SCIENCE CHINA
Information Sciences

April 2021, Vol. 64 140303:1–140303:14
https://doi.org/10.1007/s11432-020-2985-8

• RESEARCH PAPER •
Special Focus on Wireless Communications with Unmanned Aerial Vehicles (UAVs)

Energy-efficient design for mmWave-enabled NOMA-UAV networks

Xiaowei PANG[1], Jie TANG[2], Nan ZHAO[1*], Xiuyin ZHANG[2] & Yi QIAN[3]

[1]*School of Information and Communication Engineering, Dalian University of Technology, Dalian* 116024*, China;*
[2]*School of Electronic and Information Engineering, South China University of Technology, Guangzhou* 510641*, China;*
[3]*Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, Lincoln* 68588*, USA*

Received 16 March 2020/Revised 8 June 2020/Accepted 17 July 2020/Published online 2 March 2021

Abstract Owing to the recent advances of non-orthogonal multiple access (NOMA) and millimeter-wave (mmWave), these two technologies are combined in unmanned aerial vehicle (UAV) networks in this paper. However, energy efficiency has become a significant metric for UAVs owning to their limited energy. Thus, we aim to maximize the energy efficiency for mmWave-enabled NOMA-UAV networks by optimizing the UAV placement, hybrid precoding and power allocation. However, the optimization problem is complicated and intractable, which is decomposed into several sub-problems. First, we solve the UAV placement problem by approximating it into a convex one. Then, the hybrid precoding with user clustering is performed to better reap the multi-antenna gain. Particularly, three schemes are proposed, where the cluster head selection algorithm is adopted while considering different equivalent channels of users. Finally, the power allocation is optimized to maximize the energy efficiency, which is converted to convex and solved via an iterative algorithm. Simulation results are provided to evaluate the performance of the proposed schemes.

Keywords energy efficiency, hybrid precoding, millimeter-wave, non-orthogonal multiple access, placement optimization, power allocation, unmanned aerial vehicle

Citation Pang X W, Tang J, Zhao N, et al. Energy-efficient design for mmWave-enabled NOMA-UAV networks. Sci China Inf Sci, 2021, 64(4): 140303, https://doi.org/10.1007/s11432-020-2985-8

1 Introduction

Apart from performing combat missions [1], unmanned aerial vehicles (UAVs) also have found a wide range of applications in the civilian and commercial domains, such as cargo delivery, traffic management, aerial imaging, and so forth [2,3]. Recently, UAVs are employed in communication systems for relaying [4], collecting data [5] and wireless broadcasting [6]. Particularly, UAVs deployed as aerial base stations (BSs) are potential to provide ubiquitous coverage and satisfy users' multifarious requirements owing to their flexibility and mobility [7–9]. Compared with traditional terrestrial communication platforms which are generally fixed once deployed, UAVs can be flexibly deployed [7]. Considering the three-dimension (3D) mobility of UAVs, Sun et al. in [8] designed the 3D trajectory and resource allocation for multi-carrier solar-powered UAV communication systems. In [9], the UAV trajectory and BS precoding were jointly optimized to maximize the sum rate of users in a UAV-assisted network. In addition, there is another paradigm of UAV communications where UAVs as aerial users are served by the ground BSs, namely, cellular-connected UAV communication [10]. The research in [11] revealed that the aerial-ground interference is a severe performance bottleneck for the cellular-connected UAV communication. A promising approach for reducing interference in conventional cellular networks is coordinated multi-point transmission [12], while Mei et al. investigated the uplink aerial-ground interference management and achieved a flexible performance trade-off via jointly optimizing the UAV's association and power allocation (PA) in [13]. To maximize the security rate for a full-duplex UAV, the beamforming vetor, PA factor and artificial noise projection matrix were optimized in [14].

* Corresponding author (email: zhaonan@dlut.edu.cn)

Nevertheless, owing to the explosively increasing equipment density and frequency reuse, the resource available for UAVs is decreasing. In this context, non-orthogonal multiple access (NOMA) is recognized as a promising technology to allow UAVs to reuse the spectrum assigned to ground BSs or users [15]. Applying PA at transmitters as well as successive interference cancellation (SIC) at receivers, NOMA can not only further boost the network throughput, but also support more users to enhance the spectrum efficiency [16]. Up to now, plenty of studies have contributed to the combination of NOMA for UAV communication systems [17–20]. Specifically, a novel framework of UAV networks supported by NOMA was investigated in [17], where three cases were studied. Zhao et al. [18] presented two schemes to guarantee the secure transmission in NOMA-UAV networks by PA and beamforming optimization, respectively. Sohail et al. [19] maximized the users' sum rate via optimizing the UAV altitude and PA in a UAV-enabled NOMA system. Furthermore, Ref. [20] investigated the uplink NOMA transmission in cellular-connected UAV communication networks with beamforming design and interference management.

On the other hand, current wireless spectrum is overcrowded, which encourages the migration to new spectrum resource. Fortunately, the emerging millimeter-wave (mmWave) communication offers multi-gigahertz available bandwidth and hence can cope with the explosive data demand for 5G and beyond [21]. Moreover, the short wavelength of mmWave enables massive antennas to be packed into a small physical space to support multiple-input multiple-output (MIMO) [22, 23]. To strike a balance between the performance and complexity of MIMO systems, a general hybrid transceiver utilizing matrix-monotonic optimization was designed in [22]. Shen et al. [23] proposed a phase shifter-aided beam selection network, where the energy-max beam selection method and the SIC-based beamspace precoding were carried out. Furthermore, Wang et al. [24] proposed a spectrum and energy efficient mmWave transmission scheme integrating NOMA with beamspace MIMO to maximize the sum rate. Recently, extensive research has been conducted to utilize mmWave in UAV-based networks [25–28]. In particular, the studies in [25, 26] introduced the characteristics and provided valuable insights of mmWave UAV communications. Zhao et al. [27] presented an efficient channel tracking method for a MIMO mmWave-UAV system. Rupasinghe et al. [28] rigorously evaluated the outage probability and sum rate performance of NOMA transmission with a UAV serving multiple users in mmWave.

Although a large amount of research has contributed to integrating mmWave or NOMA with UAV communications, respectively, few of them investigated on mmWave-enabled NOMA-UAV networks. Moreover, energy efficiency is of vital importance for UAVs with limited energy storage. Therefore, in this paper, we intend to develop an energy-efficient transmission design for mmWave-enabled NOMA-UAV networks by optimizing the UAV placement, hybrid precoding and PA. Particularly, we break the overall energy efficiency maximization problem down into several sub-problems. First, we optimize the position for the UAV considering the total channel strength of all UAV-served users. To group the users efficiently, a channel correlation based algorithm is proposed to select a head for each cluster. Based on this, three hybrid precoding schemes with user clustering are devised. Last, we optimize the PA among users to maximize the energy efficiency with users' quality of service (QoS) requirements, for which an efficient algorithm is proposed to solve the approximated convex problem iteratively.

The rest of this paper is organized as follows. In Section 2, the system model is introduced and the energy-efficient optimization problem is formulated. Section 3 focuses on the UAV placement problem by solving an approximated convex problem. Then, hybrid precoding is designed in Section 4, with three schemes presented. In Section 5, PA is optimized to maximize the energy efficiency, followed by numerical results provided in Section 6. Finally, we conclude the work of this paper in Section 7.

Notation. $\boldsymbol{X}$ represents a matrix while $\boldsymbol{x}$ denotes a vector. $\boldsymbol{x}^{\mathrm{T}}$, $\boldsymbol{x}^{\mathrm{H}}$ and $\|\boldsymbol{x}\|_2$ refer to the transpose, conjugate transpose and 2-norm of $\boldsymbol{x}$, respectively. $\mathrm{E}\{\cdot\}$ represents the expectation, $\mathbb{C}^{M\times N}$ is the space of $M \times N$ complex matrices and $\mathcal{CN}(\boldsymbol{x}, \boldsymbol{X})$ indicates the complex Gaussian distribution with mean $\boldsymbol{x}$ and covariance $\boldsymbol{X}$. Additionally, $\otimes$ denotes the Kronecker product and we use $\emptyset$ to denote the empty set.

2 System model and problem formulation

In this section, we first introduce the system model. Then, an optimization problem is formulated to maximize the energy efficiency by optimizing the UAV placement, hybrid precoding and PA.

