SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2022, Vol. 65 202301:1–202301:15
https://doi.org/10.1007/s11432-021-3320-8

Joint channel estimation and beam selection NOMA system for satellite-based Internet of Things

Zeqiong CHEN[1,2], Jian JIAO[1,2*], Shaohua WU[1,2] & Qinyu ZHANG[1,2]

[1]*Communication Engineering Research Centre, Harbin Institute of Technology (Shenzhen), Shenzhen* 518055*, China;*
[2]*Peng Cheng Laboratory, Shenzhen* 518055*, China*

Received 19 March 2021/Revised 13 June 2021/Accepted 12 August 2021/Published online 27 September 2022

Abstract Multibeam low earth orbit (LEO) millimeter wave (mmWave) band high throughput satellite (HTS) is regarded as a key component in the upcoming satellite-based Internet of Things (S-IoT). The multibeam LEO HTS with non-orthogonal multiple access (NOMA) can support global coverage and low latency broadband access over a wide geographical area in a cost-efficient manner. However, the performance is constrained by the channel state information and subsequent beam selection in multibeam LEO HTS. We design a joint channel estimation and beam selection NOMA system for S-IoT in this paper, including the pseudo random ergodic (PRE)-throughput maximum (TM) scheme for massive multiple access, and the asynchronous CSI-assisted (AIA)-TM scheme for high throughput services. Moreover, we propose the PRE-fairness guaranteed (FG) and AIA-FG schemes for achieving better fairness with slightly loss of throughput. Simulation results show that the PRE-TM and AIA-TM schemes can adaptively reduce the required channel measurements, while approaching the optimal system throughput performance of the exhaust searching (ES) scheme with a much lower computational complexity.

Keywords satellite-based Internet of Things, non-orthogonal multiple access, channel estimation, system throughput, fairness

Citation Chen Z Q, Jiao J, Wu S H, et al. Joint channel estimation and beam selection NOMA system for satellite-based Internet of Things. Sci China Inf Sci, 2022, 65(10): 202301, https://doi.org/10.1007/s11432-021-3320-8

1 Introduction

The tremendous growth of the Internet of the Things (IoT) has prompted ubiquitous and ever growing requirements for global coverage and high throughput services, including smart agriculture, natural disaster prevention, climate monitoring and Internet of Vehicles [1, 2]. Due to the lack of base stations (BS) in oceans, forests, deserts, and other extreme topographies, the millimeter wave (mmWave) band high throughput satellite (HTS) is introduced to support broadband access and global coverage in a cost-efficient manner. Hence, the satellite-based IoT (S-IoT) has been considered as one of the technology driven paradigm shifts for beyond 5G and even 6G [3], which can support the massive multiple access and high throughput services via the HTS [4].

The HTS usually utilizes multibeam to better concentrate the power on its covered user equipment (UE) [5,6]. Furthermore, the multibeam low earth orbit (LEO) HTS can provide much lower propagation latency and path loss compared to the geostationary earth orbit satellite, which is regarded as a key component in the forthcoming S-IoT and has aroused universal attention in both industry and academia. In recent years, several giant LEO constellation projects such as Starlink, OneWeb and Telesat, have launched thousands of multibeam LEO HTS, which can cooperate with the terrestrial infrastructures to realize the Internet of Everything.

The channel estimation is critical for improving the spectrum efficiency of multibeam mmWave band HTS [7] and ensuring the quality of service (QoS) in S-IoT. Considering the sparseness of the mmWave band channel in the angle domain, the number of proper propagation path directions from the HTS to

* Corresponding author (email: jiaojian@hit.edu.cn)

UE is limited [8]. Hence, many studies have modeled the mmWave band channel as the sparse geometric-based channel [9], which can also be adopted in the satellite-terrestrial link of multibeam mmWave band HTS in S-IoT [10]. Moreover, Ref. [11] developed a compressed sensing (CS)-based algorithm to estimate the mmWave band channel for multiple UE with multibeam.

In addition, the subsequent beam selection for multibeam mmWave band HTS can significantly affect the system performance. Note that the demand for multiple access motivates the S-IoT to utilize non-orthogonal multiple access (NOMA) [12], since HTS is able to serve multiple UE simultaneously and reduce the transmission phases of S-IoT via NOMA scheme, which can make full use of the ubiquitous broadband access capability of mmWave band HTS. Moreover, although the UE within the coverage of HTS have similar distances, we can observe significant differences in their channel gains because of diverse path losses on mmWave band signals, such as weather conditions and masking effects. Ref. [13] validated that the power domain NOMA can be applied in the S-IoT system. Further, to maximize long-term throughput [14], the NOMA S-IoT system inclines to select the beam with the highest channel gain of the strong UE that own "good" channel state, which leads to a serious unequal throughput between UE and deteriorates the system fairness.

Therefore, the capacity of the mmWave NOMA S-IoT system can be significantly enhanced via multibeam LEO HTS, but the fact that the performance is constrained by the precise channel state information (CSI) and high mobility makes it essential to design a fast channel estimation for multibeam LEO HTS. Moreover, there is still lack of work on the joint design of fast channel estimation and subsequent beam selection in multibeam mmWave band LEO HTS for the NOMA S-IoT system. Therefore, we propose a joint channel estimation and beam selection NOMA S-IoT system in this paper. Moreover, the contributions of this paper can be twofold as follows.

• Joint channel estimation and beam selection NOMA S-IoT system. It is the first work to design a joint channel estimation and beam selection NOMA S-IoT system. The satellite-terrestrial link is modeled as the sparse geometric-based channel because of the high-gain and high-directivity properties of mmWave band beamforming [15]. Specially, the angle of departure (AOD) and angle of arrival (AOA) of candidate beams are narrowed by taking the altitude and elevation angle of satellite communication into account. Then, considering the sparseness feature of sparse geometric-based mmWave band channel, we utilize a CS-based algorithm to achieve a sub-optimal channel estimation with less complexity, and two random selected beamforming channel estimation algorithms are proposed for the mmWave NOMA S-IoT system. One is the pseudo random ergodic (PRE) algorithm for massive multiple access, and the other is the asynchronous CSI-assisted (AIA) algorithm for high throughput services. Moreover, two subsequent beam selection schemes are designed according to different QoS requirements, named throughput maximum (TM) and fairness guaranteed (FG) beam selection schemes.

• Parametrical optimization and performance analysis. Inspired by the rateless coding methods [16], our channel estimation algorithm randomly selects the combination from all candidate beams at the transmitter and receivers to obtain the channel matrix, which can adaptively reduce the required measurements with the increasing of signal-noise-ratio (SNR) compared with the conventional predetermined number of measurements (PNM) algorithm [17]. Then, we design a PRE algorithm to accelerate the channel estimation by modifying the probability to guarantee that all the candidate beam combinations can be spanned at least once as soon as possible. Moreover, we propose an AIA algorithm by exploiting the asynchronous CSI from the previous measurements to further accelerate channel estimation for long-term high throughput services. To this end, we finally design four joint channel estimation and beam selection schemes: the PRE-TM and AIA-TM schemes, which can achieve the maximum throughput as the exhaust searching (ES) benchmark scheme [7]. The PRE-FG and AIA-FG schemes have better fairness performance at the cost of slightly loss of throughput compared with the PRE-TM and AIA-TM schemes. Finally, we analyze the computational complexity of PRE-TM and PRE-FG schemes, and compare the system performance of our four schemes with the ES, PNM-TM and PNM-FG schemes, including the system throughput, number of channel measurements, measurement efficiency, error probability and UE fairness.

The remainder of this paper is organized as follows. In Section 2, we introduce the mmWave NOMA S-IoT downlink system and model the mmWave band channel. In Section 3, we formulate the channel estimation problem, and design two improved randomly selected beamforming channel estimation algorithms. In Section 4, we propose a UE grouping scheme, and design two subsequent beam selection algorithms to further improve the system performance. Then, we analyze the computational complexity of joint channel estimation and beam selection schemes. Finally, simulation results and comparisons

between the proposed algorithms and benchmark schemes are given in Section 5, and the conclusion is drawn in Section 6.

2 System model

We first introduce the mmWave NOMA S-IoT system, and illustrate the sparse geometric-based channel model in detail. Then, by taking account of the altitude and elevation angle of satellite communications, we narrow the AOD and AOA of mmWave band channel.

2.1 mmWave NOMA S-IoT downlink system

Assume that a NOMA S-IoT downlink system consists of a mmWave band multibeam HTS S with N_S beams, and D UE, where the i-th UE U_i is equipped with N_{U_i} beams, $i = 1, 2, \ldots, D$. Assume that S is equipped with R_S radio frequency (RF) chains ($N_S \gg R_S$) and U_i is equipped with R_{U_i} RF chains ($N_{U_i} \gg R_{U_i}$) [11].

