data that can be reliably transmitted during $T$. On the other hand, the associated total energy consumption is given by the sum of the energies consumed for signal transmission at the users and relay plus the circuit power that is dissipated in all terminals to operate the devices. The first component is expressed as $\phi P_s + \psi P_r$, with $\phi$ and $\psi$ being the inefficiencies of the users and relay transmit amplifiers,\(^2\) while the second component is modeled by the constant term $P_c$.

In the cooperative scenario, the users and the relay cooperate and jointly allocate $P_s$ and $P_r$ to maximize the common performance metric (11). Given the achievable sum rate expressions from Section III, (11) takes two different functional forms. For the outer bound and for the DF case we have $\text{GEE}_1$, and for the AF and NNC cases we have $\text{GEE}_2$ as defined next:

$$\text{GEE}_1 = \frac{a_1 C \left( \frac{P_s}{N_s} \right)}{\phi P_s + \psi P_r + P_c}, \quad \text{GEE}_2 = \alpha C \left( \frac{P_s P_r}{aP_s + bP_r + c} \right),$$

$$\quad \phi P_s + \psi P_r + P_c. \quad (12)$$

\(^2\)Specifically, $\psi \geq 1$ is the inefficiency of the relay amplifier, while $\phi \geq 3$, accounts for the inefficiency of the amplifiers of the three users, assumed equal for notational convenience
with $a_1$, $a_2$, $\alpha_1$, $\alpha_2$, $\alpha$, $a$, $b$, and $c$ non-negative parameters.

Being fractional functions, the considered GEEs are in general non-concave functions and conventional convex programming tools can not be used. Instead, fractional programming provides a framework to optimize fractional functions. In particular, we will exploit the following result from fractional programming theory.

**Proposition 3:** Let $S \subseteq \mathbb{R}^n$, $f, g : S \rightarrow \mathbb{R}$, with $f(x) \geq 0$ and $g(x) > 0$. Solving the problem
\[
\max_{x \in S} \frac{f(x)}{g(x)}
\]
is equivalent to finding the unique zero of the function $F(\lambda) = \max_{x \in S} (f(x) - \lambda g(x))$.

**Proof:** See [25], [26].

Proposition 3 provides a way to maximize a fractional function by finding the zero of the auxiliary function $F(\lambda)$. This can be accomplished by means of Dinkelbach’s algorithm [26], which is reported in Algorithm 1.

**Algorithm 1 Dinkelbach’s Algorithm**

1: Initialize $\lambda_0$, such that $F(\lambda_0) \geq 0$. Set $n = 0$.
2: while $F(\lambda_n) > \epsilon$ do
3: \[ x^*_n \leftarrow \arg \max_{x \in S} f(x) - \lambda_n g(x) \] (13)
4: \[ \lambda_{n+1} \leftarrow \frac{f(x_n^*)}{g(x_n^*)} \]
5: \[ n \leftarrow n + 1 \]
6: end while
7: Output $(x_n^*, \lambda_n)$

Dinkelbach’s algorithm exhibits super-linear convergence and only requires the solution of a sequence of convex problems, provided $f(x)$ and $g(x)$ are concave and convex, respectively, and that $S$ is a convex set. Indeed, in this case the subproblem to be solved in each iteration to find $x^*_n$ is convex. Moreover, we stress that Dinkelbach’s algorithm converges to the global solution.
of the associated fractional problem also when $f(x)$ is not concave and/or $g(x)$ is not convex, even though in this case a non-convex problem must be globally solved in each iteration.

1) Maximization of $\text{GEE}_1$: The optimization problem is formulated as

$$\max_{P_S, P_R} \min \left\{ a_1 C \left( \frac{P_R}{N_S} \right), a_2 C \left( \frac{P_S}{N_R} \right) \right\} \frac{\phi P_S + \psi P_R + P_c}{P_S \in [0; P_{S_{\text{max}}}], \quad P_R \in [0; P_{R_{\text{max}}}]}. \tag{14}$$

Problem (14) is an instance of a non-concave and non-smooth fractional problem. However, it can be efficiently solved by means of Dinkelbach’s algorithm because the minimum of concave functions is concave. Therefore, the objective of (14) has a concave numerator and an affine denominator. Moreover, it is possible to reformulate (14) into a smooth problem by introducing the auxiliary variable $t$ as follows.

$$\max_{t, P_S, P_R} \frac{t}{\phi P_S + \psi P_R + P_c} \left\{ t \geq 0, \quad a_1 C \left( \frac{P_R}{N_S} \right) - t \geq 0, \quad a_2 C \left( \frac{P_S}{N_R} \right) - t \geq 0. \tag{15} \right\}$$

In (15) the numerator and denominator of the objective are both linear, while the constraints are convex. Then, (15) can be solved by means of Dinkelbach’s algorithm with an affordable complexity.

2) Maximization of $\text{GEE}_2$: In this case, the optimization problem is formulated as

$$\max_{P_S, P_R} \frac{\alpha C \left( \frac{P_S P_R}{a P_S + b P_R + c} \right)}{\phi P_S + \psi P_R + P_c} \left\{ t \geq 0, \quad a_1 C \left( \frac{P_R}{N_S} \right) - t \geq 0, \quad a_2 C \left( \frac{P_S}{N_R} \right) - t \geq 0. \tag{16} \right\}$$