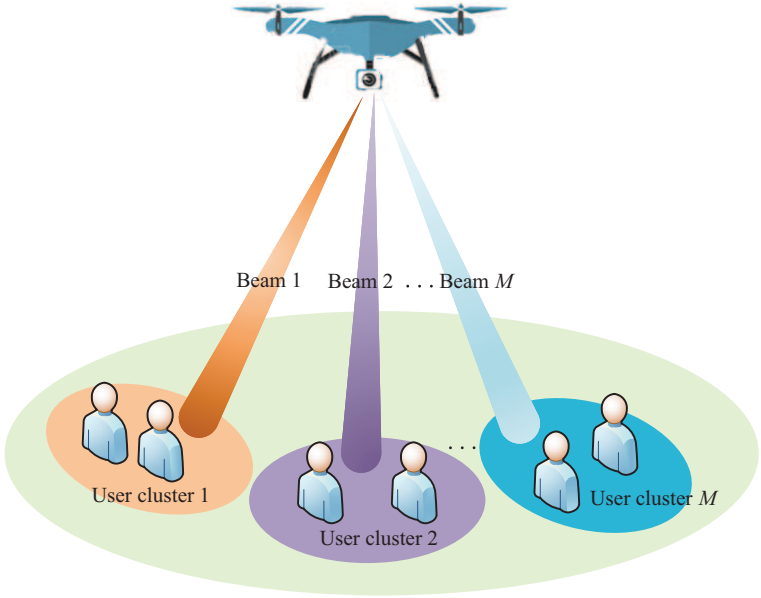

Figure 1 (Color online) Illustration of the mmWave-enabled NOMA-UAV network.

2.1 System model

Consider a mmWave NOMA-UAV network, where the UAV equipped with $N_{\rm u}$ antennas serves K single-antenna ground users in the downlink as shown in Figure 1. Owing to the fact that fully digital precoding requires one dedicated radio frequency (RF) chain per antenna, we employ a hybrid precoding (HP) architecture to reduce the hardware cost and energy consumption. As illustrated in Figure 2, the fully-connected HP is composed of a low-dimension digital precoder and an analog precoder, where all the antennas are connected to $N_{\rm RF}$ RF chains through phase shifters, with the number of RF chains much less than that of antennas. In the network, we group all the users into M clusters according to their channel conditions and each cluster is supported by one beam. The number of clusters (beams) M is assumed to be equal to that of RF chains $N_{\rm RF}$, i.e., $M = N_{\rm RF}$. To further improve spectrum efficiency and support more users, we employ NOMA in each beam to serve the users simultaneously. For $m = 1, 2, \ldots, M$, assume that there are K_m users in the mth cluster, which holds $\sum_{m=1}^{M} K_m = K$. The UAV is deployed as an arieal BS hovering above the ground at the height H, whose horizontal location is denoted as $\boldsymbol{W}_{\rm u} = (X_{\rm u}, Y_{\rm u})$. For a ground user k in the mth cluster, we define its coordinate as $\boldsymbol{W}_{m,k} = (X_{m,k}, Y_{m,k})$.

The UAV is assumed to hover at a relatively high altitude and the scattering in mmWave is limited. In addition, the LoS path is significantly stronger than Non-LoS (NLoS) paths in mmWave. Thus, similar to [28], we consider a single LoS path for the channel between the UAV and user k in the mth cluster as

$$\boldsymbol{h}_{m,k} = \sqrt{N_{\rm u}}\alpha_{m,k}\boldsymbol{f}(\phi_{m,k}, \theta_{m,k}) \in \mathbb{C}^{N_{\rm u}\times 1}, \tag{1}$$

where $\alpha_{m,k}$ denotes the complex channel gain of the LoS path and $\boldsymbol{f}(\phi_{m,k}, \theta_{m,k})$ is the steering vector related to the azimuth angle of departure (AoD) $\phi_{m,k}$ and elevation AoD $\theta_{m,k}$ for the user. Despite the fact that the presence of wind and UAV flying leads to body jittering [29], the AoDs between the UAV and users are assumed to be accurate captured as the UAV is hovering above the ground in this paper. Moreover, the complex channel gain with the free space attenuation considered can be given as

$$\alpha_{m,k} = \beta_0/d_{m,k}\mathrm{e}^{-\mathrm{j}\frac{2\pi d_{m,k}}{\lambda}}, \tag{2}$$

where β_0 denotes the channel gain at the reference distance $d_0 = 1$ m, $d_{m,k} = \sqrt{(X_{\rm u} - X_{m,k})^2 + (Y_{\rm u} - Y_{m,k})^2 + H^2}$ represents the distance between the UAV and the user, and λ is the wavelength.

Specifically, we consider a uniform planar array (UPA) in the x-y plane for the UAV with N_1 elements in the x direction and N_2 elements in the y direction as shown in Figure 3, where $N_1 N_2 = N_{\rm u}$. Then the array response vector can be expressed as

$$\boldsymbol{f}(\phi, \theta) = \boldsymbol{a}_x(\phi, \theta)^{\rm T} \otimes \boldsymbol{a}_y(\phi, \theta)^{\rm T}, \tag{3}$$

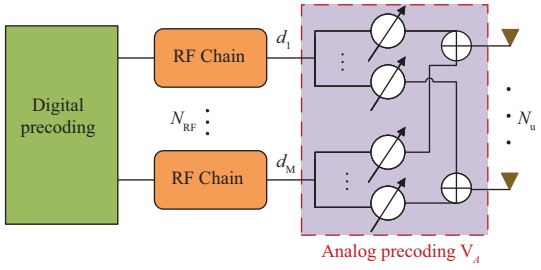

Figure 2 (Color online) Fully-connected hybrid precoding for mmWave NOMA-UAV networks.

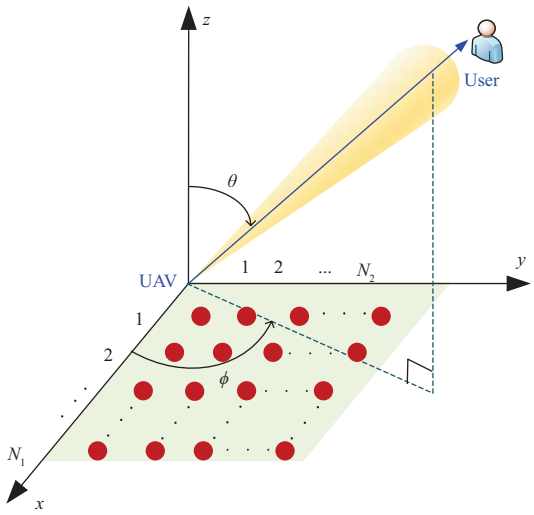

Figure 3 (Color online) 3D geometric illustration of the mmWave link from the UAV with UPA to a specific ground user.

where

$$\boldsymbol{a}_x(\phi,\theta)=\sqrt{1/N_1}\left[1,\mathrm{e}^{-\mathrm{j}2\pi(\widetilde{d}_1/\lambda)\sin\theta\cos\phi},\ldots,\mathrm{e}^{-\mathrm{j}2\pi(\widetilde{d}_1/\lambda)(N_1-1)\sin\theta\cos\phi}\right], \tag{4}$$

$$\boldsymbol{a}_y(\phi,\theta)=\sqrt{1/N_2}\left[1,\mathrm{e}^{-\mathrm{j}2\pi(\widetilde{d}_2/\lambda)\sin\theta\sin\phi},\ldots,\mathrm{e}^{-\mathrm{j}2\pi(\widetilde{d}_2/\lambda)(N_2-1)\sin\theta\sin\phi}\right], \tag{5}$$

where $\widetilde{d}_1$ and $\widetilde{d}_2$ denote the antenna spacing in axises x and y, respectively. We usually have $\widetilde{d}_1=\widetilde{d}_2=\lambda/2$.

As shown in Figure 3, the coordinate system is established with the UAV located at the origin, in which the ground users have the same z-coordinate, H. The antenna elements are marked in red circles. Particularly, the position information can be presented through the existing positioning systems or using a barometric sensor which is usually equipped in many vehicles. As a result, the corresponding azimuth and elevation AoDs for the kth user in the mth cluster can be given as

$$\phi_{m,k}=\arctan\frac{Y_{m,k}-Y_{\mathrm{u}}}{X_{m,k}-X_{\mathrm{u}}},\quad \theta_{m,k}=\arctan\frac{\sqrt{(X_{\mathrm{u}}-X_{m,k})^2+(Y_{\mathrm{u}}-Y_{m,k})^2}}{H},\quad \forall m,k. \tag{6}$$

To enhance the network performance, hybrid precoding is adopted, including the analog precoding and digital precoding design. Particularly, the analog precoding matrix $\boldsymbol{V}_{\mathrm{A}}\in\mathbb{C}^{N_{\mathrm{u}}\times M}$ and the digital precoding matrix $\boldsymbol{V}_{\mathrm{D}}\in\mathbb{C}^{M\times M}$ can be denoted as

$$\boldsymbol{V}_{\mathrm{A}}=[\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_M],\quad \boldsymbol{V}_{\mathrm{D}}=[\boldsymbol{d}_1,\boldsymbol{d}_2,\ldots,\boldsymbol{d}_M]. \tag{7}$$

Since phase shifters are implemented for analog precoding, all the elements of $\boldsymbol{a}_m\in\mathbb{C}^{N_{\mathrm{u}}\times 1}$ have the same amplitude $1/\sqrt{N_{\mathrm{u}}}$ but with different phases. The digital precoding vector for the mth cluster, $\boldsymbol{d}_m\in\mathbb{C}^{M\times 1}$, should satisfy $\|\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m\|_2=1$ for $m=1,2,\ldots,M$.

