Resource Allocation and EE-SE Tradeoff for H-CRAN with NOMA-Based D2D Communications

Jingpu Wang[1], Xin Song[1*] and Li Dong[1]
[1] College of Computer Science and Engineering, Northeastern University
NO. 3-11, Wenhua Road, Heping District, Shenyang - China
San Diego, CA 92093 - USA
[e-mail: 1610533@stu.neu.edu.cn]
[e-mail: sxin78916@neuq.edu.cn]
[e-mail: dongll0801@163.com]
*Corresponding author: Xin Song

Abstract

We propose a general framework for studying resource allocation problem and the tradeoff between spectral efficiency (SE) and energy efficiency (EE) for downlink traffic in power domain—non-orthogonal multiple access (PD-NOMA) and device to device (D2D) based heterogeneous cloud radio access networks (H-CRANs) under imperfect channel state information (CSI). The aim is jointly optimize radio remote head (RRH) selection, spectrum allocation and power control, which is formulated as a multi-objective optimization (MOO) problem that can be solved with weighted Tchebycheff method. We propose a low-complexity algorithm to solve user association, spectrum allocation and power coordination separately. We first compute the CSI for RRHs. Then we study allocating the cell users (CUs) and D2D groups to different subchannels by constructing a bipartite graph and Hungrarian algorithm. To solve the power control and EE-SE tradeoff problems, we decompose the target function into two subproblems. Then, we utilize successive convex program approach to lower the computational complexity. Moreover, we use Lagrangian method and KKT conditions to find the global optimum with low complexity, and get a fast convergence by subgradient method. Numerical simulation results demonstrate that by using PD-NOMA technique and H-CRAN with D2D communications, the system gets good EE-SE tradeoff performance.

1. Introduction

$\mathbf{P}$ower domain—non-orthogonal multiple access (PD-NOMA) has become a key issue of the novel energy and spectrum efficient technologies due to a higher network capacity compared with orthogonal multiple access (OMA) in the fifth generation (5G) environment [1]. The energy efficiency (EE) is studied in a NOMA enabled heterogeneous cloud radio access network (H-CRAN) in [2]. Apart from NOMA, the device-to-device (D2D) communication has been an essential way to alleviate the upcoming traffic pressure. Unlike the traditional concept of "D2D pair", the concept of "D2D group" in which several D2D receivers are capable of receiving information from one D2D transmitter is presented in [3]. .

In this work, we introduce multi-tier heterogeneous networks to construct a novel heterogeneous cloud radio network (H-CRAN) with PD-NOMA and D2D groups. The main parts of the H-CRAN architecture are the remote radio heads (RRHs), fiber links and baseband unit (BBU). The RRHs include a high power node (HPN) which transmits high power to the users and several lower power nodes (LPNs) which accommodate fewer users than HPN. A centralized signal processing is used in BBU pool which reduces the manufacturing and operating cost [4], [5]. However, enormous data can be transmitted over backhaul links. Hence, we investigate the backhaul bottleneck of H-CRAN.

In practice, the BS always has imperfect CSI. To deal with it, we assume that a channel estimation error model where the BS only knows the estimated channel gain and a prior knowledge of the variance of the estimation error [6]. In this paper, we consider D2D group aided PD-NOMA based C-RAN with imperfect CSI, in which D2D groups can reuse the same subchannel occupied by cellular users to improve the spectrum utilization. To the best of our knowledge, the existing works cannot use joint user association and power allocation algorithms with EE-SE tradeoff in NOMA and D2D enhanced C-RAN with channel estimation. The proposed algorithm first transforms the probabilistic multi-objective optimization (MOO) problem into a non-probabilistic single-objective (SOO) optimization problem by weighted Tchebycheff method. With successive convex programming (SCP), the non-convex optimization problem can be transformed into a convex optimization problem. Then it iteratively computes the sub-optimal power allocation coefficients for the optimization problem by Lagrangian dual multiplier method and KKT conditions. The main contributions of this work can be summarized as follows:

1) To the best of our knowledge, the proposed algorithm is firstly used in H-CRAN for the D2D underlaying cellular network combining with the NOMA technology, which is a candidate technology for future networks. By successive interference cancellation (SIC), users can achieve higher data rates in PD-NOMA than in OMA communication systems.

2) Considering imperfect CSI, we firstly transform the probabilistic problem to a non-probabilistic problem. By using statistical methods to more precisely estimate parameters in the objective function, the outage probability constrains in the optimization problem can be solved. Considering the backhaul capacity of H-CRAN, we propose a delay constraint by computing user average service time through queue theory. Then, we transform the MOO non-convex problem into a SOO convex problem by using weighted Tchebycheff method and SCP. At last, by the Lagrangian dual multiplier method with KKT conditions and the gradient descent method, the power allocation coefficients can be obtained iteratively.

3) The proposed algorithm is evaluated by numerous times of Monte Carlo simulations, which has a fast converge speed. And the power allocation scheme used in an NOMA and

D2D aided cooperative relay system with imperfect CSI can achieve a higher sum data rate compared with that used in networks without D2D or that in OMA systems. Also, the proposed algorithm gives the maximum EE and SE points in different circumstances.

The rest of the paper is organized as follows. The channel model and problem formulation are introduced in Section II. The proposed resource allocation algorithm is elaborated in Section III. In Section IV, the simulation results are presented, while Section V finally draws conclusions of the paper.

2. System Model

2.1 Downlink Network Model

We consider a downlink H-CRAN scenario involving one high-power node (HPN) and six low-power nodes (LPNs) as RRHs, i.e., $m \in M, M = \{0, \ldots, m, \ldots, 6\}$, where the first RRH represents the HPN. Each RRH has L OFDM subchannels (SCs) for cell users (CUs). Each D2D group (DG) uses the same SC as a CU in NOMA. For simplicity, we suppose that there are at most two D2D users (DUs) in one DG. We also suppose that there are N CUs, i.e., $n \in N, N = \{0, \ldots, n, \ldots, N\}$, and K DUs in the network, where $K \leq 2N$. The parameters and symbols used in the work are summarized in **Table 1**.

In **Fig. 1**, we place the LPNs at the edge of the cell, and they can be reused by adjacent H-CRANs. In the cell, each LPN occupies $L/2$ SCs; thus, signals from adjacent LPNs do not have mutual interference. In addition, all the users in small cells receive interference from the HPN, and if a CU is sufficiently close to a DU that is in the same SC, it also receives interference from the DU.
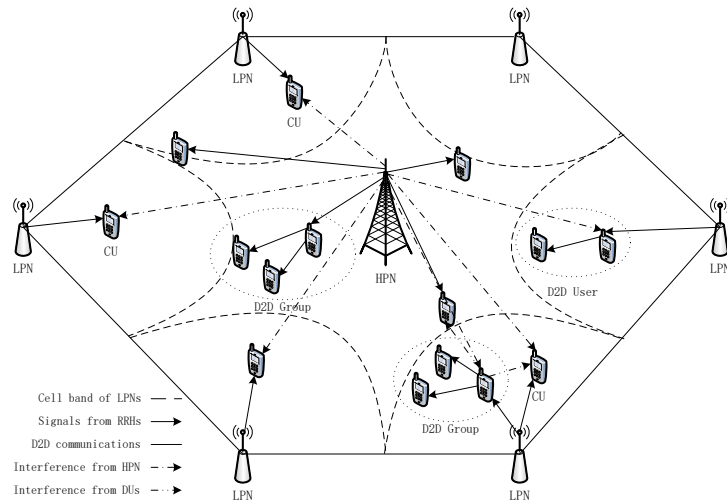

Fig. 1. System model

We assume that the m-th RRH, the n-th CU, and the k-th DU are multiplexed on the l-th SC. The signal of CUs transmitted by the RRHs through SC l can be expressed as $X_n = A_{m,n} x_{m,n}^{(l)} \sqrt{P_{m,n}} s_{m,n}$. The signal of DUs transmitted by the RRHs through SC l can be expressed as $Y_k = y_{n,k}^{(l)} \sqrt{P_{n,k}} s_{n,k}$, where $s_{m,n}$ and $s_{n,k}$ are the modulated symbols transmitted from RRH m to CU n and from CU n to DU k, respectively. Further, $P_{m,n}$ and $P_{n,k}$ represent the power

transmitted from RRH m to CU n and from CU n to DU k, respectively. In addition, $x_{m,n}^{(l)}$ is a binary variable of CU n on SC l, where $x_{m,n}^{(l)} = 1$ if CU n is allocated to SC l, while $y_{n,k}^{(l)}$ is a binary variable of DU k on SC l, where $y_{n,k}^{(l)} = 1$ if DU k is allocated to SC l. $A_{m,n}$ is a binary variable of the RRH selection indicator, where $A_{m,n} = 1$ implies that CU n is selected by RRH m.

Table 1. Table of symbols.

