

Fig. 1: N -hop MIMO AF system model.

The rest of the paper is organized as follows. Section II defines the achievable sum-rate, power consumption, and energy-per-bit consumption of the multi-hop MIMO-AF system. In Section III, we introduce the joint source and multiple relays optimization problem and prove that its objective function is not generally quasiconvex, but this function can be lower-bounded by a convex function; moreover, all the minima of this function are strict minima. We then simplify the original multivariate optimization problem into a single variable problem and design a novel approach, based on a simple unidimensional root finding method, for optimally solving it in the unconstrained and power constraint cases; we also provide a sub-optimal approach with reduced complexity. In Section IV, we discuss the accuracy and computational complexity of our novel approaches in numerous settings. The results show that our sub-optimal approach is near-optimal and has a lower complexity (up to 40 times) than our optimal approach. As an application, we utilize our novel approaches for comparing the optimal energy-per-bit consumption of multi-hop MIMO-AF and MIMO systems; results show that multi-hop MIMO-AF can be useful for saving energy when the direct link quality is poor. Finally, conclusions are drawn in Section V.

II. MULTI-HOP MIMO-AF EE FRAMEWORK

A. System model

In this paper, we consider a classic N -hop MIMO AF system, where an SN with t_1 antennas transmits a signal to a destination node (DN) with t_{N+1} antennas via $N - 1$ nonregenerative RN with t_i antennas, $i \in \{2, \dots, N\}$, as it is depicted in Fig. 1. The signal transmission is performed over N phases of equal duration and each node operates in half-duplex mode, as in [29] and [6], such that the aggregate mutual information (over N time slots) of this N -hop MIMO-AF system can be expressed as $I(\mathbf{y}_N; \mathbf{y}_0) =$

$$\log_2 \left| \mathbf{I}_{t_{N+1}} + \prod_{i=0}^{N-1} (\mathbf{H}_{N-i} \mathbf{F}_{N-i}) \prod_{i=1}^N (\mathbf{F}_i^\dagger \mathbf{H}_i^\dagger) \mathbf{R}_N^{-1} \right|, \quad (1)$$

where $\mathbf{H}_i \in \mathbb{C}^{t_{i+1} \times t_i}$ represents the MIMO channel between the i -th and $i + 1$ -th nodes, $\mathbf{F}_i \in \mathbb{C}^{t_i \times t_i}$ is the i -th node precoding matrix, \mathbf{I}_x is a $x \times x$ identity matrix, $|\cdot|$ is the matrix determinant, $(\cdot)^\dagger$ denotes the conjugate transpose, and the notation $\prod_{i=a}^b (A_i)$ is equivalent to $A_a A_{a+1} \dots A_{b-1} A_b$. In addition, the i -th noise covariance matrix, \mathbf{R}_i , with $i = N$ in (1), is defined as

$$\mathbf{R}_i = \sigma_i^2 \mathbf{I}_{t_{i+1}} + \mathbf{H}_i \mathbf{F}_i \mathbf{R}_{i-1} \mathbf{F}_i^\dagger \mathbf{H}_i^\dagger,$$

for any $i \in \mathcal{N} = \{1, \dots, N\}$, where σ_i^2 is the variance of the Gaussian noise vector $\mathbf{n}_i \in \mathbb{C}^{t_{i+1} \times 1}$. Note that $\mathbf{R}_0 = \mathbf{0}_{t_1}$ with $\mathbf{0}_x$ being a $x \times x$ matrix of zeros.

According to the Hadamard determinant theorem [30], an optimal precoder structure diagonalizes the matrix within the determinant in (1). In the two-hop scenario, such a structure has been proved to be optimal for maximizing the SE, minimizing the transmit power, and optimizing the EE in [17], [31] and [8], respectively. Assuming as in [29] that each node knows its previous and next links' CSI (only next and previous link CSI for the SN and DN, respectively), an optimal precoder structure is one of the form [29]

$$\mathbf{F}_i = \mathbf{V}_i \widehat{\mathbf{F}}_i \mathbf{U}_{i-1}^\dagger, \quad (2)$$

for any $i \in \mathcal{N}$, where \mathbf{V}_i and \mathbf{U}_{i-1} are unitary matrices that contain the t_i right-singular vectors of \mathbf{H}_i and t_i left-singular vectors of \mathbf{H}_{i-1} , respectively, with $\mathbf{U}_0 = \mathbf{I}_{t_1}$. In addition $\widehat{\mathbf{F}}_i = \text{diag}(f_{i,1}, \dots, f_{i,t_i})$ is a $t_i \times t_i$ diagonal matrix, where $f_{i,m} = \sqrt{p_{i,m}(\sigma_{i-1}^2 + p_{i-1,m} \lambda_{i-1,m})^{-1}}$. Inserting (2) into (1) and assuming perfect CSI knowledge, the sum-rate (sum of instantaneous rates, i.e. $WI(\mathbf{y}_N; \mathbf{y}_0)$) can be expressed as

$$R_\Sigma(\mathbf{P}) = W \sum_{m=1}^M \log_2 \left(\prod_{i=1}^N (1 + p_{i,m} \lambda_{i,m} \sigma_i^{-2}) \right) - \log_2 \left(\prod_{i=1}^N (1 + p_{i,m} \lambda_{i,m} \sigma_i^{-2}) - \prod_{i=1}^N p_{i,m} \lambda_{i,m} \sigma_i^{-2} \right), \quad (3)$$

where W is the channel bandwidth, $\lambda_{i,m}$ denotes the m -th eigenvalue of $\mathbf{H}_i \mathbf{H}_i^\dagger$ and M is the number of available subchannels for each hop. Note that $M = t$ in flat channel condition or $M = Kt$ in frequency selective channel condition [18], with K and $t \triangleq \min_{i \in \{1, \dots, N+1\}} \{t_i\}$ being the numbers of frequency-flat subchannels and spatial subchannels, respectively, for each hop. In addition, $p_{i,m}$ represents the transmit power of the m -th subchannel of node i , such that $\mathbf{P} = [p_{1,1}, \dots, p_{i-1,M}, p_{i,1}, \dots, p_{N,M}] \succeq \mathbf{0}$. Hence, the total transmit power of each node is given by [29]

$$P_i(\mathbf{P}) = \mathbb{E}\{\|\mathbf{F}_i \mathbf{y}_{i-1}\|_F^2\} = \sum_{m=1}^M p_{i,m}, \quad (4)$$

where $\mathbf{y}_i = \mathbf{H}_i \mathbf{F}_i \mathbf{y}_{i-1} + \mathbf{n}_i$, $\forall i \in \mathcal{N}$, $\mathbb{E}\{\cdot\}$ stands for the expectation and $\|\cdot\|_F^2$ denotes the Frobenius norm. By defining $\mathcal{C}_{i,m} = \log_2(1 + p_{i,m} \lambda_{i,m} \sigma_i^{-2})$ as the achievable SE of the m -th subchannel of node i , the total transmit power per node and sum-rate in (4) and (3), respectively, can then be re-expressed as a function of the SE, such that

$$P_i(\mathbf{C}) = \Delta_i^{-1} \sum_{m=1}^M A_{i,m} (2^{\mathcal{C}_{i,m}} - 1), \quad \text{and} \quad (5a)$$

$$R_\Sigma(\mathbf{C}) = W \sum_{m=1}^M \sum_{i=1}^N \mathcal{C}_{i,m} - \log_2 \left(\sum_{k=0}^1 j^{2k} \prod_{i=1}^N (2^{\mathcal{C}_{i,m}} - k) \right), \quad (5b)$$

respectively, where $\mathbf{C} = [\mathcal{C}_{1,1}, \dots, \mathcal{C}_{i-1,M}, \mathcal{C}_{i,1}, \dots, \mathcal{C}_{N,M}] \succeq \mathbf{0}$, $A_{i,m} \triangleq \Delta_i \sigma_i^2 \lambda_{i,m}^{-1}$ and $j = \sqrt{-1}$. Note that Δ_i is a power consumption parameter (see next subsection for more details).

B. Power consumption model

During the propagation of the signal \mathbf{y}_0 from the SN to the DN via the $N - 1$ relays (see Fig. 1), three types of

power consumption modes can be identified for the nodes, i.e. transmission, P_{Tx} , reception, P_{Rx} , and sleep, P_{SI} , such that the total consumed power of the MIMO-AF system over N time slots can be formulated as follows

$$P_{\Sigma} = P_{\text{Tx}}^{\text{SN}} + P_{\text{Rx}}^{\text{DN}} + t(N-1)(P_{\text{SI}}^{\text{SN}} + P_{\text{SI}}^{\text{DN}}) + \sum_{i=1}^{N-1} P_{\text{Tx}}^{\text{RN}_i} + P_{\text{Rx}}^{\text{RN}_i} + t(N-2)P_{\text{SI}}^{\text{RN}_i}, \quad (6)$$

when assuming that the N transmission phases have equal duration. In transmission mode, a node transmits data to another node; the transmitting node consumes power for preparing the information to be sent (e.g. baseband processing, RF transceiver chain) and sending it (power amplifier). It has been established that the power consumption of the most common types of equipments in a relay-based MIMO cooperative system, e.g. BS, RN or user equipment (UE), grows linearly with their transmit power in transmission mode (see [5], [32] or [24], respectively). Hence, the power consumption of a node in transmission mode can be expressed via a generic linear MIMO power model [33], such that

$$P_{\text{Tx}} = \Delta_i P_i(\mathbf{C}) + tP_{\text{CipA}} + P_{\text{Ci}}, \forall i \in \mathcal{N}, \quad (7)$$

where Δ_i , P_{Ci} and P_{CipA} are parameters modeling the power consumption of various elements, e.g. power amplifier, DC-DC conversion, baseband processing, RF transceiver chain [5]; Δ_i models the inefficiency of the i -th transmitting node power amplifier, P_{CipA} models the circuit power consumption that scales with the number of antennas (e.g. RF transceiver chain power consumption), and P_{Ci} models the other types of circuit power consumption (e.g. DC-DC conversion, baseband processing). In reception mode, a node receive data from a transmitting node; the receiving node consumes power for receiving the information (e.g. RF transceiver chain) and processing it (e.g. baseband processing), such that $P_{\text{Rx}} = \varsigma[tP_{\text{CipA}} + P_{\text{Ci}}]$, where $0 \leq \varsigma \leq 1$; ς characterizes the fact that reception usually consumes less circuit power than transmission. Finally, given that the information propagates over N time slots and only two nodes are active (i.e. one in transmission and one in reception mode) for a given time slot, it implies that the other $N-1$ nodes are inactive, i.e. in sleep mode. In sleep mode, a node waits to transmit/receive and does not perform any processing, such that only a fraction of the circuit power is consumed [5]. By inserting the definitions of P_{Tx} and P_{Rx} into (6), the total power consumption of the N -hops MIMO-AF system can be re-expressed as

$$P_{\Sigma}(\mathbf{C}) = P_c + \sum_{i=1}^N \Delta_i P_i(\mathbf{C}), \quad (8)$$

where $\Delta_1 = \Delta^{\text{SN}}$, and $\Delta_i = \Delta^{\text{RN}_{i-1}}$, for any $i \in \{2, \dots, N\}$. Moreover,

$$P_c = tP_{\text{CipA}}^{\text{SN}} + P_{\text{Ci}}^{\text{SN}} + \varsigma(tP_{\text{CipA}}^{\text{DN}} + P_{\text{Ci}}^{\text{DN}}) + t(N-1)(P_{\text{SI}}^{\text{SN}} + P_{\text{SI}}^{\text{DN}}) + \sum_{i=1}^{N-1} (1 + \varsigma)(tP_{\text{CipA}}^{\text{RN}_i} + P_{\text{Ci}}^{\text{RN}_i}) + t(N-2)P_{\text{SI}}^{\text{RN}_i} \quad (9)$$

according to (6), (7) and (8); P_c accounts for all the fixed circuit consumed powers.

