Short-Packet Communications in Multi-Hop Networks with WET: Performance Analysis and Deep Learning-Aided Optimization

Toan-Van Nguyen, *Member, IEEE,* Van-Dinh Nguyen, *Member, IEEE,* Daniel Benevides da Costa, *Senior Member, IEEE,* Thien Huynh-The, *Member, IEEE,* Rose Qingyang Hu, *Fellow, IEEE,* and Beongku An, *Member, IEEE*

***Abstract*—In this paper, we study short-packet communications in multi-hop networks with wireless energy transfer, where relay nodes harvest energy from power beacons to transmit short packets to multiple destinations. It is proposed a novel cooperative beamforming relay selection (CRS) scheme which incorporates partial relay selection and distributed multiuser beamforming to achieve a high-reliable transmission in two consecutive hops. A closed-form expression for the average block error rate (BLER) of the CRS scheme is derived, based on which an asymptotic analysis is also carried out. To achieve optimal channel uses allocation, we formulate a fairness end-to-end throughput maximization problem which is generally NP-hard due to the non-concavity of the objective function and mixed-integer constraints. To solve this challenging problem efficiently, we first relax channel uses to be continuous and transform the relaxed problem into an equivalent non-convex one, but with a more tractable form. We then develop a low-complexity iterative algorithm relying on inner approximation framework to convexify non-convex parts that converges to at least a locally optimal solution. Towards real-time settings, we design an efficient deep convolutional neural network (CNN) with multiscale-accumulation connections to achieve the sub-optimal solution of the relaxed problem via real-time inference processes. Numerical results are presented to verify the analytical derivations and to demonstrate performance improvements of the CRS scheme over the benchmark ones in terms of BLER, reliability, latency, and throughput in various settings. Moreover, the designed CNN provides the lowest root-mean-square error compared to the state-of-the-art deep learning approaches while the CNN-aided optimization framework estimates accurately the optimal channel uses allocation with low execution time.**

***Index Terms*—Deep learning, energy harvesting, finite block-length, multi-hop networks, nonconvex optimization.**

The work of T.-V. Nguyen and R. Q. Hu was supported by the National Science Foundation under the grants NSF CNS-2007995 and NSF EEC-1941524. The work of B. An was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2B5B01001190). The work of V.-D. Nguyen was supported in part by the ERC AGNOSTIC project, ref. H2020/ERC2020POC/957570. Part of this paper was accepted for presentation at the IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, December 07-11, 2021. *(Corresponding author: Beongku An.)*

T.-V. Nguyen and R. Q. Hu are with Utah State University, USA (emails: vannguyentoan@gmail.com; rose.hu@usu.edu).

V.-D. Nguyen is with the Interdisciplinary Centre for Security, Reliability and Trust (SnT) – University of Luxembourg, L-1855 Luxembourg (email: dinh.nguyen@uni.lu).

D. B. da Costa is with the Digital Science Research Center, Technology Innovation Institute, 9639 Masdar City, Abu Dhabi, United Arab Emirates (email: danielbcosta@ieee.org).

T. Huynh-The is with Kumoh National Institute of Technology, Republic of Korea (email: thienht@kumoh.ac.kr).

B. An is with Dept. of Software and Communications Engineering, Hongik University, Republic of Korea (e-mail: beongku@hongik.ac.kr).

I. Introduction

Short-packet communication (SPC) has been specified as a key technology for the fifth generation (5G) and beyond 5G wireless networks since it enables ultra-reliable and low-latency communications (URLLCs) with stringent requirements on the delay and reliability [1]. SPC supports a wide range of high-reliable applications, such as factory automation, intelligent transportation systems, high-speed trains, drones, and future Internet-of-Things (IoT) networks [1], [2]. For example, factory automation applications require a latency of less than a millisecond with very low block error rate (BLER) varying from 10^{-9} to 10^{-5} [1], [3]. Applying SPC for multi-hop networks can further improve communication reliability while achieving the benefit of network coverage extension by reducing the path-loss effects [4], [5]. However, the communication systems operating with finite blocklength face serious problems in the information-theoretic performance evaluation and pilot transmission perspectives [6], where the Shannon capacity formula is no longer applicable.

To prolong lifetime of IoT devices in harsh environments and inaccessible locations, which are often equipped with a (very) small battery, wireless energy transfer (WET) technique is an attractive solution for being applied in SPC-based systems because of its characteristics in providing stable energy sources [7]. Thus, SPCs with WET have been investigated in various network scenarios to realize mission-critical IoT applications [8] and to achieve high effective throughput in IoT networks [3]. However, in the context of SPC, the performance analysis in these works has been restricted to single-hop or dual-hop transmissions without considering relay selection schemes. Moreover, delivering SPCs represents challenges for designing resource allocation and network management [9], and that becomes more difficult when deploying SPC and WET in multi-hop networks, possibly resulting in an NP-hard problem. This high complexity may restrict the URLLC applications, where emerging 5G services typically require stringent requirements on latency and reliability for real-time decision making tasks.

A. Literature Survey

URLLCs in cell-free massive multiple-input multiple-output (MIMO) networks were studied in [10], where the max-

min rate and energy efficiency optimization problems were formulated to show characteristics of URLLCs under improper Gaussian signaling. In [11], WET and wireless information transmission (WIT) processes were sequentially performed in the downlink and uplink over quasi-static Nakagami-*m* fading channels assuming URLLCs. It was revealed in [12] that the full-duplex relaying system provides better BLER and latency than the half-duplex one in a relatively low relay transmit power. A cooperative relaying system with finite battery energy, finite blocklength, and imperfect channel state information (CSI) conditions was characterized in [13]. Designing of resource allocation in finite blocklength systems is more challenging than in infinite blocklength ones because of the complicated function of URLLC capacity formula [14]. To overcome this issue, an efficient iterative algorithm was proposed in [3] relying on the block coordinate descent and concave-convex procedure to solve the throughput maximization problem in wireless-powered IoT networks. A joint power control and rate allocation strategy was formulated in [15] to maximize the system energy efficiency. To further improve the output signal-to-interference-and-noise ratio (SINR), a high-reliability receiver structure in full-duplex URLLC systems was designed in [6] to show the improvement of the proposed system over the state-of-the-art methods in terms of frame error rate.

On the other hand, cooperative multi-hop transmission has been extensively studied in infinite blocklength systems to enhance the radio coverage and network performance [16], [17]. Providing SPCs in multi-hop networks is promising to achieve the expansion of radio range while coping with the high delay of delivering heterogeneous services in 5G wireless networks. In [18], a cooperative automatic-repeat-request sharing strategy for URLLC in multi-hop networks was studied, where the reliability communication was improved without compromising on the latency requirement. Recently, SPCs with WET were studied in multi-hop transmissions [19] and in dual-hop transmissions [20], where a deep neural network (DNN) was designed to evaluate the BLER performance. However, simple designs of DNN with one output having high root-mean-square-error (RMSE) may reduce the accuracy of performance prediction when deploying complex network scenarios. Moreover, more advanced relay selection schemes are needed to further improve performance of multi-hop networks.

Data generated by 5G mobile wireless networks is heterogeneously increasing in huge volumes with different types and formats since these data are usually collected from various sources, different devices, and exhibited complex correlations in highly dynamic mobile environments [21]. Therefore, several problems in wireless communications such as resource allocation, channel estimation, and channel decoding [22] become extremely difficult and impractical to be tackled relying on traditional machine learning techniques with shallow neural network models. In contrast, convolutional neural network (CNN) is designed with multiple levels of representation, obtained by composing simple but non-linear modules which transform the representation at one level into a representation at a higher level. With the composition of enough transformations, CNN has the ability to estimate highly non-linear functions at considerably low-complexity from a set of noisy data [22]. Also, CNN has enabled various interesting applications in wireless networks including beamforming optimization [22], channel estimation [23], and power control in cellular massive MIMO systems [24]. However, these works mainly focused on designing simple CNN architectures, which may cause a high estimating error in emerging wireless networks inherent with high dimensional data and complex radio conditions. In this paper, we design a novel CNN architecture to address this issue and to expedite real-time configurations in multi-hop networks with WET since CNN models can precisely estimate desired performance metrics from high dimensional raw data even with complex radio environments.

B. Motivation and Contributions

Most of aforementioned works mainly focused on performance analysis of multi-hop networks with WET under infinite blocklength (IBL) regimes [4], [17] or devoted to studying SPC in dual-hop cooperative networks [11], [13], [20]. In finite blocklength systems, the outage probability obtained by using the traditional Shannon capacity overestimates the communication reliability because it is strictly lower than the average BLER at finite blocklength regimes. Therefore, the assumption of long blocklength transmission appears impractical for future IoT networks, where a large number of smart devices with dense deployment commonly demand short bursts to communicate with each others towards ultra-reliable communications. Moreover, the capacity expression under SPCs is shown as a non-convex and complicated function of the channel uses coefficients [25], making its optimization design in multi-hop multi-relay with WET become more computationally challenging. Such issues have not been addressed in the literature yet.

Against the above background, this paper provides a novel DL-aided real-time optimization approach to maximize the end-to-end (E2E) throughput subject to total channel uses constraints with short execution time and less computational-intensive. By considering multi-hop transmission with the complex capacity expression under SPCs, the formulated optimization problem has high non-convex characteristics, making it quite difficult to solve optimally. Moreover, DL has emerged as a powerful tool to model highly non-linear functions and representing hard-to-model problems in modern wireless communication systems at low computation and cost-efficiency [20], [26]. Thus, we leverage the recent achievements of DL to design a novel deep CNN with multiple convolutional blocks, where each block is equipped with a structural multiscale-accumulation (MA) connection to effectively gather the highly relevant features of each block at multiple signal resolutions. Such versatile structures can provide a high accurate inference for optimal channel uses and throughput without running any complex optimization algorithms, as considered in [9], [27]. The proposed DL-based optimization framework gives a fresh and practical perspective on real-time configurations for multi-hop multi-relay networks with SPCs.

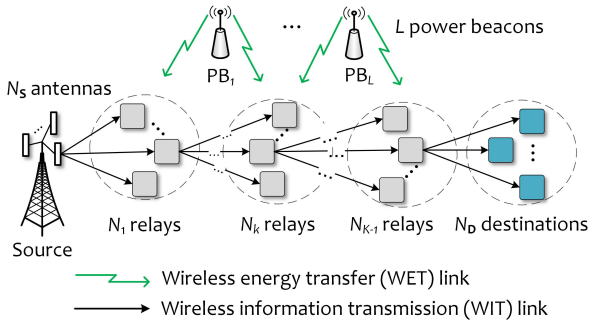
To analyze practical IoT-aided URLLCs, this paper studies SPCs in multi-hop networks with WET, where multiple smart devices can harvest and accumulate energy from multiple power beacons for packet transmissions. With respect to our earlier work on multi-hop SPCs [19], this paper aims to design a novel cooperative beamforming relay selection (CRS) scheme, formulate a new optimization problem, and develop an efficient CNN-aided optimization framework for enhancing system performances by incorporating the following contributions:

- We propose a cooperative beamforming relay selection scheme to improve the average BLER, reliability, and latency in multi-hop networks with WET, which has not studied in [19]. The best relay selection (BRS) and random relay selection (RRS) are also investigated to serve as benchmark schemes for comparison purposes.
- We derive a closed-form expression for the average BLER of the proposed CRS under the considered system setup. Asymptotic analysis is carried out to obtain more computationally tractable forms, which also provide more insights into the system design.
- We formulate the problem of max-min E2E throughput optimization subject to channel uses constraints for the considered system. This problem is highly non-convex with mixed-integer constraints involved, which is very challenging to solve optimally. By leveraging inner approximation method [28], we first relax channel uses (integer variables) to be continuous, and transform the relaxed problem into an equivalent non-convex one but with a more tractable form. We then develop a low-complexity iterative algorithm to generate a sequence of improved feasible points which rapidly converges to at least a locally optimal solution.
- Towards real-time settings, we develop a novel CNN framework to obtain the optimal solution of throughput maximization based on the developed inner approximation method. The CNN is designed with multiple convolutional blocks with structural multiscale-accumulation connection to effectively gather the highly relevant features of each block at multiple signal resolutions, which accurately predicts the real-time channel uses solution and system throughput.
- Numerical results are provided to verify the analytical results and show the performance improvement of the CRS scheme over BRS and RRS ones in terms of average BLER, reliability, and latency, where the significant impact of channel uses, message size, and the number of hops on the system performance is revealed. Moreover, the CRS scheme with optimal allocation of channel uses is more robust than its counterpart with equal allocation in terms of E2E throughput, especially at high signal-to-noise ratios (SNRs) and large number of PBs deployments. Furthermore, the designed CNN exhibits the lowest RMSE compared to state-of-the-art DL approaches and the DL-based optimization framework can achieve the same optimal channel uses solution as the inner approximation method but with very low execution time, which enables real-time decision-making tasks in the future wireless networks.

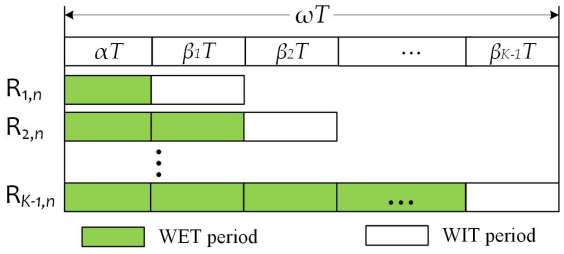
Mathematical Notations and Functions: Boldface lower- and upper-case letters represent vectors and matrices, respectively. $\|.\|$ and $(\cdot)^H$ symbolize the Frobenius norm and transpose conjugate, respectively. Trace$(\mathbf{X})$ denotes the trace of a matrix $\mathbf{X}$. $\langle \mathbf{X}, \mathbf{Y} \rangle$ presents the inner product which is defined as Trace$(\mathbf{X}^H\mathbf{Y})$. $\nabla_{\mathbf{x}} f(\mathbf{x})$ denotes the gradient of $f(.)$ with respect to vector $\mathbf{x}$. $\mathbb{E}\{\cdot\}$ and $\mathbb{N}$ are the expectation operator and the non-negative integer set, respectively. $\Gamma(.)$ and $\gamma(.,.)$ represent the Gamma function [29, Eq. 8.310.1] and the lower incomplete Gamma function [29, Eq. (8.350.1)], respectively. $G_{a,b}^{c,d}[.|.]$ stands for the Meijer's G-function [29, Eq. 9.301] and $\mathcal{K}_v(.)$ designates the v-order modified Bessel function of second kind [29, Eq. (8.432)].