Symbol	Definition / Description
$M = \{0, \ldots, m, \ldots, 6\}$	RRHs set
$N = \{0, \ldots, n, \ldots, N\}$	CU set
N	Number of CUs
L	Number of subchannels
K	Number of DUs
X_n, Y_k	Signals of CU n and DU k
$A_{m,n}$	RRH allocation indicator
$x_{m,n}^{(l)}, y_{n,k}^{(l)}$	Subchannel allocation indicator
$s_{m,n}, s_{n,k}$	Modulated symbols for CUs and DUs
$P_{m,n}, P_{n,k}$	Power transmitted from RRH m to CU n
$P_{n,k}$	Power transmitted from RRH m to DU k
$\phi_{m,n,l}$	SINR of user n on subchannel l in RRH m
$h_{m,n,l}$	Channel gain of user n on subchannel l in RRH m
$g_{m,n,l}$	Rayleigh fading channel gain
$\Gamma_{m,n,l}$	Distance between transmitter m and receiver n
$\sigma_{m,n,l}^2$	Noise power
$I_{m,n,l}$	Total interference received by CU n
I_{out}^C	Inter-cell interference from HPN
I_{in}^D	Intra-cell interference from DUs
I_{out}^D	Intra-cell interference from DUs
$e_{m,n,l}$	Estimated error
$r_{m,n,l}$	scheduled data rate of CU n selected by RRH m on SC l
$C_{m,n}$	Upper bound of $r_{m,n,l}$
R^C, R^D	Outage sum rate of CUs and DUs
η_{SE}	Network SE
η_{EE}	Network EE
P_{tot}	Network total power consumption
P_f^H, P_f^L	Fiber backhaul power of HPN and LPN
P_{cir}^H, P_{cir}^L	Circuit power of HPN and LPN
β_m	Power amplifier efficiency
$\varepsilon_{out}^C, \varepsilon_{out}^D$	Outage requirements for CUs and DUs
$\phi_{thr}^C, \phi_{thr}^D$	SINR thresholds for the CUs and DUs

$P_m^{\max}$	RRH m maximum allowable transmit power
s_{k1}, s_{k2}	power allocation coefficients for DUs
$\overline{z}$	Packet size
T_n	Maximum delay requirement
G_n	Poisson-distribution arrival rate of CU n
ω	Weighting parameter
W	Weight matrix of transmission rates
θ	Iteration index
$\lambda_1, \lambda_2, \mu_1, \mu_2$	Lagrangian multipliers vectors
Δ	Convergence threshold

For the general case, the signal-to-interference-plus-noise ratio (SINR) of the n-th CU can be written as $\phi_{m,n,l} = P_{m,n,l} h_{m,n,l}^2 / \left(\sigma_{m,n,l}^2 + I_{m,n,l} \right)$, where $h_{m,n,l}^2$ is the desired CU signal channel gain on SC l. We assume that $h_{m,n,l} = \Gamma_{m,n,l} g_{m,n,l}$, where $\Gamma_{m,n,l}$ is the distance between transmitter m and receiver n on SC l and $g_{m,n,l} : \mathcal{N}_C \left(0, \ \sigma_{m,n,l}^2 \right)$ represents the Rayleigh fading channel gain between transmitter m and receiver n. Further, $\sigma_{m,n,l}^2$ is the noise power while $I_{m,n,l}$ denotes the total interference received by CU n, including the inter-cell and intra-cell interference from the CUs and DUs. Thus, $I_{m,n,l} = I_{out}^C + I_{in}^D + I_{out}^D$, where I_{out}^C represents the interference from the HPN (j=0) and $I_{out}^C = \sum_{\substack{j \in M, j \neq m \\ i \in N, i \neq n}} A_{j,i} x_{j,i}^{(l)} P_{j,i} h_{j,n}^2$. Further, I_{in}^D denotes the intra-cell interference from DG k to CU n. $I_{in}^D = y_{n,k}^{(l)} P_{n,k} h_{k,n}^2$. In addition, I_{out}^D denotes the inter-cell interference from DG $k^{'}$ to CU n. $I_{out}^D = \sum_{\substack{k^{'} \in K, k^{'} \neq k \\ i \in N, i \neq n}} y_{i,k^{'}}^{(l)} P_{i,k} h_{k^{'},n}^2$.

As for the DUs and DGs, the SINR of the k-th DU can be written as $\phi_{n,k,l} = P_{n,k,l} h_{n,k,l}^2 / \left(\sigma_{n,k,l}^2 + I_{n,k,l}^D \right)$, where $h_{n,k,l}^2$ is the desired DU signal channel gain on SC l and $I_{n,k,l}^D$ denotes the total interference received by DU k, including the inter-cell interference from the CUs and DUs. Further, $I_{n,k,l}^D = I_{out}^C + I_{out}^D$, where I_{out}^C represents the interference from the HPN (j=0) to the k-th DU. $I_{out}^C = \sum_{j \in N, j \neq n} A_{j,k} x_{j,k}^{(l)} P_{j,k} h_{j,k}^2$. In addition, I_{out}^D denotes the inter-cell interference from D2D user $k^{'}$ to DU k. $I_{out}^D = \sum_{k^{'} \in K, k^{'} \neq k} y_{k^{'},k}^{(l)} P_{k^{'},k} h_{k^{'},k}^2$. Moreover, if the D2D users form a group as DG k, the SINR of the k-th DG can be $\phi_{n,k,l} = P_{n,k,l} h_{n,k,l}^2 / \left(\sigma_{n,k,l}^2 + I_{n,k,l} \right)$, where $h_{n,k,l}^2$ is the desired DG signal channel gain on SC l and $I_{n,k,l}$ denotes the total interference received by DG k. Each DG has two DUs ($k1$ and $k2$), by the PD–NOMA and SIC techniques, if the SINR of one of the two users ($\phi_{k2}(k2)$) is not better than that of the other, i.e., $\phi_{n,k1,l}(k2) \geq \phi_{n,k2,l}(k2)$, then the transmission powers of the two users are expressed as $P_{k2} \geq P_{k1}$. The SINR of user $k1$ of the k-th DG can be written as $\phi_{n,k1,l} = P_{n,k1,l} h_{n,k1,l}^2 / \left(\sigma_{n,k1,l}^2 + I_{n,k1,l} \right)$, where $h_{n,k1,l}^2$ is the desired channel gain of user $k1$ of the k-th DG signal on SC l while $I_{n,k1,l}$ denotes the total interference received by user $k1$ of the k-th DG. $I_{n,k1,l} = I_{out}^D + I_{out}^C$, where $I_{out}^C = \sum_{j \in N, j \neq n} A_{j,k1} x_{j,k1}^{(l)} P_{j,k1} h_{j,k1}^2$ represents the interference

from the HPN (j=0) to DU $k1$. $I_{out}^{D} = \sum_{k' \in K, k' \neq k1} y_{k',k1}^{(l)} P_{k',k1} h_{k',k1}^{2}$ denotes the inter-cell interference from D2D user k' to DU $k1$ on the same SC. The interference received by user $k1$ is exactly the same as that received by a DU that does not form any DG. Hence, in the remainder of this article, we assume that every DU belongs to a DG and one DG can involve one or two DUs, i.e., $k \in K, K = \{0, \ldots, k, \ldots, K\}$, $K \leq N$. The SINR of user $k2$ of the k-th DG can be written as $\phi_{n,k2,l} = P_{n,k2,l} h_{n,k2,l}^{2} / (\sigma_{n,k2,l}^{2} + I_{n,k2,l})$, where $h_{n,k2,l}^{2}$ is the desired channel gain of user $k2$ of the k-th DG signal on SC l while $I_{n,k2,l}$ denotes the total interference received by user $k2$ of the k-th DG. Thus, $I_{n,k2,l} = I_{out}^{D} + I_{out}^{C} + I_{in}^{C}$, where I_{out}^{C} represents the interference from the HPN (j=0) to DU $k2$. $I_{out}^{C} = \sum_{j \in N, j \neq n} A_{j,k2} x_{j,k2}^{(l)} P_{j,k2} h_{j,k2}^{2}$. Further, I_{out}^{D} denotes the inter-cell interference from D2D user k' to DU $k2$ on the same SC. $I_{out}^{D} = \sum_{k' \in K, k' \neq k2} y_{k',k2}^{(l)} P_{k',k2} h_{k',k2}^{2}$. In addition, I_{in}^{D} denotes the intra-cell interference from DU $k1$ to DU $k2$ of the k-th DG on the same SC. $I_{in}^{D} = \sum_{k1 \in K, k1 \neq k2} y_{k1,k2}^{(l)} P_{k1,k2} h_{k1,k2}^{2}$.

2.2 Channel Model

In practice, it is challenging to acquire the perfect CSI of each user. Using the minimum mean square error for the channel estimation error model [7], we can establish the model of the Rayleigh fading coefficient between RRH m and CU n as $g_{m,n,l} = \hat{g}_{m,n,l} + e_{m,n,l}$, where $\hat{g}_{m,n,l} : \mathcal{N}_{C}(0, 1-\sigma_{e}^{2})$ is the estimated channel gain and $e_{m,n,l}$ is the estimated error, which follows a complex Gaussian distribution with mean zero and variance σ_{e}^{2}. We also assume that $\hat{g}_{m,n,l}$ and $e_{m,n,l}$ are uncorrelated. Similarly, we can establish model of the Rayleigh fading coefficient between CU n and DG k as $g_{n,k,l} = \hat{g}_{n,k,l} + e_{n,k,l}$, where $\hat{g}_{n,k,l} : \mathcal{N}_{C}(0, 1-\sigma_{e}^{2})$ is the estimated channel gain and $e_{n,k,l}$ is the estimated error, which follows a complex Gaussian distribution with mean zero and the variance σ_{e}^{2}. We also assume that $\hat{g}_{n,k,l}$ and $e_{n,k,l}$ are uncorrelated.

With perfect CSI, according to Shannon's formula, the maximum achievable channel capacity of CU n allocated to RRH m on SC l can be written as $C_{m,n,l} = B_{sc} \log_{2}(1 + \phi_{m,n,l})$, where B_{sc} is the bandwidth of SC l. However, if $\hat{h}_{m,n,l} = \Gamma_{m,n,l} \hat{g}_{m,n,l}$, where $\Gamma_{m,n,l}$ is the distance between RRH m and CU n, and the scheduled data rate of CU n selected by RRH m on SC l can be rewritten as $r_{m,n,l} = B_{sc} \log_{2}(1 + \hat{\phi}_{m,n,l})$, where $B_{sc} = 1$ is assumed in this study and $\hat{\phi}_{m,n,l} = P_{m,n,l} \hat{h}_{m,n,l}^{2} / (\sigma_{m,n,l}^{2} + I_{m,n,l})$. Thus, the network's outage sum rate of CUs can be $R^{C}(A,x,P) = \sum_{m \in M} \sum_{n \in N} A_{m,n} x_{m,n}^{(l)} r_{m,n} \Pr[r_{m,n} \leq C_{m,n} | \hat{g}_{m,n}]$, where $\Pr[r_{m,n} \leq C_{m,n} | \hat{g}_{m,n}]$ is the probability of $r_{m,n} \leq C_{m,n}$ and $C_{m,n}$ represents the upper bound of $r_{m,n}$. Similarly, the maximum achievable channel capacity of DG k allocated to the n-th CU on SC l can be $C_{n,k,l} = B_{sc} \log_{2}(1 + \phi_{n,k,l})$, where B_{sc} is the bandwidth of SC l. If $\hat{h}_{n,k,l} = \Gamma_{n,k,l} \hat{g}_{n,k,l}$, the scheduled data rate of DG k selected by the n-th CU on SC l can be as $r_{n,k,l} = B_{sc} \log_{2}(1 + \hat{\phi}_{n,k,l})$,

where $\hat{\phi}_{n,k,l} = P_{n,k,l}\hat{h}_{n,k,l}^2 / \left(\sigma_{n,k,l}^2 + I_{n,k,l}\right)$. Thus, the outage sum rate of DGs can be $R^D\left(y,P\right) = \sum_{\substack{n\in N \\ k\in K}} y_{n,k}^{(l)} r_{n,k} \Pr\left[r_{n,k} \le C_{n,k} \mid \hat{g}_{n,k}\right]$, where $\Pr\left[r_{n,k} \le C_{n,k} \mid \hat{g}_{n,k}\right]$ is the probability of $r_{n,k} \le C_{n,k}$. $C_{n,k}$ represents the upper bound of $r_{n,k}$.

