Towards Optimized Network Capacity in Emerging Integrated Terrestrial-Satellite Networks

Shulei Gong, Hong Shen, Kanglian Zhao, Wenfeng Li, Haibo Zhou, Ruhai Wang, Zhili Sun, and Xinggan Zhang

Abstract—In this paper, we investigate the transmission schemes of space data systems for optimized network capacity in an integrated terrestrial-satellite network (ITSN) with a two-layered space segment. First, a theoretical model of the network capacity is developed to evaluate the strategy of utilizing both direct and relayed transmissions. Second, we consider the ideal and the resource-constrained scenarios in which the corresponding network capacity is modeled with respect to the scheduling scheme. In particular, closed form and semi-closed form solutions to the difficult integer programs are achieved via rigorous mathematical analysis. The proposed model is general for exploring the capacity of various satellite network deployments whose solutions have not been obtained in prior studies. Furthermore, we verify the potential capacity of the different transmission schemes based on the proposed solutions and prove that the system's network capacity can be significantly improved by the hybrid transmission scheme. The theoretical framework proposed in this paper is expected to provide constructive insights in the design for the future space segments of ITSN.

Index Terms—Satellite network capacity, hybrid transmission, scheduling optimization, integer program.

I. INTRODUCTION

Space data systems (SDSs) are used for data acquisition, transferring and processing in various applications including earth observation, scientific space exploration and data collection. Traditionally, SDSs are operated with dedicated ground stations (GSs) for telemetry tracking command (TT & C) and data transmission, which means that they can only be operated when they are in line-of-sight of the GS. With limited number of ground infrastructures, most of the SDSs, even dedicated for data collection and communications, such as the Orbcomm constellation, can only work in a store-and-forward mode for onboard data acquisition.

With the fast and continuous improvement in the capabilities of payload, more and more data are generated or relayed from

a single satellite. Meanwhile, the number of satellites is also increasing quickly [1]. It is getting more and more difficult to acquire the huge amount of data in near real time with a limited number of GSs. On-board processing, federated GS networks and satellite relaying are several potential solutions. For instance, federated GS network has been used to improve the performance of information acquisition from CubeSat [2].

Satellite downlink scheduling is a very common problem [3] [4], and it faces a tough challenge due to the growing need to download the ever-increasing data. The goal of satellite downlink scheduling problem is to maximize the amount of data downloaded from the satellites during a given time period. There are various practical constraints such as the opportunities to communicate with a GS, the limited on-board data, power and thermal control, coordination among the satellites, and so forth [5], which make the scheduling more complicated. In the literature, greedy algorithm [6], multi-phase approach [7], heuristic approaches [5] [8] and genetic algorithms [9] [10] can be used to solve the SDS scheduling problem.

Generally, satellites can directly deliver the data to the GSs or transfer the data with the aid of relayed satellites. The direct transmission approach was studied in previous works such as [11]–[15]. With the internetworked GSs and/or the dedicated space relay systems, such as Tracking and Data Relay Satellite System (TDRSS) [16], [17], Tianlian System [18], and European Data Relay Satellite (EDRS) system [19], more and more SDSs are now operated as part of an integrated space-terrestrial network. Then, a natural question arises-how to combine the direct and relayed transmissions in order to achieve higher network capacity. In this paper, we aim to address this problem under a classical two-layer satellite network scenario with geosynchronous earth orbit (GEO) relay satellites and low earth orbit (LEO) SDSs. Considering that the onboard data in SDSs can be downloaded to the GSs directly or relayed by GEO satellites, we propose a hybrid transmission scheduling framework. We then develop an analytical model of the hybrid transmission scheme, which is applicable to various network deployments. Furthermore, we optimize the network capacity under ideal and resource-constrained scenarios. The capacity optimization problems under both scenarios belong to integer programming that is difficult to be solved in general. Nevertheless, by performing theoretical analysis, we obtain closed form and semi-closed form optimal solutions, which, to the best of our knowledge, have not been achieved in literature. Finally, these solutions are utilized to verify the potential capacities of different transmission schemes, and it is observed that the system capacity is significantly improved

This paper is dedicated to the memory of Professor Naitong Zhang. This work was supported by the National Natural Science Foundation of China under Grant 61871108, the Fundamental Research Funds for the Central Universities under Grant 2242018K40037, and the Natural Science Foundation of Jiangsu Province under Grant BK20150635.

S. Gong, K. Zhao, W. Li, H. Zhou, and X. Zhang are with the School of Electronic Science and Engineering, Nanjing University, Nanjing, 210093, China (e-mail: gongshulei@smail.nju.edu.cn; zhaokanglian@nju.edu.cn; leewf_cn@hotmail.com; haibozhou@nju.edu.cn; zhxg@nju.edu.cn). S. Gong is also with the China Mobile Group Jiangsu Co., Ltd, Nanjing, 210029, China. H. Shen is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (email: shhseu@seu.edu.cn). R. Wang is with the Phillip M. Drayer Department of Electrical Engineering, Lamar University, Beaumont, TX 77710, USA (email: rwang@lamar.edu). Z. Sun is with the Institute of Communication Systems, University of Surrey, Guildford GU2 7XH, U.K (email: z.sun@surrey.ac.uk). (Corresponding authors: Ruhai Wang and Xinggan Zhang.)

by the proposed hybrid transmission scheme.

The remainder of this paper is structured as follows. Section II summarizes the prior related works. Section III illustrates the system model as well as the definition of system capacity. In Section IV, the capacity maximization problems and the corresponding solutions are presented. Section V provides the simulation results, and the conclusions are drawn in Section VI.

II. RELATED WORK

There have been a number of prior works studying the scheduling optimization for satellite networks. Most of them focused on the direct transmission scheme, i.e., the satellites communicate with the GSs directly. Specifically, the authors of [11] investigated the single-satellite multi-GS communication scheduling problem, which aims at maximizing the data downloaded from the satellite. The problem was formulated as a constrained mixed integer program (MIP), based on which the cases of linear and non-linear dynamics were both considered. Concerning the difficult MIP, the authors applied the off-the-shelf optimization software package to find its optimal solution. Furthermore, the multi-satellite multi-GS scheduling optimization was concerned in [12], whose design objective is to maximize the total data amount downloaded from the multiple satellites. A MIP was also formulated for this problem and solved with the optimization software package. In [13], the authors dealt with the problem of minimizing the total cost associated with the pairings between the GSs and satellites, where they proposed two formulations, i.e., linear two-index assignment and axial three-index assignment. The coupled selection equations were adopted to find near-optimal scheduling solutions. The scheduling optimization for the direct transmission was also studied in [7], [8], where near-optimal solutions were found based on a Lagrangian heuristic method and hybrid ant colony optimization combined with local search, respectively. Regarding the relay transmission scheme, the authors of [20] considered maximizing the profit-weighted number of scheduled missions and minimizing the total setup time, and they developed a two-stage heuristic algorithm for the scheduling problem.

