Physical-Layer Security in Cache-Enabled Cooperative Small Cell Networks Against Randomly Distributed Eavesdroppers

Tong-Xing Zheng, Member, IEEE, Hui-Ming Wang, Senior Member, IEEE, and Jinhong Yuan, Fellow, IEEE,

Abstract—This paper explores the physical-layer security in a small cell network (SCN) with cooperative cache-enabled small base stations (SBSs) in the presence of randomly distributed eavesdroppers. We propose a joint design on the caching placement and the physical-layer transmission to improve the secure content delivery probability (SCDP). We first put forward a hybrid caching placement strategy in which a proportion of the cache unit in each SBS is assigned to store the most popular files (MPFs), while the remaining is used to cache the disjoint subfiles (DSFs) of the less popular files in different SBSs as a means to enhance transmission secrecy and content diversity. We then introduce two coordinated multi-point (CoMP) techniques, namely, joint transmission (JT) and orthogonal transmission (OT), to deliver the MPFs and DSFs, respectively. We derive analytical expressions for the SCDP in each transmission scheme, considering both non-colluding and colluding eavesdropping scenarios. Based on the obtained analytical results, we jointly design the optimal transmission rates and the optimal caching assignment for maximizing the overall SCDP. Various insights into the optimal transmission and caching designs are further provided. Numerical results are also presented to verify our theoretical findings and to demonstrate the superiority of the proposed caching and transmission strategies.

Index Terms—Physical-layer security, wireless caching, small cell networks, cooperative transmissions, stochastic geometry.

I. INTRODUCTION

S MALL cell network (SCN) is a promising approach to improving network capacity and achieving seamless wireless coverage in the 5G wireless network. Nevertheless, the increasingly dense deployment of SCNs poses a tremendous challenge to the backhaul links and the backhaul capacity has become the major system bottleneck. To alleviate such bottleneck, wireless caching technique emerges. By pre-storing popular content at the edge of an SCN such as small base stations (SBSs) and reusing the cached content to meet frequent requests from local users, wireless caching is envisioned as an effective solution for relaxing the challenging demand of small cell backhauling and reducing the end-to-end latency [1].

Wireless caching has a serious security vulnerability as any wireless network due to the broadcast nature of the electromagnetic signal propagation. For example, although wireless caching has a great potential to meet the soaring

video-on-demand (VoD) streaming traffic in the 5G network [2], [3], the broadcast streaming data by caching nodes are susceptible to potential eavesdroppers such as non-paying subscribers and malicious attackers. It is of great significance to propose security-wise caching schemes that can guarantee both data secrecy and quality of service (QoS). Nonetheless, safeguarding the security for a cache-enabled wireless network is confronted with two major challenges. Specifically, the vast majority of current wireless data services still rely on endto-end encryption to ensure data secrecy, e.g., the hypertext transfer protocol secure (HTTPS) applied for video streaming applications such as YouTube and Netflix [4]. However, such encryption schemes might counteract the benefits of wireless caching in terms of high flexibility and large multiplexing gains since the encrypted content is uniquely defined for each user request and cannot be reused to serve other user requests [5]. Moreover, using the encryption methods will inevitably introduce a large amount of additional operations in the storage, management and distribution of secret keys, thus degrading the efficiency of content placement and delivery. Fortunately, physical-layer security (PLS) [6], an informationtheoretic approach which has been proven to gain a remarkable secrecy enhancement in various wireless networks [7]-[18], provides a new opportunity to overcome the above limitations. PLS achieves wireless secrecy by using the wiretap channel encoding instead of the source encryption such that the cached content still can be reused. Moreover, PLS exploits the randomness inherent to the wireless channels without necessarily relying on secret keys. All these advantages make PLS and wireless caching easily integrated in a low-complexity and high-flexibility way.

1

A. Previous Works

The cache-enabled SCN is first investigated in [19], where wireless caching has been shown to significantly reduce the average downloading delay. In [2], [20], wireless caching has been exploited to improve the energy efficiency of cellular networks. In [21], a joint caching and buffering strategy has been proposed to overcome the backhaul capacity bottleneck and the half-duplex transmission constraint simultaneously. Considering cooperative transmissions via distributed caching helpers, an optimal caching placement has been designed in [22] as a means to balance the file diversity gain and the cooperation gain. In [23], [24], the cache-enabled small cell cooperation in the caching placement (i.e., cache-level

T.-X. Zheng and H.-M. Wang are with the School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (zhengtx@mail.xjtu.edu.cn, xjbswhm@gmail.com).

J. Yuan is with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: j.yuan@unsw.edu.au).

cooperation) and the physical-layer transmission (i.e., signallevel cooperation) has been discussed. In [25], the cache- and signal-level cooperation has been leveraged under a combined caching placement scheme to improve the cache service performance and the energy efficiency.

Recently, the security issue in the cache-enabled wireless networks has attracted a stream of research, motivated by the coded caching scheme introduced in [26]. For example, a coded caching scheme based on Shannon's one-time pad method has been proposed in [27] to guarantee information secrecy against eavesdroppers. However, achieving secrecy requires a sufficiently large size of random secret keys, and secure sharing of such massive secret keys will cause a considerable system overhead. This scheme has been extended to the device-to-device networks in [28], where a sophisticated key generation and encryption scheme has been designed. In [29], [30], the security-oriented content placement has been studied based on the maximal distance separable encoding. All these endeavors have dealt with the security issue from an information-theoretic point of view, but few has considered the characteristics of the physical-layer media.

PLS in a cache-enabled wireless network was not considered until recently. Specifically, the caching-enabled cooperative multi-input multi-output (MIMO) transmission has been first exploited as an effective PLS mechanism to increase the secrecy rate of content delivery, confronting either a single malicious eavesdropper [3] or multiple untrusted cache helpers [31]. Nevertheless, the signaling design therein heavily relies on the estimated instantaneous channel state information (CSI) of the eavesdropper. In practice, it is difficult to estimate such CSI in real time since the eavesdropper usually listens passively. Dynamically adjusting the transmission parameters will also increase the system complexity and the end-to-end latency. Moreover, the fading feature of the wireless channels and the randomness of the eavesdroppers' locations have significant impacts on the security performance, and meanwhile they can also be exploited to facilitate secure transmissions. However, these aspects have not yet been investigated in [3], [31]. It is worth mentioning that, there has been substantial research on the PLS in random wireless networks with both channel and location uncertainties [11]-[18]. Stochastic geometry theory has provided a powerful tool to study the network security performance by modeling the positions of network nodes including the eavesdroppers according to a spatial distribution such as a Poisson point process (PPP) [32]. To the best of our knowledge, the potential of PLS in securing the content delivery for a cache-enabled SCN against randomly distributed eavesdroppers is still elusive, and even a fundamental mathematical framework for security performance analysis and optimization from a stochastic geometry perspective is lacking. This has motivated our work.

B. Our Work and Contributions

In this paper, we will explore the potential of the physicallayer secure transmission in conjunction with caching placement in realizing secure content delivery against randomly distributed eavesdroppers. The main contributions of this paper are summarized as follows.

- We study a combination of cooperations in both cachelevel and signal-level. By cache-level cooperation, we propose a hybrid caching placement strategy based on file partition, where every SBS assigns a proportion of its cache space to store the most popular files (MPFs), while using the remaining to cache the disjoint subfiles (DSFs) of the less popular files as a means to improve content diversity and secrecy. By signal-level cooperation, we put forward two coordinated multi-point (CoMP) techniques, namely, joint transmission (JT) and orthogonal transmission (OT), to deliver the cached MPFs and DSFs, respectively.
- We assess the secure content delivery probability (SCDP), which measures the probability that the reliability and secrecy of content delivery can be guaranteed simultaneously. We provide analytical results for the SCDP in each transmission scheme for both non-colluding and colluding wiretap scenarios. We show that the JT scheme outperforms the OT scheme in terms of transmission reliability, whereas the latter provides a higher level of secrecy.
- We jointly design the optimal transmission rates and the optimal caching assignment proportion to maximize the overall SCDP under the proposed caching and transmission strategies. The whole maximization procedure is decomposed into two phases. First, the optimal transmission rates that globally maximize the SCDP in the JT and OT schemes are obtained by solving a scalar convex problem and by addressing a vector convex problem through an alternating optimization, respectively; subsequently, the optimal caching assignment proportion that maximizes the overall SCDP is derived in a closed-form expression. Various insights into the optimal transmission and caching designs are further provided.

C. Organization and Notations

The remainder of this paper is organized as follows. In Section II, we describe the system model and the underlying optimization problem. In Section III, we analyze the connection probability and the secrecy probability for both the JT and OT schemes. In Section IV, we design the optimal transmission rates and the optimal caching assignment proportion to maximize the overall SCDP. In Section V, we present numerical results to validate our theoretical analysis. In Section VI, we conclude our work and provide several future research directions.

Notations: Bold lowercase letters denote column vectors. $(\cdot)^{\mathrm{T}}$, $|\cdot|$, $||\cdot||_1$, $||\cdot|$, $\mathbb{P}\{\cdot\}$, $\mathbb{E}_A[\cdot]$ denote the operations of transpose, absolute value, L-1 norm, round down, probability and mathematical expectation taken over the random variable (RV) A, respectively. $\Gamma(z) = \int_0^{\infty} e^{-t}t^{z-1}dt$, $\Gamma(a,x) = \int_x^{\infty} e^{-t}t^{a-1}dt$ and $\operatorname{erf}(z) = 2/\sqrt{\pi}\int_0^z e^{-t^2}dt$ are the gamma function [40, Eqn. (8.310.1)], the incomplete gamma function [40, Eqn. (8.350.2)], and the error function [40, Eqn. (8.250.1)], respectively. $S \setminus s_k$ denotes the set obtained by canceling the subset s_k from the superset S.



Fig. 1: Illustration of a cache-enabled cooperative SCN. The ongoing content delivery from the SBSs to an intended user is overheard by randomly located eavesdroppers.

II. SYSTEM MODEL

We consider security issues in a cache-enabled SCN where a set of SBSs, $\mathcal{K} = [1, \cdots, K]$, cooperatively deliver paid content to a subscriber coexisting with randomly located eavesdroppers (e.g., non-paying subscribers). All the SBSs are connected to a central server located at the edge of the core network via wired backhaul links. A cache unit is deployed at each SBS for pre-fetching popular content during off-peak hours from the central server. Once a file requested by a local user is pre-stored in the cache units, i.e., a cache hit event takes place, the file can be delivered to the user by the SBSs directly. Otherwise, a cache miss event is deemed to occur and the SBSs will fetch the requested file from the central server before conveying it to the user. We assume that all the network nodes including the SBSs, legitimate users and eavesdroppers each have a single antenna, and only one user can be served in each time slot. The locations of the SBSs and of the users are assumed deterministic, while those of eavesdroppers are random and unknown. Without loss of generality, we place a typical user at the origin of the polar coordinate, and denote the locations of the k-th nearest SBS and of the j-th nearest eavesdropper to the typical user as $\{b_k : (r_{b,k}, \theta_{b,k})\}$ and $\{e_j : (r_{e,j}, \theta_{e,j})\}$, with r and θ being the corresponding distance and angle, respectively. The distance between the k-th SBS and the j-th eavesdropper is expressed as $r_{j,k} = \sqrt{r_{b,k}^2 + r_{e,j}^2 - 2r_{b,k}r_{e,j}\cos(\theta_{b,k} - \theta_{e,j})}$. Since the eavesdroppers are possibly randomly located over the entire network, we model their positions as a stationary PPP Φ_e on the two-dimensional plane \mathbb{R}^2 with density λ_e , i.e., $e_i \in \Phi_e$ [11]-[18].