As a result, the received signal at the kth user in the mth cluster can be written as

$$\begin{aligned}y_{m,k}=\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\sum_{i=1}^{M}\sum_{j=1}^{K_i}\boldsymbol{d}_i\sqrt{p_{i,j}}x_{i,j}+n_{m,k}=\underbrace{\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m\sqrt{p_{m,k}}x_{m,k}}_{\text{Desired signal}}\\ \underbrace{+\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m\left(\sum_{j=1}^{k-1}\sqrt{p_{m,j}}x_{m,j}+\sum_{j=k+1}^{K_m}\sqrt{p_{m,j}}x_{m,j}\right)}_{\text{Intra-cluster interference}}+\underbrace{\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\sum_{i\neq m}\sum_{j=1}^{K_i}\boldsymbol{d}_i\sqrt{p_{i,j}}x_{i,j}}_{\text{Inter-cluster interference}}+n_{m,k},\end{aligned} \tag{8}$$

where $x_{i,j}$ with $\mathrm{E}\{|x_{i,j}|^2\}=1$, $p_{i,j}$ and $n_{i,j}\sim\mathcal{CN}\left(0,\sigma^2\right)$ denote the transmit signal, transmit power and additive white Gaussian noise (AWGN) for the jth user in the ith cluster, respectively.

In each cluster, NOMA is employed to improve the spectrum efficiency, where SIC is adopted at the receivers. Without loss of generality, assume $|\boldsymbol{h}_{m,1}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m|^2\geqslant|\boldsymbol{h}_{m,2}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m|^2\geqslant\cdots\geqslant|\boldsymbol{h}_{m,K_m}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m|^2, m=1,2,\ldots,M$. After SIC, the kth user can eliminate the interference from the users whose indices $j>k$ in the same cluster. Therefore, the achievable rate for user k in the mth cluster can be calculated as

$$R_{m,k}=\log_2\left(1+\frac{p_{m,k}|\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m|^2}{|\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m|^2\sum_{j=1}^{k-1}p_{m,j}+\sum_{i\neq m}|\boldsymbol{h}_{m,k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_i|^2\sum_{j=1}^{K_i}p_{i,j}+\sigma^2}\right). \tag{9}$$

2.2 Problem formulation

Without loss of generality, the energy consumption of UAVs mainly includes three components: the transmit power, the circuit energy consumption and the propulsion energy consumption. Particularly, the propulsion energy consumption is usually much higher than the other two, which can be regarded as a constant owing to the fact that the UAV hovers above the area during its serving time. To this end, we aim to make a balance between the sum rate and power consumption with only the communication-related power consumption taken into consideration. The energy efficiency of the system can be defined as

$$\eta_{EE}=\frac{\sum_{m=1}^{M}\sum_{k=1}^{K_m}R_{m,k}}{\sum_{m=1}^{M}\sum_{k=1}^{K_m}p_{m,k}+N_{\mathrm{RF}}P_{\mathrm{RF}}+P_{\mathrm{BB}}}, \tag{10}$$

where $\sum_{m=1}^{M}\sum_{k=1}^{K_m}p_{m,k}$ is the total transmit power, P_{RF} represents the power consumed by each RF chain and P_{BB} represents the baseband power consumption.

To save energy for the UAV, we aim to maximize the energy efficiency by jointly optimizing the UAV horizontal location $\boldsymbol{W}_{\mathrm{u}}$, the analog precoding matrix $\boldsymbol{V}_{\mathrm{A}}$, the digital precoding matrix $\boldsymbol{V}_D$, and the transmit power $\{p_{m,k}\}$, which can be formulated as

$$\max_{\boldsymbol{W}_{\mathrm{u}},\boldsymbol{V}_{\mathrm{A}},\boldsymbol{V}_D,\{p_{m,k}\}}\ \eta_{\mathrm{EE}} \tag{11a}$$

$$\text{s.t.}\quad R_{m,k}\geqslant R_{\min},\ \ \forall m,k, \tag{11b}$$

$$|\boldsymbol{V}_{\mathrm{A}}(i,j)|=\frac{1}{\sqrt{N_{\mathrm{u}}}},\ \ \forall i,j, \tag{11c}$$

$$\|\boldsymbol{V}_{\mathrm{A}}\boldsymbol{d}_m\|_2=1,\ \ \forall m, \tag{11d}$$

$$p_{m,k}\geqslant 0,\ \ \forall m,k, \tag{11e}$$

$$\sum_{m=1}^{M}\sum_{k=1}^{K_m}p_{m,k}\leqslant P_{\max}, \tag{11f}$$

$$X_{\min}\leqslant X_{\mathrm{u}}\leqslant X_{\max}, \tag{11g}$$

$$Y_{\min}\leqslant Y_{\mathrm{u}}\leqslant Y_{\max}, \tag{11h}$$

where Eq. (11b) guarantees that the rate of each ground user is greater than a threshold $R_{\min}$ and Eq. (11f) ensures that the total transmit power of UAV cannot exceed $P_{\max}$. Moreover, Eqs. (11g) and (11h) specify the minimum and maximum values allowed for X_{u} and Y_{u}, respectively.

However, Eq. (11) is difficult to solve owing to a large number of variables and non-convex constraints. Thus, we solve the problems of the UAV placement, the hybrid precoding and the PA sequentially in Sections 3–5, but working together to obtain a sub-optimal solution to the energy-efficient design for the system.

3 UAV placement optimization

Owing to the LoS-dominated channels, the UAV placement can significantly affect the transmission performance. We use $\|\boldsymbol{h}_{m,k}\|_2^2$ to denote the channel strength of user k in the mth cluster, and the total channel strength of the users can be maximized via optimizing the UAV's horizontal location as

$$\max_{\boldsymbol{W}_{\mathrm{u}}}\ \sum_{m=1}^{M}\sum_{k=1}^{K_m}\|\boldsymbol{h}_{m,k}\|_2^2. \tag{12}$$

Owing to the fact that $\|\boldsymbol{f}(\phi_{m,k},\theta_{m,k})\|_2^2=1$, we have $\|\boldsymbol{h}_{m,k}\|_2^2\leqslant N_{\rm u}(\beta_0/d_{m,k})^2$ according to (1) and (2).

By omitting the constants β_0 and $N_{\rm u}$, Eq. (12) can be approximated to the objective function as

$$\max_{\boldsymbol{W}_{\rm u}}\sum_{m=1}^{M}\sum_{k=1}^{K_m}\frac{1}{{d_{m,k}}^2}=\max_{\boldsymbol{W}_{\rm u}}\sum_{m=1}^{M}\sum_{k=1}^{K_m}\frac{1}{\|\boldsymbol{W}_{\rm u}-\boldsymbol{W}_{m,k}\|_2^2+H^2}. \tag{13}$$

According to [30], when it satisfies $H>\sqrt{3}\max(d_{m,k})$, the objective function in (13) is concave and thus can achieve the optimal solution. However, this condition cannot always be satisfied, which enables us to find a sub-optimal solution of $\boldsymbol{W}_{\rm u}$ for UAV. Note that the objective function is convex with respect to $\|\boldsymbol{W}_{\rm u}-\boldsymbol{W}_{m,k}\|_2^2$, and thus, we can adopt the successive convex optimization to approximate it into a convex one in each iteration. Specifically, with given $\boldsymbol{W}_{\rm u}^r$, the convex function with respect to $\|\boldsymbol{W}_{\rm u}-\boldsymbol{W}_{m,k}\|_2^2$ is lower-bounded by its first-order Taylor expansion at the local point as

$$\frac{1}{\|\boldsymbol{W}_{\rm u}-\boldsymbol{W}_{m,k}\|_2^2+H^2}\geqslant A_{m,k}^r-(A_{m,k}^r)^2\left(\|\boldsymbol{W}_{\rm u}-\boldsymbol{W}_{m,k}\|_2^2-\|\boldsymbol{W}_{\rm u}^r-\boldsymbol{W}_{m,k}\|_2^2\right), \tag{14}$$

where $A_{m,k}^r$ can be denoted as

$$A_{m,k}^r=\frac{1}{\|\boldsymbol{W}_{\rm u}^r-\boldsymbol{W}_{m,k}\|_2^2+H^2}. \tag{15}$$

