This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

Ergodic Sum Rate Capacity Achieving Transmit Design for Massive MIMO LEO Satellite Uplink Transmission

Ke-Xin Li, Member, IEEE
National Mobile Communications Research Laboratory (NCRL), Southeast University, Nanjing, China and Purple Mountain Laboratories (PML), Nanjing, China

Xiqi Gao, Fellow, IEEE
National Mobile Communications Research Laboratory (NCRL), Southeast University, Nanjing, China and Purple Mountain Laboratories (PML), Nanjing, China

Xiang-Gen Xia, Fellow, IEEE
Department of Electrical and Computer Engineering, University of Delaware, Newark, USA

***Abstract*— In this paper, we investigate the ergodic sum rate (ESR) capacity achieving uplink (UL) transmit design for massive multiple-input multiple-output (MIMO) low-earth-orbit (LEO) satellite communications with statistical channel state information at the user terminals (UTs). The UL massive MIMO LEO satellite channel model with uniform planar array configurations at the satellite and UTs is presented. We prove that the rank of each UT's optimal transmit covariance matrix does not exceed that of its channel correlation matrix at the UT side, which reveals the maximum number of independent data streams transmitted from each UT to the satellite. We then prove that the transmit covariance matrix design can be transformed into the lower-dimensional matrix design without loss of optimality. We also obtain a necessary and sufficient condition when single data stream transmission from each UT to the satellite can achieve the ESR capacity. A conditional gradient (CG) method is developed to compute the ESR capacity achieving transmit covariance matrices. Furthermore, to avoid the exhaustive sample average, we utilize an asymptotic expression of the ESR and devise a simplified CG method to compute the transmit covariance matrices, which can approximate the ESR capacity. Simulations demonstrate the effectiveness of the proposed approaches.**

***Index Terms*— LEO satellites, massive MIMO, ergodic sum rate capacity, satellite uplink transmission.**

Ke-Xin Li and Xiqi Gao are with the National Mobile Communications Research Laboratory (NCRL), Southeast University, Nanjing 210096, and are also with the Purple Mountain Laboratories (PML), Nanjing 211111, China (e-mail: likexin3488@seu.edu.cn, xqgao@seu.edu.cn). Xiang-Gen Xia is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: xianggen@udel.edu). *(Corresponding author: Xiqi Gao.)*

0018-9251 © 2023 IEEE

I. Introduction

To serve vast areas with insufficient terrestrial network coverage, satellite communications (SATCOM) will be an indispensable part of next generation wireless network [1], and have been extensively investigated in the non-terrestrial networks (NTN) of 5G new radio (NR) [2]. In particular, low earth orbit (LEO) satellites, deployed between 200 km to 2000 km altitudes, have been recognized as a promising infrastructure to provide low-latency and ubiquitous broadband services for the user terminals (UTs) which have no access to the terrestrial network [3], [4]. Nowadays, commercial plans have been proposed to construct the mega-constellation by launching a huge number of LEO satellites into the space, e.g., Starlink [5]. LEO SATCOM has become a hotspot of research in both academia and industry.

In the uplink (UL) SATCOM, multiple UTs on ground send messages to the satellite simultaneously. The conventional UL multibeam SATCOM with single antenna configuration at the UT sides has been investigated in the literature, most of which focus on the capacity analysis, e.g., [6]–[8]. The lower and upper bounds on the ergodic sum rate (ESR) capacity of the UL multibeam SATCOM were derived in [6] and [7], respectively. In [8], the ergodic capacity and outage capacity were analyzed by considering the spatially correlated rain attenuation in the UL multibeam satellite channels. In addition, the user scheduling combined with modulation and coding scheme selection for the UL of multibeam satellites with two-color reuse was investigated in [9]. The performance of UL multibeam SATCOM with multiple terrestrial relays was analyzed in [10], by taking account of the hardware impairments.

More flexible payloads have always been the pursuit of advanced SATCOM systems. In the last decades, massive multiple-input multiple-output (MIMO) has made great success in the terrestrial 5G system, and will continue to play an important role in the future 6G system [11]. Recently, a massive MIMO transmission framework for LEO SATCOM was proposed in [12], in which the channel model, multi-user precoder and detector, and the user grouping strategy were investigated. The multi-user precoder and detector in [12] are based on the statistical channel state information (sCSI) instead of the instantaneous CSI (iCSI). The reason is that iCSI is usually challenging to be acquired at the transmitter because of the inherent LEO satellite channel impairments, e.g., large propagation delays and Doppler effects. For time division duplexing (TDD) mode, the estimated downlink (DL) iCSI is directly used for UL transmission, which may be outdated when the UL signals arrive at the satellite. For frequency division duplex (FDD) mode, the UL iCSI is first estimated at the satellite and then fed back to the UTs, which can bring considerable feedback overhead. In addition, the feedback can be outdated because of the large propagation delays in SATCOM. Compared with the rapidly changing iCSI, sCSI can keep stable for longer

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

time intervals, which makes it easier to be acquired by the UTs in UL LEO SATCOM. In addition, thanks to the statistical reciprocity between the DL channel and UL channel, each UT can acquire its own sCSI based on the DL pilot signals sent by the satellite. Until now, there have been several works focusing on the DL transmission in massive MIMO LEO SATCOM, e.g., see [13]–[15]. The authors in [13] compared the performance of different precoding techniques with the multibeam selection scheme, and devised a resource allocation approach to further improve the throughput. In [14], the single data stream transmission for each UT was shown to be optimal in the sense of maximizing the DL ESR, even though each UT is equipped with multiple antennas. The authors in [15] investigated the distributed linear precoder and ground station (GS) equalizer for multi-satellite communications, which only rely on the positional information of the satellites and GS. In addition, a distributed massive MIMO was also introduced into LEO SATCOM, e.g., see [16], [17], which can further enhance the system performance compared with the collocated counterparts.

Unlike the DL transmission, the massive MIMO LEO satellite UL transmission receives much less attention. In [12], an UL multi-user detector at the satellite side was designed with each UT equipped with a single antenna. The authors in [18] studied the joint channel estimation and device activity detection for grant-free random access in massive MIMO LEO SATCOM, by using the orthogonal time-frequency space (OTFS) modulation to combat the large delays and Doppler shifts. In [19], the active terminal identification, channel estimation and multi-user detection were jointly considered for grant-free non-orthogonal multiple access (NOMA) with OTFS modulation in massive MIMO LEO SATCOM systems. However, to our best knowledge, the ESR capacity achieving UL transmit design for massive MIMO LEO satellite systems with multiple antenna configurations at both the satellite and UTs has not been investigated.

In this paper, we study the UL transmit design that uses sCSI at the transmitter (sCSIT) in massive MIMO LEO SATCOM systems, where both the satellite and the UTs are equipped with uniform planar arrays (UPAs). First, we present the UL massive MIMO LEO satellite channel model. The large propagation delays and Doppler shifts are pre-compensated at the UTs, to keep the received signals at the satellite synchronized, thereby supporting the orthogonal frequency division multiplexing (OFDM) based transmission. Then, we investigate the optimal UL transmit design for achieving the ESR capacity, by using the long-term sCSIT. We prove that the rank of each UT's transmit covariance matrix that achieves the UL ESR capacity is no larger than that of its own channel correlation matrix at the UT side, thus revealing the maximum number of independent data streams delivered from each multi-antenna UT to the satellite. Moreover, we prove that each UT's transmit covariance matrix can be represented by a lower-dimensional matrix, so that the transmit covariance matrix design can be transformed into the lower-dimensional matrix design without any loss of optimality. We also obtain a necessary and sufficient condition when single data stream transmission from each UT to the satellite can achieve the ESR capacity. We then develop a conditional gradient (CG) method to compute the ESR capacity achieving transmit covariance matrices. Furthermore, to avoid the complicated sample average, we utilize an asymptotic expression of the ESR, and devise a simplified CG method to compute the transmit covariance matrices, which can attain a near performance to the ESR capacity.

The remainder of this paper is organized as follows. Section II introduces the system model, where the UL channel model is presented for the satellite and the UTs both equipped with UPAs. Section III presents our main results on the UL transmit design, including the rank property of transmit covariance matrices, lower-dimensional matrix representation of transmit covariance matrices, and the CG methods to compute the transmit covariance matrices. Section IV provides the simulation results, and Section V concludes this paper.

Notations: Throughout this paper, lower case letters denote scalars, and boldface lower (upper) letters denote vectors (matrices). The set of all n-by-m complex (real) matrices is denoted as $\mathbb{C}^{n\times m}$ ($\mathbb{R}^{n\times m}$). The trace, determinant, rank, conjugate, transpose, and conjugate transpose of a matrix are represented by $\mathrm{tr}(\cdot)$, $\det(\cdot)$, $\mathrm{rank}(\cdot)$, $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$, respectively. The Euclidean norm of vector $\mathbf{x}$ is denoted as $\|\mathbf{x}\| = \sqrt{\mathbf{x}^H\mathbf{x}}$. The identity matrix is represented by $\mathbf{I}$ or $\mathbf{I}_n$. $\mathbf{1}$ and $\mathbf{0}$ denote all-one and all-zero vectors, respectively. Denote $\otimes$ and $\odot$ as the Kronecker product and Hadamard product, respectively. Let $[\mathbf{A}]_{n,m}$ represent the (n,m)th element of matrix $\mathbf{A}$. $\Upsilon_{\max}(\mathbf{A})$ denotes the maximum eigenvalue of $\mathbf{A}$. The diagonal matrix with $\mathbf{x}$ along its main diagonal is denoted as $\mathrm{diag}(\mathbf{x})$. $\mathbb{E}\{\cdot\}$ means mathematical expectation. $\mathcal{CN}(\mathbf{m},\mathbf{C})$ denotes the proper complex Gaussian random vector with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{C}$. The uniform distribution on interval $[a,b]$ is denoted as $\mathrm{U}(a,b)$.

II. System Model

In this section, we first present the system configuration for UL massive MIMO LEO SATCOM with OFDM modulation. Then, we derive the signal and channel models in the frequency domain of OFDM transmission after performing the Doppler and delay pre-compensation at the UT sides. The statistical properties of satellite channels are also provided.

A. System Configuration

A massive MIMO LEO SATCOM system operating at the lower frequency bands, e.g., L/S/C bands, is considered. As depicted in Fig. 1, the mobile UTs on the ground send messages to the LEO satellite at an altitude of H. The satellite and the mobile UTs are all equipped with UPAs. The UPA at the satellite has M_{x} and M_{y} directional

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

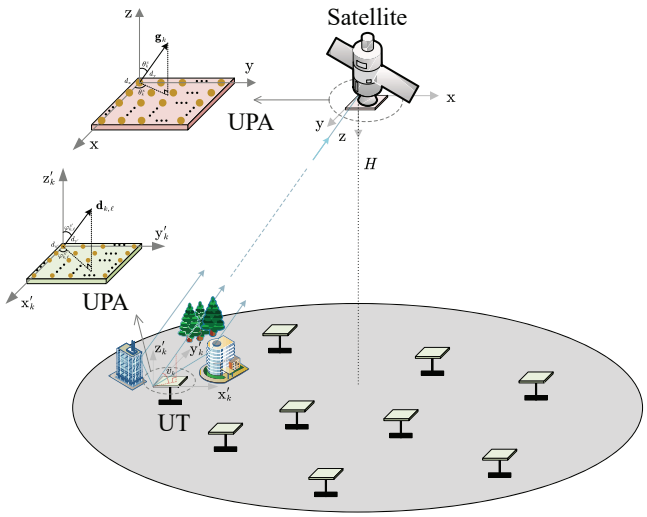

Fig. 1: A massive MIMO LEO SATCOM system.

elements in the x-axis and y-axis, respectively. Thus, the satellite has $M_{\rm x}M_{\rm y} \triangleq M$ antennas. Meanwhile, each UT k uses the UPA consisting of $N_{\rm x'}$ and $N_{\rm y'}$ omnidirectional antenna elements in the $\rm x'_k$-axis and $\rm y'_k$-axis, respectively. Hence, there are $N_{\rm x'}N_{\rm y'} \triangleq N$ antennas at each UT. Note that the antenna configurations can be extended to the case when the UPAs of the UTs have different numbers of antenna elements.

The OFDM modulation is used for wideband transmission in the LEO SATCOM system. The number of subcarriers and cyclic prefix (CP) length are represented by $N_{\rm sc}$ and $N_{\rm cp}$, respectively. The subcarrier spacing is denoted by Δf, and then $T_{\rm s} = 1/(N_{\rm sc}\Delta f)$ is the system sampling period. The time duration of CP is given by $T_{\rm cp} = N_{\rm cp}T_{\rm s}$. The time intervals of one OFDM symbol excluding and including CP are given by $T_{\rm sc} = N_{\rm sc}T_{\rm s}$ and $T = T_{\rm cp} + T_{\rm sc}$, respectively.

B. Signal and Channel Models

We consider the UL massive MIMO transmission. Let $\mathbf{x}_k(t) \in \mathbb{C}^{N\times 1}$ denote the transmit signal of UT k. The received signal at the satellite at time instant t is given by

$$\mathbf{y}(t) = \sum_k \int_{-\infty}^{\infty} \check{\mathbf{H}}_k(t,\tau)\mathbf{x}_k(t-\tau)\,{\rm d}\tau + \mathbf{z}(t), \quad (1)$$

where $\check{\mathbf{H}}_k(t,\tau) \in \mathbb{C}^{M\times N}$ is the channel impulse response of UT k, and $\mathbf{z}(t) \in \mathbb{C}^{M\times 1}$ is the additive noise at the satellite. The time-varying channel impulse response $\check{\mathbf{H}}_k(t,\tau)$ can be written as

$$\check{\mathbf{H}}_k(t,\tau) = \sum_{\ell=0}^{L_k-1} \check{a}_{k,\ell} e^{j2\pi f_{k,\ell}t}\delta(\tau-\tau_{k,\ell})\mathbf{g}_{k,\ell}\mathbf{d}_{k,\ell}^H, \quad (2)$$

where $\delta(x)$ is the Dirac delta function, $j \triangleq \sqrt{-1}$, L_k is the multipath number, $\check{a}_{k,\ell}$, $f_{k,\ell}$ and $\tau_{k,\ell}$ are the complex channel gain, the Doppler shift and the propagation delay for the ℓth path of UT k's channel. In addition, $\mathbf{g}_{k,\ell} \in \mathbb{C}^{M\times 1}$ and $\mathbf{d}_{k,\ell} \in \mathbb{C}^{N\times 1}$ are the array response vectors for the ℓth path of UT k's channel at the satellite and the UT sides, respectively.