To summarize, the aforementioned works studied the scheduling design for only direct or relay transmission, which can be further improved by the proposed hybrid transmission strategy. Moreover, the scheduling optimization problems were solved via the software package or near-optimal numerical algorithms in these works, while we obtain the optimal solutions in closed form or semi-closed form in our work.

III. SYSTEM MODEL

Considering an integrated terrestrial-satellite system with M GEO satellites, N LEO SDSs and K GSs, as depicted in Fig. 1. Each GEO satellite has its own GS. When an SDS is right above a GS, it can deliver the data directly to the station or transfer the data to the relaying GEO satellite. More than one GEO satellites or GSs may be in view of an SDS concurrently, but the SDS can only communicate with one GEO satellite or GS at a time. This assumption makes sense since

Fig. 1. An integrated terrestrial-satellite system consisting of GEO satellites, LEO SDSs and GSs.

Fig. 2. LEO-GEO, LEO-GS and GEO-GS communication links.

it is appropriate to deploy one set of communication device on the SDS with size, weight and power (SWaP) constraints. We consider the GS networks and the GEO relaying satellites as the infrastructure for various SDSs and the SDSs differ from each other. Hence, it is reasonable to assume that there are no inter-satellite links (ISLs) between adjacent SDSs. If we consider the ISLs between adjacent SDSs and assume that SDS can communicate with more than one GEO satellite or GS at a time, the resultant capacity maximization problem will be more complicated and worthwhile further studying in the ongoing research.

A. Communication Links

Communication links are categorized by ISLs and satellite-ground links. ISLs include the links in the same layer and the links between different layers. Focusing on the relaying GEO satellites, we only consider ISLs between GEO layer and LEO layer denoted by $LGLs_{i,n,m} \in \{0,1\}$, $1 \leq i \leq I$, $1 \leq n \leq N$, $1 \leq m \leq M$, where I is the number of time intervals. If SDS n is in the view of GEO satellite m during interval i , $LGLs_{i,n,m} = 1$. Otherwise, $LGLs_{i,n,m} = 0$. There exist two kinds of satellite-ground links, namely LEO-GS links (denoted by $LSLs_{i,n,k} \in \{0,1\}$, $1 \leq i \leq I$, $1 \leq n \leq N$, $1 \leq k \leq K$) and GEO-GS links (denoted by $GSLs_{i,m} \in \{0,1\}$, $1 \leq i \leq I$, $1 \leq m \leq M$). In this case, $\{LGLs_{i,n,m}, LSLs_{i,n,k}, GSLs_{i,m}\}$ constitutes the system communication link state table Θ , as shown in Fig. 2. It is further assumed that the satellite orbits and the locations of GSs are deterministic and known, i.e., we have prior knowledge of Θ .

B. Network Capacity

We define network capacity C as the total amount of data downloaded from LEO SDSs during a given time period

$$C = \sum_{n=1}^N (C_n^L + C_n^G), \quad (1)$$

where C_n^L represents the capacity of the direct link, i.e., the direct-to-Earth link, and C_n^G represents the capacity of the relay link, i.e., the relaying link with the GEO satellite as a relay. Let $\phi_{i,n,k}$ and $\phi_{i,n,m}$ denote the data transfer rates of the direct link and relay link during interval i . Then, the amounts of data downloaded from SDS n on these two links are

$$q_{i,n,k}^L = \Delta t_i \phi_{i,n,k} \cdot \text{LSLs}_{i,n,k} \cdot \chi_{i,n,k}, \quad (2)$$

$$q_{i,n,m}^G = \Delta t_i \phi_{i,n,m} \cdot \text{LGLs}_{i,n,m} \cdot \text{GSLs}_{i,m} \cdot \xi_{i,n,m} \cdot \tau_{i,m}, \quad (3)$$

where Δt_i is the duration of interval i , measured in minutes. $\chi_{i,n,k}$, $\xi_{i,n,m}$ and $\tau_{i,m} \in \{0, 1\}$ are the binary values respectively representing the decisions of downloading data during interval i . Specifically, if the decision is to transfer from SDS n to GS k , $\chi_{i,n,k} = 1$, otherwise, $\chi_{i,n,k} = 0$. Assume that GEO satellite is always connected with its GS, if SDS n chooses GEO satellite m , then $\xi_{i,n,m} = \tau_{i,m} = 1$. Otherwise, $\xi_{i,n,m} = \tau_{i,m} = 0$. Hence the relay link capacity $q_{i,n,m}^G$ can be simplified into $\Delta t_i \phi_{i,n,m} \cdot \text{LGLs}_{i,n,m} \cdot \text{GSLs}_{i,m} \cdot \xi_{i,n,m}$.

Let $\eta_{i,n,k} \in [0, 1)$ and $\varepsilon_{i,n,m} \in [0, 1)$ denote the transfer efficiency of the direct and relay links, respectively. In particular, $\eta_{i,n,k} = 0$ and $\varepsilon_{i,n,m} = 0$ mean that the direct and relay links are disconnected due to the atmospheric effects such as cloud blockage. Then, we have

$$C_n^L = \sum_{i=1}^I \sum_{k=1}^K \eta_{i,n,k} q_{i,n,k}^L \triangleq \sum_{i=1}^I \sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k}, \quad (4)$$

$$C_n^G = \sum_{i=1}^I \sum_{m=1}^M \varepsilon_{i,n,m} q_{i,n,m}^G \triangleq \sum_{i=1}^I \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m}, \quad (5)$$

where $\alpha_{i,n,k} = \eta_{i,n,k} \Delta t_i \phi_{i,n,k} \cdot \text{LSLs}_{i,n,k}$ and $\beta_{i,n,m} = \varepsilon_{i,n,m} \Delta t_i \phi_{i,n,m} \cdot \text{LGLs}_{i,n,m} \cdot \text{GSLs}_{i,m}$.

IV. CAPACITY OPTIMIZATION FOR A TWO-LAYER SATELLITE NETWORK

In order to download the on-board data of LEO satellites efficiently, we need to carefully investigate the choice of specific GEO satellite or GS. Our primal goal is to determine a feasible decision table $\{\chi_{i,n,k}, \xi_{i,n,m}\}$ that maximizes the system capacity during a certain period to promote user experience. More specifically, in Section IV-A, the capacity based scheduling optimization with respect to GEO satellites and GSs is investigated in a relatively ideal scenario. In this scenario, all the nodes in the system have unlimited energy, SDSs have sufficient data for transmission, and GEO satellites and GSs can access an arbitrary number of SDSs. In Section IV-B, we study a more practical scenario where the number of LEO satellites accessed by GEO satellites and GSs is limited. Finally, in Section IV-C, we further include the constraint that characterizes the limited data on SDSs.