Let $B_{m,k}^r=(A_{m,k}^r)^2$. Since $A_{m,k}^r$ only depends on the obtained $\boldsymbol{W}_{\rm u}^r$, it is not related to the optimization variable $\boldsymbol{W}_{\rm u}$. Thus, with given $\boldsymbol{W}_{\rm u}^r$, the optimization problem can be approximately formulated as

$$\min_{\boldsymbol{W}_{\rm u}}\ \sum_{m=1}^{M}\sum_{k=1}^{K_m}B_{m,k}^r\left(\|\boldsymbol{W}_{\rm u}-\boldsymbol{W}_{m,k}\|_2^2-\|\boldsymbol{W}_{\rm u}^r-\boldsymbol{W}_{m,k}\|_2^2\right) \tag{16a}$$

$$\text{s.t.}\ \ (11\text{g})\ \&\ (11\text{h}). \tag{16b}$$

It is obvious that Eq. (16) is convex, which owns a global optimal solution. As a result, the optimal solution of $\boldsymbol{W}_{\rm u}=(X_{\rm u},Y_{\rm u})$ in the $(r+1)$th iteration with $\boldsymbol{W}_{\rm u}^r$ obtained in the rth iteration can be given by

$$X_{\rm u}^*=\frac{\sum_{m=1}^{M}\sum_{k=1}^{K_m}B_{m,k}^rX_{m,k}}{\sum_{m=1}^{M}\sum_{k=1}^{K_m}B_{m,k}^r},\quad Y_{\rm u}^*=\frac{\sum_{m=1}^{M}\sum_{k=1}^{K_m}B_{m,k}^rY_{m,k}}{\sum_{m=1}^{M}\sum_{k=1}^{K_m}B_{m,k}^r}. \tag{17}$$

Owning to the approximation in (16), we can obtain a sub-optimal solution by iteratively solving the transformed convex problem (16) with updated local points until convergence. Thus, optimization tools such as CVX can be utilized to solve this convex problem with low computational complexity.

4 Multi-beam hybrid precoding design

As the number of users is much larger than that of RF chains, the users are grouped into M clusters and a channel correlation based algorithm is presented to select a cluster head for each cluster. In this section, we propose three hybrid precoding schemes with user clustering to improve the equivalent channel gains.

4.1 Cluster head selection

To perform clustering effectively, we choose M cluster heads according to the users' channel conditions. First, define the channel correlation coefficient between the ith and jth users as[1)] $\rho_{ij}=|\boldsymbol{h}_i^{\rm H}\boldsymbol{h}_j|/\|\boldsymbol{h}_i\|_2\|\boldsymbol{h}_j\|_2$. To mitigate the inter-cluster interference, a head selection algorithm is proposed to minimize the channel correlation among them. In detail, we pick out the user with the strongest channel strength as the first cluster head. After that, the users whose correlations with the selected head are less than the threshold ρ_0 are considered as candidates for the next cluster head. Particularly, the candidate with the strongest channel strength will be selected as the new cluster head. When there is no candidate, the threshold is updated by adding a small increment $\Delta\rho_0$. Based on the same idea, the set of candidates is updated until all the M cluster heads are obtained. The proposed algorithm is summarized in Algorithm 1.

1) Here, we omit the index representing the cluster number, m, before the users are grouped into clusters.

Algorithm 1 Correlation-based cluster head selection

Input: K, M, and the channel vectors $\boldsymbol{h}_k, k = 1, 2, \ldots, K$. The channel correlation threshold ρ_0.
Output: The cluster head set $\mathcal{U}$.
$\overline{\mathcal{H}} = [\|\boldsymbol{h}_1\|_2, \|\boldsymbol{h}_2\|_2, \ldots, \|\boldsymbol{h}_K\|_2]$; $[\sim, \Theta] = \text{sort}(\overline{\mathcal{H}}, \text{'descend'})$;
$\mathcal{U} = \Theta(1)$; $\mathcal{U}^c = \Theta/\mathcal{U}$; $\Lambda = \mathcal{U}$;
$C_{\text{idx}} = 2$;
while $C_{\text{idx}} \leqslant M$ **do**
  if $\Lambda = \emptyset$ **then**
    while $\Lambda = \emptyset$ **do**
      $\rho_0 = \rho_0 + \Delta\rho_0$;
      $\Lambda = \{j \in \mathcal{U}^c | \rho_{i,j} < \rho_0, \forall i \in \mathcal{U}\}$;
    end while
  end if
  $\Lambda = \{j \in \mathcal{U}^c | \rho_{i,j} < \rho_0, \forall i \in \mathcal{U}\}$;
  $\mathcal{U} = \mathcal{U} \cup \Lambda(1)$; $\mathcal{U}^c = \Theta/\mathcal{U}$;
  $C_{\text{idx}} = C_{\text{idx}} + 1$;
end while
return $\mathcal{U}$.

4.2 Intra-cluster strongest channel based precoding scheme

Based on the selected cluster heads, we focus on the hybrid precoding and propose an intra-cluster strongest channel based scheme. To be specific, analog precoding is first devised according to the channels of cluster heads. Afterwards, the remaining users are assigned to corresponding clusters in terms of their channel correlations. Last, the inter-cluster interference is suppressed by digital precoding considering the strongest user's channel in each cluster. The details of the first scheme are described as follows.

4.2.1 *Analog precoding with finite resolution phase shifters*

It is noteworthy that the components demanded for accurate phase control of phase shifters can be expensive. To reduce the hardware cost, finite resolution phase shifters are considered to be implemented. Therfore, the elements of the analog precoding matrix can be drawn from the following set:

$$\boldsymbol{V}_{\text{A}}(i,j) \in \mathcal{F}_{\text{A}} = \frac{1}{\sqrt{N_{\text{u}}}} \left\{ \mathrm{e}^{\mathrm{j}\frac{2\pi l}{2^b}}, l = 0, 1, \ldots, 2^b - 1 \right\}, \tag{18}$$

where b denotes the number of realizable bits in terms of the resolution of phase shifters.

With the obtained cluster heads, we can design the analog precoding via maximizing the array gain $|\boldsymbol{h}_{\mathcal{U}(m)}^{\text{H}} \boldsymbol{a}_m|^2$ for each cluster m. Specifically, we align the phases of precoding vector with that of the cluster head in each cluster, and then quantize the elements to their nearest points in the feasible set $\mathcal{F}_{\text{A}}$ regarding Euclidean distance. The idea behind this is to equivalently minimize the phase offset between $\boldsymbol{a}_m$ and $\boldsymbol{h}_{\mathcal{U}(m)}$. In this light, the elements of analog precoding vector in cluster m can be expressed as

$$\boldsymbol{a}_m(i) = \frac{1}{\sqrt{N_{\text{u}}}} \mathrm{e}^{\mathrm{j}\frac{2\pi \hat{l}}{2^b}}, \quad i = 1, 2, \ldots, N_{\text{u}}, \tag{19}$$

where $\hat{l}$ is selected based on the following criterion:

$$\hat{l} = \arg \min_{l \in \{0,1,\ldots,2^b-1\}} \left| \text{angle}\left(\boldsymbol{h}_{\mathcal{U}(m)}(i)\right) - 2\pi l / 2^b \right|. \tag{20}$$

4.2.2 *Clustering based on equivalent correlations*

With the obtained analog precoding vectors, the equivalent channel vectors of users can be expressed as

$$\overline{\boldsymbol{h}}_{m,k}^{\text{H}} = \boldsymbol{h}_{m,k}^{\text{H}} \boldsymbol{V}_{\text{A}}, \quad k = 1, 2, \ldots, K. \tag{21}$$

It is reasonable to believe that the users' equivalent channels in the same cluster are strongly correlated, provided they enjoy the same analog precoding vector. Motivated by the defined channel correlation coefficient, we group the kth user into the m_k^*th cluster according to

$$m_k^* = \arg \max_{m \in \{1,2,\ldots,M\}} \frac{|\overline{\boldsymbol{h}}_k^{\text{H}} \overline{\boldsymbol{h}}_{\mathcal{U}(m)}|}{\|\overline{\boldsymbol{h}}_k\|_2 \left\|\overline{\boldsymbol{h}}_{\mathcal{U}(m)}\right\|_2}. \tag{22}$$

In consequence, the users of the same cluster will own high correlation of equivalent channels to benefit their channel gains, while sharing low correlation in different clusters to diminish the inter-cluster interference.

