Resource Management for Multiplexing eMBB and URLLC Services Over RIS-Aided THz Communication

Hosein Zarini, Narges Gholipoor, Mohammad Robat Mili,
Mehdi Rasti, *Senior Member, IEEE*, Hina Tabassum, *Senior Member, IEEE*,
and Ekram Hossain, *Fellow, IEEE*

Abstract—Integrating the multitude of emerging internet of things (IoT) applications with diverse requirements in beyond fifth generation (B5G) networks necessitates the coexistence of enhanced mobile broadband (eMBB) and ultra-reliable low-latency communication (URLLC) services. However, bandwidth limited and congested sub-6GHz bands are incapable of fulfilling this coexistence. In this paper, we consider a reconfigurable intelligent surface (RIS)-aided wideband terahertz (THz) communication system to this end. In specific, we formulate a resource management problem, aiming at jointly optimizing the reflection coefficient of the RIS elements and the transmit power of the base station, as well as the wideband THz resource block allocation. To solve this problem, we adopt a supervised learning approach relying on optimization, deep learning and ensemble learning methods. Simulation results show that for an RIS of size 11×11, up to 49% spectral efficiency gain is achieved for the eMBB service compared to the counterparts, while ensuring the reliability and latency requirements of the URLLC service. Further, the ensemble learning model can perform real-time resource management at the expense of up to 1% performance loss, compared to the optimization approach.

Index Terms—Internet of things (IoT), enhanced mobile broadband (eMBB), ultra-reliable low latency communication (URLLC), reconfigurable intelligent surface (RIS), Terahertz (THz) communication, supervised learning approach.

I. INTRODUCTION

Towards supporting a variety of internet of things (IoT) applications with diverse performance requirements ranging from ultra-reliable and extremely low end-to-end (E2E) delay to a high data rate transmission, service-based radio resource management is becoming more and more critical [1]. In this regard, coexisting the enhanced mobile broadband (eMBB) and ultra-reliable low-latency communication (URLLC) services, turns out to be indispensable in next generation of wireless networks. However, supporting high data rates for the eMBB service as well as ultra reliability and latency requirements for the URLLC service brings about a fundamental challenge for the conventional in-use sub-6GHz bands.

Recently, [2]–[8] have investigated this coexistence problem and presented resource allocation frameworks to simultaneously support eMBB and URLLC services using the conventional sub-6GHz, i.e., radio-frequency (RF) bands, with a fixed frame structure, where the duration of each time slot is 0.5 ms [9]. This frame structure cannot meet an IoT-driven URLLC service with an E2E delay of 1 ms [9]. Further, the optimization-driven approaches adopted in [2]–[8] for this coexistence is computationally expensive and results in significant delay, which in turn negatively impact the E2E delay requirement of the URLLC service. In this paper, we overcome the aforementioned issues by considering *wideband Terahertz (THz) transmissions* as well as a *supervised learning-based resource management framework*.

The THz spectrum enjoys large vacant bandwidths and is a potential substitute for bandwidth-limited and interference-prone RF bands to support the IoT-driven URLLC and eMBB services. Unlike the conventional RF bands employed in [2]–[8], the THz channels are characterized by a dynamic frame structure [10], which results in a shorter transmission time [11]. Due to significant signal attenuation however, the THz communication is capable of providing a short coverage only. In recent years, large reconfigurable intelligent surfaces (RISs) have been investigated to overcome this limitation [12].

From the complexity perspective further, the computational burden imposed by conventional optimization-driven resource allocation frameworks in [2]–[8] makes them ill-suited for co-existing the services. Alternatively, learning-aided techniques can offer low-complexity solutions. In this regard, the authors of [13]–[15] have recently adopted a deep reinforcement learning (DRL) approach as a real-time strategy to this end. Despite being a computationally efficient solution, the DRL strategy is however highly prone to the scalability challenge and encounters a non-trivial performance loss owing to lack of generalization in large-scale environments.