To model the wireless channels, including the main channels spanning from the SBSs to the typical user and the wiretap channels spanning from the SBSs to the eavesdroppers, we consider a standard distance-based path loss governed by the exponent α along with Rayleigh fading. Hence, the channel gains from the k-th SBS to the typical user and to the *j*-th eavesdropper can be expressed as $h_{b,k}r_{b,k}^{-\alpha/2}$ and $h_{j,k}r_{j,k}^{-\alpha/2}$, respectively, where $h_{b,k}$ and $h_{j,k}$ denote the independent and identically distributed (i.i.d.) fading coefficients obeying the

circularly symmetric complex Gaussian distribution with zero mean and unit variance. We assume that the SBSs only know the statistic CSI (as opposed to the instantaneous) of the main and the wiretap channels.¹

A. Caching Placement Strategy

The central server owns a library of N equal-size files with F_n being the *n*-th most popular file. We assume that users make request for files independently with some probabilities according to a commonly adopted popularity pattern, i.e., Zipf distribution [19]. With the Zipf distribution, the request probability for the *m*-th most popular file is given by

$$f_m = \frac{m^{-\gamma}}{\sum_{n=1}^N n^{-\gamma}},\tag{1}$$

where γ models the skewness of the popularity profile. In particular, the popularity profile is uniform over files for $\gamma = 0$, and becomes more concentrated as γ increases.

We consider a finite caching capacity that the K SBSs each can store up to L files and the total capacity KL is less than the library size N, i.e., KL < N. To make efficient use of the cache units, we propose a hybrid caching placement strategy to distribute files of different popularities to the SBSs, where each file is partitioned into K equal-size subfiles. To be specific, we divide the cache unit in every SBS into two portions, where a proportion ϕ of the cache unit is used for storing the most popular files (i.e., MPF caching) and the rest proportion $1 - \phi$ is reserved for caching the disjoint subfiles of the less popular files in different SBSs (i.e., DSF caching). Under such a hybrid caching strategy, file F_n with $1 \leq n \leq \lfloor \phi L \rfloor$ belongs to the MPF group where all the SBSs possess a copy of it; file F_n with $|\phi L| < n \le |\phi L| + K(L - |\phi L|)$ belongs to the DSF group where the SBSs each store a unique subfile of it; file F_n with $n > \lfloor \phi L \rfloor + K(L - \lfloor \phi L \rfloor)$ is not cached and only can be fetched from the central server.

It is worth noting that, the DSF caching mode helps to enhance content diversity, i.e., allowing more files to be cached with a finite caching capacity. This can increase the cache hit probability and avoid frequently fetching content from the central server, thus making the communication more efficient. Furthermore, the different subfiles of each file are elaborately encoded such that one only can decode them in a sequential manner [33]. Hence, the risk of the entire file being illegally intercepted can be significantly lowered since it is difficult to decode all the subfiles simultaneously. In other words, content diversity can guarantee a higher level of secrecy.

B. Cooperative Transmission Schemes

When the SBSs receive a file request, they adopt different transmission schemes depending on the popularity of the requested file. We propose two CoMP schemes, namely, JT and OT schemes, which are described as below.

¹This is a generic assumption in literature on PLS [7]-[18]. In particular, if the potential eavesdroppers are also regular users but the ongoing content delivery should be kept secret to them, it is possible for the SBSs to acquire the statistic CSI of the eavesdroppers and the distribution of their locations through a large amount of information exchange during other time slots.

1) Joint Transmission (JT): When the requested file F_n belongs to the MPF group, the SBSs each have a copy of it. In order to enhance transmission reliability, the SBSs deliver the same file to the user simultaneously. We refer to this scheme as the JT scheme.² The received signal-to-noise ratios (SNRs) at the user and at the *j*-th eavesdropper respectively can be given by

$$\operatorname{SNR}_{b} = \rho \left| \sum_{k=1}^{K} h_{b,k} r_{b,k}^{-\alpha/2} \right|^{2}, \qquad (2)$$

$$\operatorname{SNR}_{j} = \rho \left| \sum_{k=1}^{K} h_{j,k} r_{j,k}^{-\alpha/2} \right|^{2}, \quad \forall e_{j} \in \Phi_{e},$$
(3)

where $\rho = P/(WN_0)$ denotes the normalized SNR, with P, W and N_0 being the SBS transmit power, the total available bandwidth and the noise spectral density, respectively. We consider identical noise spectral density at all the receivers.

2) Orthogonal Transmission (OT): When the requested file F_n falls within the DSF group, the SBSs have different subfiles of it. In order to avoid the co-channel interference, the SBSs use orthogonal frequency spectrum to deliver these subfiles, i.e., each SBS occupies 1/K of the overall bandwidth. We name this scheme the OT scheme. For the subfile delivered from the k-th SBS, the received SNRs at the user and at the *j*-th eavesdropper respectively can be given by

$$\operatorname{SNR}_{b,k} = K\rho \left| h_{b,k} \right|^2 r_{b,k}^{-\alpha}, \tag{4}$$

$$\operatorname{SNR}_{j,k} = K\rho \left| h_{j,k} \right|^2 r_{j,k}^{-\alpha}, \quad \forall e_j \in \Phi_e.$$
(5)

Note that the factor K exists due to a 1/K decrement of the available bandwidth for each SBS.

3) Cache Miss (CM): When the requested file F_n is not cached by the SBSs, i.e., a cache miss event occurs, all the SBSs fetch the requested file from the central server and then deliver it to the user simultaneously. Hence, the received SNRs at the user and at the *j*-th eavesdropper share the same expressions as (2) and (3), respectively. For ease of statement, we call this scheme the CM scheme. The wired backhauling process is assumed to be secure, whereas it causes extra delivery delay compared with the cache hit case, which will result in a lower end-to-end rate. This negative impact will be considered in the subsequent performance analysis and optimization.

C. Performance Metrics

To securely deliver the content, the well-known Wyner's wiretap encoding scheme is adopted to encode the confidential information, where redundant information is intentionally embedded to confuse eavesdroppers [6]. The transmission rates of the confidential information and the redundant information are referred to as the secrecy rate R_s and the redundant rate R_e , and the wiretap codeword rate is $R_t = R_s + R_e$. We

denote the achievable rates of the main channel and the wiretap channel as $C_b = \log_2(1 + \text{SNR}_b)$ and $C_e = \log_2(1 + \text{SNR}_e)$, respectively, with SNR_b and SNR_e being the corresponding SNRs.³ If C_b is larger than or equal to R_t , the legitimate user can recover the confidential information. The probability that this event happens is called the connection probability, which is defined as $p_c \triangleq \mathbb{P}\{C_b \ge R_t\}$ [16]. If C_e is lower than or equal to R_e , no confidential information will be leaked to the eavesdropper. The probability of this event occurs is termed the secrecy probability, which is defined as $p_s \triangleq \mathbb{P}\{C_e \le R_e\}$ [16].

The content delivery is secure only if the reliability (connection) and the secrecy are guaranteed simultaneously. In this paper, we employ the metric named SCDP to quantify the probability that a secure content delivery event occurs. For a specific transmission scheme, the SCDP can be defined as the product of the connection probability and the secrecy probability, i.e., $\mathcal{P}_{scd}^{S} \triangleq p_{c}^{S} p_{s}^{S}$ for $S \in \{JT, OT, CM\}$. Therefore, the overall SCDP under the proposed caching and transmission strategies can be expressed as

$$\mathcal{P}_{scd} = \sum_{\mathbf{S} \in \{\mathbf{JT}, \mathbf{OT}, \mathbf{CM}\}} p_{tr}^{\mathbf{S}}(\phi) \ \mathcal{P}_{scd}^{\mathbf{S}},\tag{6}$$

where $p_{tr}^{S}(\phi)$ denotes the probability of the scheme $S \in \{JT, OT, CM\}$ being adopted for content delivery, which can be calculated from (1), i.e.,

$$p_{tr}^{\mathrm{JT}}(\phi) = \sum_{n=1}^{\lfloor \phi L \rfloor} f_n, \quad p_{tr}^{\mathrm{OT}}(\phi) = \sum_{\substack{|\phi L| + 1 \\ |\phi L| + 1}}^{\lfloor \phi L \rfloor + K(L - \lfloor \phi L \rfloor)} f_n, \quad (7a)$$

$$p_{tr}^{\rm CM}(\phi) = 1 - p_{tr}^{\rm JT}(\phi) - p_{tr}^{\rm OT}(\phi).$$
 (7b)

We emphasize that the redundant rate R_e (for a target secrecy rate R_s) and the caching assignment proportion ϕ play critical roles in increasing the SCDP. Specifically, R_e triggers a trade-off between transmission reliability and secrecy. Choosing a larger R_e increases the secrecy probability p_s but decreases the connection probability p_c as the wiretap codeword rate R_t also increases. Likewise, ϕ strikes a nontrivial trade-off between transmission reliability, secrecy and content diversity. On one hand, devoting a larger proportion ϕ for MPFs increases the probability of adopting the JT scheme and thus enhancing transmission reliability. On the other hand, assigning a larger proportion for DSFs (i.e., a smaller ϕ) increases the probability of using the OT scheme and thus improving transmission secrecy and content diversity. The overall balance of these opposite impacts on the SCDP should be carefully addressed.

In the following sections, we will first derive the connection probability p_c and the secrecy probability p_s for various transmission schemes, and then we will design the optimal redundant rate R_e and the optimal caching assignment proportion ϕ to maximize the overall SCDP \mathcal{P}_{scd} .

III. CONNECTION AND SECRECY PROBABILITIES ANALYSIS

Since the SCDP is represented as the product of the connection probability p_c and the secrecy probability p_s , this section

²Without the instantaneous CSI of the main channels, the JT scheme corresponds to a non-coherent multi-point joint transmission. If the instantaneous CSI is available, a coherent joint transmission or distributed beamforming can be realized to further improve transmission reliability, which however will increase the system overhead.

³All the rate parameters in this paper are measured in the unit: bits/s/Hz.

will analyze p_c and p_s for the JT and OT schemes, respectively. The analysis for the CM scheme is similar as that for the JT scheme, and the only difference lies in the backhaul delay caused by the content fetching process in the former. Due to the extra delay, the actual delivery time in the CM scheme reduces and hence the required secrecy rate increases from R_s (for the JT scheme) to δR_s , where $\delta > 1$ captures the impact of backhaul delay.