As shown in Figure 1(a), S transmits multiple beamformed pilot signals s to its covered UE to estimate multiple channels simultaneously. Concretely, the beams are selected from all candidate beams corresponding to the beamforming vectors f_m and w_{n_i} at S and U_i, respectively, where $m = 1, 2, \ldots, N_S$ and $n_i = 1, 2, \ldots, N_{U_i}$. We define the candidate beamforming matrix of S and U_i as $F_d = [f_1, f_2, \ldots, f_{N_S}]$ and $W_{d_i} = [w_1, w_2, \ldots, w_{N_{U_i}}]$, comprising of the beamforming vectors of all candidate beams at S and U_i, respectively. Considering that the entries of candidate beamforming matrix F_d and W_{d_i} are all possible orthogonal beamforming vectors. In addition, since each beamforming vector is constrained with a RF phase shifter, all beamforming vectors have constant modulus such that $\|f_m\| = 1$ and $\|w_{n_i}\| = 1$, for any m and n_i, respectively. Moreover, we assume that the phase shifter of RF chains is constantly discrete to reduce the hardware cost, and the quantized angles can be represented as $\alpha_m = \pi - 2\pi\frac{m-1}{N_S}$ and $\beta_{n_i} = \pi - 2\pi\frac{n_i-1}{N_{U_i}}$. Therefore, f_m and w_{n_i} can be given by

$$\begin{cases} f_m = \frac{1}{\sqrt{N_S}}\left[\mathrm{e}^{\mathrm{j}\alpha_m}, \mathrm{e}^{\mathrm{j}2\alpha_m}, \ldots, \mathrm{e}^{\mathrm{j}N_S\alpha_m}\right]^{\mathrm{T}}, \quad m = 1, 2, \ldots, N_S, \\ w_{n_i} = \frac{1}{\sqrt{N_{U_i}}}\left[\mathrm{e}^{\mathrm{j}\beta_{n_i}}, \mathrm{e}^{\mathrm{j}2\beta_{n_i}}, \ldots, \mathrm{e}^{\mathrm{j}N_{U_i}\beta_{n_i}}\right]^{\mathrm{T}}, \quad n_i = 1, 2, \ldots, N_{U_i}. \end{cases} \tag{1}$$

Thus, each measurement of channel estimation for the mmWave NOMA S-IoT system needs select R_S and R_{U_i} beams from the $N_S \times N_S$ candidate beamforming matrix F_d and $N_{U_i} \times N_{U_i}$ candidate beamforming matrix W_{d_i}, respectively. Let f_p denote the $N_S \times 1$ transmitting beamforming vector, which is generated by the phase shifters of the p-th RF chain at S ($p = 1, 2, \ldots, R_S$), and w_{q_i} is the $N_{U_i} \times 1$ receiving beamforming vector generated by the phase shifters of the q_i-th RF chain at U_i ($q_i = 1, 2, \ldots, R_{U_i}$). Therefore, these R_S transmitting beamforming vectors at S form the transmitting beamforming matrix $F = [f_1, f_2, \ldots, f_{R_S}]$, and the R_{U_i} receiving beamforming vectors at U_i form the receiving beamforming matrix $W_i = \left[w_1, w_2, \ldots, w_{R_{U_i}}\right]$.

Therefore, the broadcast pilot signal s is the vector of transmitted pilot symbols, such that $\mathrm{E}\left[ss^{\mathrm{H}}\right] = I_{R_S}$, which is processed by the RF chains with the $N_S \times R_S$ transmitting beamforming matrix F. Thus, the transmitted signals x is given by

$$x = \frac{1}{\sqrt{R_S}} Fs, \tag{2}$$

and the signal observed by U_i can be represented as

$$r_i = \sqrt{\frac{P_t}{R_S}} H_i F s + u_i, \tag{3}$$

where H_i is the $N_{U_i} \times N_S$ mmWave band channel matrix between S and U_i, P_t is the transmit power of S, u_i is the complex additive white Gaussian noise (AWGN) vector follows $\mathcal{CN}\left(0, N_0 I_{N_{U_i}}\right)$.

At the UE side, U_i will process the observed signals with its RF beamformer corresponding to the $N_{U_i} \times R_{U_i}$ matrix W_i. Hence, the received signal of U_i can be written as

$$y_i = \sqrt{\frac{P_t}{R_S}} W_i^{\mathrm{H}} H_i F s + W_i^{\mathrm{H}} u_i, \tag{4}$$

where W_i^{H} represents the conjugate transpose matrix corresponding to W_i.

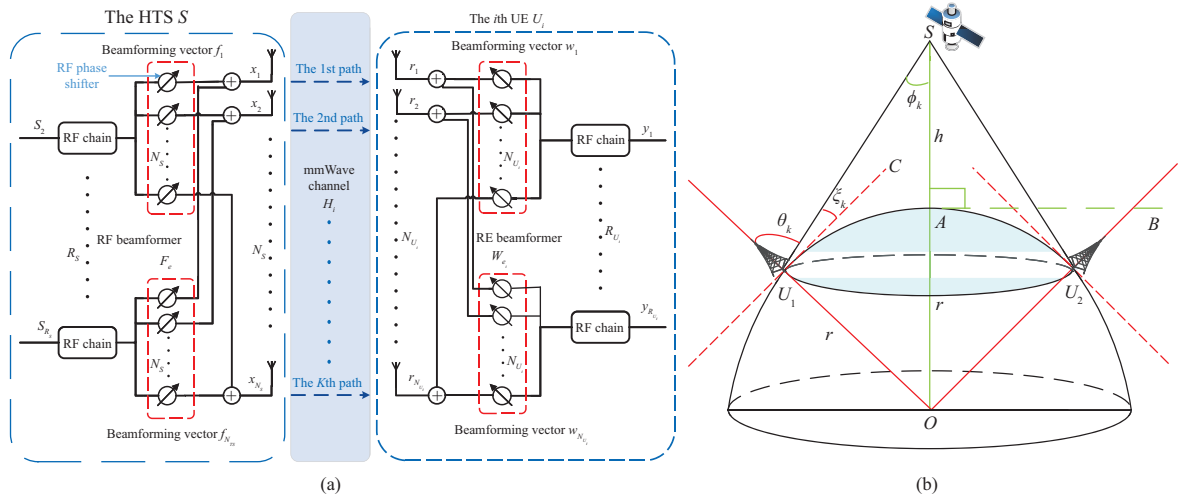

Figure 1 (Color online) The mmWave NOMA S-IoT downlink system. (a) The signal transmission process in mmWave NOMA downlink system; (b) the AOD and AOA of mmWave band channel.

2.2 Sparse geometric-based channel model

We establish a widely used sparse angle domain geometric-based channel model [9, 10] to characterize the inherent sparseness feature of mmWave band satellite-terrestrial link in angle domain. Assume that there are K paths between S and U_i as shown in Figure 1(b), we define the angle between the k-th path and array normal direction of S as AOD ϕ_k with $k = 1, 2, \ldots, K$, and the angle between the k-th path and array normal direction of U_i is AOA θ_{k_i}. Moreover, considering the hundreds of kilometers orbit altitude and mobility of LEO HTS, the Doppler shifts can be assumed to be identical and relieved by utilizing proper guard bands in the mmWave broadband channel [3, 4]. Thus, the channel matrix H_i is given by [15]

$$H_i = \sqrt{N_S N_{U_i}} \sum_{k=1}^{K} \delta_{k_i} \boldsymbol{a}_{U_i}\left(\theta_{k_i}\right)\left(\boldsymbol{a}_S\left(\phi_k\right)\right)^{\mathrm{H}}, \tag{5}$$

where $\delta_{k_i} = \bar{\delta}_{k_i} \cdot \mathrm{e}^{\mathrm{j}2\pi\tau_d}$ denotes a complex fading gain of the k-th path between S and U_i, and $\bar{\delta}_{k_i}$ is the fading coefficient following a Gaussian distribution $\mathcal{CN}\left(0, \sigma_i\right)$, which is related to the lognormal shadowing coefficient, τ_d is the Doppler shift which can be viewed as a constant multiplicator factor. Moreover, assume that the antenna spacing of S and U_i is uniform linear arrays (ULA), and the terms $\boldsymbol{a}_S(\phi_k)$ and $\boldsymbol{a}_{U_i}(\theta_{k_i})$ can be given by

$$\begin{cases} \boldsymbol{a}_S(\phi_k) = \dfrac{1}{\sqrt{N_S}}\left[1, \mathrm{e}^{\mathrm{j}\frac{\pi d \cos\phi_k}{\lambda}}, \ldots, \mathrm{e}^{\mathrm{j}\frac{\pi d(N_S-1)\cos\phi_k}{\lambda}}\right]^{\mathrm{T}}, \\ \boldsymbol{a}_{U_i}(\theta_{k_i}) = \dfrac{1}{\sqrt{N_{U_i}}}\left[1, \mathrm{e}^{\mathrm{j}\frac{\pi d \cos\theta_{k_i}}{\lambda}}, \ldots, \mathrm{e}^{\mathrm{j}\frac{\pi d(N_{U_i}-1)\cos\theta_{k_i}}{\lambda}}\right]^{\mathrm{T}}, \end{cases} \tag{6}$$

where the superscript T represents the transpose of row vector, λ is the wavelength of mmWave band signal, and d denotes the distance between antennas such that $d = \lambda/2$.

In addition, in the mmWave NOMA S-IoT downlink system, it is necessary to reconsider the range of AOA and AOD according to the elevation angle and orbit altitude of the mmWave band LEO HTS, which are different from the conventional terrestrial communication system ranging from 0 to 2π. As shown in Figure 1(b), assume that S transmits signals to U_1 and U_2, and A is a substellar point. The angle between the k-th transmitting path and horizon of U_1 is defined as elevation angle ξ_k.

First, we derive the relationship between AOA θ_{k_i}, AOD ϕ_k and elevation angle ξ_k. Here, we approximate the Earth to be a sphere globe with a radius of r, and the altitude of HTS is h, i.e., $SA = h$ and $OA = OU_1 = r$. By exploiting the law of sines, we have the relationship between ϕ_k and ξ_k, such that $\phi_k = \arcsin(\frac{r}{r+h}\cos\xi_k)$. Moreover, as U_1C represents the horizontal line of U_1, we can get $\theta_k = \frac{\pi}{2} - \xi_k$.

Without loss of generality, the elevation angle of HTS ranges from $\frac{\pi}{4}$ to $\frac{\pi}{2}$ [18], i.e., $\xi_k \in \left(\frac{\pi}{4}, \frac{\pi}{2}\right)$. Thus, the ranges of AOD and AOA are $\phi_k \in (0, \arcsin(\frac{\sqrt{2}}{2}\frac{r}{r+h}))$ and $\theta_k \in \left(0, \frac{\pi}{4}\right)$, respectively. In the following

sections, we simplify the term $\arcsin(\frac{\sqrt{2}}{2}\frac{r}{r+h})$ as $\phi_{\max}$ for convenience.