The existing works on coexistence of eMBB and URLLC services, i.e., [2]–[8] and [13]–[15] therefore are limited to conventional sub-6GHz bands and are incapable of satisfying the URLLC and eMBB service requirements. Additionally, the aforementioned works use either high-complexity optimization

H. Zarini is with the Department of Computer Engineering at Sharif University of Technology, Tehran, Iran (email: hosein.zarini68@sharif.edu). N. Gholipoor is with the Department of Electrical and Computer Engineering at Tarbiat Modares University, Tehran, Iran (email: gholipoor.narges@modares.ac.ir). M. Robat Mili is with Pasargad Institute for Advanced Innovative Solutions (PIAIS), Tehran, Iran (email: mohammad.robatmili@piais.ir). M. Rasti is with the University of Oulu, Oulu, Finland (email: mehdi.rasti@Oulu.fi). H. Tabassum is with the Department of Electrical Engineering and Computer Science (EECS) at York University, Toronto, Canada (email: Hina.Tabassum@lassonde.yorku.ca). E. Hossain is with the University of Manitoba, Canada (email: Ekram.Hossain@umanitoba.ca). The work of M. Rasti was supported by the University of Oulu and the Academy of Finland Profi6 336449.

methods [2]–[8], or non-generalizable DRL methods [13]–[15] to this end. Hence, we propose a comprehensive resource allocation framework in this paper based on a RIS-aided THz communication as well as a supervised learning approach in order to efficiently support this coexistence.

The main contributions of this paper are as follows:

- We propose an RIS-aided THz communication system to support the coexistence of URLLC and eMBB services. To do so, a puncturing approach is adopted in favour of the URLLC service, whereas the wideband THz resource blocks (RBs) for the eMBB service are allocated based on a non-orthogonal multiple access (NOMA) manner.
- We formulate a resource management problem for jointly optimizing the key system parameters including the RIS reflection coefficients, the base station (BS) transmit power and the wideband THz RBs, while ensuring the reliability and the transmission/queuing latency for URLLC users, as well as the quality of service (QoS) for eMBB users. Since the resource management optimization problem is non-convex and extremely complicated to solve, a supervised learning approach is devised relying on optimization, deep learning, and ensemble learning. Specifically, different from [4], where the authors propose an optimization-driven method for coexisting the eMBB and URLLC services in sbu-6GHz band, we investigate the problem for wideband THz band by using a supervised learning mechanism.
- We use an alternative decomposition technique to decompose the joint problem and optimize the key system parameters. By invoking a quadratic transformation, we reformulate the power control and RIS reflection design sub-problems into convex problems of minimizing the mean square error (MMSE). Later, by leveraging the relaxation technique and incorporating binary-forcing constraints, we efficiently solve the RB allocation sub-problem. We prove that the proposed alternative decomposition policy converges to a sub-optimal solution for the joint optimization problem.
- After optimizing the system parameters via the alternative decomposition technique, we train a long short term memory (LSTM) [16] through a Levenberg-Marquardt policy [17] for tracing and predicting the key system parameters in future. Based on an interpolation between the Gauss-Newton and gradient descent method for minimizing the squared errors, the Levenberg-Marquardt policy precisely determines the weights and biases in the LSTM. Meanwhile, we fine-tune the training procedure by optimizing the LSTM learning rate hyperparameter using a Bayesian optimizer [20]. As a result, fast convergence and limited prediction errors are observed. By sequentially accommodating multiple LSTMs, we finally form an ensemble model, called bootstrap aggregation [21], that improves the precision of a single LSTM.
- Simulations reveal that the proposed alternative decomposition solution achieves up to 49% eMBB spectral efficiency (SE) gain over the DRL-based counterpart [15] for coexisting eMBB and URLLC services. Besides,

the fine-tuned LSTM performs a real-time prediction of the key system parameters in future, converges within limited training epochs, exhibits negligible prediction error, and achieves up to 34% eMBB SE gain over the DRL-based scheme [15]. Further, the proposed ensemble model not only exhibits the same complexity order of training a single LSTM, but also improves the achievable spectral efficiency for eMBB services over the DRL-based strategy [15] by as much as 48%.

The remainder of this paper is organized as follows: Section II describes the system setup, the problem statement and the solution methodology. We investigate the alternative decomposition policy in Section III. In Section IV, the details related to training the LSTM, its fine-tuning procedure and the ensemble learning technique are elaborated. Section V provides a comprehensive complexity analysis of the proposed solution approach. The simulation results and conclusions are presented in Sections VI and VII, respectively.

