

Energy-Efficiency Maximization in Downlink Clustered NOMA Networks with Energy-Harvesting Relays

Mohammed W. Baidas

Department of Electrical Engineering, College of Engineering and Petroleum, Kuwait University, Kuwait
(email: m.baidas@ku.edu.kw)

Abstract—In this paper, the problem of joint relay assignment and energy-efficiency maximization (J-RA-EE-MAX) in energy-harvesting downlink clustered non-orthogonal multiple-access (NOMA) networks is considered. Specifically, the aim is to perform relay assignment to user clusters, while simultaneously maximizing energy-efficiency over each relay via multi-objective optimization, and satisfying users' quality-of-service (QoS) constraints. However, problem J-RA-EE-MAX happens to be non-convex (i.e. computationally-prohibitive). Alternatively, a low-complexity solution procedure is devised to solve problem J-RA-EE-MA by: (1) optimally solving the energy-efficiency maximizing power allocation (EE-MAX-PA) for each (user cluster, relay) combination to construct the relays' preference profile, and (2) performing relay assignment via Gale's top trading cycles (TTC) matching mechanism. Simulation results are presented to validate the proposed solution procedure, which is shown to yield comparable energy-efficiency value per relay to the J-RA-EE-MAX scheme, while satisfying users' QoS constraints.

Index Terms—Energy-efficiency, matching, non-orthogonal multiple-access, power allocation, relay assignment

I. INTRODUCTION

The recent explosive spread of smart wireless devices and services, along with the rise of the Internet-of-Things (IoT) and 5G networks, have pushed the limits for higher spectrum and energy-efficiency [1,2]. In turn, non-orthogonal multiple-access (NOMA) has been proposed, whereby multiple users can be multiplexed in the power-domain and then detected via successive interference cancellation (SIC), while satisfying the users' quality-of-service (QoS) demands [3]. Additionally, energy-harvesting (EH) has emerged as a promising technology to extend the lifetime of cellular networks. Cooperative relay communications have also appeared as a viable means to exploit diversity gains, and extend network coverage. Hence, designing transmission schemes with EH relays for energy-efficiency (EE) maximization in NOMA cellular networks is of paramount importance.

Several works have focused on energy-efficient relay communications in NOMA networks. For instance, in [4], optimal power allocation for global energy-efficiency maximization is considered, and a limited-complexity scheme is proposed, and shown to outperform its orthogonal multiple-access (OMA) counterpart. The EE-delay tradeoff for a cooperative downlink NOMA system is studied in [5], for the scenario of a base-station and two users, where a user may act as a relay for the other user. It has been demonstrated that the proposed system yields better EE-delay tradeoff than the OMA-based one. In [6], the problem of ergodic weighted sum-rate maximization of RF-EH cooperative NOMA systems is considered, and a low-complexity search algorithm is proposed, and shown to significantly improve the system's weighted sum-rate performance. Joint subcarrier assignment and weighted-sum energy-efficient

power allocation (J-SA-WSEE-PA) in uplink NOMA relay networks is studied in [7]. Particularly, a low-complexity algorithm is devised, yielding comparable sum energy-efficiency to the global optimal scheme.

This paper considers joint relay assignment and energy-efficiency maximization (J-RA-EE-MAX) in EH downlink clustered NOMA networks. The aim is to perform relay assignment for clustered NOMA users, while simultaneously maximizing EE over each EH relay and satisfying users' QoS constraints. However, problem J-RA-EE-MAX happens to be non-convex (i.e. computationally-prohibitive). In turn, a low-complexity solution procedure is devised to solve problem J-RA-EE-MAX by: (1) optimally solving the EE-maximizing power allocation (EE-MAX-PA) per (user cluster, relay) combination to construct the relays' preference profile, and (2) performing relay assignment via Gale's top trading cycles (TTC) matching mechanism [8]. Particularly, the optimal solution to the EE-MAX-PA problem is obtained by transforming it into a concave maximization problem [9], while the TTC mechanism is executed in linear time-complexity to obtain a stable relay assignment. Simulation results are presented to validate the proposed solution procedure, which will be shown to yield comparable energy-efficiency value per relay to the J-RA-EE-MAX scheme (solved via a global optimization package); however, at lower computational-complexity, while satisfying users' QoS requirements.

In the rest of this paper, Section II presents the system model, while Section III presents the joint relay assignment and energy-efficiency maximization problem formulation. In Section IV, the energy-efficiency maximizing power allocation over each relay is discussed, while Section V presents the relay assignment algorithm. In Section VI, the simulation results are presented, while Section VII draws the conclusions.

II. SYSTEM MODEL

Consider a single-cell NOMA network with a base-station (BS), K half-duplex EH amplify-and-forward (AF) relays, and N users. The users' set is denoted $\mathcal{U} = \{U_1, \dots, U_n, \dots, U_N\}$, while the relays' set is $\mathcal{R} = \{R_1, \dots, R_k, \dots, R_K\}$. Let there be a set of M user clusters¹ denoted $\mathcal{C} = \{C_1, \dots, C_m, \dots, C_M\}$. The user clusters partition \mathcal{U} (i.e. $C_m \cap C_{m'} = \emptyset$ for $m \neq m'$, and $\bigcup_{m=1}^M C_m = \mathcal{U}$). The base-station communicates the data symbols x_n to the users in each cluster via a relay $R_k \in \mathcal{R}$, which is allocated a dedicated channel. The communication channels between each user $U_n \in C_m$ and relay R_k , and also between relay R_k and

¹In social and IoT networks, users and/or devices may be clustered according to their geographical locations and density, mobility patterns, traffic similarity, resource requirements, etc. [10].

the base-station follow narrowband Rayleigh fading with zero-mean N_0 -variance additive white Gaussian noise (AWGN). Let $h_{n,m,k} \sim \mathcal{CN}(0, d_{n,m,k}^{-\nu})$ and $h_{k,bs} \sim \mathcal{CN}(0, d_{k,bs}^{-\nu})$ be the channel coefficients of the user-relay and relay-base-station links, respectively. Moreover, $d_{n,m,k}$ and $d_{k,bs}$ are the link distances, while ν is the path-loss exponent. Additionally, let $E_{n,m,k}$ be the transmit energy allocated to user $U_n \in \mathcal{C}_m$ over relay $R_k \in \mathcal{R}$. Also, let E_{\max} be the maximum transmit energy per time-slot, such that $\sum_{U_n \in \mathcal{C}_m} E_{n,m,k} \leq E_{\max}$, $\forall \mathcal{C}_m \in \mathcal{C}$, and $\forall R_k \in \mathcal{R}$.