Moreover, since ULA is applied at S and U_i, the AOD and AOA of the k-th path in mmWave band satellite-terrestrial link are uniformly distributed on their ranges such that $\phi_k = \pi - \phi_{\max}(k-1)/K$ and $\theta_k = \pi - \frac{\pi}{4}(k-1)/K$.

3 Channel estimation for mmWave NOMA S-IoT system

In this section, we formulate a CS-based channel estimation for mmWave NOMA S-IoT system, and propose two improved random selected beamforming channel estimation algorithms for the massive multiple access and high throughput services, respectively.

3.1 Channel estimation

3.1.1 *Probabilistic beam selection for channel measurements*

For the channel estimation, we aim to quickly obtain precise CSI of all D UE for the later beam selection and transmission. Since the beamforming vectors are constrained with the RF phase shifters at S and U_i, we randomly select R_S and R_{U_i} beams from the candidate beamforming matrices F_d and W_{d_i} to form the transmitting beamforming matrix F_e and the receiving beamforming matrix W_{e_i} at the e-th measurement, respectively, and we have $F_e = [f_1^e, f_2^e, \ldots, f_{R_S}^e]$ and $W_{e_i} = [w_1^e, w_2^e, \ldots, w_{R_{U_i}}^e]$.

Then, we perform the e-th channel measurement by transmitting and receiving the pilot signal s with the randomly selected beams in F_e and W_{e_i}. Hence, the received signal y_i^e of U_i at the e-th measurement can be represented as

$$y_i^e = \sqrt{\frac{P_t}{R_S}} W_{e_i}^{\mathrm{H}} H_i F_e s + W_{e_i}^{\mathrm{H}} u_i^e, \tag{7}$$

where u_i^e is the complex AWGN vector of U_i such that $u_i^e \sim \mathcal{CN}\left(0, N_0 I_{N_{U_i}}\right)$.

Consider that the $N_{U_i} \times N_S$ mmWave band channel matrix H_i can be written as follows [19]:

$$H_i = \sqrt{N_S N_{U_i}} W_{d_i} H_{\nu_i} F_d^{\mathrm{H}}, \tag{8}$$

where H_{ν_i} is an $N_{U_i} \times N_S$ virtual channel matrix of H_i. Substituting (8) into (7), we have

$$y_i^e = \sqrt{\frac{N_S N_{U_i} P_t}{R_S}} W_{e_i}^{\mathrm{H}} W_{d_i} H_{\nu_i} F_d^{\mathrm{H}} F_e s + W_{e_i}^{\mathrm{H}} u_i^e. \tag{9}$$

By utilizing $\mathrm{vec}(UVW) = \left(W^{\mathrm{T}} \otimes U\right) \mathrm{vec}(V)$, we can represent (9) as

$$y_i^e = \sqrt{\frac{N_S N_{U_i} P_t}{R_S}} \left(\left(F_d^{\mathrm{H}} F_e s\right)^{\mathrm{T}} \otimes \left(W_{e_i}\right)^{\mathrm{H}} W_{d_i}\right) \mathrm{vec}\left(H_{v_i}\right) + W_{e_i}^{\mathrm{H}} u_i^e. \tag{10}$$

Therefore, F_e and W_{e_i} at S and U_i are formed by selecting beams from the candidate beamforming matrices F_d and W_{d_i} at the e-th channel measurement, respectively.

3.1.2 *Sparse estimation problem formulation*

To estimate the virtual channel matrix H_{ν_i} via the CS-based techniques to obtain H_i [20], the received signal y_i^e is required to follow the standard-form expression:

$$y_i^e = A_c A_i^e v_i + \varepsilon_i^e, \tag{11}$$

where A_c is a constant, and A_i^e denotes a sensing matrix. Comparing (10) with (11), we have $A_c = \sqrt{\frac{N_S N_{U_i} P_t}{R_S}}$, $A_i^e = ((F_d^{\mathrm{H}} F_e s)^{\mathrm{T}} \otimes ((W_{e_i})^{\mathrm{H}} W_{d_i}))$, and $v_i = \mathrm{vec}(H_{v_i})$, which is a vectorized virtual channel matrix. Since v_i is sufficiently sparse, and the $R_{U_i} \times N_S N_{U_i}$ sensing matrix A_i^e is well-conditioned, we can adopt the CS-based techniques in the following. Note that in (5) the channel fading coefficient $\delta_{k_i} \sim \mathcal{CN}(0, \sigma_i)$, let $l = 1, \ldots, N_S N_{U_i}$, and $\mu = \frac{K}{N_S N_{U_i}}$, therefore, the l-th entry of v_i follows a Bernoulli-Gaussian distribution, and we have

$$v_i^l \sim \begin{cases} 0, & \text{with probability } 1-\mu, \\ \mathcal{CN}(0, \sigma_i), & \text{with probability } \mu. \end{cases} \tag{12}$$

Then, we focus on the noise term ε_i^e in (11). Recalling (4), we can get the complex AWGN noise vector $\varepsilon_i^e = W_{e_i}^{\mathrm{H}} u_i^e \sim \mathcal{CN}(0, N_0 W_{e_i}^{\mathrm{H}} W_{e_i})$. Since the entries of W_{e_i} are mutually orthogonal, the noise term can be simplified to $\varepsilon_i^e \sim \mathcal{CN}(0, N_0 I_{R_{U_i}})$.

Therefore, we convert the channel estimation into a CS problem to estimate H_{ν_i} on the basis of matrix $y_{(e,i)} = [y_i^1, y_i^2, \ldots, y_i^e]^{\mathrm{T}}$, which is composed of the received signals obtained in former e measurements. First, we formulate this solution to a maximum-likelihood problem as follows:

$$\hat{v}_{(e,i)} = \underset{v_i}{\operatorname{argmax}} \left[p\left(y_{(e,i)} \mid v_i \right) \right], \tag{13}$$

where $\hat{v}_{(e,i)}$ represents the estimation of v_i between S and U_i based on $y_{(e,i)}$. We leverage the Lasso method [21] to complete the channel estimation based on the sparse feature of the mmWave band channel, and $\hat{v}_{(e,i)}$ becomes

$$\hat{v}_{(e,i)} = \underset{v_i}{\operatorname{argmin}} \left[\left\| y_{(e,i)} - A_c A_{(e,i)} v_i \right\|_2^2 + \kappa \| v_i \|_1 \right], \tag{14}$$

where the vector $A_{(e,i)} = \left[A_i^1, A_i^2, \ldots, A_i^e \right]^{\mathrm{T}}$. The first term $\| y_{(e,i)} - A_c A_{(e,i)} v_i \|_2^2$ makes the estimation values fit the received signals, the second term $\| v_i \|_1$ ensures the solution contains a limited number of nonzero values, and κ is the weight coefficient. Obviously, the optimal solution of (14) is extremely complicated as the dimension of v_i is $|N_{U_i} \times N_S|$. Thus, we utilize a generalized approximate message passing (GAMP) algorithm to achieve a sub-optimal solution of (14) [22]. The input and estimation output for the GAMP estimator are denoted as ϱ_{in} and ϱ_{out}, respectively. In our channel estimation algorithm, ϱ_{in} is related to v_i following (12), A_c, $A_{(e,i)}$ and $\varepsilon_i^e \sim \mathcal{CN}(0, N_0 I_{R_{U_i}})$ in (10), and ϱ_{out} is the estimation result $\hat{v}_{(e,i)}$. Then, we can estimate v_i via the GAMP estimator in [23].

3.1.3 *UE stopping criterion*

The channel estimation is considered to be completed when the channel estimation result has converged. Recall (12), we convert the l-th entry of the vectorized virtual channel matrix $\hat{v}_{(e,i)}$ obtained from the e-th channel measurement into a binarization form $\bar{v}_{(e,i)}^l$ as follows:

$$\bar{v}_{(e,i)}^l = \begin{cases} 0, & \text{if } \left| \hat{v}_{(e,i)}^l \right| < \Lambda \sigma_i, \\ 1, & \text{otherwise}, \end{cases} \tag{15}$$

where $\hat{v}_{(e,i)}^l$ denotes the l-th entry of $\hat{v}_{(e,i)}$, and Λ is a sufficiently small coefficient to ensure a reliable transmission between S and U_i. When this binarized vector satisfies $\bar{v}_{(e,i)} = \bar{v}_{(e-1,i)}$, the estimated virtual channel matrix $\hat{H}_{v_i}$ is regarded to be convergent.

Moreover, we set the maximum allowed number of channel measurements $E_{\max}$ to prevent an endless estimation, when the channel is varying in deep fading and leading to an outage.

3.2 Improved channel estimation algorithms

To accelerate the channel estimation, we propose two improved random selected beamforming channel estimation algorithms. One is called PRE algorithm, which can accomplish the channel estimation faster by modifying the probability to guarantee that each candidate beam is selected about $\frac{E_{\max}}{\lceil N_S \cdot N_{U_i} / (R_S \cdot R_{U_i}) \rceil}$ times in $E_{\max}$ measurements. The other is named AIA algorithm, which can further accelerate the channel estimation by exploiting asynchronous CSI from the previous channel measurements for long-term high throughput services.

3.2.1 *Pseudo random ergodic channel estimation algorithm*

There are generally dozens to hundreds of beams at S and U_i, and we need a small $E_{\max}$ to accomplish the channel estimation as soon as possible. However, when F_e and W_{e_i} are uniformly random selected from F_d and W_{d_i}, it may lead several f_m in F_d and w_{n_i} in W_{d_i} are not spanned in a limited $E_{\max}$ measurements, which would deteriorate the accuracy of estimated CSI or require a larger $E_{\max}$ to go through all beam combinations.