Problem (16) is more challenging than Problem (14) because the numerator of the objective function is not jointly concave in the optimization variables. Therefore, directly applying Dinkelbach’s algorithm requires solving a sequence of non-convex problem, which is computationally demanding. In general, no optimization tool is available to maximize a fractional function with non-concave numerator with limited complexity [25]. As a consequence, finding the global solution of (16) with limited complexity appears difficult. In the following, we propose two approaches to solve (16). The first approach will be based on the alternating maximization
algorithm [27] and will find a local optimum of (16) with a limited computational complexity. Instead, the second approach will leverage the theory of monotonic optimization, determining the global solution of (16) with an exponential complexity. Interestingly, numerical evidence to be provided in Section VI will show that the two algorithms achieve virtually equal performance.

a) *Alternating maximization approach:* This approach is based on the observation that the numerator of the objective, although not being jointly concave in $P_S$ and $P_R$, is separately concave in the two variables. To be more specific, the numerator of the objective is concave in $P_S$ for fixed $P_R$ and vice versa. Therefore, if we fix either $P_S$ or $P_R$, we can solve (16) with respect to the other variable by applying Dinkelbach’s algorithm and solving a sequence of convex problems. The formal algorithm is reported next and labeled Algorithm 2.

**Algorithm 2** Alternating maximization for Problem (16)

Initialize $P_R^{(0)} \in [0,P_{max}^R]$. Set a tolerance $\epsilon$. Set $n = 0$.

while $|\text{EE}_2^{(n)} - \text{EE}_2^{(n-1)}| \geq \epsilon$ do

$P_S^{(n+1)} \leftarrow$ Solve Problem (16) with respect to $P_S$ for fixed $P_R^{(n)}$.

$P_R^{(n+1)} \leftarrow$ solve Problem (16) with respect to $P_R$ for fixed $P_S^{(n+1)}$.

$n \leftarrow n + 1$

end while

Output $(P_S^{(n)}, P_R^{(n)})$.

The following proposition holds.

**Proposition 4:** Algorithm 2 converges to a stationary point of Problem (16).

**Proof:** After each iteration of Algorithm 2 the objective is not decreased. Hence, convergence follows since the objective is upper bounded. Convergence to a stationary point follows by leveraging [27, Proposition 2.7.1], which states that alternating maximization converges to a stationary point if: 1) the feasible set is the Cartesian product of closed and convex sets; 2) the objective is continuously differentiable on the feasible set; 3) the solution to each subproblem is unique. In our case, 1) and 2) are apparent. As for Assumption 3), it follows from the properties of fractional functions. In particular, when either $P_S$ or $P_R$ is fixed, the numerator of the objective is strictly concave in the other variable, as can be seen by computing the second derivative (see also [12]). Then, since the ratio between a strictly concave and an affine function is known to be strictly pseudo-concave, each subproblem in Algorithm 2 has a strictly pseudo-concave objective. Finally, the thesis follows because a strictly pseudo-concave function admits a unique maximizer.

■
b) Monotonic optimization approach: As mentioned, when the concave/convex structure of the objective does not hold, Dinkelbach’s algorithm still converges to the global solution of the original fractional problem, provided one can globally solve the non-convex Problem (13) in each iteration. For the case at hand, Problem (13) takes the form

\[
\begin{array}{ll}
\max_{P_S, P_R} & \alpha C \left( \frac{P_S P_R}{a P_S + b P_R + c} \right) - \lambda (\phi P_S + \psi P_R + P_c) \\
\text{s.t.} & P_S \in [0; P_S^{\max}], \quad P_R \in [0; P_R^{\max}].
\end{array}
\]  

(17)

Although being non-convex, Problem (17) can be globally solved by means of monotonic optimization theory. For a detailed review on monotonic optimization we refer to [28], [29]. Here we only remark that monotonic optimization provides a framework to globally maximize monotone functions over normal sets, and recall the following facts.

**Definition 1:** Let \( x_{\min}, x_{\max} \in \mathbb{R}^n_+ \) with \( 3x_{\min} \leq x_{\max} \). Then, \([x_{\min}, x_{\max}] = \{x \in \mathbb{R}^n_+: x_{\min} \leq x \leq x_{\max}\}\) is a hyper-rectangle in \( \mathbb{R}^n_+ \) and the set \( S \subset \mathbb{R}^n_+ \) is normal if \([0, \bar{x}] \in S, \forall \bar{x} \in S\).

**Proposition 5:** The set \( S = \{x \in \mathbb{R}^n_+: g(x) \leq 0\}\) is normal and closed if \( g: \mathbb{R}^n_+ \to \mathbb{R} \) is lower semi-continuous and increasing.

The feasible set of (17) is normal, being the cartesian product of two closed convex intervals in \( \mathbb{R} \), but the objective function is not monotone, as it is the difference of two increasing functions. However, the following proposition shows how to reformulate (17) into a monotonic problem.

**Proposition 6:** Problem (17) can be equivalently reformulated as

\[
\begin{array}{ll}
\max_{P_S, P_R, t} & \alpha C \left( \frac{P_S P_R}{a P_S + b P_R + c} \right) + t \\
\text{s.t.} & P_S \in [0; P_S^{\max}], \quad P_R \in [0; P_R^{\max}], \quad t \geq 0, \quad t + \lambda (\phi P_S + \psi P_R) \leq \lambda (\phi P_S^{\max} + \psi P_R^{\max}).
\end{array}
\]  

(18)

Problem (18) is a monotonic problem in canonical form.

**Proof:** Introducing the auxiliary variable \( t \) and using the substitution \( \lambda (\phi P_S + \psi P_R + P_c) + t = \lambda (\phi P_S^{\max} + \psi P_R^{\max} + P_c) \), yields the objective of (18), up to the inessential constant \(-\lambda (\phi P_S^{\max} + \psi P_R^{\max} + P_c)\). Next, since \( \lambda (\phi P_S + \psi P_R + P_c) \) is non-negative and increasing.