2.3 Tradeoff between EE and SE

The multi-objective optimization (MOO) problem can be formulated as a joint optimization problem of user association and power allocation, which can be written as

$$\min_{A,x,y,p} -\eta_{SE}(A,x,y,p),$$

$$\min_{A,x,y,p} P_{tot}(A,x,y,p),$$

$$s.t. \ C1: A_{m,n} \in \{0,1\}, \forall m,n,$$

$$C2: \sum_m A_{m,n} \le 1, \forall n,$$

$$C3: x_{m,n} \in \{0,1\}, \forall m,n,$$

$$C4: \sum_n x_{m,n} \le L, m = 0,$$

$$C5: \sum_n x_{m,n} \le L/2, \forall m, m \neq 0,$$

$$C6: y_{n,k} \in \{0,1\}, \forall n,k,$$

$$C7: \sum_n y_{n,k} \le 1, \forall k,$$

$$C8: \sum_k y_{n,k} \le 2, \forall n,$$

$$C9: \phi_n^C \ge \phi_{thr}^C, \ \phi_k^D \ge \phi_{thr}^D, \ \forall n,k,$$

$$C10: \sum_{n\in N} A_{m,n} x_{m,n}^{(l)} P_{m,n} \le P_m^{\max}, \forall m,$$

$$C11: \phi_{n,k1} \ge \phi_{n,k1}^{thr}, \ \phi_{n,k2} \ge \phi_{n,k2}^{thr}, \ \forall n,k1,k2,$$

$$C12: s_{k1} \ge 0, \ s_{k2} \ge 0, \ \forall k1,k2,$$

$$C13: 2^{s_{k1}} + 2^{s_{k2}} \le 1, \ \forall k1,k2,$$

$$C14: \Pr\left[C_{m,n} < r_{m,n} \mid \hat{g}_{m,n}\right] \le \varepsilon_{out}^C, \ \forall m,n,$$

$$C15: \Pr\left[C_{n,k} < r_{n,k} \mid \hat{g}_{n,k}\right] \le \varepsilon_{out}^D, \ \forall n,k,. \tag{1}$$

where the network SE (bits/s/Hz) is calculated by the system sum data rate per unit bandwidth from the channel model, which can be represented by $\eta_{SE} = R^C + R^D$. The system total power consumption is $P_{tot} = P_{tot}^C + P_{tot}^D$, where $P_{tot}^C = P_f^H + 6P_f^L + \beta_m \sum_{m\in M} \sum_{n\in N} A_{m,n} x_{m,n}^{(l)} P_{m,n} + P_{cir}^H + 6P_{cir}^L$, where P_f^H is the fiber-based backhaul power of the HPN, P_f^L is the fiber-based backhaul power of the LPN, β_m is the power amplifier efficiency of the m-th RRH to each user in the network, P_{cir}^H is the circuit power of the HPN, and P_{cir}^L is the circuit power of the LPN. Further, $P_{tot}^D = \sum_{\substack{n\in N \\ k\in K}} y_{n,k}^{(l)} P_{n,k}$.

For green communication, we study the EE of the network, which can be expressed as $\eta_{EE} = \eta_{SE} / P_{tot}$. The SE and EE are conflicting parameters in terms of optimizing the performance of a wireless network with limited radio resources [8], [9]. Hence, it is necessary

to achieve a tradeoff between the EE and the SE of the PD–NOMA and D2D-communication-enhanced H-CRAN. According to the η_{EE} formula, we must simultaneously maximize η_{SE} and minimize the total power consumption P_{tot} to get the maximum η_{EE}.

In (1), $C1$ and $C2$ specify that each CU can be allocated to only one RRH in every time phase. Further, $C3-C5$ imply that each CU can select only one SC. In addition, the LPNs have $L/2$ SCs while the HPN has L SCs. This can guarantee that adjacent LPNs occupy different SCs to avoid interference at the same frequency. Moreover, $C6-C8$ indicate that each DU can select only one CU, while one CU can be associated with a DG. The constraint $C9$ guarantees the QoS by limiting the network service delay. Note that ϕ_{thr}^{C} and ϕ_{thr}^{D} are the SINR thresholds for the CUs and DUs, respectively. The constraint $C10$ ensures that the transmission power for RRH m in the downlink network is not higher than its maximum transmitted power, which is defined by $P_m^{\max}$. The constraint $C11$ gives the minimum SINR for each DU, where $\phi_{n,k1}^{thr} = \phi_{n,k2}^{thr} = \phi_{thr}^{D}$, while the constrain $C12$ implies that the power allocation coefficient s_{k1} and s_{k2} are non-negative. $C13$ limits the upper bound of the transmission power for the DGs. ε_{out}^{C} and ε_{out}^{D} are the channel outage probability requirements for the CUs and DUs as indicated by $C14$ and $C15$.

As ϕ_{thr}^{C} and ϕ_{thr}^{D} are related to the network data transmission delay in the fiber links. By the scheduled streaming data of CU n and its traffic characteristic (T_n, G_n), the delay constraint is $r_n \geq 2G_n\overline{z} / \left((2 + 2G_nT_n - \sqrt{(2 + 2G_nT)^2 - 8G_nT_n})B_{sc} \right)$ (bits/s/Hz) in here $\overline{z}$ is a random variable of the packet size in bits, while T_n and G_n are the desired maximum delay requirement and the Poisson-distribution arrival rate of CU n [4]. ϕ_{thr}^{C} and ϕ_{thr}^{D} are from Shannon's capacity formula.

For consistent comparison, we normalize the two target functions in Eq. (1) as follows:

$$\min_{A,x,y,p} F_1(A,x,y,p) = \frac{\eta_{SE}^{\max} - \eta_{SE}(A,x,y,p)}{\eta_{SE}^{\max} - \eta_{SE}^{\min}},$$

$$\min_{A,x,y,p} F_2(A,x,y,p) = \frac{P_{tot}}{P_{\max}}, \tag{2}$$

where $P_{\max}$ represents the maximum total power consumption while $\eta_{SE}^{\max}$ and $\eta_{SE}^{\min}$ are the maximum and minimum SE (system sum data rate), respectively.

The MOO problem (1) is a mixed-integer problem because of the integer variables $x_{m,n}$ and $y_{n,k}$. Owing to the existing inter-cell and intra-cell interference described in Section 2.1, the problem is a non-concave problem, which is generally NP-hard. Hence, we first transform Eq. (1) into a single-objective optimization (SOO) problem with lower computational complexity by using the weighted Tchebycheff method as follows:

$$\min_{A,x,y,p} \max \left\{ \omega F_1(A,x,y,p), (1-\omega)F_2(A,x,y,p) \right\},$$

$$s.t. \ C1-C15, \tag{3}$$

where the weighting parameter $\omega \in [0,1]$ represents the relative importance of the two objectives. The method has been shown to be sufficient and necessary for achieving Pareto

optimality and obtaining the unique optimal solution (A^*, x^*, y^*, p^*) [10]. After introducing an additional adaptive variable φ, and the transformed problem is written as

$$\min_{A,x,y,p,\varphi} \varphi,$$

$$s.t. \ \ C1 - C15,$$

$$C16: \omega \left(\frac{\eta_{SE}^{\max} - \eta_{SE}(A, x, y, p)}{\eta_{SE}^{\max} - \eta_{SE}^{\min}} \right) \le \varphi,$$

$$C17: (1 - \omega) \left(\frac{P_{tot}}{P_{\max}} \right) \le \varphi. \tag{4}$$

2.4 Optimization Problem Transformation

The optimization problem given by Eq. (4) is a non-convex problem with probabilistic constraints $C1$ and $C2$. In practice, it is rather difficult to obtain the global optimal power allocation with affordable, systematic, and computationally efficient methods, and the optimal problem cannot be solved with acceptable complexity in polynomial time. To decrease the complexity, we transform the probabilistic mixed-integer problem into a non-probabilistic problem, and we then use a series of algorithms to efficiently solve the joint resource allocation optimization problem.

According to [7], we first rewrite the data rate in section 2.2 as

$$r_{m,n} = \log_2 \left(1 + \hat{\phi}_{m,n} \right) = \log_2 \left(1 + \frac{\alpha_{m,n}}{\beta_{m,n}} \right), \forall m, n, \tag{5}$$

where $\hat{\phi}_{m,n} = \alpha_{m,n} / \beta_{m,n} = 2^{r_{m,n}} - 1$ and $\alpha_{m,n} = \beta_{m,n} \left(2^{r_{m,n}} - 1 \right)$.

The channel capacity for CU n is given by $C_{m,n} = \log_2 \left(1 + \phi_{m,n} \right) = \log_2 \left(1 + \mu_{m,n} / \upsilon_{m,n} \right), \forall m, n$.