4.2.3 *Zero-forcing based digital precoding*

As for the digital precoding, we adopt a zero-forcing (ZF) based beamforming to manage the inter-cluster interference. However, the inter-cluster interference cannot be totally eliminated through the digital precoding owing to that the number of users is much larger than that of clusters (RF chains). Considering the high similarity of equivalent channel vectors in the same cluster, the digital precoding is devised according to the equivalent channel with the strongest strength in each cluster. Assume that $\|\overline{\boldsymbol{h}}_{m,1}\|_2^2 \geqslant \|\overline{\boldsymbol{h}}_{m,2}\|_2^2 \geqslant \cdots \geqslant \|\overline{\boldsymbol{h}}_{m,K_m}\|_2^2, m = 1, 2, \ldots, M$. The concerning channel matrix can be written as

$$\overline{\boldsymbol{H}} = \left[\overline{\boldsymbol{h}}_{1,1}, \overline{\boldsymbol{h}}_{2,1}, \ldots, \overline{\boldsymbol{h}}_{M,1}\right]. \tag{23}$$

To suppress the inter-cluster interference, the ZF digital precoding matrix can be generated by

$$\overline{\boldsymbol{V}}_D = \overline{\boldsymbol{H}}(\overline{\boldsymbol{H}}^{\mathrm{H}}\overline{\boldsymbol{H}})^{-1} = [\overline{\boldsymbol{d}}_1, \overline{\boldsymbol{d}}_2, \ldots, \overline{\boldsymbol{d}}_M]. \tag{24}$$

Then, the digital precoding vector for each cluster is normalized as

$$\boldsymbol{d}_m = \overline{\boldsymbol{d}}_m / \left\|\boldsymbol{V}_{\mathrm{A}}\overline{\boldsymbol{d}}_m\right\|_2, \quad m = 1, 2, \ldots, M. \tag{25}$$

4.3 Singular value decomposition based precoding scheme

The aforementioned scheme is premised on that the equivalent channels of users in each cluster are highly correlated, which may be not satisfied in some cases. To this end, we propose a second precoding scheme with all the users' channel vectors taking into consideration when performing the digital precoding, adopting the same analog precoding and user clustering as the previous one. Specifically, for the mth cluster, we define a channel matrix comprised of K_m users' equivalent channel vectors as

$$\widetilde{\boldsymbol{H}}_m = \left[\overline{\boldsymbol{h}}_{m,1}, \overline{\boldsymbol{h}}_{m,2}, \ldots, \overline{\boldsymbol{h}}_{m,K_m}\right]. \tag{26}$$

Subsequently, take the singular value decomposition (SVD) of matrix $\widetilde{\boldsymbol{H}}_m$ as

$$\widetilde{\boldsymbol{H}}_m^{\mathrm{H}} = \boldsymbol{U}_m \Sigma_m \boldsymbol{V}_m^{\mathrm{H}}, \tag{27}$$

where $\boldsymbol{U}_m$ is the $K_m \times K_m$ unitary matrix, Σ_m is the diagonal and positive semi-definite matrix of size $K_m \times N_{\mathrm{RF}}$, and $\boldsymbol{V}_m$ denotes the $N_{\mathrm{RF}} \times N_{\mathrm{RF}}$ unitary matrix. In this way, the equivalent channel vector for each cluster can be obtained as

$$\overline{\boldsymbol{h}}_m = \widetilde{\boldsymbol{H}}_m \boldsymbol{u}_m^*, \quad m = 1, 2, \ldots, M, \tag{28}$$

where $\boldsymbol{u}_m$ is the first column of $\boldsymbol{U}_m$, corresponding to the maximum singular value. Afterwards, the effective channel matrix can be denoted as

$$\overline{\boldsymbol{H}} = \left[\overline{\boldsymbol{h}}_1, \overline{\boldsymbol{h}}_2, \ldots, \overline{\boldsymbol{h}}_M\right]. \tag{29}$$

Similar to the first scheme, the digital precoding can also be obtained by employing (24) and (25).

4.4 Phase centroid of clusters based precoding scheme

We propose a third precoding scheme in this subsection, with the following steps.

4.4.1 *User clustering depending on channel correlations*

In this scheme, we first perform user clustering with the cluster heads $\mathcal{U}$ obtained by Algorithm 1. Specifically, we can assign the kth user into the m_k^*th cluster according to the criterion as

$$m_k^* = \arg \max_{m \in \{1,2,\ldots,M\}} \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{h}_{\mathcal{U}(m)}|}{\|\boldsymbol{h}_k\|_2 \|\boldsymbol{h}_{\mathcal{U}(m)}\|_2}, \tag{30}$$

which means to group the users of high channel correlation into the same cluster.

4.4.2 *Hybrid precoding design*

Unlike the two precoding schemes proposed above, the analog precoding vectors are designed in this scheme to match the overall channels of all the users in the same cluster. Particularly, the elements of analog precoding vector for the mth cluster can be obtained as

$$\boldsymbol{a}_m(i) = \frac{1}{\sqrt{N_{\mathrm{u}}}} \mathrm{e}^{\mathrm{j}\frac{2\pi\hat{l}}{2^b}}, \quad i = 1, 2, \ldots, N_{\mathrm{u}}, \tag{31}$$

where

$$\hat{l} = \arg \min_{l \in \{0,1,\ldots,2^b-1\}} \left| \mathrm{mean}_{k=1,\ldots,K_m} \{\mathrm{angle}(\boldsymbol{h}_{m,k})\} - 2\pi l/2^b \right|. \tag{32}$$

Herein, the phase of precoding vector $\boldsymbol{a}_m$ is intended to align to the phase centroid of users' channel vectors in the mth cluster.

As for the digital precoding, we take the SVD of equivalent channel matrix according to (21), (26)–(29), and also consider a ZF-based method as performed in (24) and (25).

Remark 1. In this section, we have proposed three hybrid precoding schemes, all of which perform user clustering and design the analog and digital precodings to improve the multiplexing gains and suppress the inter-user interference. However, they all have their own distinctive features as follows.

- The first scheme only considers the equivalent channels of specific users in each cluster for precoding design, which is of low complexity. Nevertheless, the rate fairness among users cannot be guaranteed.
- The second scheme considers the SVD of the aggregate of users' equivalent channels when designing the digital precoding to balance the complexity and user fairness.
- The third scheme first conducts user clustering, and then performs the analog and digital precoding considering all users' channel vectors. This scheme can guarantee user fairness at the cost of a slightly higher complexity.

5 Energy-efficient power allocation

With the sub-optimal UAV placement and the hybrid precoding schemes, we perform PA among users to maximize the energy efficiency in this section.

5.1 Problem formulation

We aim to maximize the network energy efficiency under users' QoS requirements. To this end, the PA problem can be separated from (11) as

$$\max_{\{p_{m,k}\}} \quad \eta_{\mathrm{EE}} = \frac{\sum_{m=1}^{M}\sum_{k=1}^{K_m} R_{m,k}}{\sum_{m=1}^{M}\sum_{k=1}^{K_m} p_{m,k} + P_o} \tag{33a}$$

$$\text{s.t.} \quad R_{m,k} \geqslant R_{\min}, \quad \forall m, k, \tag{33b}$$

$$p_{m,k} \geqslant 0, \quad \forall m, k, \tag{33c}$$

$$\sum_{m=1}^{M}\sum_{k=1}^{K_m} p_{m,k} \leqslant P_{\max}, \tag{33d}$$

where $P_o = N_{\mathrm{RF}} P_{\mathrm{RF}} + P_{\mathrm{BB}}$. Obviously, the problem is non-convex owing to the non-concave objective function (33a). To tackle this problem, we make approximations to transform it into a convex one.

5.2 Approximate transformations

The problem (33) can be classified as a fractional programming, which can be equivalently converted into a set of subproblems in a parametric subtractive form as

$$\max_{\{p_{m,k}\}} \quad \sum_{m=1}^{M}\sum_{k=1}^{K_m} R_{m,k} - \lambda^{(r)} \left(\sum_{m=1}^{M}\sum_{k=1}^{K_m} p_{m,k} + P_o \right) \tag{34a}$$

$$\text{s.t.} \quad \text{(33b) \& (33c) \& (33d)}, \tag{34b}$$

where $\lambda^{(r)}(r \geqslant 0)$ is a nonnegative parameter with an initial value $\lambda^{(0)} = 0$, which can be updated by

$$\lambda^{(r)} = \frac{\sum_{m=1}^{M}\sum_{k=1}^{K_m} R_{m,k}^{(r)}}{\sum_{k=1}^{K_m} p_{m,k}^{(r)} + P_o}. \tag{35}$$

In (35), $R_{m,k}^{(r)}$ can be calculated according to (9) with given points $\{p_{m,k}^{(r)}, \forall m, k\}$.