II. SYSTEM MODEL, ASSUMPTIONS, AND PROBLEM FORMULATION

A. Network Structure

Consider a BS equipped with N_t antennas to serve a set $\mathcal{K}$ of $K = \{1, ..., |\mathcal{K}|\}$ eMBB users, as well as a set $\mathcal{U}$ of $U = \{1, ..., |\mathcal{U}|\}$ URLLC users (called teleoperators from now on) in a downlink wideband THz communication over a set $\mathcal{S}$ of $S = \{1, 2, ...|\mathcal{S}|\}$ RBs. Due to the short-range coverage of THz transmissions, the communicating network is assisted by an RIS with M passive elements and a controller module to reconfigure the reflection coefficients of its elements. The channel state information (CSI) is assumed to be available at both the BS and the RIS [16]. We also assume that the RIS controller acquires all the CSI information during the channel estimation phase. If the CSI is partially available, our results will serve as theoretical performance upper-bound for the considered system. For more practical scenarios, refer to [18], [19] on effective channel estimation of both direct and reflecting channels in a RIS-aided communication environment. We consider T_s LTE time slots denoted by $\mathcal{T}_s = \{1, 2, ..., T_s\}$ to support the services, each with a length Δ (divided into F mini-slots each with a duration of δ).

B. Channel Model

We denote the channel matrix from the BS to the RIS over the RB s by $\mathbf{G}[s] \in \mathbb{C}^{M \times N_t}$. For a clustered ray-based wideband THz channel with Rician fading[1], according to the well-known Saleh-Valenzuela geometric model [22], $\mathbf{G}[s]$ is given by

$$\mathbf{G}[s] = \sum_{d=0}^{D_{\text{samp}}-1} \sum_{l=1}^{N_{\text{cl}}^{\{1\}}} \sum_{q=1}^{N_{\text{ray}}^{\{1\}}} \alpha_{l,q} \delta(dT_{\text{samp}} - \tau_{l,q}) \tag{1}$$
$$\mathbf{a}_{\text{RIS}}\left(\phi_{\text{RIS}}^{l,q}[s]\right) \mathbf{a}_{\text{BS}}^H\left(\phi_{\text{BS}}^{l,q}[s]\right) e^{-j2\pi d \frac{s}{S}},$$

[1]Note that this distribution considers both line-of-sight (LoS) and non-line-of-sight (NLoS) links, where the NLoS components conventionally use Rayleigh distribution, while the LoS components are spawned by uniform linear array (ULA). Since the NLoS links are not dominant in the high frequency bands such as the THz bands, we only consider the LoS components.

with $N_{\text{cl}}^{\{1\}}$ clusters, $N_{\text{ray}}^{\{1\}}$ rays and the complex path gain α. Besides, in (1), j denotes the imaginary unit, δ is the band-limited pulse-shaping filter, T_{samp} is the cycle prefix length and D_{samp} is the sampling time. Moreover, for a typical user, $\mathbf{a}_{\text{BS}}(\phi) \in \mathbb{C}^{N \times 1}$ and $\mathbf{a}_{\text{RIS}}(\phi) \in \mathbb{C}^{M \times 1}$ are the antenna array response vectors for ULA at the BS and the RIS, respectively, as given by

$$\mathbf{a}_{\text{BS}}(\phi_{\text{BS}}) = \frac{1}{\sqrt{N}}\left[1, e^{-j2\pi\phi^{l,q}[s]}, ..., e^{-j2\pi(N-1)\phi^{l,q}[s]}\right], \quad (2)$$

and

$$\mathbf{a}_{\text{RIS}}(\phi_{\text{RIS}}) = \frac{1}{\sqrt{M}}\left[1, e^{-j2\pi\phi^{l,q}[s]}, ..., e^{-j2\pi(M-1)\phi^{l,q}[s]}\right], \quad (3)$$

wherein ϕ_{BS} and ϕ_{RIS} are the spatial angle-of-departure (AoD) directions for the BS and the RIS, defined as $\phi_{\text{BS}} = \phi_{\text{RIS}} = \frac{b}{\lambda[s]}$, with the antenna/element spacing b and the wavelength of the RB s, denoted by $\lambda[s]$.