\(^3\)Inequalities between vectors are component-wise.
in both $P_S$ and $P_R$, the considered substitution is equivalent to $t \geq 0$, $t + \lambda (\phi P_S + \psi P_R) \leq \lambda (\phi P_S^{\text{max}} + \psi P_R^{\text{max}})$, and we obtain Problem (18). Finally, we have to show that (18) is a monotonic problem. The monotonicity of the objective follows because the function $C(\cdot)$ is increasing and its argument $\frac{P_S P_R}{a P_S + b P_R + c}$ can be shown to be increasing in both $P_S$ and $P_R$ by direct computation of first-order derivatives. Moreover, since the constraint function $t + \lambda (\phi P_S + \psi P_R)$ is increasing, the feasible set is contained in the hyper-rectangle $[0; P_S^{\text{max}}] \times [0; P_R^{\text{max}}] \times [0; \lambda(\phi P_S^{\text{max}} + \psi P_R^{\text{max}})]$. Moreover, the feasible set is also normal, by virtue of Proposition 5.

As a consequence of Proposition 6, we can apply Dinkelbach’s algorithm to globally solve (16), where in each iteration the solution to the subproblem (17) is found by solving the equivalent problem (18) by standard monotonic optimization algorithms, such as the polyblock algorithm [29]. The drawback of this approach is that the complexity of monotone programming is in general exponential. However, we stress two points. First, in our scenario we only have three variables, $P_S$, $P_R$, and the auxiliary variable $t$, which makes the use of monotonic optimization viable. Second, the proposed method provides a benchmark for the proposed low-complexity algorithm based on alternating optimization. As already mentioned, our numerical results will show that the two methods enjoy very similar performance.

**B. Competitive resource allocation**

After analyzing the cooperative case, we turn our attention to the competitive scenario in which a distributed power allocation is performed. This scenario is particularly relevant for distributed networks in which no central coordination is employed and power allocation needs to be performed in a distributed fashion. More in detail, in this section we consider the case in which $P_S$ and $P_R$ are not jointly allocated to maximize the GEE, but rather are separately allocated in a competitive fashion for individual EE maximization. We observe that in the considered symmetric scenario where the three users use a common power level $P_S$, the competition occurs between the relay, which allocates $P_R$, and the collective of the three transmitters, which together allocate $P_S$. This competitive power control problem can be formulated as a non-cooperative game in normal form $G$, whose players are the relay $\mathcal{R}$, with strategy set $[0; P_R^{\text{max}}]$, and the three the sources together $\{S_k\}_{k \in K}$, with strategy set $[0; P_S^{\text{max}}]$. The players’ utility functions are the
individual EEs, defined as
\[ u_S = \frac{R_{\Sigma}}{P_S + P_{c,S}}, \quad u_R = \frac{R_{\Sigma}}{P_R + P_{c,R}}, \]  (19)
with \( P_{c,S} \) and \( P_{c,R} \) the energy consumptions of the three sources and of the relay, respectively.\(^4\)
We observe that both the EEs in (19) have the achievable sum rate at the numerator. This is motivated by the following observation: since the users allocate the same power level \( P_S \), they should tune \( P_S \) so as to optimize their collective benefit, i.e. the achievable sum rate \( R_{\Sigma} \). On the other hand, the relay benefit is also represented by the achievable sum rate, since the relay typically does not favor any particular user. Moreover, in many circumstances, the relay is deployed by a third-party operator, which loans the relay to the users, and whose revenue is proportional to the total amount of conveyed data.

The BRD of \( G \) is expressed as the following two coupled problems \( \max_{P_i} \frac{R_S(P_i,P_{-i})}{P_i + P_{c,i}} \), with \( P_i \in [0; P_i^{\text{max}}] \), \( i \in \{S,R\} \), and \( P_{-i} \) denoting the strategy of the player other than \( i \). The optimization problem for \( i = S \) is the best response (BR) of the sources to the relay strategy, whereas the problem for \( i = R \) is the BR of the relay to the sources strategy. Depending on the relay operation, \( R_{\Sigma} \) can take one of the two following general forms
\[ R_{\Sigma,1} = \min \left\{ a_1 C\left(\frac{P_R}{N_S}\right), a_2 C\left(\frac{P_S}{N_R}\right) \right\}, \quad R_{\Sigma,2} = C\left(\frac{P_SP_R}{aP_S + bP_R + c}\right). \] (20)

In the following, we will denote by \( G_1 \) and \( G_2 \) the game with \( R_{\Sigma,1} \) and \( R_{\Sigma,2} \), respectively. Moreover, for ease of notation, let us define \( \gamma_1 = \frac{a_1 P_R}{N_S} \), \( \gamma_2 = \frac{a_2 P_S}{N_R} \), and \( \gamma = \frac{P_SP_R}{aP_S + bP_R + c} \). We start our analysis with the following result.

**Proposition 7:** \( G_1 \) and \( G_2 \) always admit an NE. Moreover, denote by \( \text{BR}_S(P_R) \) and \( \text{BR}_R(P_S) \) the BRs of the sources and of the relay, respectively. In case of \( G_1 \) it holds
\[ \text{BR}_S(P_R) = \min(\widetilde{P}_S^{\text{max}}, \overline{P}_S^{\text{max}}, \overline{P}_S), \quad \text{BR}_R(P_S) = \min(\widetilde{P}_R^{\text{max}}, \overline{P}_R^{\text{max}}, \overline{P}_R), \] (21)
with \( \overline{P}_S^{\text{max}} = \frac{N_R}{a_2}\left((1 + \frac{a_1 P_R}{N_S})^{a_1/a_2} - 1\right) \), \( \overline{P}_R^{\text{max}} = \frac{N_S}{a_1}\left((1 + \frac{a_2 P_S}{N_R})^{a_2/a_1} - 1\right) \), \( \overline{P}_S \) the unique solution of the equation \( \frac{N_R C(\gamma_2)}{a_2 C(\gamma_2)} - P_S = P_{c,S} \), and \( \overline{P}_R \) the unique solution of the equation \( \frac{N_S C(\gamma_1)}{a_1 C(\gamma_1)} - P_R = P_{c,R} \).