A. Ideal Scenario

In this scenario, we assume that all the satellites and GSs have unlimited energy and there are abundant on-board data for transfer all the time. GEO satellites and GSs have the ability to access any number of SDSs. It is very likely that more than one GEO satellite or GS is in view of one SDS, but one SDS can only contact with one GEO satellite or GS at a time. The capacity optimization problem under the above scenario is given by

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} \sum_{i=1}^I \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \quad (6a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad (6b)$$

$$\chi_{i,n,k} \in \{0, 1\}, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (6c)$$

$$\xi_{i,n,m} \in \{0, 1\}, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M. \quad (6d)$$

The problem in (6) belongs to linear integer programming (LIP) [21]. Constraint (6b) ensures that the SDS can choose at most one GEO satellite or GS to transfer data. We can readily show that the optimal scheduling strategy in the following theorem.

Theorem 1: Define $\mathbf{u}_n = \{\alpha_{i,n,1}, \dots, \alpha_{i,n,K}, \beta_{i,n,1}, \dots, \beta_{i,n,M}\}$, $\hat{u}_n = \max\{\mathbf{u}_n\}$ and $\omega = \text{argmax}\{\mathbf{u}_n\}$ for each n , where $\max\{\cdot\}$ and $\text{argmax}\{\cdot\}$ return the maximum value of the input vector and the corresponding index, respectively. The optimal solution to problem (6) is

$$\begin{cases} \hat{\chi}_{i,n,\omega} = 1, \quad \hat{\chi}_{i,n,k \neq \omega} = \hat{\xi}_{i,n,m} = 0, & \omega \leq K \\ \hat{\xi}_{i,n,\omega-K} = 1, \quad \hat{\chi}_{i,n,k} = \hat{\xi}_{i,n,m \neq \omega-K} = 0, & \omega > K. \end{cases} \quad (7)$$

Proof: According to the constraints of problem (6), at most one $\chi_{i,n,k}$ or $\xi_{i,n,m}$ equals to 1 for given i and n . Moreover, all the coefficients in the objective function are non-negative. Therefore, it is readily to verify that the solution in (7) is optimal since it is feasible and also maximizes the objective. ■

Clearly, the optimal scheduling scheme is to choose the largest amount of data directly transferred to GSs or relayed by GEO satellites for each SDS during each interval. It is worthwhile noting that the above solution has a closed form, which can be calculated easily. We also need to point out that studying the ideal scenario enables us to quantify the potential capacity of the system, which is beneficial to later analysis.

B. Resource Constrained Scenario With Limited SDS Access Capability

The previous section describes the potential maximum capacity of the system. In this subsection, we take resource constraints into account to characterize the practical system more accurately. With increasing the number of satellites, the probability of overlapping grows, i.e., multiple satellites

are in view of a single GEO satellite or GS. The overlap effects reduce the total available line of sight opportunities in consequence. Let $\rho_{GS}^{\max}$ and $\rho_{GEO}^{\max}$ denote the maximum numbers of SDSs connected to each GS and each GEO satellite, respectively. Then, we obtain the newly added constraints by

$$\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{GS}^{\max}, \quad 1 \leq i \leq I, \quad 1 \leq k \leq K, \quad (8)$$

$$\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{GEO}^{\max}, \quad 1 \leq i \leq I, \quad 1 \leq m \leq M. \quad (9)$$

Therefore, the capacity optimization problem under the current resource constrained scenario is given by

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} \sum_{i=1}^I \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \quad (10a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad (10b)$$

$$\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{GS}^{\max}, \quad 1 \leq i \leq I, \quad 1 \leq k \leq K, \quad (10c)$$

$$\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{GEO}^{\max}, \quad 1 \leq i \leq I, \quad 1 \leq m \leq M, \quad (10d)$$

$$\chi_{i,n,k} \in \{0, 1\}, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (10e)$$

$$\xi_{i,n,m} \in \{0, 1\}, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M, \quad (10f)$$

which also belongs to LIP. Since the optimization variables corresponding to different time intervals are decoupled, we can decompose problem (10) into I independent subproblems without loss of optimality. The i th subproblem is

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \quad (11a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (11b)$$

$$\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{GS}^{\max}, \quad 1 \leq k \leq K, \quad (11c)$$

$$\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{GEO}^{\max}, \quad 1 \leq m \leq M, \quad (11d)$$

$$\chi_{i,n,k} \in \{0, 1\}, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (11e)$$

$$\xi_{i,n,m} \in \{0, 1\}, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M. \quad (11f)$$

Different from the ideal scenario, the above problem does not admit a simple closed-form solution due to the newly added constraints. Nonetheless, we are able to derive a semi-closed form solution to problem (11) as follows.

Theorem 2: Define $\mathbf{r}_n = \{\alpha_{i,n,1} - \hat{\mu}_{i,1}, \dots, \alpha_{i,n,K} - \hat{\mu}_{i,K}, \beta_{i,n,1} - \hat{\nu}_{i,1}, \dots, \beta_{i,n,M} - \hat{\nu}_{i,M}\}$, $\hat{r}_n = \max\{\mathbf{r}_n\}$ and $\zeta = \arg\max\{\mathbf{r}_n\}$ for each n , where $\hat{\mu}_{i,k}$ and $\hat{\nu}_{i,m}$ are the

optimal dual variables associated with the constraints (11c) and (11d), respectively. Then, the optimal solution to problem (11) is expressed by

$$\begin{cases} \hat{\chi}_{i,n,k} = \hat{\xi}_{i,n,m} = 0, & \hat{r}_n \leq 0 \\ \hat{\chi}_{i,n,\omega} = 1, \quad \hat{\chi}_{i,n,k \neq \omega} = \hat{\xi}_{i,n,m} = 0, & \hat{r}_n > 0 \text{ and } \omega \leq K \\ \hat{\xi}_{i,n,\omega-K} = 1, \quad \hat{\chi}_{i,n,k} = \hat{\xi}_{i,n,m \neq \omega-K} = 0, & \hat{r}_n > 0 \text{ and } \omega > K. \end{cases} \quad (12)$$

The optimal dual variables can be determined via the subgradient algorithm where the update rule is given by

$$\mu_{i,k}(t+1) = \left[\mu_{i,k}(t) - \delta_{\mu}(t) \left(\rho_{GS}^{\max} - \sum_{n=1}^N \chi_{i,n,k}(t) \right) \right]^+, \quad (13)$$

$$\nu_{i,m}(t+1) = \left[\nu_{i,m}(t) - \delta_{\nu}(t) \left(\rho_{GEO}^{\max} - \sum_{m=1}^M \xi_{i,n,m}(t) \right) \right]^+, \quad (14)$$

where $[x]^+$ returns the maximum of x and zero, t represents the iteration index, and $\delta_{\mu}(t) > 0$ and $\delta_{\nu}(t) > 0$ denote the step sizes for updating $\mu_{i,k}$ and $\nu_{i,m}$, respectively.

Proof: See Appendix A. ■

It is worthwhile noting that, compared to the ideal scenario, the optimal scheduling scheme under the considered resource constrained scenario also depends on the optimal dual variables associated with the constraints (11c) and (11d), which can be efficiently calculated.