In this work, it is assumed that there is no direct link between the base-station and the users. Thus, the communication between the users and the base-station is performed over two phases (of one time-slot each), namely the broadcasting phase, and the cooperation phase. Let a transmission frame be made up of the broadcasting and the cooperation phases (i.e. of two time-slots) [11], and be denoted by τ .

Remark 1. For $\tau = 0$, no transmission takes place, as the relays harvest energy from the surrounding environment. Data transmission occurs in frames $\tau \geq 1$, by utilizing the harvested (and any leftover) energy in the previous frames.

For convenience, define the binary decision variable

$$\mathcal{I}_{m,k}^\tau = \begin{cases} 1, & \text{if } \mathcal{C}_m \text{ is assigned } R_k \text{ in frame } \tau, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Also let $\mathcal{I}_k^\tau \triangleq \sum_{\mathcal{C}_m \in \mathcal{C}} \mathcal{I}_{m,k}^\tau$, $\forall R_k \in \mathcal{R}$.

Remark 2. In this work, each user cluster is assigned one relay, while each relay can be assigned to at most one user cluster. That is, $\sum_{R_k \in \mathcal{R}} \mathcal{I}_{m,k}^\tau = 1$, $\forall \mathcal{C}_m \in \mathcal{C}$, and $\sum_{\mathcal{C}_m \in \mathcal{C}} \mathcal{I}_{m,k}^\tau \leq 1$, $\forall R_k \in \mathcal{R}$.

A. Relays' Energy-Harvesting Model and Battery Dynamics

The harvested energy \mathcal{E}_k^τ in each frame τ is modeled as an independent uniformly distributed random variable, as $\mathcal{E}_k^\tau \sim \mathbb{U}(0, \mathcal{E}_k^{\max})$, with \mathcal{E}_k^{\max} being the maximum harvested energy value per frame [12]. Also, let B_k^{\max} be the battery capacity of relay $R_k \in \mathcal{R}$, where any leftover harvested energy from a previous frame is stored for later use.

Let E_k^τ be the harvested energy utilized in the downlink transmission in frame $\tau \geq 1$. Particularly, $E_k^\tau = \min(\mathcal{E}_k^{\tau-1}, E_{\max})$, $\forall R_k \in \mathcal{R}$. Now, if $\mathcal{E}_k^{\tau-1} > E_{\max}$, then the leftover energy (i.e. $\max(0, \mathcal{E}_k^{\tau-1} - E_{\max})$) is stored in the battery. It should be noted that if a relay was not assigned to a user cluster in frame τ (i.e. $\mathcal{I}_k^\tau = 0$), then its allocated transmit energy is stored again in the battery for use in later transmission frames along with any leftover energy, while satisfying the maximum battery capacity B_k^{\max} , $\forall R_k \in \mathcal{R}$. Therefore, the total leftover energy can be compactly expressed as [12]

$$\Delta \mathcal{E}_k^\tau = (1 - \mathcal{I}_k^\tau) E_k^\tau + \max(0, \mathcal{E}_k^{\tau-1} - E_{\max}), \quad (2)$$

where $\Delta \mathcal{E}_k^0 = 0$. Hence, the total amount of harvested (and any leftover) energy at each relay $R_k \in \mathcal{R}$ is defined as

$$\mathcal{E}_k^\tau = \begin{cases} \min(\mathcal{E}_k^0, B_k^{\max}), & \text{for } \tau = 0, \\ \min(\mathcal{E}_k^\tau + \Delta \mathcal{E}_k^\tau, B_k^{\max}), & \text{for } \tau \geq 1. \end{cases} \quad (3)$$

Remark 3. From this point onwards, and for notational convenience, the superscript τ is dropped, while implicitly taking into account battery dynamics during network operation.

B. Transmission Model

1) *Broadcasting Phase:* The BS broadcasts—via downlink NOMA—the data symbol of each user in each cluster $\mathcal{C}_m \in \mathcal{C}$, which is received at a relay $R_k \in \mathcal{R}$ as

$$y_{m,k} = h_{k,bs} \sum_{U_n \in \mathcal{C}_m} \sqrt{E_{n,m,k}} x_n + \eta_k, \quad (4)$$

where η_k is the received AWGN sample.

2) *Cooperation Phase:* Each assigned relay R_k amplifies-and-forwards its received signal to the users in cluster \mathcal{C}_m as

$$x_k = \sqrt{E_k G_{m,k}} y_{m,k}, \quad (5)$$

where $G_{m,k}$ is the AF normalization factor, given by

$$G_{m,k} = \frac{1}{|h_{k,bs}|^2 \sum_{U_n \in \mathcal{C}_m} E_{n,m,k} + N_0}. \quad (6)$$

Thus, the received signal at a user $U_n \in \mathcal{C}_m$ is expressed as

$$y_{n,m,k} = \sqrt{E_k G_{m,k}} h_{n,m,k} h_{k,bs} \sum_{U_n \in \mathcal{C}_m} \sqrt{E_{n,m,k}} x_n + \sqrt{E_k G_{m,k}} h_{n,m,k} \eta_k + \eta_{n,m,k}, \quad (7)$$

where $\eta_{n,m,k}$ is the received AWGN at user $U_n \in \mathcal{C}_m$. As per the principle of NOMA, the users in \mathcal{C}_m are ordered with respect to each relay $R_k \in \mathcal{R}$ according to their channel gains, as $|h_{1,m,k}|^2 \leq \dots \leq |h_{n,m,k}|^2 \leq \dots \leq |h_{U_{|\mathcal{C}_m|},m,k}|^2$ [13]. Thus, the transmit energies of the users in \mathcal{C}_m are ordered as $E_{1,m,k} \geq \dots \geq E_{n,m,k} \geq \dots \geq E_{U_{|\mathcal{C}_m|},m,k}$. For convenience, let the subset of users in \mathcal{C}_m with better channel conditions than user U_n (with respect to relay R_k) be denoted by