Motivated by this, in our PRE algorithm, the beam selection probability distribution ζ_S of HTS S can be calculated by the inverse proportion of the number of times that each candidate beam has been

chosen in former measurements, and the beam selection probability distribution $\zeta_S^{(e+1)}$ of HTS S for the $(e+1)$-th measurement can be represented as follows:

$$\zeta_S^{(e+1)} = \left[\frac{b_\zeta^{(e+1)}}{N_f^{(e)}(1)+\vartheta_\zeta}, \frac{b_\zeta^{(e+1)}}{N_f^{(e)}(2)+\vartheta_\zeta}, \ldots, \frac{b_\zeta^{(e+1)}}{N_f^{(e)}(N_S)+\vartheta_\zeta} \right], \tag{16}$$

where $N_f^{(e)}(m)$ denotes the total number of measurements in which the m-th beam of S has been selected in former e measurements, ϑ_ζ denotes a small positive constant to avoid the case of a zero denominator, and $b_\zeta^{(e+1)} = (\sum_m (N_f^{(e)}(m) + \vartheta_\zeta)^{-1})^{-1}$ guarantees the sum of all elements in $\zeta_S^{(e+1)}$ equals one.

To ensure that all the beam combinations are spanned as soon as possible, the beam selection probability distribution γ_{U_i} of U_i is related to ζ_S of S. Specially, $\gamma_{U_i}^{(e+1)}$ of U_i for the $(e+1)$-th measurement is affected by the $(e+1)$-th beam selection of S, and we have

$$\gamma_{U_i}^{(e+1)} = \left[\frac{b_\gamma^{(e+1)}}{\min\limits_{f\in F_{e+1}} N_w^{(e)}(1|f)+\vartheta_\gamma}, \frac{b_\gamma^{(e+1)}}{\min\limits_{f\in F_{e+1}} N_w^{(e)}(2|f)+\vartheta_\gamma}, \ldots, \frac{b_\gamma^{(e+1)}}{\min\limits_{f\in F_{e+1}} N_w^{(e)}(N_{U_i}|f)+\vartheta_\gamma} \right], \tag{17}$$

where f denotes the beamforming vector of selected beam at S in the $(e+1)$-th measurement, $N_w^{(e)}(n_i \mid f)$ is the number of measurements in which the n-th beam of U_i is selected together in previous e measurements with f, and $\min_{f\in F_{e+1}} N_w^{(e)}(n_i \mid f)$ finds the beam in f that has been selected the least times with the n-th beam of U_i in former e measurements. Similarly, ϑ_γ is a positive value to avoid zero denominator, and $b_\gamma^{(e+1)}$ is a constant to ensure the elements of $\gamma_{U_i}^{(e+1)}$ add up to one.

Thus, we update ζ_S and γ_{U_i} in each measurement in the PRE algorithm and then perform the random selected beamforming channel estimation as shown in Subsection 3.1, and all candidate beam combinations could be spanned one time in the first $\lceil N_S \cdot N_{U_i}/(R_S \cdot R_{U_i}) \rceil$ measurements.

3.2.2 *Asynchronous CSI-assisted channel estimation algorithm*

By exploiting the estimated CSI from previous e channel measurements, we further propose the AIA algorithm for long-term high throughput services. Define that the received signal power of U_i in the $(e+1)$-th measurement of each beamforming vector w_{n_i} is $p_{i,n_i}^{e+1} = w_{n_i}^{\mathrm{H}} \hat{r}_i^{e+1} (\hat{r}_i^{e+1})^{\mathrm{H}} w_{n_i}$, where w_{n_i} with higher power indicates to have higher channel gain. Therefore, we modify the beam selection probability distribution φ_{U_i} of U_i according to $p_{i,n}^{e+1}$, and U_i inclines to select w_{n_i} that has higher received signal power $p_{i,n}$ in the AIA algorithm, which narrows the range of beam selection to accelerate channel estimation.

Specifically, in our AIA algorithm, the HTS S updates its beam selection probability distribution ζ_S as the same in PRE algorithm in (16). Then, the beam selection probability distribution $\varphi_{U_i}^{(e+1)}$ of U_i in the $(e+1)$-th measurement is calculated to be directly proportional to $p_{i,n}^{e+1}$, and we have

$$\varphi_{U_i}^{(e+1)} = c_\varphi^{(e+1)} \left[p_{i,1}^{e+1}, \ldots, p_{i,N_{U_i}}^{e+1} \right] = c_\varphi^{(e+1)} \operatorname{diag}\left(W_{d_i}^{\mathrm{H}} \hat{r}_i^{e+1} (\hat{r}_i^{e+1})^{\mathrm{H}} W_{d_i} \right), \tag{18}$$

where $\hat{r}_i^{e+1}$ denotes the observed signal of U_i at the $(e+1)$-th measurement, and $c_\varphi^{(e+1)}$ is a scalar constant that guarantees the elements of $\varphi_{U_i}^{(e+1)}$ add up to one.

After e channel measurements, U_i can predict the transmitting beamforming matrix F_{e+1} selected by S at the $(e+1)$-th channel measurement, and obtain an estimation of channel matrix $\hat{H}_i^{(e)}$ based on former e channel measurements. Hence, we can calculate $\hat{r}_i^{e+1}$ by

$$\hat{r}_i^{e+1} = \sqrt{\frac{P_t}{R_S}} \hat{H}_i^{(e)} F_{e+1} s. \tag{19}$$

By substituting (19) into (18), and recalling (4), $\varphi_{U_i}^{(e+1)}$ can be represented as

$$\varphi_{U_i}^{(e+1)} = c_\varphi^{(e+1)} \frac{P_t}{R_S} \operatorname{diag}\left(W_{d_i}^{\mathrm{H}} \hat{H}_i^{(e)} F_{e+1} s (\hat{H}_i^{(e)} F_{e+1} s)^{\mathrm{H}} W_{d_i} \right) = c_\varphi^{(e+1)} A_c^2 \operatorname{diag}\left[\hat{H}_{\nu_i}^{(e)} \Gamma_{e+1} (\hat{H}_{\nu_i}^{(e)})^{\mathrm{H}} \right], \tag{20}$$

where $\Gamma_{e+1} = F_d^{\mathrm{H}} F_{e+1} s s^{\mathrm{H}} F_{e+1}^{\mathrm{H}} F_d$ is a sparse diagonal matrix, which only has R_S nonzero entries.

Therefore, we update φ_{U_i} in each measurement as (20), and the rest part of AIA algorithm is the same as PRE algorithm.

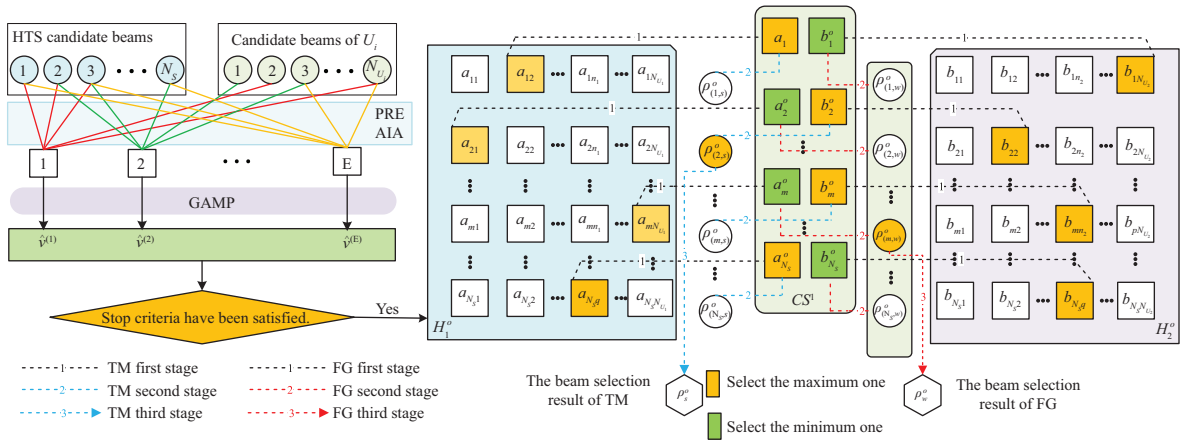

Figure 2 (Color online) Joint channel estimation and beam selection NOMA system.

4 Beam selection for mmWave NOMA S-IoT system

In this section, we first propose a UE grouping scheme to group UE into multiple two-UE groups according to their estimated channel gains for the subsequent beam selection transmission. Then, two beam selection algorithms are proposed to further improve the system throughput and UE fairness. The joint channel estimation and beam selection schemes are shown in Figure 2.

4.1 UE grouping scheme

Note that all D UE in the coverage of S still can be observed significant differences in their channel gains due to the various path losses on mmWave band signals. Therefore, based on the previous channel estimation in mmWave NOMA S-IoT system, we can obtain the mmWave band channel matrix $\hat{H}_i$ of all D UE and the beam with highest channel gain $\hat{h}_i$ in each $\hat{H}_i$. Then, we sort these D channel gain $\hat{h}_i$ in a descending order, and the corresponding UE set $U^{\dagger}$ can be expressed as

$$U^{\dagger} = \{U_1^{\dagger}, U_2^{\dagger}, \ldots, U_D^{\dagger}\}. \tag{21}$$

Assume that D is even, and we can equally divide $U^{\dagger}$ into a strong UE group and a weak UE group, i.e., $U^s = \{U_1^{\dagger}, U_2^{\dagger}, \ldots, U_{\frac{D}{2}}^{\dagger}\}$ and $U^w = \{U_{\frac{D}{2}+1}^{\dagger}, U_{\frac{D}{2}+2}^{\dagger}, \ldots, U_D^{\dagger}\}$, and sequentially select one UE from the two groups to form $O = \frac{D}{2}$ NOMA groups [24], and the o-th group can be expressed as

$$U^o = \left(U_o^{\dagger}, U_{o+\frac{D}{2}}^{\dagger}\right) = (U_1^o, U_2^o), \quad o = 1, 2, \ldots, O, \tag{22}$$

where U_1^o has a beam with the highest channel gain ρ_s^o in $\hat{H}_1^o$, and ρ_w^o is the highest channel gain of a beam in $\hat{H}_2^o$ of U_2^o, and in the o-th two-UE NOMA group, we have $\rho_s^o > \rho_w^o$.