And $\phi_{m,n} = \mu_{m,n} / \upsilon_{m,n} = 2^{C_{m,n}} - 1, \mu_{m,n} = \upsilon_{m,n} \left(2^{C_{m,n}} - 1 \right)$. The bound of the outage probability is

$$\Pr \left[C_{m,n} < r_{m,n} \mid \hat{g}_{m,n} \right] \le \varepsilon_{out}^C = \Pr \left[\phi_{m,n} < \hat{\phi}_{m,n} \mid \hat{g}_{m,n} \right] \le \varepsilon_{out}^C. \tag{6}$$

According to the total probability theorem, the outage probability can be given by

$$\Pr \left[\phi_{m,n} < \hat{\phi}_{m,n} \mid \hat{g}_{m,n} \right] \le \varepsilon_{out}^C = \Pr \left[\frac{\mu_{m,n}}{\upsilon_{m,n}} < 2^{r_{m,n}} - 1 \mid \hat{g}_{m,n} \right]$$

$$= \Pr \left[E1 \right] \cdot \Pr \left[\mu_{m,n} \le \alpha_{m,n} \mid \hat{g}_{m,n} \right]$$

$$+ \Pr \left[E2 \right] \cdot \Pr \left[\mu_{m,n} > \alpha_{m,n} \mid \hat{g}_{m,n} \right] \tag{7}$$

where

$$\Pr \left[E1 \right] = \Pr \left[\frac{\mu_{m,n}}{\upsilon_{m,n}} < 2^{r_{m,n}} - 1 \mid \mu_{m,n} \le \alpha_{m,n}, \ \hat{g}_{m,n} \right], \tag{8}$$

$$\Pr \left[E2 \right] = \Pr \left[\frac{\mu_{m,n}}{\upsilon_{m,n}} < 2^{r_{m,n}} - 1 \mid \mu_{m,n} > \alpha_{m,n}, \ \hat{g}_{m,n} \right]. \tag{9}$$

According to [11], we can transform the outage probability constraint $C14$ into the following formulas: $\Pr\left[\mu_{m,n} \leq \alpha_{m,n} \mid \hat{g}_{m,n}\right] = \varepsilon_{out}^C / 2$ and $\Pr\left[\upsilon_{m,n} \geq \beta_{m,n} \mid \hat{g}_{m,n}\right] \leq \varepsilon_{out}^C / 2$.

By substituting the probabilistic constraints above into Eq. (4), we can transform the target problem into a revised optimization problem as follows. For CU n, with the Markov inequality, we can get

$$\Pr\left[\upsilon_{m,n} \geq \beta_{m,n} \mid \hat{g}_{m,n}\right] = \Pr\left[\Gamma_{m,n}^2 \left|g_{m,n}\right|^2 P_{m,n} \geq \beta_{m,n} - \sigma^2 \mid \hat{g}_{m,n}\right]$$

$$\leq \frac{E\left[\Gamma_{m,n}^2 \left|g_{m,n}\right|^2 P_{m,n}\right]}{\beta_{m,n} - \sigma^2}$$

$$= \frac{\Gamma_{m,n}^2 \left|g_{m,n}\right|^2 P_{m,n}}{\beta_{m,n} - \sigma^2}. \tag{10}$$

According to Eq. (9), let the right-hand side of Eq. (10) be equal to $\varepsilon_{out} / 2$; then, we can get $\Gamma_{m,n}^2 \left|g_{m,n}\right|^2 P_{m,n} / \left(\beta_{m,n} - \sigma^2\right) = \varepsilon_{out}^C / 2$. Because $\left|g_{m,n}\right|^2 : \mathcal{N}_C\left(\hat{g}_{m,n}, \sigma_e^2\right)$ is a non-central chi-squared distributed random variable with two degrees of freedom, the left-hand side of $\Pr\left[\mu_{m,n} \leq \alpha_{m,n} \mid \hat{g}_{m,n}\right] = \varepsilon_{out}^C / 2$ is

$$\Pr\left[\mu_{m,n} \leq \alpha_{m,n} \mid \hat{g}_{m,n}\right] = \Pr\left[\Gamma_{m,n}^2 \left|g_{m,n}\right|^2 P_{m,n} \leq \alpha_{m,n} \mid \hat{g}_{m,n}\right]$$

$$= \Pr\left[\left|g_{m,n}\right|^2 \leq \frac{\alpha_{m,n}}{\Gamma_{m,n}^2 P_{m,n}}\right]$$

$$= F_{\left|g_{m,n}\right|^2}\left(\frac{\alpha_{m,n}}{\Gamma_{m,n}^2 P_{m,n}}\right)$$

$$= 1 - Q_{m,n}\left(\sqrt{\frac{2\left|\hat{g}_{m,n}\right|^2}{\sigma_e^2}}, \sqrt{\frac{2}{\sigma_e^2} \frac{\alpha_{m,n}}{\Gamma_{m,n}^2 P_{m,n}}}\right), \tag{11}$$

where $Q_1(a,b) = \exp\left(-\left(a^2 + b^2\right)/2\right) \sum_{k=0}^{\infty} (a/b)^k I_k(ab)$ is the first-order Marcum Q-function and $I_k(ab)$ is the k-th order modified Bessel function of the first type. Let Eq. (11) be equal to $\varepsilon_{out} / 2$; then, we can get

$$\alpha_{m,n} = F_{\left|g_{m,n}\right|^2}^{-1}\left(\frac{\varepsilon_{out}^C}{2}\right) \cdot \Gamma_{m,n}^2 P_{m,n}, \tag{12}$$

From $\left|g_{m,n}\right|^2 = \left|\hat{g}_{m,n}\right|^2 + \sigma_e^2$, $\beta_{m,n} = \alpha_{m,n} / \left(2^{r_{m,n}} - 1\right)$, we can get

$$\frac{\Gamma_{m,n}^2 \left|g_{m,n}\right|^2 P_{m,n}}{\alpha_{m,n} / \left(2^{r_{m,n}} - 1\right) - \sigma^2} = \frac{\Gamma_{m,n}^2 \left(\left|\hat{g}_{m,n}\right|^2 + \sigma_e^2\right) P_{m,n}}{\dfrac{F_{\left|g_{m,n}\right|^2}^{-1}\left(\varepsilon_{out}^C / 2\right) \cdot \Gamma_{m,n}^2 P_{m,n}}{2^{r_{m,n}} - 1} - \sigma^2} = \frac{\varepsilon_{out}^C}{2}. \tag{13}$$

By following the same process, we can obtain the data rate of the devices in the network as follows: $\tilde{r}_{m,n,l} = \log_2\left(1 + \tilde{\phi}_{m,n,l}\right)$ where $\tilde{r}_{m,n,l}$ is similar to $r_{m,n,l}$ and $\tilde{\phi}_{m,n,l}$ denotes the received SINR from the HPN to CU n.

$$\tilde{\phi}_{0,n,l} = \frac{P_{0,n,l}\Gamma_{0,n}^2 \varepsilon_{out}^C F_{0,n,l}^{-1}(\frac{\varepsilon_{out}^C}{2})}{\left(\varepsilon_{out}^C\right)^2 \sigma_z^2 + \sum_{\substack{k' \in K, k' \neq k \\ i \in N, i \neq n}} 2(|\hat{g}_{i,n}|^2 + \sigma_e^2)\Gamma_{i,n}^2 P_{i,k,l}}, \tag{14}$$

where σ_z^2 is the variance of a random zero-mean complex additive white Gaussian noise (AWGN) variable and $\tilde{\phi}_{m,n,l}$ is the received SINR from LPN m to CU n

$$\tilde{\phi}_{m,n,l} = \frac{P_{m,n,l}\Gamma_{m,n}^2 \varepsilon_{out}^C F_{m,n,l}^{-1}(\frac{\varepsilon_{out}^C}{2})}{\left(\varepsilon_{out}^C\right)^2 \sigma_z^2 + \sum_{\substack{k' \in K, k' \neq k \\ i \in N, i \neq n}} 2\left[(|\hat{g}_{0,n,l}|^2 + \sigma_e^2)\Gamma_{0,n,l}^2 P_{0,i,l} + (|\hat{g}_{i,n,l}|^2 + \sigma_e^2)\Gamma_{i,n,l}^2 P_{i,k',l}\right]}. \tag{15}$$

$\tilde{\phi}_{n,k,l}$ is the received SINR of the DUs:

$$\begin{aligned}
\tilde{\phi}_{n,k,l} = P_{n,k,l}\Gamma_{n,k}\varepsilon_{out}^C F_{n,k,l}^{-1}(\frac{\varepsilon_{out}^C}{2}) / \{&\left(\varepsilon_{out}^C\right)^2 \sigma_z^2 \\
+ \sum_{\substack{k' \in K, k' \neq k \\ i \in N, i \neq n}} &2\left[(|\hat{g}_{m,k,l}|^2 + \sigma_e^2)\Gamma_{m,k,l}^2 P_{m,n,l}\right. \\
+ (|\hat{g}_{0,k,l}|^2 &+ \sigma_e^2)\Gamma_{0,k,l}^2 P_{0,i,l} \\
+ (|\hat{g}_{i,k,l}|^2 &+ \left.\sigma_e^2)\Gamma_{i,k,l}^2 P_{i,k',l}\right]\}.
\end{aligned} \tag{16}$$

The transformed average sum rate of the network can be written as

$$\eta_{SE}(A, x, y, p) = \sum_{n=1}^{N}\left(1 - \varepsilon_{out}^C\right)\tilde{r}_n + \sum_{k=1}^{K}\left(1 - \varepsilon_{out}^D\right)\tilde{r}_k. \tag{17}$$

For simplicity, we define $\varepsilon_{out} = \varepsilon_{out}^C = \varepsilon_{out}^D$.

The optimization problem can be reformulated as

$$\begin{aligned}
&\min_{A, x, y, p, \varphi} \varphi, \\
&s.t. \quad C1 - C13, \\
&C14 : \omega\left(\frac{\eta_{SE}^{max} - \eta_{SE}(A, x, y, p)}{\eta_{SE}^{max} - \eta_{SE}^{min}}\right) \leq \varphi, \\
&C15 : (1 - \omega)\left(\frac{P_{tot}}{P_{max}}\right) \leq \varphi.
\end{aligned} \tag{18}$$

As the non-probabilistic optimization problem involves co-channel interference, the transformed problem is still non-convex. In the next section, we propose several resource allocation algorithms for deriving optimal solutions to the problem given by Eq. (18).

3. Proposed Algorithms for Resource Management

In this section, we divide the optimal problem into two parts. In the first part, we focus on the user scheduling and spectrum allocation subproblem. In the second part, we solve the power control subproblem for the CUs and DGs, respectively. Finally, we summarize the joint EE-SE tradeoff algorithm and present its complexity analysis.

3.1 User Association and Spectrum Allocation Algorithms

The RRH selection is based on the locations of the CUs as shown in **Fig. 1**, where the users selecting LPNs and the users selecting the HPN are divided by dashed lines.