$$\bar{\mathcal{C}}_{n,m,k} = \left\{ U_i \in \mathcal{C}_m \setminus U_n \mid |h_{i,m,k}|^2 \geq |h_{n,m,k}|^2 \right\}, \quad (8)$$

with $\bar{\mathcal{C}}_{|\mathcal{C}_m|,m,k} = \emptyset$. Assuming perfect SIC, the end-to-end signal-to-interference-plus-noise ratio (SINR) of the ordered user $U_n \in \mathcal{C}_m$ over relay R_k can be shown to be [13]

$$\gamma_{n,m,k}(\mathbf{E}_{m,k}) = \frac{\bar{\gamma}_{n,m,k}(\mathbf{E}_{m,k})}{\underline{\gamma}_{n,m,k}(\mathbf{E}_{m,k})}, \quad (9)$$

where

$$\bar{\gamma}_{n,m,k}(\mathbf{E}_{m,k}) \triangleq E_{n,m,k} \xi_{n,m,k} E_k \xi_{k,bs}, \quad (10)$$

and

$$\underline{\gamma}_{n,m,k}(\mathbf{E}_{m,k}) \triangleq E_k \xi_{n,m,k} \xi_{k,bs} \sum_{U_i \in \bar{\mathcal{C}}_{n,m,k}} E_{i,m,k} + \xi_{k,bs} \sum_{U_i \in \mathcal{C}_m} E_{i,m,k} + E_k \xi_{n,m,k} + 1. \quad (11)$$

In addition, $\xi_{n,m,k} = |h_{n,m,k}|^2 / N_0$, $\xi_{k,bs} = |h_{k,bs}|^2 / N_0$, and $\mathbf{E}_{m,k} = [E_{1,m,k}, \dots, E_{n,m,k}, \dots, E_{U_{|\mathcal{C}_m|},m,k}]$ is transmit energy vector of user cluster \mathcal{C}_m over relay R_k . Hence, the achievable rate of user $U_n \in \mathcal{C}_m$ over relay R_k is given by

$$\mathbb{R}_{n,m,k}(\mathbf{E}_{m,k}) = \frac{1}{2} \log_2 \left(1 + \gamma_{n,m,k}(\mathbf{E}_{m,k}) \right). \quad (12)$$

To ensure QoS, each user must satisfy

$$\mathbb{R}_{n,m,k}(\mathbf{E}_{m,k}) \geq \mathbb{R}_{\min}, \quad \forall U_n \in \mathcal{C}_m, \quad (13)$$

where \mathbb{R}_{\min} is the target minimum rate.

Remark 4. $\gamma_{n,m,k}(\mathbf{E}_{m,k})$ is a linear-fractional function in $\mathbf{E}_{m,k}$, with $\underline{\gamma}_{n,m,k}(\mathbf{E}_{m,k}) > 0$. In turn, $\gamma_{n,m,k}(\mathbf{E}_{m,k})$ is a pseudo-linear function (i.e. both pseudo-concave and pseudo-convex), $\forall U_n \in \mathcal{C}_m$, $\forall \mathcal{C}_m \in \mathcal{C}$, and $\forall R_k \in \mathcal{R}$ [14].

C. Energy-Efficiency

Let the sum-rate of a user cluster \mathcal{C}_m over a relay R_k be

$$\mathbb{R}_{m,k}(\mathbf{E}_{m,k}) = \sum_{U_n \in \mathcal{C}_m} \mathbb{R}_{n,m,k}(\mathbf{E}_{m,k}), \quad (14)$$

while the total energy consumption of cluster \mathcal{C}_m over R_k is

$$\mathbb{E}_{m,k}(\mathbf{E}_{m,k}) = \sum_{U_n \in \mathcal{C}_m} E_{n,m,k}. \quad (15)$$

Hence, the energy-efficiency over each relay R_k is defined as

$$\mathbb{EE}_k(\mathbf{E}_k, \mathcal{I}_k) = \frac{\sum_{\mathcal{C}_m \in \mathcal{C}} \mathcal{I}_{m,k} \mathbb{R}_{m,k}(\mathbf{E}_{m,k})}{\sum_{\mathcal{C}_m \in \mathcal{C}} \mathcal{I}_{m,k} \mathbb{E}_{m,k}(\mathbf{E}_{m,k}) + E_k}, \quad (16)$$

where $\mathbf{E}_k \triangleq [\mathbf{E}_{1,k}, \dots, \mathbf{E}_{m,k}, \dots, \mathbf{E}_{M,k}]$ is the transmit energy matrix over relay R_k , while $\mathcal{I}_k \triangleq [\mathcal{I}_{1,k}, \dots, \mathcal{I}_{m,k}, \dots, \mathcal{I}_{M,k}]$ is the user cluster-relay assignment vector over relay R_k .

III. JOINT RELAY ASSIGNMENT AND ENERGY-EFFICIENCY MAXIMIZATION

The joint relay assignment and energy-efficiency maximization (J-RA-EE-MAX) problem over all relays is formulated as a multi-objective optimization problem, as

$$\mathbf{J}\text{-RA-EE-MAX}(\mathbf{E}, \mathcal{I}): \quad (17)$$

$$\max (\mathbb{EE}_1(\mathbf{E}_1, \mathcal{I}_1), \dots, \mathbb{EE}_k(\mathbf{E}_k, \mathcal{I}_k), \dots, \mathbb{EE}_K(\mathbf{E}_K, \mathcal{I}_K)) \quad (17a)$$

$$\text{s.t. } \sum_{R_k \in \mathcal{R}} \mathcal{I}_{m,k} = 1, \quad \forall \mathcal{C}_m \in \mathcal{C}, \quad (17b)$$

$$\sum_{\mathcal{C}_m \in \mathcal{C}} \mathcal{I}_{m,k} \leq 1, \quad \forall R_k \in \mathcal{R}, \quad (17c)$$

$$\mathcal{I}_{m,k} \cdot (\mathbb{R}_{n,m,k}(\mathbf{E}_{m,k}) - \mathbb{R}_{\min}) \geq 0, \quad \forall U_n \in \mathcal{C}_m, \forall \mathcal{C}_m \in \mathcal{C}, \forall R_k \in \mathcal{R}, \quad (17d)$$