We also denote by $\boldsymbol{h}_{d,k}[s] \in \mathbb{C}^{N_t \times 1}$ and $\boldsymbol{h}_{r,k}[s] \in \mathbb{C}^{M \times 1}$ the direct channel from the BS to the eMBB user k and the reflecting channel from the RIS to this user, both over the RB s, respectively, as given by

$$h_{d,k}[s] = \quad (4)$$
$$\sum_{d=0}^{D_{\text{samp}}-1} \sum_{l=1}^{N_{\text{cl}}^{\{2\}}} \sum_{u=1}^{N_{\text{ray}}^{\{2\}}} \alpha_{l,u}\delta(dT_{\text{samp}} - \tau_{l,u})\mathbf{a}_{\text{BS}}\left(\bar{\phi}_{\text{BS},k}^{l,u}[s]\right)e^{-j2\pi d\frac{s}{S}},$$

and

$$h_{r,k}[s] = \quad (5)$$
$$\sum_{d=0}^{D_{\text{samp}}-1} \sum_{l=1}^{N_{\text{cl}}^{\{3\}}} \sum_{u=1}^{N_{\text{ray}}^{\{3\}}} \alpha_{l,u}\delta(dT_{\text{samp}} - \tau_{l,u})\mathbf{a}_{\text{RIS}}\left(\bar{\phi}_{\text{RIS},k}^{l,u}[s]\right)e^{-j2\pi d\frac{s}{S}},$$

where $N_{\text{cl}}^{\{2\}}$ ($N_{\text{ray}}^{\{2\}}$) indicates the number of clusters (rays in each cluster) for $h_{d,k}$, while $N_{\text{cl}}^{\{3\}}$ ($N_{\text{ray}}^{\{3\}}$) is the number of clusters (rays in each cluster) for $h_{r,k}$. Note that $\bar{\phi}_{\text{BS},k}^{l,u}[s]$ and $\bar{\phi}_{\text{RIS},k}^{l,u}[s]$ in (4) and (5), respectively, denote the spatial AoD of the BS and the RIS towards user k. Similarly, for URLLC service, the direct channel from the BS to the uth teleoperator and the reflecting channel from the RIS to this user, over RB s within the fth mini-slot are denoted by $\tilde{\boldsymbol{h}}_{d,u}^{f}[s]$ and $\tilde{\boldsymbol{h}}_{r,u}^{f}[s]$, respectively. Considering that LoS propagation is dominant in the THz band, for the BS-to-RIS, RIS-to-user, and BS-to-user links, channel fading distribution only for LoS propagation is considered.

C. RIS Configuration

For designing optimal reflection coefficient for an RIS with M elements, coherent combining at the user yields a high array gain proportional to M^2, which is a major motivation for RIS-assisted communications [23]. Let $\boldsymbol{\Theta} = \text{diag}(\theta_1, \theta_2, ..., \theta_M)$ be the diagonal RIS reflection coefficient matrix, where $\theta_m = \beta_m e^{j\varphi_m}$, with $\beta_m \in [0,1]$ and $\varphi_m \in [0, 2\pi]$ representing the amplitude and the phase of the m-th RIS reflecting element, respectively, and j is the imaginary unit. For a more realistic and practical modelling of the RIS, the amplitude

coefficient is a function of the phase shift and the reflection coefficient can be stated as [26]: $\theta_m = \beta_m(\varphi_m)e^{j\varphi_m}$, where $\beta_m(\varphi_m) = (1 - \beta_{\min})\left(\frac{\sin(\varphi_m - \phi) + 1}{2}\right)^k + \beta_{\min}$. Meanwhile, $\beta_{\min} \geq 0$ is the minimum amplitude, $\phi \geq 0$ is the horizontal distance between $-\frac{\pi}{2}$ and $\beta_{\min}$, and $k \geq 0$ controls the steepness of the function curve. Here, $\beta_{\min}$, ϕ and k are constants related to the specific circuit implementation. The feasible set for reflection coefficient of the RIS is given as $\mathcal{Q} \triangleq \{\theta_m, |\theta_m| \leq 1\} \forall m \in M$, which satisfies [25]

$$\textbf{C1:} \ |\theta_m| \leq 1, \forall m. \quad (6)$$

Note that more realistic models for the RIS architecture such as [26], [27] can also be used in our system model and may further enhance the results of this study.