In LEO satellite channels, the Doppler shifts $f_{k,\ell}$'s and the propagation delays $\tau_{k,\ell}$'s are much larger than those in terrestrial wireless channels. Hence, the dominant parts in $f_{k,\ell}$'s and $\tau_{k,\ell}$'s need to be well pre-compensated at each UT, so that the received OFDM signals of the UTs can be synchronized at the satellite. The Doppler shift $f_{k,\ell}$ can be separated as $f_{k,\ell} = f_{k,\ell}^{\rm sat} + f_{k,\ell}^{\rm ut}$, where $f_{k,\ell}^{\rm sat}$ and $f_{k,\ell}^{\rm ut}$ are the Doppler shifts caused by the motion of the satellite and UT k, respectively. In addition, the Doppler shifts $f_{k,\ell}^{\rm sat}$, $0 \le \ell \le L_k - 1$, tend to be identical for different paths of UT k's channel [2], [20], [21]. Thus, we can rewrite $f_{k,\ell}^{\rm sat} = f_k^{\rm sat}$, $0 \le \ell \le L_k - 1$. Likewise, the propagation delay $\tau_{k,\ell}$ can also be written as $\tau_{k,\ell} = \tau_k^{\rm sat} + \tau_{k,\ell}^{\rm ut}$, where $\tau_k^{\rm sat}$ is the large propagation delay due to the long distance between the satellite and UT k, and $\tau_{k,\ell}^{\rm ut}$ is the residual propagation delay depending on the scatter distribution around UT k. Notice that $f_k^{\rm sat}$ and $\tau_k^{\rm sat}$ are the dominant parts in $f_{k,\ell}$ and $\tau_{k,\ell}$, respectively, which can be determined by the locations of the satellite and UT k.

Denote $\boldsymbol{\theta}_{k,\ell} = (\theta_{k,\ell}^{\rm x}, \theta_{k,\ell}^{\rm z})$ and $\boldsymbol{\varphi}_{k,\ell} = (\varphi_{k,\ell}^{\rm x'}, \varphi_{k,\ell}^{\rm z'})$ as the paired angles-of-arrival (AoAs) and angles-of-departure (AoDs) related to the ℓth path of UT k's channel. Then, $\mathbf{g}_{k,\ell}$ and $\mathbf{d}_{k,\ell}$ in (2) can be expressed as $\mathbf{g}_{k,\ell} = \mathbf{g}(\boldsymbol{\theta}_{k,\ell})$ and $\mathbf{d}_{k,\ell} = \mathbf{d}(\boldsymbol{\varphi}_{k,\ell})$, respectively. The vectors $\mathbf{g}(\boldsymbol{\theta})$ and $\mathbf{d}(\boldsymbol{\varphi})$ for arbitrary $\boldsymbol{\theta} = (\theta_{\rm x}, \theta_{\rm z})$ and $\boldsymbol{\varphi} = (\varphi_{\rm x'}, \varphi_{\rm z'})$ are defined as $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{a}_{M_{\rm x}}(\sin\theta_{\rm z}\cos\theta_{\rm x}) \otimes \mathbf{a}_{M_{\rm y}}(\sin\theta_{\rm z}\sin\theta_{\rm x})$ and $\mathbf{d}(\boldsymbol{\varphi}) = \mathbf{a}_{N_{\rm x'}}(\sin\varphi_{\rm z'}\cos\varphi_{\rm x'}) \otimes \mathbf{a}_{N_{\rm y'}}(\sin\varphi_{\rm z'}\sin\varphi_{\rm x'})$, respectively. Here, $\mathbf{a}_{n_{\rm v}}(x) = \frac{1}{\sqrt{n_{\rm v}}}\left[1\ e^{-j\frac{2\pi d_{\rm v}}{\lambda}x}\ \cdots\ e^{-j\frac{2\pi d_{\rm v}}{\lambda}(n_{\rm v}-1)x}\right]^T \in \mathbb{C}^{n_{\rm v}\times 1}$, where $n_{\rm v} \in \{M_{\rm x}, M_{\rm y}, N_{\rm x'}, N_{\rm y'}\}$, $\lambda = c/f_c$ is the carrier wavelength, c is the speed of the light, f_c is the carrier frequency, $d_{\rm v} \in \{d_{\rm x}, d_{\rm y}, d_{\rm x'}, d_{\rm y'}\}$ is the spacing between adjacent antennas. Moreover, owing to the high altitude of the satellite, the paired AoAs for different paths of UT k's channel are nearly identical, i.e., $\boldsymbol{\theta}_{k,\ell} = \boldsymbol{\theta}_k$, $0 \le \ell \le L_k - 1$ [2], [22]. In other words, the angular spread of each UT's channel observed at the satellite side is zero [2], [12], [14]. Thus, we can discard the subscript of the path ℓ in $\mathbf{g}_{k,\ell}$ and rewrite it as $\mathbf{g}_{k,\ell} = \mathbf{g}_k = \mathbf{g}(\boldsymbol{\theta}_k)$, where $\boldsymbol{\theta}_k = (\theta_k^{\rm x}, \theta_k^{\rm z})$ is referred to as the physical AoA pair of UT k. The AoAs and AoDs associated with the ℓth path of UT k's satellite channel are illustrated in Fig. 1. Notice that $\theta_k^{\rm z}$ is also known as the nadir angle of UT k [23]. Furthermore, we define $\boldsymbol{\zeta}_k = (\zeta_k^{\rm x}, \zeta_k^{\rm y})$ as the space angle pair of UT k, where $\zeta_k^{\rm x} = \sin\theta_k^{\rm z}\cos\theta_k^{\rm x}$ and $\zeta_k^{\rm y} = \sin\theta_k^{\rm z}\sin\theta_k^{\rm x}$. Since the satellite is far away from the UTs, the space angle pairs $\{\boldsymbol{\zeta}_k\}_{k=1}^K$ vary rather slowly and only depend on the locations of the satellite and the UTs. As soon as the UTs acquire their own location information, e.g., via the global navigation satellite system (GNSS), they can immediately derive the channel parameters $\{f_k^{\rm sat}, \tau_k^{\rm sat}, \boldsymbol{\zeta}_k\}_{k=1}^K$ by utilizing the ephemeris.

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

Let $\{\mathbf{x}_{k,s,r}\}_{r=0}^{N_{\rm sc}-1}$ denote the frequency-domain transmit signal of UT k within the sth OFDM symbol. Then, the time-domain transmit signal is given by [24]

$$\mathbf{x}_{k,s}(t) = \sum_{r=0}^{N_{\rm sc}-1} \mathbf{x}_{k,s,r} e^{j2\pi r\Delta f\cdot t}, \tag{3}$$

where $-T_{\rm cp} \le t - sT < T_{\rm sc}$. Let $f_k^{\rm cps} = f_k^{\rm sat}$ and $\tau_k^{\rm cps} = \tau_k^{\min}$. By resorting to the time and frequency compensation techniques [12], the transmit signal of UT k in the sth OFDM symbol is given by

$$\mathbf{x}_{k,s}^{\rm cps}(t) = \mathbf{x}_{k,s}\left(t+\tau_k^{\rm cps}\right) e^{-j2\pi f_k^{\rm cps}\left(t+\tau_k^{\rm cps}\right)}. \tag{4}$$

The time-domain received signal at the satellite in the sth OFDM symbol can be written as

$$\begin{aligned}\mathbf{y}_s^{\rm cps}(t) &= \sum_k \int_{-\infty}^{\infty} \breve{\mathbf{H}}_k(t,\tau)\mathbf{x}_{k,s}^{\rm cps}(t-\tau)\,{\rm d}\tau + \mathbf{z}_s(t)\\ &= \sum_k \mathbf{g}_k \sum_{\ell=0}^{L_k-1} \breve{a}_{k,\ell} e^{j2\pi f_{k,\ell}^{\rm ut} t}\mathbf{d}_{k,\ell}^H\cdot \mathbf{x}_{k,s}(t-\tau_{k,\ell}^{\rm ut}) + \mathbf{z}_s(t)\\ &= \sum_k \int_{-\infty}^{\infty} \breve{\mathbf{H}}_k(t,\tau)\mathbf{x}_{k,s}(t-\tau)\,{\rm d}\tau + \mathbf{z}_s(t),\end{aligned} \tag{5}$$

where $\breve{a}_{k,\ell} = \breve{a}_{k,\ell} e^{j2\pi f_k^{\rm cps}\tau_{k,\ell}^{\rm ut}}$, and $\mathbf{z}_s(t) \in \mathbb{C}^{M\times 1}$ is the additive noise. In addition, $\breve{\mathbf{H}}_k(t,\tau)$ is the effective channel impulse response of UT k, and it can be written as

$$\breve{\mathbf{H}}_k(t,\tau) = \mathbf{g}_k \cdot \sum_{\ell=0}^{L_k-1} \breve{a}_{k,\ell} e^{j2\pi f_{k,\ell}^{\rm ut} t}\delta(\tau-\tau_{k,\ell}^{\rm ut})\mathbf{d}_{k,\ell}^H. \tag{6}$$

In contrast to the original channel impulse response $\breve{\mathbf{H}}_k(t,\tau)$, the Doppler shifts and propagation delays in the effective channel impulse response $\breve{\mathbf{H}}_k(t,\tau)$ have been mitigated to a large extent. The residual propagation delay and Doppler frequency need to satisfy $\tau_{k,\ell}^{\rm ut} \le T_{\rm cp}$ and $f_{k,\ell}^{\rm ut} \ll \Delta f$, $\forall k,\ell$, respectively, so that the time and frequency at the satellite and the UTs can be assumed as perfectly synchronized.

Let $\mathbf{H}_k(t,f)$ denote the effective channel frequency response of UT k, which can be given by

$$\begin{aligned}\mathbf{H}_k(t,f) &= \int_{-\infty}^{\infty} \breve{\mathbf{H}}_k(t,\tau) e^{-j2\pi f\tau}\,{\rm d}\tau\\ &= \mathbf{g}_k\left(\mathbf{d}_k(t,f)\right)^H,\end{aligned} \tag{7}$$

where $\mathbf{d}_k(t,f) = \sum_{\ell=0}^{L_k-1} \breve{a}_{k,\ell}^* e^{-j2\pi(f_{k,\ell}^{\rm ut}t - f\tau_{k,\ell}^{\rm ut})}\mathbf{d}_{k,\ell} \in \mathbb{C}^{N\times 1}$. Thus, the frequency-domain received signal at the satellite on the rth subcarrier in the sth OFDM symbol can be written as [24]

$$\begin{aligned}\mathbf{y}_{s,r} &= \frac{1}{T_{\rm sc}}\int_{sT}^{sT+T_{\rm sc}} \mathbf{y}_s^{\rm cps}(t) e^{-j2\pi r\Delta f\cdot t}\,{\rm d}t\\ &= \sum_k \mathbf{H}_{k,s,r}\mathbf{x}_{k,s,r} + \mathbf{z}_{s,r},\end{aligned} \tag{8}$$

where $\mathbf{H}_{k,s,r}$ is the channel matrix of UT k, $\mathbf{z}_{s,r}$ is the additive Gaussian noise, both on the rth subcarrier of the sth OFDM symbol. Furthermore, $\mathbf{H}_{k,s,r}$ in (8) can be written as

$$\mathbf{H}_{k,s,r} = \mathbf{H}_k(sT, r\Delta f) = \mathbf{g}_k\mathbf{d}_{k,s,r}^H, \tag{9}$$

where $\mathbf{d}_{k,s,r} = \mathbf{d}_k(sT, r\Delta f)$. Notice that the channel matrices $\mathbf{H}_{k,s,r}$ may vary from OFDM symbol to symbol.

C. Statistical Properties of Satellite Channels

In this subsection, we briefly describe the statistical properties of the massive MIMO LEO satellite UL channels. For convenience, we omit the subscripts of OFDM symbol s and subcarrier r in $\mathbf{H}_{k,s,r} = \mathbf{g}_k\mathbf{d}_{k,s,r}^H$ and denote $\mathbf{H}_k = \mathbf{g}_k\mathbf{d}_k^H$ as the flat fading channel matrix of UT k on a specific subcarrier over an OFDM symbol. Based on the effective physical channel models in Section B, $\mathbf{g}_k$ is a non-random vector, and $\mathbf{d}_k$ can be modeled according to the Rician distribution as

$$\mathbf{d}_k = \sqrt{\frac{\kappa_k\beta_k}{\kappa_k+1}}\mathbf{d}_{k,0} + \sqrt{\frac{\beta_k}{\kappa_k+1}}\tilde{\mathbf{d}}_k, \tag{10}$$

where κ_k is the Rician factor, $\beta_k = \mathbb{E}\left\{{\rm tr}(\mathbf{H}_k\mathbf{H}_k^H)\right\} = \mathbb{E}\left\{\|\mathbf{d}_k\|^2\right\}$ is the average channel power, $\mathbf{d}_{k,0}$ represents the direction of the line-of-sight (LoS) path seen at UT k's side, and $\tilde{\mathbf{d}}_k$ is distributed as $\tilde{\mathbf{d}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_k)$ with ${\rm tr}(\mathbf{\Sigma}_k) = 1$. In addition, we assume that the random vectors $\mathbf{d}_k$'s are independent for different k's. The channel matrix $\mathbf{H}_k$ can be written as

$$\mathbf{H}_k = \mathbf{g}_k\mathbf{d}_k^H = \sqrt{\frac{\kappa_k\beta_k}{\kappa_k+1}}\mathbf{H}_k^{\rm LoS} + \sqrt{\frac{\beta_k}{\kappa_k+1}}\mathbf{H}_k^{\rm NLoS}, \tag{11}$$

where $\mathbf{H}_k^{\rm LoS} = \mathbf{g}_k\mathbf{d}_{k,0}^H$ is the deterministic LoS component, $\mathbf{H}_k^{\rm NLoS} = \mathbf{g}_k\tilde{\mathbf{d}}_k^H$ is the random scattering component. The channel correlation matrices of UT k at the satellite and the UT sides are given by

$$\mathbf{R}_k^{\rm sat} = \mathbb{E}\left\{\mathbf{H}_k\mathbf{H}_k^H\right\} = \beta_k\mathbf{g}_k\mathbf{g}_k^H \tag{12}$$

$$\begin{aligned}\mathbf{R}_k^{\rm ut} &= \mathbb{E}\left\{\mathbf{H}_k^H\mathbf{H}_k\right\}\\ &= \mathbb{E}\left\{\mathbf{d}_k\mathbf{d}_k^H\right\} = \frac{\kappa_k\beta_k}{\kappa_k+1}\mathbf{d}_{k,0}\mathbf{d}_{k,0}^H + \frac{\beta_k}{\kappa_k+1}\mathbf{\Sigma}_k,\end{aligned} \tag{13}$$

respectively. It is worth noting that $\mathbf{R}_k^{\rm sat}$ is rank-one indicating that the arriving signals on different antennas at the satellite are highly correlated. Meanwhile, the rank of $\mathbf{R}_k^{\rm ut}$ is determined by the scatter distribution around UT k. The long-term sCSI consists of the channel parameters in the set $\mathcal{H} = \{\beta_k, \kappa_k, \mathbf{g}_k, \mathbf{d}_{k,0}, \mathbf{\Sigma}_k\}_{\forall k}$. In this work, we consider that the satellite and UTs move within such a range that the channel parameters in $\mathcal{H}$ can be regarded as nearly unchanged. Whenever the satellite or some UT moves out of this range, the channel parameters in $\mathcal{H}$ should be updated accordingly.