$$\sum_{U_n \in \mathcal{C}_m} E_{n,m,k} \leq E_{\max} \cdot \mathcal{I}_{m,k}, \quad \forall \mathcal{C}_m \in \mathcal{C}, \forall R_k \in \mathcal{R}, \quad (17e)$$

$$E_{1,m,k} \geq \dots \geq E_{n,m,k} \geq \dots \geq E_{U_{|\mathcal{C}_m|},m,k}, \quad \forall \mathcal{C}_m \in \mathcal{C}, \forall R_k \in \mathcal{R}, \quad (17f)$$

$$0 \leq E_{n,m,k} \leq E_{\max}, \quad \forall U_n \in \mathcal{C}_m, \forall \mathcal{C}_m \in \mathcal{C}, \forall R_k \in \mathcal{R}, \quad (17g)$$

$$\mathcal{I}_{m,k} \in \{0, 1\}, \quad \forall \mathcal{C}_m \in \mathcal{C}, \forall R_k \in \mathcal{R}, \quad (17h)$$

where $\mathbf{E} \triangleq [\mathbf{E}_1, \dots, \mathbf{E}_k, \dots, \mathbf{E}_K]$, and $\mathcal{I} \triangleq [\mathcal{I}_1, \dots, \mathcal{I}_k, \dots, \mathcal{I}_K]$. Moreover, (17a) is the multi-objective function. Constraint (17b) ensures that each user cluster is assigned to one relay, while Constraint (17c) ensures that each relay is assigned to at most one user cluster. Constraint (17d) ensures that if a relay R_k is assigned to a user cluster \mathcal{C}_m (i.e. $\mathcal{I}_{m,k} = 1$), then each user in \mathcal{C}_m must satisfy \mathbb{R}_{\min} . Constraint (17e) ensures that if a user cluster is assigned a relay, then the total transmit energy of its users does not exceed E_{\max} . Constraint (17f) enforces the SIC decoding order within each user cluster. The last two constraints define the range of values the decision variables take.

Remark 5. Problem **J-RA-EE-MAX** is a multi-objective mixed-integer non-linear fractional programming problem, which is non-convex and NP-complete [15].

Alternatively, problem **J-RA-EE-MAX** can be efficiently solved by splitting it into two problems: (1) energy-efficiency maximizing optimal power allocation over each relay, and (2) relay assignment via one-to-one matching.

IV. ENERGY-EFFICIENCY MAXIMIZATION

Assume that relay R_k is assigned to user cluster \mathcal{C}_m (i.e. $\mathcal{I}_{m,k} = 1$, while $\mathcal{I}_{m',k} = 0$, $\forall \mathcal{C}_{m'} \in \mathcal{C}$, and $m' \neq m$). In turn, the energy-efficiency function in (16) reduces to

$$\mathbb{EE}_{m,k}(\mathbf{E}_{m,k}) \triangleq \frac{\mathbb{R}_{m,k}(\mathbf{E}_{m,k})}{\mathbb{E}_{m,k}(\mathbf{E}_{m,k}) + E_k}. \quad (18)$$

Therefore, the energy-efficiency maximizing power allocation (EE-MAX-PA) problem of each user cluster $\mathcal{C}_m \in \mathcal{C}$ over each relay $R_k \in \mathcal{R}$ is expressed as

$$\mathbf{EE}\text{-MAX-PA}(\mathbf{m}, \mathbf{k}): \quad (19)$$

$$\max \mathbb{EE}_{m,k}(\mathbf{E}_{m,k}) \quad (19a)$$

$$\text{s.t. } \mathbb{R}_{n,m,k}(\mathbf{E}_{m,k}) \geq \mathbb{R}_{\min}, \quad \forall U_n \in \mathcal{C}_m, \quad (19b)$$

$$\sum_{U_n \in \mathcal{C}_m} E_{n,m,k} \leq E_{\max}, \quad (19c)$$

$$E_{1,m,k} \geq \dots \geq E_{n,m,k} \geq \dots \geq E_{U_{|\mathcal{C}_m|},m,k}, \quad (19d)$$

$$0 \leq E_{n,m,k} \leq E_{\max}, \quad \forall U_n \in \mathcal{C}_m. \quad (19e)$$

Based on **Remark 4**, the rate function of each user is not concave in $\mathbf{E}_{m,k}$ [16], implying that problem **EE-MAX-PA**(\mathbf{m}, \mathbf{k}) is not a concave maximization problem. Thus, consider the following *e*pi-multiple transformation [9]

$$\mathbb{EE}_{m,k}(\mathbf{Q}_{m,k}, \lambda) \triangleq \lambda \mathbb{R}_{m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right), \quad (20)$$

where

$$\lambda = \frac{1}{\mathbb{E}_{m,k}(\mathbf{E}_{m,k}) + E_k}, \quad \text{and } \mathbf{Q}_{m,k} = \lambda \mathbf{E}_{m,k}, \quad (21)$$

while $\mathbf{Q}_{m,k}$ is the transformed transmit energy vector. Hence, problem **EE-MAX-PA**(\mathbf{m}, \mathbf{k}) can be transformed into [9]

$$\mathbf{T}\text{-EE-MAX-PA}(\mathbf{m}, \mathbf{k}): \quad (22)$$

$$\max \lambda \mathbb{R}_{m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) \quad (22a)$$

$$\text{s.t. } \lambda \left(\mathbb{E}_{m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) + E_k \right) \leq 1, \quad (22b)$$

$$\mathbb{R}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) \geq \mathbb{R}_{\min}, \quad \forall U_n \in \mathcal{C}_m, \quad (22c)$$

$$\sum_{U_n \in \mathcal{C}_m} Q_{n,m,k} \leq \lambda E_{\max}, \quad (22d)$$

$$Q_{1,m,k} \geq \dots \geq Q_{n,m,k} \geq \dots \geq Q_{U_{|\mathcal{C}_m|},m,k}, \quad (22e)$$

$$0 \leq Q_{n,m,k} \leq \lambda E_{\max}, \quad \forall U_n \in \mathcal{C}_m, \quad (22f)$$