4.2 Beam selection for NOMA systems

Then, we can carry out the beam selection in each two-UE NOMA group to select out beams with high channel gains at S and U_i for the transmission. In the o-th NOMA group, U_1^o is the strong UE and U_2^o is the weak UE, and their channel matrices are denoted by H_1^o and H_2^o, respectively. Let h_{m,n_1}^o and h_{m,n_2}^o denote the channel gain between the m-th beam of S and the n-th beam of U_1^o and U_2^o, respectively, and we define an indicator ω_m^o for the beam selection as follows:

$$\omega_m^o = \begin{cases} 1, & h_{m,n_1}^o \geqslant h_{m,n_2}^o, \\ 0, & h_{m,n_1}^o < h_{m,n_2}^o. \end{cases} \tag{23}$$

In terms of NOMA principle [25], the data signal that S transmits to the two-UE NOMA group can be defined as [26]

$$\alpha_m^o = (1 - \omega_m^o)\left(\sqrt{c_1}\alpha_1 + \sqrt{c_2}\alpha_2\right) + \omega_m^o\left(\sqrt{c_2}\alpha_1 + \sqrt{c_1}\alpha_2\right), \tag{24}$$

where c_1 and c_2 denote the power allocation coefficients of U_1^o and U_2^o with $c_1 + c_2 = 1$, and α_1 and α_2 denote the data signals that S transmits to U_1^o and U_2^o, respectively.

In the beam selection, the received signals at U_1^o and U_2^o can be expressed as

$$\begin{cases} r_1^o = \sqrt{P_t} h_{m,n_1}^o \alpha_m^o + z_1, \\ r_2^o = \sqrt{P_t} h_{m,n_2}^o \alpha_m^o + z_2, \end{cases} \tag{25}$$

where z_1 and z_2 denote the AWGN of U_1^o and U_2^o, respectively. For simplicity, we assume that their variances σ_{z1}^2 and σ_{z2}^2 satisfy $\sigma_{z1}^2 = \sigma_{z2}^2 = \sigma_z^2$. Following (24) and (25), by utilizing the successive interference cancellation (SIC), we can decode the signals of U_1^o and U_2^o, and the system throughput of U_1^o and U_2^o can be represented as

$$\begin{cases} R_1 = (1 - \omega_m^o) \log_2 \left(1 + \dfrac{c_1 |h_{m,n_1}^o|^2}{c_2 |h_{m,n_1}^o|^2 + \frac{1}{\tau}} \right) + \omega_m^o \log_2 (1 + c_2 |h_{m,n_1}^o|^2 \tau), \\ R_2 = (1 - \omega_m^o) \log_2 (1 + c_2 |h_{m,n_2}^o|^2 \tau) + \omega_m^o \log_2 \left(1 + \dfrac{c_1 |h_{m,n_2}^o|^2}{c_2 |h_{m,n_2}^o|^2 + \frac{1}{\tau}} \right), \end{cases} \tag{26}$$

where $\tau = P_t / \sigma_z^2$ is the SNR. Moreover, we introduce $\rho_{(m,s)}^o$ and $\rho_{(m,w)}^o$ to represent the higher channel gain and lower channel gain of h_{m,n_1}^o and h_{m,n_2}^o, respectively, where $\rho_{(m,s)}^o = \max(h_{m,n_1}^o, h_{m,n_2}^o)$ and $\rho_{(m,w)}^o = \min(h_{m,n_1}^o, h_{m,n_2}^o)$. Thus, we can utilize $\rho_{(m,s)}^o$ and $\rho_{(m,w)}^o$ to replace h_{m,n_1}^o and h_{m,n_2}^o in (26), and the system throughput of the o-th two-UE NOMA group can be calculated as follows:

$$R\left(\rho_{(m,s)}^o, \rho_{(m,w)}^o\right) = R_1 + R_2 = \log_2 \left(1 + \frac{c_1 |\rho_{(m,w)}^o|^2}{c_2 |\rho_{(m,w)}^o|^2 + \frac{1}{\tau}} \right) + \log_2 \left(1 + c_2 |\rho_{(m,s)}^o|^2 \tau \right). \tag{27}$$

To maximize the system throughput of the o-th two-UE NOMA group, we can formulate the beam selection problem as

$$R: \quad \{m^o, n_1^o, n_2^o\} = \underset{m \in F_d, n_1 \in W_{d_1}^o, n_2 \in W_{d_2}^o}{\arg\max} R\left(\rho_{(m,s)}^o, \rho_{(m,w)}^o\right), \tag{28}$$

where F_d is the beamforming matrix of S, $W_{d_1}^o$ and $W_{d_2}^o$ are the beamforming matrices of U_1^o and U_2^o, respectively.

4.3 Beam selection algorithm

In the subsequent beam selection, recall $R(\rho_{(m,s)}^o, \rho_{(m,w)}^o)$ denotes the system throughput of the o-th two-UE NOMA group, which is an increasing function of $\rho_{(m,s)}^o$ and $\rho_{(m,w)}^o$, and we can maximize $R(\rho_{(m,s)}^o, \rho_{(m,w)}^o)$ by maximizing $\rho_{(m,s)}^o$ and $\rho_{(m,w)}^o$. Note that U_1^o has a strongest beam with higher channel gain ρ_s^o than that of ρ_w^o in U_2^o, and U_2^o is the weak UE with its highest channel gain ρ_w^o, where $\rho_s^o > \rho_w^o$. Hence, we propose the TM beam selection algorithm to achieve maximum system throughput by selecting ρ_s^o in priority, and design the FG beam selection algorithm to guarantee better fairness at the cost of slightly loss of system throughput by first dealing with ρ_w^o.

4.3.1 *Throughput maximum beam selection algorithm*

In the TM beam selection algorithm, we maximize $\rho_{(m,s)}^o$ to pursue a high system throughput performance [27, 28]. We summarize the stages of the o-th two-UE NOMA group beam selection as follows.

(1) For each f_m at S, we select the receiving beam with highest channel gain at U_1^o and U_2^o, respectively, and we have

$$\begin{cases} a_m^o = \max\{a_{m,1}^o, a_{m,2}^o, \ldots, a_{m,N_{U_1}}^o\}, \\ b_m^o = \max\{b_{m,1}^o, b_{m,2}^o, \ldots, b_{m,N_{U_2}}^o\}, \end{cases} \tag{29}$$

where a_{m,n_1}^o and b_{m,n_2}^o are the channel gains between the m-th transmitting beam of S and the n-th receiving beam of U_1^o and U_2^o, respectively, and we denote a channel gain set χ^1 for the candidate beams as follows:

$$\chi^1 = \{(a_1^o, b_1^o), (a_2^o, b_2^o), \ldots, (a_{N_S}^o, b_{N_S}^o)\}. \tag{30}$$

(2) Then, we select the beam with higher channel gain from each pair (a_m^o, b_m^o) of candidate beams in χ^1, denoted by $\rho_{(m,s)}^o$, and we have

$$\rho_{(m,s)}^o = \max(a_m^o, b_m^o), \tag{31}$$

and we can denote a strong channel gain set χ_s^2 as follows:

$$\chi_s^2 = \{\rho_{(1,s)}^o, \rho_{(2,s)}^o, \ldots, \rho_{(N_S,s)}^o\}. \tag{32}$$

(3) Further, we can find the highest channel gain ρ_s^o in χ_s^2 as follows:

$$\rho_s^o = \max(\chi_s^2). \tag{33}$$

Since U_1^o is the strong UE, we can find ρ_s^o is in H_1^o; then we can determine the location of ρ_s^o in H_1^o and denote the row index and column index of ρ_s^o as m_{TM}^o and $n_{\mathrm{TM},1}^o$, respectively. Moreover, we need to find the other candidate beam which is in pair with ρ_s^o in (30), and determine its column index $n_{\mathrm{TM},2}^o$ in H_2^o. Finally, we select the m_{TM}^o-th beam of S, the $n_{\mathrm{TM},1}^o$-th beam of U_1^o, and the $n_{\mathrm{TM},2}^o$-th beam of U_2^o for the o-th NOMA group in the TM beam selection algorithm.

4.3.2 *Fairness guaranteed beam selection algorithm*

In the FG beam selection algorithm, we introduce the widely utilized Jain's fairness index to measure the UE fairness performance of our joint channel estimation and beam selection NOMA S-IoT system. The Jain's fairness index for a two-UE NOMA group can be defined as

$$J = \frac{(R_1 + R_2)^2}{2\,(R_1^2 + R_2^2)}. \tag{34}$$

Note that the Jain's fairness index is a real number between 0 and 1, i.e., $J \in [0, 1]$, which reaches its maximum value $J = 1$ when the system throughput of U_1^o and U_2^o is equal, i.e., $R_1 = R_2$.

Moreover, note that in the TM beam selection algorithm, if we first select $\rho_s^o = a_{m,n}^o$ in U_1^o and then find the corresponding beam in its pair as $b_{m,n}^o$ in U_2^o as in (30), which can achieve the maximum system throughput as the ES scheme. However, $b_{m,n}^o$ may be much lower than $a_{m,n}^o$, which would lead a serious unequal throughput between R_1 and R_2 and result in a fairness problem. Hence, we enhance the UE fairness performance by selecting ρ_w^o in priority in our FG beam selection algorithm, which can achieve better fairness at the cost of slightly loss of system throughput.