II. System Model

A. System Description and Operation



(a) Illustration of the proposed multi-hop network with WET.



(b) Time block structure for WET and WIT processes.

Fig. 1: Illustration of system model and time block structure of the proposed network.

We consider a multi-hop network with WET, where a source (S) equipped with N_{S} antennas transmits data packets to multiple destination (D) nodes via multiple IoT devices located in $(K-1)$ clusters, as shown in Fig. 1(a). We assume that each IoT device acts as a relay and uses decode-and-forward protocol to transmit short packets over multi-hop network. The relay, denoted by $\mathsf{R}_{k,n}$, is the n-th relay node in the k-th cluster belonging to a set of relays $\mathcal{R} = \{\mathsf{R}_{k,n} \mid k = 1, \ldots, K, n = 1, \ldots, N_k\}$, where $\mathsf{R}_{K,n} \equiv \mathsf{D}_n$. Since IoT devices have limited power supplies, they can harvest energy from the radio-frequency (RF) signals radiated by a set of power beacons $\mathcal{P} = \{\mathsf{PB}_l \mid l = 1, \ldots, L\}$. In this system, S can be seen as an access point and uses its own

energy while IoT devices are equipped with a single-antenna and operate on half-duplex mode. It is assumed that perfect knowledge of channel state information (CSI) is available at each receiver terminal [1]. Moreover, IoT nodes are usually deployed in rich-scattering environments; thus, all wireless links are assumed to follow Rayleigh fading channels [2] [9], [13], [15], [20].

We consider a channel from transmitter X to receiver Y, where $\mathsf{X} \in \{\mathsf{S}, \mathsf{PB}_l, \mathsf{R}_{k,n}\}$ and $\mathsf{Y} \in \{\mathsf{S}, \mathsf{R}_{k+1,j}, \mathsf{D}_n\}$. Let u_{XY} be the channel coefficient of the link $\mathsf{X} \rightarrow \mathsf{Y}$. The channel gain $|u_{\mathsf{XY}}|^2$ follows an exponential distribution, whose cumulative distribution function (CDF), $F_{|u_{\mathsf{XY}}|^2}(x)$, and probability density function (PDF), $f_{|u_{\mathsf{XY}}|^2}(x)$, are $F_{|u_{\mathsf{XY}}|^2}(x) = 1 - \exp(-\frac{x}{\lambda_{\mathsf{XY}}})$ and $f_{|u_{\mathsf{XY}}|^2}(x) = \frac{1}{\lambda_{\mathsf{XY}}} \exp(-\frac{x}{\lambda_{\mathsf{XY}}})$, respectively, where λ_{XY} is the mean of $|u_{\mathsf{XY}}|^2$. Taking into account large-scale pathloss, λ_{XY} can be modeled as $\lambda_{\mathsf{XY}} = \mathcal{L}(d_{\mathsf{XY}}/d_0)^{-\sigma_{\mathrm{PL}}}$ [5], where d_{XY}, σ_{PL}, d_0, and $\mathcal{L}$ present the distance (in meter) between X and Y, the pathloss exponent, the reference distance, and the measured pathloss at d_0, respectively.

B. Wireless Energy Transfer Process

It is considered SPCs in multi-hop networks with two consecutive phases including WET and WIT, respectively [7], [11]. We denote by ω and T the total number of channel uses and duration of each channel use, respectively; hence, the entire period of multi-hop transmission is ωT. Moreover, the channel use of WET is denoted by α; thus, the period of αT is spent for WET and the remaining one, $(\omega - \alpha)T$, is used for WIT process, as sketched in Fig. **??**(b). During the WET period, all relay nodes simultaneously harvest energy from L PBs while the remaining one, $(\omega - \alpha)T$, is equally allocated to K hops transmission, thanks to the time-division multiple access protocol in medium access control layer. To enhance the energy harvesting (EH) capability at each relay, the *accumulate-then-transmit* protocol is employed, where the last relays have more opportunities to continuously harvest and accumulate energy before turning to data transmission. The last time period $\beta_k T$ at relay $\mathsf{R}_{k,i}$ is preserved for its data transmission. Without loss of generality, the total time period of harvesting and accumulating energy at $\mathsf{R}_{k,i}$ can be expressed as $\alpha T + \sum_{m=1}^{k-1} \beta_m T$. Thus, the harvested energy at $\mathsf{R}_{k,i}$ from L PBs can be expressed as

$$E_{k,i} = \sum_{l=1}^{L} \eta \left(\alpha + \sum_{m=1}^{k-1} \beta_m \right) T P |g_{k,l,u}|^2, \tag{1}$$

where $\eta \in (0, 1)$ and P denote the energy conversion efficiency and transmit power of each power beacon, respectively. At each relay employing *accumulate-and transmit* protocol, we assume that the input RF circuit power operates in the linear regions of the rectifier, and consider a linear EH model has been widely adopted for WET in the literature [3], [7], [11], [13], [17].

C. Wireless Information Transmission Process

As mentioned earlier, the WIT process is performed at the end of WET process and takes place during the time period of $(\omega - \alpha)T/K$. The transmit power of $\mathsf{R}_{k,i}$ over the WIT period can be expressed as

$$P_{k,i} = \frac{E_{k,i}}{\beta_k T} = \sum_{l=1}^{L} \chi_k P |g_{k,l,i}|^2, \tag{2}$$

where $\chi_k \triangleq \eta\big(\alpha + \sum_{m=1}^{k-1} \beta_m\big)/\beta_k$. In the first-hop transmission, the S employs an efficient maximal-ratio transmission (MRT) criterion based on beamforming design to transmit its packets to the relay $\mathsf{R}_{1,i}$ located in cluster 1. The instantaneous SNR at $\mathsf{R}_{1,i}$ can be expressed as

$$\gamma_{1,i} = \bar{\gamma}_1 |\mathbf{h}_{1,i}^H \mathbf{w}|^2, \tag{3}$$

where $\mathbf{w} \in \mathbb{C}^{N_{\mathsf{S}}}$ and $\mathbf{h}_{1,i} \in \mathbb{C}^{N_{\mathsf{S}}}$ represent the beamforming and channel coefficient vectors between S and $\mathsf{R}_{1,i}$, respectively, $\bar{\gamma}_1 \triangleq P_{\mathsf{S}}/\sigma^2$, with P_{S} and σ^2 being transmit power at S and the noise variance, respectively. Based on [4], [8], the beamforming vector is designed based on the channel between S and $\mathsf{R}_{1,i}$, which is formulated as $\mathbf{w} = \mathbf{h}_{1,i}/\|\mathbf{h}_{1,i}\|$.

D. Cooperative Beamforming Relay Selection Scheme

The CRS scheme tries to improve the diversity transmission of two consecutive hops, where the cooperative beamforming is sequentially employed with the partial relay selection. We assume that all relays in each cluster successfully decode and store the received short-packets in its memory cache with infinite buffer size and they are represented as an active relay set. We consider that at any two consecutive hops, which includes three clusters, where the partial relay selection and the distributed beamforming are sequentially performed in two sub-time slots. The partial relay selection is used in the first sub-time slot to select the best available relay among active relays in the second cluster for improving diversity of data reception. In the second sub-time slot, all active relays in the second cluster collaborate together to form a distributed multiuser beamforming to transmit packets to relays located in the third cluster. Based on the channel state information exchanged locally in the third cluster, the relays in the third cluster collaborate to select the best relay which has the strongest link between the active relay set and itself. It is noted that the relay in the first cluster has already been selected in the previous transmissions. Based on this strategy, two practical scenarios can be accounted: (i) the number of cluster is odd; and (ii) the number of clusters is even. This paper considers the latter one since it is a general scenario and covers the former one. Therefore, the number of hops in scenario (ii) is odd and we will consider the packet transmission in the first single-hop and the remaining $(K-1)/2$ dual-hops. In the first-hop transmission, the relay in cluster 1 providing the highest data channel gain from the source is selected to be the forwarder

[1]The perfect CSI may not available in practice, which results in the error and delay feedback. However, this error can be compensated by increasing the average transmit feedback power [30].

[2]It is worth noting that the Rayleigh fading is the worst wireless fading channel compared to the Rician and Nakagami ones, if the system can operate under the Rayleigh fading channels, it also works well with other fading channels.

of the next hop. Thus, the instantaneous SNR at the selected relay R_{1,i^*} can be expressed as

$$\gamma_{1,i^*} = \max_{i=1,\ldots,N_1} \bar{\gamma}_1 |\mathbf{h}_{1,i}^H \mathbf{w}|^2 = \max_{i=1,\ldots,N_1} \bar{\gamma}_1 \|\mathbf{h}_{1,i}\|^2. \quad (4)$$

In the remaining $(K-1)/2$ dual-hop transmissions, we assume that the best relay at cluster $k-2$ is R_{k-2,n^*}; thus, the criterion can be expressed as

$$(\mathsf{R}_{k-1,i^*}, \mathsf{R}_{k,j^*}) = \arg\min \Big\{ \max_{i=1,\ldots,N_{k-1}} P_{k-1,i} |h_{k-1,n^*,i}|^2, \max_{j=1,\ldots,N_k} P_{k,j} |\mathbf{h}_{k,j}^H \mathbf{w}_k|^2 \Big\}. \quad (5)$$

Based on [4], we consider the optimal beamforming design as $\mathbf{w}_k = \mathbf{h}_{k,j}/\|\mathbf{h}_{k,j}\|$. The respective instantaneous SNR at hop c, with $c = 2, \ldots, (K+1)/2$ can be expressed as

$$\gamma_c = \min \Big\{ \max_{i=1,\ldots,N_{k-1}} P_{k-1,i} |h_{k-1,n^*,i}|^2, \max_{j=1,\ldots,N_k} P_{k,j} \|\mathbf{h}_{k,j}\|^2 \Big\}. \quad (6)$$

It is noted that the CRS scheme requires the same number of CSI estimations at relays as the best relay selection in [19] but it achieves a high coding gain and reliable transmission.

III. BLER Performance Analysis

In this section, the performance of SPCs in multi-hop networks with WET is analyzed under finite blocklength transmission condition. Considering SPC at hop k, where a packet of m information bits at $\mathsf{R}_{k-1,i}$ is transmitted to $\mathsf{R}_{k,j}$ over the blocklength β_k (the number of channel uses), with $\beta_k > 100$ [7], received SNR equals to γ_k, and a fixed transmission rate given by $R_k = m/\beta_k$ can be accurately approximated as [8], [11]

$$R_k \triangleq C(\gamma_k) - \sqrt{V(\gamma_k)/\beta_k} Q^{-1}(\phi_k), \quad (7)$$

where ϕ_k presents the instantaneous BLER at hop k, $Q^{-1}(.)$ denotes the inverse Gaussian Q-function, $C(x) \triangleq \log_2(1+x)$ and $V(x) \triangleq (1-(1+x)^{-2})(\log_2 e)^2$ denote the Shannon capacity and the channel dispersion, respectively.

A. BLER Analysis

The instantaneous E2E BLER of the considered multi-hop multirelay in IoT networks can be expressed as [5] $\phi_{e2e} = 1 - \prod_{k=1}^{K} \big[1 - \phi_k(\gamma_k)\big]$. Thus, the average E2E BLER can be calculated as

$$\varepsilon_{e2e} = \mathbb{E}[\phi_{e2e}(\gamma_k)] \approx 1 - \prod_{k=1}^{K} \big[1 - \varepsilon_k\big], \quad (8)$$

where $\varepsilon_k = \mathbb{E}[\phi_k(\gamma_k)]$. It is intractable to directly evaluate an exact closed-form expression for ε_{e2e} in (8) since it involves a complicated Gaussian Q-function. This motivates us to apply a tight approximation approach to calculate the E2E BLER, which will be presented hereafter. From the instantaneous SNRs presented as in (4) and (6), the BLER of CRS scheme can be calculated by the following theorem.

Theorem 1. *The closed-form expression for the BLER of CRS scheme can be derived as*

$$\varepsilon_{e2e} = 1 - [1-\varepsilon_1] \prod_{h=2}^{(K+1)/2} \big[1 - \varepsilon_k\big], \quad (9)$$

where ε_1 *is expressed as*

$$\varepsilon_1 = 1 + \delta_1 \sqrt{\beta_1} \sum_{t=1}^{N_1} \sum_{q=0}^{t(N_S-1)} \binom{N_1}{t} \frac{(-1)^t \Theta(q, N_{k-1}, t)}{(\bar{\gamma}_1 \lambda_{D,1})^{-1} t^{q+1}} \times \Big[\gamma\Big(1+q, \frac{tv_1}{\bar{\gamma}_1 \lambda_{D,1}}\Big) - \gamma\Big(1+q, \frac{tu_1}{\bar{\gamma}_1 \lambda_{D,1}}\Big) \Big], \quad (10)$$

while ε_k *is expressed as*

$$\begin{aligned} \varepsilon_k =& 1 + \delta_k \sqrt{\beta_k} \sum_{p=1}^{N_{k-1}} \sum_{t=1}^{N_k} \sum_{q=0}^{t(N_{k-1}-1)} \frac{(-1)^{t+p-1} \Theta(q, N_{k-1}, t)}{\Gamma(L)^2 t^q} \\ & \times \binom{N_{k-1}}{p} \binom{N_k}{t} \sum_{n_1=0}^{I} \sum_{i_1=0}^{n_1} \Lambda(L, n_1, i_1) \\ & \times \sum_{n_2=0}^{I} \sum_{i_2=0}^{n_2} \frac{[\Lambda(2, n_2, i_2) - 2\Lambda(1, n_2, i_2)]}{2^{-i_1-i_2+L-1}} \\ & \times \left(\frac{p}{\psi_{k-1}}\right)^{\frac{i_1}{2}} \left(\frac{t}{\psi_k}\right)^{\frac{L+q+i_2-2}{2}} \Bigg[v^{\frac{L+q+i_1+i_2}{2}} \\ & \times E_{-1-L-q-i_1-i_2}\left(2\sqrt{\frac{pv}{\psi_{k-1}}} + 2\sqrt{\frac{tv}{\psi_k}}\right) \\ & - u^{\frac{L+q+i_1+i_2}{2}} E_{-1-L-q-i_1-i_2}\left(2\sqrt{\frac{pu}{\psi_{k-1}}} + 2\sqrt{\frac{tu}{\psi_k}}\right) \Bigg], \end{aligned} \quad (11)$$