C. EE formulation

The existence of a trade-off between EE and SE [34] implies that these two quantities can only be jointly optimized by using the explicit expression of this trade-off as an objective function. In the general case, it has been shown in [34] that an explicit expression of this trade-off can be obtained through the ratio between the sum-rate and total consumed power, expressed as a function of the SE, \mathcal{C} , which are respectively given in (5b) and (8) & (5a). Consequently, the EE-SE trade-off of the N -hop MIMO-AF system with CSI knowledge can be expressed as $E_b(\mathbf{C}) = P_{\Sigma}(\mathbf{C})/R_{\Sigma}(\mathbf{C}) =$

$$\frac{P_c + \sum_{i=1}^N \sum_{m=1}^M A_{i,m} (2^{C_{i,m}} - 1)}{W \sum_{m=1}^M \sum_{i=1}^N C_{i,m} - \log_2 \left(\sum_{k=0}^1 j^{2k} \prod_{i=1}^N (2^{C_{i,m}} - k) \right)}, \quad (10)$$

where E_b stands for the energy-per-bit, i.e. 1/EE.

III. MULTI-HOP MIMO-AF EE OPTIMIZATION

We have recently proposed in [28], a near-optimal joint resource allocation scheme for multi-hop MIMO-AF systems by assuming that E_b was quasiconvex for any $\mathbf{C} \succeq \mathbf{0}$ and regardless of N . However, it turns out that E_b is not necessarily quasiconvex in the general case, as it is discussed in section A of the Appendix; thus, a different approach must be followed for optimizing the EE of N -hop MIMO-AF systems. Even though E_b is not necessarily quasiconvex, by knowing that it can be lower bounded by a convex function, we design here algorithms for optimally solving the following problem

$$E_b^* = \min_{\mathbf{C}} E_b(\mathbf{C}) \quad (11)$$

$$\text{s.t. } \mathbf{C} \succeq \mathbf{0}, \quad (12a)$$

$$P_{i,n}(\mathbf{C}) \leq P_{i,n}^{\text{max}}, i \in \mathcal{N}, n \in \{1, \dots, t\}, \quad (12b)$$

i.e. finding the optimal energy-efficient joint resource allocation for N -hop MIMO-AF systems in the unconstrained as well as power constraint cases, where $P_{i,n}(\mathbf{C})$ and $P_{i,n}^{\text{max}}$ are respectively the transmit and maximum transmit powers per antenna of node i . We also provide alternative sub-optimal procedures for solving this problem in a simplified manner.

Proposition 1: The function E_b in (10) can be lower bounded as

$$E_b(\mathbf{C}) \geq \underline{E}_b(\mathbf{C}) = \frac{P_c + \sum_{m=1}^M A_m^{-1} (2^{C_m} - 1)}{W \sum_{m=1}^M C_m}, \quad (13)$$

for any $\mathbf{C} \succeq \mathbf{0}$; see section B of the Appendix for the proof.

In (13), $A_m = \left(\sum_{i=1}^N \sqrt{A_{i,m}} \right)^{-2}$,

$$C_m = \sum_{i=1}^N C_{i,m} - \log_2 \left(\prod_{i=1}^N 2^{C_{i,m}} - \prod_{i=1}^N (2^{C_{i,m}} - 1) \right) \quad (14)$$

is the end-to-end SE per subchannel, and conversely,

$$\prod_{i=1}^N (1 - 2^{-C_{i,m}}) = 1 - 2^{-C_m}. \quad (15)$$

Note that the formulation of (13) is equivalent to the formulation of a classic MIMO system with CSI [35]. Contrary to

Algorithm 1 E_b^- : Lower-bound of E_b^*

-
- 1: **Inputs:** M, P_c, W , and A_m , for any $m \in \mathcal{M}$
 - 2: Set $x = 0$, $E_b^- = \max_{m \in \mathcal{M}} A_m^{-1}$, and $\epsilon = 10^{-6}$;
 - 3: **while** ($|E_b^- - x| \geq \epsilon$) **do**
 - 4: Obtain C_m^\bullet by inserting E_b^- into (19), for any $m \in \mathcal{M}$;
 - 5: Set $x = E_b^-$;
 - 6: Update $E_b^- = \underline{E}_b(\mathbf{C}^\bullet)$ by inserting C_m^\bullet into (13).
 - 7: **end while**
 - 8: **Outputs:** E_b^- and $C_m^\bullet, \forall m \in \mathcal{M}$.
-

(10), \underline{E}_b in (13) is a convex function (see proof in [35]), such that

$$\begin{aligned} E_b^- &= \min_{\mathbf{C}} \underline{E}_b(\mathbf{C}) \\ &\text{s.t. constraints in (12)} \end{aligned} \quad (16)$$

is a convex optimization problem with a unique solution that occurs at $\mathbf{C} = \mathbf{C}^\bullet$.

Corollary 1: According to Proposition 1 and (11), $E_b^- = \underline{E}_b(\mathbf{C} = \mathbf{C}^\bullet) \leq E_b^* = E_b(\mathbf{C} = \mathbf{C}^*) \leq E_b^+ = E_b(\mathbf{C} = \mathbf{C}^\bullet)$.

Proof: Since $\underline{E}_b(\mathbf{C}) \leq E_b(\mathbf{C})$, for any $\mathbf{C} \succeq \mathbf{0}$, and both functions are continuous, it implies that $\min_{\mathbf{C}} \underline{E}_b(\mathbf{C}) \leq \min_{\mathbf{C}} E_b(\mathbf{C})$, i.e. $E_b^- \leq E_b^*$. On the other hand, since $E_b^* = E_b(\mathbf{C} = \mathbf{C}^*)$ is the global minimum of E_b , then $E_b(\mathbf{C}^*) \leq E_b(\mathbf{C}^\bullet)$, i.e. $E_b^* \leq E_b^+$, holds for any $\mathbf{C}^\bullet \succeq \mathbf{0}$. ■

Corollary 2: The value of E_b at $\mathbf{C} = \mathbf{C}^\bullet$, $E_b^+ = E_b(\mathbf{C} = \mathbf{C}^\bullet)$, can be obtained by solving the following convex optimization problem (in abstract form)

$$\begin{aligned} \min_{\mathbf{C}} & \sum_{i=1}^N \sum_{m=1}^M A_{i,m} (2^{C_{i,m}} - 1) \\ \text{s.t.} & \prod_{i=1}^N (1 - 2^{-C_{i,m}}) = 1 - 2^{-C_m^\bullet}, \text{ if } C_m^\bullet > 0, \forall m \in \mathcal{M}, \\ & C_{i,m} = 0, \text{ if } C_m^\bullet = 0, \forall m \in \mathcal{M}, \end{aligned} \quad (17)$$

where C_m^\bullet is first obtained by solving (16) via Algorithm 1. According to (17), we obtain

$$C_{i,m}^\bullet = \begin{cases} -1 + \log_2 \left(1 + \sqrt{1 + 4A_{i,m}^{-1} \mu_m^\bullet} \right) & \text{if } C_m^\bullet > 0, \forall m \in \mathcal{M}, \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

where μ_m^\bullet can be obtained in a low-complexity and optimal manner by means of the Newton-Raphson method [36], for any $m \in \mathcal{M} = \{1, \dots, M\}$, as it is detailed in Algorithm 2; the initial value of μ_m^\bullet at line 5 of Algorithm 2 is based on the following inequality $A_m^{-1} (2^{C_m^\bullet} - 1) \leq \sum_{i=1}^N A_{i,m} (2^{C_{i,m}^\bullet} - 1) \leq \sum_{i=1}^N \sqrt{A_{i,m} \mu_m^\bullet}$, such that $\mu_m^\bullet \geq A_m^{-1} (2^{C_m^\bullet} - 1)^2$.

A. Unconstrained EE Optimization

EE optimization is a generalization of both sum-rate maximization and transmit power minimization [35]. Hence, enforcing rate or power constraints on EE provides either a power or sum-rate optimal solution, which is however suboptimal in terms of EE. The sole EE-optimal solution is the optimal unconstrained EE solution [35] and, consequently, we first aim at finding a solution to the problem in (11) for the unconstrained scenario, i.e. with constraint (12a) only.

Algorithm 2 E_b^+ : Upper-bound of E_b^* (Unconstrained)

-
- 1: **Inputs:** $N, M, P_c, W, C_m^\bullet$ and $A_{i,m}, \forall i \in \mathcal{N}$ and $m \in \mathcal{M}$;
 - 2: Set $C_{i,m}^\bullet = 0$, for any $i \in \mathcal{N}$ and $m \in \mathcal{M}$, and $\epsilon = 10^{-6}$;
 - 3: **for** $m = 1 : M$ **do**
 - 4: **if** $C_m^\bullet > 0$ **then**
 - 5: Set $\mu_m^\bullet = A_m^{-1} (2^{C_m^\bullet} - 1)^2$;
 - 6: Compute $f = \sum_{i=1}^N \ln \left(\sqrt{1 + 4A_{i,m}^{-1} \mu_m^\bullet} - 1 \right) - \ln \left(\sqrt{1 + 4A_{i,m}^{-1} \mu_m^\bullet} + 1 \right) - \ln \left(1 - 2^{-C_m^\bullet} \right)$;
 - 7: **while** $|f| > \epsilon$ **do**
 - 8: Compute $\partial f = \sum_{i=1}^N 2A_{i,m}^{-1} (1 + 4A_{i,m}^{-1} \mu_m^\bullet)^{-\frac{1}{2}} \times \left[\left(\sqrt{1 + 4A_{i,m}^{-1} \mu_m^\bullet} - 1 \right)^{-1} - \left(\sqrt{1 + 4A_{i,m}^{-1} \mu_m^\bullet} + 1 \right)^{-1} \right]$;
 - 9: Set $\mu_m^\bullet = \mu_m^\bullet - f/\partial f$;
 - 10: Compute f as in line 6;
 - 11: **end while**
 - 12: Obtain $C_{i,m}^\bullet$, for any $i \in \mathcal{N}$, by inserting μ_m^\bullet into (18);
 - 13: **end if**
 - 14: **end for**
 - 15: Obtain E_b^+ by inserting $C_{i,m}^\bullet$ into (10);
 - 16: **Outputs:** E_b^+ and $C_{i,m}^\bullet, \forall i \in \mathcal{N} \ \& \ m \in \mathcal{M}$.
-

1) *Lower and upper bounds:* In this scenario, C_m^\bullet can be expressed in closed-form as

$$C_m^\bullet = [\log_2(W E_b^- A_m / \ln(2))]_+, \quad (19)$$

where $[x]_+ = \max\{x, 0\}$. Thus, the problem in (16) for getting a lower bound of E_b^* can be solved in a low-complexity and optimal manner by using a unidimensional search based on the Dinkelbach method [37] (see Algorithm 1). Moreover, an upper bound of E_b^* can then be obtained via Algorithm 2, by using the values of C_m^\bullet returned via Algorithm 1.