(1) The first step of the FG beam selection algorithm is to find χ^1, which is the same as in the TM beam selection algorithm.

(2) We select the beam with lower channel gain from each pair (a_m^o, b_m^o) of candidate beams in χ^1, denoted by $\rho_{(m,w)}^o$, and we have

$$\rho_{(m,w)}^o = \min(a_m^o, b_m^o). \tag{35}$$

Then, we construct the weak channel gain set χ_w^2 as follows:

$$\chi_w^2 = \{\rho_{(1,w)}^o, \rho_{(2,w)}^o, \ldots, \rho_{(N_S,w)}^o\}. \tag{36}$$

(3) Then, we can determine the highest channel gain ρ_w^o of U_2^o by finding the maximum channel gain of weak channel gain set χ_w^2 as follows:

$$\rho_w^o = \max(\chi_w^2). \tag{37}$$

Similar to the TM beam selection algorithm, since U_2^o is the weak UE, we can find that ρ_w^o is in H_2^o. Then, we can ascertain the location of ρ_w^o in H_2^o and denote the row index and column index of ρ_w^o as m_{FG}^o and $n_{\mathrm{FG},2}^o$, respectively. Further, we have to find the corresponding candidate beam, which is in pair with ρ_w^o in (30), and determine its column index $n_{\mathrm{FG},1}^o$ in H_1^o. Finally, we select the m_{FG}^o-th beam of S, $n_{\mathrm{FG},1}^o$-th beam of U_1^o, and $n_{\mathrm{FG},2}^o$-th beam of U_2^o for the o-th NOMA group in FG beam selection algorithm.

Therefore, we have four joint channel estimation and beam selection schemes: the PRE-TM and AIA-TM schemes, which can achieve the maximum throughput as the ES benchmark scheme; the PRE-FG and AIA-FG schemes, which have better fairness at the cost of slightly loss of throughput compared with the PRE-TM and AIA-TM schemes. In addition, as shown in Algorithm 1, we illustrate the detailed process of these four schemes.

Algorithm 1 Joint channel estimation and beam selection schemes

Input: N_S, R_S, N_{U_i}, R_{U_i}, where $i = 1, 2, \ldots, D$; the candidate beamforming matrices F_d and W_{d_i}.

Output: The channel matrix $\hat{H}_i$ between S and U_i, the number of NOMA groups $O = D/2$, and the index of selected beam in the NOMA beam selections: $\{m^o_{\mathrm{FG}}, n^o_{\mathrm{FG},1}, n^o_{\mathrm{FG},2}\}$, where $o = 1, 2, \ldots, O$.

1: **for** $i = 1, 2, \ldots, D$, **do**
2: **for** $e = 1$; $m \leqslant E$; $e++$ **do**
3: **if** we adopt the PRE channel estimation algorithm **then**
4: $F_d \xrightarrow{\zeta^e_S} F_e$, $W_{d_i} \xrightarrow{\gamma^e_{U_i}} W_{e_i}$; //Updates ζ^e_S and $\gamma^e_{U_i}$ according to (24) and (25), respectively;
5: **else**
6: **if** we adopt the AIA channel estimation algorithm **then**
7: $F_d \xrightarrow{\zeta^e_S} F_e$, $W_{d_i} \xrightarrow{\varphi^e_{U_i}} W_{e_i}$; //Updates ζ^e_S and $\varphi^e_{U_i}$ according to (24) and (20), respectively;
8: **end if**
9: **end if**
10: $A_c = \sqrt{\frac{N_S N_{U_i} P_t}{R_S}}$, $A^e_i = ((F^{\mathrm{H}}_d F_e s)^{\mathrm{T}} \otimes (W_{e_i})^{\mathrm{H}} W_{d_i})$, v_i, $\varepsilon^e_i = W^{\mathrm{H}}_{e_i} u^e_i \sim \mathcal{CN}(0, N_0 I_{R_{U_i}})$.
11: **Input** ϱ_{in}: v_i, A_c, $A_{(e,i)}$, ε^e_i;
12: **Perform** the GAMP from [23];
13: **Output** the eth measurement ϱ_{out}: $\hat{v}_{(e,i)}$.
14: Convert $\hat{v}_{(e,i)}$ into a binarization form $\bar{v}_{(e,i)}$.
15: **if** $\bar{v}_{(e,i)} = \bar{v}_{(e-1,i)}$; //Check the stop criteria. **then**
16: Break;
17: **else**
18: $e = e + 1$; return to 2;
19: **end if**
20: **end for**
21: **end for**
22: $\hat{v}_{(e,i)} \to \hat{H}_{\nu_i} \to \hat{H}_i = \sqrt{N_S N_{U_i}} W_d \hat{H}_{\nu_i} {F_d}^{\mathrm{H}}$. //Initialization of the subsequent beam selection.
23: $U^\dagger = \{U^\dagger_1, U^\dagger_2, \ldots, U^\dagger_D\}$, $U^s = \{U^\dagger_1, U^\dagger_2, \ldots, U^\dagger_{\frac{D}{2}}\}$, $U^w = \{U^\dagger_{\frac{D}{2}+1}, U^\dagger_{\frac{D}{2}+2}, \ldots, U^\dagger_D\}$;
24: $U^o = (U^\dagger_o, U^\dagger_{o+\frac{D}{2}}) = (U^o_1, U^o_2), o = 1, 2, \ldots, O$. //Construct $O = D/2$ two-UE NOMA groups.
25: **for** $m = 1$; $m \leqslant N_S$; $m++$; //Perform the beam selection in each two-UE NOMA group. **do**
26: $a^o_m = \max(a^o_{m,1}, a^o_{m,2}, \ldots, a^o_{m,N_{U_1}}), b^o_m = \max(b^o_{m,1}, b^o_{m,2}, \ldots, b^o_{m,N_{U_2}}), m \in \{1, 2, \ldots, N_S\}$;
27: $\rho^o_{(m,w)} = \min(a^o_m, b^o_m)$, $\rho^o_{(m,s)} = \max(a^o_m, b^o_m)$;
28: **end for**
29: **If** perform the TM algorithm: $\rho^s_m = \max\{\rho^o_{(1,s)}, \rho^o_{(2,s)}, \ldots, \rho^o_{(N_S,s)}\}$; find the location of ρ^s_m;
30: **If** perform the FG algorithm: $\rho^w_m = \max\{\rho^o_{(1,w)}, \rho^o_{(2,w)}, \ldots, \rho^o_{(N_S,w)}\}$; find the location of ρ^w_m.
31: **Return** $\hat{H}_i$, $\{m^o_{\mathrm{TM}}, n^o_{\mathrm{TM},1}, n^o_{\mathrm{TM},2}\}$, $\{m^o_{\mathrm{FG}}, n^o_{\mathrm{FG},1}, n^o_{\mathrm{FG},2}\}$.

4.4 Computational complexities comparisons

Since S has limited computation resources, we analyze the computational complexity for each two-UE NOMA group in our proposed joint channel estimation and beam selection schemes. For notation simplicity, we omit the index of two-UE NOMA group in the following description. In the ES scheme, the total computational complexity of spanning over all possible beam combinations in NOMA system is $\mathcal{O}(N_S N_{U_1} N_{U_2})$. In contrast, in our joint channel estimation and beam selection schemes, there are two levels of "for" loop as shown in Algorithm 1 as follows.

(1) In the first "for" loop from line 2 to line 20, we span all the beam combinations by updating ζ_S and γ_{U_i}, introducing a computational complexity of $\mathcal{O}(\lceil \frac{N_S}{R_S} \cdot (\frac{N_{U_1}}{R_{U_1}} + \frac{N_{U_2}}{R_{U_2}}) \rceil)$.

(2) In the second "for" loop from line 25 to line 28, we successively select the receiving beam of U_1 and U_2 for each transmitting beam at S, introducing a complexity of $\mathcal{O}(N_S N_{U_1} + N_{U_2})$.

Therefore, the total computational complexity of the joint channel estimation and beam selection schemes is $\mathcal{O}(N_S N_U)$, where $N_U = \max\{\min(N_{U_1}, N_{U_2}), \frac{1}{R_S} \cdot (\frac{N_{U_1}}{R_{U_1}} + \frac{N_{U_2}}{R_{U_2}})\}$.

5 Simulation results

In this section, we simulate and analyze the results for the system performance of above schemes, including the number of channel measurements, system throughput, measurement efficiency and UE fairness, and compare with the existing benchmark schemes including ES scheme, PNM-TM scheme and PNM-FG scheme. We summarize the important simulation parameters in Table 1.