It is worth mentioning that, for a fair comparison between the JT and OT schemes, we consider identical target secrecy rate R_s in both schemes. In practice, the secrecy rate can correspond to the end-to-end rate of a specific service requested by subscribers such that its value may be pre-established. Furthermore, all the SBSs should set a same wiretap codeword rate R_t and a same redundant rate R_e in the JT scheme due to the joint transmission of a same file; whereas they can choose different rates in the OT scheme due to the orthogonal transmission of different subfiles. This also means that, the wiretap codeword rates (also the redundant rates) in the JT and OT schemes are not necessarily the same and actually their values should be properly designed for maximizing the SCDP as will be discussed in Sec. IV.

A. Connection Probability

1) JT Scheme: In this case, the achievable rate of the main channel is $C_b = \log_2 (1 + \text{SNR}_b)$ with SNR_b given in (2). Let $R_t = \log_2 (1 + \beta_t)$ be the wiretap codeword rate with $\beta_t \triangleq 2^{R_t} - 1$. The connection probability can be expressed as $p_c^{\text{JT}} = \mathbb{P}\{\text{SNR}_b \ge \beta_t\}$, which is calculated as

$$p_c^{\rm JT} = \mathbb{P}\left\{ \left| \sum_{k=1}^K h_{b,k} r_{b,k}^{-\alpha/2} \right|^2 \ge \frac{\beta_t}{\rho} \right\} \stackrel{(a)}{=} e^{-\frac{\beta_t/\rho}{\sum_{k=1}^K r_{b,k}^{-\alpha}}}, \quad (8)$$

where step (a) follows from the fact that $\left|\sum_{k=1}^{K} h_{b,k} r_{b,k}^{-\alpha/2}\right|^2$ is exponentially distributed with mean $\sum_{k=1}^{K} r_{b,k}^{-\alpha}$.

2) OT Scheme: In this case, the achievable rate of the main channel of the k-th SBS is $C_{b,k} = \log_2 (1 + \text{SNR}_{b,k})$ with $\text{SNR}_{b,k}$ given in (4). Let $R_{t,k} = \log_2 (1 + \beta_{t,k})$ be the wiretap codeword rate of the k-th SBS with $\beta_{t,k} = 2^{R_{t,k}} - 1$. Since the entire file can be recovered only if all the subfiles have already been decoded. Therefore, the connection probability can be expressed as $p_c^{\text{OT}} = \mathbb{P} \{ \bigcap_{k \in \mathcal{K}} \text{SNR}_{b,k} \ge \beta_{t,k} \}$, which is calculated as

$$p_{c}^{\text{OT}} = \prod_{k=1}^{K} \mathbb{P}\left\{ \left| h_{b,k} \right|^{2} r_{b,k}^{-\alpha} \ge \frac{\beta_{t,k}}{K\rho} \right\} = e^{-\frac{\sum_{k=1}^{K} r_{b,k}^{\alpha} \beta_{t,k}}{K\rho}}.$$
 (9)

Comparing (9) with (8), we find that if $\beta_t = \beta_{t,k}$ for $k \in \mathcal{K}$, i.e., the same wiretap codeword rate is used in the JT and OT schemes, we have $p_c^{\text{JT}} \ge p_c^{\text{OT}}$ as $1/\sum_{k=1}^{K} r_{b,k}^{-\alpha} \le \min_{k \in \mathcal{K}} r_{b,k}^{\alpha} \le \sum_{k=1}^{K} r_{b,k}^{\alpha}/K$. This shows the superiority of the JT scheme in terms of transmission reliability.

B. Secrecy Probability

We consider both non-colluding and colluding eavesdropping (NCE/CE) scenarios. In the NCE case, eavesdroppers individually decode the confidential information and thus the content can be delivered secretly if only the achievable rate of the most deteriorate eavesdropper $\max_{e_j \in \Phi_e} C_j$ does not exceed the redundant rate R_e , i.e., $\max_{e_j \in \Phi_e} C_j \leq R_e$. In the CE case, eavesdroppers jointly decode the confidential information using the maximal ratio combination (MRC) method. The content delivery is deemed to be secret only if the achievable rate of the equivalent wiretap channel C_e does not lie beyond the redundant rate R_e , i.e., $C_e \leq R_e$.

1) JT Scheme for NCE Case: In this case, the achievable rate of the *j*-th eavesdropper is $C_j = \log_2 (1 + \text{SNR}_j)$ with SNR_j given in (3). Let $R_e = \log_2 (1 + \beta_e)$ be the redundant rate with $\beta_e \triangleq 2^{R_e} - 1$. The secrecy probability can be expressed as $p_{s,nce}^{\text{JT}} = \mathbb{P} \{ \max_{e_j \in \Phi_e} \text{SNR}_j \leq \beta_e \}$, which is calculated as

$$p_{s,nce}^{\mathrm{JT}} = \mathbb{E}_{\Phi_e} \left[\prod_{e_j \in \Phi_e} \mathbb{P}\left\{ \left| \sum_{k=1}^{K} h_{j,k} r_{j,k}^{-\alpha/2} \right|^2 \leq \frac{\beta_e}{\rho} \right\} \right]$$

$$\stackrel{(\mathrm{b})}{=} \mathbb{E}_{\Phi_e} \left[\prod_{e_j \in \Phi_e} \left(1 - e^{-\frac{\beta_e/\rho}{\sum_{k=1}^{K} r_k^{-\alpha}}} \right) \right]$$

$$\stackrel{(\mathrm{c})}{=} \exp\left(-2\lambda_e \int_0^\infty \int_0^\pi e^{-\frac{\beta_e/\rho}{\sum_{k=1}^{K} r_k^{-\alpha}}} r dr d\theta \right), \quad (10)$$

where step (b) follows from knowing that $\left|\sum_{k=1}^{K} h_{j,k} r_{j,k}^{-\alpha/2}\right|^2$ is an exponential RV with mean $\sum_{k=1}^{K} r_{j,k}^{-\alpha}$, and step (c) holds by using the probability generating functional (PGFL) over a PPP [39] with $r_k = \sqrt{r_{b,k}^2 + r^2 - 2r_{b,k}r \cos(\theta_{b,k} - \theta)}$. Although the result in (10) does not appear in a closed form, the integral therein is fairly easy to compute.

2) OT Scheme for NCE Case: In this case, the achievable rate of the wiretap channel from the k-th SBS to the j-th eavesdropper is $C_{j,k} = \log_2(1 + \operatorname{SNR}_{j,k})$ with $\operatorname{SNR}_{j,k}$ given in (5). Let $R_{e,k} = \log_2(1 + \beta_{e,k})$ be the redundant rate of the k-th SBS with $\beta_{e,k} \triangleq 2^{R_{e,k}} - 1$. Note that a file is intercepted only if all its subfiles have already been intercepted. Then, the secrecy probability can be given by $p_{s,nce}^{\operatorname{OT}} = \mathbb{P}\left\{\bigcup_{k\in\mathcal{K}} \operatorname{SNR}_{j,k} \leq \beta_{e,k}, \forall e_j \in \Phi_e\right\}$, which is calculated as

$$p_{s,nce}^{\text{OT}} = \mathbb{E}_{\Phi_e} \left[\prod_{e_j \in \Phi_e} \left(1 - \prod_{k=1}^K \mathbb{P} \{ \text{SNR}_{j,k} > \beta_{e,k} \} \right) \right]$$
$$= \mathbb{E}_{\Phi_e} \left[\prod_{e_j \in \Phi_e} \left(1 - e^{-\frac{\sum_{k=1}^K r_k^{\alpha} \beta_{e,k}}{K_{\rho}}} \right) \right]$$
$$= \exp \left(-2\lambda_e \int_0^\infty \int_0^\pi e^{-\frac{\sum_{k=1}^K r_k^{\alpha} \beta_{e,k}}{K_{\rho}}} r dr d\theta \right). \quad (11)$$

Comparing (11) with (10), if the same redundant rate is employed in the JT and OT schemes, i.e., $\beta_e = \beta_{e,k}$ for $k \in \mathcal{K}$, we have $p_s^{\text{OT}} \ge p_s^{\text{JT}}$ since $1/\sum_{k=1}^{K} r_k^{-\alpha} \le \min_{k \in \mathcal{K}} r_k^{\alpha} < \sum_{k=1}^{K} r_k^{\alpha}/K$. This means the OT scheme provides a higher level of secrecy than does the JT scheme. For the special case with $\alpha = 2$, we can obtain a more concise expression for $p_{s,nce}^{\text{OT}}$ as given below

$$p_{s,\alpha=2}^{\text{OT}} = \exp\left(-\rho\lambda_e \left(\pi + Z_K\right)e^{-\frac{\sum_{k=1}^{K} r_{b,k}^2 \beta_{e,k}}{K\rho}}\right), \quad (12)$$

where $Z_K = \int_0^{\pi} \sqrt{\pi} z e^{z^2} [1 + \operatorname{erf}(z)] d\theta$ with $z = \sum_{k=1}^{K} r_{b,k} \cos(\theta_{b,k} - \theta) / (K \sqrt{\rho}).$

3) JT Scheme for CE Case: In this case, the achievable rate of the equivalent wiretap channel is $C_e = \log_2\left(1 + \sum_{e_j \in \Phi_e} \text{SNR}_j\right)$ with SNR_j given in (3). Let $I_e = \sum_{e_j \in \Phi_e} \text{SNR}_j$. Then, the secrecy probability can be expressed as the cumulative distribution function (CDF) of I_e , i.e., $p_{s,ce}^{\text{JT}} = \mathbb{P}\left\{I_e \leq \beta_e\right\}$. In order to compute $p_{s,ce}^{\text{JT}}$, we first calculate the Laplace transform of I_e .

Lemma 1: The Laplace transform of I_e evaluated at value s is given by

$$\mathcal{L}_{I_e}(s) = \exp\left(-2\lambda_e \int_0^\infty \int_0^\pi \frac{s\rho \sum_{k=1}^K r_k^{-\alpha}}{1 + s\rho \sum_{k=1}^K r_k^{-\alpha}} r dr d\theta\right),\tag{13}$$

where r_k shares the same expression as (10).

Proof 1: Recalling (3), the Laplace transform $\mathcal{L}_{I_e}(s) = \mathbb{E}_{I_e}[e^{-sI_e}]$ can be calculated as

$$\mathcal{L}_{I_e}(s) \stackrel{(d)}{=} \mathbb{E}_{\Phi_e} \left[\prod_{e_j \in \Phi_e} \mathbb{E}_{h_{j,k}} \left[e^{-s\rho \left| \sum_{k=1}^K h_{j,k} r_{j,k}^{-\alpha/2} \right|^2} \right] \right]$$
$$\stackrel{(e)}{=} \mathbb{E}_{\Phi_e} \left[\prod_{e_j \in \Phi_e} \frac{1}{1 + s\rho \sum_{k=1}^K r_{j,k}^{-\alpha}} \right], \quad (14)$$

where step (d) follows from the independence between channel fading and the PPP such that the expectation over h can be moved inside the product; step (e) is obtained by calculating the Laplace transform of an exponential RV $\left|\sum_{k=1}^{K} h_{j,k} r_{j,k}^{-\alpha/2}\right|^2$. Applying the PGFL over a PPP with (14) completes the proof.