2) *Optimal solution for (11):*

Proposition 2: Let \mathbf{C}^* be a stationary point of $E_b(\mathbf{C})$, i.e. $\nabla E_b(\mathbf{C} = \mathbf{C}^*) = \mathbf{0}$, then $E_b(\mathbf{C}^*)$ is a strict local minimum of E_b , i.e. $E_b(\mathbf{C}^* + \mathbf{z}) > E_b(\mathbf{C}^*)$ for $\mathbf{z} \rightarrow \mathbf{0}$ and $\mathbf{z} \neq \mathbf{0}$; see section C of the Appendix for the proof. Note that local extrema of differentiable functions can only occur at stationary points (see p.194 of [38]). However, stationary points are not necessarily local extrema (see p.195 of [38]), which highlights the importance of Proposition 2.

Proposition 3: According to Corollary 1, (11) has at least one solution in the unconstrained scenario, i.e. it has at least one local minimum (occurring at a stationary point); and according to Proposition 2, this local minimum is a strict local minimum. Let E_b^* be a strict minimum of E_b occurring at $\mathbf{C} = \mathbf{C}^*$ then it can be expressed as

$$E_b^* = \frac{\ln(2) A_{i,m} 2^{C_{i,m}^*} (2^{C_{i,m}^*} - 1)}{W (2^{C_m^*} - 1)}, \quad (20)$$

for any $i \in \mathcal{N}$ and $m \in \mathcal{M}^*$, where $\mathcal{M}^* = \{m \in \mathcal{M} | C_m^* > 0\}$ is a set of allocated subchannel indices; see section D of the Appendix for the proof of this proposition. Consequently, $C_{i,m}^*$ can be expressed as a function of E_b^* and C_m^* such that

$$C_{i,m}^* = \left[-1 + \log_2 \left(1 + \sqrt{1 + \frac{4W (2^{C_m^*} - 1) E_b^*}{\ln(2) A_{i,m}}} \right) \right]_+, \quad (21)$$

for any $i \in \mathcal{N}$ and $m \in \mathcal{M}$.

```

1: function  $C_{\text{UNC}}^*(N, M, U, W, E_b^*, A_{i,m})$ 
2:   Set  $C_0^* = 1$ , and  $C_m^* = 0$ , for any  $m \in \mathcal{M}$ ;
3:   Set  $m = 1$  and  $\epsilon = 10^{-6}$ ;
4:   while ( $C_{m-1}^* > 0$ ) and ( $m \leq U$ ) do
5:     Compute  $X_m^{\max}$  in (23);
6:     Set  $X_m = X_m^{\max}$ , and obtain  $F$  in (22);
7:     while  $|F| > \epsilon$  do
8:       Compute  $\partial F$  in (24);
9:       Set  $X_m = \min\{[X_m - F/\partial F]_+, X_m^{\max}\}$ ;
10:      Compute  $F$  in (22);
11:    end while
12:    Set  $C_m^* = [\log_2(1 + X_m^2)]_+$  and  $m = m + 1$ ;
13:  end while
14:  Obtain  $C_{i,m}^*$  by inserting  $E_b^*$  and  $C_m^*$  into (21), for any  $i \in \mathcal{N}$ ;
15:  return  $\mathcal{C}^*$ ;
16: end function

```

Corollary 3: According to (20) and (21), C_m^* can be obtained by solving $F(X_m) = 0$, for any $m \in \mathcal{M}^*$, where

$$F(X_m) = \prod_{i=1}^N \left(\sqrt{a_{i,m}^*} + \sqrt{a_{i,m}^* + X_m^2} \right) - X_m^{N-1} \sqrt{X_m^2 + 1}, \quad (22)$$

$a_{i,m}^* = \frac{\ln(2)A_{i,m}}{4WE_b^*}$, and $X_m = \sqrt{2C_m^* - 1}$. In turn, any root of F in (22) belongs to $[0, X_m^{\max}]$, where X_m^{\max} is given by

$$X_m^{\max} = \frac{1 - 2 \sum_{i=1}^N \sum_{k=i+1}^N \sqrt{a_{i,m}^* a_{k,m}^*}}{2 \sum_{i=1}^N \sqrt{a_{i,m}^*}}. \quad (23)$$

See section E of the Appendix for the proof of this corollary.

As a result of Proposition 3 and Corollary 3, we can obtain $C_{i,m}^*$, for any $i \in \mathcal{N}$ and $m \in \mathcal{M}$, in a low-complexity (based on the Newton-Raphson method [36]) by using the function “ C_{UNC}^* ” that is detailed at the top of this page, with

$$\begin{aligned} \partial F = X_m \prod_{i=1}^N \left(\sqrt{a_{i,m}^*} + \sqrt{a_{i,m}^* + X_m^2} \right) \sum_{i=1}^N \left(\sqrt{a_{i,m}^* + a_{i,m}^* X_m^2} \right. \\ \left. + a_{i,m}^* + X_m^2 \right)^{-1} - X_m^{N-2} \frac{NX_m^2 + N - 1}{\sqrt{X_m^2 + 1}}. \end{aligned} \quad (24)$$

Corollary 4: Let $E_b^* = E_b(\mathcal{C}^* = \mathcal{C}^*)$ be the minimum of the strict local minima, i.e. the global minimum, it then can also be expressed as in (20) such that C_m^* can be obtained via (22).

Even though E_b in (10) is not a convex function, we can obtain the optimal solution to the problem in (11), i.e. E_b^* , by using Corollaries 1 and 4; on the one hand, since $E_b^- \leq E_b^* \leq E_b^+$ in Corollary 1, it exists a unique $\alpha \in [0, 1]$ such that

$$\begin{aligned} E_b^* = E_b(\mathcal{C}^*, \alpha) &= \underline{E}_b(\mathcal{C}^*) + \alpha(E_b(\mathcal{C}^*) - \underline{E}_b(\mathcal{C}^*)) \\ &= E_b^- + \alpha(E_b^+ - E_b^-), \end{aligned} \quad (25)$$

with $E_b^+ - E_b^- \leq \frac{\sum_{i=1}^N \sum_{k=i+1}^N \sqrt{A_{i,m} A_{k,m}}}{W \sum_{m=1}^M C_m^*}$. On the other hand, Corollary 4 states that C_m^* can be obtained via (22) when $E_b^* = E_b^*$. Accordingly, the following procedure can be used to find E_b^* with a precision of ϵ (see details in Algorithm 3):

E_b^- and E_b^+ are first obtained by using Algorithms 1 and 2, respectively; then $E_b(\mathcal{C}^*, \alpha)$ is computed via (25), where $\alpha = 0$ is chosen as a starting value; next, C_m^* is obtained by solving (22) for $E_b = E_b(\mathcal{C}^*, \alpha)$ and, in turn, $C_{i,m}^*$ is calculated through (21). Note that (20) is a function of \mathcal{M}^* and, hence,

Algorithm 3 E_b^* : Optimal solution

```

1: Inputs:  $N, M, P_c, W, E_b^-, E_b^+$ , and  $A_{i,m}$ , for any  $i \in \mathcal{N}$  and  $m \in \mathcal{M}$ 
2: Set  $A_m = \left( \sum_{i=1}^N \sqrt{A_{i,m}} \right)^{-2}$ ,  $\forall m \in \mathcal{M}$ ;
3: Set  $\alpha_{\min} = 0$ ,  $\alpha_{\max} = 1$ ,  $\eta = 1$ ,  $u = 4$ ,  $v = 2$ ,  $V = 10$ ,  $\epsilon = 10^{-6}$ ;
4: while ( $\eta > \epsilon$ ) and ( $v < V$ ) do
5:    $\alpha = \alpha_{\min}$ ;
6:   while ( $\eta > 0$ ) and ( $\alpha \leq \alpha_{\max}$ ) do
7:     Set  $E_b^* = E_b^- + \alpha(E_b^+ - E_b^-)$ , based on (25);
8:     Obtain  $\mathcal{C}^*$  via  $C_{\text{UNC}}^*$  for  $U = M$ ;
9:     Set  $\mathcal{C}^m$  by using (26) for  $\mathcal{C} = \mathcal{C}^*$ ,  $\forall m \in \mathcal{M}$ ;
10:    Obtain  $m^* = \arg \min_{m \in \mathcal{M}} \{E_b(\mathcal{C}^m)\}$ ,  $E_b$  in (10);
11:    Set  $\mathcal{C}^* = \mathcal{C}^{m^*}$ ,  $E_b^* = E_b(\mathcal{C}^*)$  and  $\eta = E_b^* - E_b^-$ ;
12:    Set  $\alpha = \alpha + 10^{-v}$ ;
13:  end while
14:  Set  $\alpha_{\min} = [\alpha - 10^{-v} - 10^{1-v}/u]_+$ ,  $\alpha_{\max} = \min\{\alpha - 10^{-v} + 10^{1-v}/u, 1\}$  and  $v = v + 1$ ;
15: end while
16: Outputs:  $E_b^*$  and  $\mathcal{C}^*$ .

```

E_b^* can only be obtained for $m \in \mathcal{M}^* = \{m \in \mathcal{M} | C_m^* > 0\}$; \mathcal{M}^* is obtained by using $E_b^* = \min_{m \in \mathcal{M}} E_b(\mathcal{C}^m)$, such that \mathcal{M}^* contains only the index of subchannels minimizing (10). If the subchannels are sorted in descending order according to their strength, it is only necessary to find $m^* \in \mathcal{M}$ such that $E_b(\mathcal{C}^{m^*-1}) \geq E_b(\mathcal{C}^{m^*}) \leq E_b(\mathcal{C}^{m^*+1})$, where

$$\begin{aligned} \mathcal{C}^m = [C_{1,1}, \dots, C_{1,m}, 0, \dots, 0, C_{2,1}, \dots, C_{2,m}, 0, \dots, \\ 0, \dots, C_{N,1}, \dots, C_{N,m}, 0, \dots, 0]. \end{aligned} \quad (26)$$

Next, E_b^* is computed by setting $\mathcal{C}^* = \mathcal{C}^{m^*}$ and inserting it into (10). As long as $E_b^* - E_b(\mathcal{C}^*, \alpha) > \epsilon$, $\epsilon \ll 1$, the same procedure is repeated for an incremented α . Since $E_b(\mathcal{C}^*, 0) \leq E_b^*$ and \mathcal{C}^* is unique when E_b^* is a strict minimum (Propositions 2 and 3), it ensures that the first point \mathcal{C}^* verifying (20) is the lowest of the minimum points, i.e. $\mathcal{C}^* = \mathcal{C}^*$.

3) *Low-complexity solution:* Algorithm 3 returns the optimal solution to the problem in (11) with a precision of ϵ ; however, its computational complexity is high since it is based on an exhaustive search method. A low-complexity alternative to Algorithm 3 is to use the Dinkelbach approach, as in Algorithm 1; however, this approach is not necessarily optimal since E_b may have more than one minimum. Given that E_b^* is obtained for $m \in \mathcal{M}^*$, the likelihood of finding E_b^* can be increased by repeating the Dinkelbach method at least once to ensure that $E_b(\mathcal{C}^{m^*}) \leq E_b(\mathcal{C}^{m^*-1})$, as it is further detailed in Algorithm 4.