C. Resource Constrained Scenario With Limited Data on SDSs

When the SDS orbits the Earth, the amount of on-board data varies constantly. Regarding this fact, we aim to maximize system capacity by considering the data dynamics on SDSs, as well as the constraints introduced in Section IV-B. We model the data amount on the n th SDS at the end of interval i by

$$d_{i+1,n} = d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N, \quad (15)$$

where $\delta_{i,n}^b$ represents the burst data amount on SDS n during interval i and $\delta_{i,n}^l$ represents the total amount of lost data on SDS n during interval i caused by data expiration, etc. Both $\delta_{i,n}^b$ and $\delta_{i,n}^l$ are constants. Denote the minimum and maximum amounts of on-board data that can be stored on SDSs by $d_{\min}$ and $d_{\max}$, respectively. Then, we have

$$d_{\min} \leq d_{i,n} \leq d_{\max}, \quad 1 \leq i \leq I, \quad 1 \leq n \leq N. \quad (16)$$

By including constraints (15) and (16), and also the constraints presented in subsections IV-A and IV-B, we formulate the

TABLE I
PARAMETERS OF SATELLITE CONSTELLATIONS

Constellations	Number of satellites	Number of planes	Number of satellites per plane	Altitude	Inclination
GEO relays	3	1	3	36000 km	0°
SDSs	64	8	8	822 km	98.7°

capacity maximization problem during the i -th interval by

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \quad (17a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (17b)$$

$$\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{\text{GS}}^{\max}, \quad 1 \leq k \leq K, \quad (17c)$$

$$\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{\text{GEO}}^{\max}, \quad 1 \leq m \leq M, \quad (17d)$$

$$d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l \geq d_{\min}, \quad 1 \leq n \leq N, \quad (17e)$$

$$d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l \leq d_{\max}, \quad 1 \leq n \leq N, \quad (17f)$$

$$\chi_{i,n,k} \in \{0, 1\}, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (17g)$$

$$\xi_{i,n,m} \in \{0, 1\}, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M. \quad (17h)$$

The constraints on the data amount make the problem more complicated. However, by exploiting a similar approach in Appendix A, we are still able to achieve a semi-closed form solution as follows.

Theorem 3: Define $\mathbf{q}_n = \{(1 - \hat{\lambda}_{i,n} + \hat{\kappa}_{i,n})\alpha_{i,n,1} - \hat{\mu}_{i,1}, \dots, (1 - \hat{\lambda}_{i,n} + \hat{\kappa}_{i,n})\alpha_{i,n,K} - \hat{\mu}_{i,K}, (1 - \hat{\lambda}_{i,n} + \hat{\kappa}_{i,n})\beta_{i,n,1} - \hat{\nu}_{i,1}, \dots, (1 - \hat{\lambda}_{i,n} + \hat{\kappa}_{i,n})\beta_{i,n,M} - \hat{\nu}_{i,M}\}$, $\hat{q}_n = \max\{\mathbf{q}_n\}$ and $\gamma = \arg\max\{\mathbf{q}_n\}$ for each n , where $\hat{\mu}_{i,k}$, $\hat{\nu}_{i,m}$, $\hat{\lambda}_{i,n}$ and $\hat{\kappa}_{i,n}$ are the optimal dual variables associated with the constraints (17c), (17d), (17e) and (17f), respectively. Then, the optimal solution to problem (17) is

$$\begin{cases} \hat{\chi}_{i,n,k} = \hat{\xi}_{i,n,m} = 0, & \hat{q}_n \leq 0 \\ \hat{\chi}_{i,n,\gamma} = 1, \hat{\chi}_{i,n,k \neq \gamma} = \hat{\xi}_{i,n,m} = 0, & \hat{q}_n > 0 \text{ and } \gamma \leq K \\ \hat{\xi}_{i,n,\gamma-K} = 1, \hat{\xi}_{i,n,k \neq \gamma-K} = \hat{\chi}_{i,n,k} = 0, & \hat{q}_n > 0 \text{ and } \gamma > K. \end{cases} \quad (18)$$

The optimal dual variables can be determined via the subgra-

dient algorithm where the update rule is given by

$$\mu_{i,k}(t+1) = \left[\mu_{i,k}(t) - \delta_{\mu}(t) \left(\rho_{\text{GS}}^{\max} - \sum_{n=1}^N \chi_{i,n,k}(t) \right) \right]^+, \quad (19)$$

$$\nu_{i,m}(t+1) = \left[\nu_{i,m}(t) - \delta_{\nu}(t) \left(\rho_{\text{GEO}}^{\max} - \sum_{n=1}^M \xi_{i,n,m}(t) \right) \right]^+, \quad (20)$$

$$\lambda_{i,n}(t+1) = \left[\lambda_{i,n}(t) - \delta_{\lambda}(t) \left(d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l - d_{\min} \right) \right]^+, \quad (21)$$

$$\kappa_{i,n}(t+1) = \left[\kappa_{i,n}(t) - \delta_{\kappa}(t) \left(d_{\max} + \delta_{i,n}^l - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} - d_{i,n} - \delta_{i,n}^b \right) \right) \right]^+, \quad (22)$$

where $\delta_{\lambda}(t) > 0$ and $\delta_{\kappa}(t) > 0$ denote the step sizes for updating $\lambda_{i,n}$ and $\kappa_{i,n}$, respectively.

Proof: See Appendix B. ■

From the above theorem, we find that the optimal scheduling solution depends on the optimal dual variables associated with the constraints (17c)–(17f), which is analogous to the conclusion drawn in Theorem 2.

Remark 1: The proposed hybrid transmission scheme adaptively schedules the direct and relaying transmission schemes, both of which have already been applied in real systems. Furthermore, the proposed semi-closed form scheduling solution has low computational complexity and can be readily implemented.

Remark 2: We note that the delay introduced by GEO satellite does not impact our obtained results since the above considered capacity maximization problems are irrelevant to the delay.