For notation convenience, we define a function as

$$G_{m,k}^{i}(\boldsymbol{h}_{m,k}, \boldsymbol{d}_i) = \left\|\boldsymbol{h}_{m,k}^{\mathrm{H}} \boldsymbol{V}_{\mathrm{A}} \boldsymbol{d}_i\right\|_2^2. \tag{36}$$

Then, the achievable rate $R_{m,k}$ can be equivalently rewritten as

$$\begin{aligned} R_{m,k} &= \log_2\left(1 + \frac{p_{m,k}G_{m,k}^{m}}{\sum_{j=1}^{k-1} p_{m,j}G_{m,k}^{m} + \sum_{i\neq m}\sum_{j=1}^{K_i} p_{i,j}G_{m,k}^{i} + \sigma^2}\right) \\ &= \underbrace{\log_2\left(\sum_{j=1}^{k} p_{m,j}G_{m,k}^{m} + \sum_{i\neq m}\sum_{j=1}^{K_i} p_{i,j}G_{m,k}^{i} + \sigma^2\right)}_{\widehat{R}_{m,k}^{1}} - \underbrace{\log_2\left(\sum_{j=1}^{k-1} p_{m,j}G_{m,k}^{m} + \sum_{i\neq m}\sum_{j=1}^{K_i} p_{i,j}G_{m,k}^{i} + \sigma^2\right)}_{\widehat{R}_{m,k}^{2}}. \end{aligned} \tag{37}$$

It is evident that $\widehat{R}_{m,k}^{1}$ and $\widehat{R}_{m,k}^{2}$ are both concave with respect to the transmit power, while the difference of them is not concave. Therefore, Eq. (34) is still not convex owing to the non-concave objective function in (34a). To this end, we adopt the successive convex optimization on $\widehat{R}_{m,k}^{2}$ to approximate it into convex. Particularly, the approximate convex substitution of $\widehat{R}_{m,k}^{2}$ is introduced in Lemma 1.

Lemma 1. $\widehat{R}_{m,k}^{2}$ can be approximately substituted by its upper bound, which is convex, as follows:

$$\widehat{R}_{m,k}^{2} \leqslant \widehat{R}_{m,k}^{2ub} = \log_2(S_{m,k}^{r}) + \frac{\log_2(e)}{S_{m,k}^{r}}\sum_{j=1}^{k-1} G_{m,k}^{m}(p_{m,j} - p_{m,j}^{r}) + \frac{\log_2(e)}{S_{m,k}^{r}}\sum_{i\neq m}\sum_{j=1}^{K_i} G_{m,k}^{i}(p_{i,j} - p_{i,j}^{r}), \tag{38}$$

where

$$S_{m,k}^{r} = \sum_{j=1}^{k-1} p_{m,j}^{r}G_{m,k}^{m} + \sum_{i\neq m}\sum_{j=1}^{K_i} p_{i,j}^{r}G_{m,k}^{i} + \sigma^2. \tag{39}$$

Proof. Because $\widehat{R}_{m,k}^{2}$ is a concave function, it is globally upper-bounded by its first-order Taylor series expansion at any specific point. Thus, by taking first-order Taylor expansion operation on $\widehat{R}_{m,k}^{2}$ at the given points $\{p_{i,j}^{r}, \forall i, j\}$, the corresponding expression in (38) can be easily derived.

Based on Lemma 1, we can convert (34) into a convex problem in Theorem 1.

Theorem 1. With given PA $\{p_{i,j}^{r}, \forall i, j\}$ and calculated $\lambda^{(r)}$, Eq. (34) can be reformulated as

$$\max_{\{p_{m,k}\}} \widehat{R}_1 - \widehat{R}_2^{ub} - \lambda^{(r)}\left(\sum_{m=1}^{M}\sum_{k=1}^{K_m} p_{m,k} + P_o\right) \tag{40a}$$

$$\text{s.t.} \sum_{j=1}^{k-1} p_{m,j}G_{m,k}^{m} + \sum_{i\neq m}\sum_{j=1}^{K_i} p_{i,j}G_{m,k}^{i} + \sigma^2 \leqslant \frac{p_{m,k}G_{m,k}^{m}}{2^{R_{\min}} - 1}, \tag{40b}$$

$$\text{(33c) \& (33d)}, \tag{40c}$$

where

$$\widehat{R}_1 = \sum_{m=1}^{M}\sum_{k=1}^{K_m} \widehat{R}_{m,k}^{1}, \quad \widehat{R}_2^{ub} = \sum_{m=1}^{M}\sum_{k=1}^{K_m} \widehat{R}_{m,k}^{2ub}. \tag{41}$$

Proof. By substituting $\widehat{R}_{m,k}^{2}$ in the objective function with its upper bound derived in (38) and adopting simple mathematical manipulation on (33b), Eq. (34) can be transformed into the convex (40).

As expected, Eq. (40) is convex, which can be solved efficiently by available convex optimization tools.

Table 1 Simulation parameters

Parameter	Value	Parameter	Value
UAV altitude	$H = 120$ m	Channel gain at reference distance	$\beta_0 = -30$ dB
Noise power	$\sigma^2 = -100$ dBm	The channel correlation threshold	$\rho_0 = 0.05$
Number of RF chains	$N_{\mathrm{RF}} = 3$	Maximum transmit power of UAV	$P_{\max} = 500$ mW
Number of antennas at UAV	$N_{\mathrm{u}} = 64$	Power consumed by one RF chain	$P_{\mathrm{RF}} = 34.4$ mW
Baseband power consumption	$P_{\mathrm{BB}} = 200$ mW	Realizable bits of quantized phase shifters	$b = 4$

5.3 Iterative algorithm for energy-efficient PA

To maximize the energy efficiency, we propose to solve the convex problem (40) iteratively to obtain a sub-optimal solution to (34). The process of the proposed algorithm is outlined as Algorithm 2.

Algorithm 2 Energy-efficient PA algorithm

Input: Initialize feasible PA set $\{p_{m,k}^{(0)}\}$. Given the maximum tolerance ϵ^*. Denote the index of iteration as $r = 0$, $\lambda^{(0)} = 0$, $\epsilon^{(0)} = 10\epsilon^*$.

1: **while** $|\epsilon^r| > \epsilon^*$ **do**

2: With given $\lambda^{(r)}$ and $\{p_{m,k}^{(r)}\}$, solve the problem (40) iteratively to converge to the new PA set denoted as $\{p_{m,k}^{(r+1)}\}$.

3: Update $r = r + 1$.

4: Calculate $\lambda^{(r)} = \sum_{m=1}^{M}\sum_{k=1}^{K_m} R_{m,k}^{(r)}/(\sum_{k=1}^{K_m} p_{m,k}^{(r)} + P_o)$, $\epsilon^{(r)} = \sum_{m=1}^{M}\sum_{k=1}^{K_m} R_{m,k}^{(r)} - \lambda^{(r-1)}(\sum_{m=1}^{M}\sum_{k=1}^{K_m} p_{m,k}^{r} + P_o)$;

5: Denote the optimal solution as $\{p_{m,k}^*\} = \{p_{m,k}^{(r)}\}$;

6: **end while**

Output: The final solution of PA $\{p_{m,k}^*\}$.

Particularly, Algorithm 2 includes outer iteration and inner iteration. The outer iteration is the fractional programming, which has been demonstrated to converge to the stationary and optimal solution in [31]. Since Eq. (34) belongs to the D.C. (difference of two concave functions) programming, in the inner iteration, we solve its approximated problem (40) iteratively. The work in [32] verified that the D.C. programming can always converge with finite iterations. As a result, the proposed algorithm will surely converge to a sub-optimal solution to the problem (33).

Assume that the maximum number of outer iteration and inner iteration is denoted as I_{o} and I_{i}, respectively. Then, the overall computational complexity of Algorithm 2 can be calculated as $\mathcal{O}(I_{\mathrm{o}}I_{\mathrm{i}}C^2)$, where C denotes the number of dual variables for solving (40). Thus, we can conclude that the energy-efficient PA problem in (33) can be efficiently solved by Algorithm 2 with polynomial complexity.