B. Transmit Power Constrained EE Optimization

In the power constraint scenario, the optimization problem is defined as in (11) but with both constraints (12a) and (12b).

Whenever the transmit power per antenna of node i

$$P_{i,n}(\mathcal{C}) = \Delta_i^{-1} \sum_{k=1}^K A_{i,(n-1)K+k} (2^{C_{i,(n-1)K+k}} - 1)$$

is strictly lower than $P_{i,n}^{\max}$, for $\mathcal{C} = \mathcal{C}^*$ and any $i \in \mathcal{N}$ as well as $n \in \{1, \dots, t\}$, this problem is equivalent to the unconstrained EE optimization. At the other extreme, whenever all the power constraints are enforced, the optimization

Algorithm 4 \tilde{E}_b^* : Low-complexity solution

1: **Inputs:** $N, M, P_c, W, E_b^-,$ and $A_{i,m},$ for any $i \in \mathcal{N}$ and $m \in \mathcal{M}$
 2: Set $A_m = \left(\sum_{i=1}^N \sqrt{A_{i,m}} \right)^{-2}, \forall m \in \mathcal{M};$
 3: Set $\eta = 0, U = M, E_b^* = E_b^-, \tilde{E}_b^* = 10^{50},$ and $\epsilon = 10^{-6};$
 4: **while** ($\eta == 0$) and ($U > 0$) **do**
 5: Obtain \mathbf{C}^* via $C_{\text{UNC}}^*;$
 6: Set $x = E_b^*;$
 7: Obtain E_b^* by inserting \mathbf{C}^* into (10);
 8: **if** $|E_b^* - x| \leq \epsilon$ **then**
 9: **if** $E_b^* \geq \tilde{E}_b^*$ **then**
 10: Set $\eta = 1;$
 11: **else**
 12: **if** $\tilde{E}_b^* == 10^{50}$ **then** $U = \sum_{m=1}^M (C_{1,m}^* > 0);$
 13: Set $\tilde{E}_b^* = E_b^*$ and $\tilde{\mathbf{C}}^* = \mathbf{C}^*;$
 14: Set $U = U - 1;$
 15: **end if**
 16: **end if**
 17: **end while**
 18: **Outputs:** \tilde{E}_b^* and $\tilde{\mathbf{C}}^*.$

problem of (11) subject to (12a) and (12b) reverts to a sum-rate maximization problem.

1) *Lower and upper bounds:* In the intermediate case, i.e. when some power constraints are enforced on some antennas, some elements of (10) would become fixed while some others would remain variables. For instance, the total consumed power in the nominator of (10) would be re-expressed as

$$P_{\Sigma}(\mathbf{C}) = \hat{P}_c + \sum_{n=1}^t \sum_{i \in \underline{\mathcal{N}}_n} P_{i,n}(\mathbf{C}),$$

where $\hat{P}_c = P_c + \sum_{n=1}^t \sum_{i \in \underline{\mathcal{N}}_n} P_{i,n}^{\max}.$ In addition, $\underline{\mathcal{N}}_n = \{i \in \mathcal{N} | P_{i,n}(\mathbf{C}) < P_{i,n}^{\max}\}$ and $\overline{\mathcal{N}}_n = \{i \in \mathcal{N} | P_{i,n}(\mathbf{C}) \geq P_{i,n}^{\max}\}$ are sets of node indices for which the n -th antenna transmit power is unconstrained or power constrained, respectively. Given that $P_{i,n}(\mathbf{C})$ only depends of the elements $\mathcal{C}_{i,(n-1)K+[1,\dots,K]}$ of $\mathbf{C},$ the problem in (11) becomes equivalent to finding the optimal unconstrained elements of \mathbf{C} knowing the constrained elements. Modifying the constrained elements without updating the unconstrained elements is clearly not optimal; hence, a simple upper bound for the power constrained case can be obtained by computing the constrained elements of \mathbf{C} (by solving a classic water-filling problem) while keeping unchanged the unconstrained elements of $\mathbf{C},$ which results from the unconstrained search, as it is detailed in Algorithm 5. Note that the $\mathbf{C}_{i,m}^{\bullet}$ values in line 1 of Algorithm 5 need first to be obtained from Algorithm 2.

As for the lower bound, it is obvious that the optimal E_b obtained from the unconstrained search is always lower than any constrained E_b and, thus, E_b^- is also a lower bound (but a looser bound) for the power constrained case.

2) *Optimal solution for the flat channel case:* As it has previously been mentioned, in the power constrained scenario, the problem in (11) becomes equivalent to finding the optimal unconstrained elements of \mathbf{C} knowing the constrained ones. For a given antenna $n,$ if $i \in \underline{\mathcal{N}}_n,$ then relation (20) holds where $C_{i,m}^*$ is replaced by $C_{i,(n-1)K+k}^*$ and, hence, $C_{i,(n-1)K+k}^*, \forall k \in [1, \dots, K],$ can be expressed as in (21);

Algorithm 5 $E_b^+:$ Upper-bound of E_b^* (power constrained)

1: **Inputs:** $N, t, K, P_c, W, \Delta_i, P_{i,n}^{\max}, \mathbf{C}_{i,m}^{\bullet}$ and $A_{i,m}, \forall i \in \mathcal{N}$ and $m \in \mathcal{M};$
 2: **for** $i = 1 : N$ **do**
 3: **for** $n = 1 : t$ **do**
 4: Set $\gamma_k = \Delta_i A_{i,(n-1)K+k}^{-1},$ for any $k \in [1, \dots, K];$
 5: Set $P = \sum_{k=1}^K \gamma_k^{-1} \left(2^{C_{i,(n-1)K+k}^{\bullet}} - 1 \right)$
 6: **if** $P > P_{i,n}^{\max}$ **then**
 7: Obtain \mathbf{p} via water-filling by knowing $\gamma, p_k = [\nu - \gamma_k^{-1}]_+$ (ν : water level);
 8: Set $\mathbf{C}_{i,(n-1)K+k}^{\bullet} = \log_2(1 + \gamma_k p_k), \forall k \in [1, \dots, K];$
 9: **end if**
 10: **end for**
 11: **end for**
 12: Obtain E_b^+ by inserting $\mathbf{C}_{i,(n-1)K+k}^{\bullet}$ into (10);
 13: **Outputs:** E_b^+

otherwise, if $i \in \overline{\mathcal{N}}_n,$ then $P_{i,n}(\mathbf{C}^*)$ is equal to $P_{i,n}^{\max}$ such that the power-constrained elements of \mathbf{C}^* are expressed as

$$C_{i,m}^* = \log_2 \left(1 + \Delta_i P_{i,m}^{\max} A_{i,m}^{-1} \right), \quad (27)$$

in the multi-hop MIMO-AF case with flat channel condition, i.e. $K = 1$ and $m = n.$ Then, knowing the constrained $C_{i,m}^*$ values from (27), C_m^* as well as the unconstrained $C_{i,m}^*$ values can be obtained by inserting (27) into (21) such that (22) is modified as

$$F(X_m) = \prod_{i \in \underline{\mathcal{N}}_m} \left(\sqrt{a_{i,m}^*} + \sqrt{a_{i,m}^* + X_m^2} \right) - B X_m^{N - |\overline{\mathcal{N}}_m| - 1} \times \sqrt{X_m^2 + 1}, \quad (28)$$

in the power constrained scenario when $K = 1,$ and where $B = \sqrt{\prod_{i \in \overline{\mathcal{N}}_m} 1 - 2^{-C_{i,m}^*}}$ can be obtained via (27). In addition, $|\overline{\mathcal{N}}_m|$ denotes the cardinal of $\overline{\mathcal{N}}_m,$ i.e. its number of elements. In turn, any root of F in (28) belongs to $[0, X_m^{\max}],$ where X_m^{\max} is given by $X_m^{\max} =$

$$\begin{cases} -\frac{\sqrt{a_{i,m}^*}}{1 - B^2} + \sqrt{\frac{a_{i,m}^*}{(1 - B^2)^2} + \frac{B^2 - a_{i,m}^*}{1 - B^2}}, i \in \underline{\mathcal{N}}_m, \text{ if } |\underline{\mathcal{N}}_m| = 1; \\ -\frac{\sum_{i \in \underline{\mathcal{N}}_m} \sqrt{a_{i,m}^*}}{2(1 - B)} + \left[\frac{B - 2 \sum_{i \in \underline{\mathcal{N}}_m} \sum_{\substack{k \in \underline{\mathcal{N}}_m \\ k > i}} \sqrt{a_{i,m}^* a_{k,m}^*}}{2(1 - B)} \right. \\ \left. + \left(\frac{\sum_{i \in \underline{\mathcal{N}}_m} \sqrt{a_{i,m}^*}}{2(1 - B)} \right)^2 \right]^{\frac{1}{2}}, \text{ if } |\underline{\mathcal{N}}_m| > 1. \end{cases} \quad (29)$$

Note that if $|\underline{\mathcal{N}}_m| = 0,$ C_m^* can directly be obtained by inserting $C_{i,m}^*$ in (27), $\forall i \in \mathcal{N},$ into (14). Hence, we can obtain $C_m^*,$ for any $i \in \mathcal{N}$ and $m \in \mathcal{M},$ in the transmit power constraint case (for $K = 1$) by using the function “ C_{PWC}^* ” that is detailed thereafter, where ∂F is given by $\partial F =$

$$X_m \prod_{i \in \underline{\mathcal{N}}_m} \left(\sqrt{a_{i,m}^*} + \sqrt{a_{i,m}^* + X_m^2} \right) \sum_{i \in \underline{\mathcal{N}}_m} \left(\sqrt{a_{i,m}^* + a_{i,m}^* X_m^2} + a_{i,m}^* + X_m^2 \right)^{-1} - B \frac{(N - |\overline{\mathcal{N}}_m|) X_m^2 + N - |\overline{\mathcal{N}}_m| - 1}{X_m^{-N + |\overline{\mathcal{N}}_m| + 2} \sqrt{X_m^2 + 1}}. \quad (30)$$

```

1: function  $C_{\text{PWC}}^*(N, t, U, W, E_b^*, A_{i,m})$ 
2:   Set  $M = t$ ,  $C_0^* = 1$ , and  $C_m^* = 0$ , for any  $m \in \mathcal{M}$ ;
3:   Set  $m = 1$  and  $\epsilon = 10^{-6}$ ;
4:   while  $(C_{m-1}^* > 0)$  and  $(m \leq U)$  do
5:      $\underline{\mathcal{N}}_m = \mathcal{N}$ ,  $\overline{\mathcal{N}}_m = \emptyset$ ,  $k = 0$ ,  $\eta = 2$ ;
6:     while  $\eta > 1$  do
7:       if  $k > 0$  then Add  $k$  to  $\overline{\mathcal{N}}_m$  and remove it from  $\underline{\mathcal{N}}_m$ ;
8:       if  $\underline{\mathcal{N}}_m == \emptyset$  then ▷ fully constrained
9:         Obtain  $C_{i,m}^*$  via (27),  $\forall i \in \mathcal{N}$ ;
10:        Set  $C_m^* = -\log_2(1 - \prod_{i=1}^N 1 - 2^{-C_{i,m}^*})$ ;
11:       else
12:         if  $\overline{\mathcal{N}}_m == \emptyset$  then ▷ unconstrained
13:           Set  $B = 1$  and  $X_m^{\max}$  via (23);
14:         else
15:           Obtain  $C_{i,m}^*$  via (27),  $\forall i \in \overline{\mathcal{N}}_m$ ;
16:           Set  $B = \sqrt{\prod_{i \in \overline{\mathcal{N}}_m} 1 - 2^{-C_{i,m}^*}}$  and  $X_m^{\max}$  via
17:             (29);
18:         end if
19:         Same as lines 6 to 12 of " $C_{\text{UNC}}^*$ ", but where  $F$  and
20:          $\partial F$  are given in (28) and (30), respectively.
21:       end if
22:       if  $C_m^* == 0$  then  $C_{i,m}^* = 0$ ,  $\forall i \in \mathcal{N}$ ; else Obtain
23:          $C_{i,m}^*$  by inserting  $E_b^*$  and  $C_m^*$  into (21),  $\forall i \in \underline{\mathcal{N}}_m$ ;
24:         Compute  $\chi_i = P_{i,m}(C^*)/P_{i,m}^{\max}$ ,  $\forall i \in \mathcal{N}$ ;
25:          $\eta = \max_i \{\chi\}$  and  $k = \arg \max_i \{\chi\}$ ;
26:       end while
27:       Set  $m = m + 1$ ;
28:     end while
29:     return  $C^*$ .
30: end function

```

The optimum solution for (11) in the flat channel scenario with power constraint can finally be obtained by using Algorithm 3, but where E_b^+ is computed via Algorithm 5 and by using the function " C_{PWC}^* " instead of " C_{UNC}^* " for getting C^* .