D. URLLC Service

Let us define the binary variable $\tilde{\rho}_u^f[s]$, which is 1 if RB s is allocated to the uth teleoperator within the fth mini-slot and zero, otherwise. To guarantee the feasibility of URLLC service, each user must be allocated at least one RB; therefore,

$$\textbf{C2:} \ \sum_{s \in \mathcal{S}} \sum_{f \in \mathcal{F}} \tilde{\rho}_u^f[s] \geq 1, \forall u \in \mathcal{U}. \quad (7)$$

We adopt a puncturing method to multiplex the services on a shared RB. By doing so, one can guarantee that the URLLC transmission will not interfere with the eMBB transmissions [4]. The received signal at uth URLLC user within the fth mini-slot of RB s is thus represented by

$$\tilde{\mathbf{y}}_u^f[s] = \left(\tilde{\boldsymbol{h}}_{d,u}^f[s] + \mathbf{G}[s]\boldsymbol{\Theta}\tilde{\boldsymbol{h}}_{r,u}^f[s]\right)\tilde{\mathbf{x}}_u^f[s] + \tilde{\mathbf{n}}_u^f[s], \quad (8)$$

where $\tilde{\mathbf{x}}_u^f[s] = \tilde{w}_u^f[s]\sqrt{\tilde{p}_u^f[s]}\tilde{\mathbf{c}}_u^f[s]$ is the transmitted signal at the BS with $\tilde{c}_u^f[s]$ being the transmit data symbol, $\tilde{w}_u^f[s] \in \mathbb{C}^{M \times 1}$ being the corresponding transmit beamforming vector ($\| \tilde{\mathbf{n}}_u^f[s] \|_2 = 1$) and $\tilde{\mathbf{n}}_u^f[s]$ being the unit variance complex additive white Gaussian noise (AWGN), all for the uth tele-operator within the fth mini-slot of RB s. Accordingly, the received signal-to-noise-ratio (SNR) at URLLC user u within the fth mini-slot of RB s can be expressed as

$$\tilde{\gamma}_u^f[s] = \left| \left(\tilde{\boldsymbol{h}}_{d,u}^f[s] + \mathbf{G}[s]\boldsymbol{\Theta}\tilde{\boldsymbol{h}}_{r,u}^f[s]\right)\tilde{w}_u^f[s] \right|^2 \tilde{p}_u^f[s]\tilde{\sigma}^{-2}. \quad (9)$$

With respect to the short blocklength regime of the URLLC transmissions, the achievable rate for the uth URLLC user is calculated as $\tilde{R}_u = \sum_{s \in \mathcal{S}} \sum_{f \in \mathcal{F}} \tilde{\rho}_u^f[s]\tilde{R}_u^f[s]$ [4], in which

$$\tilde{R}_u^f[s] = \frac{\delta}{\ln 2}\left(B_0 \log_2(1 + \tilde{\gamma}_u^f[s]) - \sqrt{\frac{\Gamma_u^f[s]}{\Omega_u^f[s]}}Q^{-1}(\varepsilon_u^f[s])\right). \quad (10)$$

In (10), δ indicates the duration of a mini-slot, $\varepsilon_u^f[s]$ is the decoding error probability, $\Omega_u^f[s]$ denotes the blocklength of the channel code as $\Omega_u^f[s] = B_0\delta$, $Q^{-1}(.)$ is the inverse of the Gaussian Q-function and $\Gamma_u^f[s]$ represents the channel dispersion, defined as $\Gamma_u^f[s] = 1 - (1 + \tilde{\gamma}_u^f[s])^{-2}$.

By simple manipulations, one can reformulate (10) in the following steps:

$$\frac{\tilde{R}_u^f[s]\ln 2}{\delta} = \left(B_0\log_2(1+\tilde{\gamma}_u^f[s]) - \sqrt{\frac{\Gamma_u^f[s]}{\Omega_u^f[s]}}Q^{-1}(\varepsilon_u^f[s])\right),$$
(11)

$$\sqrt{\frac{\Gamma_u^f[s]}{\Omega_u^f[s]}}Q^{-1}(\varepsilon_u^f[s]) = \left(B_0\log_2(1+\tilde{\gamma}_u^f[s]) - \frac{\tilde{R}_u^f[s]\ln 2}{\delta}\right),$$
(12)

$$Q^{-1}(\varepsilon_u^f[s]) = \sqrt{\frac{\Omega_u^f[s]}{\Gamma_u^f[s]}}\left(B_0\log_2(1+\tilde{\gamma}_u^f[s]) - \frac{\tilde{R}_u^f[s]\ln 2}{\delta}\right),$$
(13)

which therefore yields

$$\varepsilon_u^f[s] = Q\left[\sqrt{\frac{\Omega_u^f[s]}{\Gamma_u^f[s]}}\left(B_0\log_2(1+\tilde{\gamma}_u^f[s]) - \frac{\tilde{R}_u^f[s]\ln 2}{\delta}\right)\right].$$
(14)