III. UL Transmit Design

In this section, with the above established massive MIMO LEO satellite channel model, we investigate the transmit covariance matrix design that achieves the UL

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

ESR capacity. First, we prove that the rank of each UT's transmit covariance matrix should be no larger than that of its channel correlation matrix at the UT side. Then, we prove that each UT's transmit covariance matrix can be represented by a low-dimensional matrix, so that the transmit covariance matrix design can be transformed into the lower-dimensional matrix design without any loss of optimality. We further derive a necessary and sufficient condition when single data stream transmission from each UT to the satellite can achieve the ESR capacity. Afterwards, a CG method is developed to compute the ESR capacity achieving transmit covariance matrices with guaranteed convergence. Further, in order to avoid the complicated sample average, we resort to an asymptotic expression of the ESR, and then devise a simplified CG method to compute the transmit covariance matrices, which can attain a near performance to the ESR capacity.

A. Rank Property of Transmit Covariance Matrices

For convenience, we omit the subscripts of OFDM symbol s and subcarrier r in $\mathbf{x}_{k,s,r}$, and denote $\mathbf{x}_k \in \mathbb{C}^{N\times 1}$ as the transmit signal of UT k on a specific subcarrier over an OFDM symbol. In this paper, we consider a general design of the transmit signals $\mathbf{x}_k$'s, where $\mathbf{x}_k$ is a circularly symmetric complex Gaussian (CSCG) random vector with zero mean and covariance matrix $\mathbf{Q}_k = \mathbb{E}\{\mathbf{x}_k\mathbf{x}_k^H\} \in \mathbb{C}^{N\times N}$. Let us denote the eigenvalue decomposition (EVD) of $\mathbf{Q}_k$ as $\mathbf{Q}_k = \mathbf{V}_{\mathrm{Q},k}\mathbf{\Lambda}_{\mathrm{Q},k}\mathbf{V}_{\mathrm{Q},k}^H$, where the column vectors in $\mathbf{V}_{\mathrm{Q},k} \in \mathbb{C}^{N\times N}$ are the eigenvectors and $\mathbf{\Lambda}_{\mathrm{Q},k} \in \mathbb{R}^{N\times N}$ is a diagonal matrix with the non-negative eigenvalues as the diagonal elements. Then, the transmit signal $\mathbf{x}_k$ can be written as

$$\mathbf{x}_k = \mathbf{V}_{\mathrm{Q},k}\mathbf{\Lambda}_{\mathrm{Q},k}^{1/2}\mathbf{s}_{\mathrm{Q},k}, \tag{14}$$

where $\mathbf{s}_{\mathrm{Q},k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ denotes the data vector of UT k. Moreover, we also consider the sum power constraint $\mathrm{tr}(\mathbf{Q}_k) \le P_k$ for each UT k in the UL transmission. We assume that K mobile UTs send messages to the satellite simultaneously. The UT index set is denoted by $\mathcal{K} = \{1, \ldots, K\}$. Thus, the received signal $\mathbf{y} \in \mathbb{C}^{M\times 1}$ at the satellite is expressed as

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k\mathbf{x}_k + \mathbf{z}, \tag{15}$$

where $\mathbf{z} \in \mathbb{C}^{M\times 1}$ is the additive Gaussian noise at the satellite distributed as $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_M)$.

We assume that perfect iCSI is known by the receiver at the satellite side, while only sCSI is known by the transmitters at the UTs' side. Thanks to the statistical reciprocity between the DL and UL channels [25], each UT can acquire its own sCSI based on the DL pilot signals sent by the satellite. The UL ESR capacity C is given by

$$C = \max_{\mathbf{Q}_k\succeq\mathbf{0},\ \mathrm{tr}(\mathbf{Q}_k)\le P_k,\ \forall k} \mathbb{E}\left\{\log\det\left(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^{K}\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H\right)\right\}. \tag{16}$$

By substituting $\mathbf{H}_k = \mathbf{g}_k\mathbf{d}_k^H$ into (16), C can be further written as

$$C = \max_{\mathbf{Q}_k\succeq\mathbf{0},\ \mathrm{tr}(\mathbf{Q}_k)\le P_k,\ \forall k} \mathbb{E}\left\{\log\det\left(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^{K}\mathbf{d}_k^H\mathbf{Q}_k\mathbf{d}_k\cdot\mathbf{g}_k\mathbf{g}_k^H\right)\right\}. \tag{17}$$

The ESR capacity depends on the long-term sCSI consisting of the channel parameters in the set $\mathcal{H} = \{\beta_k, \kappa_k, \mathbf{g}_k, \mathbf{d}_{k,0}, \mathbf{\Sigma}_k\}_{\forall k}$.

Although the problem in (17) is a convex optimization problem [26], the mathematical expectation in the ESR makes it challenging to manifest the solution of the problem in (17).

In fact, the rank of the matrix $\mathbf{Q}_k$ is the number of independent data streams that are delivered through UT k's channel. In the following theorem, we show the rank property of the optimal transmit covariance matrices $\{\mathbf{Q}_k\}_{k=1}^K$, which paves the way for the lower-dimensional representation of $\{\mathbf{Q}_k\}_{k=1}^K$.

Theorem 1:
The optimal transmit covariance matrices $\{\mathbf{Q}_k^\star\}_{k=1}^K$ that achieve the UL ESR capacity should satisfy

$$\mathrm{rank}(\mathbf{Q}_k^\star) \le \mathrm{rank}\left(\mathbf{R}_k^{\mathrm{ut}}\right),\ \forall k \in \mathcal{K}. \tag{18}$$

Proof:
Please refer to Appendix A. ∎

The rank property of $\mathbf{Q}_k$ for the kth UT in (18) holds independently of other UTs' channel correlation matrices. From Theorem 1, the maximum number of independent data streams transmitted from UT k to the satellite should be no larger than the rank of UT k's channel correlation matrix $\mathbf{R}_k^{\mathrm{ut}}$. It can be anticipated that if there are only sparse scatterers distributed around UT k, $\mathrm{rank}(\mathbf{Q}_k)$ may be much less than the number of antennas N at UT k. Moreover, because the assumption on each UT's channel distribution is not invoked in the proof procedure of Theorem 1, the rank property in (18) is applicable to a large set of channel distributions as long as each UT's channel is independently distributed.

Next, we show that in some extreme cases including the low signal-to-noise ratio (SNR) case and the high Rician factor case, the optimal $\{\mathbf{Q}_k\}_{k=1}^K$ to the problem in (17) are of rank-one.

1. *Low SNR Case*

By following the derivation techniques in [27], if $P_k \to 0$, $\forall k \in \mathcal{K}$, holds, the UL ESR is reduced into $\frac{1}{\sigma^2}\sum_{k=1}^K \mathrm{tr}\left(\mathbf{R}_k^{\mathrm{ut}}\mathbf{Q}_k\right)$. Then, the problem in (17) can be

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

simplified into

$$C^{\mathrm{LS}} = \max_{\mathbf{Q}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{Q}_k) \le P_k,\ \forall k \in \mathcal{K}} \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{R}_k^{\mathrm{ut}} \mathbf{Q}_k\right). \quad (19)$$

The optimal matrices $\{\mathbf{Q}_k\}_{k=1}^K$ to the problem in (19) are given by

$$\mathbf{Q}_k = P_k \cdot \mathbf{b}_{\mathrm{Q},k} \mathbf{b}_{\mathrm{Q},k}^H,\ \forall k \in \mathcal{K}, \quad (20)$$

where $\mathbf{b}_{\mathrm{Q},k} \in \mathbb{C}^{N\times 1}$ is the unit-norm eigenvector of $\mathbf{R}_k^{\mathrm{ut}}$ associated with its maximum eigenvalue. Notice that the elements of $\mathbf{b}_{\mathrm{Q},k}$ are different not only in phase but also in magnitude. Therefore, to implement the beamforming vector $\mathbf{b}_{\mathrm{Q},k}$ at UT k's side, both magnitude and phase of the antennas equipped at UT k's side should be able to be manipulated. Specifically, the data symbol of UT k should be multiplied by different elements of $\mathbf{b}_{\mathrm{Q},k}$, and then transmitted from the antennas at UT k's side.

2. High Rician Factor Case

If $\kappa_k \to \infty$ holds for each UT $k \in \mathcal{K}$, the problem in (17) is reduced into

$$C^{\mathrm{HR}} = \max_{\mathbf{Q}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{Q}_k) \le P_k,\ \forall k \in \mathcal{K}} \log\det\left(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^{K} \beta_k \mathbf{d}_{k,0}^H \mathbf{Q}_k \mathbf{d}_{k,0} \cdot \mathbf{g}_k \mathbf{g}_k^H\right). \quad (21)$$

The optimal matrices $\{\mathbf{Q}_k\}_{k=1}^K$ to the problem in (21) can be derived as follows

$$\mathbf{Q}_k = P_k \cdot \mathbf{d}_{k,0} \mathbf{d}_{k,0}^H,\ \forall k \in \mathcal{K}. \quad (22)$$

The optimal transmit strategy for the high Rician factor case is to perform the transmit beamforming along each UT's LoS direction seen at the UT side. In this case, owing to $\mathbf{d}_{k,0} = \mathbf{d}(\boldsymbol{\varphi}_{k,0})$, only the paired AoDs $\boldsymbol{\varphi}_{k,0} = (\varphi_{k,0}^{\mathrm{x}'}, \varphi_{k,0}^{\mathrm{z}'})$ for the LoS path is required to be known at UT k. Moreover, the relatively simple phased array antennas (PAAs) can be used at the UT sides to implement the beamformers $\{\sqrt{P_k}\mathbf{d}_{k,0}\}_{k=1}^K$, which can significantly reduce the implementation cost and complexity.

B. Lower-Dimensional Matrix Representation of Transmit Covariance Matrix

In this subsection, we show that each UT k's transmit covariance matrix can be represented by a lower-dimensional matrix, and the transmit covariance matrix design can be transformed into the lower-dimensional matrix design. Let us denote the EVD of $\boldsymbol{\Sigma}_k$ in (13) as $\boldsymbol{\Sigma}_k = \mathbf{U}_k \mathrm{diag}(\boldsymbol{\lambda}_k)\mathbf{U}_k^H$. The columns of $\mathbf{U}_k = [\mathbf{u}_{k,1}\ \cdots\ \mathbf{u}_{k,S_k}] \in \mathbb{C}^{N\times S_k}$ are the eigenvectors with $S_k = \mathrm{rank}(\boldsymbol{\Sigma}_k)$. The elements of $\boldsymbol{\lambda}_k = [\lambda_{k,1}, \ldots, \lambda_{k,S_k}]^T$ are the corresponding positive eigenvalues in non-increasing order, which means $\lambda_{k,1} \ge \cdots \ge \lambda_{k,S_k}$. Let us further denote the linear subspace spanned by the columns in $\mathbf{U}_k$ as $\mathrm{span}(\mathbf{U}_k)$. We can separate $\mathbf{d}_{k,0}$ into two orthogonal terms as

$$\mathbf{d}_{k,0} = \underline{\mathbf{u}}_{k,0} + \mathbf{U}_k \boldsymbol{\xi}_{k,0}, \quad (23)$$

where $\underline{\mathbf{u}}_{k,0} \triangleq (\mathbf{I} - \mathbf{U}_k\mathbf{U}_k^H)\mathbf{d}_{k,0}$ and $\boldsymbol{\xi}_{k,0} \triangleq \mathbf{U}_k^H \mathbf{d}_{k,0}$. The first term in (23) is orthogonal to the linear subspace $\mathrm{span}(\mathbf{U}_k)$ and the second term lies in $\mathrm{span}(\mathbf{U}_k)$. By using the Rician fading channel assumption in (10), we can rewrite $\tilde{\mathbf{d}}_k$ as $\tilde{\mathbf{d}}_k = \mathbf{U}_k \underline{\tilde{\mathbf{c}}}_k$, where the elements in $\underline{\tilde{\mathbf{c}}}_k = \left[\tilde{c}_{k,1}, \ldots, \tilde{c}_{k,S_k}\right]^T \in \mathbb{C}^{S_k\times 1}$ are independent CSCG random variables with distinct variances. Indeed, $\underline{\tilde{\mathbf{c}}}_k$ is distributed as $\underline{\tilde{\mathbf{c}}}_k \sim \mathcal{CN}\left(\mathbf{0}, \mathrm{diag}(\boldsymbol{\lambda}_k)\right)$. Henceforth, $\mathbf{d}_k$ can be rewritten as

$$\begin{aligned}\mathbf{d}_k &= \sqrt{\frac{\kappa_k\beta_k}{\kappa_k+1}}\mathbf{d}_{k,0} + \sqrt{\frac{\beta_k}{\kappa_k+1}}\mathbf{U}_k\underline{\tilde{\mathbf{c}}}_k \\ &= \sqrt{\frac{\kappa_k\beta_k}{\kappa_k+1}}\underline{\mathbf{u}}_{k,0} + \sqrt{\frac{\kappa_k\beta_k}{\kappa_k+1}}\mathbf{U}_k\boldsymbol{\xi}_{k,0} + \sqrt{\frac{\beta_k}{\kappa_k+1}}\mathbf{U}_k\underline{\tilde{\mathbf{c}}}_k \\ &= \sqrt{\frac{\beta_k}{\kappa_k+1}}\left(\sqrt{\kappa_k}\underline{\mathbf{u}}_{k,0} + \mathbf{U}_k\left(\sqrt{\kappa_k}\boldsymbol{\xi}_{k,0} + \underline{\tilde{\mathbf{c}}}_k\right)\right).\end{aligned} \quad (24)$$

Let $\eta_{k,0} = \|\underline{\mathbf{u}}_{k,0}\|^2$ and $\mathbf{u}_{k,0} = \frac{\underline{\mathbf{u}}_{k,0}}{\|\underline{\mathbf{u}}_{k,0}\|}$. Define $\mathbf{B}_k \in \mathbb{C}^{N\times\tilde{S}_k}$ and $\mathbf{c}_k \in \mathbb{C}^{\tilde{S}_k\times 1}$ as follows

$$\mathbf{B}_k = \begin{cases}\left[\mathbf{u}_{k,0}\ \mathbf{U}_k\right], & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0}, \\ \mathbf{U}_k, & \text{if } \underline{\mathbf{u}}_{k,0} = \mathbf{0},\end{cases} \quad (25a)$$

$$\mathbf{c}_k = \begin{cases}\sqrt{\frac{\beta_k}{\kappa_k+1}}\begin{bmatrix}\sqrt{\kappa_k\eta_{k,0}} \\ \sqrt{\kappa_k}\boldsymbol{\xi}_{k,0} + \underline{\tilde{\mathbf{c}}}_k\end{bmatrix}, & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0}, \\ \sqrt{\frac{\beta_k}{\kappa_k+1}}\left(\sqrt{\kappa_k}\boldsymbol{\xi}_{k,0} + \underline{\tilde{\mathbf{c}}}_k\right), & \text{if } \underline{\mathbf{u}}_{k,0} = \mathbf{0},\end{cases} \quad (25b)$$

respectively, where $\tilde{S}_k$ is defined as

$$\tilde{S}_k = \begin{cases}S_k + 1, & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0}, \\ S_k, & \text{if } \underline{\mathbf{u}}_{k,0} = \mathbf{0}.\end{cases} \quad (26)$$