3) *Low-complexity solution*: Similarly to the optimal case, Algorithm 4 in conjunction with the function " C_{PWC}^* ", instead of " C_{UNC}^* ", can be used for obtaining a low-complexity solution to the problem in (11) with constraints (12a) and (12b).

C. Energy-efficient multi-hop MIMO-AF procedure

As it has been previously mentioned, the sole EE-optimal solution is the optimal unconstrained EE solution; consequently, the various algorithms developed in this paper must be sequentially utilized for solving (11) subject to the constraints in (12), as it is depicted in Fig. 2. Algorithm 1 is first used to obtain E_b^- and C_m^* , since the former is an input of both Algorithms 3 and 4, and the latter is an input of Algorithm 2. In order to obtain the sub-optimal solution in Algorithm 4, the function " C_{PWC}^* " is utilized, which is a generalization of " C_{UNC}^* " that can be used for both the unconstrained and power constraint cases. Whereas, in order to obtain the optimal solution in Algorithm 3, Algorithm 2 is used to compute E_b^+ , which is further refined in Algorithm 5 if some per-antenna power constraints are not met.

From an implementation perspective, similar to cooperative multi-point (CoMP) communication, a central processing unit needs to first collect the CSI of all the channels/subchannels, and then performs the joint optimization (by using our algorithms) before sending back the results of this optimization process to each relevant transmitting node.

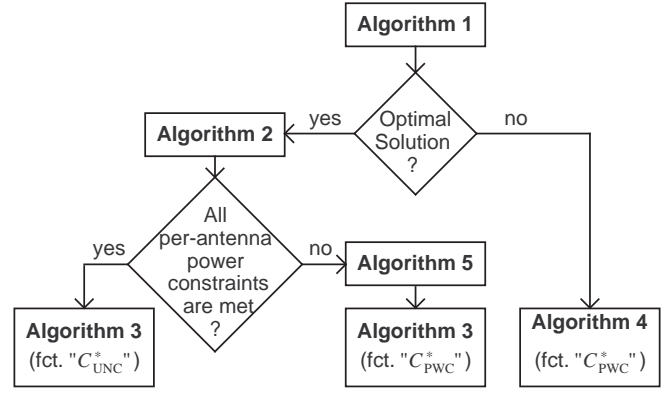


Fig. 2: Energy-efficient multi-hop MIMO-AF procedure flowchart.

TABLE I: Power parameter values

Parameters	Δ	P_{CipA} (W)	P_{Ci} (W)	P_{Si} (W)
BS	4.7 [5]	100 [5]*	180 [5]*	75 [5]
RN	6.3 [32]	4 [32]*	4.9 [32]*	3.45 [32]*
UE	—	0.03 [24]*	0.07 [24]*	0.02 [24]*

IV. NUMERICAL RESULTS AND DISCUSSION

In order to demonstrate the reliability of our various algorithms for jointly optimizing the resource usage of multiple nodes in an energy-efficient manner, we compare their results, averaged over 1000 runs. In our simulations, we assume a downlink transmission of the N -hop MIMO AF system, such that the SN is a BS and the DN is a UE, and utilize the power model parameters of Table I accordingly, when considering a MIMO flat or frequency selective Rayleigh fading channel between each node, $W = 1$, $\varsigma = 1/2$, and $\sigma_1^2 = [-20 : 20]$ dB. In table I, $[\#]^*$ indicates that these values have been extrapolated from reference $[\#]$, since all the references in Table I do not differentiate between antenna dependent and independent circuit powers. In addition, 24 different settings, which are listed in Table II, of the parameters N (number of hops), t (number of antennas), K (number of subchannels), $P_{i,n}^{\max}$ (per-antenna constraint at each node), and σ_i^2 , $\forall i \in \{2, \dots, N\}$ are simulated. For simplicity reason, we consider that all $P_{i,n}^{\max}$ are equal $\forall i \in \mathcal{N}$, $n \in \{1, \dots, t\}$. We also assume, as in [8], [17], that the eigenvalues $\lambda_{i,m}$ are sorted in descending order for each link prior to run our algorithms.

A. Accuracy results and discussion

In Figs. 3 and 4, we compare the energy-per-bit performance of our lower bound, upper bound, optimal and sub-optimal approaches as a function of the noise power at RN_1 in the unconstrained and power constrained scenarios, respectively, for the various settings of Table II. Note that the lower bound results have been omitted in Fig. 4 since they are the same as in Fig. 3. The results first show a good match between our sub-optimal and optimal approaches in both figures, which emphasizes that the sub-optimal approach can achieve near-optimal performance. Moreover, they confirm that the performance of

TABLE II: Simulation settings

Settings	A	B	C	D	E	F	G	H	I	J	K	L	M
N	3	7	11	3	3	3	3	3	3	3	3	3	3
t	4	4	4	4	4	4	4	4	16	4	64	4	256
K	4	4	4	4	4	4	4	16	4	64	4	256	4
σ_i^2 (dB)	[0, 0]	[0, ..., 0]	[0, ..., 0]	[-20, -20]	[-20, 20]	[20, -20]	[20, 20]	[0, 0]	[0, 0]	[0, 0]	[0, 0]	[0, 0]	[0, 0]
$P_{i,n}^{\max}$ (W)	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞

Settings	N	O	P	Q	R	S	T	U	V	W	X
N	3	7	11	3	3	3	3	3	3	3	3
t	4	4	4	4	4	4	4	4	4	64	256
K	1	1	1	1	1	1	1	1	1	1	1
σ_i^2 (dB)	[0, 0]	[0, ..., 0]	[0, ..., 0]	[-20, -20]	[-20, 20]	[20, -20]	[20, 20]	[0, 0]	[0, 0]	[0, 0]	[0, 0]
$P_{i,n}^{\max}$ (W)	20	20	20	20	20	20	20	20	5	10	20

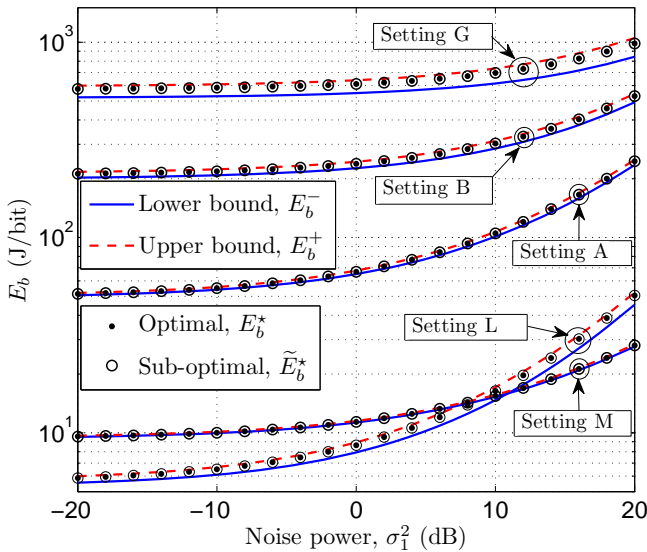


Fig. 3: Comparison of the energy-per-bit performances of our lower bound, upper bound, optimal and sub-optimal approaches in the unconstrained scenario for various settings.

these two approaches are bounded by E_b^- and E_b^+ , as it has been explained in Sections III-A1 and III-B1. It can also be remarked that the upper bound, E_b^+ , is quite a tight bound of the optimal E_b , E_b^* . This is due to the fact that C_m^* converges towards C_m^\bullet in (19) when $C_m^* \gg \max_{i \in \mathcal{N}} \{A_{i,m}\} (E_b^*)^{-1}$.

Regarding the variation of E_b in relation to the various settings, it can be seen from Figs. 3 and 4 that increasing the number of nodes (when comparing setting A with B in Fig. 3 and setting N and O in Fig. 4), increases the energy-per-bit consumption. Indeed, adding more nodes increases the power consumption which in turn increases the energy consumption, unless it helps to improve the channel quality between the nodes. Moreover, increasing the number of antennas (when comparing setting A with M in Fig. 3 and settings N with W/X in Fig. 4) reduces the energy consumption; indeed, having more antennas helps to improve the channel quality, which in turn decrease the energy consumption. Similarly, increasing the number of frequency flat subchannels (when comparing setting A with L in Fig. 3) reduces the energy consumption.

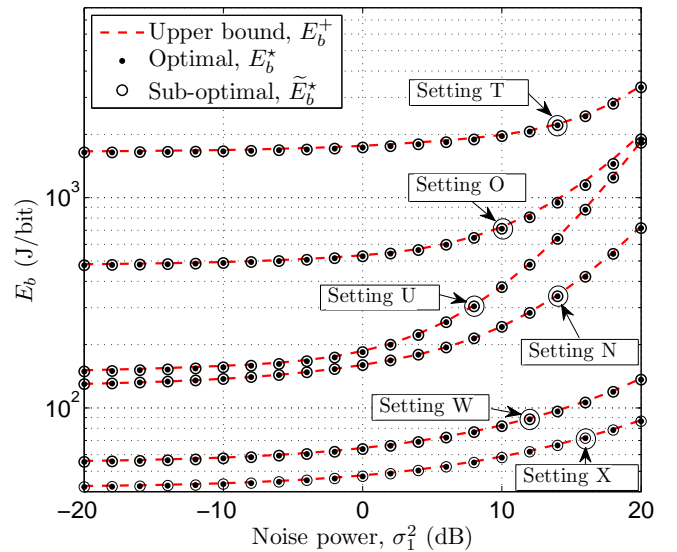


Fig. 4: Comparison of the energy-per-bit performances of our upper bound, optimal and sub-optimal approaches in the power constrained scenario for various settings.

Finally, lowering the power constraint value (when comparing settings N and U in Fig. 4) degrades E_b performance.