\(^4\)We have not considered any scaling coefficients in front of the transmit powers. This entails no loss of generality since any scaling coefficient can be included in the circuit power terms without affecting the maximizers of the utilities.
In case of $G_2$ it holds

$$
\text{BR}_S(P_R) = \min(P_{S}^{\max}, \bar{P}_S), \quad \text{BR}_R(P_S) = \min(P_{R}^{\max}, \bar{P}_R),
$$

(22)

with $\bar{P}_S$ the unique solution of the equation
$$
\frac{C(\gamma)}{C(\gamma)d\gamma/dP_S} - P_S = P_{c,S}
$$

and $\bar{P}_R$ the unique solution of the equation
$$
\frac{C(\gamma)}{C(\gamma)d\gamma/dP_R} - P_R = P_{c,R}.
$$

**Proof:** The existence follows by observing that for both $G_1$ and $G_2$ the strategy sets are closed and compact [30], the utility functions are continuous in $P_S$ and $P_R$, and that $u_S$ is pseudo-concave in $P_S$, while $u_R$ is pseudo-concave in $P_R$.

Next, let us focus on $G_1$ and on the derivation of $\text{BR}_S(P_R)$. For any fixed $P_R$, the $P_S$ where $a_1 C(\gamma_1)$ intersects $a_2 C(\gamma_2)$ is $\bar{P}_S^{\max}$. A key observation is that the optimal $P_S$ will never be larger than $\bar{P}_S^{\max}$, because for $P_S \geq \bar{P}_S^{\max}$ the numerator of $u_S$ keeps constant, while the denominator increases. On the other hand, in the range $[0; \bar{P}_S^{\max}]$, we have $u_S = a_2 \frac{C(\gamma_2)}{P_S^{\max} + P_{c,S}}$, which is increasing in $[0; \bar{P}_S]$, with $\bar{P}_S$ its unique maximizer. This proves $\text{BR}_S(P_R)$ in (21) and by a similar reasoning we can prove $\text{BR}_R(P_S)$.

Finally, let us focus on $G_2$ and on the derivation of $\text{BR}_S(P_R)$. In this case, the utility functions are differentiable, which allows to obtain $\bar{P}_S$ from the first-order optimality condition
$$
\frac{C(\gamma)}{C(\gamma)d\gamma/dP_S} - P_S = P_{c,S}
$$

Since $u_S$ is also pseudo-concave in $P_S$, first-order conditions are also sufficient and $\bar{P}_S$ is the global maximizer of $u_S$ for any fixed $P_R$. Accounting for the maximum power constraint we obtain $\text{BR}_S(P_R)$ in (22), while $\text{BR}_R(P_S)$ follows by a similar argument.

After establishing the existence of at least one NE, a natural question to ask is whether a unique NE exists and whether the BRD converges to an NE. Unlike previous related results on non-cooperative power control in two-hop systems [31], [12], which show the uniqueness of the NE for the scenario in which the relay is not one player of the game, for the case at hand uniqueness does not hold. To see this, first observe that the strategy profile $(P_S = 0, P_R = 0)$ is an NE. Indeed, if either $P_S$ or $P_R$ is set to zero, then the utility functions are both identically zero and the other player has no incentive to use a non-zero transmit power. This circumstance is a direct consequence of the fact that the relay has been included as a player of the game, a scenario which was not considered in previous works where the relay was assumed to have no EE concerns. Next, it is easy to find numerical examples in which also non-trivial NE exist, thus proving that more than one NE exists in general. However, while the uniqueness of the NE is
lost if the relay takes part in the game, the convergence of the BRD still holds. To show this, the following lemma is required, which proves the monotonicity of the BR functions of \( G_1 \) and \( G_2 \).

**Lemma 6:** For \( G_1 \) and \( G_2 \), BR\(_S(P_R)\) is increasing in \( P_R \) and BR\(_R(P_S)\) is increasing in \( P_S \).

**Proof:** Let us consider \( G_1 \) first. We will show the monotonicity of BR\(_S(P_R)\). The monotonicity of BR\(_R(P_S)\) can be obtained by a similar argument. From Proposition 7 we know that BR\(_S(P_R) = \min(P_{S_{\text{max}}}, \bar{P}_{S_{\text{max}}}, \bar{P}_S)\). The only argument of the minimum function that depends on \( P_R \) is \( \bar{P}_{S_{\text{max}}} \), since the equation whose solution defines \( \bar{P}_S \) does not actually depend on \( P_R \). Moreover, \( \bar{P}_{S_{\text{max}}} \), whose expression has been provided in Proposition 7, is an increasing function of \( P_R \). Consequently, recalling that the \( \min \) function is increasing, we conclude that BR\(_S(P_R)\) is increasing in \( P_R \).

Let us now consider \( G_2 \). The BR of the sources to the relay strategy is BR\(_S(P_R) = \min(P_{S_{\text{max}}}, P_S)\), with \( P_S \) obtained by solving

\[
\frac{C(\gamma)}{C'(\gamma)d\gamma/dP_S} - P_S = P_{e,S}
\]  

Let us define the LHS of (23) as the function \( g(P_S, P_R) \). The first step of the proof is to show that \( g \) is increasing in \( P_S \).

\[
\frac{\partial g}{\partial P_S} = -\frac{C(\gamma) \left[ C''(\gamma)(P_S) + \gamma''(P_S) C'(\gamma) \right]}{(C'(\gamma)\gamma'(P_S))^2} \geq 0,
\]  

where we have exploited the fact that \( C(\gamma) \) and \( \gamma(P_S) \) are positive, increasing, and concave.