6 Numerical results

In this section, simulation results are provided to demonstrate the performance of our proposed schemes. Assume that the ground users are randomly and uniformly distributed within an area of 100 m $\times$100 m. In each case, the results are obtained by averaging over 2000 random distributions. Unless stated otherwise, the default simulation parameters are listed in Table 1. Consider the UAV transmits to $K = 6$ users simultaneously and the users are grouped into $M = 3$ clusters. The performance of the three proposed precoding schemes is compared in Figures 4–6, with the UAV placement and PA optimized as proposed in the paper. For simplicity, we use "proposed scheme I", "proposed scheme II" and "proposed scheme III" to denote the three schemes in the order that they are mentioned in Section 4.

In Figure 4, we depict the energy efficiency and sum rate of the three schemes versus the minimum rate requirement, $R_{\min}$. The curves of energy efficiency demonstrate that proposed scheme I achieves significantly higher energy efficiency than the other two schemes and proposed scheme III yields the lowest energy efficiency. Moreover, the energy efficiency all decreases with $R_{\min}$ in the three schemes owing to the fact that larger $R_{\min}$ means stricter constraint on the PA optimization. Meanwhile, the result of sum rate is similar to that of energy efficiency, where proposed scheme I achieves the highest sum rate and proposed scheme III performs worst. In addition, the sum rate also decreases when $R_{\min}$ increases. In Figure 5, we compare the sum power for UAV under these three schemes. In contrast to the sum rate, the sum power increases as $R_{\min}$ increases, and proposed scheme I exploits the highest transmit power while proposed scheme III consumes the lowest. The result reveals that proposed scheme I realizes a higher sum rate at the cost of consuming more power, which consequently achieves a higher energy efficiency.

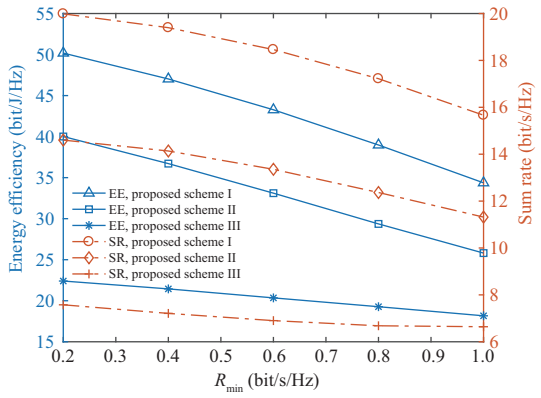

Figure 4 (Color online) Energy efficiency and sum rate against the minimum rate threshold $K = 6$.

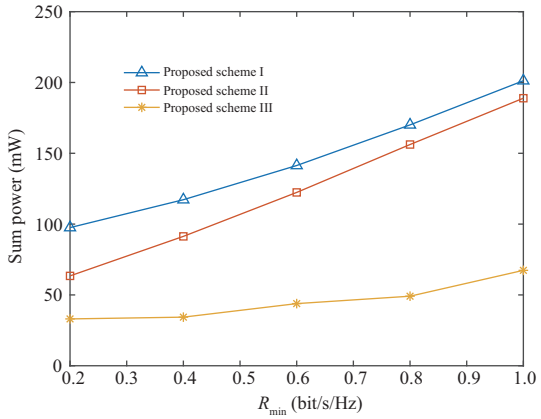

Figure 5 (Color online) Sum power consumed by UAV of the three proposed schemes $K = 6$.

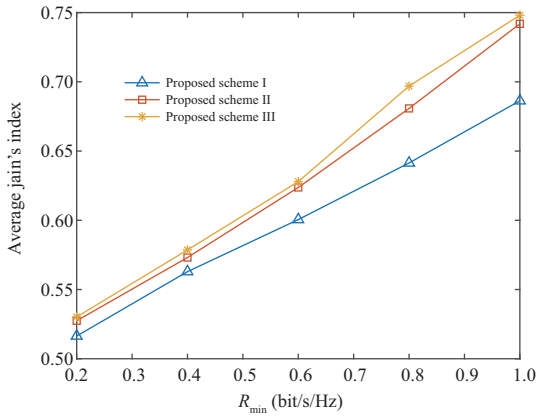

Figure 6 (Color online) Average Jain's index comparison of the three proposed schemes $K = 6$.

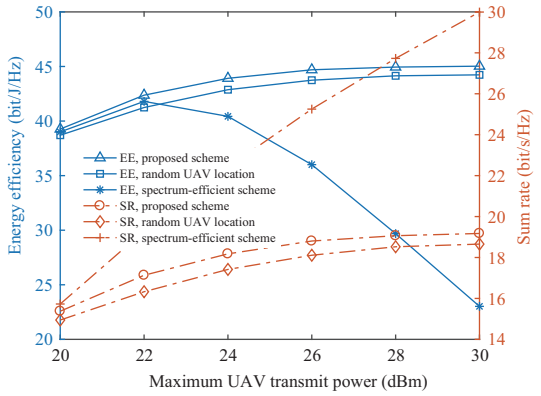

Figure 7 (Color online) Energy efficiency and sum rate comparison between the proposed scheme and two benchmarks.

The fairness of the three schemes concerned with the transmission rate is evaluated via Jain's fairness index in Figure 6. For the transmission rate vector $\boldsymbol{R}$ of K users, Jain's fairness index can be given by [33]

$$J(\boldsymbol{R}) = \left(\sum_{k=1}^{K} R_k\right)^2 \Big/ \left(K \sum_{k=1}^{K} R_k^2\right). \tag{42}$$

From Figure 6, we can see that the fairness of the second and third schemes is close and much better than the first one. Moreover, Jain's fairness index increases with $R_{\min}$ because of the tighter constraint on the rate of users. To sum up the results in Figures 4–6, proposed scheme I can obtain a higher energy efficiency but consumes more transmit power and leads to the unfairness among users. By comparison, the third scheme requires much less transmit power to guarantee the users' QoS, which enhances the performance of fairness but achieves a lower energy efficiency. Nevertheless, it can save power with the QoS satisfied. The second scheme can balance the energy efficiency and user fairness. In practice, we can adopt a proper scheme according to the available resource and performance demands.

Furthermore, to certify the effectiveness of our proposed energy-efficient scheme, we compare three schemes: (1) "proposed scheme", where the first hybrid precoding scheme is adopted; (2) "random UAV location", where the UAV is randomly deployed without placement optimization; (3)"spectrum-efficient scheme", which aims to maximize the sum rate of users. The three schemes are compared in terms of the energy efficiency and sum rate in Figure 7, where $R_{\min} = 0.5$ bit/s/Hz and $K = 6$. We can see that the proposed scheme achieves the highest energy efficiency for all the feasible values of $P_{\max}$, which validates the effectiveness of our proposed energy-efficient design. Comparing the proposed scheme and random

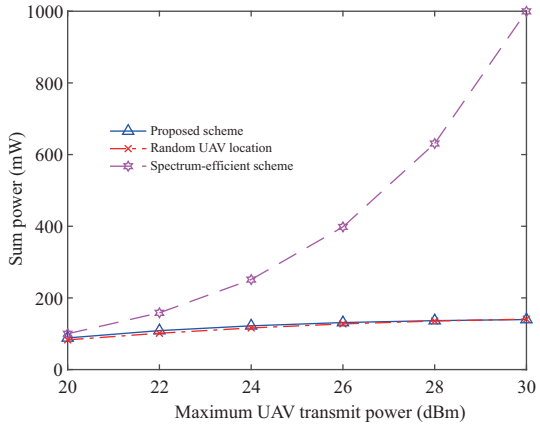

Figure 8 (Color online) Sum transmit power comparison between the proposed scheme and two benchmarks, $R_{\min} = 0.5$ bit/s/Hz.

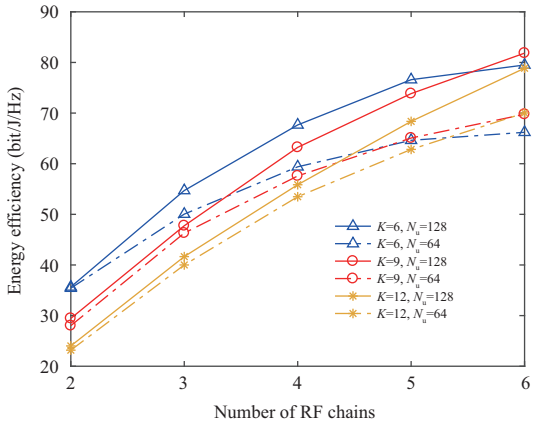

Figure 9 (Color online) Energy efficiency of proposed scheme I versus the different number of ground users and RF chains, $R_{\min} = 0.2$ bit/s/Hz.