After providing network access to the CUs, we consider matching the DUs and CUs. As we assume that the number of DUs is larger than the number of CUs in the system, some of the DUs may form DGs. In this section, we also assume that the power allocated to each user is a fixed value, which leads to a fixed F_2. Thus, we aim to optimize F_1 and maximize η_{SE} in this algorithm as follows:

$$\max_{y} \ \eta_{SE}(A^*, x^*, y, p^*),$$
$$s.t. \ \ C6 - C8. \tag{19}$$

In contrast to [12], which assumes a one-to-one D2D link, in this network, the number of DUs, K, is larger than the number of CUs, N. Thus, there are K DGs in it. Let W be the weight matrix of the transmission rates when the DUs reuse the SCs of different CUs. When $N \geq K$,

$$W_1 = \begin{pmatrix} \Omega_{1,1} & ... & \Omega_{1,N} \\ ... & ... & ... \\ \Omega_{N,1} & ... & \Omega_{N,N} \\ \Omega_{N+1,1} & ... & \Omega_{N+1,N} \\ ... & ... & ... \\ \Omega_{N+K,1} & ... & \Omega_{N+K,N} \end{pmatrix}, \tag{20}$$

where $\Omega_{n,k} = P\!\!\!/_{m,n} + P\!\!\!/_{n,k}$. We employ W_1 to get N one-to-one D2D pairs using the well-known Hungarian algorithm [13]. Then, we reuse the Hungarian algorithm to allocate spectrum resources to K DUs that do not match any CU, to get K DGs. As $N \geq K$,

$$W_2 = \begin{pmatrix} \Omega_{1,1} & ... & \Omega_{1,K} \\ ... & ... & ... \\ \Omega_{N,1} & ... & \Omega_{N,K} \end{pmatrix}. \tag{21}$$

3.2 Power Allocation Algorithm

Having already addressed the user association and spectrum allocation problem, we now need to determine the unknown variable p. As the optimization problem is still a non-concave problem, it is challenging to find the global optimum; hence, we first need to further decompose it to reduce its computational complexity and then use the corresponding Lagrange function to solve it. Eq. (18) can be rewritten as

$$L(p, \varphi, \mu_1, \mu_2) = \varphi + \mu_1(\omega F_1(p) - \varphi) + \mu_2((1-\omega)F_2(p) - \varphi). \tag{22}$$

where μ_1 and μ_2 are the Lagrangian multipliers. By rearranging (22) to get

$$L\left(p,\varphi,\mu_1,\mu_2\right) = -\frac{\mu_1\omega\eta_{SE}(p)}{\eta_{SE}^{\max}-\eta_{SE}^{\min}} + \frac{\mu_2(1-\omega)P_{tot}(p)}{P_{\max}}$$
$$+ (1-\mu_1-\mu_2)\varphi + \frac{\mu_1\omega U_{\max}}{U_{\max}-U_{\min}}, \tag{23}$$

Eq. (18) can be decomposed into two suboptimal problems, i.e., a power allocation subproblem for the CUs and a power allocation subproblem for the DGs:

$$\min_p L_1\left(p,\mu_1,\mu_2\right)$$
$$s.t.\ C9-C13 \tag{24}$$

where

$$L_1\left(p,\mu_1,\mu_2\right) = -\frac{\mu_1\omega\eta_{SE}(p)}{\eta_{SE}^{\max}-\eta_{SE}^{\min}} + \frac{\mu_2(1-\omega)P_{tot}(p)}{P_{\max}}. \tag{25}$$

along with an adaptive φ selection subproblem

$$\min_\varphi L_2\left(\varphi,\mu_1,\mu_2\right) \tag{26}$$

where $L_2\left(\varphi,\mu_1,\mu_2\right) = (1-\mu_1-\mu_2)\varphi$. Further, Eq. (24) can be rewritten as

$$\max_p -L_1\left(p,\mu_1,\mu_2\right)$$
$$s.t.\ C9-C13 \tag{27}$$

where

$$-L_1\left(p,\mu_1,\mu_2\right) = \frac{\mu_1\omega\eta_{SE}(p)}{\eta_{SE}^{\max}-\eta_{SE}^{\min}} - \frac{\mu_2(1-\omega)P_{tot}(p)}{P_{\max}}, \tag{28}$$

To find the optimal power allocation solution of Eq. (27), we first jointly solve it for the CUs and DGs. Then, we allocate power to each DG by PD–NOMA if it has two receivers, i.e.,

$$\max_p f_1\left(p,\mu_1,\mu_2\right) = \frac{\mu_1\omega\eta_{SE}(p)}{\eta_{SE}^{\max}-\eta_{SE}^{\min}} - \frac{\mu_2(1-\omega)P_{tot}(p)}{P_{\max}},$$
$$s.t.\ C9,\ C10. \tag{29}$$

As the total power allocated to the DGs is obtained, we only need to solve the power allocation optimization problem for maximizing the sum data rate of each DG, which is given by

$$\max_p f_2\left(p,\mu_1,\mu_2\right) = \max_p R^D,$$
$$s.t.\ C11-C13. \tag{30}$$

First, we focus on Eq. (29). It is difficult to obtain the global optimal solution of a non-convex problem. Hence, we use sequential convex programming (SCP) to derive local optimal solutions by solving a sequence of problems to reduce the computational complexity. The lower bound of the logarithmic function is given by [14], $\log_2(1+\widetilde{\phi}) \geq a\log_2(\widetilde{\phi}+b$, where $a = \phi'/\left(1+\phi'\right)$ and $\log_2(1+\phi')-a\log_2(\phi')$. When $\widetilde{\phi}=\phi'$, the equality holds. Thus, the lower bound of $f_1\left(p,\mu_1,\mu_2\right)$ is $\widetilde{f_1}(p,\mu_1,\mu_2)$. By the SINR of each user we get:

$$\overset{o}{\phi}_{0,n,l} = \frac{\Gamma_{0,n}^2 \varepsilon_{out} F_{0,n,l}^{-1}(\frac{\varepsilon_{out}}{2})}{\varepsilon_{out}^2 \sigma_z^2 + \sum_{\substack{k' \in K, k' \neq k \\ i \in N, i \neq n}} 2(|\hat{g}_{i,n,l}|^2 + \sigma_e^2)\Gamma_{i,n}^2 P_{i,k,l}}, \tag{31}$$

$$\overset{o}{\phi}_{0,n,l} = \frac{\Gamma_{m,n}^2 \varepsilon_{out} F_{m,n,l}^{-1}(\frac{\varepsilon_{out}}{2})}{\varepsilon_{out}^2 \sigma_z^2 + \sum_{\substack{k' \in K, k' \neq k \\ i \in N, i \neq n}} 2\left[(|\hat{g}_{0,n,l}|^2 + \sigma_e^2)\Gamma_{0,n,l}^2 P_{0,i,l} + (|\hat{g}_{i,n,l}|^2 + \sigma_e^2)\Gamma_{i,n,l}^2 P_{i,k',l}\right]}, \tag{32}$$

$$\overset{o}{\phi}_{n,k,l} = \frac{\Gamma_{n,k} \varepsilon_{out} F_{n,k,l}^{-1}(\frac{\varepsilon_{out}}{2})}{\varepsilon_{out}^2 \sigma_z^2 + \sum_{\substack{k' \in K, k' \neq k \\ i \in N, i \neq n}} 2\left[(|\hat{g}_{m,k,l}|^2 + \sigma_e^2)\Gamma_{m,k,l}^2 P_{m,n,l} + (|\hat{g}_{0,k,l}|^2 + \sigma_e^2)\Gamma_{0,k,l}^2 P_{0,i,l} + (|\hat{g}_{i,k,l}|^2 + \sigma_e^2)\Gamma_{i,k,l}^2 P_{i,k',l}\right]}. \tag{33}$$

Considering the constrains as KKT conditions, we can obtain the Lagrangian function as

$$L_n^C = \frac{\sum_n (1 - \varepsilon_{out}) a_n^C \mu_1 \omega \log_2 \overset{o}{\phi}_n + (1 - \varepsilon_{out}) \mu_1 \omega b_n^C}{\eta_{SE}^{\max} - \eta_{SE}^{\min}} - \frac{\mu_2 (1 - \omega) P_{tot}}{P_{\max}}$$

$$+ \lambda_{1,n}\left(\overset{o}{\phi}_n - \phi_{thr}^C\right) + \lambda_{2,n}\left(P_{\max}^C - \sum_n P_n\right) \tag{34}$$

for CU n and

$$L_k^D = \frac{\sum_k (1 - \varepsilon_{out}) a_k^D \mu_1 \omega \log_2 \overset{o}{\phi}_k + (1 - \varepsilon_{out}) \mu_1 \omega b_k^D}{\eta_{SE}^{\max} - \eta_{SE}^{\min}} - \frac{\mu_2 (1 - \omega) P_{tot}}{P_{\max}}$$

$$+ \lambda_{1,k}\left(\overset{o}{\phi}_k - \phi_{thr}^D\right) + \lambda_{2,k}\left(P_{\max}^D - \sum_k P_k\right) \tag{35}$$

for DG k, where a_n^C, b_n^C, a_k^D, and b_k^D are the indexes for the CUs and DGs from SCP. Further, $P_{\max}^C$ and $P_{\max}^D$ are the predefined maximum power that a user can receive for every CU and DG, respectively, while λ_1 and λ_2 are the Lagrangian multipliers.