For a typical URLLC user, the QoS constraint needs to be satisfied as follows [4]:

$$\textbf{C3: } \varepsilon_u^f[s] = Q\left[\sqrt{\frac{\Omega_u^f[s]}{\Gamma_u^f[s]}}\left(B_0\log_2(1+\tilde{\gamma}_u^f[s]) - \frac{\tilde{R}_u^f[s]\ln 2}{\delta}\right)\right] \leq \tilde{\varepsilon}_u^{\max}.$$
(15)

Besides, the reliability of the URLLC users can be assured as follows:

$$\textbf{C4: } \Pr\left(\sum_{u\in\mathcal{U}}\tilde{\rho}_u^f[s] \leq U\right) \leq \tilde{\varepsilon}_u^{\max}, \forall f\in\mathcal{F}, s\in\mathcal{S}.$$
(16)

The uth URLLC user experiences a downlink transmission delay, denoted by D_u^{T}, as well as a queuing delay, denoted by D_u^{Q}. The maximum tolerable one-way delay for the uth URLLC user, denoted by $D_u^{\max}$ is required to be considered as follows:

$$\textbf{C5: } D_u^{\mathrm{T}} + D_u^{\mathrm{Q}} \leq D_u^{\max}, \forall u\in\mathcal{U}.$$
(17)

wherein D_u^{T} and D_u^{Q} are constrained on their own. The downlink transmission delay D_u^{T} is constrained to

$$\textbf{C6: } D_u^{\mathrm{T}} \leq \frac{C_u}{\tilde{R}_u}, \forall u\in\mathcal{U},$$
(18)

where C_u denotes the total transmitted bits per packet for the uth user. The aggregation of arrival bit rate for the teleoperators can be modeled as a Poisson process [28] and thus the effective capacity of the uth user is obtained as:

$$EC_u = \tilde{\Lambda}_u\frac{(e^{\tilde{\theta}_u}-1)}{\tilde{\theta}_u}, \forall u\in\mathcal{U},$$
(19)

with $\tilde{\theta}_u$ denoting the statistical QoS exponent related to user u. A larger $\tilde{\theta}_u$ obviously represents a more stringent QoS requirement, whereas a smaller $\tilde{\theta}_u$ implies a looser QoS requirement. Moreover, $\tilde{\Lambda}_u$ in (19) is the number of bits arrived in time unit at the BS queue for the uth user and defined as $\tilde{\Lambda}_u = \tilde{R}_u, \forall u\in\mathcal{U}$. The queuing delay violation probability for the URLLC service must be ultra low and given by

$$\tilde{\epsilon}_u^Q = \Pr\{D_u^Q > D_u^{\mathrm{Q,max}}\} = \tilde{\eta}e^{-\tilde{\theta}_u EC_u D_u^Q}, \forall u\in\mathcal{U},$$
(20)

where $D_u^{\mathrm{Q,max}}$ is the maximum queuing delay, $\tilde{\eta}$ is the non-empty buffer probability in the downlink and $\tilde{\delta}$ is the violation probability. By substituting (19), the violation probability in (20) can be simplified as

$$e^{-\tilde{\theta}_u EC_u D_u^Q} = e^{-\tilde{\theta}_u\Lambda_u\frac{(e^{\tilde{\theta}_u}-1)}{\tilde{\theta}_u}D_u^Q} = e^{-\tilde{\Lambda}_u(e^{\tilde{\theta}_u}-1)D_u^Q} \leq \tilde{\delta},$$
(21)

which implies

$$\tilde{R}_u \geq \frac{\ln(1/\tilde{\delta})}{(e^{\tilde{\theta}_u}-1)D_u^Q}, \forall u\in\mathcal{U}.$$
(22)