$$\lambda \geq 0, \quad (22g)$$

where $\lambda \left(\mathbb{E}_{m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) + E_k \right) \leq 1 \implies \sum_{U_n \in \mathcal{C}_m} Q_{n,m,k} + \lambda E_k \leq 1$, which is a linear constraint. Note that Constraint (22c) is not concave in $(\mathbf{Q}_{m,k}, \lambda)$. However, it can be rewritten as

$$\bar{\gamma}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) \geq \bar{\mathbb{R}}_{\min} \cdot \gamma_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right)$$

$$\begin{aligned} Q_{n,m,k} \xi_{n,m,k} E_k \xi_{k,bs} \geq \bar{\mathbb{R}}_{\min} \cdot \left(E_k \xi_{n,m,k} \xi_{k,bs} \sum_{U_i \in \bar{\mathcal{C}}_{n,m,k}} Q_{i,m,k} \right. \\ \left. + \xi_{k,bs} \sum_{U_i \in \mathcal{C}_m} Q_{i,m,k} + \lambda (E_k \xi_{n,m,k} + 1) \right), \end{aligned} \quad (23)$$

where $\bar{\mathbb{R}}_{\min} \triangleq 2^{2\mathbb{R}_{\min}} - 1$. Hence, (22c) is thus transformed into a linear constraint in $(\mathbf{Q}_{m,k}, \lambda)$.

Remark 6. If $(\mathbf{Q}_{m,k}^*, \lambda^*)$ is an optimal solution to **T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}), then $\mathbf{E}_{m,k}^* = \frac{\mathbf{Q}_{m,k}^*}{\lambda^*}$ is an optimal solution to **EE-MAX-PA**(\mathbf{m}, \mathbf{k}). Also, if $\mathbf{E}_{m,k}^*$ is an optimal solution to **EE-MAX-PA**(\mathbf{m}, \mathbf{k}), then $(\mathbf{Q}_{m,k}^*, \lambda^*)$ is an optimal solution to **T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) [9].

Now, note that $\mathbb{R}_{n,m,k}(\mathbf{E}_{m,k})$ can be re-written as

$$\begin{aligned} \mathbb{R}_{n,m,k}(\mathbf{E}_{m,k}) &= \bar{\mathbb{R}}_{n,m,k}(\mathbf{E}_{m,k}) - \underline{\mathbb{R}}_{n,m,k}(\mathbf{E}_{m,k}) \\ &\triangleq \frac{1}{2} \log_2 \left(\bar{\gamma}_{n,m,k}(\mathbf{E}_{m,k}) + \underline{\gamma}_{n,m,k}(\mathbf{E}_{m,k}) \right) \\ &\quad - \frac{1}{2} \log_2 \left(\underline{\gamma}_{n,m,k}(\mathbf{E}_{m,k}) \right). \end{aligned} \quad (24)$$

Thus, the objective function in (22a) is re-written as

$$\begin{aligned} &\lambda \sum_{U_n \in \mathcal{C}_m} \mathbb{R}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) \\ &= \sum_{U_n \in \mathcal{C}_m} \lambda \bar{\mathbb{R}}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) - \sum_{U_n \in \mathcal{C}_m} \lambda \underline{\mathbb{R}}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) \quad (25) \\ &\triangleq \bar{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda) - \underline{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda). \end{aligned}$$

It can be verified that $\bar{\mathbb{R}}_{n,m,k}(\mathbf{E}_{m,k})$ is concave in $\mathbf{E}_{m,k}$, and thus function $\bar{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda)$ is also concave in $(\mathbf{Q}_{m,k}, \lambda)$. Also, $\underline{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda)$ is concave in $(\mathbf{Q}_{m,k}, \lambda)$ as $\underline{\mathbb{R}}_{n,m,k}(\mathbf{E}_{m,k})$ is concave in $\mathbf{E}_{m,k}$ [9]. Now, based on (25), the objective function of problem **T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) is a difference of concave functions [17]. By introducing the auxiliary variable χ , problem **T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) is reformulated as [9]

$$\mathbf{R-T-EE-MAX-PA}(\mathbf{m}, \mathbf{k}): \quad (26)$$

$$\max \quad \bar{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda) - \chi \quad (26a)$$

$$\text{s.t.} \quad \underline{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda) \leq \chi, \quad (26b)$$

$$\lambda \left(\mathbb{E}_{m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) + E_k \right) \leq 1, \quad (26c)$$

$$\bar{\gamma}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right) \geq \bar{\mathbb{R}}_{\min} \cdot \underline{\gamma}_{n,m,k} \left(\frac{\mathbf{Q}_{m,k}}{\lambda} \right), \quad \forall U_n \in \mathcal{C}_m, \quad (26d)$$

$$\sum_{U_n \in \mathcal{C}_m} Q_{n,m,k} \leq \lambda E_{\max}, \quad (26e)$$

$$Q_{1,m,k} \geq \dots \geq Q_{n,m,k} \geq \dots \geq Q_{|U_{\mathcal{C}_m}|,m,k}, \quad (26f)$$

$$0 \leq Q_{n,m,k} \leq \lambda E_{\max}, \quad \forall U_n \in \mathcal{C}_m, \quad (26g)$$

$$\lambda \geq 0, \chi \geq 0, \quad (26h)$$

where (26a) can be verified to be concave in $(\mathbf{Q}_{m,k}, \lambda)$ [18]. Moreover, the first constraint is concave in $(\mathbf{Q}_{m,k}, \lambda)$, while the remaining constraints are linear in $(\mathbf{Q}_{m,k}, \lambda)$. Intuitively, the objective is to maximize $\bar{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda) - \chi$, which in turns maximizes $\bar{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda)$ and simultaneously minimizes χ . Also, by Constraint (26b), $\underline{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda)$ is minimized, such that the difference $\bar{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda) - \underline{\mathbb{F}}(\mathbf{Q}_{m,k}, \lambda)$ in (25) is maximized.

Remark 7. Problem **R-T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) is a concave maximization problem, and thus can be solved optimally and efficiently (i.e. within polynomial-time complexity) [18].

V. RELAY ASSIGNMENT VIA ONE-TO-ONE MATCHING

The relay assignment problem is modeled as a house allocation (HA) problem [19], where each group of agents (i.e. a user cluster) wishes to be assigned a house (i.e. a relay), and each house can be assigned to at most one group. Thus, the aim is to obtain a one-to-one user cluster-relay matching via Gale's Top Trading Cycles (TTC) mechanism [8].