for $L = q$ *in function* $\mathcal{K}_{L-q}(.)$ *shown in* (44)*, and*

$$\begin{aligned} \varepsilon_k =& 1 + \delta_k \sqrt{\beta_k} \sum_{p=1}^{N_{k-1}} \sum_{t=1}^{N_k} \sum_{q=0}^{t(N_{k-1}-1)} \frac{(-1)^{t+p-1}}{\Gamma(L)^2 t^q} \\ & \times \Theta(q, N_{k-1}, t) \binom{N_{k-1}}{p} \binom{N_k}{t} \sum_{n_1=0}^{I} \sum_{i_1=0}^{n_1} \Lambda(L, n_1, i_1) \\ & \times \sum_{n_2=0}^{I} \sum_{i_2=0}^{n_2} \frac{\Lambda(|L-q|, n_2, i_2)}{2^{-2+2L+q+J}} \left(\frac{t}{\psi_k}\right)^{\frac{i_2+2\min(L,q)}{2}} \\ & \times \left(\frac{p}{\psi_{k-1}}\right)^{\frac{i_1}{2}} \Bigg[v^{\frac{1-J}{2}} E_J\left(2\sqrt{\frac{pv}{\psi_{k-1}}} + 2\sqrt{\frac{tv}{\psi_k}}\right) \\ & - u^{\frac{1-J}{2}} E_J\left(2\sqrt{\frac{pu}{\psi_{k-1}}} + 2\sqrt{\frac{tu}{\psi_k}}\right) \Bigg], \end{aligned} \quad (12)$$

for $L \neq q$ *in function* $\mathcal{K}_{L-q}(.)$ *shown in* (44)*, where* $J \triangleq -1 - i_1 - i_2 - 2\min(L, q)$ *and*

$$\Lambda(L, n, i) = \frac{(-1)^i \sqrt{\pi} \Gamma(2L) \Gamma(1/2 + n + L) \mathbf{L}(n, i)}{2^{L-i} \Gamma(1/2 - L) \Gamma(1/2 + L + n) n!} \quad (13)$$

wherein $\mathbf{L}(n, i)$ *denotes the Lah number which is defined as* $\mathbf{L}(n, i) = \binom{n-1}{i-1} \frac{n!}{i!}$, $\mathbf{L}(0, 0) = 1$, $\mathbf{L}(n, 0) = 0$*, and* $\mathbf{L}(n, 1) = n!$, $\zeta_k = 2^{r_k} - 1$ $\delta_k = [2\pi(2^{2r_k} - 1)]^{-1/2}$, $v_k = \zeta_k - 1/(2\delta_k\sqrt{\beta_k})$*, and* $u_k = \zeta_k + 1/(2\delta_k\sqrt{\beta_k})$.

Proof: The proof of Theorem 1 is in Appendix A. ■

To the best of our knowledge, the average BLER of CRS scheme derived in Theorem 1 is new. However, this expression is rather intractable and may be difficult to provide a deeper understanding of system behavior from a non-expert in the field. This motivates us to perform an asymptotic analysis with a simpler expression in the next subsection.

B. Asymptotic BLER at High SNR

Theorem 2. *The asymptotic expression for the BLER of CRS scheme can be derived as*

$$\begin{aligned}\varepsilon_{e2e}^{\mathtt{asym}} \approx \frac{K+1}{2} - \sum_{k=2}^{(K+1)/2} \sum_{p=1}^{N_{k-1}} \sum_{t=1}^{N_k} \sum_{q=0}^{t(N_{k-1}-1)} \binom{N_{k-1}}{p}\binom{N_k}{t} \\ \times \frac{4\Lambda(L,n,i)}{(-1)^{1-p-t}\Gamma(L)^2 t^q} \left(\frac{p\zeta_{k-1}}{\psi_{k-1}}\right)^{\frac{L}{2}} \mathcal{K}_L\left(2\sqrt{\frac{p\zeta_{k-1}}{\psi_{k-1}}}\right) \\ \times \left(\frac{t\zeta_k}{\psi_k}\right)^{\frac{L+q}{2}} \mathcal{K}_{L-q}\left(2\sqrt{\frac{t\zeta_k}{\psi_k}}\right) + \sum_{t=1}^{N_1} \sum_{q=0}^{t(N_{\mathtt{S}}-1)} (-1)^t \\ \times \Theta(q, N_{k-1}, t)\binom{N_1}{t}\left(\frac{\zeta_1}{\bar{\gamma}_1\lambda_{D,1}}\right)^q \exp\left(-\frac{\zeta_1}{\bar{\gamma}_1\lambda_{D,1}}\right).\end{aligned} \tag{14}$$

where $\psi_{k-1} \triangleq \chi_{k-1}\bar{\gamma}_{k-1}k\lambda_{D,k-1}\lambda_{E,k-1}$.

Proof: The proof of Theorem 2 is in Appendix B. ■

Remark 1. *The first term in the right-hand side (RHS) of* (14)*, i.e.,* $(K+1)/2$*, shows that when a large number of hops is employed, the average BLER will be increased and the outage event may occur.*

The BLER performance of the CRS scheme in Section III is analyzed by assuming equal channel uses allocation, and therefore the blocklength utilization is not optimized. This motivates us to find the optimal channel uses allocation for the considered multi-hop network by maximizing the E2E system throughput under channel uses constraints in Section IV. In other words, Section III focuses on the BLER analysis given a fixed transmission rate while Section IV focuses on throughput optimization given fixed average BLER.

IV. Maximum of The E2E Throughput

To meet the stringent requirements of 5G networks, such as low communication latency and high decoding reliability, we formulate the E2E throughput maximization problem subject to the blocklength requirements. With the allocated channel uses β_k, the effective throughput in bit per channel uses (BPCU) at hop k is expressed as [3]

$$\begin{aligned}\mathtt{R}_k(\alpha,\beta_k,\gamma_k) &= \frac{\beta_k}{\omega}(1-\varepsilon_k)R_k \\ &= \tau_k\Big[\beta_k \log_2(1+\gamma_k) - \xi\sqrt{\beta_k\Big(1-(1+\gamma_k)^{-2}\Big)}\Big],\end{aligned} \tag{15}$$

where $\tau_k \triangleq (1-\varepsilon_k)/\omega$ and $\xi \triangleq \log_2(e)Q^{-1}(\varepsilon_k)$. Based on (15), when the average BLER at hop k, ε_k, reduces, the effective throughput increases. We rely on the concept of selection criterion in (5) of the proposed CRS scheme in Section III to formulate the throughput for two consecutive hops of the CRS scheme, denoted by $\mathtt{R}_c$, which can be expressed as

$$\begin{aligned}\mathtt{R}_c = \min\Big\{\, &\mathtt{R}_{k,i}(\alpha,\beta_k,\gamma_{k,i^*}), \mathtt{R}_{k+1,j}(\alpha,\beta_{k+1},\gamma_{k+1,j^*})\Big\}, \\ &c = 2,\ldots,(K+1)/2,\end{aligned} \tag{16}$$

where $\gamma_{k,i^*} = \max\limits_{i=1,\ldots,N_k} P_{k,i}|h_{k,n^*,i}|^2$, $\gamma_{k+1,j^*} = \max\limits_{j=1,\ldots,N_k} P_{k+1,j}\|\mathbf{h}_{k+1,j}\|^2$. The optimization problem for the E2E throughput maximization of the CRS scheme can be formulated as

$$\mathtt{P}: \max_{\alpha,\boldsymbol{\beta}} \mathtt{R}_{e2e} \triangleq \mathtt{R}_c \tag{17a}$$

$$\text{s. t.} \quad \alpha + \sum_{k=1}^{K}\beta_k \le \omega, \tag{17b}$$

$$\alpha \ge 100, \beta_k \ge 100, \forall\alpha\in\mathbb{Z}, \beta_k\in\mathbb{Z}, \tag{17c}$$

where $\boldsymbol{\beta} = [\beta_1,\ldots,\beta_K]$ and constraint (17b) guarantees the total blocklength less than the total channel uses in the system. Constraint (18e) is the minimum blocklength constraints making the result in (7) to be tight. It is clear that the objective function (17a) is non-convex subject to mixed-integer constraints, which is non-convex problem and very challenging to obtain a globally optimal solution. To tackle this problem, we introduce the soft data rate $r > 0$ such that (17) can be equivalently rewritten as

$$\mathtt{P}-\mathtt{Relaxed}: \max_{\alpha,\boldsymbol{\beta}} \mathtt{R}_{e2e} \triangleq r \tag{18a}$$

$$\text{s. t. } \mathtt{R}_{1,i}(\beta_1,\gamma_{1,i^*}) \ge r, \tag{18b}$$

$$\mathtt{R}_{k,i}(\alpha,\beta_k,\gamma_{k,i^*}) \ge r, k = 2,\ldots,K-1, \tag{18c}$$

$$\mathtt{R}_{k+1,j}(\alpha,\beta_{k+1},\gamma_{k+1,j^*}) \ge r, k = 2,\ldots,K-1, \tag{18d}$$

$$\alpha \ge 100, \beta_k \ge 100, \quad \forall\alpha\in\mathbb{R}, \beta_k\in\mathbb{R}, \tag{18e}$$

$$(17b). \tag{18f}$$

We denote by $h_1 \triangleq \max\limits_{i=1,\ldots,N_1}\|\mathbf{h}_{1,i}\|^2$, $g_k^* \triangleq \sum_{l=1}^{L}|g_{l,k,n^*}|^2$, $h_{k,i^*} \triangleq \max\limits_{i=1,\ldots,N_k}|h_{k,n^*,i}|^2$, $g_{k+1}^* \triangleq \sum_{l=1}^{L}|g_{l,k+1,i^*}|^2$, and $h_{k+1,j^*} \triangleq \max\limits_{j=1,\ldots,N_{k+1}}|h_{k+1,i^*,j}|^2$. Thus, problem (18) can be rewritten as

$$\mathtt{P}-\mathtt{Equiv}: \max_{\alpha,\boldsymbol{\beta}} r \tag{19a}$$

$$\text{s. t. } \tau_1\left[\beta_1 \ln\left(1+\bar{\gamma}_{\mathtt{S}}h_1\right) - \xi\sqrt{\beta_1\Big[1-\left(1+\bar{\gamma}_{\mathtt{S}}h_1\right)^{-2}\Big]}\right] \ge r, \tag{19b}$$

$$\tau_k\left[\beta_k \ln\left(1+\frac{\eta\bar{\gamma}\varphi_k}{\beta_k}g^*h_j^*\right) - \xi\sqrt{\beta_k\Big[1-\left(1+\frac{\eta\bar{\gamma}\varphi_k}{\beta_k}g^*h_j^*\right)^{-2}\Big]}\right] \ge r, \tag{19c}$$

$$\tau_{k+1}\left[\beta_{k+1} \ln\left(1+\frac{\eta\bar{\gamma}\varphi_{k+1}}{\beta_{k+1}}g_{k+1}^*h_{k+1,j^*}\right) - \xi\sqrt{\beta_{k+1}\Big[1-\left(1+\frac{\eta\bar{\gamma}\varphi_{k+1}}{\beta_{k+1}}g_{k+1}^*h_{k+1,j^*}\right)^{-2}\Big]}\right] \ge r, \tag{19d}$$

$$(17b), (18e), \tag{19e}$$

where $\varphi_k \triangleq (\alpha + \sum_{m=1}^{k-1} \beta_m)$. For constraint (19b), the lower bound of $\sqrt{\beta_1}$ can be expressed as

$$\sqrt{\beta_1} \leq \frac{1}{2}\left(\sqrt{\beta_1^{(\kappa)}} + \frac{\beta_1}{\sqrt{\beta_1^{(\kappa)}}}\right). \tag{20}$$

Constraint (19b) can be expressed as

$$\tau_1\left[\beta_1 \ln\left(1+\bar{\gamma}_{\mathsf{S}} h_1\right) - \frac{\xi}{2}\left(\sqrt{\beta_1^{(\kappa)}} + \beta_1\sqrt{(\beta_1^{(\kappa)})^{-1}}\right) \times \sqrt{1-\left(1+\bar{\gamma}_{\mathsf{S}} h_1\right)^{-2}}\right] \geq r. \tag{21}$$

For constraint (19c), let $A(\varphi_k, \beta_k) \triangleq \beta_k \ln\left(1 + \frac{\eta\bar{\gamma}\varphi_k}{\beta_k} g^* h_j^*\right)$ and $B(\varphi_k, \beta_k) \triangleq \sqrt{\beta_k\left[1 - \left(1 + \frac{\eta\bar{\gamma}\varphi_k}{\beta_k} g^* h_j^*\right)^{-2}\right]}$.