To solve Eqs. (34) and (35), we decompose the problem into an inner layer and an outer layer. In the inner layer of the power allocation algorithm, we first fix the Lagrange multipliers, and the problem is transformed into a standard convex optimization problem with KKT conditions. By deduction and simplification, the optimal power allocation policy of CU n and DG k in the θ-th iteration can be derived as

$$P_n(\theta + 1) = \frac{(1 - \varepsilon_{out}) a_n^C(\theta) \mu_1(\theta) \omega}{\ln 2\left(\eta_{SE}^{\max} - \eta_{SE}^{\min}\right)\left[\lambda_{2,n}(\theta) + \frac{\lambda_{2,n}(\theta)(1 - \omega)}{P_{\max}^C} - \lambda_{1,n}(\theta)\overset{o}{\phi}_n(\theta)\right]}, \tag{36}$$

$$P_k(\theta + 1) = \frac{(1 - \varepsilon_{out}) a_k^D(\theta) \mu_1(\theta) \omega}{\ln 2\left(\eta_{SE}^{\max} - \eta_{SE}^{\min}\right)\left[\lambda_{2,k}(\theta) + \frac{\lambda_{2,k}(\theta)(1 - \omega)}{P_{\max}^D} - \lambda_{1,k}(\theta)\overset{o}{\phi}_k(\theta)\right]}, \tag{37}$$

As for the outer layer, using the power allocation scheme, we can obtain the variables on the basis of gradient descent, and the Lagrange multipliers can be updated by

$$\lambda_{1,n}(\theta+1) = \left[\lambda_{1,n}(\theta) + v(\theta)\left(\phi_n - \phi_{thr}^C\right)\right]^+, \tag{38}$$

$$\lambda_{2,n}(\theta+1) = \left[\lambda_{2,n}(\theta) + v(\theta)\left(P_{max}^C - \sum_k P_{n,}\right)\right]^+, \tag{39}$$

$$\lambda_{1,k}(\theta+1) = \left[\lambda_1(\theta) + v(\theta)\left(\phi_k - \phi_{thr}^D\right)\right]^+, \tag{40}$$

$$\lambda_{2,k}(\theta+1) = \left[\lambda_{2,k}(\theta) + v(\theta)\left(P_{max}^D - \sum_k P_{k,}\right)\right]^+, \tag{41}$$

along with

$$\mu_1(\theta+1) = \left[\mu_1(\theta) + v(\theta)\left(\omega\frac{\eta_{SE}^{max} - \eta_{SE}(\theta+1)}{\eta_{SE}^{max} - \eta_{SE}^{min}} - \varphi(\theta+1)\right)\right]^+, \tag{42}$$

and

$$\mu_2(\theta+1) = \left[\mu_2(\theta) + v(\theta)\left((1-\omega)\frac{P_{tot}(\theta+1)}{P_{max}} - \varphi(\theta+1)\right)\right]^+, \tag{43}$$

where $v(\theta) = 0.1/\sqrt{\theta}$. The optimal solution of Eq. (63) can be obtained by

$$\varphi(\theta+1) = \max\left\{\omega F_1(p(\theta+1)), (1-\omega)F_2(p(\theta+1))\right\}. \tag{44}$$

The proposed iterative power allocation algorithm based on SCP is outlined in **Algorithm 1**.

Algorithm 1 SCP-Based Iterative Power Allocation Algorithm

1: – *Initialization Phase:*

2: Set the iteration index $\theta = 0$.

3: Initialize $p_n(0), p_k(0), \lambda_1(0), \lambda_2(0), \mu_1(0)$ and $\mu_2(0)$.

4: Calculate $a_n^C(0), b_n^C(0), \phi_n(0)$ and $a_k^D(0), b_k^D(0), \phi_k(0)$.

5: Set a convergence threshold Δ and a maximum iteration number N_{max}.

6: – *Update Phase:*

7: **while** $\left|p_n(\theta) - p_n(\theta-1)\right| \geq \Delta$ **or** $\left|p_k(\theta) - p_k(\theta-1)\right| \geq \Delta$ **or** $\theta < N_{max}$ **do**

8: $\theta = \theta + 1$;

9: Update $p_n(\theta), p_k(\theta), \lambda_1(\theta), \lambda_2(\theta), \mu_1(\theta), \mu_2(\theta)$;

10: Update $a_n^C(\theta), b_n^C(\theta), \phi_n(\theta)$ and $a_k^D(\theta), b_k^D(\theta), \phi_k(\theta)$;

11: Get the final power allocation coefficients and solve the optimization problem in Eq. (19);

12: **end while**

13: Result: $p_n^* = p_n(\theta)$ and $p_k^* = p_k(\theta)$.

Similarly, we use SCP to solve the power allocation problem for K DGs in the network, which is also a non-convex problem.

The objective function in Eq. (30) can be rewritten as

$$\max_{s_{k1},s_{k2}} \tilde{R}_{sum} = \left(1-\varepsilon_{out}\right)\sum_{k=1}^{K}\log_2\left(1+\tilde{\phi}_k\right), \tag{45}$$

where s_{k1} and s_{k2} are the power allocation coefficients for the two receivers in DG k.

As proved in [14], we have

$$\log_2\left(1+\tilde{\phi}_k\right) \geq c_k\log_2\tilde{\phi}_k + d_k, \tag{46}$$

where c_k is defined as

$$c_k = \frac{\phi_k^{'}}{1+\phi_k^{'}}, \tag{47}$$

and d_k is defined as

$$d_k = \log_2\left(1+\phi_k^{'}\right) - \frac{\phi_k^{'}}{1+\phi_k^{'}}\log_2\phi_k^{'}. \tag{48}$$

The equality is satisfied when $\tilde{\phi}_k = \phi_k^{'}$.

From Eqs. (46), the lower bound of the target function in Eq. (45) can be expressed as

$$\tilde{R}_{sum} \geq \left(1-\varepsilon_{out}\right)\sum_{k=1}^{K}\Theta_k, \tag{49}$$

where Θ_k can be defined as $\Theta_k = c_k\log_2\tilde{\Phi}_k + d_k$, $\forall k$.

We can obtain a new optimization problem from Eqs. (30) and (46) as follows:

$$\max_p f_2\left(p,\mu_1,\mu_2\right) = \max_p R^D = \max_{s_{k1},s_{k2}}\sum_{k=1}^{K}\Theta_k,$$
$$s.t. \quad C11-C13. \tag{50}$$

It can be proved that the newly formulated problem is a concave problem [3]. For each Θ_k, we can get $\tilde{\phi}_{n,k1,l}$, which denotes the SINR of DU $k1$ in DG k according to Eq. (22).

$$\tilde{\phi}_{n,k1,l} = P_{n,k1,l}\Gamma_{n,k1}\varepsilon_{out}F_{n,k1,l}^{-1}(\frac{\varepsilon_{out}}{2}) / \{\varepsilon_{out}^2\sigma_z^2$$
$$+ \sum_{\substack{k^{'}\in K, k^{'}\neq k \\ i\in N, i\neq n}} 2\left[(\left|\hat{g}_{m,k1,l}\right|^2 + \sigma_e^2)\Gamma_{m,k1,l}^2 P_{m,n,l}\right.$$
$$+ (\left|\hat{g}_{0,k1,l}\right|^2 + \sigma_e^2)\Gamma_{0,k1,l}^2 P_{0,i,l}$$
$$+ (\left|\hat{g}_{i,k1,l}\right|^2 + \sigma_e^2)\Gamma_{i,k1,l}^2 P_{i,k^{'},l}\right]\}. \tag{51}$$

Similarly, we can get $\tilde{\phi}_{n,k2,l}$, which denotes the SINR of DU $k2$ in DG k.

$$\phi_{n,k2,l}^{\%} = P_{n,k2,l}\Gamma_{n,k2}\varepsilon_{\text{out}}F_{n,k2,l}^{-1}(\frac{\varepsilon_{\text{out}}}{2})/\{\varepsilon_{\text{out}}^2\sigma_z^2$$

$$+\sum_{\substack{k'\in K,k'\neq k\\i\in N,i\neq n}}2[(\left|\hat{g}_{m,k2,l}\right|^2+\sigma_e^2)\Gamma_{m,k2,l}^2P_{m,n,l}+(\left|\hat{g}_{0,k2,l}\right|^2+\sigma_e^2)\Gamma_{0,k2,l}^2P_{0,i,l}$$

$$+(\left|\hat{g}_{i,k2,l}\right|^2+\sigma_e^2)\Gamma_{i,k2,l}^2P_{i,k',l}+(\left|\hat{g}_{k1}\right|^2+\sigma_e^2)\Gamma_{k1,k2}^2P_{k1}]\}. \tag{52}$$

Then, we derive the SINR of each DG and get the following results:

$$\phi_{n,k1,l}^{\%} = \frac{\Gamma_{n,k1}\varepsilon_{\text{out}}F_{n,k1,l}^{-1}(\frac{\varepsilon_{out}}{2})}{\varepsilon_{out}^2\sigma_z^2+\sum_{\substack{k'\in K,k'\neq k\\i\in N,i\neq n}}2\left[(\left|\hat{g}_{m,k1,l}\right|^2+\sigma_e^2)\Gamma_{m,k1,l}^2P_{m,n,l}+(\left|\hat{g}_{0,k1,l}\right|^2+\sigma_e^2)\Gamma_{0,k1,l}^2P_{0,i,l}+(\left|\hat{g}_{i,k1,l}\right|^2+\sigma_e^2)\Gamma_{i,k1,l}^2P_{i,k',l}\right]}. \tag{53}$$

$$\phi_{n,k2,l}^{\%} = \Gamma_{n,k2}\varepsilon_{\text{out}}F_{n,k2,l}^{-1}(\frac{\varepsilon_{\text{out}}}{2})/\{\varepsilon_{\text{out}}^2\sigma_z^2$$

$$+\sum_{\substack{k'\in K,k'\neq k\\i\in N,i\neq n}}2[(\left|\hat{g}_{m,k2,l}\right|^2+\sigma_e^2)\Gamma_{m,k2,l}^2P_{m,n,l}+(\left|\hat{g}_{0,k2,l}\right|^2+\sigma_e^2)\Gamma_{0,k2,l}^2P_{0,i,l}$$

$$+(\left|\hat{g}_{i,k2,l}\right|^2+\sigma_e^2)\Gamma_{i,k2,l}^2P_{i,k',l}+(\left|\hat{g}_{k1}\right|^2+\sigma_e^2)\Gamma_{k1,k2}^2P_{k1}]\}. \tag{54}$$

Considering the constrains as KKT conditions, we obtain the Lagrangian function as

$$L_k^D = \sum_k (1-\varepsilon_{out})a_k^D\log_2\phi_k^{\%}+(1-\varepsilon_{out})b_k^D+\lambda_1\left(\phi_k^{\%}-\phi_{thr}^D\right)+\lambda_2\left(P_{\max}^D-\sum_k P_k\right) \tag{55}$$

The power of the two users in DG k in the θ-th iteration can be derived as

$$P_{k1}(\theta+1)=(1-\varepsilon_{out})a_{k1}^D(\theta)/\left(\ln 2\left(\lambda_2(\theta)+\frac{\lambda_2(\theta)(1-\omega)}{P_{\max}^D}-\lambda_{1,k1}(\theta)\phi_{k1}^{\%}(\theta)\right)\right), \tag{56}$$

$$P_{k2}(\theta+1)=(1-\varepsilon_{out})a_{k2}^D(\theta)/\left(\ln 2\left(\lambda_2(\theta)+\frac{\lambda_2(\theta)(1-\omega)}{P_{\max}^D}-\lambda_{1,k2}(\theta)\phi_{k2}^{\%}(\theta)\right)\right). \tag{57}$$

As for the outer layer, using the power allocation scheme, we can obtain the variables on the basis of gradient descent, and the Lagrange multipliers can be updated by

$$\lambda_{1,k1}(\theta+1)=\left[\lambda_{1,k1}(\theta)+\nu(\theta)\left(\phi_{k1}^{\%}(\theta+1)-\phi_{thr}^D\right)\right]^+, \tag{58}$$

$$\lambda_{1,k2}(\theta+1)=\left[\lambda_{1,k2}(\theta)+\nu(\theta)\left(\phi_{k2}^{\%}(\theta+1)-\phi_{thr}^D\right)\right]^+, \tag{59}$$

and

$$\lambda_2(\theta+1)=\left[\lambda_2(\theta)+\nu(\theta)\left(P_{\max}^D-\sum_k P_k\right)\right]^+. \tag{60}$$

The proposed iterative power allocation algorithm for DGs, which is based on SCP, is outlined in **Algorithm 2**.