From (24) and (25), we can rewrite $\mathbf{d}_k$ as

$$\mathbf{d}_k = \mathbf{B}_k\mathbf{c}_k, \quad (27)$$

where the columns in $\mathbf{B}_k$ are orthogonal to each other, i.e., $\mathbf{B}_k^H\mathbf{B}_k = \mathbf{I}$. Notice that $\mathbf{c}_k$ can be written as $\mathbf{c}_k = \sqrt{\frac{\kappa_k\beta_k}{\kappa_k+1}}\mathbf{c}_{k,0} + \sqrt{\frac{\beta_k}{\kappa_k+1}}\tilde{\mathbf{c}}_k$, where $\mathbf{c}_{k,0}$ and $\tilde{\mathbf{c}}_k$ are given by

$$\mathbf{c}_{k,0} = \begin{cases}[\sqrt{\eta_{k,0}}\ \boldsymbol{\xi}_{k,0}^T]^T, & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0}, \\ \boldsymbol{\xi}_{k,0}, & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0},\end{cases} \quad (28)$$

$$\tilde{\mathbf{c}}_k = \begin{cases}[0\ \underline{\tilde{\mathbf{c}}}_k^T]^T, & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0}, \\ \underline{\tilde{\mathbf{c}}}_k, & \text{if } \underline{\mathbf{u}}_{k,0} = \mathbf{0},\end{cases} \quad (29)$$

respectively. Moreover, $\mathbf{R}_k^{\mathrm{ut}}$ can be rewritten as

$$\mathbf{R}_k^{\mathrm{ut}} = \mathbf{B}_k\boldsymbol{\Omega}_k\mathbf{B}_k^H, \quad (30)$$

where $\boldsymbol{\Omega}_k \triangleq \mathbb{E}\{\mathbf{c}_k\mathbf{c}_k^H\} \in \mathbb{C}^{\tilde{S}_k\times\tilde{S}_k}$ is given by

$$\boldsymbol{\Omega}_k = \begin{cases}\begin{bmatrix}\frac{\kappa_k\beta_k}{\kappa_k+1}\eta_{k,0}, & \frac{\kappa_k\beta_k\sqrt{\eta_{k,0}}}{\kappa_k+1}\boldsymbol{\xi}_{k,0}^H \\ \frac{\kappa_k\beta_k\sqrt{\eta_{k,0}}}{\kappa_k+1}\boldsymbol{\xi}_{k,0}, & \frac{\beta_k}{\kappa_k+1}\tilde{\boldsymbol{\Omega}}_k\end{bmatrix}, & \text{if } \underline{\mathbf{u}}_{k,0} \ne \mathbf{0}, \\ \frac{\beta_k}{\kappa_k+1}\tilde{\boldsymbol{\Omega}}_k, & \text{if } \underline{\mathbf{u}}_{k,0} = \mathbf{0},\end{cases} \quad (31)$$

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

where $\tilde{\mathbf{\Omega}}_k = \kappa_k \boldsymbol{\xi}_{k,0}\boldsymbol{\xi}_{k,0}^H + \mathrm{diag}(\boldsymbol{\lambda}_k)$. Note that $\mathbf{\Omega}_k$ is positive definite. Thus, we have

$$\mathrm{rank}\left(\mathbf{R}_k^{\mathrm{ut}}\right) = \mathrm{rank}(\mathbf{B}_k\mathbf{\Omega}_k\mathbf{B}_k^H) \overset{\text{(a)}}{=} \mathrm{rank}(\mathbf{B}_k) \overset{\text{(b)}}{=} \tilde{S}_k, \quad (32)$$

where (a) comes from [28, Observation 7.1.8(b)], and (b) follows from the fact that $\mathbf{B}_k$ has orthogonal columns.

Theorem 2:
The optimal solution $\{\mathbf{Q}_k^\star\}_{k=1}^K$ to the problem in (17) can be obtained by

$$\mathbf{Q}_k^\star = \mathbf{B}_k\mathbf{T}_k^\star\mathbf{B}_k^H, \ \forall k \in \mathcal{K}. \quad (33)$$

Here, $\{\mathbf{T}_k^\star\}_{k=1}^K$ is the optimal solution to the following problem

$$C = \max_{\mathbf{T}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{T}_k) \le P_k,\ \forall k} \mathbb{E}\left\{\log\det\left(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^K \mathbf{c}_k^H\mathbf{T}_k\mathbf{c}_k \cdot \mathbf{g}_k\mathbf{g}_k^H\right)\right\}. \quad (34)$$

Proof:
Please refer to Appendix B. ∎

Theorem 2 reveals that each UT k's $N \times N$ transmit covariance matrix $\mathbf{Q}_k$ can be represented by an $\tilde{S}_k \times \tilde{S}_k$ lower-dimensional matrix $\mathbf{T}_k$, whose dimension is exactly equal to $\mathrm{rank}(\mathbf{R}_k^{\mathrm{ut}})$ as shown by (32). Interestingly, this is consistent with the results in Theorem 1. It is worth noting that the proof of Theorem 2 relies on the equality $\mathbf{d}_k = \mathbf{B}_k\mathbf{c}_k$ in (27), which is derived based on the Rician fading channel assumption in (10). With the aid of Theorem 2, the transmit covariance matrix design can be transformed into the lower-dimensional matrix design without any loss of optimality. After the optimal lower-dimensional matrices $\{\mathbf{T}_k^\star\}_{k=1}^K$ are obtained, the optimal transmit covariance matrices $\{\mathbf{Q}_k^\star\}_{k=1}^K$ can be obtained immediately by using (33). Henceforth, we only need to concentrate on the optimization of the lower-dimensional matrices $\{\mathbf{T}_k\}_{k=1}^K$. The problem in (34) keeps the convex property, and the optimization variables therein have lower dimensions.

Let $\mathbf{T}_k = \mathbf{V}_{\mathrm{T},k}\mathbf{\Lambda}_{\mathrm{T},k}\mathbf{V}_{\mathrm{T},k}^H$ denote the EVD of $\mathbf{T}_k$, where the column vectors of $\mathbf{V}_{\mathrm{T},k} \in \mathbb{C}^{\tilde{S}_k \times \tilde{S}_k}$ consist of eigenvectors and $\mathbf{\Lambda}_{\mathrm{T},k} \in \mathbb{R}^{\tilde{S}_k \times \tilde{S}_k}$ is a diagonal matrix including the corresponding non-negative eigenvalues along the main diagonal. The transmit signal $\mathbf{x}_k$ should be given by

$$\mathbf{x}_k = \mathbf{B}_k\mathbf{V}_{\mathrm{T},k}\mathbf{\Lambda}_{\mathrm{T},k}^{1/2}\mathbf{s}_{\mathrm{T},k}, \quad (35)$$

where $\mathbf{s}_{\mathrm{T},k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{\tilde{S}_k})$ denotes the lower-dimensional data vector of UT k. Notice that the analysis of two extreme cases including the low SNR case and the high Rician factor case in Section A after Theorem 1 also holds here.

The following results show that the solution to the problem in (34) can be further simplified under some special conditions.

Theorem 3:
If $\boldsymbol{\xi}_{k,0} = \mathbf{0}$ holds for UT k, then $\mathbf{T}_k^\star$ is a diagonal matrix.

Proof:
Please refer to Appendix C. ∎

In fact, $\boldsymbol{\xi}_{k,0} = \mathbf{0}$ indicates that $\mathbf{d}_{k,0}$ is orthogonal to $\mathrm{span}(\mathbf{U}_k)$. In this case, $\mathbf{B}_k$ reduces to $\mathbf{B}_k = \left[\mathbf{d}_{k,0}\ \mathbf{U}_k\right]$, whose column vectors actually become the eigenvectors of $\mathbf{Q}_k^\star$. As shown in Theorem 3, if $\boldsymbol{\xi}_{k,0} = \mathbf{0}$ holds, the optimal transmit strategy of UT k would be sending independent data streams along the directions determined by the columns in $\mathbf{B}_k = \left[\mathbf{d}_{k,0}\ \mathbf{U}_k\right]$.

In the following, we provide a necessary and sufficient condition when single data stream transmission from each UT to the satellite can achieve the ESR capacity. The reason why we consider this is that single data stream transmission can reduce the design and implementation complexities further.

Theorem 4:
The optimal matrix $\mathbf{T}_k$ of the problem in (34) is given by $\mathbf{T}_k^\star = P_k\mathbf{w}_k\mathbf{w}_k^H$, if and only if

$$\Upsilon_{\max}\left(\mathbb{E}\left\{\frac{G_k}{1 + G_kP_k\left|\mathbf{c}_k^H\mathbf{w}_k\right|^2}\mathbf{c}_k\mathbf{c}_k^H\right\}\right) = \mathbb{E}\left\{\frac{G_k|\mathbf{c}_k^H\mathbf{w}_k|^2}{1 + G_kP_k\left|\mathbf{c}_k^H\mathbf{w}_k\right|^2}\right\}, \quad (36)$$

with $G_k = \frac{1}{\sigma^2}\mathbf{g}_k^H\tilde{\mathbf{A}}_k^{-1}\mathbf{g}_k$, and $\tilde{\mathbf{A}}_k = \mathbf{I}_M + \frac{1}{\sigma^2}\sum_{i\neq k}\mathbf{c}_i^H\mathbf{T}_i\mathbf{c}_i \cdot \mathbf{g}_i\mathbf{g}_i^H$, where $\mathbf{w}_k$ is a unit-norm vector.

Proof:
Please refer to Appendix D. ∎

According to Theorem 2, if $\mathbf{T}_k^\star = P_k\mathbf{w}_k\mathbf{w}_k^H$, then $\mathbf{Q}_k^\star = \mathbf{B}_k\mathbf{T}_k^\star\mathbf{B}_k^H = P_k\mathbf{B}_k\mathbf{w}_k\mathbf{w}_k^H\mathbf{B}_k^H$. The unit-norm vector $\mathbf{B}_k\mathbf{w}_k$ actually plays the role of the beamforming vector of UT k. Theorem 4 provides a necessary and sufficient condition when single data stream transmission from each UT to the satellite can achieve the UL ESR capacity. In other words, the optimal matrix $\mathbf{T}_k$ of the problem in (34) has the rank-one structure if and only if there exists some vector $\mathbf{w}_k$ which fulfills the condition in (36). With the help of Theorem 2, as long as the condition in (36) is satisfied, the lower-dimensional matrix design can be further simplified into the rank-one matrix design, i.e., vector design, for UT k, which can significantly reduce the implementation complexity especially at the UT sides.

On the other hand, if there does not exist such a vector $\mathbf{w}_k$ that satisfies the condition in (36), the single data stream transmission from UT k to the satellite will be unable to achieve the UL ESR capacity. Hence, it still needs to consider the lower-dimensional matrix design for the most general circumstances. In the next subsection, we elaborate the general design approaches for the low-dimensional matrices, which can achieve or approximate

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

the UL ESR capacity for massive MIMO LEO SATCOM systems.

C. CG Method for UL Transmit Design

In this subsection, we focus on the transmit covariance matrix optimization approach to achieve the UL ESR capacity for the most general cases regardless of whether single or multiple data stream transmission from each UT to the satellite can achieve the ESR capacity. In more detail, the CG based iterative method, a.k.a., the Frank-Wolfe method [29], is resorted to solve the convex optimization problem in (34) with guaranteed convergence. CG method is applicable to the convex programming with linear constraints. In each iteration of the CG method, a feasible direction to update the optimization variables is obtained by solving a simple linear programming problem. Simulation results in Section IV show that the CG method can converge to the optimal points of the problem in (34) within very few iterations.

The gradient of the ESR $R_{\rm sum} = \mathbb{E}\{\log\det(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^K \mathbf{c}_k^H\mathbf{T}_k\mathbf{c}_k \cdot \mathbf{g}_k\mathbf{g}_k^H)\}$ with respect to $\mathbf{T}_k$ can be written as follows

$$\mathbf{M}_k = \nabla_{\mathbf{T}_k} R_{\rm sum} = \mathbb{E}\left\{\tilde{W}_k\mathbf{c}_k\mathbf{c}_k^H\right\}, \tag{37}$$

where $\tilde{W}_k = \frac{1}{\sigma^2}\mathbf{g}_k^H\mathbf{\Theta}^{-1}\mathbf{g}_k = \frac{G_k}{1+G_k\mathbf{c}_k^H\mathbf{T}_k\mathbf{c}_k}$ with $\mathbf{\Theta} = \mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^K \mathbf{c}_k^H\mathbf{T}_k\mathbf{c}_k \cdot \mathbf{g}_k\mathbf{g}_k^H$. We use $(\cdot)^{(n)}$ to denote the argument in the nth iteration. Given $\{\mathbf{T}_k^{(n)}\}_{k=1}^K$, a feasible solution to the problem in (34) can be derived by solving the following linear programming problem [29]

$$\max_{\mathbf{T}_k\succeq\mathbf{0},\ {\rm tr}(\mathbf{T}_k)\le P_k,\ \forall k\in\mathcal{K}} \sum_{k=1}^K {\rm tr}\left(\mathbf{M}_k^{(n)}\mathbf{T}_k\right). \tag{38}$$

The closed-form solution to the problem in (38) can be written as [28]

$$\tilde{\mathbf{T}}_k^{(n+1)} = P_k \cdot \mathbf{m}_{k,n}\mathbf{m}_{k,n}^H,\ \forall k\in\mathcal{K}, \tag{39}$$

where $\mathbf{m}_{k,n} \in \mathbb{C}^{\tilde{S}_k\times 1}$ is the unit-norm eigenvector of $\mathbf{M}_k^{(n)}$ corresponding to its maximum eigenvalue. In terms of the CG method, $\{\mathbf{T}_k^{(n+1)}\}_{k=1}^K$ is given by

$$\mathbf{T}_k^{(n+1)} = \mathbf{T}_k^{(n)} + \alpha_k^{(n)}\left(\tilde{\mathbf{T}}_k^{(n+1)} - \mathbf{T}_k^{(n)}\right),\ \forall k\in\mathcal{K}, \tag{40}$$

where $\alpha_k^{(n)} \in (0,1]$ is the stepsize of UT k in the nth iteration. Th values of the stepsize can be determined by the line search method [26], [29]. Once the optimal lower-dimensional $\{\mathbf{T}_k\}_{k=1}^K$ are obtained, the optimal $\{\mathbf{Q}_k\}_{k=1}^K$ can be obtained immediately by using (33). The CG method for UL transmit design is summarized in Algorithm 1.