In order to quantify more finely the accuracy of our sub-optimal approach against our optimal one, we plot in Fig. 5 the approximation error, Δ_{ac} , between \tilde{E}_b^* (obtained via Algorithm 4) and E_b^* (obtained via Algorithm 3) for various settings, such that $\Delta_{ac} = 100(1 - \tilde{E}_b^*/E_b^*)\%$. We quantify the accuracy of the original sub-optimal approach (as described in Algorithm 4) in the upper part of Fig. 5, whereas in the lower part, we quantify the accuracy of a simplified sub-optimal approach (where $U = 0$ instead of $U = U - 1$ in line 14 of Algorithm 4) such that the algorithm stopped once the first \tilde{E}_b^* is obtained, without extra recursion. The results in the upper part of Fig. 5 confirm the near-optimality of our original sub-optimal approach; indeed, since the figure is only white, it indicates (based on the legend) that \tilde{E}_b^* differs from E_b^* by less than 0.001% regardless of the settings. Whereas the simplified sub-optimal approach exhibits an approximation error that is lower than 2% for most of the settings, as it is depicted in the lower part of Fig. 5.

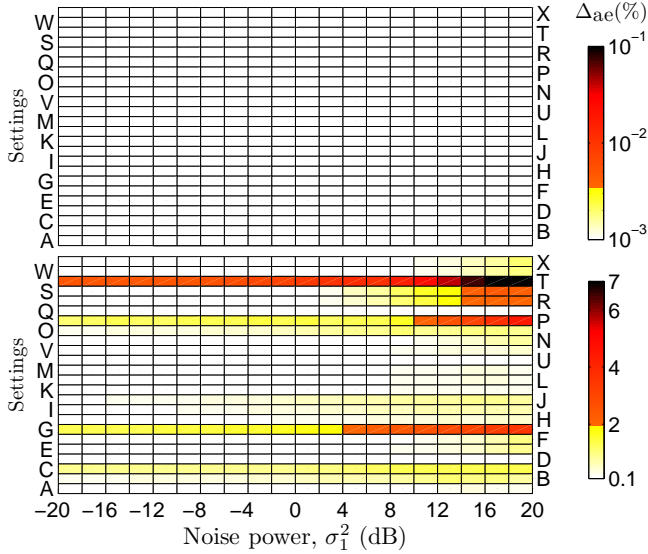


Fig. 5: Approximation error between our optimal and two sub-optimal approaches (original [top] and simplified [bottom]) for various settings.

B. Complexity discussion and results

It can be remarked that the function E_b in (10) is a NM variable function and, hence, solving the problems in (11) requires to search for NM optimal variables at the same time. Such an approach exhibits a prohibitive computational complexity for large values of N or M . Instead, given that we can express $C_{i,m}^*$ solely as function of the same variable, e.g. by using equations (21), (22) and (23) in the unconstrained scenario, the problem in (11) reverts to a single variable problem that can be solved by using a unidimensional root finding method such as the Newton-Raphson or Dinkelbach method. Concerning the complexity of our approaches:

- our lower bound E_b^- is obtained by using a simple root finding method. According to Algorithm 1, obtaining E_b^- requires few steps and a number of operations that increases linearly with M , such that Algorithm 1 exhibits a computational complexity of $O(M)$.
- our upper bound E_b^+ is also obtained via a low-complexity, i.e. water-filling approach, in Algorithm 2, but it clearly involves more operations than Algorithm 1 for computing f and df ; a number which increases not only with M , but also with N , such that Algorithm 2 exhibits a computational complexity of $O(NM)$. Moreover, it requires the output of Algorithm 1 and, hence, obtaining E_b^+ requires more operations than obtaining E_b^- .
- our sub-optimal approach in Algorithm 4 is based on a low-complexity approach, i.e. a unidimensional search based on the Dinkelbach method. Most of the computational complexity arises from the computation of $C_{i,m}^*$ in the function C_{UNC}^* and C_{PWC}^* in the unconstrained and power constrained scenarios, respectively, which grows with both N and M . As it is explained in Section III-A2, the first minimum obtained by Algorithm 4 is not necessarily the lowest minimum and up to $M - 1$ iterations can be necessary to obtain \hat{E}_b^* such that com-

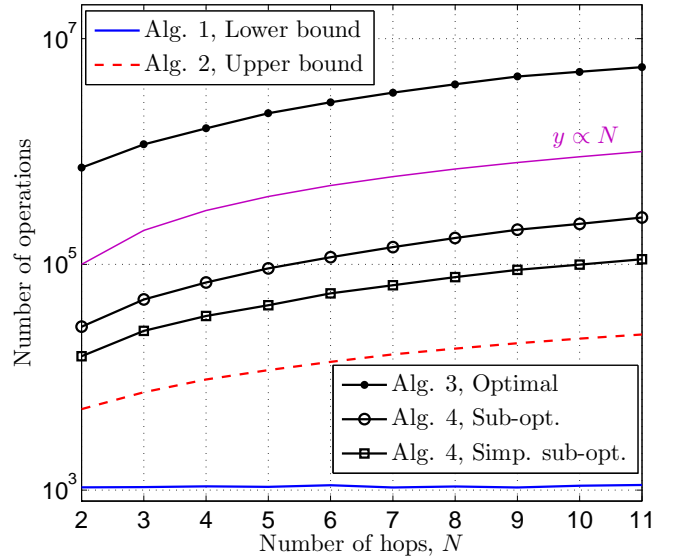


Fig. 6: Comparison of the average number of basic operations required by our various approaches in the unconstrained scenario, as a function of the number of hops.

putational complexity of Algorithm 4 is at worst of $O(NM^2)$. Whereas, in its simplified version, where only one iteration is needed, the complexity is of $O(NM)$.

- our optimal approach in Algorithm 3 is based on an iterative approach, where the maximum number of iterations for finding E_b^* is such that $N_{\text{iter}}^{\max} = 10^v \left(1 + 2 \frac{(V-v)}{u}\right)$ with u , v and V being accuracy parameters that are defined at line 3 of Algorithm 3. The larger v and V are, the more accurate is this algorithm, but the more iterations (complexity) are required. For instance, when $u = 4$, $v = 2$ and $V = 10$ as in Algorithm 3, then $N_{\text{iter}}^{\max} = 500$. The computational complexity of our optimal approach is $O(N_{\text{iter}}^{\max} NM)$ and this approach is expected to be the most complex of all our proposed algorithms.

In order to confirm the conclusions of our complexity analysis, we compare in Figs. 6 and 7 the average number of basic operations (e.g. addition, subtraction, multiplications, etc.) that are required by each of the previously mentioned approaches for returning a result in the unconstrained scenario. The number of basic operations for each approach has been averaged over σ_1^2 and plotted as a function of N with $t = K = 4$ and of K with $N = 2$ & $t = 4$ in the upper and lower parts of Fig. 6, respectively. The results in Fig. 6 confirm that the complexity of Algorithm 1 (Lower bound) is independent of N , and that the complexity of the other algorithms grows linearly with N . Whereas, the results in Fig. 7 demonstrate that the complexity of Algorithm 1, Algorithm 2 and the simplified version of Algorithm 4 grows linearly with M , i.e. K for a fixed t , whereas, the complexity of both Algorithms 3 and 4 grows as a power of M for large M . Note that the complexity results, as a function of t , for fixed N and K are similar to the ones in Fig. 7 and, thus, have been omitted for the clarity of the presentation. More generally, the results in Figs. 6 and 7 confirm that Algorithm 3 (optimal approach) is the more computationally demanding

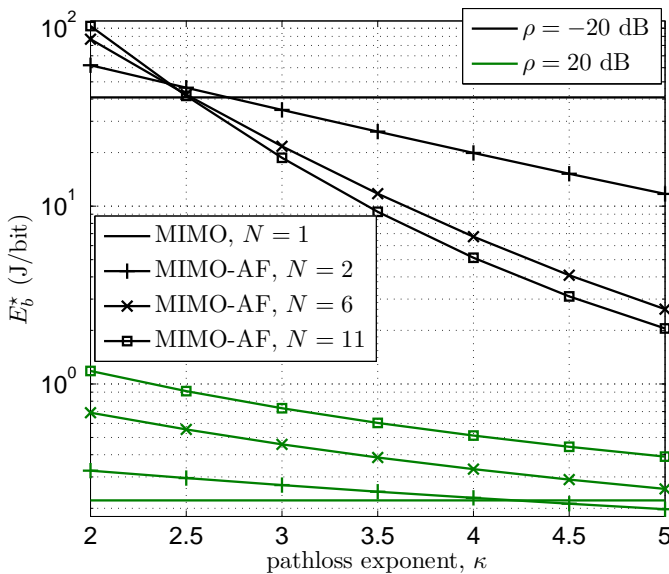


Fig. 9: Comparison of the optimal energy-per-bit consumption of N -hops MIMO-AF and MIMO systems in the unconstrained scenario as a function of the pathloss exponent.

Based on the same parameters as in Fig. 9, Fig. 10 depicts the average sum-rate and transmit power per hop as a function of κ . These results, which complement the results of Fig. 9, provide further insights on the reason behind multi-hopping being more energy efficient than direct communication when the direct link quality is poor. Indeed, the rate and transmit power results for $\rho = -20$ dB show that the channel gain improvement due to multi-hopping translates into a drastic transmit power reduction ($> 90\%$ when comparing the transmit power of 11-hops MIMO-AF with MIMO at $\kappa = 5$) as well as a significant increase in average rate per hop (> 4 times when comparing the rate of 11-hops MIMO-AF with MIMO at $\kappa = 5$). Whereas for $\rho = 20$ dB, the average rate of N -hops MIMO-AF is always worse than MIMO, but N -hops MIMO-AF still requires less transmit power per hop than MIMO.

V. CONCLUSION

In this paper, an optimal energy-efficient joint resource allocation method has been designed for the multi-hop MIMO-AF system when considering that transmit and/or receive CSI is available at each node. We have first expressed the EE-SE trade-off of the multi-hop MIMO-AF system with CSI in close-form when considering a realistic power model and, then, use it as an objective function. Even though, the latter is not generally quasiconvex, we have proved that it can be lower bounded by a convex function, and that all of its minima are strict minima. In addition, we have shown that the optimal solution to the joint resource allocation problem can be upper bounded by solving a low-complexity convex optimization. Based on these properties, we have simplified the multivariate joint resource allocation problem into a single variable problem and designed a novel approach for optimally solving this problem in both the unconstrained and power constraint energy-efficient cases. We have also provided a sub-optimal approach with reduced computational complexity. Accuracy

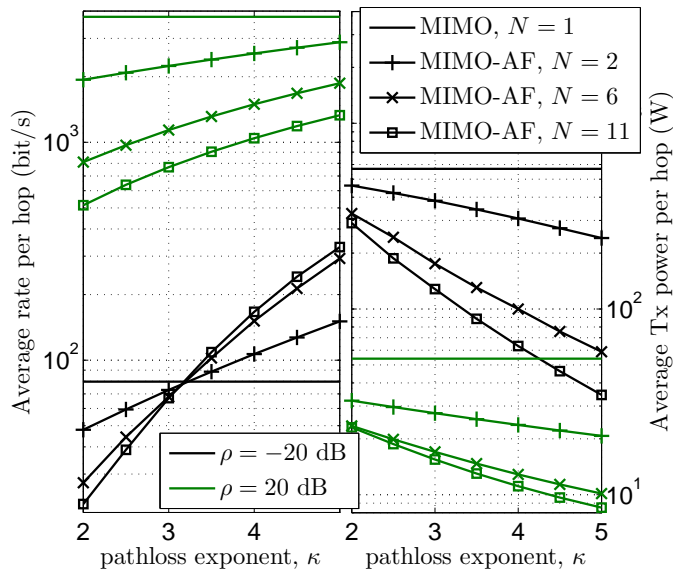


Fig. 10: Comparison of the N -hops MIMO-AF and MIMO systems in terms of the average sum-rate per hop and average transmit power per hop as a function of the pathloss exponent.

and computational complexity simulation results undertaken in numerous settings have shown that our sub-optimal approach is near-optimal with a complexity reduced by up to 40 times in comparison the optimal approach; these results have also confirmed the reliability of our bounds. As an application, we have compared the optimal energy-per-bit consumption of multi-hop MIMO-AF and MIMO systems with CSI. The results have indicated that transmitting over multiple hops can save energy-per-bit when the channel quality of the direct link is poor, i.e. for large pathloss exponent and/or at low SNR. In such channel condition, the rate improvement due to the reduced inter-node distances more than offset the extra power consumption resulting from deploying extra nodes.