First, we compare the number of measurements and error probability of PRE and AIA algorithms with PNM and ES algorithms over different SNR values in Figure 3. In Figure 3(a), note that as the SNR increases, both PRE and AIA algorithms can adaptively reduce their channel measurements, while

Table 1 Simulation setup

Parameter	Value
Number of beams in S	$N_S = 64$
Number of beams in U_1 (U_2)	N_{U_1} (N_{U_2}) $= 16$
Number of RF chains in S	$R_S = 8$
Number of RF chains in U_1 (U_2)	R_{U_1} (R_{U_2}) $= 4$
Communication distance from S to U_1 (U_2)	1500 km
Expected number of paths	$\mathrm{E}[K] = 3$
Stopping criterion coefficient	$\Lambda = 0.1$

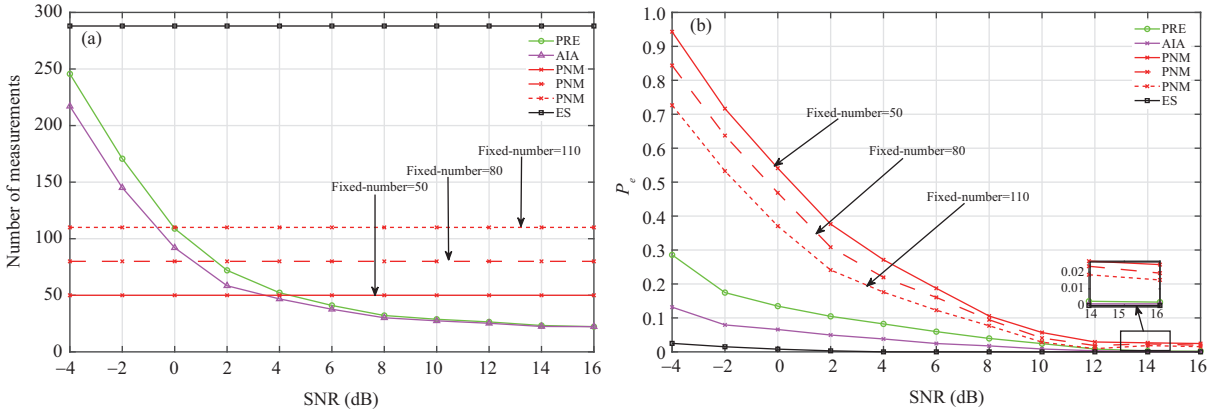

Figure 3 (Color online) The average number of measurements and error probability of our proposed algorithms comparing with the existing benchmark algorithms. (a) Average number of measurements versus SNR; (b) error probability versus SNR.

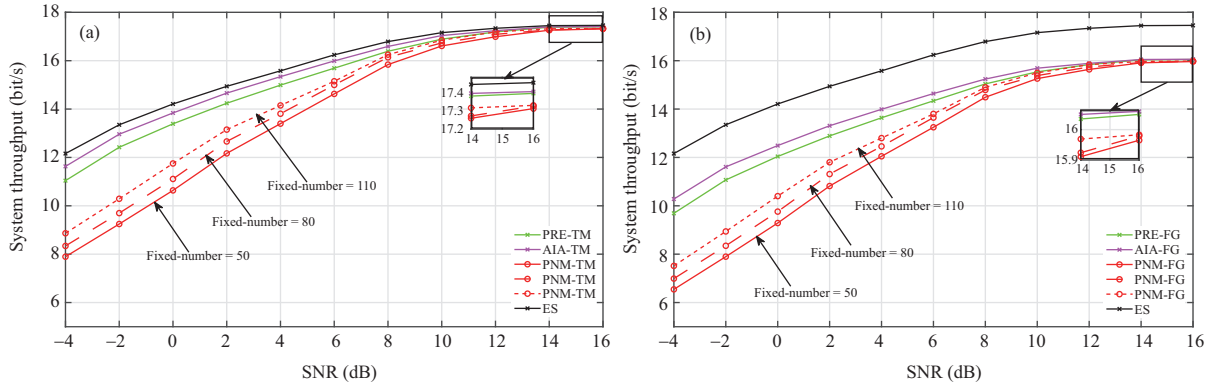

Figure 4 (Color online) The system throughput of our proposed schemes comparing with the existing benchmark schemes versus SNR. (a) System throughput of AIA-TM and PRE-TM schemes versus SNR; (b) system throughput of AIA-FG and PRE-FG schemes versus SNR.

the PNM and ES algorithms are fixed in the entire SNR region. For example, when SNR $= 4$ dB, the number of channel measurements in PRE and AIA algorithms can be reduced by 78.7% and 78.5% than that of SNR $= -4$ dB, respectively. Moreover, it is worth noting that the AIA algorithm requires fewer measurements than the PRE algorithm when SNR < 14 dB, and the number of measurements in these two algorithms can converge to 22 when the SNR $\geqslant 14$ dB. Then, we analyze the error probability in Figure 3(b). In addition, although the measurements in PRE and AIA algorithms are more than that in PNM algorithm when SNR $\leqslant 4$ dB, the error probability P_e of PRE and AIA algorithms is much lower than that of PNM algorithm. This is because the PRE and AIA algorithms can obtain accurate CSI within the adaptable channel measurements, while the PNM algorithm cannot get accurate CSI because the number of measurements in PNM algorithm would be not enough in the low SNR region. In addition, P_e of our proposed algorithms converges to the ES scheme when SNR $\geqslant 14$ dB.

Then, we compare the system throughput of PRE-TM and AIA-TM schemes with the benchmark schemes PNM-TM [17] and ES in Figure 4(a). Note that the AIA-TM scheme approaches the system

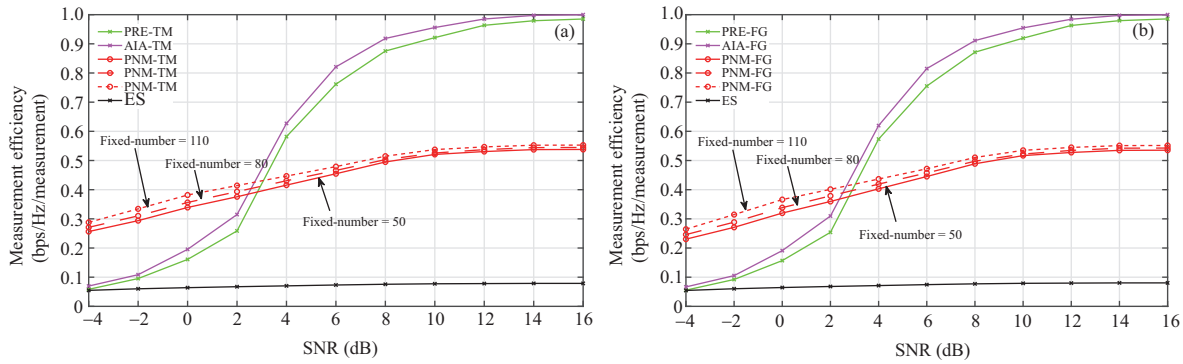

Figure 5 (Color online) The measurement efficiency performance of our proposed schemes comparing with the existing benchmark schemes versus SNR. (a) Measurement efficiency of AIA-TM and PRE-TM schemes versus SNR; (b) measurement efficiency of AIA-FG and PRE-FG schemes versus SNR.

throughput of ES scheme in the whole simulated SNR region, while the PRE-TM scheme is relatively lower about 0.4 bit/s than the AIA-TM scheme. Moreover, when the SNR is lower than 8 dB, the throughput of PNM-TM schemes are much lower than the other three schemes, since their number of measurements is not enough for obtaining accurate CSI. Besides, all the schemes converge to a similar throughput when the SNR $\geqslant$ 14 dB. Similarly, we compare the system throughput of PRE-FG and AIA-FG schemes with PNM-FG and ES schemes in Figure 4(b). It is clear that the system throughput of PRE-FG and AIA-FG schemes is lower than that of the ES scheme about 2 bit/s to pursue a better UE fairness performance. When SNR $\geqslant$ 14 dB, the PRE-TM, AIA-TM, PNM-TM and ES schemes converge to 17.45 bit/s, while that of PRE-FG, AIA-FG and PNM-FG schemes converges to 16.05 bit/s.

It is worth noting that the system throughput of our proposed PRE-TM and AIA-TM schemes approaches that of the ES scheme in the simulated SNR region, while the number of channel measurements of the proposed schemes is much lower than that of the ES scheme, especially in the high SNR region. Therefore, we define the measurement efficiency η as the ratio of system throughput to number of channel measurements in a comprehensive way, which can be viewed as the number of bits that could be successfully transmitted per channel measurement, and we have

$$\eta = \frac{\text{system throughput}}{\text{number of channel measurements}} = \frac{R}{e} \text{ bit/s/measurement}. \tag{38}$$

Then, we compare η of PRE-TM/FG, AIA-TM/FG, PNM-TM/FG and ES schemes in Figure 5. Obviously, although the ES scheme can achieve the highest system performance as shown in Figure 4, the corresponding measurement efficiency is poor due to the throughput that is obtained at the cost of a large number of channel measurements. In contrast, the throughput of our PRE-TM and AIA-TM schemes is slightly lower than that of ES scheme as shown in Figure 4(a), while their measurement efficiency is significantly increasing with the SNR, as their channel measurements are decreasing as shown in Figure 5(a). Moreover, when SNR $\geqslant$ 4 dB, the measurement efficiency of PRE-TM/FG and AIA-TM/FG schemes outperforms that of the PNM-TM/FG schemes as shown in Figure 4, which validates that our proposed schemes can further improve the measurement efficiency via adaptively reducing the required measurements in the high SNR region.

In Figure 6, we investigate the impact of power allocation coefficient c_2 on the system throughput. It shows that the optimal channel coefficient for achieving the maximum system throughput is $c_2 = 0.65$, while the system throughput with other channel coefficient values is slightly lower with a difference less than 0.5 bit/s.

Finally, we compare the UE fairness performance of our four schemes to the PNM-TM and PNM-FG schemes and ES scheme in Figure 7. It can be observed that the PRE-FG, AIA-FG and PNM-FG achieve better UE fairness performance compared with the PRE-TM, AIA-TM, PNM-TM and ES schemes, which validates that our FG beam selection algorithm can significantly improve the UE fairness performance. Moreover, we can find that when the power coefficient $c_2 = 0.5$, the Jain's fairness index of PRE-FG, AIA-FG and PNM-FG schemes is about 0.676, while that of the PRE-TM, AIA-TM, PNM-TM schemes and ES scheme is about 0.572. This is because the FG beam selection algorithm enhances the UE fairness by selecting the channel gain of the weak UE in priority, which narrows the difference between R_1 and R_2.