The major computational burden of Algorithm 1 comes from the computation of $\mathbf{M}_k$, which needs the inverse of an $M\times M$ matrix followed by the multiplication with an $M\times 1$ vector for each UT k. In terms of the number of multiplications, the computational complexity of Algorithm 1 is nearly given by $\mathcal{O}\left(N_S(K^3+K^2M+M^2K)\right)$,

Algorithm 1 CG method for UL transmit design.

Input: Initialize matrices $\mathbf{T}_k^{(0)} = \frac{P_k}{\tilde{S}_k}\cdot\mathbf{I}$, $\forall k\in\mathcal{K}$, and iteration index $n=0$.
Output: Transmit covariance matrices $\{\mathbf{Q}_k\}_{k=1}^K$.
while 1 **do**
  Compute $\mathbf{M}_k^{(n)}$ and corresponding $\mathbf{m}_{k,n}$, $\forall k\in\mathcal{K}$.
  Update $\{\mathbf{T}_k^{(n+1)}\}_{k=1}^K$ according to (40).
  if $n\ge N_{\rm iter}-1$ or $|R_{\rm sum}^{(n+1)} - R_{\rm sum}^{(n)}| < \epsilon$ **then**
    Set $\mathbf{T}_k := \mathbf{T}_k^{(n+1)}$, $\forall k\in\mathcal{K}$, **break**.
  else
    Set $n := n+1$.
  end if
end while
Compute $\mathbf{Q}_k = \mathbf{B}_k\mathbf{T}_k\mathbf{B}_k^H$, $\forall k\in\mathcal{K}$.

where N_S is the number of channel samples used to calculate the mathematical expectation.

Because of the difficulty to derive a closed-form expression of the ESR, Monte-Carlo method is needed to compute the mathematical expectation. In particular, the complicated sample average has to be used for the computation of $\mathbf{M}_k^{(n)}$ in Algorithm 1. Next, to avoid the exhaustive sample average, the asymptotic expression of ESR is used for the transmit design, which has a lower computational complexity and can achieve near-optimal performance.

D. Simplified CG Method for UL Transmit Design

In this subsection, we utilize an asymptotic expression of the ESR and devise a simplified CG method to compute the transmit covariance matrices, which can approximate the ESR capacity. Compared with Algorithm 1, the time-consuming sample average when computing $\mathbf{M}_k^{(n)}$ is no longer required in the simplified CG method.

Let $\boldsymbol{\omega}_k \triangleq \frac{\beta_k}{\kappa_k+1}\mathbb{E}\left\{\tilde{\mathbf{c}}_k\odot\tilde{\mathbf{c}}_k^*\right\} \in \mathbb{C}^{\tilde{S}_k\times 1}$, and define $\bar{\mathbf{H}}_{\rm e} \in \mathbb{C}^{M\times\tilde{S}}$ as follows

$$\bar{\mathbf{H}}_{\rm e} = \left[\sqrt{\frac{\kappa_1\beta_1}{\kappa_1+1}}\mathbf{g}_1\mathbf{c}_{1,0}^H\ \cdots\ \sqrt{\frac{\kappa_K\beta_K}{\kappa_K+1}}\mathbf{g}_K\mathbf{c}_{K,0}^H\right], \tag{41}$$

with $\tilde{S} = \sum_{k=1}^K \tilde{S}_k$. An asymptotic expression, a.k.a. the deterministic equivalent, of the ESR can be written as [30]–[33]

$$R_{\rm sum} \to \underline{R}_{\rm sum} = \log\det\left(\mathbf{I}+\mathbf{\Xi}\mathbf{T}\right) + \log\det\left(\mathbf{I}+\mathbf{\Phi}_{\rm R}\right) - \sum_{k=1}^K \gamma_k\boldsymbol{\omega}_k^T\boldsymbol{\psi}_k, \tag{42}$$

where $\mathbf{T} = {\rm diag}(\mathbf{T}_1,\ldots,\mathbf{T}_K) \in \mathbb{C}^{\tilde{S}\times\tilde{S}}$, $\boldsymbol{\psi}_k = [\psi_{k,1}\ \cdots\ \psi_{k,\tilde{S}_k}]^T \in \mathbb{R}^{\tilde{S}_k\times 1}$. In (42), $\{(\gamma_k,\boldsymbol{\psi}_k)\}_{k=1}^K$ is the unique solution of the following equations

$$\gamma_k = \mathbf{g}_k^H\left(\mathbf{I}+\mathbf{\Psi}\right)^{-1}\mathbf{g}_k, \tag{43a}$$

$$\boldsymbol{\psi}_k = {\rm diag}\left(\mathbf{T}_k\left\langle\left(\mathbf{I}+\mathbf{\Xi}\mathbf{T}\right)^{-1}\right\rangle_k\right), \tag{43b}$$

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

$k \in \mathcal{K}$, where $\langle\cdot\rangle_k$ means the operation of taking the kth sub-block along the diagonal of the matrix argument. In addition, $\mathbf{\Xi} \in \mathbb{C}^{\tilde{S}\times\tilde{S}}$ and $\mathbf{\Psi} \in \mathbb{C}^{M\times M}$ in (43) depend on $\{(\gamma_k, \boldsymbol{\psi}_k)\}_{k=1}^K$ by

$$\mathbf{\Xi} = \mathbf{\Phi}_{\mathrm{T}} + \bar{\mathbf{H}}_{\mathrm{e}}^H \left(\mathbf{I} + \mathbf{\Phi}_{\mathrm{R}}\right)^{-1} \bar{\mathbf{H}}_{\mathrm{e}}, \tag{44a}$$

$$\mathbf{\Psi} = \mathbf{\Phi}_{\mathrm{R}} + \bar{\mathbf{H}}_{\mathrm{e}} \mathbf{T} \left(\mathbf{I} + \mathbf{\Phi}_{\mathrm{T}}\mathbf{T}\right)^{-1} \bar{\mathbf{H}}_{\mathrm{e}}^H, \tag{44b}$$

respectively, where $\mathbf{\Phi}_{\mathrm{T}}$ and $\mathbf{\Phi}_{\mathrm{R}}$ are given by

$$\mathbf{\Phi}_{\mathrm{T}} = \mathrm{diag}\left(\mathbf{\Phi}_{\mathrm{T},1}, \ldots, \mathbf{\Phi}_{\mathrm{T},K}\right), \tag{45a}$$

$$\mathbf{\Phi}_{\mathrm{R}} = \sum_{k=1}^{K} \mathbf{\Phi}_{\mathrm{R},k}, \tag{45b}$$

respectively, with $\mathbf{\Phi}_{\mathrm{T},k} = \gamma_k \cdot \mathrm{diag}(\boldsymbol{\omega}_k) \in \mathbb{C}^{\tilde{S}_k\times\tilde{S}_k}$ and $\mathbf{\Phi}_{\mathrm{R},k} = \boldsymbol{\omega}_k^T\boldsymbol{\psi}_k \cdot \mathbf{g}_k\mathbf{g}_k^H \in \mathbb{C}^{M\times M}$. It has been claimed that the unique solution $\{(\gamma_k, \boldsymbol{\psi}_k)\}_{k=1}^K$ to the equations in (43) can be obtained by performing a fixed-point iterative procedure until convergence [30]–[33]. Then, the asymptotic approximation of the problem in (34) can be formulated as follows

$$\underline{C} = \max_{\mathbf{T}_k\succeq\mathbf{0},\ \mathrm{tr}(\mathbf{T}_k)\le P_k,\ \forall k} \underline{R}_{\mathrm{sum}}. \tag{46}$$

The gradient of $\underline{R}_{\mathrm{sum}}$ with respect to $\mathbf{T}_k$ is given by [30]–[33]

$$\underline{\mathbf{M}}_k = \nabla_{\mathbf{T}_k}\underline{R}_{\mathrm{sum}} = \left\langle \left(\mathbf{I} + \mathbf{\Xi}\mathbf{T}\right)^{-1}\mathbf{\Xi} \right\rangle_k. \tag{47}$$

For given $\{\mathbf{T}_k^{(n)}\}_{k=1}^K$, a feasible solution of the problem in (46) in the nth iteration can be computed by solving the following subproblem

$$\max_{\mathbf{T}_k\succeq\mathbf{0},\ \mathrm{tr}(\mathbf{T}_k)\le P_k,\ \forall k\in\mathcal{K}} \sum_{k=1}^{K} \mathrm{tr}\left(\underline{\mathbf{M}}_k^{(n)}\mathbf{T}_k\right). \tag{48}$$

The closed-form solution to the problem in (48) can be written as [28]

$$\underline{\tilde{\mathbf{T}}}_k^{(n+1)} = P_k \cdot \underline{\mathbf{m}}_{k,n}\underline{\mathbf{m}}_{k,n}^H,\ \forall k \in \mathcal{K}, \tag{49}$$

where $\underline{\mathbf{m}}_{k,n} \in \mathbb{C}^{\tilde{S}_k\times 1}$ is the unit-norm eigenvector of $\underline{\mathbf{M}}_k^{(n)}$ corresponding to its maximum eigenvalue. Then, $\{\mathbf{T}_k^{(n+1)}\}_{k=1}^K$ is given by

$$\mathbf{T}_k^{(n+1)} = \mathbf{T}_k^{(n)} + \underline{\alpha}_k^{(n)}\left(\underline{\tilde{\mathbf{T}}}_k^{(n+1)} - \mathbf{T}_k^{(n)}\right),\ \forall k \in \mathcal{K}, \tag{50}$$

where $\underline{\alpha}_k^{(n)} \in (0, 1]$ is the stepsize of the UT k in the nth iteration. After $\{\mathbf{T}_k^{(n)}\}_{k=1}^K$ converges, $\{\mathbf{Q}_k\}_{k=1}^K$ can be obtained with (33). The simplified CG method for UL transmit design with asymptotic ESR is shown in Algorithm 2.

The major computational burden of Algorithm 2 comes from the calculation of $\underline{\mathbf{M}}_k$, in which the complexity to inverse an $\tilde{S}\times\tilde{S}$ matrix is about $\mathcal{O}(\tilde{S}^3)$, and the complexity to compute $\bar{\mathbf{H}}_{\mathrm{e}}^H\left(\mathbf{I}+\mathbf{\Phi}_{\mathrm{R}}\right)^{-1}\bar{\mathbf{H}}_{\mathrm{e}}$ is mainly $\mathcal{O}(K^3 + K^2M + M^2\tilde{S} + \tilde{S}^2M)$. Therefore, the computational complexity of Algorithm 2 is approximately given by $\mathcal{O}(\tilde{S}^3 + K^3 + K^2M + M^2\tilde{S} + \tilde{S}^2M)$.

Algorithm 2 Simplified CG method for UL transmit design.

Input: Initialize matrices $\mathbf{T}_k^{(0)} = \frac{P_k}{\tilde{S}_k}\cdot\mathbf{I}$, $\forall k \in \mathcal{K}$, and iteration index $n = 0$.
Output: Transmit covariance matrices $\{\mathbf{Q}_k\}_{k=1}^K$.

```
while 1 do
    Calculate M_k^(n) and corresponding m_{k,n}, ∀k ∈ K.
    Update {T_k^(n+1)}_{k=1}^K according to (50).
    if n ≥ N_iter − 1 or |R_sum^(n+1) − R_sum^(n)| < ϵ then
        Set T_k := T_k^(n+1), ∀k ∈ K, break.
    else
        Set n := n + 1.
    end if
end while
Compute Q_k = B_k T_k B_k^H, ∀k ∈ K.
```

TABLE I: Simulation Parameters

Parameters	Values
Earth radius R_e	6378 km
Orbit altitude H	1000 km
Central frequency f_c	2 GHz
Bandwidth B	20 MHz
Noise temperature T_{n}	273 K
Number of antennas at satellite $M_{\mathrm{x}} \times M_{\mathrm{y}}$	12×12
Number of antennas at UTs $N_{\mathrm{x}'} \times N_{\mathrm{y}'}$	6×6
Antenna spacing at satellite $d_{\mathrm{x}}(d_{\mathrm{y}})$	λ
Antenna spacing at UTs $d_{\mathrm{x}'}(d_{\mathrm{y}'})$	$\frac{\lambda}{2}$
Antenna gain at satellite G_{sat}	7 dBi
Antenna gain at UTs G_{ut}	0 dBi
Maximum nadir angle $\theta_{\max}^{\mathrm{z}}$	30°
Number of UTs K	100
Transmit power per UT	20 dBm – 40 dBm

TABLE II: Shadow fading parameter

Elevation	10°	20°	30°	40°	50°	60°	70°	80°	90°
$\sigma_{\mathrm{SF},k}$	3.5	3.4	2.9	3.0	3.1	2.7	2.5	2.3	1.2

IV. Simulation Results

In this section, we provide the simulation results to verify the performance of the proposed UL transmit designs in massive MIMO LEO SATCOM. The simulation parameters are summarized in TABLE I. We denote the maximum nadir angle of the UTs as $\theta_{\max}^{\mathrm{z}}$. The space angle pair $\boldsymbol{\zeta}_k = (\zeta_k^{\mathrm{x}}, \zeta_k^{\mathrm{y}})$ is generated according to the uniform distribution in the circle $\{(x, y)|x^2 + y^2 \le \sin^2\theta_{\max}^{\mathrm{z}}\}$. The elevation angle of UT k in Fig. 1 is given by $v_k = \cos^{-1}\left(\frac{R_s}{R_e}\sin\theta_k^{\mathrm{z}}\right)$, where R_e is the earth radius, $R_s = R_e + H$ is the orbit radius [23]. The distance between the satellite and UT k in Fig. 1 is given by $D_k = \sqrt{R_e^2\sin^2 v_k + H^2 + 2HR_e} - R_e\sin v_k$ [2]. The per-antenna gains of the UPAs at the satellite and the

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

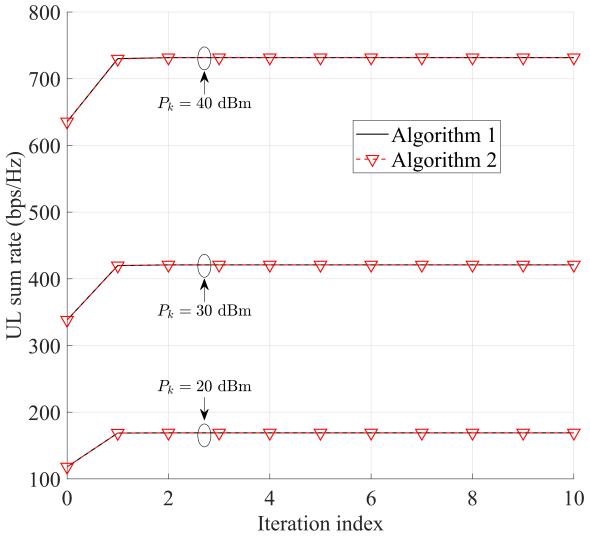

Fig. 2: Convergence of Algorithms 1 and 2 at different transmit power.