A. Definitions

Definition 1 (Matching): A matching $\mu : \mathcal{R} \rightarrow \mathcal{C}$ is a one-to-one mapping of the relays to the user clusters, where $\mu \in \mathcal{M}$, and \mathcal{M} is the set of all possible matchings. Also, $\mu(R_k)$ refers to the user cluster assigned to R_k under μ .

Definition 2 (Preference Relation): Let $\mathcal{C}_m \succ_{R_k} \mathcal{C}_{m'}$ denote that R_k strictly prefers \mathcal{C}_m to $\mathcal{C}_{m'}$ (for $m \neq m'$).

Definition 3 (Preference Profile): Let \mathcal{P}_k be the preference list of R_k , which contains the user clusters ordered from the most preferred to the least. More generally, let $\mathcal{P} \triangleq \{\mathcal{P}_1, \dots, \mathcal{P}_k, \dots, \mathcal{P}_K\}$ be the relays' preference profile.

Definition 4 (Relay Assignment Problem): The relay assignment problem is defined as a tuple $(\mathcal{C}, \mathcal{R}, \mathcal{P}, \mu_0)$, with μ_0 being some initial assignment.

Definition 5 (Pareto-Domination): Let μ and μ' be two matchings, where $\mu \neq \mu'$. Then, μ is said to *Pareto-dominate* μ' if and only if [19]: (1) $\mu \succeq_{R_k} \mu'$, $\forall R_k \in \mathcal{R}$, and (2) $\mu \succ_{R_k} \mu'$, for some $R_k \in \mathcal{R}$.

Definition 6 (Relay Assignment Core): A matching μ is in the *core* of the relay assignment problem if there exists no coalition $\bar{\mathcal{R}} \subset \mathcal{R}$, and a matching $\mu' \in \mathcal{M}$, such that:

- 1) for any $R_k \in \bar{\mathcal{R}}$, $\mu'(R_k)$ is the initial user cluster assigned to some relay in $\bar{\mathcal{R}}$, and
- 2) for relays in $\bar{\mathcal{R}}$, matching μ' Pareto-dominates μ .

Definition 7 (Core-Stability): A matching μ of the relay assignment problem is *core-stable* if no subset of relays can together improve their energy-efficiency by exchanging their user clusters [8].

B. Construction of Preference Profile

To construct the preference profile \mathcal{P} , the preference list of each relay over all user clusters must be determined. In turn, problem **R-T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) is solved for each potential (user cluster, relay) combination to get $\mathbf{E}_{m,k}^* = \frac{\mathbf{Q}_{m,k}^*}{\lambda^*}$, and the resulting optimal energy-efficiency value (i.e. $\mathbb{E}\mathbb{E}_{m,k}(\mathbf{E}_{m,k}^*)$) is determined for each combination. Based on the determined energy-efficiency values, the preference list of each relay is constructed, as outlined in Algorithm 1.

Algorithm 1: Construction of Preference Profile

Input: User cluster set \mathcal{C} , and relays set \mathcal{R} .

- 1 FOR each $\mathcal{C}_m \in \mathcal{C}$
- 2 FOR each $R_k \in \mathcal{R}$
- 3 Solve problem **R-T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) to get $\mathbf{E}_{m,k}^*$;
- 4 Determine $\mathbb{E}\mathbb{E}_{m,k}(\mathbf{E}_{m,k}^*)$;
- 5 END FOR
- 6 END FOR

Output: Preference profile $\mathcal{P} = \{\mathcal{P}_1, \dots, \mathcal{P}_k, \dots, \mathcal{P}_K\}$.

Remark 8. Algorithm 1 involves the solution of $M \cdot K$ concave maximization problems, and thus can be executed efficiently, and with minimal computational-complexity [18].

Remark 9. Since problem **R-T-EE-MAX-PA**(\mathbf{m}, \mathbf{k}) is solved for each (user cluster, relay) combination in Algorithm 1, then upon convergence, the optimal transmit energy matrix $\mathbf{E}_k^* = [\mathbf{E}_{1,k}^*, \dots, \mathbf{E}_{m,k}^*, \dots, \mathbf{E}_{M,k}^*]$ is obtained $\forall R_k \in \mathcal{R}$.

C. Gale's TTC Matching Mechanism

In the first step, each relay $R_k \in \mathcal{R}$ points to the relay already assigned to its most preferred user cluster, as per \mathcal{P}_k , and μ_0 . If a cycle exists, then each relay that is part of a cycle is assigned to the user cluster it points to, and all assigned user cluster and relays are removed from the problem [8]². If there is at least one remaining user cluster and one remaining relay, then proceed to the next step (i.e. $\ell \geq 2$), until there are no more cycles (i.e. no remaining clusters or relays), and thus a matching μ is obtained, as outlined in Algorithm 2.

Algorithm 2: Relay Assignment via the TTC Mechanism

Input: $(\mathcal{C}, \mathcal{R}, \mathcal{P}, \mu_0)$.

Step 1:

- 1.1 Each $R_k \in \mathcal{R}$ points to the relay assigned to their most preferred user cluster, based on \mathcal{P}_k ;
- 1.2 IF a cycle exists
- 1.3 Assign each relay to the user cluster it points to;
- 1.4 Remove all assigned user clusters and relays from \mathcal{C} , \mathcal{R} and \mathcal{P} ;
- 1.5 END IF
- 1.6 IF there are remaining unassigned user clusters and/or relays
- 1.7 Continue to the next step;
- 1.8 END IF

Step $\ell \geq 2$: Repeat Step 1 for the remaining relays until all are assigned.

Output: Matching μ .

Remark 10. The obtained matching μ represents the relay assignment solution \mathcal{I}_k^* , $\forall R_k \in \mathcal{R}$.

D. Properties

1) *Convergence and Complexity:* Due to the finite sets of user clusters and relays, then **Algorithm 2** terminates in at most $L = \min(|\mathcal{C}|, |\mathcal{R}|)$ steps, and thus has a linear time-complexity of $\mathcal{O}(L)$ [19].