It is intractable to give a closed form for the exact $p_{s,ce}^{JT}$. In the following theorem, we resort to a widely used approximation method [34] and derive a closed-form approximation for $p_{s,ce}^{JT}$.

Theorem 1: The secrecy probability in the JT scheme for the CE case satisfies

$$p_{s,ce}^{\rm JT} \lessapprox \sum_{m=1}^{M} \binom{M}{m} (-1)^{m+1} \mathcal{L}_{I_e} \left(\frac{m\xi}{\beta_e}\right), \qquad (15)$$

where $\xi \triangleq M(M!)^{-1/M}$ and M is the number of terms used in the approximation.

Proof 2: The secrecy probability $p_{s,ce}^{\text{JT}} = \mathbb{P}\{I_e \leq \beta_e\}$ can be calculated as follows,

$$p_{s,ce}^{\text{JT}} = \mathbb{P}\left\{I_e/\beta_e \le 1\right\} \stackrel{\text{(f)}}{\approx} \mathbb{P}\left\{I_e/\beta_e \le \iota\right\}$$
$$\stackrel{\text{(g)}}{\approx} 1 - \mathbb{E}_{I_e}\left[\left(1 - e^{-\xi I_e/\beta_e}\right)^M\right], \quad (16)$$

where the dummy variable ι in step (f) is a normalized gamma RV with the PDF $f_{\iota}(x) = x^{M-1}e^{-x}/\Gamma(M)$, and this step follows from the fact that ι converges to identity as M approaches

infinity [34]; step (g) yields an upper bound by invoking Alzer's inequality [35], i.e., $\mathbb{P}\{\iota \geq z\} \leq 1 - [1 - e^{-\xi z}]^M$ for a constant z > 0. Using the binomial expansion with (16) and substituting in the Laplace transform $\mathcal{L}_{I_e}(s) = \mathbb{E}_{I_e}[e^{-sI_e}]$ with $s = m\xi/\beta_e$ completes the proof.

4) OT Scheme for CE Case: In this case, the achievable rate of the equivalent wiretap channel from the k-th SBS to the colluding eavesdroppers is $C_{e,k} = \log_2 \left(1 + \sum_{e_j \in \Phi_e} \text{SNR}_{j,k}\right)$ with $\text{SNR}_{j,k}$ given in (5). The secrecy probability can be interpreted as the complement of the probability that all the subfiles are intercepted by the eavesdroppers, which is expressed as

$$p_{s,ce}^{\text{OT}} = 1 - \mathbb{E}_{\Phi_e} \left[\prod_{k=1}^{K} \mathbb{P} \left\{ \sum_{e_j \in \Phi_e} |h_{j,k}|^2 r_{j,k}^{-\alpha} > \frac{\beta_{e,k}}{K\rho} \right\} \right].$$
(17)

Since the locations of the SBSs are deterministic, the distances between the *j*-th eavesdropper and any two SBSs actually are not independent [36]. Hence, the expectation over the PPP Φ_e in (17) cannot be moved inside the product, which makes $p_{s,ce}^{\text{OT}}$ difficult to compute. To facilitate the calculation, we first consider a disc $\mathcal{B}(o, R)$ centered at the origin *o* with a radius *R* and let $\Phi_e^R \triangleq \Phi_e \cap \mathcal{B}_{o,R}$ denote the location set of the eavesdroppers residing in the disc $\mathcal{B}_{o,R}$. We then calculate the inner probability in (17) resorting to a common gamma approximation [37]. Specifically, we approximate the term $X_k = \sum_{e_j \in \Phi_e^R} |h_{j,k}|^2 r_{j,k}^{-\alpha}$ as a gamma RV, the probability density function (PDF) of which is given by

$$f_{X_k}\left(x_k; \upsilon_k, \tau_k\right) = \frac{x_k^{\upsilon_k - 1} e^{-\frac{x_k}{\tau_k}}}{\tau_k^{\upsilon_k} \Gamma(\upsilon_k)}.$$
(18)

The parameters v_k and τ_k can be derived from matching the first and second moments of X_k , which are given in the following lemma with the detailed calculation relegated to Appendix A.

Lemma 2: For fixed positions of eavesdroppers in the disc $\mathcal{B}(o, R)$, v_k and τ_k are given by

$$v_{k} = \frac{\left(\sum_{e_{i} \in \Phi_{e}^{R}} r_{i,k}^{-\alpha}\right)^{2}}{\sum_{e_{j} \in \Phi_{e}^{R}} r_{j,k}^{-2\alpha}}, \ \tau_{k} = \frac{\sum_{e_{i} \in \Phi_{e}^{R}} r_{i,k}^{-2\alpha}}{\sum_{e_{j} \in \Phi_{e}^{R}} r_{j,k}^{-\alpha}}.$$
 (19)

For a PPP, the probability of having J eavesdroppers inside the disc $\mathcal{B}_{o,R}$ is given by [39]

$$\mathcal{O}_J \triangleq \mathbb{P}\{n = J\} = e^{-\pi\lambda_e R^2} \frac{\left(\pi\lambda_e R^2\right)^J}{J!}.$$
 (20)

Hence, the secrecy probability in (17) can be rewritten as

$$p_{s,ce}^{\text{OT}} = 1 - \lim_{R \to \infty} \sum_{J=1}^{\infty} \mathcal{O}_J \mathbb{E}_{\Phi_e^R} \left[\prod_{k=1}^K \mathbb{P}\left\{ X_k > \frac{\beta_{e,k}}{K\rho} \right\} \left| \Phi_e^R, J \right].$$
(21)

The exact $p_{s,ce}^{OT}$ is provided by the following theorem.

Theorem 2: The secrecy probability in the OT scheme for the CE case is given by

$$p_{s,ce}^{\text{OT}} = 1 - \lim_{R \to \infty} \sum_{J=1}^{\infty} \frac{\lambda_e^J e^{-\pi\lambda_e R^2}}{J!} \times \left(\int_o^R \int_0^{2\pi} \prod_{k=1}^K \frac{\Gamma\left(\upsilon_k, \frac{\beta_{e,k}}{K\rho\tau_k}\right)}{\Gamma\left(\upsilon_k\right)} r dr d\theta \right)^J$$
(22)

where v_k and τ_k are given in (19).

Proof 3: Recalling the PDF of X_k in (18), the inner probability in (21) can be given by

$$\mathbb{P}\left\{X_k > \frac{\beta_{e,k}}{K\rho}\right\} = \frac{1}{\Gamma\left(\upsilon_k\right)}\Gamma\left(\upsilon_k, \frac{\beta_{e,k}}{K\rho\tau_k}\right).$$
(23)

Conditioned on having J eavesdroppers in the disc $\mathcal{B}_{o,R}$, the distribution of the eavesdroppers' locations follows a binomial point process (BPP). Using the i.i.d. property of a BPP, the joint PDF of the distances $r_k = [r_{1,k}, \cdots, r_{J,k}]^{\mathrm{T}}$ and angles $\boldsymbol{\theta}_k = [\theta_{1,k}, \cdots, \theta_{J,k}]^{\mathrm{T}}$ is given by

$$f_{\boldsymbol{r}_k,\boldsymbol{\theta}_k}(r_1,\cdots,r_J,\theta_1,\cdots,\theta_J) = \prod_{j=1}^J \frac{r_j}{\pi R^2}.$$
 (24)

Substituting (20), (23), and (24) into (21) completes the proof.

Although Theorem 2 does not give a closed-form expression for the secrecy probability, it yields a general and exact result without requiring time-consuming Monte Carlo simulations. Furthermore, it provides a benchmark for comparison with other approximate results.

If the distance between any two adjacent SBSs is large enough, the correlation of the distances between an eavesdropper and any two SBSs can be ignored due to the random mobility of eavesdroppers. In other words, the positions of eavesdroppers seen from different SBSs can be regarded as independent PPPs $\Phi_{e,k}$ with the same density λ_e . In this case, the secrecy probability in (17) can be recast as

$$p_{s,ce}^{\text{OT}} = 1 - \prod_{k=1}^{K} \mathbb{P}\left\{\sum_{e_j \in \Phi_{e,k}} K\rho |h_{j,k}|^2 \tilde{r}_{j,k}^{-\alpha} > \beta_{e,k}\right\}, \quad (25)$$

where $\tilde{r}_{j,k}$ denotes the distance between the *j*-th eavesdropper and the *k*-th SBS after shifting the coordinate system to place the *k*-th SBS at the origin. Note that $\tilde{r}_{j,k}$ is distinguished from $r_{j,k}$ given in (17). Since the expectation over $\Phi_{e,k}$ is moved inside the product, the computation can be greatly simplified. Let $I_k = K\rho \sum_{e_j \in \Phi_{e,k}} |h_{j,k}|^2 \tilde{r}_{j,k}^{-\alpha}$. We first give the Laplace transform of I_k in the following lemma.

Lemma 3 ([32, Eqn. (8)]): Let $\kappa \triangleq \pi \Gamma(1+2/\alpha) \Gamma(1-2/\alpha)$. The Laplace transform of I_k is

$$\mathcal{L}_{I_k}(s) = \exp\left(-\kappa\lambda_e (K\rho s)^{2/\alpha}\right).$$
(26)

Similar to Theorem 1, a closed-form expression for an upper bound of $p_{s,ce}^{\rm OT}$ is given below.

Theorem 3: The secrecy probability in the OT scheme for the CE case satisfies

$$p_{s,ce}^{\text{OT}} \lesssim 1 - \prod_{k=1}^{K} \sum_{m=0}^{M} \binom{M}{m} (-1)^{m} \mathcal{L}_{I_{k}} \left(\frac{m\xi}{\beta_{e,k}}\right), \qquad (27)$$

where ξ and M have been stated in Theorem 1.

For the special case with $\alpha = 4$, we can further derive a closed-form expression for the exact $p_{s,ce}^{OT}$. The PDF of I_k can be obtained from the inverse Laplace transform of $\mathcal{L}_{I_k}(s)$, i.e.,

$$f_{I_k}^{\alpha=4}(x) = \mathcal{L}_{I_k}^{-1}(s) = \frac{\kappa \lambda_e \sqrt{K\rho}}{2\sqrt{\pi}x^{3/2}} \exp\left(-\frac{\kappa^2 \lambda_e^2 K\rho}{4x}\right).$$
 (28)

Plugging (28) into (25) yields

$$p_{s,ce}^{\text{OT},\alpha=4} = 1 - \prod_{k=1}^{K} \int_{\beta_{e,k}}^{\infty} f_{I_{k}}^{\alpha=4}(x) dx$$

$$\stackrel{\text{(h)}}{=} 1 - \prod_{k=1}^{K} \operatorname{erf}\left(\frac{\kappa\lambda_{e}}{2}\sqrt{\frac{K\rho}{\beta_{e,k}}}\right), \quad (29)$$

where step (h) follows from the substitution $\kappa^2 \lambda_e^2 K \rho / (4x) \rightarrow t^2$. Since $\operatorname{erf}(z) < 1$ increases with z, it is apparent that $p_{s,ce}^{\operatorname{OT}}$ increases with the number of SBSs K and the redundant rate R_e , whereas decreases with the density of eavesdroppers λ_e and the normalized SNR ρ .