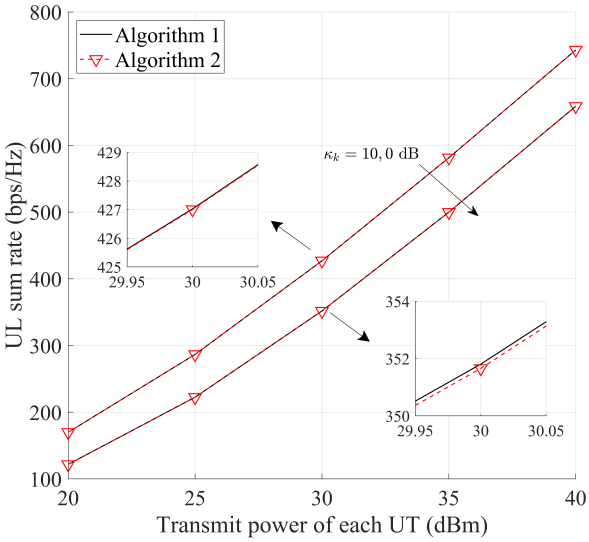

Fig. 3: Performance of Algorithms 1 and 2 for different Rician factors.

UT sides are denoted as G_{sat} and G_{ut}, respectively. The pathloss in dB between the satellite and UT k is given by $\text{PL}_k = 32.45 + 20\log_{10}(f_c) + 20\log_{10}(D_k)$, where f_c is the carrier frequency in GHz, D_k is the distance between the satellite and UT k in meter. The shadow fading in dB is assumed to follow the Gaussian distribution $\text{SF}_k \sim \mathcal{N}(0, \sigma^2_{\text{SF},k})$. For given elevation angle v_k of UT k, we use the value of $\sigma_{\text{SF},k}$ in TABLE II [2, Table 6.6.2-1] whose corresponding elevation angle is nearest to v_k. The ionospheric loss is set as 2 dB approximately [2, Section 6]. For simplicity, we assume that each UT's UPA is placed horizontally, which implies that the paired AoDs $\boldsymbol{\varphi}_{k,0} = (\varphi^{\text{x}'}_{k,0}, \varphi^{\text{z}'}_{k,0})$ associated with the LoS path of UTk's channel satisfies $\varphi^{\text{z}'}_{k,0} = 90^\circ - v_k$. In addition, the covariance matrix $\boldsymbol{\Sigma}_k$ is constructed by $\boldsymbol{\Sigma}_k = \mathbf{f}_k\mathbf{f}_k^H$ with $\mathbf{f}_k = \mathbf{a}_{N_{\text{x}'}}(\phi^{\text{x}'}_k + 2/N_{\text{x}'}) \otimes \mathbf{a}_{N_{\text{y}'}}(\phi^{\text{y}'}_k + 2/N_{\text{y}'})$ with $\phi^{\text{x}'}_k = \sin\varphi^{\text{z}'}_{k,0}\cos\varphi^{\text{x}'}_{k,0}$ and $\phi^{\text{y}'}_k = \sin\varphi^{\text{z}'}_{k,0}\sin\varphi^{\text{x}'}_{k,0}$, so that the rank of $\boldsymbol{\Sigma}_k$ is $S_k = 1$, $\forall k \in \mathcal{K}$. The noise variance is given by $\sigma^2 = k_{\text{B}}T_{\text{n}}B$, where $k_{\text{B}} = 1.38 \times 10^{-23}$ J $\cdot$ K^{-1} is the Boltzmann constant, T_{n} is the noise temperature and B is the system bandwidth.

In Fig. 2, the convergence performance of Algorithms 1 and 2 is depicted. The transmit power P_k of each UT k takes different values including 20 dBm, 30 dBm, and 40 dBm. The Rician factor κ_k of each UT k is set as $\kappa_k = 10$ dB, $\forall k \in \mathcal{K}$. To solve the fixed-point equations in (43), the number of iterations, N_{f}, is set as $N_{\text{f}} = 10$ hereafter. As we can see, both Algorithms 1 and 2 can converge to the optimal values within a very small number of iterations.

Fig. 3 shows the performance of Algorithms 1 and 2 for different values of Rician factors. It can be observed that the performance loss between Algorithms 1 and 2 is negligible for both high and low Rician factor cases. In addition, we notice that the increase of Rician factors may bring the improvement of UL ESR capacity in massive MIMO LEO SATCOM, which means that the LoS components in LEO satellite channels may be more favorable to UL transmission.

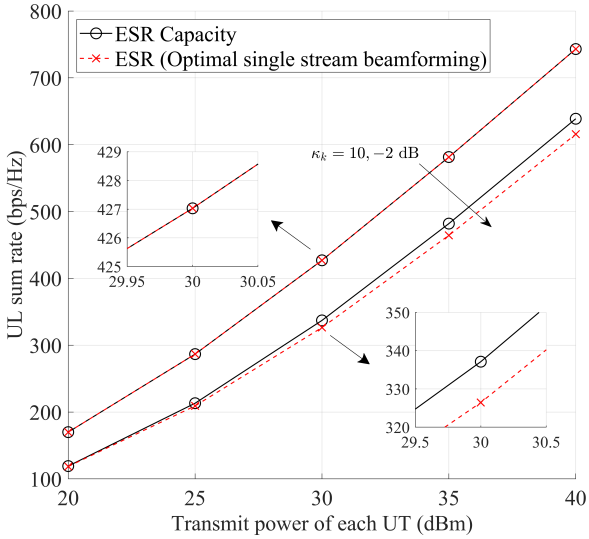

Fig. 4: Comparison of ESR capacity and ESR with optimal single-stream beamforming.

In Fig. 4, the ESR capacity and the optimal ESR attained by single data stream transmission from each UT to the satellite is compared. The ESR capacity is calculated by using Algorithm 1. Meanwhile, the optimal ESR with single data stream transmission from each UT to the satellite can be attained by using the optimal single-stream beamforming vectors of UTs denoted as $\{\mathbf{B}_k\mathbf{w}_k^{\text{opt}}\}_{k=1}^K$ with $\{\mathbf{w}_k^{\text{opt}}\}_{k=1}^K = \arg\max_{\|\mathbf{w}_k\|=1,\ \forall k\in\mathcal{K}} \mathbb{E}\{\log\det(\mathbf{I}_M + \sum_{k=1}^K \frac{P_k}{\sigma^2}|\mathbf{c}_k^H\mathbf{w}_k|^2 \cdot \mathbf{g}_k\mathbf{g}_k^H)\}$, and this lower-dimensional vector optimization problem can be solved based on the

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

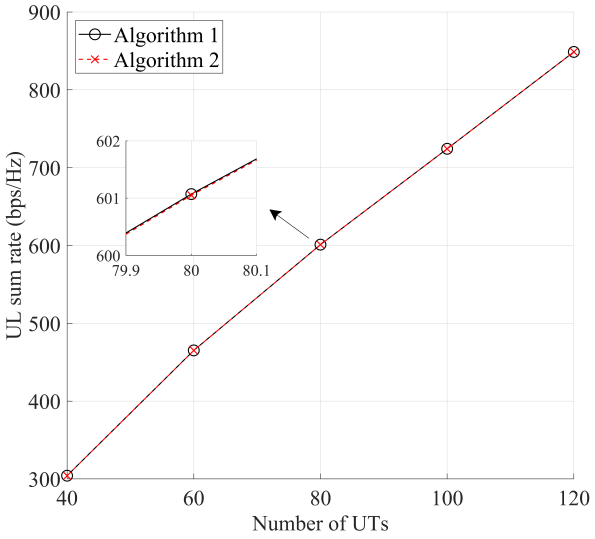

Fig. 5: Performance of Algorithms 1 and 2 for different numbers of UTs.

CG method, which is omitted for conciseness. It can be observed that when the Rician factors are relatively large, the ESR with optimal single-stream beamforming can be very close to the ESR capacity. Meanwhile, if the Rician factors are reduced to a relatively low level, which is likely to occur in some scenarios where the LoS components of UTs' channels suffer from some degree of blockage, the performance gap between the ESR capacity and the ESR with optimal single-stream beamforming will be non-negligible.

In Fig. 5, the performance of Algorithms 1 and 2 for different numbers of UTs is depicted, in which the Rician factors are set as $\kappa_k = 10$ dB, $\forall k \in \mathcal{K}$. It can be seen that the ESR attained by Algorithm 2 is still tight to the ESR capacity achieved by using Algorithm 1 for different numbers of UTs. In addition, as the number of UTs increases, the ESR capacity can be significantly improved in massive MIMO LEO SATCOM.

V. Conclusion

We have investigated the ESR capacity achieving UL transmit design with long-term sCSIT in massive MIMO LEO SATCOM systems. The UL massive MIMO LEO satellite channel model is established, where the satellite and UTs are equipped with UPAs. We prove that the rank of each UT's optimal transmit covariance matrix does not exceed that of its channel correlation matrix at the UT side. This reveals the maximum number of independent data streams transmitted from each multi-antenna UT to the satellite. We then prove that the transmit covariance matrix design can be transformed into the lower-dimensional matrix design without any loss of optimality. We obtain a necessary and sufficient condition when single data stream transmission from each UT to the satellite can achieve the ESR capacity. A CG method is developed to compute the ESR capacity achieving transmit covariance matrices with guaranteed convergence. Further, to avoid the exhaustive sample average, we resort to an asymptotic expression of the ESR and devise a simplified CG method to compute the transmit covariance matrices which can approximate the ESR capacity. The effectiveness of the proposed approaches is verified in the simulation results.

Appendix A
Proof of Theorem 1

The Lagrangian to the problem in (17) is given by

$$\mathcal{L}_{\mathrm{Q}} = \mathbb{E}\left\{\log\det\left(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^{K}\mathbf{d}_k^H\mathbf{Q}_k\mathbf{d}_k\cdot\mathbf{g}_k\mathbf{g}_k^H\right)\right\} + \sum_{k=1}^{K}\mathrm{tr}(\mathbf{S}_k\mathbf{Q}_k) - \sum_{k=1}^{K}\mu_k\left(\mathrm{tr}(\mathbf{Q}_k) - P_k\right), \tag{51}$$

where $\mu_k \geq 0$ and $\mathbf{S}_k \succeq \mathbf{0}$ are Lagrange multipliers associated with the constraints $\mathrm{tr}(\mathbf{Q}_k) \leq P_k$ and $\mathbf{Q}_k \succeq \mathbf{0}$, respectively. At the optimum to the problem in (17), there must exist the Lagrange multipliers $\mu_k^\star \geq 0$ and $\mathbf{S}_k^\star \succeq \mathbf{0}$ such that the gradient of $\mathcal{L}_{\mathrm{Q}}$ with respect to $\mathbf{Q}_k$ will vanish, i.e.,

$$\nabla_{\mathbf{Q}_k}\mathcal{L}_{\mathrm{Q}}\big|_{\mathbf{Q}_k=\mathbf{Q}_k^\star} = \mathbb{E}\left\{W_k^\star\mathbf{d}_k\mathbf{d}_k^H\right\} - \mu_k^\star\mathbf{I}_N + \mathbf{S}_k^\star = \mathbf{0}, \tag{52}$$

where $W_k^\star = \frac{\frac{1}{\sigma^2}\mathbf{g}_k^H\left(\mathbf{A}_k^\star\right)^{-1}\mathbf{g}_k}{1+\frac{1}{\sigma^2}\mathbf{g}_k^H\left(\mathbf{A}_k^\star\right)^{-1}\mathbf{g}_k\cdot\mathbf{d}_k^H\mathbf{Q}_k^\star\mathbf{d}_k}$ with $\mathbf{A}_k^\star = \mathbf{I}_M + \frac{1}{\sigma^2}\sum_{i\neq k}\mathbf{d}_i^H\mathbf{Q}_i^\star\mathbf{d}_i\cdot\mathbf{g}_i\mathbf{g}_i^H$. By multiplying $\mathbf{Q}_k^\star$ with the equation in (52) from the right side, it yields

$$\mathbb{E}\left\{W_k^\star\mathbf{d}_k\mathbf{d}_k^H\right\}\mathbf{Q}_k^\star = \mu_k^\star\mathbf{Q}_k^\star, \tag{53}$$

where the complementary slackness condition $\mathbf{S}_k^\star\mathbf{Q}_k^\star = \mathbf{0}$ in the Karush-Kuhn-Tucker (KKT) conditions is adopted [26]. We may now notice that the Lagrange multiplier $\mu_k^\star$ must be strictly positive, $\forall k \in \mathcal{K}$. Otherwise, the zero matrices $\{\mathbf{Q}_k = \mathbf{0}\}_{k=1}^K$ will be an optimal solution to the problem in (17), which is clearly not true. From the equality in (53), we can derive

$$\begin{aligned}\mathrm{rank}(\mathbf{Q}_k^\star) &= \mathrm{rank}\left(\mathbb{E}\left\{W_k^\star\mathbf{d}_k\mathbf{d}_k^H\right\}\mathbf{Q}_k^\star\right)\\ &\leq \mathrm{rank}\left(\mathbb{E}\left\{W_k^\star\mathbf{d}_k\mathbf{d}_k^H\right\}\right).\end{aligned} \tag{54}$$

In addition, for every random vector $\mathbf{d}_k$, we have

$$W_k^\star\mathbf{d}_k\mathbf{d}_k^H \overset{(a)}{\preceq} \frac{1}{\sigma^2}\mathbf{g}_k^H\left(\mathbf{A}_k^\star\right)^{-1}\mathbf{g}_k\cdot\mathbf{d}_k\mathbf{d}_k^H, \tag{55}$$

where (a) follows from $\mathbf{d}_k^H\mathbf{Q}_k^\star\mathbf{d}_k \geq 0$. By taking the mathematical expectation for all random vector $\mathbf{d}_k$ in (55), we have

$$\mathbb{E}\left\{W_k^\star\mathbf{d}_k\mathbf{d}_k^H\right\} \preceq \frac{1}{\sigma^2}\mathbb{E}\{\mathbf{g}_k^H\left(\mathbf{A}_k^\star\right)^{-1}\mathbf{g}_k\}\cdot\mathbf{R}_k^{\mathrm{ut}}, \tag{56}$$

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

where $\mathbb{E}\{\mathbf{g}_k^H \left(\mathbf{A}_k^\star\right)^{-1} \mathbf{g}_k\} > 0$ and $\mathbf{R}_k^{\rm ut} = \mathbb{E}\{\mathbf{d}_k\mathbf{d}_k^H\}$. By applying the inequality $\mathrm{rank}(\mathbf{A}) \geq \mathrm{rank}(\mathbf{B})$ for $\mathbf{A} \succeq \mathbf{B} \succeq \mathbf{0}$ [34, Theorem 7.8], the relation in (56) implies that

$$\mathrm{rank}\left(\mathbb{E}\left\{W_k^\star \mathbf{d}_k\mathbf{d}_k^H\right\}\right) \leq \mathrm{rank}\left(\mathbf{R}_k^{\rm ut}\right). \tag{57}$$

After combining the results in (54) and (57), we can complete the proof.