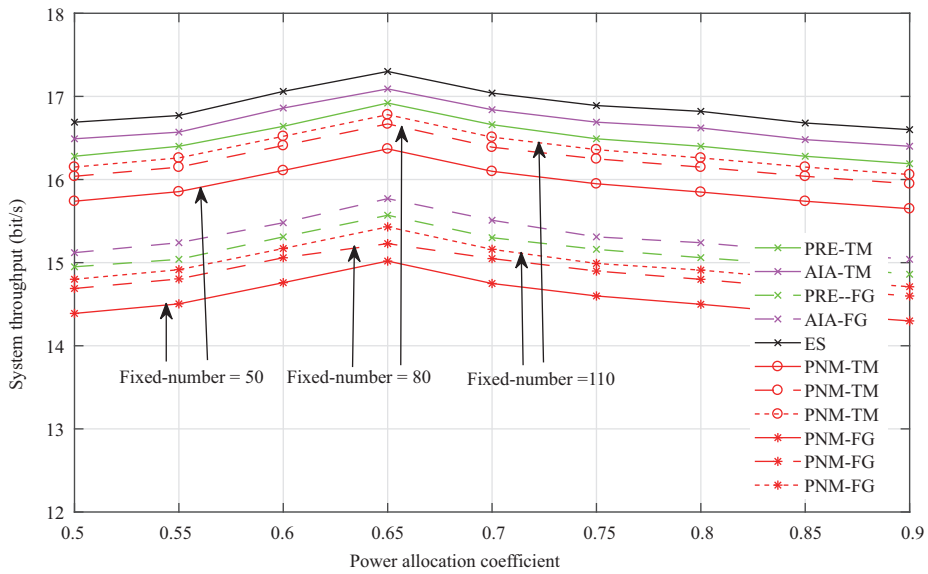

Figure 6 (Color online) The system throughput of different joint channel estimation and beam selection schemes and benchmark schemes versus power coefficient c_2.

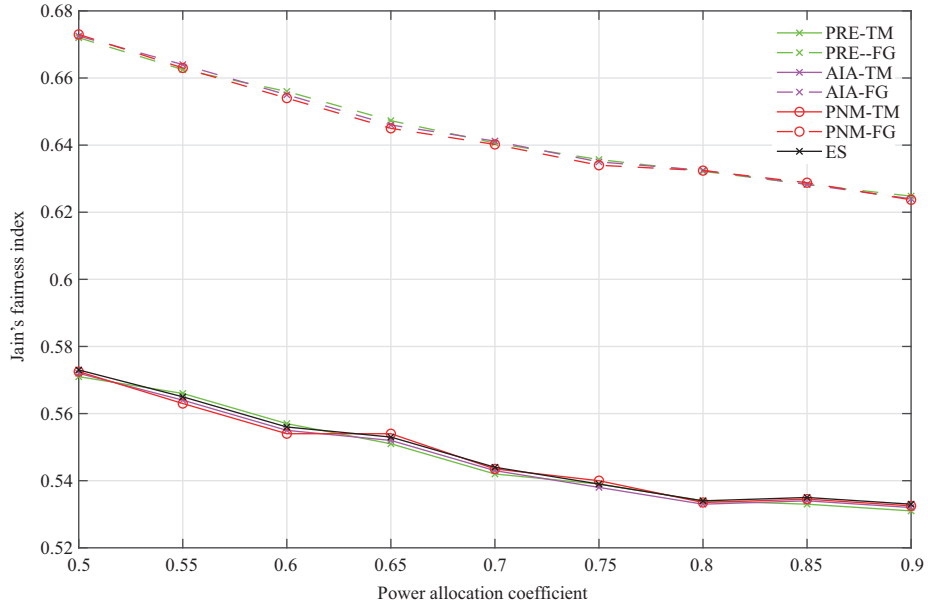

Figure 7 (Color online) The Jain's fairness of different joint channel estimation and beam selection schemes and benchmark schemes versus power coefficient c_2.

6 Conclusion

In this paper, we design a joint channel estimation and beam selection NOMA system for S-IoT to achieve better system performance. Furthermore, we propose a UE grouping scheme to group UE into multiple two-UE groups according to their estimated channel gains. To be specific, we propose two channel estimation algorithms and two beam selection algorithms for different QoS requirements, then we can combine them into four joint channel estimation and beam selection schemes: the PRE-TM and AIA-TM schemes which can achieve the maximum throughput with a significantly lower complexity compared with the ES scheme, the PRE-FG and AIA-FG schemes which have better fairness at the cost of slightly loss of the throughput compared with the PRE-TM and AIA-TM schemes. Moreover, the simulation results also validate that these four schemes can adaptively reduce the channel measurements by about 78.5% compared with that of SNR $= -4$ dB, while the ES, PNM-TM and PNM-FG schemes require fixed

number of measurements.

Acknowledgements This work was supported in part by National Natural Sciences Foundation of China (Grant Nos. 62071141, 61871147, 61831008, 62027802), Shenzhen Basic Research Program (Grant No. GXWD20201230155427003-20200822-165138001), Natural Science Foundation of Guangdong Province (Grant No. 2020A1515010505), Guangdong Science and Technology Planning Project (Grant No. 2018B030322004), and the project "The Verification Platform of Multi-tier Coverage Communication Network for Oceans" (Grant No. LZC0020).

References

1 You X, Wang C X, Huang J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. Sci China Inf Sci, 2021, 64: 110301
2 Yao H, Wang L, Wang X, et al. The space-terrestrial integrated network: an overview. IEEE Commun Mag, 2018, 56: 178–185
3 You L, Li K X, Wang J, et al. Massive MIMO transmission for LEO satellite communications. IEEE J Sel Areas Commun, 2020, 38: 1851–1865
4 Goto D, Shibayama H, Yamashita F, et al. LEO-MIMO satellite systems for high capacity transmission. In: Proceedings of IEEE Global Communications Conference (GLOBECOM), 2018. 1–6
5 Jiao J, Liao S Y, Sun Y Y, et al. Fairness-improved and QoS-guaranteed resource allocation for NOMA-based S-IoT network. Sci China Inf Sci, 2021, 64: 169306
6 Liu S, Lin J, Xu L, et al. A dynamic beam shut off algorithm for LEO multibeam satellite constellation network. IEEE Wireless Commun Lett, 2020, 9: 1730–1733
7 Hussein H S, Hussein S, Mohamed E M. Efficient channel estimation techniques for MIMO systems with 1-bit ADC. China Commun, 2020, 17: 50–64
8 Ding L, Wu S, Ding X, et al. Off-the-grid sparse imaging by one-dimensional sparse MIMO array. IEEE Sens J, 2018, 18: 9993–10001
9 Zhang D, Li A, Chen H, et al. Beam allocation for millimeter-wave MIMO tracking systems. IEEE Trans Veh Technol, 2020, 69: 1595–1611
10 Jiao J, He Y, Wu S, et al. Intelligent hybrid non-orthogonal multiple access relaying for vehicular networks in 6G. IEEE Internet Things J, 2021, 8: 14773–14786
11 Kokshoorn M, Chen H, Li Y, et al. Beam-on-graph: simultaneous channel estimation for mmWave MIMO systems with multiple users. IEEE Trans Commun, 2018, 66: 2931–2946
12 Jiao J, He Y, Wang Y, et al. Design and analysis of novel Ka band NOMA uplink relay system for Lunar farside exploration. China Commun, 2020, 17: 1–14
13 Yang K, Yang N, Ye N, et al. Non-orthogonal multiple access: achieving sustainable future radio access. IEEE Commun Mag, 2019, 57: 116–121
14 Jiao J, Sun Y, Wu S, et al. Network utility maximization resource allocation for NOMA in satellite-based Internet of Things. IEEE Internet Things J, 2020, 7: 3230–3242
15 Akdeniz M R, Liu Y, Samimi M K, et al. Millimeter wave channel modeling and cellular capacity evaluation. IEEE J Sel Areas Commun, 2014, 32: 1164–1179
16 Shokrollahi A. Raptor codes. IEEE Trans Inform Theor, 2006, 52: 2551–2567
17 Huang J, Wang C X, Chang H, et al. Multi-frequency multi-scenario millimeter wave MIMO channel measurements and modeling for B5G wireless communication systems. IEEE J Sel Areas Commun, 2020, 38: 2010–2025
18 Foust J. SpaceX's space-Internet woes: despite technical glitches, the company plans to launch the first of nearly 12000 satellites in 2019. IEEE Spectr, 2019, 56: 50–51
19 Alkhateeb A, El Ayach O, Leus G, et al. Channel estimation and hybrid precoding for millimeter wave cellular systems. IEEE J Sel Top Signal Process, 2014, 8: 831–846
20 Wang K, Wei H, He W, et al. Dynamic measurement for compressed sensing based channel estimation in OFDM systems. In: Proceedings of the 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, 2017. 106–110
21 Huang S, Tran T D. Sparse signal recovery using generalized approximate message passing with built-in parameter estimation. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017. 4321–4325
22 Rangan S. Generalized approximate message passing for estimation with random linear mixing. In: Proceedings of IEEE International Symposium on Information Theory Proceedings, 2011. 2168–2172
23 Vila J, Schniter P. Expectation-maximization bernoulli-Gaussian approximate message passing. In: Proceedings of 2011 Conference Record of the 45th Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2011. 799–803
24 Ding Z, Fan P, Poor H V. Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions. IEEE Trans Veh Technol, 2016, 65: 6010–6023
25 Pang X W, Tang J, Zhao N, et al. Energy-efficient design for mmWave-enabled NOMA-UAV networks. Sci China Inf Sci, 2021, 64: 140303
26 Zhao N, Li Y, Zhang S, et al. Security enhancement for NOMA-UAV networks. IEEE Trans Veh Technol, 2020, 69: 3994–4005
27 Ye N, Li X, Yu H, et al. Deep learning aided grant-free NOMA toward reliable low-latency access in tactile Internet of Things. IEEE Trans Ind Inf, 2019, 15: 2995–3005
28 Ye N, Li X, Yu H, et al. DeepNOMA: a unified framework for NOMA using deep multi-task learning. IEEE Trans Wireless Commun, 2020, 19: 2208–2225