IV. SECURE CONTENT DELIVERY PROBABILITY MAXIMIZATION

In this section, we jointly design the optimal redundant rate R_e and the optimal caching assignment proportion ϕ to maximize the overall SCDP \mathcal{P}_{scd} . From the definition of \mathcal{P}_{scd} given in (6), we observe that the problem of maximizing \mathcal{P}_{scd} can be decomposed into two steps: 1) designing the optimal R_e to maximize the SCDP $\mathcal{P}_{scd}^{S} = p_c^{S} p_s^{S}$ for for each scheme $S \in \{JT, OT, CM\}$; 2) designing the optimal ϕ to maximize the overall SCDP \mathcal{P}_{scd} . In what follows, we perform the optimization procedure step by step.

A. Optimization of Redundant Rate R_e

This subsection determines the optimal redundant rate R_e for a target secrecy rate R_s to maximize the SCDP for scheme $S \in \{JT, OT, CM\}$. For tractability, we focus on the NCE case. The optimization in the CE case can be operated similarly, which however would result in a considerable calculation complexity and a much more sophisticated analysis but provide no significant qualitative difference.

1) JT Scheme: The SCDP in this case is defined as the product of the connection probability $p_c^{\rm JT}$ in (8) and the secrecy probability in (10), i.e., $\mathcal{P}_{scd}^{\rm JT} = p_c^{\rm JT} p_s^{\rm JT}$, which can be written as

$$\mathcal{P}_{scd}^{\rm JT} = \exp\left(-A\beta_t - 2\lambda_e \int_0^\infty \int_0^\pi e^{-B(r,\theta)\beta_e} r dr d\theta\right),\tag{30}$$

where $A \triangleq 1/\left(\rho \sum_{k=1}^{K} r_{b,k}^{-\alpha}\right)$ and $B(r,\theta) = 1/\left(\rho \sum_{k=1}^{K} r_{k}^{-\alpha}\right)$. Since we have $R_t = R_s + R_e \Rightarrow \beta_t = \beta_s + (1 + \beta_s)\beta_e$, substituting β_t into (30) yields $\mathcal{P}_{scd}^{\text{JT}} = e^{-A\beta_s}e^{-Q(\beta_e)}$ such that maximizing $\mathcal{P}_{scd}^{\text{JT}}$ is equivalent to minimizing the auxiliary function $Q(\beta_e)$ given below,

$$Q(\beta_e) = A(1+\beta_s)\beta_e + 2\lambda_e \int_0^\infty \int_0^\pi e^{-B(r,\theta)\beta_e} r dr d\theta.$$
(31)

Hence, we focus on the following problem, and the solution is given in Theorem 4,

$$\min_{\beta_e} Q(\beta_e), \quad \text{s.t.} \quad \beta_e > 0. \tag{32}$$

Theorem 4: $Q(\beta_e)$ is convex on β_e , and the solution β_e^* to problem (32) is characterized by

$$\frac{dQ(\beta_e^{\star})}{d\beta_e^{\star}} = 0, \tag{33}$$

i.e., it is the unique zero-crossing of the derivative $dQ(\beta_e)/d\beta_e$ given below

$$\frac{dQ(\beta_e)}{d\beta_e} = A(1+\beta_s) - 2\lambda_e \int_0^\infty \int_0^\pi B(r,\theta) e^{-B(r,\theta)\beta_e} r dr d\theta.$$
(34)

Proof 4: Please refer to Appendix B.

Appendix B shows that $dQ(\beta_e)/d\beta_e$ increases from negative to positive as β_e increases from zero to infinity. Then, the value of β_e^* can be efficiently obtained via a bisection search with equation (33). The following corollary develops some insights into the behavior of β_e^* .

Corollary 1: The optimal β_e^{\star} that maximizes $\mathcal{P}_{scd}^{\text{JT}}$ increases with the eavesdropper density λ_e , and decreases with the secrecy rate R_s and the SBS-user distance $r_{b,k}$ for $k \in \mathcal{K}$.

Proof 5: Let us take λ_e as an example. Denote $dQ(\beta_e)/d\beta_e$ in (34) as $Q_1(\beta_e)$ such that $Q_1(\beta_e^*) = 0$. Using the derivative rule for implicit functions with $Q_1(\beta_e^*) = 0$ yields,

$$\frac{dQ_1(\beta_e^{\star})}{d\lambda_e} = -\frac{\partial Q_1(\beta_e^{\star})/\partial\lambda_e}{\partial Q_1(\beta_e^{\star})/\partial\beta_e} = \frac{\int_0^\infty \int_0^\pi B(r,\theta) e^{-B(r,\theta)\beta_e} r dr d\theta}{\lambda_e \int_0^\infty \int_0^\pi B^2(r,\theta) e^{-B(r,\theta)\beta_e} r dr d\theta} > 0.$$
(35)

Hence, β_e^{\star} increases with λ_e . The other conclusions can be obtained in a similar way.

Corollary 1 captures an inherent trade-off between the reliability and secrecy. We should choose a large redundant rate for dense eavesdroppers, whereas we should keep redundant rate low for a large target secrecy rate or for a remote user

2) OT Scheme: The SCDP in this case is can be obtained by computing the product of the connection probability p_c^{OT} in (9) and the secrecy probability p_s^{OT} in (11), which is given below,

$$\mathcal{P}_{scd}^{\text{OT}} = \exp\left(-\frac{1}{K\rho}\sum_{k=1}^{K}r_{b,k}^{\alpha}[\beta_{s} + (1+\beta_{s})\beta_{e,k}] - 2\lambda_{e}\int_{0}^{\infty}\int_{0}^{\pi}e^{-\frac{1}{K\rho}\sum_{k=1}^{K}r_{k}^{\alpha}\beta_{e,k}}rdrd\theta\right).$$
 (36)

Compared with the JT scheme, the SBSs in the OT scheme can use different redundant rates for maximizing the SCDP. We assume that the knowledge of SBS-user distances is known at the SBSs and can be exploited to determine the optimal redundant rates at different SBSs. In order to jointly design the redundant rates $R_{e,k}$ for $k \in \mathcal{K}$, we recast (36) into a vector form, i.e., $\mathcal{P}_{scd}^{OT} = e^{-\frac{\beta_s}{\mathcal{K}_{P}} \|\mathbf{r}_b\|_1} e^{-\Omega(\beta_e)}$, where the auxiliary function $\Omega(\beta_e)$ is given as below,

$$\Omega(\boldsymbol{\beta}_e) = \frac{1+\beta_s}{K\rho} \boldsymbol{r}_b^{\mathrm{T}} \boldsymbol{\beta}_e + \lambda_e \int_0^\infty \int_0^{2\pi} e^{-\frac{1}{K\rho} \boldsymbol{r}_e^{\mathrm{T}} \boldsymbol{\beta}_e} r dr d\theta, \quad (37)$$

with $\boldsymbol{r}_b = [r_{b,1}^{\alpha}, \cdots, r_{b,K}^{\alpha}]^{\mathrm{T}} \geq 0$, $\boldsymbol{r}_e = [r_1^{\alpha}, \cdots, r_K^{\alpha}]^{\mathrm{T}} \geq 0$, and $\boldsymbol{\beta}_e = [\boldsymbol{\beta}_{e,1}, \cdots, \boldsymbol{\beta}_{e,K}]^{\mathrm{T}} \geq 0$. Apparently, to maximize $\mathcal{P}_{scd}^{\mathrm{OT}}$ we only need to tackle the following minimization problem,

$$\min_{\boldsymbol{\beta}_e} \Omega(\boldsymbol{\beta}_e), \quad \text{s.t.} \quad \boldsymbol{\beta}_e \ge 0.$$
(38)

We point out that the objective function $\Omega(\beta_e)$ in (37) is strictly convex on β_e due to the summation of an affine function and an integral with exponential terms. Generally, problem (38) can be numerically resolved using some gradient methods, e.g., Newton's method [41]. However, Newton's method requires forming and storing the Hessian matrix repeatedly and the computation of the Newton step requires solving a set of linear equations. All these operations will bring the system a huge computational burden and thus resulting in a low system efficiency, particularly when the number of SBSs K goes large. To reduce the computational complexity, we propose to process problem (38) through an alternating optimization (AO) as described below. Denote $\left(\beta_{e,1}^{(n)}, \dots, \beta_{e,K}^{(n)}\right)$ as the AO iterate at the *n*-th iteration, and let $\hat{\beta}_{e,k}^{(n)} = \beta_e^{(n)} \setminus \beta_{e,k}^{(n)}$. We solve the following K subproblems alternatively to obtain $\left(\beta_{e,1}^{(n)}, \dots, \beta_{e,K}^{(n)}\right)$ for $n = 1, 2, \dots$

$$\beta_{e,k}^{(n)} = \arg\min_{\beta_{e,k} \ge 0} \Omega\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}\right).$$
(39)

Lemma 4: When the values of $\beta_{e,j}$ for $j \neq k$ are given, the solution to problem (39) is

$$\beta_{e,k}^{\star} = \begin{cases} 0, & \frac{dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}\right)}{d\beta_{e,k}} \big|_{\beta_{e,k}=0} \ge 0, \\ \beta_{e,k}^{\circ}, & \text{otherwise}, \end{cases}$$
(40)

where $\beta_{e,k}^{\circ}$ is the unique zero-crossing of the derivative given below,

$$\frac{dQ\left(\hat{\beta}_{e,k}^{(n)},\beta_{e,k}\right)}{d\beta_{e,k}} = \frac{(1+\beta_s)r_{b,k} - 2\lambda_e \int_0^\infty \int_0^\pi r_{e,k} e^{-\frac{r_e^T \beta_e}{K\rho}} r dr d\theta}{K\rho}$$
(41)

Proof 6: Since $Q\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}\right)$ in (37) is a convex function of $\beta_{e,k}$, it arrives at the minimal value at $\beta_{e,k} = 0$ if $\left(dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}\right)/d\beta_{e,k}\right)|_{\beta_{e,k}=0} \ge 0$ or at the zero-crossing of $dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}\right)/d\beta_{e,k}$ otherwise. The value of $\beta_{e,k}^{\circ}$ can be efficiently calculated through a

The value of $\beta_{e,k}^{\circ}$ can be efficiently calculated through a bisection search with $dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}^{\circ}\right)/d\beta_{e,k}^{\circ} = 0$. Lemma 4 suggests that the remote SBS, e.g., with a large

Lemma 4 suggests that the remote SBS, e.g., with a large distance $r_{b,k}$ such that $\left(dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}\right)/d\beta_{e,k}\right)|_{\beta_{e,k}=0} \ge 0$, should set zero redundant rate. With Lemma 4, we summarize the whole AO process in Algorithm 1. Notably, the proposed AO iterative algorithm produces descending objective values, i.e., $\Omega\left(\beta_e^{(n)}\right) < \Omega\left(\beta_e^{(n-1)}\right), \dots, < \Omega\left(\beta_e^{(0)}\right)$. Moreover, it has a theoretically provable guarantee on the global optimality of our solution and its convergence.