V. PERFORMANCE ANALYSIS

We construct a simple scenario featuring a typical two-layer satellite network by integrating evenly distributed GEO relaying satellites and a Walker constellation of LEO SDSs. Detailed configurations of these constellations are summarized in Table I. Other parameters of the network are listed in Table II, where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 . The 16 GSs are distributed on the earth in a random manner. We conduct performance evaluations for direct transmission, relay transmission and hybrid

TABLE II
PARAMETERS OF THE NETWORK

Parameters	Value
Number of GSs	16
Simulation time	27 days
Data rate (Mbps)	$\phi_{i,n,m} \sim N(5, 1)$, $\phi_{i,n,k} \sim N(30, 2)$
Downloading efficiency	$\eta_{i,n,m} = 0.95$, $\eta_{i,n,k} = 0.99$

TABLE III
CPU TIME COMPARISON ($M = 1$, $K = 8$)

Parameter	Method	CPU Time
$N = 32$	MOSEK	0.2208 s
$N = 32$	Proposed Solution	0.0527 s
$N = 64$	MOSEK	0.3127 s
$N = 64$	Proposed Solution	0.0656 s

transmission. The link state table Θ is generated by System Tool Kit (STK) [22], and the capacity is evaluated using the proposed semi-closed form solution. More specifically, we adopt simulations to calculate the proposed solution. Then, we substitute the calculated solution into the objective function and obtain the optimal capacity value. We also note that, in previous works such as [11], [12], the capacity maximized scheduling solution was obtained for the direct transmission scheme with the off-the-shelf optimization software package. Accordingly, as a benchmark, we also solve the capacity maximization problem in our work via MOSEK [23], which is usually adopted to solve integer programs. We find that the resultant capacity values coincide with those obtained by us. This is as expected since the proposed semi-closed form solution is optimal theoretically. Furthermore, compared to [11], [12], the proposed solution has a semi-closed form and thus requires lower computational complexity. To verify this, we run simulations to compare the CPU time required by the proposed solution and MOSEK. Specifically, we focus on solving problem (11). The CPU time comparison results are shown in Table III, which are acquired by using an Intel i7-8550U core computer with 1.8GHz CPU and 16GB RAM. It can be seen from Table III that the proposed solution needs much less CPU time than MOSEK. Moreover, the CPU time of both methods increases when the problem size gets larger.

Fig. 3 depicts the capacity performance of hybrid and direct transmissions. For the hybrid scheme, we set $N = K = 2$ and $\rho_{GS}^{\max} = \rho_{GEO}^{\max} = 1$ and varies the number of GEO satellites. For the direct transmission, we set $N = 2$ and $\rho_{GS}^{\max} = 1$ while varying the number of GSs. It is found that the network capacity gets larger when the number of GEO satellites increases. This is due to the fact that the connection time of SDS increases with more GEO satellites. In particular, we notice that the capacity does not increase linearly with the number of GEO satellites owing to the overlap effect. Specifically, if a certain SDS is involved in multiple available LGLs, it can only choose one link for data transfer. Another observation made from Fig. 3 is that when the hybrid and direct transmission schemes achieve the same capacity, the latter one requires more GSs. For instance, 10 GSs are needed

Fig. 3. Network capacity of hybrid and direct transmissions ($N = K = 2$ and $\rho_{GS}^{\max} = \rho_{GEO}^{\max} = 1$ for hybrid transmission; $N = 2$ and $\rho_{GS}^{\max} = 1$ for direct transmission)

Fig. 4. Network capacity of different transmission schemes under different numbers of GEO satellites ($N = K = 2$ and $\rho_{GS}^{\max} = \rho_{GEO}^{\max} = 1$ for hybrid transmission; $N = 2$ and $\rho_{GEO}^{\max} = 1$ for relay transmission; $N = K = 2$ and $\rho_{GS}^{\max} = 1$ for direct transmission)

by the direct transmission to achieve the same capacity as the hybrid transmission with 2 GEO satellites and 2 GSs.

Fig. 4 compares the performance of direct transmission, relay transmission and hybrid transmission. We set $N = K = 2$ and $\rho_{GS}^{\max} = \rho_{GEO}^{\max} = 1$ for the hybrid transmission, $N = 2$ and $\rho_{GEO}^{\max} = 1$ for relay transmission, and $N = K = 2$ and $\rho_{GS}^{\max} = 1$ for direct transmission. It is clearly observed that the performance of hybrid transmission is superior to the direct transmission and relay transmission. Moreover, the capacity of hybrid transmission is almost the sum of the capacities of the other two methods.

Fig. 5 shows the capacity of hybrid transmission under different numbers of LEO satellites and $\rho_{GEO}^{\max}$'s, where $M = 1$, $K = 2$ and $\rho_{GS}^{\max} = 1$. We observe that the capacity increases when the number of SDSs increases, which is due to the fact that the chance that the SDSs are connected to the GEO satellites or GSs gets higher. Moreover, the capacity is also improved when $\rho_{GEO}^{\max}$ gets larger. This is because the resource related constraint $\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{GEO}^{\max}$ is more relaxed with the increase of $\rho_{GEO}^{\max}$. In fact, when $\rho_{GEO}^{\max}$ approaches infinity,

Fig. 5. Network capacity of hybrid transmission under different numbers of LEO satellites and $\rho_{\text{GEO}}^{\max}$'s ($M = 1$, $K = 2$, $\rho_{\text{GS}}^{\max} = 1$)

Fig. 7. Network capacity of hybrid transmission under different $d_{\max}$'s ($N = K = 2$, $M = 1$, $\rho_{\text{GS}}^{\max} = \rho_{\text{GEO}}^{\max} = 1$)

Fig. 6. Network capacity of hybrid transmission under different number of LEO satellites and $\rho_{\text{GS}}^{\max}$'s ($M = 1$, $K = 8$, $\rho_{\text{GEO}}^{\max} = 1$)

Fig. 8. Network capacity of different transmission schemes under different numbers of GEO satellites ($N = K = 2$ and $\rho_{\text{GS}}^{\max} = \rho_{\text{GEO}}^{\max} = 1$ for hybrid transmission; $N = 2$ and $\rho_{\text{GEO}}^{\max} = 1$ for relay transmission; $N = K = 2$ and $\rho_{\text{GS}}^{\max} = 1$ for direct transmission)

this constraint becomes inactive.

Fig. 6 demonstrates the performance of hybrid transmission under different $\rho_{\text{GS}}^{\max}$'s, where we set $M = 1$, $K = 8$ and $\rho_{\text{GEO}}^{\max} = 1$. We find that, as the value of $\rho_{\text{GS}}^{\max}$ increases, the capacity of both transmission schemes becomes larger since the constraint $\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{\text{GS}}^{\max}$ gets more relaxed. Furthermore, the capacity gain is minor when $\rho_{\text{GS}}^{\max} > 2$ because the chance that more than 2 LEO satellites are in view of one GS is small during our simulation.

In Fig. 7, we show the network capacity of hybrid transmission by including the constraints on data amount. For this simulation, we set $N = K = 2$, $M = 1$, $\rho_{\text{GS}}^{\max} = \rho_{\text{GEO}}^{\max} = 1$, $d_{\min} = 0$ and $d_{1,n} = d_{\max}$. Besides, we generate $\delta_{i,n}^b - \delta_{i,n}^l$ with a random integer in the interval $[0, 10^6]$. From this figure, we discover that the network capacity is improved when $d_{\max}$ is relatively small since the constraint $d_{i,n} \leq d_{\max}$ is more relaxed. Moreover, the capacity gain becomes less significant when $d_{\max}$ is sufficiently large.

We also perform simulations by including the factor of link availability. We set $\eta_{i,n,k}$ and $\varepsilon_{i,n,m}$ to $0.99 \times \text{Bernoulli}(0.9)$ and $0.95 \times \text{Bernoulli}(0.9)$, respectively, where $\text{Bernoulli}(0.9)$ represents a Bernoulli random variable with parameter 0.9.