Algorithm 2 SCP-Based Iterative Power Allocation Algorithm for DGs

---

1: – *Initialization Phase:*

2: Set the iteration index $\theta = 0$.

3: Initialize $p_{k1}(0)$, $p_{k2}(0)$, $\lambda_{1,k1}(0)$, $\lambda_{1,k2}(0)$ and $\lambda_2(0)$.

4: Calculate $a_{k1}^D(0)$, $b_{k1}^D(0)$, $\mathcal{G}_{k1}^0(0)$ and $a_{k2}^D(0)$, $b_{k2}^D(0)$, $\mathcal{G}_{k2}^0(0)$.

5: Set a convergence threshold Δ and a maximum iteration number $N_{\max}$.

6: – *Update Phase:*

7: **while** $\left|p_{k1}(\theta) - p_{k1}(\theta-1)\right| \geq \Delta$ **or** $\left|p_{k2}(\theta) - p_{k2}(\theta-1)\right| \geq \Delta$ **or** $\theta < N_{\max}$ **do**

8: $\theta = \theta + 1$;

9: Update $p_{k1}(\theta)$, $p_{k2}(\theta)$, $\lambda_{1,k1}(\theta)$, $\lambda_{1,k2}(\theta)$ and $\lambda_2(\theta)$;

10: Update $a_{k1}^D(\theta)$, $b_{k1}^D(\theta)$, $\mathcal{G}_{k1}^0(\theta)$ and $a_{k2}^D(\theta)$, $b_{k2}^D(\theta)$, $\mathcal{G}_{k2}^0(\theta)$;

11: Get the final power allocation coefficients and solve the optimization problem in Eq. (30);

12: **end while**

13: Result: $p_{k1}^* = p_{k1}(\theta)$ and $p_{k2}^* = p_{k2}(\theta)$ along with s_{k1}, s_{k2}.

3.3 Complexity Analysis

Through the spectrum allocation problem , we compute the weights in each bipartite graph NK times. Suppose that we have I_1 and I_2 iterations for Algorithm 1 and Algorithm 2, respectively. In each iteration, the algorithms perform $4(N + K + 1)$ and $10K + 1$ computations for updating each variable, respectively. Let the maximum time for computing any variable be $t_{\max}$. Therefore, in summary, the overall computational complexity of the proposed algorithm is $o(t_{\max})(NK + 4(N + K + 1) + 10K + 1)$.

4. Simulation Results

In this section, we evaluate the performance of the proposed algorithm for the H-CRAN system with imperfect CSI. The specific parameter settings are summarized in **Table 2** [15].

Table 2. Simulation parameters

System Parameters	Default Value
Bandwidth	10 MHz
Noise power	-114 dBm
Cell radius	220 m
D2D radius	80 m
$P_{0,\max}$	45 dBm
$P_{m,\max}$	40 dBm
$P^d_{\max}$	25 dBm
L	12
M	7
K	30

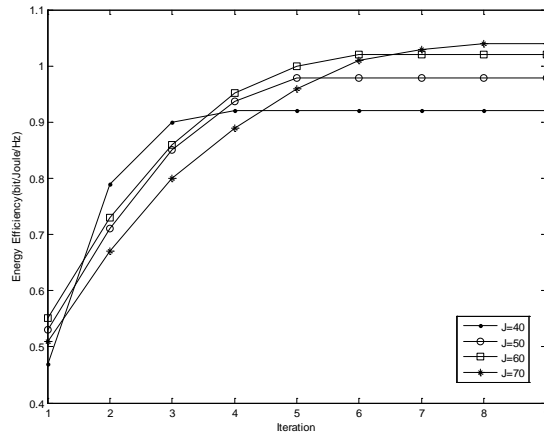| N	30
$\eta_{SE,max}$	20 Gb/s

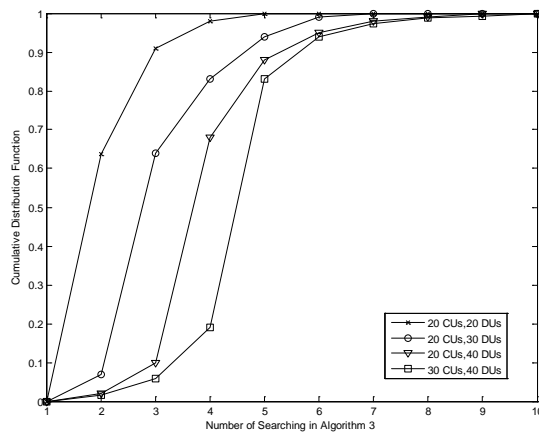**Fig. 2.** EE vs. number of iterations of Algorithm 1

Fig. 3. CDF of the number of iterations of Algorithm 2

4.1 Convergence Performance

The power allocation coefficient of each user is plotted against the number of iterations of Algorithm 1 in **Fig. 2**. The results in **Fig. 2** are averaged over 1000 independent adaptation processes with different distributions of users in one cell. It can be observed that with $\Delta = 0.01$, the proposed algorithm achieves fast converge within 10 iterations, on average.

Fig. 3 shows the cumulative distribution function (CDF) of the number of iterations of Algorithm 2 for power allocation of the DGs with imperfect CSI. As can be seen, the network EE becomes stable within around 10 iterations and the changing speed gradually decreases after 5 iterations, on average. Thus, from **Fig. 2** and **Fig. 3**, we can observe the convergence of the power allocation algorithms with different parameters.

4.2 Performance Comparison of EE-SE Tradeoff in Different Communication Networks

Fig. 4 shows the performance comparison of the EE-SE tradeoff in different communication networks, where the weighting parameter ω is adjusted from 0.1 to 1. The EEs of the proposed algorithms are plotted against the system SE in four scenarios: NOMA-enhanced H-CRAN, C-RAN without D2D communications in NOMA, H-CRAN in OMA without DGs, and C-RAN in OMA without D2D. From **Fig. 9**, the EE of the network first increases with the system SE and then decreases after reaching its highest value, which lies between 1000 bit/s/Hz and 1050 bit/s/Hz. Because of the optimal network structure, the interference between different RRHs can be managed effectively, which causes the EE of the entire system to increase. In addition, PD–NOMA based on the D2D scheme achieves a higher data rate than the conventional OMA and the scheme without D2D at the same EE.
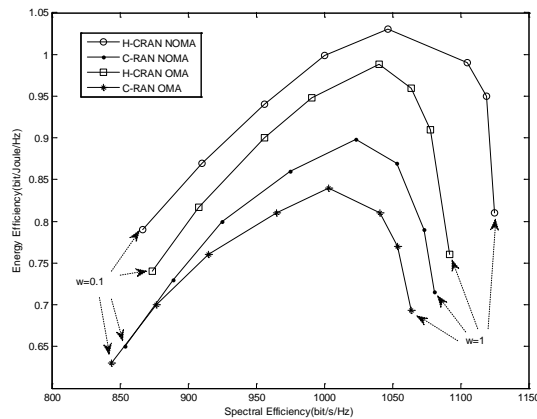

Fig. 4. EE-SE tradeoff in different types of communication networks

4.3 Performance Comparison of EE-SE Tradeoff with Different RRH Distributions

Fig. 5 shows the EE-SE tradeoff performance, where the weighting parameter ω is adjusted from 0.1 to 1 when the LPNs have different locations as shown in **Fig. 10** with 40 users in the network. Because of the optimal network structure in the proposed H-CRAN model, the interference between different RRHs is weakened, which causes the EE to increase compared to other distributions. In addition, it can be seen that among the four scenarios, PD–NOMA with perfect CSI achieves a remarkable performance gain compared to that with imperfect CSI. When the transmitted SE is 1050 bit/s/Hz, the proposed algorithm with perfect CSI achieves 33.1% and 23.5% EE gain over the algorithm with imperfect CSI under two different RRH distributions, respectively.
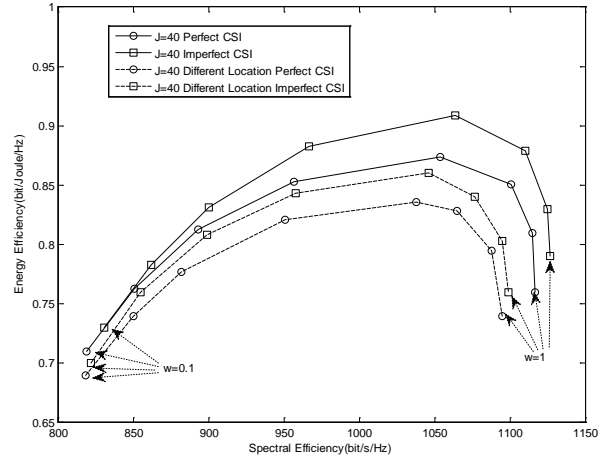
Fig. 5. EE-SE tradeoff with different LPN distributions

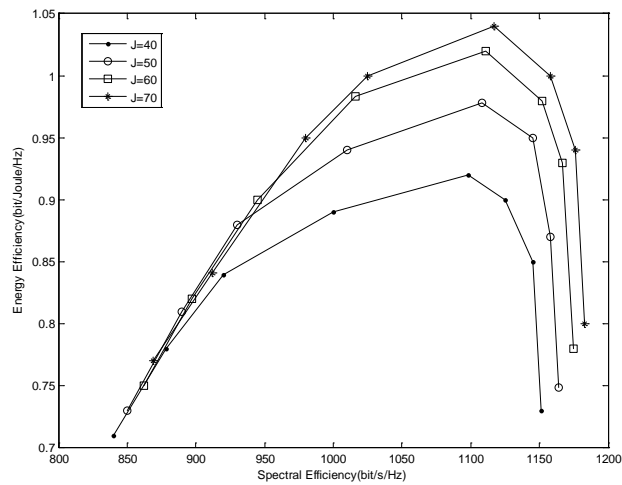## 4.4 Performance Comparison of EE-SE Tradeoff with Different Numbers of Users

Fig. 6. EE-SE tradeoff with different numbers of users

Fig. 6 plots the EE of the proposed algorithm against the SE for different numbers of users (40 to 70 devices). From **Fig. 8**, the EE decreases as the number of users in the same H-CRAN increases. Because of the increasing number of users, the probability of the DGs receiving greater interference from the CUs and other DGs increases, which leads to a monotonous decrease in the data rates with increasing N and K. Further, according to Eqs. (59), (60), and (61), there is greater interference in the SINR of each user.