Next, we will show that \( g \) is decreasing in \( P_R \).

\[
\frac{\partial g}{\partial P_R} \leq 0 \iff \left( C'(\gamma) \right)^2 - C(\gamma) C''(\gamma) \leq C(\gamma) C'(\gamma) \frac{\partial^2 \gamma}{\partial P_S P_R} \frac{1}{\partial P_R \partial P_S} \frac{1}{\gamma + \delta}
\]  

Elaborating, the last two terms of the RHS can be expressed as

\[
\frac{\partial^2 \gamma}{\partial P_S P_R} \frac{1}{\partial P_R \partial P_S} = \frac{1}{(bP_R + c)(aP_S + c)} \left( \frac{aP_S + bP_R + c}{P_SP_R} \right) = \left( 1 + \frac{c}{ab} \frac{aP_S + bP_R + c}{P_SP_R} \right) \frac{1}{\gamma} + \frac{1}{\gamma + d}
\]  

with \( d = c/ab \). Thus, (25) becomes \( \left( C'(\gamma) \right)^2 - C(\gamma) C''(\gamma) \leq C(\gamma) C'(\gamma) \left( \frac{1}{\gamma} + \frac{1}{\gamma + d} \right) \), and
plugging the expressions of $C(\gamma)$ and its derivatives, we have

$$\gamma \leq \log(1 + \gamma) + \frac{\gamma(1 + \gamma)}{\gamma + d} \log(1 + \gamma)$$

(26)

It can be seen that (26) holds if $d \leq 1$. In particular, for $d = 1$ we have $\gamma \leq (\gamma + 1) \log(1 + \gamma)$. Both functions start from zero, but the RHS has a larger derivative for all $\gamma \geq 0$. Then, (26) holds also for $d < 1$ since the RHS is decreasing in $d$. For both AF-IAN and NNC we have $d = 1$, while for AF-SND we have $d = 2/3$. Hence, for all of the considered schemes, (26) is true and if $P_R$ increases, the LHS of (23) decreases. As a consequence, in order to reach the constant level $P_{c,S}$ at the RHS, $\bar{P}_S$ must increase, since we have also shown that the LHS of (23) is increasing in $P_S$. Finally, the monotonicity of the BR follows from the fact that the min function is increasing.

**Proposition 8:** For any set of system parameters and initialization point, the BRDs of both $G_1$ and $G_2$ are guaranteed to converge.\(^5\)

**Proof:** Denote by $P_S^{(0)}$ the initial value of $P_S$. Then, $P_R^{(0)} = \text{BR}_R(P_S^{(0)})$ and $P_S^{(1)} = \text{BR}_S(P_R^{(0)})$. Let us consider three cases.

If $P_S^{(1)} = P_S^{(0)}$, we have reached convergence.

If $P_S^{(1)} > P_S^{(0)}$, then we also have $P_R^{(1)} = \text{BR}_R(P_S^{(1)}) \geq P_R^{(0)}$ by virtue of Lemma 6. But then, at the second iteration we also have $P_S^{(2)} = \text{BR}_S(P_R^{(1)}) \geq P_S^{(1)}$ and $P_R^{(2)} = \text{BR}_R(P_S^{(2)}) \geq P_R^{(1)}$. Similarly, at the $n$-th iteration we have $P_S^{(n)} \geq P_S^{(n-1)}$ and $P_R^{(n)} \geq P_R^{(n-1)}$. Since the sources and relay BRs are upper bounded by $P_S^{\text{max}}$ and $P_R^{\text{max}}$, respectively, the iteration must converge.

If $P_S^{(1)} < P_S^{(0)}$, by a similar argument it follows that after each iteration $P_S$ and $P_R$ are not increased, i.e. $P_S^{(n)} \leq P_S^{(n-1)}$ and $P_R^{(n)} \leq P_R^{(n-1)}$. Thus, convergence must eventually occur since the BRs are non negative.

Based on Proposition 8, the BRD-based, competitive, power control algorithm in which each player iteratively maximizes his own utility function is guaranteed to converge. The formal procedure is reported next and labeled Algorithm 3.

Even though it would seem that Algorithm 3 requires each player to know the other player’s strategy, this is not the case and Algorithm 3 can be actually implemented in a distributed fashion with similar techniques as explained in [12].

\(^5\)Convergence is meant in the value of the utilities.
Algorithm 3 Competitive power control algorithm

Initialize $P_R^{(0)} \in [0, P_{max}^R]$. Set a tolerance $\epsilon$.

while $\max_{i=S,R} \left\{ |EE^{(n)}_i - EE^{(n-1)}_i| \right\} \geq \epsilon$

$P_S^{(n+1)} \leftarrow \min(P_S^{max}, \bar{P}_S)$ for given $P_R^{(n)}$, with $\bar{P}_S$ defined in Proposition 7

$P_R^{(n+1)} \leftarrow \min(P_R^{max}, \bar{P}_R)$ for given $P_S^{(n+1)}$, with $\bar{P}_R$ defined in Proposition 7

end while

Output $(P_S^{(n)}, P_R^{(n)})$.

V. Extension to non-symmetric channels

In case of non-symmetric channels, obtaining closed-form sum rate expressions is very cumbersome and leads to involved expressions that do not allow to gain much analytical insight on the system achievable sum rate and energy efficiency, as in the symmetric case. For this reason, in this work we focused on symmetric channels. Nevertheless, in this section we propose an alternative approach to energy-efficient power control in asymmetric scenarios, that directly uses the rate region instead of the sum rate expression. A similar approach has been used for sum rate optimization in the downlink of a MIMO multi-cell network [32].