Lower bounding concave function approximation for $A(\beta_k, \varphi_k)$: [9, Eq. (73)]

$$\begin{aligned} A(\varphi_k,\beta_k) &\geq 2\beta_k^{(\kappa)} \ln\left(1 + \frac{\eta\bar{\gamma} g^* h_j^* \varphi_k^{(\kappa)}}{\beta_k^{(\kappa)}}\right) + \frac{\eta\bar{\gamma} g^* h_j^* \varphi_k^{(\kappa)} \beta_k^{(\kappa)}}{\eta\bar{\gamma} g^* h_j^* \varphi_k^{(\kappa)} + \beta_k^{(\kappa)}} \\ &\times \left(2 - \frac{\varphi_k^{(\kappa)}}{\varphi_k} - \frac{\beta_k}{\beta_k^{(\kappa)}}\right) - \ln\left(1 + \frac{\eta\bar{\gamma} g^* h_j^* \varphi_k^{(\kappa)}}{\beta_k^{(\kappa)}}\right) \frac{(\beta_k^{(\kappa)})^2}{\beta_k} \\ &\triangleq f^{(\kappa)}(\beta_k, \varphi_k). \end{aligned} \tag{22}$$

Upper bounding convex function approximation for $B(\varphi_k, \beta_k)$: Let $\Xi_k(\varphi_k, \beta_k) = \frac{\eta\bar{\gamma}\varphi_k}{\beta_k} g^* h_j^*$, we have

$$\begin{aligned} B(\varphi_k, \beta_k) &\leq a_k^{(\kappa)} - \frac{a_k^{(\kappa)}}{(1 + \Xi(\varphi_k, \beta_k))^2} + b_k^{(\kappa)} \beta_k \\ &= a_k^{(\kappa)} - a_k^{(\kappa)} \frac{\beta_k^2}{(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)^2} + b_k^{(\kappa)} \beta_k, \end{aligned} \tag{23}$$

where

$$\begin{aligned} a_k^{(\kappa)} &= \frac{1}{2}\sqrt{\frac{\beta_k^{(\kappa)}}{1 - 1/(1 + \Xi_k(\varphi_k^{(\kappa)}, \beta_k^{(\kappa)}))^2}}, b_k^{(\kappa)} \\ &= \frac{1}{2}\sqrt{\frac{1 - 1/(1 + \Xi_k(\varphi_k^{(\kappa)}, \beta_k^{(\kappa)}))^2}{\beta_k^{(\kappa)}}}. \end{aligned} \tag{24}$$

The RHS of (23) is not a convex function; thus, the concave lower bound of $\beta_k^2/(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)^2$ can be expressed as

$$\begin{aligned} \frac{\beta_k^2}{(\beta_k + \eta\varphi_k g^* h_j^*)^2} &= \frac{\beta_k^2}{(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)} \frac{1}{(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)} \\ &\geq \frac{\beta_k^2}{\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*}\left(\frac{2}{\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*} \right. \\ &\left. - \frac{\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*}{(\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*)^2}\right) \\ &= \frac{2}{\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*} \frac{\beta_k^2}{\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*} \\ &- \frac{\beta_k^2}{(\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*)^2}. \end{aligned} \tag{25}$$

The first term in the RHS of (25) can be lower bounded as

$$\begin{aligned} \frac{\beta_k^2}{\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*} &\geq \frac{2\beta_k^{(\kappa)}\beta_k}{\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*} \\ &- \frac{(\beta_k^{(\kappa)})^2(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)}{(\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*)^2}. \end{aligned} \tag{26}$$

By plugging (26) into (25), we have the following approximation:

$$\begin{aligned} \frac{\beta_k^2}{(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)^2} &\geq \frac{2}{\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*}\left(\frac{2\beta_k^{(\kappa)}\beta_k}{\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*} \right. \\ &\left. - \frac{(\beta_k^{(\kappa)})^2(\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^*)}{(\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*)^2}\right) - \frac{\beta_k^2}{(\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*)^2} \\ &\triangleq C^{(\kappa)}(\varphi_k, \beta_k), \end{aligned} \tag{27}$$

over the trust region constrained by

$$\beta_k + \eta\bar{\gamma}\varphi_k g^* h_j^* \leq 2(\beta_k^{(\kappa)} + \eta\bar{\gamma}\varphi_k^{(\kappa)} g^* h_j^*). \tag{28}$$

By plugging (27) into (23), we obtain the concave upper bound as

$$\begin{aligned} B(\varphi_k, \beta_k) &\leq a_k^{(\kappa)} - a_k^{(\kappa)} C^{(\kappa)}(\varphi_k, \beta_k) + b_k^{(\kappa)} \beta_k \\ &\triangleq B^{(\kappa)}(\varphi_k, \beta_k). \end{aligned} \tag{29}$$

Finally, we have

$$f^{(\kappa)}(\beta_k, \varphi_k) - \xi B^{(\kappa)}(\varphi_k, \beta_k) \geq r. \tag{30}$$

Similar to the transformation of (19c), the concave upper bound of constraint (19e) can be approximated as

$$f^{(\kappa)}(\beta_{k+1}, \varphi_{k+1}) - \xi B^{(\kappa)}(\varphi_{k+1}, \beta_{k+1}) \geq r. \tag{31}$$

From the above discussions, the problem (19) can be approximated by the following convex program at iterative $\kappa + 1$:

$$\mathtt{P-Convex}: \max_{\boldsymbol{\beta}, r, \boldsymbol{\varphi}} \quad r \tag{32a}$$

$$\text{s. t. } (17\text{b}), (18\text{e}), (21), (28), (30), (31). \tag{32b}$$

In Algorithm 1, we summarize the proposed iterative algorithm to solve the $\mathtt{P-Convex}$ problem (32). The initial points can be generated randomly, provided that $\alpha > 100$ and $\beta_m > 100$ [7], such that Algorithm 1 always initializes with a feasible solution of (32). Relying on the convergence of inner approximation-based algorithms in [28], Algorithm 1 produces the sequence $\varphi_k^{(\kappa)}, \beta_k^{(\kappa)}$ of improved solutions of (32) (hence, (17)), which converges to a Karush–Kuhn–Tucker point of the relaxed problem. Thus, the obtained solutions of Algorithm 1 for solving the relaxed problem (31) converge at least to local optimum, satisfying the Karush–Kuhn–Tucker conditions.

V. Deep Learning-Aided Real-Time Optimization

In multi-hop networks, data generated by IoT devices are constantly increased with different formats and complex correlations. Moreover, the model-based approach in modern wireless systems is facing grand challenges in resource optimization and hard-to-model problems due to heterogeneous

Algorithm 1 Proposed Iterative Algorithm for Solving Problem (17)

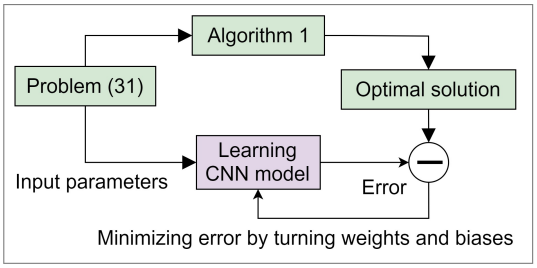
1: **Initialization:** Set $\kappa \coloneqq 0$, $(\varphi_k^\star, \beta_k^\star) \coloneqq 0$, and generate an initial feasible point $(\varphi_k^{(0)}, \beta_k^{(0)})$ for (17) by setting all the constrains in (32) to be equality.
2: **repeat**
3: Solve the convex program (32) to obtain the optimal solution $(\varphi_k^\star, \beta_k^\star)$;
4: Update $(\varphi_k^{(\kappa+1)}, \beta_k^{(\kappa+1)}) \coloneqq (\varphi_k^\star, \beta_k^\star)$;
5: Set $\kappa \coloneqq \kappa + 1$;
6: **until** Convergence
7: Update $(\varphi_k^\star, \beta_k^\star) \coloneqq (\varphi_k^{(\kappa)}, \beta_k^{(\kappa)})$;
8: Calculate $\mathtt{R}_{e2e}$ in (17a) based on $(\varphi_k^\star, \beta_k^\star)$;
9: **Output:** r and the optimal solution $(\varphi_k^\star, \beta_k^\star)$.

networks with different access technologies. To tackle these problems, we design a novel framework using CNN as an underlying technology to discover the optimal signaling inference from very large and diverse data sets that ensures reliable packet transmissions.

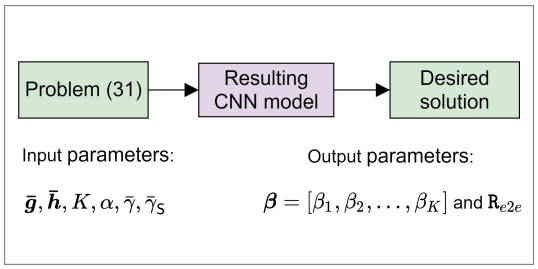
A. Optimization Framework Design and CNN Architecture

Relying on the problem statement (32), we design a deep CNN framework including training and predicting stages to obtain the optimal channel uses allocation with low-complexity and real-time manner, as shown in Fig. 2. The information for training can be typically exchanged over the network via physical broadcast channel as the basic system configurations in long-term evolution systems or 5G networks [31], [32]. Moreover, encryption and authentication protocols in application layer are used to transmit and receive information during data collection in 5G networks and IoT systems [33]; therefore, the privacy and security are guaranteed. Once the training process is done at source node, the trained CNN model consisting of weights and biases can be used to predict the channel uses allocation and throughput whenever any new information is available at the input, which can provide real-time and highly accurate predictions. In conventional systems without deploying deep learning, the source has to run optimization algorithm to obtain the optimal channel uses allocation. In our design, the source only needs the new network settings and can predict in real-time the optimal channel uses when the CNN is well trained, thus significantly reducing complexity.

In Fig. 2(a), CNN learns by off-line the mapping function between the input parameters of the problem (32) and optimal channel uses solution obtained through solving the problem (32) by Algorithm 1. During training, CNN keeps improving the deep model by updating weights and biases to minimize the error. The resulting deep model is used as a mapping function to predict online the optimal solution via low processing time inference process whenever the input parameters are available at the CNN, as shown in Fig. 2(b). In this way, the computational complexity is transferred to the offline training while the complexity during online estimation is significantly reduced. The input variables of CNN include the



(a) Training



(b) Predicting

Fig. 2: The proposed CNN-aided optimization framework for the throughput problem.

channel gains of WET, $\bar{\boldsymbol{g}} \triangleq [|g_{1,i}|^2, \ldots, |g_{K-1,i}|^2]$, and WIT, $\bar{\boldsymbol{h}} \triangleq [|h_{1,i}|^2, \ldots, |h_{K,i}|^2]$, the number of PBs with $L \in [1, 12]$, the number of antennas at S with $N_{\mathsf{S}} \in [1, 5]$, the number of relays in each cluster, $N_k \in [1, 12]$, the transmit power at the PB with $\bar{\gamma} \in [0, 50]$, and the transmit power at the S with $\bar{\gamma}_{\mathsf{S}} \in [15, 25]$. Based on these settings, the maximum throughput and optimal channel uses, which are obtained by solving problem (32) via Algorithm 1 and then arranged by a vector $\boldsymbol{\tau} \triangleq [\boldsymbol{\beta}, \mathcal{R}_{e2e}]$, serve as output variables of CNN. We consider the network with $K = 3$ hops; thus, the input layer has the size of 1×10 to process input data. The whole dataset having $140,000$ samples is divided randomly into the training and test sets with the ratio of 80-20.

In the network architecture, we design four processing blocks, denoted by block-A, block-B, block-C, and block-D, as shown in Fig. 3 to reveal the underlying relevant features as the local correlations between system variables. Each block is specified by multiple convolutional (conv) layers in the parallel connection with different one-dimensional asymmetric kernels of sizes, such as 1×9, 1×7, 1×5, 1×3, and 1×1, to capture more diversified features, where their outputs are then gathered along in the depth dimension via a depth-wise concatenation (concat) layer. Adding more kernels into our CNN design may slow down the training process. However, the source node can be seen as a base station with powerful energy and high processing capability, allowing the ability to efficiently train the deep CNN model. Moreover, recent achievements of graphics processing units and high-performance chips with advanced parallel computing further foster the training capabilities of deep CNN models to make inferences within milliseconds [34]. It is worth noting that each conv layer is followed by a batch normalization (bn) layer and a rectified linear unit (ReLU) activation layer to accelerate training convergence and nonlinearity [35]. For example, the

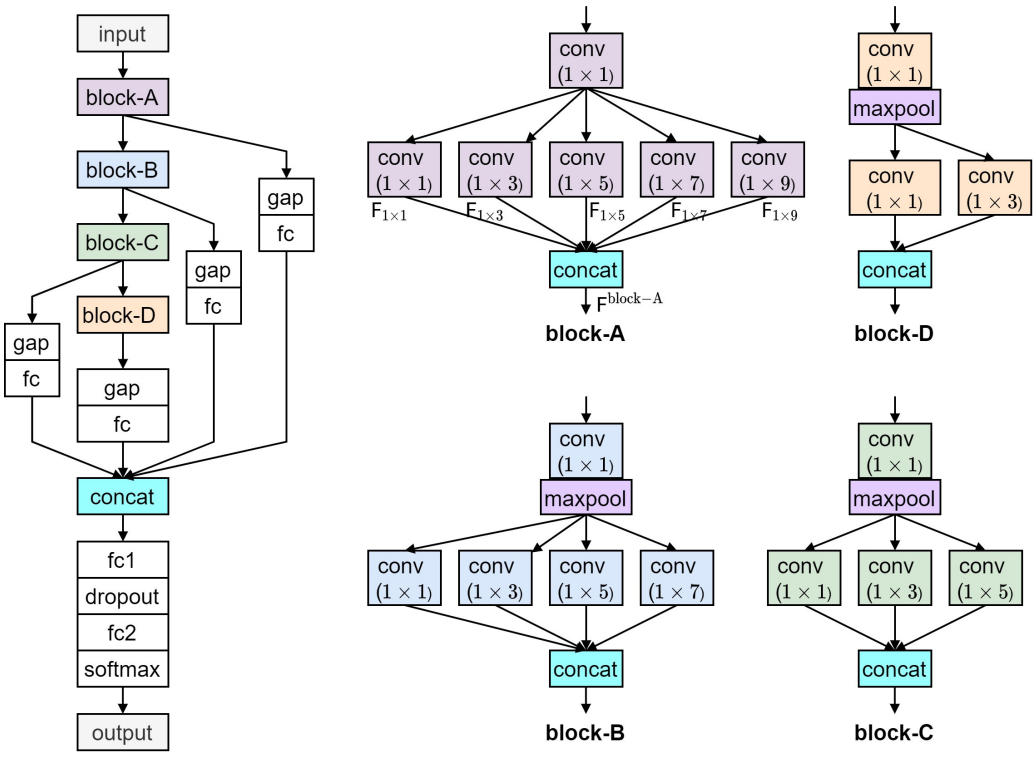


Fig. 3: Deep CNN design with the main architecture and the detailed structure of blocks.

outcome of block-A can be expressed as follows:

$$\mathsf{F}^{\mathrm{block-A}} = \mathcal{C}_{depth}\left(\mathsf{F}_{1\times1}, \mathsf{F}_{1\times3}, \mathsf{F}_{1\times5}, \mathsf{F}_{1\times7}, \mathsf{F}_{1\times9}\right), \quad (33)$$

where $\mathcal{C}_{depth}$ denotes the depth-wise concatenation operation of the (concat) layer. The elements $\mathsf{F}_{1\times1}$, $\mathsf{F}_{1\times3}$, $\mathsf{F}_{1\times5}$, $\mathsf{F}_{1\times7}$, and $\mathsf{F}_{1\times9}$ are the features computed by a sequential transformation conv−bn−ReLU with different kernels of sizes correspondingly. To save the computational cost of subsequent conv layers when going deeper, a max-pooling (maxpool) layer with the stride of $(1, 2)$ and the pool size of 1×2 is deployed in block-B, block-C, and block-D, which in turn halves the spatial size of feature maps.