Appendix B
Proof of Theorem 2

By utilizing $\mathbf{d}_k = \mathbf{B}_k\mathbf{c}_k$ in (27), the UL ESR capacity in (17) can be rewritten as

$$C = \max_{\mathbf{Q}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{Q}_k) \leq P_k,\ \forall k} \mathbb{E}\left\{\log\det\left(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^K \mathbf{c}_k^H\mathbf{B}_k^H\mathbf{Q}_k\mathbf{B}_k\mathbf{c}_k \cdot \mathbf{g}_k\mathbf{g}_k^H\right)\right\}. \tag{58}$$

For any given positive semidefinite matrix $\mathbf{T}_k$, the linear matrix equation $\mathbf{B}_k^H\mathbf{Q}_k\mathbf{B}_k = \mathbf{T}_k$ always has a positive semidefinite solution $\mathbf{Q}_k$ if and only if $\mathbf{T}_k\mathbf{B}_k^-\mathbf{B}_k = \mathbf{T}_k$ [35], [36], where $\mathbf{B}_k^- \in \mathbb{C}^{\tilde{S}_k \times M}$ is an arbitrary generalized inverse (g-inverse) of $\mathbf{B}_k$. The g-inverse $\mathbf{A}^-$ of $\mathbf{A}$ is defined as the matrix that satisfies the equation $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$. In terms of [37, pp. 47], one of the g-inverses of $\mathbf{B}_k$ is given by $\mathbf{B}_k^- = \left(\mathbf{B}_k^H\mathbf{B}_k\right)^-\mathbf{B}_k^H = \mathbf{B}_k^H$ due to $\mathbf{B}_k^H\mathbf{B}_k = \mathbf{I}$. Thus, we have

$$\mathbf{T}_k\mathbf{B}_k^-\mathbf{B}_k = \mathbf{T}_k\mathbf{B}_k^H\mathbf{B}_k \overset{\rm (a)}{=} \mathbf{T}_k, \tag{59}$$

where (a) follows from $\mathbf{B}_k^H\mathbf{B}_k = \mathbf{I}$. From the relation in (59), we can conclude that for any given positive semidefinite $\mathbf{T}_k$, there always exists a positive semidefinite solution $\mathbf{Q}_k$ to the equation $\mathbf{B}_k^H\mathbf{Q}_k\mathbf{B}_k = \mathbf{T}_k$.

According to [35], for a given positive semidefinite matrix $\mathbf{T}_k$, the positive semidefinite solution $\mathbf{Q}_k$ to the equation $\mathbf{B}_k^H\mathbf{Q}_k\mathbf{B}_k = \mathbf{T}_k$ can be represented by

$$\mathbf{Q}_{N,k} = \mathbf{B}_k\mathbf{T}_k\mathbf{B}_k^H + \left(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^H\right)\mathbf{N}_k\left(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^H\right), \tag{60}$$

where $\mathbf{N}_k \in \mathbb{C}^{N\times N}$ is an arbitrary positive semidefinite matrix. Then, the trace of matrix $\mathbf{Q}_{N,k}$ in (60) is given by

$$\mathrm{tr}\left(\mathbf{Q}_{N,k}\right) = \mathrm{tr}\left(\mathbf{T}_k\right) + \Delta_k, \tag{61}$$

where $\Delta_k = \mathrm{tr}\left(\left(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^H\right)\mathbf{N}_k\left(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^H\right)\right)$. If $\Delta_k > 0$, we can always construct another matrix $\mathbf{Q}_{T,k} = \mathbf{B}_k\mathbf{T}_k\mathbf{B}_k^H$. Then, the matrix $\frac{\mathrm{tr}(\mathbf{Q}_{N,k})}{\mathrm{tr}(\mathbf{T}_k)}\mathbf{Q}_{T,k}$ can attain a larger UL ESR $I_{\rm sum}$ in (16) under the same power constraint with matrix $\mathbf{Q}_{N,k}$. On the other hand, if $\Delta_k = 0$, which indicates that $\left(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^H\right)\mathbf{N}_k\left(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^H\right) = \mathbf{0}$, then $\mathbf{Q}_{N,k}$ is exactly equal to $\mathbf{B}_k\mathbf{T}_k\mathbf{B}_k^H$. Consequently, for a given matrix $\mathbf{T}_k$, we only need to consider the solutions $\mathbf{Q}_k$ to the linear matrix equation $\mathbf{B}_k^H\mathbf{Q}_k\mathbf{B}_k = \mathbf{T}_k$ with the following form

$$\mathbf{Q}_k = \mathbf{B}_k\mathbf{T}_k\mathbf{B}_k^H. \tag{62}$$

Hence, the power constraint $\mathrm{tr}(\mathbf{Q}_k) \leq P_k$ can be rewritten as

$$\begin{aligned}\mathrm{tr}(\mathbf{Q}_k) &= \mathrm{tr}(\mathbf{B}_k\mathbf{T}_k\mathbf{B}_k^H)\\ &= \mathrm{tr}(\mathbf{B}_k^H\mathbf{B}_k\mathbf{T}_k)\\ &\overset{\rm (a)}{=} \mathrm{tr}(\mathbf{T}_k) \leq P_k,\end{aligned} \tag{63}$$

where (a) follows from $\mathbf{B}_k^H\mathbf{B}_k = \mathbf{I}$. With the above statement, the problem in (17) can be equivalently transformed into the problem in (34). Furthermore, once the optimal solution $\{\mathbf{T}_k^\star\}_{k=1}^K$ to the problem in (34) is obtained, the optimal matrices $\{\mathbf{Q}_k^\star\}_{k=1}^K$ to the problem in (17) can be directly derived with the aid of the relation in (33). This completes the proof.

Appendix C
Proof of Theorem 3

In light of (23), if $\boldsymbol{\xi}_{k,0} = \mathbf{0}$ holds for UT k, then $\underline{\mathbf{u}}_{k,0} = \mathbf{u}_{k,0} = \mathbf{d}_{k,0}$, $\tilde{S}_k = S_k + 1$ and $\mathbf{d}_k = \mathbf{B}_k\mathbf{c}_k$ where $\mathbf{B}_k = [\mathbf{d}_{k,0}\ \mathbf{U}_k]$ and $\mathbf{c}_k$ is given by

$$\mathbf{c}_k = \sqrt{\frac{\beta_k}{\kappa_k + 1}}\begin{bmatrix}\sqrt{\kappa_k}\\ \tilde{\underline{\mathbf{c}}}_k\end{bmatrix}. \tag{64}$$

Notice that $\tilde{\underline{\mathbf{c}}}_k$ is a CSCG random vector distributed as $\tilde{\underline{\mathbf{c}}}_k \sim \mathcal{CN}\left(\mathbf{0}, \mathrm{diag}(\boldsymbol{\lambda}_k)\right)$.

Let $\mathbf{T}_{D,k}^\circ$ be the maximizer of $R_{\rm sum} = \mathbb{E}\{\log\det(\mathbf{I}_M + \frac{1}{\sigma^2}\sum_{k=1}^K \mathbf{c}_k^H\mathbf{T}_k\mathbf{c}_k \cdot \mathbf{g}_k\mathbf{g}_k^H)\}$ for any $\mathbf{T}_{D,k} \in \mathcal{D}_k$, where $\mathcal{D}_k = \{\mathbf{T}_k | \mathbf{T}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{T}_k) \leq P_k$ and $\mathbf{T}_k$ is diagonal$\}$. For any $\mathbf{T}_{D,k} \in \mathcal{D}_k$, the first order optimality condition must hold at $\mathbf{T}_{D,k}^\circ$ as follows

$$\mathrm{tr}\left(\nabla_{\mathbf{T}_k} R_{\rm sum}\big|_{\mathbf{T}_k = \mathbf{T}_{D,k}^\circ}\left(\mathbf{T}_{D,k} - \mathbf{T}_{D,k}^\circ\right)\right) \leq 0. \tag{65}$$

Let $G_k = \frac{1}{\sigma^2}\mathbf{g}_k^H\tilde{\mathbf{A}}_k^{-1}\mathbf{g}_k$ with $\tilde{\mathbf{A}}_k = \mathbf{I}_M + \frac{1}{\sigma^2}\sum_{i\neq k}\mathbf{c}_i^H\mathbf{T}_i\mathbf{c}_i \cdot \mathbf{g}_i\mathbf{g}_i^H$. Then, $\nabla_{\mathbf{T}_k} R_{\rm sum}$ is given by

$$\nabla_{\mathbf{T}_k} R_{\rm sum} = \mathbb{E}\left\{\frac{G_k}{1 + G_k\mathbf{c}_k^H\mathbf{T}_k\mathbf{c}_k}\mathbf{c}_k\mathbf{c}_k^H\right\}. \tag{66}$$

Consequently, the inequality in (65) can be further written as

$$\begin{aligned}&\mathrm{tr}\left(\mathbb{E}\left\{\frac{G_k}{1 + G_k\mathbf{c}_k^H\mathbf{T}_{D,k}^\circ\mathbf{c}_k}\mathbf{c}_k\mathbf{c}_k^H\right\}\left(\mathbf{T}_{D,k} - \mathbf{T}_{D,k}^\circ\right)\right)\\ &= \mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H(\mathbf{T}_{D,k} - \mathbf{T}_{D,k}^\circ)\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H\mathbf{T}_{D,k}^\circ\mathbf{c}_k}\right\} \leq 0.\end{aligned} \tag{67}$$

Then, we show that $\mathbf{T}_{D,k}^\circ$ is also the maximizer of $R_{\rm sum}$ for all $\mathbf{T}_k \in \mathcal{N}_k$, where $\mathcal{N}_k = \{\mathbf{T}_k | \mathbf{T}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{T}_k) \leq P_k\}$. To explain this, we separate $\mathbf{T}_k$ as $\mathbf{T}_k = \mathbf{T}_{G,k} + \mathbf{T}_{F,k}$, where $\mathbf{T}_{G,k}$ and $\mathbf{T}_{F,k}$ contain the diagonal and off-diagonal entries in $\mathbf{T}_k$, respectively. Then, we can obtain that

$$\mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H(\mathbf{T}_k - \mathbf{T}_{D,k}^\circ)\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H\mathbf{T}_{D,k}^\circ\mathbf{c}_k}\right\}$$

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

$$= \mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H(\mathbf{T}_{G,k} - \mathbf{T}_{D,k}^{\circ})\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H \mathbf{T}_{D,k}^{\circ}\mathbf{c}_k}\right\} + \mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H \mathbf{T}_{F,k}\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H \mathbf{T}_{D,k}^{\circ}\mathbf{c}_k}\right\}. \quad (68)$$

Notice that $\mathbf{T}_{G,k} \in \mathcal{D}_k$ holds, because of $\mathrm{tr}(\mathbf{T}_k) = \mathrm{tr}(\mathbf{T}_{G,k}) \leq P_k$. Combining (67) and $\mathbf{T}_{G,k} \in \mathcal{D}_k$, we can conclude that the first term in (68) is non-positive. On the other hand, the second term in (68) can be rewritten as

$$\mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H \mathbf{T}_{F,k}\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H \mathbf{T}_{D,k}^{\circ}\mathbf{c}_k}\right\} = \sum_{p=1}^{\tilde{S}_k}\sum_{q=1,q\neq p}^{\tilde{S}_k} [\mathbf{T}_{F,k}]_{p,q} \cdot \mathbb{E}\left\{J_k\left([\mathbf{c}_k]_p, [\mathbf{c}_k]_q\right)\right\}, \quad (69)$$

where $J_k\left([\mathbf{c}_k]_p, [\mathbf{c}_k]_q\right)$ is defined as

$$J_k\left([\mathbf{c}_k]_p, [\mathbf{c}_k]_q\right) = \frac{G_k \cdot [\mathbf{c}_k]_p^* [\mathbf{c}_k]_q}{1 + G_k \cdot \sum_{i=1}^{\tilde{S}_k} [\mathbf{T}_{D,k}^{\circ}]_{i,i} \left|[\mathbf{c}_k]_i\right|^2}, \quad (70)$$

with $[\mathbf{c}_k]_p$ denoting the pth element of $\mathbf{c}_k$. From (64), $[\mathbf{c}_k]_p$ is either a constant or a CSCG random variable. Because $J_k\left([\mathbf{c}_k]_p, [\mathbf{c}_k]_q\right)$ is an odd function with respect to $[\mathbf{c}_k]_p$ or $[\mathbf{c}_k]_q$, at least one of which is a CSCG random variable, $\mathbb{E}\{J_k\left([\mathbf{c}_k]_p, [\mathbf{c}_k]_q\right)\} = 0$ must hold for any $q \neq p$, so that (69) is equal to zero. Therefore, the following inequality holds for all $\mathbf{T}_k \in \mathcal{N}_k$

$$\begin{aligned}
&\mathrm{tr}\left(\nabla_{\mathbf{T}_k} R_{\mathrm{sum}}\big|_{\mathbf{T}_k = \mathbf{T}_{D,k}^{\circ}} \left(\mathbf{T}_k - \mathbf{T}_{D,k}^{\circ}\right)\right) \\
&= \mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H(\mathbf{T}_k - \mathbf{T}_{D,k}^{\circ})\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H \mathbf{T}_{D,k}^{\circ}\mathbf{c}_k}\right\} \\
&= \mathbb{E}\left\{\frac{G_k \cdot \mathbf{c}_k^H(\mathbf{T}_{G,k} - \mathbf{T}_{D,k}^{\circ})\mathbf{c}_k}{1 + G_k \cdot \mathbf{c}_k^H \mathbf{T}_{D,k}^{\circ}\mathbf{c}_k}\right\} \leq 0. \quad (71)
\end{aligned}$$

The condition in (71) means that $\mathbf{T}_k^{\star}$ will be a diagonal matrix under the condition $\boldsymbol{\xi}_{k,0} = \mathbf{0}$. This completes the proof.

Appendix D
Proof of Theorem 4

Let $\mathbf{K}_k \in \mathbb{C}^{\tilde{S}_k \times \tilde{S}_k}$ be an arbitrary feasible matrix for the problem in (34), which means $\mathbf{K}_k \succeq \mathbf{0}$ and $\mathrm{tr}(\mathbf{K}_k) \leq P_k$. Define a set of lower-dimensional matrices as follows

$$\mathcal{W}_k(\mathbf{K}_k) = \{(1-\rho)P_k \mathbf{w}_k \mathbf{w}_k^H + \rho \mathbf{K}_k | \rho \in [0,1]\}. \quad (72)$$

By taking all the possible feasible matrix $\mathbf{K}_k$, the union of the above sets will constitute the entire feasible matrix set $\{\mathbf{T}_k | \mathbf{T}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{T}_k) \leq P_k\}$ of UT k for the problem in (34). Therefore, for each UT k, the ESR maximization over the feasible matrix set $\{\mathbf{T}_k | \mathbf{T}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{T}_k) \leq P_k\}$ can be transformed into that over the set $\mathcal{W}_k(\mathbf{K}_k)$ for any feasible matrix $\mathbf{K}_k$.