APPENDIX

A. Non-quasiconvexity of E_b in the general case

According to Example 3.38 of [39], if $P_{\Sigma}(\mathcal{C}) \geq 0$ and $R_{\Sigma}(\mathcal{C}) > 0$ would be convex and concave, respectively, for any $\mathcal{C} \succ 0$, then E_b would be quasiconvex. Clearly, according to (8) & (5a), $P_{\Sigma}(\mathcal{C}) \geq 0$ and is strictly convex for any $\mathcal{C} \succ 0$. Whereas, $R_{\Sigma}(\mathcal{C}) > 0$ for any $\mathcal{C} \succ 0$, but it is not always concave, as it is shown in the following. The function $R_{\Sigma}(\mathcal{C})$ in (5b) is twice differentiable such that the elements of its gradient and Hessian are expressed as $\{\nabla R_{\Sigma}(\mathcal{C})\}_{\{i,m\}} = \frac{\partial R_{\Sigma}(\mathcal{C})}{\partial c_{i,m}} = W \frac{2^{c_m} - 1}{2^{c_{i,m}} - 1}$ and $\{\nabla^2 R_{\Sigma}(\mathcal{C})\}_{\{i,m\},\{k,l\}} = \frac{\partial^2 R_{\Sigma}(\mathcal{C})}{\partial c_{i,m} \partial c_{k,l}} = W \ln(2)$

$$\times \begin{cases} (2^{c_m} - 1) \frac{2^{c_m} - 2^{c_{i,m}}}{(2^{c_{i,m}} - 1)^2} & \text{if } k = i \text{ and } l = m, \\ \frac{2^{c_m} (2^{c_m} - 1)}{(2^{c_{i,m}} - 1)(2^{c_{i,m}} - 1)} & \text{if } k = \bar{i} \text{ and } l = m, \\ 0 & \text{otherwise,} \end{cases} \quad (31)$$

respectively. Let $\mathbf{z} \in \mathbb{R}^{NM}$, then, according to (31), $\mathbf{z}^\top \nabla^2 R_\Sigma(\mathbf{C}) \mathbf{z}^\top = W \ln(2) \sum_{m=1}^M (2^{C_m} - 1) \times \left[2^{C_m} \left(\sum_{i=1}^N \frac{z_{i,m}}{2^{C_{i,m}-1}} \right)^2 - \sum_{i=1}^N \frac{z_{i,m}^2 2^{C_{i,m}}}{(2^{C_{i,m}-1})^2} \right]$. In the case that $C_{i,m} \ll 1$, then $\mathbf{z}^\top \nabla^2 R_\Sigma(\mathbf{C}) \mathbf{z}^\top \sim \frac{W}{\ln(2)} \sum_{m=1}^M (2^{C_m} - 1) \left[\left(\sum_{i=1}^N \frac{z_{i,m}}{C_{i,m}} \right)^2 - \sum_{i=1}^N \left(\frac{z_{i,m}}{C_{i,m}} \right)^2 \right]$ such that $\mathbf{z}^\top \nabla^2 R_\Sigma(\mathbf{C}) \mathbf{z}^\top$ can either be positive or negative, depending on the sign of the elements of \mathbf{z} . Thus, $R_\Sigma(\mathbf{C})$ is not always concave and, hence, E_b in (10) is not necessarily quasiconvex.

In addition, according to (3.20) of [39], if $E_b(\mathbf{C} + \mathbf{z}) \leq E_b(\mathbf{C}) \Rightarrow \nabla E_b(\mathbf{C}) \mathbf{z}^\top \leq 0$, for all $\mathbf{C} \succ 0$ and $\mathbf{C} + \mathbf{z} \succ 0$, then E_b would be quasiconvex. However, it exists some points $\mathbf{C} \succ 0$ and $\mathbf{C} + \mathbf{z} \succ 0$ that contradicts this statement. For instance, if $M = 2$, $N = 3$, $P_c = 100$, $W = 1$, $\mathbf{A} = 0.01 \times \mathbf{1}^{NM}$, $\mathbf{C} = [2, 1, 3, 3, 4, 5]$ and $\mathbf{z} = [-1, 1, -1, 1, -1, 1]$, then $E_b(\mathbf{C} + \mathbf{z}) - E_b(\mathbf{C}) = -1.91 \leq 0$ but $\nabla E_b(\mathbf{C}) \mathbf{z}^\top = 0.12 > 0$. Similarly, the points $\mathbf{C} = [2, 1, 3, 3, 4, 5, 5, 7]$ and $\mathbf{z} = [-1, 1, -1, 1, -1, 1, -1, 1]$ or $\mathbf{C} = [2, 1, 3, 3, 4, 5, 5, 7, 6, 9]$ and $\mathbf{z} = [-1, 1, -1, 1, -1, 1, -1, 1, -1, 1]$ for $N = 4$ or $N = 5$, respectively, fail the quasiconvexity condition as well. Based on these numerical examples, we can conclude that E_b in (10) is not quasiconvex in the general case.

B. Proof for (13) in Proposition 1

Proof: By inserting (14) into equations (5b) and (5a), the latter can be re-expressed as follows

$$R_\Sigma(\mathbf{C}) = W \sum_{m=1}^M C_m, \text{ and} \quad (32a)$$

$$P_i(\mathbf{C}) = \Delta_i^{-1} \sum_{m=1}^M A_{i,m} \left[(2^{C_m} - 1) + 2^{C_m} 2^{C_{i,m}} (2^{C_{i,m}} - 1) \times \frac{\prod_{\substack{k=1 \\ k \neq i}}^N 2^{C_{k,m}} - \prod_{\substack{k=1 \\ k \neq i}}^N (2^{C_{k,m}} - 1)}{\prod_{k=1}^N 2^{C_{k,m}}} \right], \quad (32b)$$

respectively, for any $\mathbf{C} \succeq 0$. Given that

$$\prod_{i=1}^N 2^{C_{i,m}} - \prod_{i=1}^N (2^{C_{i,m}} - 1) = \sum_{i=1}^N \prod_{\substack{k=1 \\ k \neq i}}^N \left(2^{C_{k,m}} - \frac{1}{2} \right) + \varepsilon_N,$$

with $\varepsilon_N \geq 0$, i.e. $\varepsilon_2 = 0$, $\varepsilon_3 = \frac{1}{4}$, $\varepsilon_4 = \frac{1}{4} \sum_{k=1}^4 (2^{C_{k,m}} - \frac{1}{2})$, etc., it implies that $\sum_{i=1}^N \Delta_i P_i(\mathbf{C}) - \sum_{m=1}^M \sum_{i=1}^N A_{i,m} (2^{C_m} - 1)$ can be lower bounded by

$$\sum_{m=1}^M \frac{2^{C_m}}{\prod_{k=1}^N 2^{C_{k,m}}} \sum_{i=1}^N \sum_{k=i+1}^N [A_{i,m} 2^{C_{i,m}} (2^{C_{i,m}} - 1) + A_{k,m} 2^{C_{k,m}} (2^{C_{k,m}} - 1)] \prod_{\substack{l=1 \\ l \neq k \neq i}}^N \left(2^{C_{l,m}} - \frac{1}{2} \right). \quad (33)$$

Moreover, given that $(2^x - \frac{1}{2}) \geq \sqrt{2^x(2^x - 1)}$, $\forall x \geq 0$ and $x + y \geq 2\sqrt{xy}$, since $(\sqrt{x} - \sqrt{y})^2 \geq 0$, $\forall x, y \geq 0$, (33) can further be lower bounded by

$$\sum_{m=1}^M 2^{C_m+1} \sum_{i=1}^N \sum_{k=i+1}^N \sqrt{A_{i,m} A_{k,m}} \prod_{l=1}^N \sqrt{1 - 2^{-C_{l,m}}}. \quad (34)$$

Then, by inserting (15) into (34), the latter can be re-expressed solely as a function of C_m as

$$2 \sum_{m=1}^M \left(\sum_{i=1}^N \sum_{k=i+1}^N \sqrt{A_{i,m} A_{k,m}} \right) \sqrt{2^{C_m} (2^{C_m} - 1)}.$$

Consequently, $\sum_{i=1}^N \Delta_i P_i(\mathbf{C})$ can be lower bounded by

$$\sum_{i=1}^N \Delta_i P_i(\mathbf{C}) \geq \sum_{m=1}^M \left(\sum_{i=1}^N \sqrt{A_{i,m}} \right)^2 (2^{C_m} - 1), \quad (35)$$

for any $\mathbf{C} \succeq 0$, since $\sqrt{2^x(2^x - 1)} \geq (2^x - 1)$, $\forall x \geq 0$, and $2 \sum_{i=1}^N \sum_{k=i+1}^N \sqrt{A_{i,m} A_{k,m}} = \left(\sum_{i=1}^N \sqrt{A_{i,m}} \right)^2 - \sum_{i=1}^N A_{i,m}$, $\forall m \in \mathcal{M}$. Equation (13) is finally obtained by substituting the numerator and denominator of (10) with (35) and (32a), respectively. ■

C. Proof for Proposition 2

Proof: Let \mathbf{C}^* be a stationary point of E_b , accordingly, $\nabla E_b(\mathbf{C} = \mathbf{C}^*) = \mathbf{0}$. Given that $\nabla E_b(\mathbf{C}) \mathbf{z}^\top \triangleq R_\Sigma(\mathbf{C})^{-2} [\nabla P_\Sigma(\mathbf{C}) \mathbf{z}^\top R_\Sigma(\mathbf{C}) - \nabla R_\Sigma(\mathbf{C}) \mathbf{z}^\top P_\Sigma(\mathbf{C})] = 0$ is equivalent to $E_b(\mathbf{C}) = \frac{\nabla P_\Sigma(\mathbf{C}) \mathbf{z}^\top}{\nabla R_\Sigma(\mathbf{C}) \mathbf{z}^\top}$, then $E_b(\mathbf{C} + \mathbf{z}) - E_b(\mathbf{C}) =$

$$\frac{\nabla P_\Sigma(\mathbf{C})(2^{\mathbf{z}} - 1)^\top}{\ln(2) R_\Sigma(\mathbf{C} + \mathbf{z})} - \frac{[R_\Sigma(\mathbf{C} + \mathbf{z}) - R_\Sigma(\mathbf{C})] \nabla P_\Sigma(\mathbf{C}) \mathbf{z}^\top}{R_\Sigma(\mathbf{C} + \mathbf{z}) \nabla R_\Sigma(\mathbf{C}) \mathbf{z}^\top}. \quad (36)$$