Furthermore, blocks are connected to each other in cascades and an effective connection structure is employed to collect meaningful features summarized at multi-scale representations. In particular, the output of each block is traveled to a global average pooling (gap) layer and a fully connected (fc) layer. Afterward, a general concat layer is used as a selective feature collection to combine the outputs from these fc layers. Finally, two fc layers, denoted by fc1 and fc2, are used to complete the CNN, where the number of hidden nodes in fc2 equals to the number of variables for estimating output values (here the throughput and channel uses). Regarding the configuration of layers, we set 32, 64, 128, and 256 kernels in each conv layer of block-A, block-B, block-C, and block-D, respectively; and 64 neurons in fc layers. The zero padding is automatically added when computing convolution to keep the spatial size of outputs unchangeable. The loss function for the regression problem represents the error between predicted and expected values over training samples, which can be generally expressed as [20]

$$\mathcal{L} = \mathbb{E}\{(\boldsymbol{\tau} - \tilde{\boldsymbol{\tau}})^2\}, \quad (34)$$

where $\boldsymbol{\tau}$ and $\tilde{\boldsymbol{\tau}} \triangleq [\tilde{\boldsymbol{\beta}}, \tilde{\mathcal{R}}_{e2e}]$ denote the expected and predicted outputs, respectively. By adopting the adaptive moment estimation optimization algorithm, the trainable parameters, including weights of the connections and biases, are repeatedly adjusted during the backpropagation process aiming to minimize the loss function of the entire training set.

B. DL-Aided Real-Time Optimization

After finishing the offline training, the resulting CNN model comprising of weights and biases can be represented in a compact mapping function, $\mathcal{F}(.)$, as shown in (35). In general, a well-trained CNN can deliver high accuracy and real-time estimation. We use the trained CNN model to estimate the throughput and optimal channel uses values whenever any new information is available at the input of network. Particularly, we input serially each sample, which is arranged as a vector $\mathbf{x} \triangleq [\bar{\boldsymbol{g}}, \bar{\boldsymbol{h}}, \alpha, \bar{\gamma}, \bar{\gamma}_{\mathsf{S}}]$, the resulting CNN will output the predicted optimal channel uses and throughput, which can be expressed as

$$[\boldsymbol{\beta}, \mathcal{R}_{e2e}] = \mathcal{F}(\mathbf{x}). \quad (35)$$

By low-latency inference process in (35), the resulting CNN model can predict the optimal throughput and channel uses in a short time. Since the learning capability of CNN can be enhanced by going a deeper (more hidden layers) or wider (more kernels/neurons) network, its architecture can be cleverly fine-tuned to obtain the smallest error during

training process. If the predicted throughput and channel uses do not close to the expected ones, our designed network can be ameliorated with new appropriate configurations until the lowest error in (34) is reached. Channel statistics or dynamics greatly affects the performance optimization of the proposed CNN. Particularly, different fading distributions lead to different average channel gains, and resulting in different prediction results of throughput and channel uses allocation. In high dynamic wireless environments, channel conditions with line-of-sight (LoS) and non-LoS links vary rapidly, which also impact on the prediction results of the CNN. If the source cannot acquire the current system configurations or the destination cannot obtain channel conditions timely due to the non-LoS links or propagation delay, the CNN may get inputs with out-of-date information, and it infers the inappropriate channel uses allocation or wrong throughput evaluation even when the inference process is occurred within a coherent interval.

VI. Numerical Results and Discussions

In this section, we evaluate the performance of the proposed relay selection scheme and the developed CNN-based optimization through illustrative numerical examples. Monte-Carlo simulations results are provided to verify the analytical expressions. We consider a bi-dimensional plane for the location of all nodes, where the position of S, $\mathsf{R}_{k,i}$, D, and PB are $(0,0)$, $(k/K,0)$, $(50,0)$, and $(7.0,15)$, respectively. Unless otherwise stated, based on [17], [36], we set in simulations the pathloss exponent $\sigma_{\rm PL}=2.7$, the energy conversion efficiency $\eta=0.8$, the target data rate $r_k=m/\beta_k$ BPCU, the total channel uses $\omega=1000$, the channel uses for WET $\alpha=150$, the number of antennas at the source $N_{\mathsf{S}}=3$, the number of PBs $L=3$, the transmit power at each PB and S $P=P_{\mathsf{S}}=20$ dB, the number of hops $K=5$, the number of relays in each cluster $N_k=3$, the reference distance $d_0=1$ m, the pathloss at d_0 $\mathcal{L}=-30$ dB, the noise variance $\sigma^2=1$, and the average BLER $\varepsilon_k=10^{-5}$ for throughput maximization.

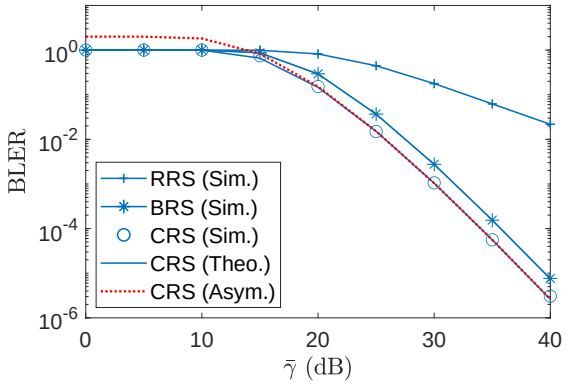


Fig. 4: The BLER of all schemes versus the average SNR.

We first present the BLER versus average SNR $\bar{\gamma}$, as shown in Fig. 4, where the outstanding performance of the proposed CRS scheme over BRS and RRS ones can be observed. Particularly, the BRS scheme has about 2.5 dB improvement in average SNR compared to BRS scheme at BLER lower than 9×10^{-2}. Multiple relays involving in relay selection process with distributed multiuser beamforming design in (5) help CRS scheme to provide high diversity transmission while the partial relay selection employing in BRS scheme results in a lower BLER performance. The single-relay selection criterion used in RRS scheme has the worst BLER performance. Moreover, the BLER reduces when $\bar{\gamma}$ increases due to the fact that relay nodes have more opportunities to harvest sufficient energy for packet transmission at high average SNRs. It is shown that the theoretical analysis results match with the simulation ones, validating the developed analysis in Theorem 1. Moreover, the asymptotic results are in tight upper bound with the theoretical analysis ones, confirming results in Theorem 2. It is noted that the CRS scheme yields higher complexity than the BRS and RRS ones when considering IoT applications. However, the CRS scheme outperforms the BRS and RRS ones in terms of reliability, latency, average BLER, and throughput, which are suitable for IoT applications to meet stringent requirements on reliability and latency.

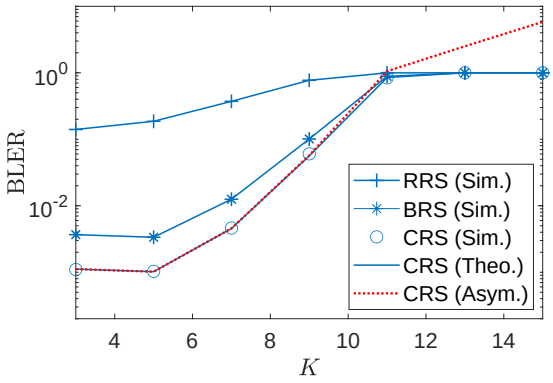


Fig. 5: The BLER of all schemes versus the number of hops K with $\bar{\gamma}=30$ dB.

In Fig. 5, we show the effect of the number of hops K on the E2E BLER of three schemes. As can be observed, the average BLER of all schemes increases when K is large because the error probability of decoding and sending short packets to users increases via multiple hops transmission, which is also aligned with the discussion in Remark 1. There exists an optimal number of hops that minimizes the average BLER. The CRS and BRS with multiple relays selection achieve the same optimal value of K as 5, while it is 3 for RRS scheme with single relay selection. The CRS scheme achieves the highest BLER performance when K is less than 11 hops and the BLER of all schemes converges at its highest error when K is lager than 11. Thus, we focus on evaluating the CRS scheme in the next figures.

In Fig. 6, we show the average BLER versus $\bar{\gamma}$ with different message sizes and transmit powers at source node. As can be observed, the average BLER is strongly influenced by the size of messages and the transmit power at source node. Particularly, with the transmit power $P_{\mathsf{S}}=20$ dB, the CRS scheme with 64 bytes has about 10 dB improvement in average

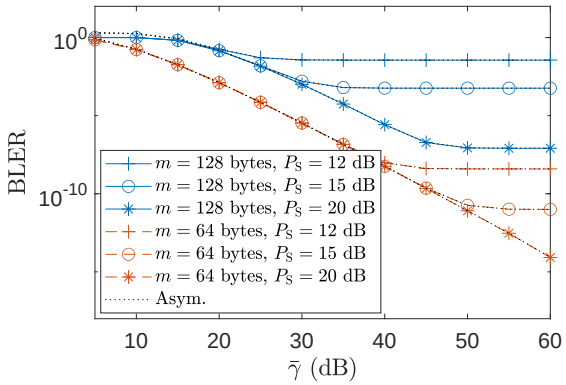


Fig. 6: Effects of message size on the BLER of CRS scheme with $\alpha = 150$ channel uses.

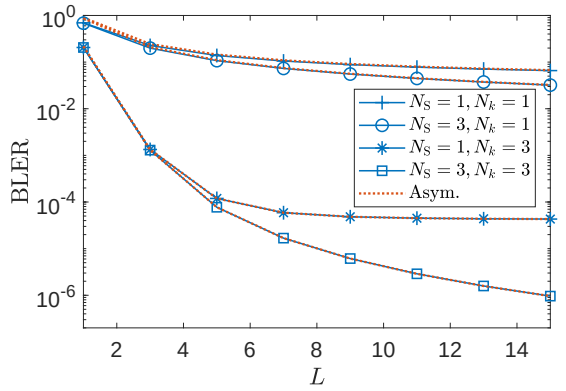


Fig. 7: Effects of N_S and N_k on the BLER of CRS scheme with $\omega = 1000$ channel uses.

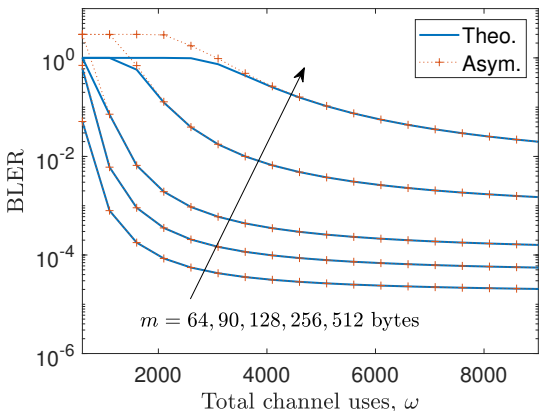


Fig. 8: Effects of message size and total channel uses on the BLER of CRS scheme with $\alpha = 150$ channel uses.

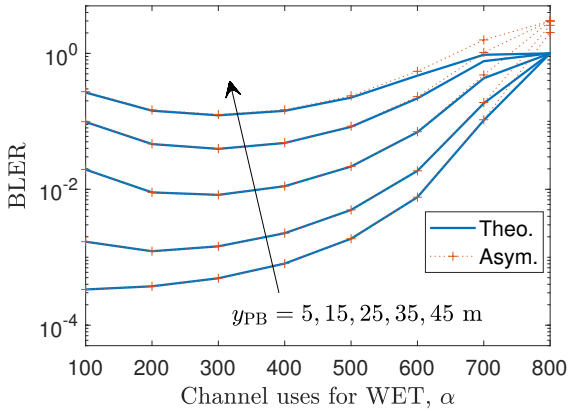


Fig. 9: Effects of PB placement on the BLER of CRS scheme with $\omega = 1000$ channel uses.

SNR compared to its counterpart with 128 bytes. The reason is that the message size is increased, the data transmission rate is reduced as in 7, hence increasing the BLER of CRS scheme. The small transmit power at source also decreases the BLER performance and makes it to converge at the error floor even though the transmit power at PB keeps increasing. This shows the importance of the first hop transmission when the source uses its own energy for data transmission.

As shown in Fig. 7, the number of antennas at PB, at source, and relays in each cluster has a considerable impact on the BLER performance of CRS scheme. When these parameters increase, the average BLER is improved. For example, the setting of $N_S = 1$ and $N_k = 1$ provides the lowest BLER performance while that of $N_S = 3$ and $N_k = 3$ is the best one. This shows the advantage of using MRT at source and relay selection at each cluster when deploying multi-antenna multi-relay system model. Moreover, when the number of PBs increases, relays have more opportunities to harvest sufficient energies from PBs for their packet transmissions. The asymptotic results are also in tight upper bound with the theoretical analysis ones.

In Fig. 8, we show the effect of message sizes and total channel uses on the average BLER of CRS scheme. When the message size increases the transmission rate reduce, leading to the decrease of BLER performance. When the total channel uses increase, the BLER is improved since the system has more channels for packet transmission. However, with small size of message, the increase of channel uses does not bring much benefit and the system only requires a small number of channel uses for its operation since the BLER is slightly reduced as the channel uses increase. For example, the BLER of system with 64 bytes is nearly equal to that of its counterpart with 8500 bytes.

In Fig. 9, the average BLER of CRS scheme is shown as a function of α with different settings of PB placement. When PBs locate nearby the multi-hop network, the BLER performance is enhanced because the pathloss effect is reduced with the small distance. Moreover, when α increases, the BLER performance of all settings decreases since most channel uses are now dedicated to WET and thus reduces channel uses for WIT, which makes the outage event occur.