The GEE optimization problem from Section IV can be stated in more general form as

$$\begin{aligned}
\max_{R_1,R_2,R_3, P_1,P_2,P_3,P_R} & \quad \frac{R_1 + R_2 + R_3}{\phi_1 P_1 + \phi_2 P_2 + \phi_3 P_3 + \psi P_R + P_c} \\
\text{s.t.} & \quad (R_1, R_2, R_3) \in \mathcal{R}(P_1, P_2, P_3, P_R), P_R \in [0, P_{max}^R], P_i \in [0, P_{max}^i], i = 1, 2, 3
\end{aligned}$$

(27)

This is a fractional programming problem and can be solved using Dinkelbach’s algorithm as we have seen in Section IV. Here, the objective is pseudo-linear and the complexity is related to the constraints. The rate region $\mathcal{R}$ of Gaussian channels usually consists of several inequalities with linear functions of the rates on the LHS and functions of the transmit powers on the RHS. In most cases, the RHSs are non-convex functions. However, in many cases they belong to the class of d.i. functions (difference of increasing functions). Thus, the auxiliary problem (13), which has to be solved for Dinkelbach’s algorithm, can be transformed into a monotonic optimization problem following the steps in [28, Section 7.3]. Then, it can be solved using monotonic optimization.

As an illustrative example, consider AF-SND. It can be easily seen from the proof to Lemma 3
that $\mathcal{R}$ contains all non-negative rate tuples $(R_1, R_2, R_3)$ that satisfy
\[
\sum_{i \in S} R_i \leq C \left( \frac{P_R \sum_{i \in S} P_i}{N_R P_R + N_q(k) \left( N_R + \sum_{j \in \mathcal{K}} P_j \right)} \right),
\]
for all $S \in \{\{k\}, \{k, l(k)\} : k \in \mathcal{K}\}$. The RHS of (28) can be written as
\[
\log \left( P_R \sum_{i \in S} P_i + N_R P_R + N_q(k) \left( N_R + \sum_{j \in \mathcal{K}} P_j \right) \right) - \log \left( N_R P_R + N_q(k) \left( N_R + \sum_{j \in \mathcal{K}} P_j \right) \right),
\]
which is a d.i. function. Thus, we have a monotonic optimization problem.

The detailed derivation of the rate regions with asymmetric channels for the other considered relaying/decoding schemes is not trivial and is left as future work. As for AF and NNC one should extend the results in Section III, accounting for arbitrary time sharing in case of AF-IAN. As for DF, our approach could be extended by leveraging [33], [34]. In any case, as already mentioned for the sum rate, the resulting expressions would be rather involved, thus preventing to gain much analytical insight compared to the symmetric case.

VI. NUMERICAL RESULTS

For a discussion and numerical evaluation of the presented transmission schemes, we consider completely symmetric channels with $N = N_S = N_R$ and $P = P_S = P_R$. For the EE evaluation, we have $P_{\text{max}} = P_{S_{\text{max}}} = P_{R_{\text{max}}}$ and define $\text{SNR}_{\text{max}} = P_{\text{max}} / N$. At first, we will assume $P_c = 1$ W and no power loss at the transmitter, i.e., $\psi = 1$ and $\phi = 3$, to compare algorithms and gain some general insight on the EE. In the next subsection, we model a wireless 200 GHz board-to-board communication system to obtain more realistic simulation parameters.

Fig. 2 shows the achievable sum rates from Section III as a function of the SNR. It reflects the analytical results derived in Section III-E. First, observe that both NNC and AF-SND scale with the outer bound in the high SNR regime. Second, DF achieves the sum capacity up to a SNR of 8.1 dB and is the best performing scheme up to 14.27 dB (see Remarks 4 and 6). In the high SNR regime, its gap to the outer bound grows unbounded. Finally, AF-IAN scales with DF in the high SNR regime. It can be shown using the same methods as in Section III-E that its gap to DF approaches exactly 2 bit as $\text{SNR} \to \infty$.

Fig. 3 illustrates the performance of the cooperative GEE maximization algorithms developed in Section IV-A. As for the GEE$_2$-type optimization problems, the performance of both monotonic
optimization and alternating optimization are reported and it is seen that they perform virtually equally, thus suggesting that alternating optimization achieves the global optimum of the GEE. This is also the case in all other numerical simulations carried out in this paper. Thus, the alternating optimization algorithm is a good alternative to monotonic optimization when low computational complexity is the primary concern.

Fig. 3a shows the GEE as a function of the SNR for a fixed circuit power $P_c = 1$ W and unit noise variance. First of all, it can be seen that the GEE saturates when $\text{SNR}^{\text{max}}$ exceeds a given value, which is lower than 5 dB for all considered schemes. This is explained recalling that, unlike the achievable rate, the GEE is not increasing with the transmit powers, but instead admits an optimum transmit power level. If $P^{\text{max}}$ is larger than such power level, then it is not optimal to transmit at full power. This also explains why DF performs significantly better than all other schemes, including NNC. Indeed, due to the saturation of the GEE, the SNR range for which NNC yields a larger achievable sum rate than DF is not reached when GEE is optimized. Finally, as expected, NNC is better than AF-SND, which is better than AF-IAN.