Proposition 1: Every limit point β_e^* of the iterates $\left\{\beta_e^{(n)}\right\}$ generated by the AO process in (39) is a Karush-Kuhn-Tucker (KKT) point of the primal problem (38).

Algorithm 1 AO Algorithm for Problem (38)

1: Initialize $n = 1, \ \beta_e^{(0)} \ge 0$, and assign ϵ a sufficiently small positive value, e.g., $\epsilon = 10^{-10}$; 2: Update $\beta_e^{(n)} \leftarrow \beta_e^{(n-1)}$; 2: Optime p_e 3: for k = 1 to K do 4: if $\left(dQ \left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k} \right) / d\beta_{e,k} \right) |_{\beta_{e,k}=0} \ge 0$ then $\beta_{e,k}^{(n)} \leftarrow 0;$ 5: 6: Calculate $\beta_{e,k}^{(n)}$ through a bisection search with the 7: equation $dQ\left(\hat{\beta}_{e,k}^{(n)},\beta_{e,k}^{(n)}\right)/d\beta_{e,k}^{(n)}=0;$ end if 8: Update $\beta_{P}^{(n)}$ 9. 10: end for 11: while $\left| \left[\Omega\left(\boldsymbol{\beta}_{e}^{(n)} \right) - \Omega\left(\boldsymbol{\beta}_{e}^{(n-1)} \right) \right] / \Omega\left(\boldsymbol{\beta}_{e}^{(n-1)} \right) \right| \geq \epsilon$ do 12: Update $n \leftarrow n+1$; Repeat step 2 to step 10; 13: 14: end while 15: Output $\boldsymbol{\beta}_{e}^{(n)}$

Proof 7: Since the objective function $\Omega(\beta_e)$ in (38) is a strictly convex function of β_e , the KKT conditions are necessary and sufficient for the solution to problem (38) [41]:

$$\nabla \Omega \left(\boldsymbol{\beta}_{e}^{\star} \right) - \boldsymbol{\lambda}^{\star} = 0, \tag{42a}$$

$$\boldsymbol{\beta}_{e}^{\star} \geq 0, \ \boldsymbol{\lambda}^{\star} \geq 0, \ \lambda_{k}^{\star} \boldsymbol{\beta}_{e,k}^{\star} = 0, \ k \in \mathcal{K},$$
 (42b)

where λ^{\star} is the Lagrange multiplier introduced for the inequality constraints given in (38). Note that since λ^* acts as a slack variable in (42a), it actually can be eliminated, leaving

$$\boldsymbol{\beta}_{e}^{\star} \geq 0, \ \nabla \Omega\left(\boldsymbol{\beta}_{e}^{\star}\right) \geq 0, \ \boldsymbol{\beta}_{e,k}^{\star} \frac{d\Omega(\boldsymbol{\beta}_{e}^{\star})}{d\boldsymbol{\beta}_{e,k}^{\star}} = 0, \ k \in \mathcal{K}.$$
(43)

Now, let us recall Lemma 4, from which we have

$$\beta_{e,k}^{\star} \ge 0, \ \frac{dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}^{\star}\right)}{d\beta_{e,k}^{\star}} \ge 0, \ \beta_{e,k}^{\star} \frac{dQ\left(\hat{\beta}_{e,k}^{(n)}, \beta_{e,k}^{\star}\right)}{d\beta_{e,k}^{\star}} = 0.$$
(44)

Evidently, (44) gives the KKT conditions for the k-th subproblem (39). The combination of the KKT conditions for the Ksubproblems is exactly the KKT conditions for problem (38).

For a simplified case where all the SBSs use the same redundant rate R_e , a more computation-convenient solution to problem (38) can be provided by the following theorem.

Theorem 5: If all the SBSs use the same redundant rate R_e , the optimal β_e^{\star} that minimizes $Q(\beta_e)$ in (38) is the unique zero-crossing of the derivative $dQ (\beta_e)/d\beta_e$ given below

$$\frac{dQ\left(\beta_{e}\right)}{d\beta_{e}} = \frac{1+\beta_{s}}{K\rho} \|\boldsymbol{r}_{b}\|_{1} - \int_{0}^{\infty} \int_{0}^{2\pi} \frac{\lambda_{e} \|\boldsymbol{r}_{e}\|_{1}}{K\rho} e^{-\frac{\|\boldsymbol{r}_{e}\|_{1}\beta_{e}}{K\rho}} r dr d\theta.$$
(45)

The value of β_{e}^{\star} can be efficiently calculated via a bisection search with equation $dQ (\beta_e)/d\beta_e = 0$. Some insights into the solution β_e^{\star} that are similar to Corollary 1 can be developed.

3) CM Scheme: The SCDP in the CM scheme has the same expression as in the JT scheme, only with R_s increasing to δR_s . Therefore, the optimal β_e that maximizes the SCDP in the CM scheme shares the same form as in Theorem 4, simply by replacing β_s with $2^{\delta R_s} - 1$.

B. Optimization of Caching Assignment Proportion ϕ

This subsection determines the optimal caching assignment proportion ϕ that maximizes the overall SCDP \mathcal{P}_{scd} in (6) with p_{tr}^{S} for $S \in \{JT, OT, CM\}$ given in (7a) and (7b). Note that the summation of discrete sequence aroused by p_{tr}^{S} hampers the optimization of the overall SCDP. Fortunately, the sum of the Zipf probabilities can be approximated as [38]

$$\sum_{n=1}^{M} f_n \approx \frac{M^{1-\gamma} - 1}{N^{1-\gamma} - 1}.$$
 (46)

Invoking (46) with (7a) and (7b) and plugging the obtained results into (6) with the integer $|\phi L|$ replaced with the continuous quantity ϕL , \mathcal{P}_{scd} can be simplified as a continuous function of ϕ ,

$$\mathcal{P}_{scd} \approx \frac{\hat{\mathcal{P}}_{jo}\phi^{1-\gamma} + \hat{\mathcal{P}}_{oc}[K - K\phi + \phi]^{1-\gamma} - \hat{\mathcal{P}}_{jc}L^{\gamma-1}}{L^{\gamma-1}(N^{1-\gamma} - 1)} + \mathcal{P}_{scd}^{\mathrm{CM}},$$
(47)

where $\hat{\mathcal{P}}_{jo} = \mathcal{P}_{scd}^{\text{JT}} - \mathcal{P}_{scd}^{\text{OT}}$, $\hat{\mathcal{P}}_{oc} = \mathcal{P}_{scd}^{\text{OT}} - \mathcal{P}_{scd}^{\text{CM}}$, and $\hat{\mathcal{P}}_{jc} = \mathcal{P}_{scd}^{\text{JT}} - \mathcal{P}_{scd}^{\text{CM}}$ denote the SCDP differences, with $\mathcal{P}_{scd}^{\text{S}}$ being the SCDP for the scheme $S \in \{JT, OT, CM\}$. Before proceeding to derive the optimal ϕ that maximizes \mathcal{P}_{scd} , we give the following lemma.

Lemma 5: The JT scheme gives a larger SCDP than does

the CM scheme, i.e., $\mathcal{P}_{scd}^{\mathrm{JT}} > \mathcal{P}_{scd}^{\mathrm{CM}}$. *Proof 8:* Note that $\mathcal{P}_{scd}^{\mathrm{CM}}$ shares the same expression as $\mathcal{P}_{scd}^{\mathrm{JT}}$ in (30) only with β_s increasing from $2^{R_s} - 1$ to $2^{\delta R_s} - 1$. Then, to complete the proof we only need to prove that the maximal $\mathcal{P}_{scd}^{\mathrm{JT}}$ with the optimal β_{e}^{\star} given in Theorem 4 decreases with β_{s} . By re-expressing $\mathcal{P}_{scd}^{\mathrm{JT}} = e^{-W(\beta_{s})}$ where $W(\beta_{s}) = A\beta_{s} + \frac{1}{2} \sum_{scd} \frac{1}{2} \sum_{scd}$ $Q(\beta_e^{\star}(\beta_s))$ with A and $Q(\beta_e^{\star}(\beta_s))$ given in (31), it is also equivalent to proving that $W(\beta_s)$ is an increasing function of β_s . The derivative $dW(\beta_s)/d\beta_s$ is given by

$$\frac{dW(\beta_s)}{d\beta_s} = A(1+\beta_e^*) + \frac{d\beta_e^*}{d\beta_s} \frac{dQ(\beta_e^*)}{d\beta_e^*}.$$
 (48)

From (33) we know that $dQ(\beta_e^{\star})/d\beta_e^{\star} = 0$. Hence, we have $dW(\beta_s)/d\beta_s > 0.$

Theorem 6: With the proposed hybrid caching placement strategy, the optimal proportion ϕ that maximizes \mathcal{P}_{scd} in (47) is given by

$$\phi^{\star} = \begin{cases} 1, & \hat{\mathcal{P}}_{jo} > (K-1)\hat{\mathcal{P}}_{oc}, \\ 0, & \hat{\mathcal{P}}_{jo} < 0, \\ \frac{1}{1 + \frac{1}{K} \left[\left(\frac{(K-1)\hat{\mathcal{P}}_{oc}}{\hat{\mathcal{P}}_{jo}} \right)^{\frac{1}{\gamma}} - 1 \right]}, & \text{otherwise.} \end{cases}$$
(49)

Proof 9: Please refer to Appendix C.

Theorem 6 shows that the SCDP difference between different transmission schemes is critical to the optimal caching assignment. Specifically, when \mathcal{P}_{scd}^{CM} exceeds \mathcal{P}_{scd}^{OT} or when $\mathcal{P}_{scd}^{JT} - \mathcal{P}_{scd}^{OT}$ is K-1 times larger than $\mathcal{P}_{scd}^{OT} - \mathcal{P}_{scd}^{CM}$, we have



Fig. 2: Model used for experiments. K SBSs are placed along the horizontal axis with an identical distance D, and the k-th nearest SBS to the origin is located at (0, (k-1)D). The user moves along the horizontal direction with a location $(X_u, d_0/2)$. Eavesdroppers are randomly distributed according to a PPP.



Fig. 3: Secrecy probability $p_{s,nce}$ vs. R_e for different values of λ_e , with K = 3.

 $\phi^* = 1$, meaning that caching the MPFs is more conductive. As \mathcal{P}_{scd}^{OT} increases, the optimal ϕ^* becomes smaller, i.e., a larger proportion of the cache unit should be assigned for the DSFs. We also can prove that the optimal ϕ^* increases with the content popularity skewness γ . The reason behind is that as the content popularity becomes more concentrated (i.e., a larger γ), the benefit of caching different files becomes limited.