The reliability and latency of three considered schemes are shown in Fig. 10, where they can be defined as [11], [20]

$$\text{Reliability} = (1 - \varepsilon_{e2e}) \times 100\%, \tag{36}$$

$$\text{Latency} = (\omega - \alpha)T/(1 - \varepsilon_{e2e}), \tag{37}$$

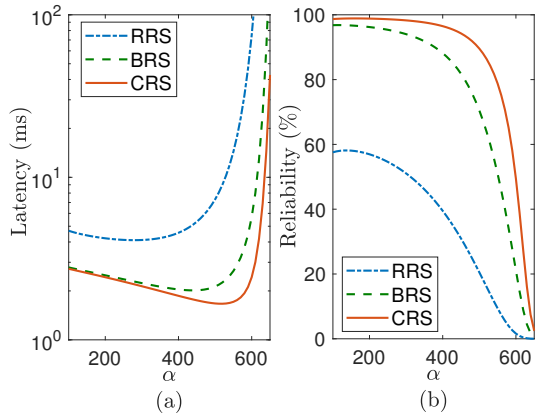


Fig. 10: The latency and reliability of three schemes versus α with $m = 90$ bytes.

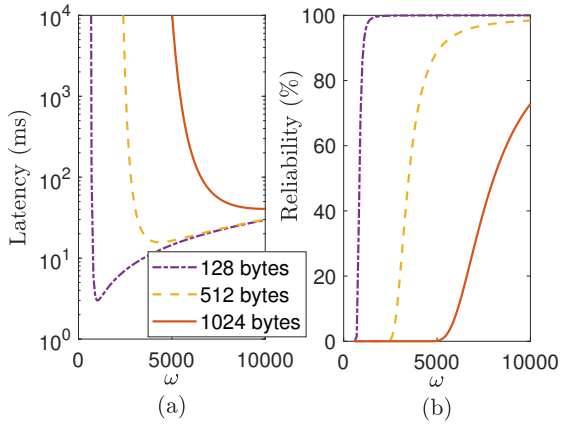


Fig. 11: The latency and reliability of CRS scheme versus ω.

where $T = 3\mu s$ [11]. As can be observed, the CRS scheme achieves lower latency and higher reliability than that of BRS and RRS ones because it has lower BLER than the BRS and RS schemes. As shown in Fig. 10(a), there exists an optimal channel uses for WET, α, that minimizes the latency of all schemes. When α increases, the latency of all schemes first gradually reduces and reaches its minimum value before rapidly raising. The reason is that small values of α make the relays to not harvest sufficient energies while large values of α narrow the channel uses for WIT. These scenarios increase the BLER, leading to the high latency as in (37). Large values of α also reduce the reliability of all schemes, as shown in Fig. 10(b).

In Fig. 11, the message with 128 bytes has higher reliability and lower latency than long ones (i.e., with 512 or 1024 bytes). To achieve high reliability (over 99.99%), the 512-byte message can be framed into packets by using 6000 channel uses, as shown in Fig. 11(a), but its latency is about 18s, as shown in Fig. 11(b), which exceeds the maximum latency budget for URLLC applications [37]. The 1024-byte message is transmitted with the latency of 20s and reliability of 90% at 6000 channel uses, which does not meet the stringent requirements of latency and reliability. Thus, the long-messages are not able to support URLLCs in IoT networks and low-latency transmission in factory automation.

Fig. 12 shows the convergence behavior of Algorithm 1 with different schemes, where the results are obtained from a random channel realization with $\bar{\gamma}$ = 30 dB. For each realization, we successively solve problem (32) until it converges. As can be observed, Algorithm 1 quickly converges to its highest values after less than 10 iterations. This is attributed to the fact that an improved solution over the whole feasible set can be found in each iteration, which demonstrates the effectiveness of the proposed iterative algorithm.

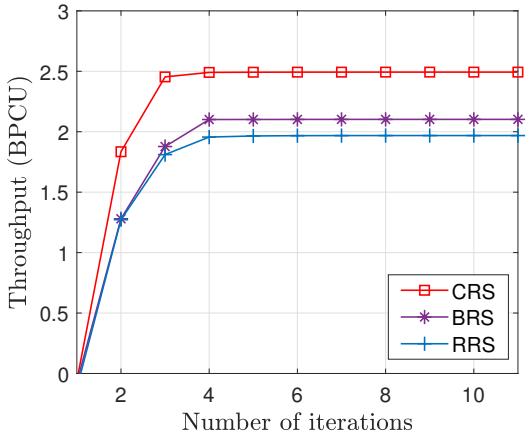


Fig. 12: The convergence of Algorithm 1 for different schemes.

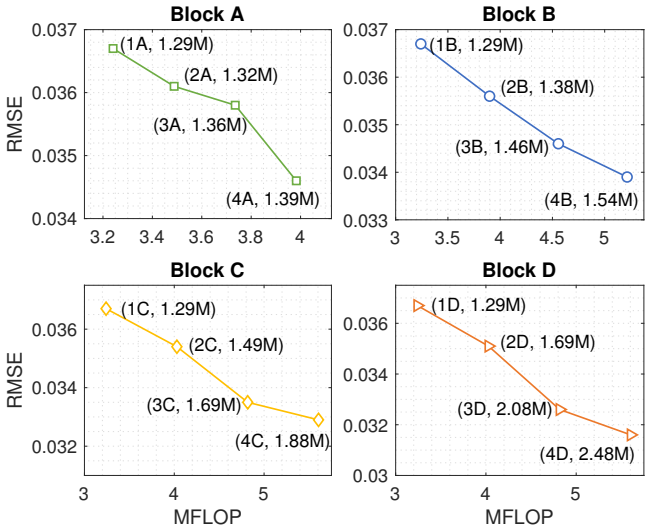


Fig. 13: The RMSE of different number of processing blocks.

Based on CNN design in Fig. 3, the effects of different number of processing blocks on the RMSE are shown in Fig. 13. The label "3A, 1.36M" in sub-figure **Block A** indicates the proposed CNN with 3 blocks A, 1 block B, 1 block C, and 1 block D having 1.36M trainable parameters (sum of all weights and biases). Other labels are followed in the same manner. These results are obtained by averaging RMSE on the test set with multiple independent running times. The RMSE performance improves as the number of blocks increases.

To achieve the best trade-off between estimated accuracy and complexity, the CNN architecture should be designed considering the prediction accuracy as well as the number of trainable parameters (network size) and the number of floating point operations (FLOPs) [38]. For example, considering the CNN architecture with blocks B, C, and D, when the number of blocks A increases (the green line), the proposed CNN with 4 blocks A has 1.39M parameters and 3.98 mega FLOPs (MFLOPs), which yields the RMSE of around 0.0346. To achieve the same RMSE result, the proposed CNN should be equipped with 3 blocks B (the blue line) with the increase of complexity requiring 1.46M parameters and 4.55 MFLOPs. Thus, in this case, the proposed CNN with 4 blocks A achieves a better trade-off between accuracy and complexity than its counterpart with 3 blocks B.

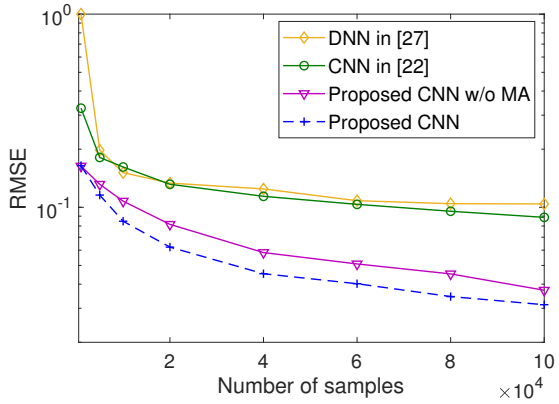


Fig. 14: RMSE versus the number of samples for training.

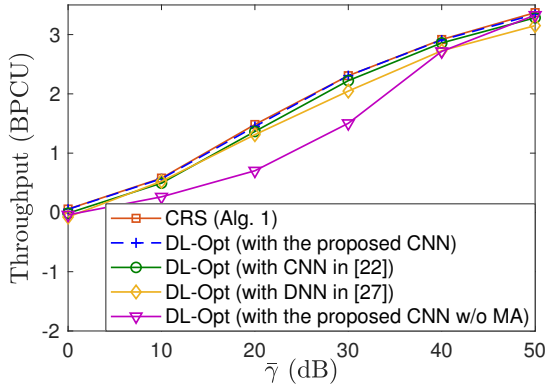


Fig. 15: Throughput of three considered schemes using Algorithm 1 with $K = 3$.

To evaluate the estimation performance of the proposed CNN for predicting the optimal channel uses and throughput, the RMSE is used to indicate the difference between the expected values and predicted ones over the entire test set. The RMSE can be calculated by using the samples in the test set as follows: $\text{RMSE} = \sqrt{\mathbb{E}\{(\mathbf{y} - \tilde{\mathbf{y}})^2\}}$, where $\mathbf{y}$ and $\tilde{\mathbf{y}}$ are the expected and predicted values in the test set, respectively. We compare the RMSE of the proposed CNN-based model with that of other DL-based approaches, such as DNN [27], the conventional CNN [22], and the proposed CNN without MA connection. These DL-based regression models are evaluated on the same dataset with the proposed CNN one. In Fig. 14, we show the RMSE versus the number of samples with different regression models. As can be observed, the DNN and conventional CNN model have the largest RMSE, while the proposed CNN model gets the smallest RMSE, showing the best performer. Since the DNN and conventional CNN models are infeasible to estimate a moderate-to-high-dimensional dataset, they provide the lowest performance. The proposed CNN model has the ability to map the original dataset into a higher dimensional space through the versatile design of block-A, block-B, block-C, and block-D and attention connection. This reveals the beneficial feature of deep architecture design for learning models in effectively dealing with big communication datasets. On the contrary, the RMSE of the DNN, the conventional CNN, and the proposed CNN without MA connection models are slightly changed when the number of samples is increased, showing that they have poor predictive ability in large datasets.

The effects of different DL method in the proposed DL framework on the ability to achieve the throughput and optimal channel uses are shown in Fig. 15. It is observed that the proposed DL-Opt with the proposed CNN provides the best fit curve of the throughput to the CRS scheme while the DL-Opt with the proposed CNN without MA connection has a high error prediction which cannot attain the optimal solutions and fails to estimate the throughput.

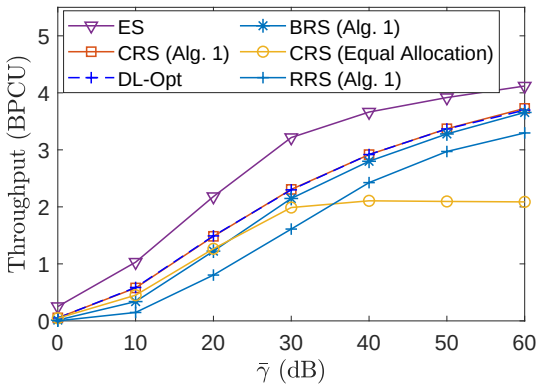


Fig. 16: Throughput versus $\bar{\gamma}$ with $K = 3$.

Fig. 16 plots the throughput versus the average SNR, $\bar{\gamma}$, with different schemes. Simulation results of the exhaustive search (ES) algorithm are also provided, where the ES can find the best optimal channel uses allocation among $(\omega - \alpha)^K$ possibilities. The throughput increases with the increase of $\bar{\gamma}$ since the high transmit power at PBs can enhance the harvested energy at each relay. As expected, the ES algorithm provides the highest throughput among other existing ones since it persistently finds the best channel uses allocation among all possibilities. However, when the number of hops or channel uses increases, the search method may become computationally prohibitive even for medium-sized networks. Therefore, it is utilized typically for benchmarking purposes. On the

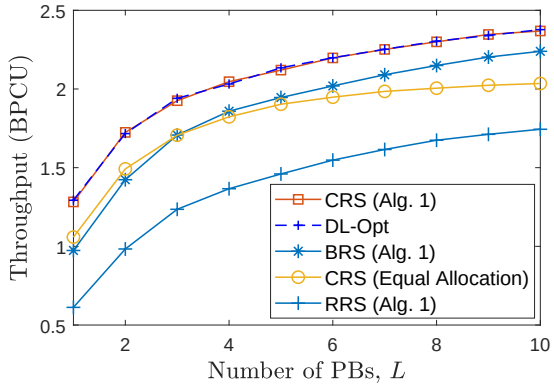


Fig. 17: Effects of L on the throughput.

other hand, the per-iterative complexity of Algorithm 1 is $\mathcal{O}(n_C^2 m_C^{2.5} + m_C^{3.5})$, where $n_C = 2K+1$ and $m_C = K+2$ are the numbers of scalar optimization variables and linear/SOC constraints, respectively. Given this complexity, Algorithm 1 is capable of finding a feasible solution in the system with a large number of hops and channel uses. The performance gap between Algorithm 1 and exhaustive search is large due to the relaxation of channel uses variables, i.e., the integer numbers in constraint (17c) are relaxed to be continuous ones in (18e). Then, the channel uses are rounded after obtaining the optimal solution to recover exact integer numbers which are feasible to the original problem. Moreover, this gap is reduced and narrower when the average SNR increases. Importantly, the CRS scheme based Algorithm 1 clearly outperforms the BRS and RRS based Algorithm 1 ones, which by their turn outperform the CRS based equal channel uses allocation. At high SNR regions, the CRS based equal allocation converges to its maximum throughput due to the fixed channel uses strategy. Finally, the DL-Opt achieves a similar throughput to that of the CRS one based Algorithm 1, confirming the accuracy of our efficient CNN design.

In Fig. 17, we show the effect of the number of PBs on the throughput with different schemes. As expected, the throughput improves with the increase of the number of PBs because the relays have more opportunities to harvest sufficient energies for their operation. Moreover, the CRS scheme based Algorithm 1 provides the highest performance while the RS based Algorithm 1 is the worst performer because CRS employs the optimal channel uses allocation strategy. Compared to the BRS scheme based Algorithm 1, the CRS based equal allocation has lower throughput due to the inefficient of deploying the equal channel uses allocation. The DL-Opt achieves the same performance as the CRS based Algorithm 1 one, which shows the efficiency of the CNN design.

Finally, we evaluate the execution time of the optimal throughput prediction in Table I. Each sample of CRS based Algorithm 1 is obtained by using Monte-Carlo simulations with an average of 10^3 independent channel realizations. This experiment is measured on a hardware platform of Ryzen 7 processor, 32GB RAM, and a single GeForce GTX 1080Ti graphics processing unit. The resulting CNN model is used to estimate the optimal channel uses and throughput of the CRS scheme on a per case basis. When the network increases in terms of the number of devices (N_k, L) and number of antennas (N_{S}), the DL-Opt still guarantees an execution time of less than 3 seconds, while the execution time of CRS based Algorithm 1 takes a long period to attain the optimal solution for over 2×10^3 seconds in all scenarios. These results reveal the excellent ability of the proposed CNN framework in dealing with large-scale network settings.