In Fig. 3b the same simulation parameters as in Fig. 3a have been used but with different noise variance $N = 0.1 \text{ mW}$. This results in a much higher saturation point which is approximately 30 dB and produces a different behavior of the curves. In the low SNR regime, DF achieves the outer bound, while in the high SNR regime NNC and AF-SND perform best while DF has
Next, Figs. 4 and 5 consider the competitive scenario of Section IV-B, also in comparison with the performance of the cooperative algorithms. The same simulation parameters as in Fig. 3a have been used, and it is assumed that every node consumes the same circuit power. Thus, \( P_{c,S} = \frac{3}{4} P_c \) and \( P_{c,R} = \frac{1}{4} P_c \). In particular, Fig. 4 compares the GEE achieved by the competitive power control algorithms to that obtained by the cooperative allocation of Section IV. It is seen that competitive and cooperative power control perform virtually the same if DF is used, whereas a limited gap appears if the other relaying strategies are employed. This can be intuitively explained recalling that a more significant coupling in a non-cooperative game results in a higher price of anarchy, and in our scenario, the game for DF is more coupled than the game for the other schemes. Indeed, the numerator of the players EEs for DF is given by the equation on the left in (20), while for the other schemes it is given by the equation on the right. In the latter case, the two variables \( P_S \) and \( P_R \) are more heavily coupled than in the former one because they appear significantly lower GEE. Due to the different DoFs, it can be expected that for even higher saturation points the gap between NNC and DF grows further, while the gaps between NNC and AF-SND, and between DF and AF-IAN stay nearly the same.
in the same $C(\cdot)$ expression.

Instead, Fig. 5 shows the behavior of the competitive algorithms for DF and NNC for different initialization values. It is observed that DF is more sensitive to the initialization point than NNC. In particular, the closer the initialization point to $P^{\text{max}}$, the sooner the GEE reaches the saturation level. This can be explained recalling that for DF the BRs are expressed as in (21), where the variables $\bar{P}_{S}^{\text{max}}$ and $\bar{P}_{R}^{\text{max}}$ appear, which are increasing in the strategy of the other player. If the initialization values of the transmit power is low with respect to $P^{\text{max}}$, the corresponding values $\bar{P}_{S}^{\text{max}}$ and $\bar{P}_{R}^{\text{max}}$ will also be low and will likely prevail in the $\min$ operation, preventing from transmitting at the saturation levels $\bar{P}_{S}$ and $\bar{P}_{R}$. This does not occur for NNC and AF since the BR has a different expression which does not depend on similar variables as $\bar{P}_{S}^{\text{max}}$ and $\bar{P}_{R}^{\text{max}}$.

Comparing previous illustrations and especially Fig. 3, we learn that the energy-efficient performance of the different transmission and relaying scheme highly depends on the saturation point of the GEE function, which in turn is heavily affected by the the system parameters. Thus, for a sensible comparison of the different schemes, it is necessary to choose physically meaningful values for the noise variance and the circuit power consumption. Furthermore, assuming the same circuit power consumption for all schemes is unfair since they have different hardware complexity. For this reason, in the following subsection we consider a specific communication system and develop a simple model for its power consumption which allows for a fair comparison of the considered schemes.

A. mmWave Board-to-Board Communications

We consider a wireless board-to-board communication system operating at 200 GHz carrier frequency with 25 GHz bandwidth. Channel measurements and a link budget for such a system with 4-by-4 antenna arrays are given in [2]. Since there is no mobility in this scenario, the beams have a fixed direction and the antenna array combined with the beamforming network can be regarded as a single antenna with a gain of 7 dB. With a link length of 0.1 m, the resulting channel gain is -65.8 dB as reported in [2, Table I].\(^6\)\(^7\) As for the noise, we have $N_{S} = N_{R} = k_{B}T_{c}$ where $k_{B}$ is the Boltzmann constant and $T_{c}$ the absolute temperature of operation.

\(^6\)Computed without considering the polarization mismatch since we do not consider polarization here.

\(^7\)In previous sections, unit channel gain and bandwidth have been assumed in the sum rate expressions. However, previous results and algorithms can be straightforwardly extended to the case in which the powers $P_{S}$ and $P_{R}$ in $R_{S}$ are scaled by a channel coefficient and divided by the bandwidth.
Next, we develop a model for the circuit power consumption $P_c$ that reflects the different hardware complexities of the considered transmission schemes. We assume that $P_c$ depends neither on the power allocation nor on the transmission rate for simplicity. The analog part of each receiver consists of an Rx frontend and two analog-to-digital converters (ADCs) consuming a total power of $P_{ADC}$. Similarly, each transmitter consists of a Tx frontend, two digital-to-analog converters (DACs) (consuming a power $P_{DAC}$), and a power amplifier (PA) with efficiency $\eta$, i.e., $\psi = \frac{1}{\eta}$ and $\phi = \frac{3}{\eta}$. The Rx frontend has 16 low noise amplifiers (LNAs), one for each antenna, two mixers and one LO-driver (local oscillator) for those mixers resulting in a power $P_{rx}$. Similarly, the Tx frontend consists of two mixers and a LO-driver consuming a power $P_{tx}$.\(^8\) Realistic numbers for hardware-dissipated power and PA efficiencies are reported in Table II, and the resulting model parameters are given in Table III. We define $P_{c,\text{analog}} = P_{rx} + P_{ADC} + P_{DAC} + P_{tx}$. The digital signal processing (DSP) power consumption $P_{DSP}$ of each node is modeled as a multiple of the power $P_{\text{dec}}$ required by a single user decoder. From the LDPC decoder model in [35]\(^9\) we get a decoding power of 267.6 mW. Adding some power to account for other processing such as encoding, we round $P_{\text{dec}}$ to 300 mW.

As for the users, when interference is treated as noise, only one message needs to be decoded,

\(^8\)We do not include the PA in the Tx frontend since it is modelled separately.