From the results in Fig. 8, we find that the link outage leads to a minor capacity loss for different transmission schemes. Nevertheless, the proposed hybrid transmission scheme still achieves higher capacity than both direct and relay transmission schemes.

VI. CONCLUSION

In this paper, we have proposed a hybrid transmission scheduling method aiming at maximizing the network capacity for a two-layer space network. We have developed an analytical model of the hybrid transmission scheme for which we have also characterized the capacity in both the ideal and resource constrained scenarios. Simulations have been performed to calculate the proposed semi-closed form scheduling solutions. The results demonstrate that the proposed hybrid transmission strategy achieves a larger network capacity than the conventional direct or relaying transmission scheme. Ongoing work includes identifying economic models to evaluate the trade-offs between federated GS network-based infrastructures and the GEO satellite-based space infrastructures.

APPENDIX A PROOF OF THEOREM 2

Problem (11) is non-convex due to the binary constraints. We first relax them by $0 \leq \chi_{i,n,k} \leq 1$ and $0 \leq \xi_{i,n,m} \leq 1$, which are convex constraints. Then, problem (11) becomes

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \quad (23a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (23b)$$

$$\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{\text{GS}}^{\max}, \quad 1 \leq k \leq K, \quad (23c)$$

$$\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{\text{GEO}}^{\max}, \quad 1 \leq m \leq M, \quad (23d)$$

$$0 \leq \chi_{i,n,k} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (23e)$$

$$0 \leq \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M. \quad (23f)$$

The above problem is convex and satisfies Slater's condition. Hence, strong duality holds. Although we can apply the well-established interior-point method [24] to address this convex problem, it can yield a fractional solution which does not satisfy the binary constraints of problem (11). To handle this, we address problem (23) via solving the dual problem [24] and show that the resultant optimal solution must be binary. Concretely, we first express its partial Lagrangian by

$$\begin{aligned} L(\chi_{i,n,k}, \xi_{i,n,m}, \mu_{i,k}, \nu_{i,m}) &= \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \\ &\quad - \sum_{k=1}^K \mu_{i,k} \left(\sum_{n=1}^N \chi_{i,n,k} - \rho_{\text{GS}}^{\max} \right) - \sum_{m=1}^M \nu_{i,m} \left(\sum_{n=1}^N \xi_{i,n,m} - \rho_{\text{GEO}}^{\max} \right) \\ &= \sum_{n=1}^N \sum_{k=1}^K (\alpha_{i,n,k} - \mu_{i,k}) \chi_{i,n,k} + \sum_{n=1}^N \sum_{m=1}^M (\beta_{i,n,m} - \nu_{i,m}) \xi_{i,n,m} \\ &\quad + \sum_{k=1}^K \rho_{\text{GS}}^{\max} \mu_{i,k} + \sum_{m=1}^M \rho_{\text{GEO}}^{\max} \nu_{i,m}, \end{aligned} \quad (24)$$

where $\mu_{i,k} \geq 0$ and $\nu_{i,m} \geq 0$ represent the dual variables corresponding to the constraints (23c) and (23d), respectively. Then, we obtain the corresponding Lagrangian dual function by

$$f(\mu_{i,k}, \nu_{i,m}) = \max_{\chi_{i,n,k}, \xi_{i,n,m}} L(\chi_{i,n,k}, \xi_{i,n,m}, \mu_{i,k}, \nu_{i,m}) \quad (25a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (25b)$$

$$0 \leq \chi_{i,n,k} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (25c)$$

$$0 \leq \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M. \quad (25d)$$

The dual problem is given by

$$\min_{\mu_{i,k}, \nu_{i,m}} f(\mu_{i,k}, \nu_{i,m}) \quad (26a)$$

$$\text{s.t.} \quad \mu_{i,k} \geq 0, \quad (26b)$$

$$\nu_{i,m} \geq 0. \quad (26c)$$

Denote the dual optimal solution by $(\hat{\mu}_{i,k}, \hat{\nu}_{i,m})$. Then, the primal optimal solution can be determined by solving the following problem:

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} L(\chi_{i,n,k}, \xi_{i,n,m}, \hat{\mu}_{i,k}, \hat{\nu}_{i,m}) \quad (27a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (27b)$$

$$0 \leq \chi_{i,n,k} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (27c)$$

$$0 \leq \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M, \quad (27d)$$

where the objective function has been given in (24). Similar to the proof for Theorem 1, we can also claim that there exists at most one $\chi_{i,n,k}$ or $\xi_{i,n,m}$ being 1. However, the optimal solution in Theorem 1 does not apply here since the coefficients, i.e., the entries of $\mathbf{r}_n$, can be negative. If all the entries of $\mathbf{r}_n$ are negative, then all $\chi_{i,n,k}$'s and $\xi_{i,n,m}$'s should be zero at optimality such that the objective is maximized. Otherwise, the optimization variable corresponding to the largest entry of $\mathbf{r}_n$ should be 1. Therefore, we have proved that (12) is the optimal solution to problem (23).

To determine the dual optimal solution, we need to apply the subgradient algorithm since the Lagrangian dual function $f(\mu_{i,k}, \nu_{i,m})$ is not differentiable. Specifically, following [25], we calculate the subgradients with respect to $\mu_{i,k}$ and $\nu_{i,m}$ in the t -th iteration by

$$s_{\mu}(t) = \rho_{\text{GS}}^{\max} - \sum_{n=1}^N \chi_{i,n,k}(t), \quad (28)$$

$$s_{\nu}(t) = \rho_{\text{GEO}}^{\max} - \sum_{m=1}^M \xi_{i,n,m}(t). \quad (29)$$

The corresponding update rule is thus given by (13) and (14), where the operator $[\cdot]^+$ is due to the non-negative constraints on $\mu_{i,k}$ and $\nu_{i,m}$.

Since problem (23), which relaxes the binary constraints of problem (11), admits binary solutions to $\chi_{i,n,k}$ and $\xi_{i,n,m}$, they must also be the optimal solutions to the original integer program in (11). Therefore, we complete the proof.