V. SIMULATION RESULTS

In this section, we present simulation results to validate our theoretical analysis. For simplicity, we consider a twodimensional system model as illustrated in Fig. 2. Without loss of generality, we set a reference distance $d_0 = 100$ m and a reference density $\lambda_0 = 10^{-6}$ nodes/m². Unless specified otherwise, we fix the normalized SNR $\rho = P/(WN_0) = 10$ dB, the path-loss exponent $\alpha = 4$, the distance between two adjacent SBSs $D = d_0$, the content library size N = 100, and the SBS caching capacity L = 20.

Fig. 3 and Fig. 4 depict the secrecy probabilities versus the redundant rates R_e in the NCE and CE cases, respectively. The Monte-Carlo simulation results match well with the theoretical values. Both figures verify the superiority of the OT scheme over the JT scheme in terms of transmission secrecy. As expected, the secrecy probability increases with R_e and



Fig. 4: Secrecy probability $p_{s,ce}$ vs. R_e for different values of M, with K = 3, $D = 6d_0$, and $\lambda_e = 3\lambda_0$.



Fig. 5: SCDP \mathcal{P}_{scd}^{JT} vs. R_e for different values of λ_e and R_s , with K = 3, and $X_u = 3d_0$.

decreases with λ_e . Fig. 4 shows that the approximate results given in (15) and (27) coincide well with the real ones when M = 5. We also find that the results in (29) approach the simulated ones in (22), particularly for a large distance D, e.g., $D = 6d_0$. This is because, for a sufficiently large D, the correlation in the distances between an eavesdropper and any two SBSs can be ignored.

Fig. 5 and Fig. 6 plot the SCDP \mathcal{P}_{scd}^{S} as a function of the redundant rate R_e in the JT and OT schemes, respectively. Fig. 5 shows that the SCDP \mathcal{P}_{scd}^{JT} in the JT scheme first increases and then decreases with R_e , just as proved in Theorem 4. We also find that the optimal R_e that maximizes \mathcal{P}_{scd}^{JT} increases with a larger λ_e or a smaller R_s , verifying the insights obtained in Corollary 1. Fig. 6 shows how the SCDP \mathcal{P}_{scd}^{OT} in the OT scheme varies with the redundant rates $R_{e,k}$ at different SBSs when the user is located differently. Sub-figures (a) to (d) show that \mathcal{P}_{scd}^{OT} is dominated by the redundant rate at the closer SBS to the user. This suggests that in order to improve the SCDP, we should set zero redundant rates at those neighboring SBSs, just as indicated in Lemma 4.

Fig. 7 illustrates how the SCDP is affected by the geographical relationship between the user and the SBSs in various



Fig. 6: SCDP \mathcal{P}_{scd}^{OT} vs. $R_{e,1}$ and $R_{e,2}$ for $X_u = \{0.3, 0.5, 0.8, 1.5\}d_0$ in (a)-(d), with K = 2, $\lambda_e = \lambda_0$, and $R_s = 1$ bit/s/Hz.



Fig. 7: SCDP $\mathcal{P}_{scd}^{\mathrm{S}}$ for $\mathrm{S} \in \{\mathrm{JT}, \mathrm{OT}, \mathrm{CM}\}$ vs. X_u for different values of λ_e , with K = 3, $\delta = 2$, and $R_s = 1$ bit/s/Hz.

transmission schemes. With the experimental model described in Fig. 2 in mind, we find that for the JT scheme, if only the user moves close to one of the SBSs, the SCDP can be remarkably improved. Whereas for the OT scheme, only if the user is located approximately at the center of the SBSs, or there is no significant difference among the distances between the user and different SBSs, a high SCDP can be achieved; otherwise the SCDP performance is severely degraded. In addition, Fig. 7 shows that the JT scheme outperforms the CM scheme, just as proved in Lemma 5. However, whether the JT or the OT scheme is superior depends on the specific transmission environment. For example, as can be seen from Fig. 7, in a sparse eavesdropper scenario, the JT scheme outperforms the OT scheme; whereas in a dense eavesdropper case and when the user is located at the center of the SBSs, the OT scheme provides a higher benefit for the secure content delivery.

Fig. 8 shows how the optimal caching placement is influenced by the user's location and the eavesdropper density. When eavesdroppers are distributed sparsely (a small λ_e) or when the user is located far away from the SBSs, the SCDP is bottlenecked by the connection probability. This suggests that the SBSs should adopt the JT scheme to improve transmission



Fig. 8: Optimal caching proportion ϕ^* vs. X_u for different values of λ_e and γ , with K = 3, $\delta = 2$, and $R_s = 1$ bit/s/Hz.



Fig. 9: Optimal caching assignment proportion ϕ^* vs. R_s for different K and λ_e , with $\gamma = 1.2$, and $\delta = 3$.

reliability. Hence, the optimal ϕ^* goes to one, meaning that caching the MPFs is more beneficial. When the eavesdropper density λ_e increases or the user moves close to the SBSs, the secrecy probability becomes the major bottleneck for improving the SCDP. In order to guarantee a certain level of secrecy, the OT scheme is preferred. Hence, the optimal ϕ^* approaches zero, suggesting that the SBSs should store the DSFs geographically. For a moderate density of eavesdroppers or SBS-user distance, the optimal ϕ^* lies between zero and one, showing that there exists a trade-off between caching the MPFs and the DSFs. We also find that the optimal ϕ^* increases with the content popularity skewness γ , which coincides with the finding given in Theorem 6.

Fig. 9 depicts the optimal caching proportion ϕ^* versus the target secrecy rate R_s . Monte-Carlo simulated results match well with the theoretical values, verifying the accuracy of the approximation in (47). We show that ϕ^* initially is small at the small R_s region and then increases with R_s . This is because, to support a large R_s , transmission reliability should be adequately ensured and thus the JT scheme becomes favorable. We also observe that ϕ^* increases with the number of SBSs K. This means, with more cooperative SBSs, the benefit from caching the MPFs along with the JT scheme



Fig. 10: Overall SCDP \mathcal{P}_{scd} vs. R_s for different λ_e , with K = 3, $\gamma = 1.2$, and $\delta = 3$.

would be more pronounced.

Fig. 10 compares the overall SCDPs for different caching strategies. We show that the proposed hybrid caching strategy outperforms either the MPF-only or the DSF-only scheme. In particular, hybrid caching can provide a remarkable performance gain over MPF-only caching for a small secrecy rate R_s and a large eavesdropper density λ_e , or over DSF-only caching for a large R_s and a small λ_e . In addition, the SCDP performance without caching severely deteriorates.

VI. CONCLUSIONS AND FUTURE WORK

This paper studies the security issue in a cache-enabled cooperative SCN against randomly located eavesdroppers. We propose a hybrid MPF and DSF caching strategy along with the JT and OT schemes. We derive analytical expressions for the SCDP in each transmission scheme for both NCE and CE scenarios, based on which the optimal transmission rates and the optimal caching assignment are jointly designed for maximizing the overall SCDP. We also develop various insights into the optimal designs. Numerical results demonstrate the superiority of the proposed hybrid caching strategy over the MPF- and DSF-only ones in terms of the SCDP.

This paper opens up several interesting research directions. For example, the proposed analysis and design framework can be extended to investigate the cooperative multi-antenna SBSs in cache-enabled SCNs, where artificial jamming signals can be exploited to confound eavesdroppers. The potential of PLS can be further tapped by performing adaptive designs leveraging the instantaneous CSI of the main channels. Studying the secure content delivery from a network perspective, e.g., considering a multi-cell cellular network, is also an interesting issue. Nevertheless, this might be much more sophisticated, since we should analyze the influence of both random interferers and eavesdroppers. Another possible direction is to consider diverse secrecy attributes for different files, and to explore more intelligent caching strategies.

APPENDIX

A. Proof of Lemma 2

The *i*-th cumulant of the RV X_k is defined as

$$Q_{X_k}^{(i)} = \frac{d^i \mathbb{E}_{X_k} \left[e^{\omega X_k} \right]}{d\omega^i} \Big|_{\omega=0}.$$
 (50)

The mean and variance of X_k are $\mu_{X_k} = Q_{X_k}^{(1)}$ and $\sigma_{X_k}^2 = Q_{X_k}^{(2)} - \left(Q_{X_k}^{(1)}\right)^2$. Hence, the parameters v_k and τ_k in (18) can be calculated as $v_k = \mu_{X_k}^2 / \sigma_{X_k}^2$ and $\tau_k = \sigma_{X_k}^2 / \mu_{X_k}$. Recall that $X_k = \sum_{e_j \in \Phi_e^R} |h_{j,k}|^2 r_{j,k}^{-\alpha}$. Due to the mutual independence among $\{h_{j,k}\}$ for $e_j \in \Phi_e^R$, we can express $\mathbb{E}_{X_k} \left[e^{\omega X_k} \right]$ for a fixed Φ_e^R as follows,

$$\mathbb{E}_{X_k}\left[e^{\omega X_k}\right] = \prod_{e_j \in \Phi_e^R} \mathbb{E}_{h_{j,k}}\left[e^{-\omega |h_{j,k}|^2 r_{j,k}^{-\alpha}}\right].$$
 (51)

Denote $g_{j,k}(\omega) = \mathbb{E}_{h_{j,k}} \left[e^{-\omega |h_{j,k}|^2 r_{j,k}^{-\alpha}} \right]$ such that $\mathbb{E}_{X_k} \left[e^{\omega X_k} \right] = \prod_{e_j \in \Phi_e^R} g_{j,k}(\omega)$ and $Q_{X_k}^{(i)} = \left(\frac{d^i g_{j,k}(\omega)}{d\omega^i} \right)|_{\omega=0}$. Then, we have $g_{j,k}(\omega)|_{\omega=0} = 1$, $\left(\frac{d g_{j,k}(\omega)}{d\omega} \right)|_{\omega=0} = r_{j,k}^{-\alpha}$, and $\left(\frac{d^2 g_{j,k}(\omega)}{d\omega^2} \right)|_{\omega=0} = 2r_{j,k}^{-2\alpha}$. Based on these results, the first and second cumulants of X_k respectively can be calculated as

$$Q_{X_k}^{(1)} = \sum_{e_j \in \Phi_e^R} \prod_{e_l \in \Phi_e^R \setminus e_j} \left(g_{l,k}(\omega) \frac{dg_{j,k}(\omega)}{d\omega} \right) \Big|_{\omega=0} = \sum_{e_j \in \Phi_e^R} r_{j,k}^{-\alpha}$$
(52)

$$Q_{X_{k}}^{(2)} = \sum_{e_{j}\in\Phi_{e}^{R}} \left[\prod_{e_{l}\in\Phi_{e}^{R}\setminus e_{j}} \left(g_{l,k}(\omega) \frac{d^{2}g_{j,k}(\omega)}{d\omega^{2}} \right) + \frac{dg_{j,k}(\omega)}{d\omega} \times \right] \right]$$
$$\sum_{e_{l}\in\Phi_{e}^{R}\setminus e_{j}} \prod_{e_{q}\in\Phi_{e}^{R}\setminus \{e_{j},e_{l}\}} \left(g_{q,k}(\omega) \frac{dg_{l,k}(\omega)}{d\omega} \right) \right] \Big|_{\omega=0}$$
$$= \sum_{e_{j}\in\Phi_{e}^{R}} \left[2r_{j,k}^{-2\alpha} + r_{j,k}^{-\alpha} \sum_{e_{l}\in\Phi_{e}^{R}\setminus e_{j}} r_{l,k}^{-\alpha} \right].$$
(53)

After some algebraic manipulations, we can obtain $\mu_{X_k} = \sum_{e_j \in \Phi_e^R} r_{j,k}^{-\alpha}$ and $\sigma_{X_k}^2 = \sum_{e_j \in \Phi_e^R} r_{j,k}^{-2\alpha}$, and substituting them into the expressions of v_k and τ_k completes the proof.