\(^9\)For our model, the following parameters are suitable: $E_{\text{edge}} = 10^8 \cdot k_B \cdot T$, $r_{\text{COD}} = \frac{1}{2}$, $R = 50$ GBit/s, $\lambda = 3$, $l = 2$. 
TABLE II
EXEMPLARY POWER CONSUMPTIONS FOR 200 GHz RF COMPONENTS.

<table>
<thead>
<tr>
<th>Component</th>
<th>Value</th>
<th>Unit</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixer</td>
<td>17</td>
<td>mW</td>
<td>Unpublished measurements(^{10})</td>
</tr>
<tr>
<td>LO-driver</td>
<td>24</td>
<td>mW</td>
<td></td>
</tr>
<tr>
<td>LNA</td>
<td>18</td>
<td>mW</td>
<td>[36]</td>
</tr>
<tr>
<td>ADC</td>
<td>406</td>
<td>mW</td>
<td>[37]</td>
</tr>
<tr>
<td>DAC</td>
<td>400</td>
<td>mW</td>
<td>Estimate(^{10})</td>
</tr>
<tr>
<td>PA efficiency</td>
<td>6.2</td>
<td>%</td>
<td>[38]</td>
</tr>
</tbody>
</table>

TABLE III
PARAMETERS FOR CIRCUIT POWER MODELLING OF BOARD-TO-BOARD COMMUNICATIONS.

<table>
<thead>
<tr>
<th>Component</th>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx Frontend</td>
<td>(P_{rx})</td>
<td>346</td>
<td>mW</td>
</tr>
<tr>
<td>Analog-to-digital converter</td>
<td>(P_{ADC})</td>
<td>812</td>
<td>mW</td>
</tr>
<tr>
<td>Single User Decoder</td>
<td>(P_{dec})</td>
<td>300</td>
<td>mW</td>
</tr>
<tr>
<td>Digital-to-analog converter</td>
<td>(P_{DAC})</td>
<td>800</td>
<td>mW</td>
</tr>
<tr>
<td>Tx Frontend</td>
<td>(P_{tx})</td>
<td>58</td>
<td>mW</td>
</tr>
<tr>
<td>Power amplifier</td>
<td>(\eta)</td>
<td>6.2</td>
<td>%</td>
</tr>
</tbody>
</table>

hence \(P_{c,S}^{IAN} = P_{c,analog} + P_{dec}\). For SND, we model the power consumption as if the two messages were decoded sequentially, hence, \(P_{DSP} = 2P_{dec}\).\(^{11}\) A similar assumption is used for receive decoding if DF relaying is used because also in this case the receiver has to decode two messages. Thus, \(P_{c,S}^{SND} = P_{c,S}^{DF} = P_{c,analog} + 2P_{dec}\).

As for the relay, DF uses a 3-user MAC decoder. Hence, \(P_{c,R}^{DF} = P_{c,analog} + 3P_{dec}\). The NNC relay does not decode the received messages, and, in our scenario, does not have a message to transmit on its own. Instead it uses a vector quantizer to compress the received sequence. Following [39], we assume that the (scalar) ADC in our model roughly models the quantization operation of NNC. The only DSP operation then is encoding, and we have \(P_{c,R} = P_{c,analog} + 0.1P_{dec}\).\(^{12}\) In contrast to the other relaying schemes, AF directly amplifies the analog signal, so neither DSP nor ADC and DAC are necessary. Thus, \(P_{c,R}^{AF} = P_{rx} + P_{tx}\). Finally, the circuit power \(P_c = 3P_{c,S} + P_{c,R}\).

Fig. 6 was obtained using the developed power consumption model with the algorithms from Section IV-A. It can be seen that in the low SNR regime DF performs best but that starting from approximately 10 dB AF-SND achieves the best GEE. This continues to hold for higher SNR’s, where DF is outperformed by all other schemes. Since there are currently no NNC implementations available, we might have underestimated its power consumption. Let us assume its decoding complexity is actually higher than that of DF and add an additional \(P_{dec}\) for each

\(^{10}\)Those values have been provided by project partners from the DFG CRC 912 “Highly Adaptive Energy-Efficient Computing”.

\(^{11}\)Recall that NNC uses SND at the receivers.

\(^{12}\)We assume \(P_{DSP} = 0.1P_{dec}\), which is roughly the power we accounted for other processing in \(P_{dec}\) before.
user and set $P_{DSP} = 4P_{dec}$ for the relay (instead of $0.1P_{dec}$). The resulting GEE is shown as a dashed line in Fig. 6. It can be seen that it still outperforms DF in the high SNR regime despite its much higher power consumption.

VII. CONCLUSION

In this paper, we studied both achievable sum rates and the EE of the symmetric 3-user MWRC with a partial message exchange. We provided analytic sum rate expressions for the most common relaying schemes and characterized the sum capacity for certain SNR regimes. Next, we discussed the problem of energy-efficient power control both in the cooperative and competitive scenario. The results indicate that competitive optimization suffers a limited gap with respect to cooperative resource allocation, and that the proposed low-complexity cooperative allocation achieves the global optimum.

Moreover with reference to a realistic mmWave board-to-board communication system, we have shown that AF performs best for SNRs above 10 dB, closely followed by NNC, while DF suffers from a higher hardware power consumption. However, as discussed in Section VI, the performance of the different scheme is significantly affected by the simulation parameters. Further analysis is required to understand which is the most energy-efficient relaying scheme in other communication systems.
ACKNOWLEDGEMENT

The authors would like to thank C. Carta, G. Tretter, D. Fritsche, and N. ul Hassan from Technische Universität Dresden for valuable discussions on the circuit power consumption.

REFERENCES