$$\begin{aligned}
& L(\chi_{i,n,k}, \xi_{i,n,m}, \mu_{i,k}, \nu_{i,m}, \lambda_{i,n}, \kappa_{i,n}) \\
&= \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \sum_{k=1}^K \mu_{i,k} \left(\sum_{n=1}^N \chi_{i,n,k} - \rho_{\text{GS}}^{\max} \right) - \sum_{m=1}^M \nu_{i,m} \left(\sum_{n=1}^N \xi_{i,n,m} - \rho_{\text{GEO}}^{\max} \right) \\
&+ \sum_{n=1}^N \lambda_{i,n} \left(d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l - d_{\min} \right) \\
&- \sum_{n=1}^N \kappa_{i,n} \left(d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l - d_{\max} \right) \\
&= \sum_{n=1}^N \sum_{k=1}^K ((1 - \lambda_{i,n} + \kappa_{i,n}) \alpha_{i,n,k} - \mu_{i,k}) \chi_{i,n,k} + \sum_{n=1}^N \sum_{m=1}^M ((1 - \lambda_{i,n} + \kappa_{i,n}) \beta_{i,n,m} - \nu_{i,m}) \xi_{i,n,m} \\
&+ \sum_{k=1}^K \rho_{\text{GS}}^{\max} \mu_{i,k} + \sum_{m=1}^M \rho_{\text{GEO}}^{\max} \nu_{i,m} + \sum_{n=1}^N \lambda_{i,n} (d_{i,n} + \delta_{i,n}^b - \delta_{i,n}^l - d_{\min}) - \sum_{n=1}^N \kappa_{i,n} (d_{i,n} + \delta_{i,n}^b - \delta_{i,n}^l - d_{\max}) \quad (31)
\end{aligned}$$

APPENDIX B PROOF OF THEOREM 3

Similar to problem (11), we relax the problem in (17) by

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} \sum_{n=1}^N \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) \quad (30a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (30b)$$

$$\sum_{n=1}^N \chi_{i,n,k} \leq \rho_{\text{GS}}^{\max}, \quad 1 \leq k \leq K, \quad (30c)$$

$$\sum_{n=1}^N \xi_{i,n,m} \leq \rho_{\text{GEO}}^{\max}, \quad 1 \leq m \leq M, \quad (30d)$$

$$d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l \geq d_{\min}, \quad 1 \leq n \leq N, \quad (30e)$$

$$d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k} + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m} \right) - \delta_{i,n}^l \leq d_{\max}, \quad 1 \leq n \leq N, \quad (30f)$$

$$0 \leq \chi_{i,n,k} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (30g)$$

$$0 \leq \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M, \quad (30h)$$

which is convex and satisfies Slater's condition. Similar as Appendix A, we also apply the duality based method to solve this problem in order to avoid fractional solutions. We first express its partial Lagrangian by (31), where $\lambda_{i,n} \geq 0$ and $\kappa_{i,n} \geq 0$ are the dual variables corresponding to the constraints (30e) and (30f), respectively. Accordingly, the Lagrangian dual

function takes the form

$$g(\mu_{i,k}, \nu_{i,m}, \lambda_{i,n}, \kappa_{i,n}) = \max_{\chi_{i,n,k}, \xi_{i,n,m}} L(\chi_{i,n,k}, \xi_{i,n,m}, \mu_{i,k}, \nu_{i,m}, \lambda_{i,n}, \kappa_{i,n}) \quad (32a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (32b)$$

$$0 \leq \chi_{i,n,k} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (32c)$$

$$0 \leq \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M. \quad (32d)$$

The dual problem is given by

$$\min_{\mu_{i,k}, \nu_{i,m}, \lambda_{i,n}, \kappa_{i,n}} g(\mu_{i,k}, \nu_{i,m}, \lambda_{i,n}, \kappa_{i,n}) \quad (33a)$$

$$\text{s.t.} \quad \mu_{i,k} \geq 0, \quad (33b)$$

$$\nu_{i,m} \geq 0, \quad (33c)$$

$$\lambda_{i,n} \geq 0, \quad (33d)$$

$$\kappa_{i,n} \geq 0. \quad (33e)$$

Denote the dual optimal solution by $(\hat{\mu}_{i,k}, \hat{\nu}_{i,m}, \hat{\lambda}_{i,n}, \hat{\kappa}_{i,n})$. Then, the primal optimal solution can be determined by solving the following problem:

$$\max_{\chi_{i,n,k}, \xi_{i,n,m}} L(\chi_{i,n,k}, \xi_{i,n,m}, \hat{\mu}_{i,k}, \hat{\nu}_{i,m}, \hat{\lambda}_{i,n}, \hat{\kappa}_{i,n}) \quad (34a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \chi_{i,n,k} + \sum_{m=1}^M \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad (34b)$$

$$0 \leq \chi_{i,n,k} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq k \leq K, \quad (34c)$$

$$0 \leq \xi_{i,n,m} \leq 1, \quad 1 \leq n \leq N, \quad 1 \leq m \leq M, \quad (34d)$$

where the objective function has been shown in (31). Similar to the proof for Theorem 2, we can prove that (18) is the optimal solution to problem (30).

To find the optimal solutions to the dual variables, we also apply the subgradient algorithm. In particular, based on the proof for Theorem 2, we only need to compute the subgradients with respect to $\lambda_{i,n}$ and $\kappa_{i,n}$ in the t -th iteration,

which are

$$s_{\lambda}(t) = d_{i,n} + \delta_{i,n}^b - \left(\sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k}(t) + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m}(t) \right) - \delta_{i,n}^l - d_{\min}, \quad (35)$$

$$s_{\kappa}(t) = d_{\max} + \delta_{i,n}^l - d_{i,n} - \delta_{i,n}^b + \sum_{k=1}^K \alpha_{i,n,k} \chi_{i,n,k}(t) + \sum_{m=1}^M \beta_{i,n,m} \xi_{i,n,m}(t). \quad (36)$$

Therefore, the corresponding update rules for the dual variables are given in (21) and (22), respectively.

Note that problems (17) and (30) are equivalent since problem (30) admits a binary solution although it relaxes the binary constraints of problem (17). Thus far, we have proved Theorem 3.