2) *Uniqueness:* The obtained matching μ is the unique matching in the core of the relay assignment problem [19].

3) *Stability:* The matching μ is core-stable, such that no relay would change its assigned user cluster [19].

E. Solution Procedure

The first step in the solution procedure for joint relay assignment and energy-efficiency maximization (SP-J-RA-EE-MAX) is to construct the preference profile via **Algorithm 1**. Then, **Algorithm 2** is executed to obtain the relay assignment solution \mathcal{I}_k^* , $\forall R_k \in \mathcal{R}$. Based on the obtained transmit energy and relay assignment solutions, the energy-efficiency value of each relay can be computed, as outlined in Algorithm 3.

VI. SIMULATION RESULTS

The simulations assume a network consisting of $N = 12$ users clustered into $M = 4$ clusters (i.e. each of 3 users), and $K = 6$ EH AF relays, as shown in Fig. 1. The maximum transmit energy per time-slot is $E_{\max} = 0.5$ J, while the noise variance is to $N_0 = 10^{-9}$ J, and the path-loss exponent is $\nu = 3$. The minimum rate per user is $\mathbb{R}_{\min} = 1.5$ bit/s/Hz.

²A cycle exists if a relay points to a relay assigned to its most preferred user cluster, and the latter relay also points to the former relay's already assigned user cluster (this also includes self-cycles) [19].

Algorithm 3: SP-J-RA-EE-MAX

Input: User cluster set \mathcal{C} and relays set \mathcal{R} .

Step 1: Construct preference profile via **Algorithm 1**.

- Store the optimal transmit energy solutions \mathbf{E}_k^* for each relay $R_k \in \mathcal{R}$, determined via **Algorithm 1**.

Step 2: Determine the relay assignment solution \mathcal{I}_k^* , $\forall R_k \in \mathcal{R}$, via **Algorithm 2**.

- Calculate the energy-efficiency value $\mathbb{E}\mathbb{E}_k^*(\mathbf{E}_k^*, \mathcal{I}_k^*)$ of each relay.

Output: \mathbf{E}_k^* , \mathcal{I}_k^* , and $\mathbb{E}\mathbb{E}_k^*(\mathbf{E}_k^*, \mathcal{I}_k^*)$, $\forall R_k \in \mathcal{R}$.

The maximum harvested energy per transmission frame at each relay is $\mathcal{E}_k^{\max} = 0.5$ J, while the battery capacity is $B_k^{\max} = 5$ J, $\forall R_k \in \mathcal{R}$. The simulations are averaged over 10^3 independent network instances, each of 10 transmission frames. The randomly generated channel coefficients vary from one network instance to another, with randomly generated energy arrivals in each transmission frame.

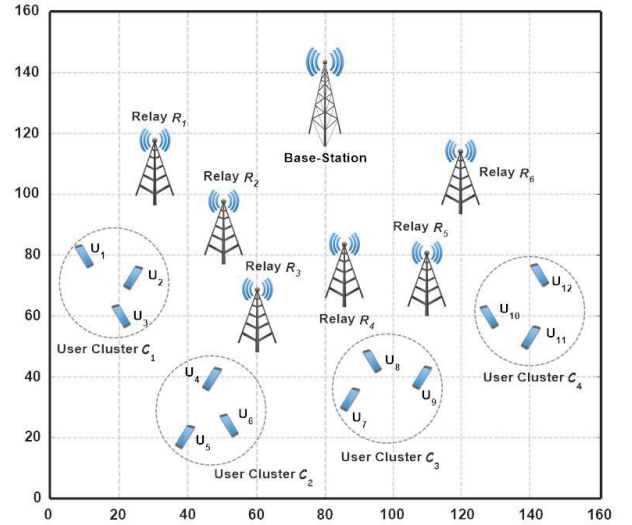


Fig. 1. Simulated Network Topology

In the simulations, the proposed **SP-J-RA-EE-MAX** scheme is compared with the following schemes³:

J-RA-EE-MAX:

In this scheme, the **J-RA-EE-MAX** problem is solved via a global optimization package, and serves an upper-bound benchmark scheme.

Random RA and EE-MAX (R-RA-EE-MAX):

This scheme randomly assigns a relay to each user cluster, followed by energy-efficiency maximization. This scheme serves a lower-bound benchmark.

In Fig. 2a, the average rate per user is illustrated, where it is evident that the **SP-J-RA-EE-MAX** scheme yields comparable user rate to the **J-RA-EE-MAX** scheme; however, both are superior to the **R-RA-EE-MAX** scheme. Moreover, the **J-RA-EE-MAX** and **SP-J-RA-EE-MAX** schemes satisfy the minimum rate of $\mathbb{R}_{\min} = 1.5$ bits/s/Hz for all users, which is not the case for the **R-RA-EE-MAX** scheme. A similar observation can be made in Fig. 2b for the average sum-rate per user cluster.

³All optimization problems are solved via the global optimization package MIDACO [20], with tolerance set to 0.001.