B. Proof of Theorem 4

(

Recall (34), and the second derivative $d^2Q(\beta_e)/d\beta_e^2$ can be given by

$$\frac{d^2 Q(\beta_e)}{d\beta_e^2} = \lambda_e \int_0^\infty \int_0^{2\pi} B^2(r,\theta) e^{-B(r,\theta)\beta_e} r dr d\theta > 0.$$
(54)

This means that $Q(\beta_e)$ is strictly convex on β_e . Next, we determine the signs of $dQ(\beta_e)/d\beta_e$ at $\beta_e \to \infty$ and $\beta_e = 0$. We can prove that $(dQ(\beta_e)/d\beta_e)|_{\beta_e\to\infty} = A(1+\beta_s) > 0$ from (34), and $(dQ(\beta_e)/d\beta_e)|_{\beta_e=0} < 0$ by realizing that $\mathcal{P}_{scd}|_{\beta_e=0} = 0$ and $\mathcal{P}_{scd}|_{\beta_e>0} > 0$ from (30), respectively. Since $dQ(\beta_e)/d\beta_e$ increases with β_e , there exists a unique β_e that satisfies $dQ(\beta_e)/d\beta_e = 0$, which is the solution to (32).

C. Proof of Theorem 6

Recall (47), and the first and second derivatives of \mathcal{P}_{scd} respectively can be given by

$$\frac{d\mathcal{P}_{scd}}{d\phi} = \frac{\phi^{-\gamma}\hat{\mathcal{P}}_{jo} - (K-1)[K - \phi(K-1)]^{-\gamma}\hat{\mathcal{P}}_{oc}}{\Delta}, \quad (55)$$

$$\frac{d^2 \mathcal{P}_{scd}}{d\phi^2} = -\frac{\phi^{-\gamma - 1} \hat{\mathcal{P}}_{jo} + (K - 1)^2 [K - \phi(K - 1)]^{-\gamma - 1} \hat{\mathcal{P}}_{oc}}{\Delta \gamma^{-1}}$$
(56)

where $\Delta \triangleq L^{\gamma-1} (N^{1-\gamma} - 1)/(1-\gamma) > 0$. Since $\mathcal{P}_{scd}^{JT} > \mathcal{P}_{scd}^{CM}$ always holds, we can determine the optimal ϕ that maximizes \mathcal{P}_{scd} by distinguishing three cases: 1) If $\mathcal{P}_{scd}^{OT} < \mathcal{P}_{scd}^{CM}$, we have $d\mathcal{P}_{scd}/d\phi > 0$, i.e., \mathcal{P}_{scd} monotonically increases with ϕ . Hence, \mathcal{P}_{scd} reaches the maximal value at $\phi^* = 1$; 2) If $\mathcal{P}_{scd}^{OT} > \mathcal{P}_{scd}^{JT}$, we have $d\mathcal{P}_{scd}/d\phi < 0$, i.e., \mathcal{P}_{scd} monotonically decreases with ϕ . Then, the maximal \mathcal{P}_{scd} is achieved at $\phi^* = 0$; 3) If $\mathcal{P}_{scd}^{JT} \ge \mathcal{P}_{scd}^{OT} \ge \mathcal{P}_{scd}^{CM}$, we have $d^2\mathcal{P}_{scd}/d\phi^2 < 0$, i.e., \mathcal{P}_{scd} is concave on ϕ . Since $(d\mathcal{P}_{scd}/d\phi)|_{\phi=0} > 0$, \mathcal{P}_{scd} arrives at the maximal value at $\phi = 1$ if $(d\mathcal{P}_{scd}/d\phi)|_{\phi=1} > 0$ or otherwise at the zero-crossing of the derivative $d\mathcal{P}_{scd}/d\phi$. By now, we have completed the proof.

REFERENCES

- X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. Leung, "Cache in the air: Exploiting content caching and delivery techniques for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131–139, Feb. 2014.
- [2] A. Liu and V. Lau, "Cache-enabled opportunistic cooperative MIMO for video streaming in wireless systems," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 390–402, Jan. 2014.
- [3] L. Xiang, D. W. K. Ng, R. Schober, and V. W. S. Wong, "Cache-enabled physical layer security for video streaming in backhaul-limited cellular networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 736–751, Feb. 2018.
- [4] E. Rescorla, "HTTP over TLS," IETF RFC 2818, May 2000.
- [5] G. Paschos, E. Baştuğ, I. Land, G. Caire, and M. Debbah, "Wireless caching: Technical misconceptions and business barriers," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 16–22, Aug. 2016.
- [6] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [7] H.-M. Wang and T.-X. Zheng, *Physical Layer Security in Random Cellular Networks*. Singapore: Springer Press, 2016.
- [8] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, "Safeguarding 5G wireless communication networks using physical tier security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.
- [9] H.-M. Wang and X.-G. Xia, "Enhancing wireless secrecy via cooperation: Signal design and optimization," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 47–53, Dec. 2015.
- [10] Y. Zou, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.
- [11] X. Zhou, R. Ganti, J. Andrews, and A. Hjørungnes, "On the throughput cost of physical tier security in decentralized wireless networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2764–2775, Aug. 2011.
- [12] X. Zhang, X. Zhou, and M. R. McKay, "Enhancing secrecy with multiantenna transmission in wireless ad hoc networks," *IEEE Trans. Inf. Forensics and Security*, vol. 8, no. 11, pp. 1802–1814, Nov. 2013.
- [13] T.-X. Zheng, H.-M. Wang, J. Yuan, D. Towsley, and M. H. Lee, "Multiantenna transmission with artificial noise against randomly distributed eavesdroppers," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4347–4362, Nov. 2015.
- [14] T.-X. Zheng, H.-M. Wang, Q. Yang, and M. H. Lee, "Safeguarding decentralized wireless networks using full-duplex jamming receivers," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 278–292, Jan. 2017.
- [15] T.-X. Zheng, H.-M. Wang, J. Yuan, Z. Han, and M. H. Lee, "Physical layer security in wireless ad hoc networks under a hybrid full-/halfduplex receiver deployment strategy," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3827–3839, Jun. 2017.

- [16] H.-M. Wang, T.-X. Zheng, J. Yuan, D. Towsley, and M. H. Lee, "Physical layer security in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1204–1219, Mar. 2016.
- [17] Y. Deng, L. Wang, S. A. R. Zaidi, J. Yuan, and M. Elkashlan, "Artificialnoise aided secure transmission in large scale spectrum sharing networks," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2116–2129, May 2016.
- [18] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1656– 1672, Mar. 2017.
- [19] K. Shanmugam, N. Golrezaei, A. Dimakis, A. Molisch, and G. Caire, "FemtoCaching: Wireless content delivery through distributed caching helpers," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.
- [20] D. Liu and C. Yang, "Will caching at base station improve energy efficiency of downlink transmission?" in *Proc. IEEE GlobalSIP*, Atlanta, GA, Dec. 2014.
- [21] L. Xiang, D. W. K. Ng, T. Islam, R. Schober, V. W. S. Wong, and J. Wang, "Cross-layer optimization of fast video delivery in cache- and buffer-enabled relaying networks," *IEEE Trans. Veh. Tech.*, vol. 66, no. 12, pp. 11366–11382, Jun. 2017.
- [22] S. H. Chae, J. Y. Ryu, T. Q. S. Quek, and W. Choi, "Cooperative transmission via caching helpers," in *Proc. IEEE GLOBECOM*, San Diego, CA, USA, Dec. 2015, pp. 1-6
- [23] W. C. Ao and K. Psounis, "Distributed caching and small cell cooperation for fast content delivery," in *Proc. ACM MobiHoc*, Hangzhou, China, Jun. 2015, pp. 127–136.
- [24] X. Peng, J.-C. Shen, J. Zhang, and K. B. Letaief, "Backhaul-aware caching placement for wireless networks," in *Proc. IEEE GLOBECOM*, San Diego, CA, USA, Dec. 2015, pp. 1–6.
- [25] Z. Chen, J. Lee, and M. Kountouris, "Cooperative caching and transmission design in cluster-centric small cell networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3401–3415, Mar. 2017.
- [26] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.
- [27] A. Sengupta, R. Tandon, and T. C. Clancy, "Fundamental limits of caching with secure delivery," *IEEE Trans. Inf. Forensics and Security*, vol. 10, no. 2, pp. 355–370, Feb. 2015.
- [28] Z. H. Awan and A. Sezgin, "Fundamental limits of caching in D2D networks with secure delivery," in *Proc. IEEE ICC Workshops*, London, UK, Jun. 2015.
- [29] M. Gerami, M. Xiao, S. Salimi, and M. Skoglund, "Secure partial repair in wireless caching networks with broadcast channels," in *Proc. IEEE Conf. CNS*, Sep. 2015, pp. 353–360.
- [30] F. Gabry, V. Bioglio, and I. Land, "On edge caching with secrecy constraints," in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May 2016.
- [31] L. Xiang, D. W. K. Ng, R. Schober, and V. W. S. Wong, "Secure video streaming in heterogeneous small cell networks with untrusted cache helpers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2645–2661, Apr. 2018.
- [32] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.
- [33] J.-R. Ohm, "Advances in scalable video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 42–56, Jan. 2005.
- [34] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2196–2211, Oct. 2015.
- [35] H. Alzer, "On some inequalities for the incomplete gamma function," *Math. Comput.*, vol. 66, no. 218, pp. 771–778, Apr. 1997.
- [36] J. Zhang, L. Xiang, D. W. K. Ng, M. Jo, and M. Chen, "Energy efficiency evaluation of multi-tier cellular uplink transmission under maximum power constraint," *IEEE Trans. Wireless Commun.*, vol. 16, no, 11, pp. 7092–7107, Nov. 2017.
- [37] R. W. Heath, M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using Poisson point processes," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4114-4126, Aug. 2013.
- [38] M. Taghizadeh, K. Micinski, S. Biswas, C. Ofria, and E. Torng, "Distributed cooperative caching in social wireless networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 6, pp. 1037–1053, Jun. 2013.
- [39] S. N. Chiu, D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and its Applications, 3rd ed.* John Wiley and Sons, 2013.
- [40] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey, D. Zwillinger, and S. Technica, *Table of Integrals, Series, and Products, 7th ed.* New York: Academic Press, 2007.

[41] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.