For a given feasible matrix $\mathbf{K}_k$, define the ESR over the set $\mathcal{W}_k(\mathbf{K}_k)$ as follows

$$\begin{aligned}
F_k(\rho) &= \mathbb{E}\left\{\log\det\left(\tilde{\mathbf{A}}_k + \frac{1}{\sigma^2}\mathbf{c}_k^H\left(P_k \mathbf{w}_k \mathbf{w}_k^H + \rho\left(\mathbf{K}_k - P_k \mathbf{w}_k \mathbf{w}_k^H\right)\right)\mathbf{c}_k \cdot \mathbf{g}_k \mathbf{g}_k^H\right)\right\} \\
&= \mathbb{E}\left\{\log\det\left(\tilde{\mathbf{A}}_k\right)\right\} + \mathbb{E}\left\{\log\left(1 + G_k \cdot \mathbf{c}_k^H\left(P_k \mathbf{w}_k \mathbf{w}_k^H + \rho\left(\mathbf{K}_k - P_k \mathbf{w}_k \mathbf{w}_k^H\right)\right)\mathbf{c}_k\right)\right\}, \quad (73)
\end{aligned}$$

where $0 \leq \rho \leq 1$, $\tilde{\mathbf{A}}_k = \mathbf{I}_M + \frac{1}{\sigma^2}\sum_{i \neq k} \mathbf{c}_i^H \mathbf{T}_i \mathbf{c}_i \cdot \mathbf{g}_i \mathbf{g}_i^H$ and $G_k = \frac{1}{\sigma^2}\mathbf{g}_k^H \tilde{\mathbf{A}}_k^{-1} \mathbf{g}_k$. It can be seen that $F_k(\rho)$ is a concave function of ρ [26]. A necessary condition for the optimality of $\mathbf{T}_k^{\star} = P_k \mathbf{w}_k \mathbf{w}_k$ regarding the problem in (34) is $\frac{\mathrm{d}F_k(\rho)}{\mathrm{d}\rho}\Big|_{\rho=0} \leq 0$ for any feasible matrix $\mathbf{K}_k$, which means that

$$\max_{\mathbf{K}_k \succeq \mathbf{0},\ \mathrm{tr}(\mathbf{K}_k) \leq P_k} \left(\frac{\mathrm{d}F_k(\rho)}{\mathrm{d}\rho}\Big|_{\rho=0}\right) \leq 0. \quad (74)$$

In addition, due to the concavity of $F_k(\rho)$, if (74) holds, the point $\rho = 0$ will be the global optimum of $F_k(\rho)$, for any feasible matrix $\mathbf{K}_k$. Furthermore, because $F_k(\rho)$ represents the ESR over the set $\mathcal{W}_k(\mathbf{K}_k)$, by recalling the definition of $\mathcal{W}_k(\mathbf{K}_k)$ in (72), if $\rho = 0$ is the global optimum of $F_k(\rho)$ for any feasible matrix $\mathbf{K}_k$, $\mathbf{T}_k = P_k \mathbf{w}_k \mathbf{w}_k$ will be the optimal solution to the problem in (34). Therefore, (74) is a necessary and sufficient condition when $\mathbf{T}_k = P_k \mathbf{w}_k \mathbf{w}_k$ is optimal for the problem in (34) . In (74), $\frac{\mathrm{d}F_k(\rho)}{\mathrm{d}\rho}\Big|_{\rho=0}$ is given by

$$\frac{\mathrm{d}F_k(\rho)}{\mathrm{d}\rho}\Big|_{\rho=0} = \mathbb{E}\left\{\frac{G_k \mathbf{c}_k^H \mathbf{K}_k \mathbf{c}_k - G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}{1 + G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}\right\}. \quad (75)$$

Let $\mathbf{K}_k = \sum_{i=1}^{\tilde{S}_k} \xi_i \mathbf{p}_i \mathbf{p}_i^H$ denote the EVD of $\mathbf{K}_k$, where $\{\xi_i\}_{i=1}^{\tilde{S}_k}$ and $\{\mathbf{p}_i\}_{i=1}^{\tilde{S}_k}$ are the non-negative eigenvalues and the corresponding orthogonal eigenvectors, respectively. Then, the condition in (74) can be further written as

$$\begin{aligned}
0 &\geq \max_{\xi_i,\ \mathbf{p}_i,\ \forall i} \mathbb{E}\left\{\frac{\sum_{i=1}^{\tilde{S}_k} \xi_i G_k \mathbf{c}_k^H \mathbf{p}_i \mathbf{p}_i^H \mathbf{c}_k - G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}{1 + G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}\right\} \\
&= \max_{\xi_i,\ \mathbf{p}_i,\ \forall i} \sum_{i=1}^{\tilde{S}_k} \xi_i \mathbf{p}_i^H \mathbb{E}\left\{\frac{G_k \mathbf{c}_k \mathbf{c}_k^H}{1 + G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}\right\} \mathbf{p}_i - \mathbb{E}\left\{\frac{G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}{1 + G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}\right\} \\
&\overset{(a)}{=} P_k \left(\Upsilon_{\max}\left(\mathbb{E}\left\{\frac{G_k \mathbf{c}_k \mathbf{c}_k^H}{1 + G_k P_k |\mathbf{c}_k^H \mathbf{w}_k|^2}\right\}\right)\right.
\end{aligned}$$

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

$$-\mathbb{E}\left\{\frac{G_k|\mathbf{c}_k^H\mathbf{w}_k|^2}{1+G_kP_k|\mathbf{c}_k^H\mathbf{w}_k|^2}\right\}\Bigg), \tag{76}$$

where (a) follows from $\mathbf{p}_i^H\mathbf{X}\mathbf{p}_i \leq \Upsilon_{\max}(\mathbf{X})$ and $\sum_{i=1}^{\tilde{S}_k}\xi_i \leq P_k$. Thus, the condition in (76) can be written as

$$\Upsilon_{\max}\left(\mathbb{E}\left\{\frac{G_k\mathbf{c}_k\mathbf{c}_k^H}{1+G_kP_k|\mathbf{c}_k^H\mathbf{w}_k|^2}\right\}\right) \leq \mathbb{E}\left\{\frac{G_k|\mathbf{c}_k^H\mathbf{w}_k|^2}{1+G_kP_k|\mathbf{c}_k^H\mathbf{w}_k|^2}\right\}. \tag{77}$$

In addition, since $\mathbf{w}_k$ is a unit-norm vector, we have

$$\begin{aligned}&\mathbb{E}\left\{\frac{G_k|\mathbf{c}_k^H\mathbf{w}_k|^2}{1+G_kP_k|\mathbf{c}_k^H\mathbf{w}_k|^2}\right\}\\ &=\mathbf{w}_k^H\mathbb{E}\left\{\frac{G_k\mathbf{c}_k\mathbf{c}_k^H}{1+G_kP_k|\mathbf{c}_k^H\mathbf{w}_k|^2}\right\}\mathbf{w}_k\\ &\leq\Upsilon_{\max}\left(\mathbb{E}\left\{\frac{G_k\mathbf{c}_k\mathbf{c}_k^H}{1+G_kP_k|\mathbf{c}_k^H\mathbf{w}_k|^2}\right\}\right).\end{aligned} \tag{78}$$

Consequently, by combining the results in (77) and (78), the condition in (36) can be obtained. This completes the proof.

REFERENCES

[1] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, J. Querol, L. Lei, T. X. Vu, and G. Goussetis, "Satellite communications in the new space era: A survey and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 70–109, 1st Quart 2021.

[2] 3GPP, "Study on new radio (NR) to support non-terrestrial networks (Release 15)," TR 38.811, V15.3.0, July 2020.

[3] B. Di, L. Song, Y. Li, and H. V. Poor, "Ultra-dense LEO: Integration of satellite access networks into 5G and beyond," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 62–69, Apr. 2019.

[4] Z. Xiao, J. Yang, T. Mao, C. Xu, R. Zhang, Z. Han, and X.-G. Xia, "LEO satellite access network (LEO-SAN) towards 6G: Challenges and approaches," *IEEE Wireless Commun.*, pp. 1–8, 2022, early access.

[5] S. Liu, Z. Gao, Y. Wu, D. W. K. Ng, X. Q. Gao, K.-K. Wong, S. Chatzinotas, and B. Ottersten, "LEO satellite constellations for 5G and beyond: How will they reshape vertical domains?" *IEEE Commun. Mag.*, vol. 59, no. 7, pp. 30–36, July 2021.

[6] D. Christopoulos, S. Chatzinotas, M. Matthaiou, and B. Ottersten, "Capacity analysis of multibeam joint decoding over composite satellite channels," in *Proc. IEEE ASILOMAR*, Pacific Grove, CA, USA, Nov. 2011, pp. 1795–1799.

[7] Y. Yang, X. Q. Gao, and X.-G. Xia, "A closed-form capacity upper bound of multibeam GEO MSC uplink channel," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 576–579, Dec. 2016.

[8] J. Arnau, D. Christopoulos, S. Chatzinotas, C. Mosquera, and B. Ottersten, "Performance of the multibeam satellite return link with correlated rain attenuation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6286–6299, Nov. 2014.

[9] Y. Couble, C. Rosenberg, E. Chaput, J.-B. Dupé, C. Baudoin, and A.-L. Beylot, "Two-color scheme for a multi-beam satellite return link: Impact of interference coordination," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 5, pp. 993–1003, May 2018.

[10] K. Guo, K. An, B. Zhang, Y. Huang, D. Guo, G. Zheng, and S. Chatzinotas, "On the performance of the uplink satellite multiterrestrial relay networks with hardware impairments and interference," *IEEE Syst. J.*, vol. 13, no. 3, pp. 2297–2308, Sep. 2019.

[11] C.-X. Wang, X. You, X. Q. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2023, early access.

[12] L. You, K.-X. Li, J. Wang, X. Q. Gao, X.-G. Xia, and B. Ottersten, "Massive MIMO transmission for LEO satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1851–1865, Aug. 2020.

[13] P. Angeletti and R. De Gaudenzi, "A pragmatic approach to massive MIMO for broadband communication satellites," *IEEE Access*, vol. 8, pp. 132 212–132 236, 2020.

[14] K.-X. Li, L. You, J. Wang, X. Q. Gao, C. G. Tsinos, S. Chatzinotas, and B. Ottersten, "Downlink transmit design for massive MIMO LEO satellite communications," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1014–1028, Feb. 2022.

[15] M. Röper, B. Matthiesen, D. Wübben, P. Popovski, and A. Dekorsy, "Beamspace MIMO for satellite swarms," in *Proc. IEEE WCNC*, Austin, TX, USA, Apr. 2022, pp. 1307–1312.

[16] M. Y. Abdelsadek, G. K. Kurt, and H. Yanikomeroglu, "Distributed massive MIMO for LEO satellite networks," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2162–2177, 2022.

[17] M. Y. Abdelsadek, G. Karabulut-Kurt, H. Yanikomeroglu, P. Hu, G. Lamontagne, and K. Ahmed, "Broadband connectivity for handheld devices via LEO satellites: Is distributed massive MIMO the answer?" *IEEE Open J. Commun. Soc.*, vol. 4, pp. 713–726, 2023.

[18] B. Shen, Y. Wu, J. An, C. Xing, L. Zhao, and W. Zhang, "Random access with massive MIMO-OTFS in LEO satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2865–2881, Oct. 2022.

[19] X. Zhou, K. Ying, Z. Gao, Y. Wu, Z. Xiao, S. Chatzinotas, J. Yuan, and B. Ottersten, "Active terminal identification, channel estimation, and signal detection for grant-free NOMA-OTFS in LEO satellite Internet-of-Things," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2847–2866, Apr. 2023.

[20] H. Chougrani, S. Kisseleff, W. A. Martins, and S. Chatzinotas, "NB-IoT random access for nonterrestrial networks: Preamble detection and uplink synchronization," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14 913–14 927, Aug. 2022.

[21] L. You, X. Qiang, K.-X. Li, C. G. Tsinos, W. Wang, X. Q. Gao, and B. Ottersten, "Hybrid analog/digital precoding for downlink massive MIMO LEO satellite communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 5962–5976, Aug. 2022.

[22] Z. Gao, A. Liu, C. Han, and X. Liang, "Sum rate maximization of massive MIMO NOMA in LEO satellite communication system," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1667–1671, Aug. 2021.

[23] G. Maral, M. Bousquet, and Z. Sun, *Satellite Communications Systems: Systems, Techniques and Technology*, 6th ed. Chichester, UK: Wiley, 2020.

[24] Y. Zuo, M. Yue, M. Zhang, S. Li, S. Ni, and X. Yuan, "OFDM-based massive connectivity for LEO satellite Internet of Things," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2023, early access.

[25] G. Barriac and U. Madhow, "Space-time precoding for mean and covariance feedback: application to wideband OFDM," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 96–107, Jan. 2006.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[27] X. Li, S. Jin, M. R. McKay, X. Q. Gao, and K.-K. Wong, "Capacity of MIMO-MAC with transmit channel knowledge in the low

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3361885

Accepted Manuscript - Version of Record at: https://doi.org/10.1109/TAES.2024.3361885

SNR regime," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 926–931, Mar. 2010.

[28] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge Univ. Press, 2013.

[29] D. P. Bertsekas, *Convex Optimization Algorithms*. Belmont, MA, USA: Athena Scientific, 2015.

[30] C.-K. Wen, S. Jin, and K.-K. Wong, "On the sum-rate of multiuser MIMO uplink channels with jointly-correlated Rician fading," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2883–2895, Oct. 2011.

[31] A.-A. Lu, X. Q. Gao, and C. Xiao, "Free deterministic equivalents for the analysis of MIMO multiple access channel," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4604–4629, Aug. 2016.

[32] K. Xu, J. Zhang, X. Yang, S. Ma, and G. Yang, "On the sum-rate of RIS-assisted MIMO multiple-access channels over spatially correlated Rician fading," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8228–8241, Dec. 2021.

[33] L. You, J. Xiong, D. W. K. Ng, C. Yuen, W. Wang, and X. Q. Gao, "Energy efficiency and spectral efficiency tradeoff in RIS-aided multiuser MIMO uplink transmission," *IEEE Trans. Signal Process.*, vol. 69, pp. 1407–1421, 2021.

[34] F. Zhang, *Matrix Theory: Basic Results and Techniques*, 2nd ed. New York, NY, USA: Spinger, 2011.

[35] J. Groß, "Nonnegative-definite and positive-definite solutions to the matrix equation $\mathbf{AXA}^* = \mathbf{B}$ - revisited," *Linear Algebra Appl.*, vol. 321, pp. 123–129, 2000.

[36] W. Wu, X. Q. Gao, C. Sun, and G. Y. Li, "Shallow underwater acoustic massive MIMO communications," *IEEE Trans. Signal Process.*, vol. 69, pp. 1124–1139, 2021.

[37] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd ed. New York, NY, USA: Springer-Verlag, 2003.

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.