TABLE I: Execution time for obtaining optimal throughput between CRS based Algorithm 1 and the proposed DL-Opt, with $\alpha = 220$, $P_{\mathsf{S}} = 20$ dB, and $P = 20$ dB.

Scenarios $\{N_{\mathsf{S}}, N_k, L\}$	CRS-based Algorithm 1	Proposed DL-Opt	RMSE
$\{3, 4, 5\}$	2321.8s	2.665s	0.0855
$\{4, 5, 6\}$	2326.3s	2.774s	0.0955
$\{5, 6, 7\}$	2338.5s	2.625s	0.0922

VII. Conclusions

In this paper, we proposed the CRS scheme in multi-hop networks with WET to achieve high BLER performance, high-reliable, and low-latency communications under short-packet regime. The closed-form expression for the E2E average BLER of the CRS scheme was derived from which the respective asymptotic analysis was also carried out. To achieve an optimal channel uses allocation, a max-min throughput optimization was proposed, where the formulated problem was non-convex under the class of mixed-integer constraints, that was very challenging to solve optimally. By relaxing channel uses variables to be continuous and then transforming the relaxed problem into an equivalent non-convex one, but with a more tractable form. A low-complexity iterative algorithm was developed to generate a sequence of improved objective, which converged to a local optimum by using inner approximation method. To achieve real-time settings, we designed an efficient deep CNN for the relaxed optimization problem which achieved the sub-optimal solution via a real-time inference process. Numerical results presented effectiveness of the proposed CRS scheme in improving the reliability and latency compared to the BRS and RRS ones. Furthermore, it was also showed the lowest RMSE of the designed CNN and throughput improvement of the CRS scheme over state-of-the-art ones.

Appendix A
Proof of Theorem 1

For the first-hop transmission, from (7), the instantaneous BLER at hop k can be approximated as [8], [11] $\phi_k(\gamma_k) \approx Q\Big(\frac{C(\gamma_k)-R_k}{\sqrt{V(\gamma_k)/\beta_k}}\Big)$, where the respective average BLER can be calculated as

$$\varepsilon_k = \mathbb{E}[\phi_k(\gamma_k)] = \int_0^\infty \left[Q\left(\frac{C(\gamma_k) - R_k}{\sqrt{V(\gamma_k)/\beta_k}}\right)\right] f_{\gamma_k}(x)dx. \quad (38)$$

In order to calculate the integral (38), we apply a linear approximation for Q-function as $Q\Big(\frac{C(\gamma_k)-R_k}{\sqrt{V(\gamma_k)/\beta_k}}\Big) \approx \Xi(\gamma_k)$, where $\Xi(\gamma_k)$ can be expressed as [7], [11]

$$\Xi(\gamma_k)=\begin{cases} 1, & \gamma_k \le v_k, \\ 1/2-\delta_k\sqrt{\beta_k}(\gamma_k-\zeta_k), & v_k \le \gamma_k \le u_k, \\ 0, & \gamma_k \ge u_k. \end{cases} \tag{39}$$

With the above tight approximation, ε_{e2e} can be expressed as $\varepsilon_{e2e} = \int_0^\infty \Xi(\gamma_k) f_{\gamma_k}(x)dx$, where it is further calculated by utilizing the partial integration method as

$$\varepsilon_{e2e} = \delta_k\sqrt{\beta_k}\int_{v_k}^{u_k} F_{\gamma_k}(x)dx. \tag{40}$$

We next derive the CDFs of γ_1 and γ_k. Considering (4) with $\|\mathbf{h}_{1,i}\|$ being a Gamma random variable (RV), and consequently γ_1 is the maximum of N_k Gamma RVs. Thus, the CDF of γ_1 can be expressed as [4, Eq. (25)]

$$\begin{aligned} F_{\gamma_1}(x) =& 1+\sum_{t=1}^{N_1}\sum_{q=0}^{t(N_{\mathsf{S}}-1)}(-1)^t\binom{N_1}{t}\Theta(q,N_{\mathsf{S}},t)\Big(\frac{x}{\bar{\gamma}_1\lambda_{D,1}}\Big)^q \\ &\times\exp\Big(-\frac{tx}{\bar{\gamma}_1\lambda_{D,1}}\Big), \end{aligned} \tag{41}$$

where coefficient $\Theta(q,N_{\mathsf{S}},t)$ can be expressed as

$$\begin{aligned} &\Theta(0,N_{\mathsf{S}},t)=1, \Theta(N_{\mathsf{S}}-1,N_{\mathsf{S}},t)=[(N_{\mathsf{S}}-1)!]^{-t}, \\ &\Theta(q,N_{\mathsf{S}},t)=\frac{1}{q}\sum_{j=1}^{J}\frac{j(t+1)-q}{j!}b_{q-j}^t, 2\le q\le t(N_{\mathsf{S}}-1)-1, \\ &\Theta(1,N_{\mathsf{S}},t)=t, J=\min(q,N_{\mathsf{S}}-1). \end{aligned} \tag{42}$$

Plugging (41) into (40) with the index $k=1$, it follows that

$$\begin{aligned} \varepsilon_1 =& 1+\sum_{t=1}^{N_1}\sum_{q=0}^{t(N_{\mathsf{S}}-1)}(-1)^t\binom{N_1}{t}\Theta(q,N_{\mathsf{S}},t) \\ &\times\int_{v_k}^{u_k}\Big(\frac{x\bar{\gamma}^{-1}}{\lambda_{D,1}}\Big)\exp\Big(-\frac{x\bar{\gamma}^{-1}}{\lambda_{D,1}}\Big)dx. \end{aligned} \tag{43}$$

After solving the integral in (43), ε_1 can be obtained as shown in (10). For the remaining dual-hop transmission, from (6), the CDF of γ_k can be expressed as

$$\begin{aligned} F_{\gamma_k}(x) =& 1+\sum_{p=1}^{N_{k-1}}\binom{N_{k-1}}{p}\Big(\frac{px}{\chi_{k-1}\bar{\gamma}\lambda_{D,k-1}\lambda_{E,k-1}}\Big)^{\frac{L}{2}} \\ &\times\frac{2(-1)^{p-1}}{\Gamma(L)}\mathcal{K}_L\Big(2\sqrt{\frac{px}{\chi_{k-1}\bar{\gamma}\lambda_{D,k-1}\lambda_{E,k-1}}}\Big) \\ &\times\sum_{t=1}^{N_k}\sum_{q=0}^{t(N_{k-1}-1)}\binom{N_k}{t}\frac{2(-1)^t\Theta(q,N_{k-1},t)}{\Gamma(L)t^q} \\ &\times\Big(\frac{tx}{\chi_k\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}\Big)^{\frac{L+q}{2}}\mathcal{K}_{L-q}\Big(2\sqrt{\frac{tx}{\chi_k\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\Big). \end{aligned} \tag{44}$$

It is difficult to further calculate the BLER for the CRS scheme with the above expression since (44) shows the product of two BesselK function. Therefore, we represent $\mathcal{K}_{L-q}(x)$ in (44) into the exponential forms by using the following Lemma.

Lemma 1. *The equivalent representation of $\mathcal{K}_{L-q}(x)$ can be expressed as* (45) *and* (46)*, given at the top of the next page, where $n_{\max}$ indicates the limited upper bound of the summation, which can be suitably set satisfying the desired accuracy level.*

Proof: For the case $L \neq q$, we apply [39, Eq. (3)], which results in (45). For the case $L = q$, we utilize the recursive relationship of $\mathcal{K}_v(t) = \mathcal{K}_{v-1}(t) + \frac{2(v-1)}{t}\mathcal{K}_{v-1}(t)$, along with the identity $\mathcal{K}_v(t) = \mathcal{K}_{-v}(t)$, we can obtain the desired results as (46). ■

Recalling Lemma 1, $\mathcal{K}_L(2\sqrt{ax})$ can be approximated as

$$\mathcal{K}_L(2\sqrt{ax}) = \exp(-2\sqrt{ax})\sum_{n=0}^{n_{\max}}\sum_{i=0}^{n}\Lambda(L,n,i)(2\sqrt{ax})^{i-L}, \tag{47}$$

Plugging (47) and (45) into (44), the CDF of γ_k can be further rewritten as

$$\begin{aligned} F_{\gamma_k}(x) =& 1+\sum_{p=1}^{N_{k-1}}\sum_{t=1}^{N_k}\sum_{q=0}^{t(N_{k-1}-1)}\binom{N_{k-1}}{p}\binom{N_k}{t} \\ &\times\frac{(-1)^{p+t-1}\Theta(q,N_{k-1},t)}{\Gamma(L)^2t^q2^{2L+q-2}}\sum_{n_1=0}^{n_{max}}\sum_{i_1=0}^{n_1}\Lambda(L,n_1,i_1) \\ &\times\sum_{n_2=0}^{n_{max}}\sum_{i_2=0}^{n_2}\Lambda(|L-q|,n_2,i_2)\Big(2\sqrt{\frac{p\chi_{k-1}^{-1}}{\bar{\gamma}\lambda_{D,k-1}\lambda_{E,k-1}}}\Big)^{i_1} \\ &\times\Big(2\sqrt{\frac{\chi_k^{-1}t}{\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\Big)^{i_2+2\min(L,q)}x^{\frac{i_1+i_2+2\min(L,q)}{2}} \\ &\times\exp\Big(-2\Big(\sqrt{\frac{\chi_{k-1}^{-1}p}{\bar{\gamma}\lambda_{D,k-1}\lambda_{E,k-1}}}+\sqrt{\frac{\chi_k^{-1}t}{\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\Big)\sqrt{x}\Big). \end{aligned} \tag{48}$$

Substituting (48) into (40), we obtain the desired result as (11). Similarly, by plugging (47) and (46) into (44), the CDF of γ_k can be further rewritten as

$$\begin{aligned} F_{\gamma_k}(x) =& 1+\sum_{p=1}^{N_{k-1}}\sum_{t=1}^{N_k}\sum_{q=0}^{t(N_{k-1}-1)}\binom{N_{k-1}}{p}\binom{N_k}{t} \\ &\times\frac{(-1)^{p+t-1}\Theta(q,N_{k-1},t)}{\Gamma(L)^2t^q2^{2L+q-2}}\sum_{n_1=0}^{n_{max}}\sum_{i_1=0}^{n_1}\Lambda(L,n_1,i_1) \\ &\times\sum_{n_2=0}^{n_{max}}\sum_{i_2=0}^{n_2}[\Lambda(2,n_2,i_2)-2\Lambda(1,n_2,i_2)] \\ &\times\Big(2\sqrt{\frac{\chi_{k-1}^{-1}p}{\bar{\gamma}\lambda_{D,k-1}\lambda_{E,k-1}}}\Big)^{i_1}\Big(2\sqrt{\frac{\chi_k^{-1}t}{\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\Big)^{L+q+i_2-2} \\ &\times\exp\Big(-2\Big(\sqrt{\frac{\chi_{k-1}^{-1}p}{\bar{\gamma}\lambda_{D,k-1}\lambda_{E,k-1}}}+\sqrt{\frac{\chi_k^{-1}t}{\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\Big)\sqrt{x}\Big) \\ &\times x^{\frac{L+q+i_1+i_2-2}{2}}. \end{aligned} \tag{49}$$

Substituting (49) into (40), we obtain the result as (12). The proof of Theorem 1 is concluded.

$$\mathcal{K}_{L-q}(x)=\begin{cases}\exp(-x)\sum\limits_{n=0}^{n_{\max}}\sum\limits_{i=0}^{n}\Lambda(|L-q|,n,i)x^{i-|L-q|}, & \text{if} \quad L\neq q, \quad (45)\\ \exp(-x)\sum\limits_{n=0}^{n_{\max}}\sum\limits_{i=0}^{n}[\Lambda(2,n,i)-2\Lambda(1,n,i)]x^{i-2}, & \text{if} \quad L=q, \quad (46)\end{cases}$$

APPENDIX B
PROOF OF THEOREM 2

The average BLER of each hop is obtained by taking the integral over $[v_k, u_k]$, with $u_k - v_k = 1/(\delta_k\sqrt{\beta_k})$ being a small number. Thus, we apply the first order Riemann integral approximation $\int_{v_k}^{u_k} f(x)dx = (u_k - v_k)f\left(\frac{v_k+u_k}{2}\right)$ for the first hop transmission, where ε_1 is obtained in a compact form as

$$\varepsilon_1 = 1 + \sum_{t=1}^{N_1}\sum_{q=0}^{t(N_{\mathsf{S}}-1)}(-1)^t\binom{N_1}{t}\Theta(q,N_{\mathsf{S}},t)\left(\frac{\zeta_1}{\bar{\gamma}_1\lambda_{D,1}}\right)^q \times \exp\left(-\frac{\zeta_1}{\bar{\gamma}_1\lambda_{D,1}}\right). \quad (50)$$

For the remaining hops, ε_k is obtained in a compact form as

$$\varepsilon_k = 1 - \sum_{t=1}^{N_k}\binom{N_k}{t}\frac{2(-1)^{t-1}}{\Gamma(L)}\left(\frac{t\zeta_k}{\chi_k\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}\right)^{\frac{L}{2}} \times \mathcal{K}_L\left(2\sqrt{\frac{t\zeta_k}{\chi_k\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\right) \quad (51)$$

Then, invoking Lemma 1 to equivalently rewrite BesselK function in (51), which results as

$$\varepsilon_k = 1 - \sum_{t=1}^{N_k}\sum_{n=0}^{I_{\max}}\sum_{i=0}^{n}\binom{N_k}{t}\frac{(-1)^{t-1}\Lambda(L,n,i)}{\Gamma(L)2^{L-i-1}} \times \left(\frac{t\zeta_k}{\chi_k\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}\right)^{\frac{i}{2}}\exp\left(-2\sqrt{\frac{t\zeta_k}{\chi_k\bar{\gamma}\lambda_{D,k}\lambda_{E,k}}}\right). \quad (52)$$

Furthermore, the asymptotic expression for the BLER of multi-hop system can be attained by recalling the following expression:

$$\prod_{k=1}^{K}(1-x_k) \approx 1 - \sum_{k=1}^{K}x_k, \quad (53)$$

which is valid for small x_k. By plugging (50) and (52) into (53), the asymptotic expression for the BLER of CRS scheme can be asymptotically expressed as (14), which concludes the proof of Theorem 2.