REFERENCES

- [1] K. Woellert, P. Ehrenfreund, A. J. Ricco, and H. Hertzfeld, "Cubesats: Cost-effective science and technology platforms for emerging and developing nations," *Advances in Space Research*, vol. 47, no. 4, pp. 663–684, Feb. 2011.
- [2] J. Culter, P. Linder, and A. Fox, "A federated ground station network," in *Proceedings of 5th International Symposium on Space Mission Operations and Ground Data Systems*, Houston, Texas, Oct. 2002, pp. 1–8.
- [3] G. Verfaillie, C. Pralet, and M. Lemaître, "How to model planning and scheduling problems using constraint networks on timelines," *The Knowledge Engineering Review*, vol. 25, no. 3, pp. 319–336, Sep. 2010.
- [4] S. C. Spangelo, J. W. Cutler, A. T. Klesh, and D. R. Boone, "Models and tools to evaluate space communication network capacity," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2387–2404, Jul. 2012.
- [5] D. Zhou, M. Sheng, X. Wang, C. Xu, R. Liu, and J. Li, "Mission aware contact plan design in resource-limited small satellite networks," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2451–2466, Jun. 2017.
- [6] W.-C. Lin, D.-Y. Liao, C.-Y. Liu, and Y.-Y. Lee, "Daily imaging scheduling of an Earth observation satellite," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 2, pp. 213–223, Mar. 2005.
- [7] G. Wu, J. Liu, M. Ma, and D. Qiu, "A two-phase scheduling method with the consideration of task clustering for Earth observing satellites," *Computers & Operations Research*, vol. 40, no. 7, pp. 1884–1894, Jul. 2013.
- [8] F. Marinelli, S. Nocella, F. Rossi, and S. Smriglio, "A Lagrangian heuristic for satellite range scheduling with resource constraints," *Computers & Operations Research*, vol. 38, no. 11, pp. 1572–1583, Nov. 2011.
- [9] Z. Zhang, N. Zhang, and Z. Feng, "Multi-satellite control resource scheduling based on ant colony optimization," *Expert Systems with Applications*, vol. 41, no. 6, pp. 2816–2823, May 2014.
- [10] G.-G. Wang, A. H. Gandomi, A. H. Alavi, and S. Deb, "A hybrid method based on krill herd and quantum-behaved particle swarm optimization," *Neural Computing & Applications*, vol. 27, no. 4, pp. 989–1006, May 2016.
- [11] S. Spangelo, J. Cutler, K. Gilson, and A. Cohn, "Optimization-based scheduling for the single-satellite, multi-ground station communication problem," *Computers & Operations Research*, vol. 57, no. C, pp. 1–16, May 2015.
- [12] J. Castaing, "Scheduling downloads for multi-satellite multi-ground station missions," in *Proceedings of Small Satellite Conference*, Utah, Logan, Aug. 2014, pp. 1–12.
- [13] G. Mushet, G. Mingotti, C. Colombo, and C. McInnes, "Self-organising satellite constellation in geostationary Earth orbit," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 910–923, Apr. 2015.
- [14] X. Jia, T. Lv, F. He, and H. Huang, "Collaborative data downloading by using inter-satellite links in LEO satellite networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1523–1532, Mar. 2017.

- [15] M. Porretta, B. K. Schlarmann, A. Ballereau, and M. Crisci, "A novel uplink scheduling algorithm for the Galileo system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 2, pp. 819–833, Apr. 2018.
- [16] D. L. Brandel, W. A. Watson, and A. Weinberg, "NASA's advanced tracking and data relay satellite system for the years 2000 and beyond," *Proceedings of the IEEE*, vol. 78, no. 7, pp. 1141–1151, Jul. 1990.
- [17] S. Horan, T. O. Minnix, and J. S. Vigil, "Small satellite access of the space network," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 4, pp. 1173–1182, Oct. 1999.
- [18] J. Wang, "China's data relay satellite system and its application prospect," *Spacecraft Engineering*, vol. 22, no. 1, pp. 1–6, Feb. 2013.
- [19] G. Berretta, A. Agostini, and A. Dickinson, "The European data relay system: Present concept and future evolution," *Proceedings of the IEEE*, vol. 78, no. 7, pp. 1152–1164, Jul. 1990.
- [20] L. Wang, C. Jiang, L. Kuang, S. Wu, L. Fei, and H. Huang, "Mission scheduling in space network with antenna dynamic setup times," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 1, pp. 31–45, Feb. 2019.
- [21] A. Schrijver, *Theory of Linear and Integer Programming*, Hoboken, NJ, USA: Wiley, 1998.
- [22] Analytical Graphics, Inc. (AGI), "Systems Tool Kit (STK)," *AGI Product Literature*, 2014 [Online]. Available: <https://www.agi.com/products/engineering-tools>.
- [23] APS Mosek, "The MOSEK optimization software," 2010. [Online]. Available: <http://www.mosek.com>
- [24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [25] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

Shulei Gong received her B.S. and M.S. degrees in information and communication engineering from Southeast University, Nanjing, China, in 2010 and 2013, respectively. She is currently pursuing the Ph.D. degree in communication and information systems at Nanjing University, China. She is also an engineer with the China Mobile Group Jiangsu Co., Ltd. Her current research interests include space information networking and wireless communications.

Hong Shen (S'11–M'14) received the B.E., M.S., and Ph.D. degrees in information and communication engineering from Southeast University, Nanjing, China, in 2009, 2011, and 2014, respectively. During his doctoral studies, he conducted cooperative research with the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA, USA. He is currently an Associate Professor with the National Mobile Communications Research Lab, Southeast University. His research interests include optimization in multiple-input multiple-output systems, full-duplex communications, visible light communications, and source localization in wireless sensor networks.

Kanglian Zhao (M'11) received the B.Sc. and Ph.D. degrees from Nanjing University (China), in 2003 and 2014 respectively. He was with IMEC and K. U. Leuven as an international scholar during Aug. 2009 and Aug. 2010. He is now an associate professor in School of Electronic Science and Engineering and the Institute of Space-Terrestrial Intelligent Networks, Nanjing University. His current research interests include network architecture and emulation for space internetworking, satellite-terrestrial integrated communications.

Xinggan Zhang received the B.S. and M.S. degrees in electronic engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1982 and 1988, respectively, and the Ph.D. degree in signal and information processing from the Nanjing University of Aeronautics and Astronautics, in 2001. He is currently a Professor with the School of Electronic Science and Engineering, Nanjing University, China. His research interests include radar imaging and array signal processing.

Wenfeng Li (M'08) received his Ph.D. degree at National Mobile Communications Research Laboratory, Southeast University, China, in 2012. From August 2012 to April 2017, he was a Postdoctoral Fellow at the School of Electronic Science and Engineering, Nanjing University, China. In April 2017, he joined Nanjing University, where he is an Associate Research Fellow. His current research interests include space network architecture, satellite communications and networks, wireless and mobile communications.

Haibo Zhou (M'14–SM'18) received the Ph.D. degree in information and communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. From 2014 to 2017, he has worked a Post-Doctoral Fellow with the Broadband Communications Research Group, ECE Department, University of Waterloo. Currently, he is an Associate Professor with the School of Electronic Science and Engineering, Nanjing University. His research interests include resource management and protocol design in cognitive radio networks and vehicular

networks.

Ruhai Wang (M'03–SM'17) received the Ph.D. degree in electrical engineering from New Mexico State University, Las Cruces, NM, USA, in 2001.

He is a Professor with the Phillip M. Drayer Department of Electrical Engineering, Lamar University, Beaumont, TX, USA. He currently also serves as a Principal Scientist with Nanjing University, Nanjing, China and a Chair Professor with Soochow University, Suzhou, China. He has published nearly 100 research papers in international journals and conferences proceedings. His current research

interests include space communications and networks, and cybersecurity.

Dr. Wang is currently an Associate Editor for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS and IEEE AEROSPACE AND ELECTRONIC SYSTEMS MAGAZINE.

Zhili Sun (SM'09) is a Chair Professor with Institute of Communication Systems (Formerly Centre for Communication Systems Research CCSR), University of Surrey, UK. He received his BSc in mathematics from Nanjing University and PhD from the Department of Computing, Lancaster University. His research interests include satellite communications and networks, wireless mobile and sensor networks, mobile operating systems, traffic engineering, IP networks and security. He has authored 3 books and published over 240 papers in international journals

and conferences.