4.5 Performance Comparison of EE-SE Tradeoff with Different Energy Consumptions

Fig. 7 shows the performance comparison of the EE-SE tradeoff with different transmission powers from the RRHs. After achieving the maximum EE of the proposed algorithm with different transmission powers of H-CRAN, the system suffers a sudden decrease in EE as the SE increases, because the system resources for optimizing the network structure and interference management have been consumed. If the users aim to further improve the network SE, the energy consumption and circuit loss will become extremely high, which will result in the scenario described above. Thus, the system soon achieves its maximum SE.
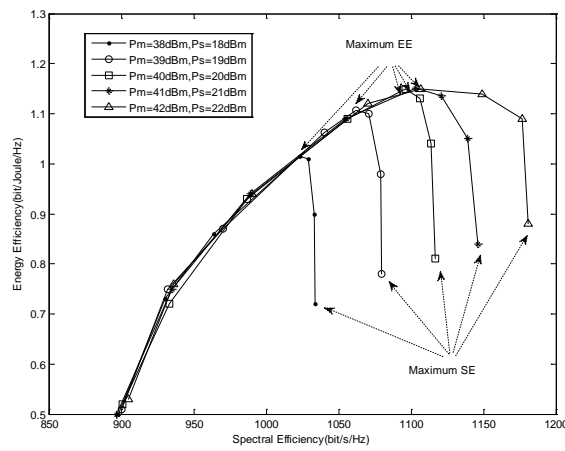

Fig. 7. EE-SE tradeoff with different energy consumptions

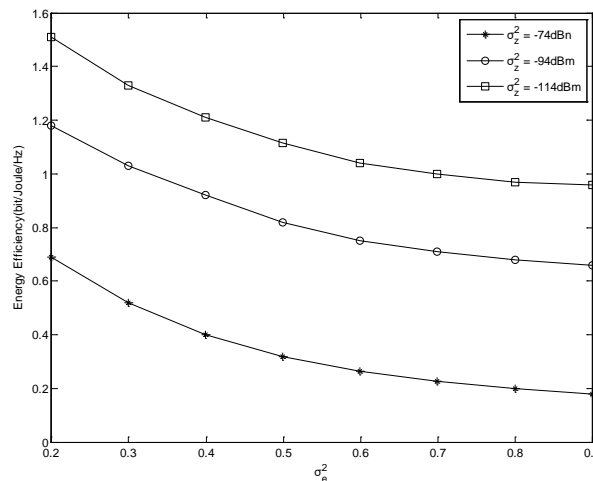## 4.6 Performance Comparison of EE-SE Tradeoff with Different Noise Power

Fig. 8. EE versus σ_e^2

Fig. 8 plots the EE of the proposed algorithm against σ_e^2 with three different σ_z^2 values. It can be seen that the EE of the network gradually decreases as σ_e^2 increases. Further, the speed changes from high to low because of the increasing σ_e^2 and σ_z^2, which indicates that the probability of the users receiving additional interference from the network increases. Consequently, the EE of the entire system decreases.

5. Conclusion

This paper proposed a general framework for studying the resource allocation problem and the EE-SE tradeoff for downlink traffic in PD-NOMA and D2D-based H-CRANs under imperfect CSI. The problem was formulated as an MOO problem that can be solved using the weighted Tchebycheff method. As the problem is a mixed-integer non-convex problem with outage probability constraints, we transformed it into a non-probability problem and separately solved the problems of user association and power coordination. First, we proposed fully optimized user association and spectrum allocation algorithms based on the Hungarian algorithm. Then, two power allocation algorithms with SCP were proposed to reduce the computational complexity. In addition, we used the Lagrangian method with KKT conditions to find the global optimum with fast convergence based on the subgradient method. Finally, the results of numerical simulations in different scenarios demonstrated that the PD-NOMA technique and H-CRAN with D2D communications enable the system to achieve a good EE-SE tradeoff.

References

[1] H. Haci, H. Zhu and J. Wang, "Performance of Non-orthogonal Multiple Access With a Novel Asynchronous Interference Cancellation Technique," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1319-1335, March 2017. Article (CrossRef Link).

[2] F. Zhou, Y. Wu, R. Q. Hu, Y. Wang and K. K. Wong, "Energy-Efficient NOMA Enabled Heterogeneous Cloud Radio Access Networks," *IEEE Network,* vol. 32, no. 2, pp. 152-160, March-April 2018. Article (CrossRef Link).

[3] J. Zhao, Y. Liu, K. K. Chai, Y. Chen and M. Elkashlan, "Joint Subchannel and Power Allocation for NOMA Enhanced D2D Communications," *IEEE Transactions on Communications,* vol. 65, no. 11, pp. 5081-5094, Nov 2017. Article (CrossRef Link).

[4] A. Mokdad, P. Azmi, N. Mokari, M. Moltafet and M. Ghaffari-Miab, "Cross-Layer Energy Efficient Resource Allocation in PD-NOMA Based H-CRANs: Implementation via GPU," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1246-1259, 1 June 2019. Article (CrossRef Link).

[5] A. Paul, A. Banerjee and S. P. Maity, "Throughput maximisation in cognitive radio networks with residual bandwidth," *IET Communications*, vol. 13, no. 10, pp. 1327-1335, 25 June 2019. Article (CrossRef Link).

[6] T. Yang, R. Zhang, X. Cheng and L. Yang, "Secure Massive MIMO Under Imperfect CSI: Performance Analysis and Channel Prediction," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1610-1623, June 2019. Article (CrossRef Link)..

[7] F. Fang, H. Zhang, J. Cheng, S. Roy and V. C. M. Leung, "Joint User Scheduling and Power Allocation Optimization for Energy-Efficient NOMA Systems With Imperfect CSI," *IEEE Journal on Selected Areas in Communications,* vol. 35, no. 12, pp. 2874-2885, Dec. 2017. Article (CrossRef Link).

[8]   Maity S.P., Paul A. "On Joint Maximization in Energy and Spectral Efficiency in Cooperative Cognitive Radio Networks," *Multi-Objective Optimization, Springer, Singapore*, pp. 141-157, 2018. Article (CrossRef Link).

[9]   Y. Hao, Q. Ni, H. Li and S. Hou, "On the Energy and Spectral Efficiency Tradeoff in Massive MIMO-Enabled HetNets With Capacity-Constrained Backhaul Links," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4720-4733, Nov. 2017. Article (CrossRef Link)..

[10] L. Kleinrock, *Queuing Systems*, Hoboken, Wiley, NJ, 1975. Article (CrossRef Link).

[11] X. Wang, F. Zheng, P. Zhu and X. You, "Energy-Efficient Resource Allocation in Coordinated Downlink Multicell OFDMA Systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1395-1408, Mar. 2016. Article (CrossRef Link).

[12] L. Wang, H. Tang, H. Wu and G. L. Stüber, "Resource Allocation for D2D Communications Underlay in Rayleigh Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1159-1170, Feb. 2017. Article (CrossRef Link).

[13] A. Gjendemsjø, D. Gesbert, G. E. Øien and S. G. Kiani, "Binary Power Control for Sum Rate Maximization over Multiple Interfering Links," *IEEE Transactions on Wireless Communications*, vol. 7, no. 8, pp. 3164-3173, August 2008. Article (CrossRef Link).

[14] J. Papandriopoulos and J. S. Evans, "Low-Complexity Distributed Algorithms for Spectrum Balancing in Multi-User DSL Networks," in *Proc. of 2006 IEEE International Conference on Communications, Istanbul*, pp. 3270-3275, 2006. Article (CrossRef Link).

[15] S. Chatterjee, S. P. Maity and T. Acharya, "Energy-Spectrum Efficiency Trade-Off in Energy Harvesting Cooperative Cognitive Radio Networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 295-303, June 2019. Article (CrossRef Link).

Jingpu Wang received the B.S. degree and the M.S. degree in Communications Engineering from Dalian Maritime University (DMU), Dalian, China, in 2013 and 2016, respectively. He is currently pursuing the Ph.D. degree in Communication and Information System with Northeastern University (NEU), Shenyang, China. His research interests include device-to-device communications, nonorthogonal multiple access and wireless resource allocation.
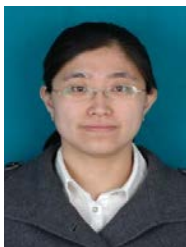
Xin Song was born in Jilin, China, in 1978. She received her PhD degree in Communication and Information System in Northeastern University in China in 2008. She is now a teacher working in Northeastern University at Qinhuangdao, China. Her research interests are in the area of robust adaptive beamforming and wireless communication.

Li Dong received the B.S. degree in Communications Engineering from Shenyang University of Technology in 2017. She is currently pursuing the M.S. degree in Communication and Information System with the Northeastern University, Qinhuangdao, China. Her research interests include heterogeneous networks, non-orthogonal multiple access technology, and wireless resource allocation.