REFERENCES

[1] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5G: RAN, core network and caching solutions," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 4, pp. 3098–3130, 2018.

[2] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sept. 2016.

[3] J. Chen, L. Zhang, Y.-C. Liang, X. Kang, and R. Zhang, "Resource allocation for wireless-powered IoT networks with short packet communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1447–1461, Feb. 2019.

[4] T.-V. Nguyen and B. An, "Cognitive multihop wireless powered relaying networks over Nakagami-*m* fading channels," *IEEE Access*, vol. 7, pp. 154 600–154 616, Oct. 2019.

[5] T.-V. Nguyen, T.-N. Do, V. N. Q. Bao, D. B. da Costa, and B. An, "Performance analysis of multihop cognitive WPCNs with imperfect CSI," in *2019 IEEE Global Commun. Conf., Waikoloa, HI, USA*. IEEE, 2019, pp. 1–6.

[6] Y. Liu, X. Zhu, E. G. Lim, Y. Jiang, and Y. Huang, "Fast iterative semi-blind receiver for URLLC in short-frame full-duplex systems with CFO," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 839–853, Apr. 2019.

[7] B. Makki, T. Svensson, and M. Zorzi, "Wireless energy and information transmission using feedback: Infinite and finite block-length analysis," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5304–5318, Dec. 2016.

[8] H.-M. Wang, Q. Yang, Z. Ding, and H. V. Poor, "Secure short-packet communications for mission-critical IoT applications," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2565–2578, May 2019.

[9] A. A. Nasir, H. Tuan, H. H. Nguyen, M. Debbah, and H. V. Poor, "Resource allocation and beamforming design in the short blocklength regime for URLLC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1321–1335, Feb. 2021.

[10] A. A. Nasir, H. D. Tuan, H. Q. Ngo, T. Q. Duong, and H. V. Poor, "Cell-free massive MIMO in the short blocklength regime for URLLC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5861–5871, Sept. 2021.

[11] O. L. A. López, H. Alves, R. D. Souza, and E. M. G. Fernández, "Ultrareliable short-packet communications with wireless energy transfer," *IEEE Signal Process. Lett.*, vol. 24, no. 4, pp. 387–391, Apr. 2017.

[12] Y. Gu, H. Chen, Y. Li, and B. Vucetic, "Ultra-reliable short-packet communications: Half-duplex or full-duplex relaying?" *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 348–351, Jun. 2018.

[13] O. L. A. López, E. M. G. Fernández, R. D. Souza, and H. Alves, "Ultra-reliable cooperative short-packet communications with wireless energy transfer," *IEEE Sens. J.*, vol. 18, no. 5, pp. 2161–2177, Mar. 2018.

[14] C. She, C. Yang, and T. Q. Quek, "Cross-layer optimization for ultra-reliable and low-latency radio access networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 127–141, Jan. 2018.

[15] O. L. A. López, H. Alves, and M. Latva-aho, "Joint power control and rate allocation enabling ultra-reliability and energy efficiency in SIMO wireless networks," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5768–5782, Aug. 2019.

[16] S. Atapattu, N. Ross, Y. Jing, Y. He, and J. S. Evans, "Physical-layer security in full-duplex multi-hop multi-user wireless network with relay selection," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1216–1232, Feb. 2019.

[17] T. V. Nguyen, T.-N. Do, V. N. Q. Bao, D. B. da Costa, and B. An, "On the performance of multihop cognitive wireless powered D2D communications in WSNs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2684–2699, Mar. 2020.

[18] J. Goel and J. Harshan, "Minimal overhead ARQ sharing strategies for URLLC in multi-hop networks," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), Helsinki, Finland*. IEEE, 2021, pp. 1–7.

[19] T.-V. Nguyen, V.-D. Nguyen, D. B. da Costa, and B. An, "Short-packet communications in multi-hop WPINs: Performance analysis and deep learning design," in *2021 IEEE Global Commun. Conf., Madrid, Spain*. IEEE, 2021, pp. 1–6.

[20] C. D. Ho, T.-V. Nguyen, T. Huynh-The, T.-T. Nguyen, D. B. da Costa, and B. An, "Short-packet communications in wireless-powered cognitive IoT networks: Performance analysis and deep learning evaluation," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2894–2899, Mar. 2021.

[21] M. A. Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, "Mobile big data analytics using deep learning and apache spark," *IEEE Network*, vol. 30, no. 3, pp. 22–29, May-Jun. 2016.

[22] J. Zhang, W. Xia, M. You, G. Zheng, S. Lambotharan, and K.-K. Wong, "Deep learning enabled optimization of downlink beamforming under per-antenna power constraints: Algorithms and experimental demonstration," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3738–3752, Jun. 2020.

[23] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, Apr. 2019.

[24] T. Van Chien, T. N. Canh, E. Björnson, and E. G. Larsson, "Power control in cellular massive MIMO with varying user activity: A deep learning solution," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5732–5748, Sept. 2020.

[25] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, p. 2307, May 2010.

[26] T.-V. Nguyen, T.-N. Tran, K. Shim, T. Huynh-The, and B. An, "A deep-neural-network-based relay selection scheme in wireless-powered cognitive IoT networks," *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7423–7436, May 2020.

[27] T. X. Vu, S. Chatzinotas, V.-D. Nguyen, D. T. Hoang, D. N. Nguyen, M. Di Renzo, and B. Ottersten, "Machine learning-enabled joint antenna selection and precoding design: From offline complexity to online performance," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3710–3722, Jun. 2021.

[28] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, Aug. 1978.

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic Press, 2007.

[30] H. A. Suraweera, M. Soysa, C. Tellambura, and H. K. Garg, "Performance analysis of partial relay selection with feedback delay," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 531–534, Jun. 2010.

[31] R. P. Jover, "LTE security, protocol exploits and location tracking experimentation with low-cost software radio," *arXiv preprint arXiv:1607.05171*, 2016.

[32] A. Hoglund, D. P. Van, T. Tirronen, O. Liberg, Y. Sui, and E. A. Yavuz, "3GPP release 15 early data transmission," *IEEE Commun. Stand. Mag.*, vol. 2, no. 2, pp. 90–96, Jun. 2018.

[33] D. Xu, K. Yu, and J. A. Ritcey, "Cross-layer device authentication with quantum encryption for 5G enabled IIoT in industry 4.0," *IEEE Trans. Ind. Inform.*, vol. Early Access, Nov. 2021.

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 25, 2012.

[35] T.-V. Nguyen, T. Huynh-The, and B. An, "A deep CNN-based relay selection in EH full-duplex IoT networks with short-packet communications," in *2021-IEEE International Conference on Communications (ICC), Montreal, QC, Canada*. IEEE, Aug. 2021, pp. 1–6.

[36] X. Sun, S. Yan, N. Yang, Z. Ding, C. Shen, and Z. Zhong, "Short-packet downlink transmission with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4550–4564, Jul. 2018.

[37] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE network*, vol. 34, no. 3, pp. 134–142, May 2019.

[38] T. Huynh-The, C.-H. Hua, V.-S. Doan, Q.-V. Pham, and D.-S. Kim, "Accurate deep CNN-based waveform recognition for intelligent radar systems," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 2938–2942, Sept. 2021.

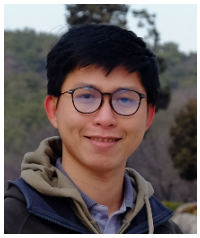
[39] M. M. Molu, P. Xiao, M. Khalily, L. Zhang, and R. Tafazolli, "A novel equivalent definition of modified Bessel functions for performance analysis of multi-hop wireless communication systems," *IEEE Access*, vol. 5, pp. 7594–7605, May 2017.

Toan-Van Nguyen (M'21) received the B.S. degree in electronics and telecommunications engineering and the M.S. degree in electronics engineering from HCMC University of Technology and Education, Vietnam, in 2011 and 2014, respectively, and the Ph.D. degree in electronics and computer engineering from Hongik University, Sejong, Republic of Korea, in 2021. He is currently a postdoctoral researcher at Utah State University, UT, USA. His current research activity is focused on the mathematical modeling of 6G networks and machine learning for wireless communications.

Van-Dinh Nguyen (S'14-M'19) received the B.E. degree in electrical engineering from Ho Chi Minh City University of Technology, Vietnam, in 2012 and the M.E. and Ph.D. degrees in electronic engineering from Soongsil University, Seoul, South Korea, in 2015 and 2018, respectively. He is currently a Research Associate with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg. He was a Postdoc Researcher and a Lecturer with Soongsil University, a Postdoctoral Visiting Scholar with University of Technology Sydney, AUS (July-August 2018) and a Ph.D. Visiting Scholar with Queen's University Belfast, U.K. (June-July 2015 and August 2016). His current research activity is focused on the mathematical modeling of 5G cellular networks and machine learning for wireless communications.

Dr. Nguyen received several best conference paper awards, the Exemplary Editor Award of IEEE COMMUNICATIONS LETTERS 2019, IEEE TRANSACTION ON COMMUNICATIONS Exemplary Reviewer 2018 and IEEE GLOBECOM Student Travel Grant Award 2017. He has authored or co-authored in some 40 papers published in international journals and conference proceedings. He has served as a reviewer for many top-tier international journals on wireless communications, and has also been a Technical Programme Committee Member for several flag-ship international conferences in the related fields. He is an Editor for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and IEEE COMMUNICATIONS LETTERS.

Dr. Daniel Benevides da Costa was born in Fortaleza, Ceará, Brazil, in 1981. He received the B.Sc. degree in Telecommunications from the Military Institute of Engineering (IME), Rio de Janeiro, Brazil, in 2003, and the M.Sc. and Ph.D. degrees in Electrical Engineering, Area: Telecommunications, from the University of Campinas, SP, Brazil, in 2006 and 2008, respectively. His Ph.D thesis was awarded the Best Ph.D. Thesis in Electrical Engineering by the Brazilian Ministry of Education (CAPES) at the 2009 CAPES Thesis Contest. From 2008 to 2009, he was a Postdoctoral Research Fellow with INRS-EMT, University of Quebec, Montreal, QC, Canada. From 2010 to 2022, he was with the Federal University of Ceará, Brazil. From January 2019 to April 2019, he was Visiting Professor at Lappeenranta University of Technology (LUT), Finland, with financial support from Nokia Foundation. He was awarded with the prestigious Nokia Visiting Professor Grant. From May 2019 to August 2019, he was with King Abdullah University of Science and Technology (KAUST), Saudi Arabia, as a Visiting Faculty, and from September 2019 to November 2019, he was a Visiting Researcher at Istanbul Medipol University, Turkey. From 2021 to 2022, he was Full Professor at the National Yunlin University of Science and Technology (YunTech), Taiwan. Since 2022, he is with the Digital Science Research Center at the Technology Innovation Institute (TII), a global research center and the applied pillar of Abu Dhabi's Advanced Technology Research Council. He is Editor of several IEEE journals and has acted as Symposium/Track Co-Chair in numerous IEEE flagship conferences.



Thien Huynh-The (Member, IEEE) received the B.S. degree in electronics and telecommunication engineering from Ho Chi Minh City University of Technology and Education, Vietnam, in 2011, and the Ph.D. degree in computer science and engineering from Kyung Hee University (KHU), South Korea, in 2018. From March to August 2018, he was a Postdoctoral Researcher with KHU.

He is currently a Postdoctoral Research Fellow with ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea. He is awarded with the Superior Thesis Prize by KHU in 2018 and the Golden globe award 2022 for Vietnamese young scientists by Ministry of Science and Technology of Vietnam in 2020. His current research interest includes radio signal processing, digital image processing, computer vision, wireless communications, machine learning, and deep learning.

Rose Qingyang Hu (F'20) received the B.S. degree from the University of Science and Technology of China, the M.S. degree from New York University, and the Ph.D. degree from the University of Kansas. Besides a decade academia experience, she has more than 10 years of R&D experience with Nortel, Blackberry, and Intel as a Technical Manager, a Senior Wireless System Architect, and a Senior Research Scientist, actively participating in industrial 3G/4G technology development, standardization, system level simulation, and performance evaluation. She is a Professor with the Electrical and Computer Engineering Department and Associate Dean for research of College of Engineering at Utah State University. She also directs Communications Network Innovation Lab at Utah State University. Her current research interests include next generation wireless system design and optimization, Internet of Things, Cyber Physical system, Mobile Edge Computing, V2X communications, artificial intelligence in wireless networks, wireless system modeling and performance analysis. She has published extensively in top IEEE journals and conferences and also holds numerous patents in her research areas. She is currently serving on the editorial boards of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE Communications Magazine and IEEE WIRELESS COMMUNICATIONS. She also served as the TPC Co-Chair for the IEEE ICC 2018. She is an IEEE Communications Society Distinguished Lecturer Class 2015-2018 and a recipient of prestigious Best Paper Awards from the IEEE GLOBECOM 2012, the IEEE ICC 2015, the IEEE VTC Spring 2016, and the IEEE ICC 2016. She is member of Phi Kappa Phi Honor Society.

Beongku An Beongku An received the B.S. degree in electronic engineering from Kyungpook National University, Republic of Korea, in 1988, the M.S. degree in electrical engineering from Polytechnic University(NYU), NY, USA, in 1996 and Ph.D. degree from New Jersey Institute of Technology (NJIT), NJ, USA, in 2002, respectively. After graduation, he joined the Faculty of the Department of Software and Communications Engineering, Hongik University, Republic of Korea, where he is currently a Professor. From 1989 to 1993, he was a senior researcher in RIST, Republic of Korea. He also was lecturer and RA in NJIT from 1997 to 2002. He was a president of IEIE Computer Society in 2012. From 2013, he also works as a General Chair in the International Conference, ICGHIT. His current research interests include mobile wireless networks and communications such as ad-hoc networks, sensor networks, cognitive radio networks, cellular networks, and IoT. In particular, he is interested in cooperative transmission & QoS routing, QoS multicast routing, energy harvesting, physical layer security(PLS), visible light communication(VLC), cross-layer technology, 5G/Beyond 5G, Machine Learning/Deep Learning applications. Professor An was listed in Marquis Who's Who in Science and Engineering, and Marquis Who's Who in the World, respectively.