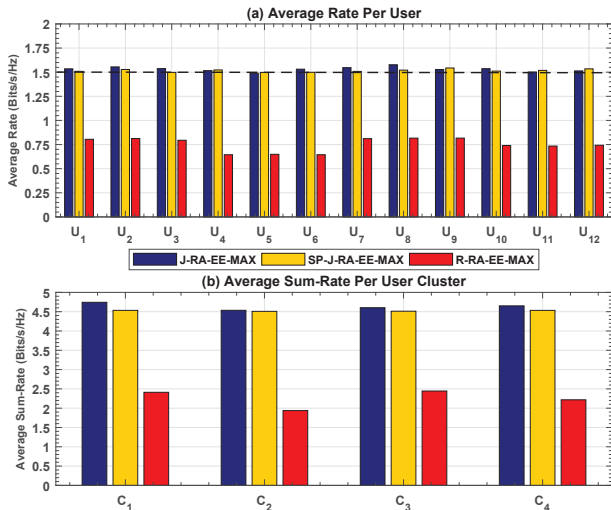


Fig. 2. Average (a) Rate Per User, and (b) Sum-Rate Per User Cluster

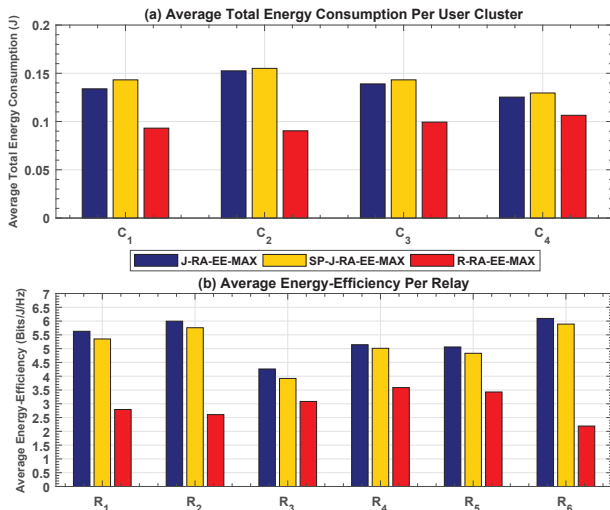


Fig. 3. Average (a) Total Energy Consumption Per User Cluster, and (b) Energy-Efficiency Per Relay

Fig. 3a shows the average total energy consumption per user cluster, where one can see that the **SP-J-RA-EE-MAX** scheme consumes slightly higher energy than the **J-RA-EE-MAX** scheme. In general, the aforementioned two schemes consume higher energy than the **R-RA-EE-MAX** scheme to satisfy the minimum rate and also maximize energy-efficiency. Fig. 3b illustrates the average energy-efficiency per relay, where it is clear that the **SP-J-RA-EE-MAX** scheme yields slightly less but comparable energy-efficiency to the **J-RA-EE-MAX** scheme; however, both are superior to **R-RA-EE-MAX**.

It has been determined via the simulations that **Algorithm 2** requires—on average—less than 3 iterations to converge, while **Algorithm 1** requires 24 iterations, as per **Remark 8**. Hence, the average number of iterations involved in the two steps of **Algorithm 3** averages to less than 27 iterations. Hence, the proposed **SP-J-RA-EE-MAX** scheme is executed efficiently, and with minimal computational-complexity.

VII. CONCLUSIONS

In this paper, the problem of joint relay assignment and energy-efficiency maximization in downlink clustered NOMA

networks has been studied. Particularly, a low-complexity solution procedure has been devised, which performs optimal energy-efficiency maximizing power allocation for each (user cluster, relay) combination, and assigns relays to user clusters in linear time-complexity. Simulation results have been presented, where the proposed solution procedure has been shown to yield comparable energy-efficiency per relay to the **J-RA-EE-MAX** scheme, while satisfying the users' QoS constraints.

ACKNOWLEDGEMENTS

This work was partially supported by the Kuwait Foundation for the Advancement of Sciences (KFAS), under project code PN17-15EE-02. An extended version of this paper can be found in [21].

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [2] S. Li, L. D. Xu, and S. Zhao, "The Internet of Things: A survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, Apr. 2015.
- [3] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C.-L. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Comms. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.
- [4] Q. Liu, T. Lv, and Z. Lin, "Energy-efficient transmission design in cooperative relaying systems using NOMA," *IEEE Communications Letters*, vol. 22, no. 3, pp. 594–597, Mar. 2018.
- [5] B. Ning, W. Hao, A. Zhang, J. Zhang, and G. Giu, "Energy efficiency-delay tradeoff for a cooperative NOMA system," *IEEE Communications Letters*, vol. 23, no. 4, pp. 732–735, Apr. 2019.
- [6] B. V. Nguyen, Q.-D. Vu, and K. Kim, "Analysis and optimization for weighted sum-rate in energy harvesting cooperative NOMA systems," *IEEE Trans. Veh. Tech.*, vol. 67, no. 12, pp. 12 379–12 383, Dec. 2018.
- [7] M. W. Baidas, E. Alsusa, and K. A. Hamdi, "Joint subcarrier assignment and weighted-sm energy-efficiency power allocation in multi-carrier uplink NOMA relay networks," *Phys. Commun.*, vol. 36, pp. 1–17, 2019.
- [8] L. Shapley and H. Scarf, "On cores and indivisibility," *Journal of Mathematical Economics*, vol. 1, pp. 23–37, Mar. 1974.
- [9] J. Shi, "A combined algorithm for fractional programming," *Annals of Operations Research*, vol. 103, no. 1, 2001.
- [10] A. M. Ortiz, D. Hussein, S. Park, S. N. Han, and N. Crespi, "The cluster between Internet of Things and social networks: Review and research challenges," *IEEE IoT Journal*, vol. 1, no. 3, pp. 206–215, Jun. 2014.
- [11] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, 2008.
- [12] M. W. Baidas, M. Al-Mubarak, E. Alsusa, and M. K. Awad, "Joint subcarrier assignment and global energy-efficient power allocation for energy-harvesting two-tier downlink NOMA HetNets," *IEEE Access*, vol. 7, no. 1, pp. 163 556–163 577, Nov. 2019.
- [13] S. Lee, D. B. da Costa, Q.-T. Vien, T. Q. Duong, and R. T. de Sousa, "Non-orthogonal multiple access schemes with partial relay selection," *IET Communications*, vol. 11, no. 6, pp. 846–854, May 2017.
- [14] T. Rapesak, "On pseudolinear functions," *European Journal of Operations Research*, vol. 50, pp. 353–360, 1991.
- [15] M. E.-A. Cherqui and M. Moulai, "An exact method for a discrete multi-objective linear fractional optimization," *Journal of Applied Mathematics and Decision Sciences*, vol. 12, 2008, DOI: 10.1155/2008/760191.
- [16] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3, pp. 185–399, 2014.
- [17] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," *IEEE Trans. on Wireless Comms.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2003.
- [19] A. Abdulkadrioglu and T. Sonmez, "House allocation with existing tenants," *Journal of Econ. Th.*, vol. 88, no. 2, pp. 233–260, Oct. 1999.
- [20] M. Schlueter, "MIDACO software performance on interplanetary trajectory benchmarks," *Adv. Space Res.*, vol. 54, no. 4, pp. 744–754, 2014.
- [21] M. W. Baidas, "Joint relay assignment and energy-efficiency maximization in energy-harvesting downlink/uplink clustered nonorthogonal multiple-access networks," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 6, p. e3962, Jun. 2020.