UAV location, the performance of UAV placement optimization can be verified. The energy efficiency of spectrum-efficient scheme first grows with $P_{\max}$ and then declines. This is because the increase of sum rate cannot catch that of the maximum UAV transmit power, especially when $P_{\max}$ is high. The sum rate of the two energy-efficient schemes improves with $P_{\max}$ and slows down when $P_{\max}$ is high enough, because $P_{\max}$ is not fully utilized to save energy when $P_{\max}$ is high. As for spectrum-efficient scheme, the sum rate shows nearly linear increase with $P_{\max}$, owing to the full utilization of $P_{\max}$.

In Figure 8, the sum transmit power is compared with different values of $P_{\max}$ in the three schemes. In contrast to spectrum-efficient scheme that always consumes all the provided maximum transmit power $P_{\max}$, the sum power consumed in the other two schemes increases slightly and then remains almost fixed. This is because once the value of $P_{\max}$ reaches the optimal one for PA, the transmit power will not change to maximize energy efficiency. In spectrum-efficient scheme, it always tends to exploit all the available power to maximum the sum rate.

To further demonstrate the performance of proposed scheme I, we show the energy efficiency against the number of ground users and RF chains in Figure 9. From the result, we can conclude that properly increasing the number of RF chains will enhance energy efficiency, especially for the case with more users. Specifically, the improvement in the case $K = 6$ is not significant when $N_{\mathrm{RF}} \geqslant 5$ since the improvement of sum rate cannot compensate for the power consumption caused by increasing N_{RF}. Additionally, the performance of $N_{\mathrm{u}} = 128$ is always better than that of $N_{\mathrm{u}} = 64$ because more antennas can be exploited for efficient beamforming. Seen from the curves with different number of users, the energy efficiency is not monotonically decreasing with the number of users, especially when N_{RF} is small. Taking the curves of $N_{\mathrm{u}} = 64$ for example, when $N_{\mathrm{RF}} < 5$, the curve of $K = 9$ lies below the curve of $K = 6$, while the former exceeds the latter when $N_{\mathrm{RF}} \geqslant 5$. Similar results can also be seen from the cases with $N_{\mathrm{u}} = 128$ when $N_{\mathrm{RF}} \geqslant 6$. The slopes of curves show that the increasing speed is higher with more users, which indicates that the performance enhanced by increasing N_{RF} is more obvious in cases with more users.

7 Conclusion

In this paper, we have investigated the energy efficiency maximization problem for a mmWave-enabled NOMA-UAV network. The joint optimization problem cannot be tackled directly, which is thus divided into several subproblems. Specifically, the UAV placement is first optimized to improve the overall channel strength of users. Based on this, we present three precoding schemes where users are grouped into clusters via NOMA. To maximize the energy efficiency, the PA among users is optimized, which is solved by an iterative algorithm. Numerical results are presented to validate the effectiveness of the proposed schemes and provide more insights into the three schemes.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61871065, 61971194).

References

1 Luo S X, Zhang Z S, Wang S, et al. Network for hypersonic UCAV swarms. Sci China Inf Sci, 2020, 63: 140311

2 Zeng Y, Zhang R, Lim T J. Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun Mag, 2016, 54: 36–42
3 Zhao N, Lu W D, Sheng M, et al. UAV-assisted emergency networks in disasters. IEEE Wirel Commun, 2019, 26: 45–51
4 Cheng F, Gui G, Zhao N, et al. UAV-relaying-assisted secure transmission with caching. IEEE Trans Commun, 2019, 67: 3140–3153
5 You C S, Zhang R. 3D trajectory optimization in rician fading for UAV-enabled data harvesting. IEEE Trans Wirel Commun, 2019, 18: 3192–3207
6 Wu Q Q, Xu J, Zhang R. Capacity characterization of UAV-enabled two-user broadcast channel. IEEE J Sel Areas Commun, 2018, 36: 1955–1971
7 Lyu J B, Zeng Y, Zhang R, et al. Placement optimization of UAV-mounted mobile base stations. IEEE Commun Lett, 2017, 21: 604–607
8 Sun Y, Xu D F, Ng D W K, et al. Optimal 3D-trajectory design and resource allocation for solar-powered UAV communication systems. IEEE Trans Commun, 2019, 67: 4281–4298
9 Zhao N, Pang X W, Li Z, et al. Joint trajectory and precoding optimization for UAV-assisted NOMA networks. IEEE Trans Commun, 2019, 67: 3723–3735
10 Zeng Y, Lyu J B, Zhang R. Cellular-connected UAV: potential, challenges, and promising technologies. IEEE Wirel Commun, 2019, 26: 120–127
11 Amorim R, Nguyen H, Wigard J, et al. Measured uplink interference caused by aerial vehicles in LTE cellular networks. IEEE Wirel Commun Lett, 2018, 7: 958–961
12 Xiao H L, Zhang Z S. Swarm intelligence approaches to power allocation for downlink base station cooperative system in dense cellular networks. Sci China Inf Sci, 2020, 63: 169302
13 Mei W D, Wu Q Q, Zhang R. Cellular-connected UAV: uplink association, power control and interference coordination. IEEE Trans Wirel Commun, 2019, 18: 5380–5393
14 Lu Z Y, Sun L L, Zhang S, et al. Optimal power allocation for secure directional modulation networks with a full-duplex UAV user. Sci China Inf Sci, 2019, 62: 080304
15 Mei W D, Zhang R. Uplink cooperative NOMA for cellular-connected UAV. IEEE J Sel Top Signal Process, 2019, 13: 644–656
16 Ding Z G, Lei X F, Karagiannidis G K, et al. A survey on non-orthogonal multiple access for 5G networks: research challenges and future trends. IEEE J Sel Areas Commun, 2017, 35: 2181–2195
17 Liu Y W, Qin Z J, Cai Y L, et al. UAV communications based on non-orthogonal multiple access. IEEE Wirel Commun, 2019, 26: 52–57
18 Zhao N, Li Y X, Zhang S, et al. Security enhancement for NOMA-UAV networks. IEEE Trans Veh Technol, 2020, 69: 3994–4005
19 Sohail M F, Leow C Y, Won S. Non-orthogonal multiple access for unmanned aerial vehicle assisted communication. IEEE Access, 2018, 6: 22716–22727
20 Pang X W, Gui G, Zhao N, et al. Uplink precoding optimization for NOMA cellular-connected UAV networks. IEEE Trans Commun, 2020, 68: 1271–1283
21 Xiao Z Y, He T, Xia P F, et al. Hierarchical codebook design for beamforming training in millimeter-wave communication. IEEE Trans Wirel Commun, 2016, 15: 3380–3392
22 Xing C W, Zhao X, Xu W, et al. A framework on hybrid MIMO transceiver design based on matrix-monotonic optimization. IEEE Trans Signal Process, 2019, 67: 3531–3546
23 Shen W Q, Bu X Y, Gao X Y, et al. Beamspace precoding and beam selection for wideband millimeter-wave MIMO relying on lens antenna arrays. IEEE Trans Signal Process, 2019, 67: 6301–6313
24 Wang B C, Dai L L, Wang Z C, et al. Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array. IEEE J Sel Areas Commun, 2017, 35: 2370–2382
25 Zhang C Y, Zhang W Z, Wang W, et al. Research challenges and opportunities of UAV millimeter-wave communications. IEEE Wirel Commun, 2019, 26: 58–62
26 Xiao Z Y, Xia P F, Xia X G. Enabling UAV cellular with millimeter-wave communication: potentials and approaches. IEEE Commun Mag, 2016, 54: 66–73
27 Zhao J W, Gao F F, Kuang L L, et al. Channel tracking with flight control system for UAV mmWave MIMO communications. IEEE Commun Lett, 2018, 22: 1224–1227
28 Rupasinghe N, Yapici Y, Guvenc I, et al. Angle feedback for NOMA transmission in mmWave drone networks. IEEE J Sel Top Signal Process, 2019, 13: 628–643
29 Xu D F, Sun Y, Ng D W K, et al. Multiuser MISO UAV communications in uncertain environments with no-fly zones: robust trajectory and resource allocation design. IEEE Trans Commun, 2020, 68: 3153–3172
30 Shakhatreh H, Khreishah A. Optimal placement of a UAV to maximize the lifetime of wireless devices. In: Proceedings of 2018 14th International Wireless Communications & Mobile Computing Conference, Limassol, 2018. 1225–1230
31 Dinkelbach W. On nonlinear fractional programming. Manage Sci, 1967, 13: 492–498
32 Kha H H, Tuan H D, Nguyen H H. Fast global optimal power allocation in wireless networks by local D.C. programming. IEEE Trans Wirel Commun, 2012, 11: 510–515
33 Sediq A B, Gohary R H, Schoenen R, et al. Optimal tradeoff between sum-rate efficiency and Jain's fairness index in resource allocation. IEEE Trans Wirel Commun, 2013, 12: 3496–3509