In addition, let $F : \mathbf{X} \in \mathbb{R}^{2M} \mapsto \mathbb{R}$ and $\mathbf{z} \rightarrow 0$, then the gradient of F is similar to

$$\nabla F(\mathbf{X}) \mathbf{z}^\top \simeq F(\mathbf{X} + \mathbf{z}) - F(\mathbf{X}). \quad (37)$$

Given that $\nabla P_\Sigma(\mathbf{C})(2^{\mathbf{z}} - 1)^\top > \ln(2) \nabla P_\Sigma(\mathbf{C}) \mathbf{z}^\top$, for $\mathbf{z} \neq \mathbf{0}$, it implies with (36) and (37) that $E_b(\mathbf{C}^* + \mathbf{z}) > E_b(\mathbf{C}^*)$, for $\mathbf{z} \rightarrow 0$ and $\mathbf{z} \neq 0$. ■

D. Proof for Proposition 3

Proof: From (10), solving $\nabla E_b(\mathbf{C}^*) = \mathbf{0}$ is such that

$$\nabla E_b(\mathbf{C}^*) = \frac{\nabla P_\Sigma(\mathbf{C}^*) R_\Sigma(\mathbf{C}^*) - \nabla R_\Sigma(\mathbf{C}^*) P_\Sigma(\mathbf{C}^*)}{R_\Sigma(\mathbf{C}^*)^2} = \mathbf{0}.$$

In turn, it yields, as it is summarized in equation (20),

$$E_b^* = E_b(\mathbf{C}^*) = \frac{P_\Sigma(\mathbf{C}^*)}{R_\Sigma(\mathbf{C}^*)} = \frac{\partial P_\Sigma(\mathbf{C}^*)}{\partial C_{i,m}} \left[\frac{\partial R_\Sigma(\mathbf{C}^*)}{\partial C_{i,m}} \right]^{-1},$$

for $i \in \mathcal{N}$ and $m \in \mathcal{M}^*$, where $\frac{\partial P_\Sigma(\mathbf{C})}{\partial C_{i,m}} = \ln(2) A_{i,m} 2^{C_{i,m}}$ and

$$\frac{\partial R_\Sigma(\mathbf{C})}{\partial C_{i,m}} = W \frac{\prod_{\substack{k=1 \\ k \neq i}}^N (2^{C_{k,m}} - 1)}{\prod_{k=1}^N 2^{C_{k,m}} - \prod_{k=1}^N (2^{C_{k,m}} - 1)} = W \frac{2^{C_m} - 1}{2^{C_{i,m}} - 1}. \quad \blacksquare$$

E. Proof for Corollary 3

1) Equation (22): *Proof:* Given any $m \in \mathcal{M}^*$, relation (20) holds for any $i \in \mathcal{N}$, and

$$\prod_{i=1}^N E_b^* = \prod_{i=1}^N \frac{\ln(2) A_{i,m} 2^{C_{i,m}^*} (2^{C_{i,m}^*} - 1)}{W (2^{C_m^*} - 1)}, \quad (38a)$$

$$\Leftrightarrow \frac{(2^{C_m^*} - 1)^N}{\prod_{i=1}^N (1 - 2^{-C_{i,m}^*})} = \prod_{i=1}^N 4a_{i,m}^* 2^{2C_{i,m}^*}, \quad (38b)$$

$$\Leftrightarrow \sqrt{2^{C_m^*} (2^{C_m^*} - 1)^{N-1}} = \prod_{i=1}^N 2\sqrt{a_{i,m}^*} 2^{C_{i,m}^*}, \quad (38c)$$

where $a_{i,m}^* = \frac{\ln(2)A_{i,m}}{4WE_b^*}$. Note that we have relied on equation (15) for transforming (38b) into (38c). Equation (22) is then obtained by replacing $C_{i,m}^*$ in (38c) with (21) and applying the change of variables $X_m = \sqrt{2^{C_m^*} - 1}$. ■

2) Equation (23): *Proof:*

Firstly, given that $a_{i,m}^* > 0$, it implies that $\prod_{i=1}^N (\sqrt{a_{i,m}^*} + \sqrt{a_{i,m}^* + X_m^2}) > \prod_{i=1}^N (\sqrt{a_{i,m}^*} + X_m) = X_m^N + X_m^{N-1} \sum_{i=1}^N \sqrt{a_{i,m}^*} + X_m^{N-2} \sum_{i=1}^N \sum_{k=i+1}^N \sqrt{a_{i,m}^* a_{k,m}^*} + \sigma(X_m)$, where $\sigma(X_m) \geq 0$; secondly, knowing that $\sqrt{x^2(x^2+1)} < \sqrt{x^2(x^2+1)+1/4} = x + 1/2$, $\forall x \in \mathbb{R}$, it implies that $-X_m^{N-1} \sqrt{X_m^2+1} > -X_m^{N-2} (X_m^2 + 1/2)$. Consequently, $F(X_m) > \widehat{F}(X_m) = X_m^{N-2} (\sum_{i=1}^N \sqrt{a_{i,m}^*} X_m + \sum_{i=1}^N \sum_{k=i+1}^N \sqrt{a_{i,m}^* a_{k,m}^*} - 1/2)$, $\forall X_m \geq 0$, such that the only non-zero root of $\widehat{F}(X_m)$, X_m^{\max} , is expressed as in (23).

Given that $\widehat{F}(X_m) > 0$ for $X_m > X_m^{\max}$ and $F(X_m) > \widehat{F}(X_m)$, it implies that $F(X_m) > 0$ when $X_m > X_m^{\max}$; hence, F has no roots for $X_m > X_m^{\max}$ such that any root of F lies in $[0, X_m^{\max}]$. ■

REFERENCES

- [1] H. Zhang et al., "Energy efficiency in communications," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 48–79, Nov. 2010.
- [2] —, "Energy efficiency in communications: part II," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 28–87, Jun. 2011.
- [3] —, "Energy efficiency in communications: part III," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 52–109, Aug. 2011.
- [4] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Godór, G. Auer, and L. V. D. Perre, "Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 66–72, Nov. 2010.
- [5] G. Auer et al., "How Much Energy is Needed to Run a Wireless Network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.
- [6] C.-L. Chen, W. E. Stark, and S.-G. Chen, "Energy-Bandwidth Efficiency Tradeoff in MIMO Multi-Hop Wireless Networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1537–1546, Sep. 2011.
- [7] C. Sun and C. Yang, "Energy efficiency analysis of one-way and two-way relay systems," *EURASIP J. Wirel. Commun. and Netw.*, vol. 46, no. 2, Feb. 2012.
- [8] A. Zappone, P. Cao, and E. A. Jorswieck, "Energy Efficiency Optimization in Relay-Assisted MIMO Systems with Perfect and Statistical CSI," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 443–457, Jan. 2014.
- [9] K. T. K. Cheung, S. Yang, and L. Hanzo, "Achieving maximum energy-efficiency in multi-relay OFDMA cellular networks: a fractional Programming approach," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2746–2757, Jul. 2013.
- [10] F. Hélot, "Low-complexity energy-efficient joint resource allocation for two-hop MIMO-AF systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3088–3099, Jun. 2014.
- [11] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity Part I & II- System Description / Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1948, Nov. 2003.
- [12] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded Cooperation in Wireless Communications: Space-Time Transmission and Iterative Decoding," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 362–371, Feb. 2004.
- [13] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.
- [14] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [15] "Reconfigurable OFDMA-based Cooperative Networks Enabled by Agile SpecTrum Use (ROCKET)," ICT-215282 FP7 STREP project, Tech. Rep., 2008-2009. [Online]. Available: <http://www.ict-rocket.eu>
- [16] 3GPP TS 36.216 (2010-15), "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer for relaying operation (Release 10-13)," 3GPP, Tech. Rep., 2010-2015. [Online]. Available: <http://www.3gpp.org/dynareport/36216.htm>
- [17] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear Transceiver Design in Nonregenerative Relays with Channel State Information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2593–2604, Jun. 2007.
- [18] I. Hammerström and A. Wittneben, "Power Allocation Schemes for Amplify-and-Forward MIMO-OFDM Relay Links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2798–2802, Aug. 2007.
- [19] C. Li, X. Wang, L. Yang, and W.-P. Zhu, "A Joint Source and Relay Power Allocation Scheme for a Class of MIMO Relay Systems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4852–4860, Dec. 2009.
- [20] Y. Yu and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2823–2835, May 2010.
- [21] C. Li, X. Wang, L. Yang, and W.-P. Zhu, "Joint source-and-relay power allocation in multiple-input multiple-output amplify-and-forward relay systems: a non-convex problem and its solution," *IET Signal Processing*, vol. 5, no. 6, pp. 612–622, Jun. 2011.
- [22] C. Xing, S. Ma, Z. Fei, Y.-C. Wu, and H. V. Poor, "A General Robust Linear Transceiver Design for Multi-Hop Amplify-and-Forward MIMO Relaying Systems," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1196–1209, Mar. 2013.
- [23] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An Energy-Efficient Approach to Power Control and Receiver Design in Wireless Networks," *IEEE Trans. Commun.*, vol. 5, no. 1, pp. 3306–3315, Nov. 2006.
- [24] G. Miao, N. Himayat, and G. Y. Li, "Energy-Efficient Link Adaptation in Frequency-Selective Channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545–554, Feb. 2010.
- [25] R. S. Prabhu and B. Daneshrad, "Energy-efficient power loading for a MIMO-SVD system and its performance in flat fading," in *Proc. IEEE Globecom*, Miami, USA, Dec. 2010.
- [26] F. Hélot, M. A. Imran, and R. Tafazolli, "Energy-efficient power allocation for point-to-point MIMO systems over the Rayleigh fading channel," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 304–307, Aug. 2012.
- [27] —, "Energy-efficient resource allocation for orthogonal multi-antenna multi-carrier channel," in *Proc. IEEE OnlineGreenComm 2013*, Online, Oct. 2013.
- [28] —, "Near-optimal energy-efficient joint resource allocation for multi-hop MIMO-AF systems," in *Proc. IEEE PIMRC*, London, UK, Sep. 2013.
- [29] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6068–6077, Dec. 2009.
- [30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1985.
- [31] Y. Rong, "Multihop nonregenerative MIMO relays-QoS considerations," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 290–303, Jan. 2011.
- [32] Y. Qi, F. Hélot, and M. A. Imran, *Green communications and networking: chapter 3: Green relay techniques in cellular systems*. CRC press, Dec. 2012.
- [33] F. Hélot, M. A. Imran, and R. Tafazolli, "On the energy efficiency-spectral efficiency trade-off over the MIMO Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1345–1356, May 2012.
- [34] S. Verdú, "Spectral Efficiency in the Wideband Regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

- [35] F. Hélot, M. A. Imran, and R. Tafazolli, "Low-complexity energy-efficient coordinated resource allocation in cellular systems," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2271–2281, Jun. 2013.
- [36] "Wolfram Mathworld Website," [Online]. Available: <http://mathworld.wolfram.com/NewtonsMethod.html>.
- [37] W. Dinkelbach, "On non-linear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
- [38] M. Golden, *Mathematical Methods for Neural Network Analysis and Design*. MIT Press, Dec. 1996.
- [39] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.