Accordingly, the queuing delay is constrained to

$$\textbf{C7: } D_u^Q \leq \frac{\ln(1/\tilde{\delta})}{(e^{\tilde{\theta}_u}-1)\tilde{R}_u}, \forall u\in\mathcal{U}.$$
(23)

E. eMBB Service

By adopting the puncturing approach for multiplexing URLLC and eMBB services, the inter-service interference would be eliminated. Thus, the RBs are often occupied in favour of the eMBB requests owing to the bursty nature of the URLLC requests. Inspired by this fact, we adopt a non-orthogonal RB allocation approach within the eMBB service for a higher achievable data rate [30]. However, superposition of multiple eMBB users over a specific RB, induces a non-trivial complexity for the signal decoding process. Therefore, in order to reduce the decoding complexity, the number of multiplexed eMBB users over a specific RB is constrained as follows:

$$\textbf{C8: } \sum_{k\in\mathcal{K}}\rho_k[s] \leq N^{\max}, \forall s\in\mathcal{S}.$$
(24)

Also, each eMBB user needs to get at least one RB as follows:

$$\textbf{C9: } \sum_{s\in\mathcal{S}}\rho_k[s] \geq 1, \forall k\in\mathcal{K}.$$
(25)

Furthermore, the eMBB users should be appropriately ordered in superposition to take the advantage of the NOMA. Assume that $\Upsilon(k)$ indicates the decoding order for the eMBB user k. This user would be corresponded to the k'th decoded signal, provided that $\Upsilon(k) = k'$. In this sense, the received signal at the eMBB user k over the RB s is represented by

$$\mathbf{y}_k[s] = (\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\,\mathbf{x}_k[s] + \tag{26}$$
$$\sum_{\substack{i\in\mathcal{K},i\neq k,\\\Upsilon(k)<\Upsilon(i)}} \left([\boldsymbol{h}_{d,k}[s]]^H + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\mathbf{x}_i[s] + \mathbf{n}_k[s],$$

where $\mathbf{x}_k[s] = \boldsymbol{w}_k[s]\sqrt{p_k[s]}c_k[s]$ is the transmitted signal at the BS, in which $c_k[s]$ denotes the transmitted data symbol, $\boldsymbol{w}_k[s]\in\mathbb{C}^{M\times 1}$ is the corresponding transmit beamforming vector with $||\boldsymbol{w}_k[s]||_2 = 1$ and $\mathbf{n}_k[s]$ represents the complex unit variance AWGN, all for the kth eMBB user over the RB s. The received signal-to-interference-plus-noise ratio (SINR) at the eMBB user k over the RB s, would be accordingly expressed as

$$\gamma_k[s] = \frac{\left|(\boldsymbol{h}_{d,k}[s]+\mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\,\boldsymbol{w}_k[s]\right|^2 p_k[s]}{\sum_{\substack{i\in\mathcal{K},i\neq k,\\\Upsilon(k)<\Upsilon(i)}}\left|([\boldsymbol{h}_{d,k}[s]]^H+\mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\right|^2\rho_i[s]p_i[s]+1}.$$
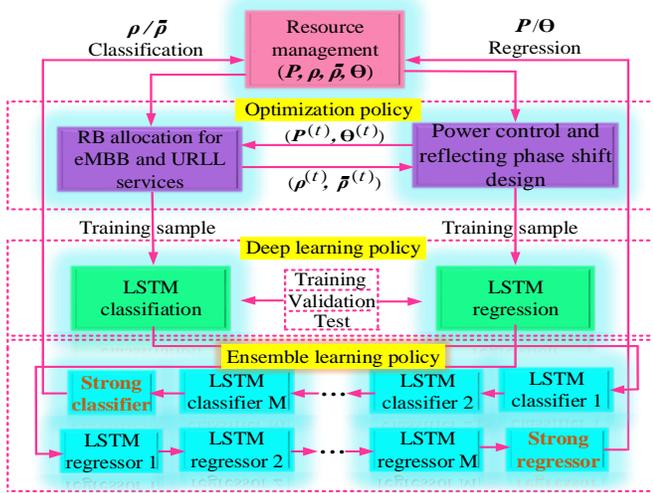(27)

Fig. 1: The schematic of the proposed resource management framework with optimization, followed by deep learning and ensemble learning.

Therefore, the achievable SE for the eMBB user k, according to the Shannon-Hartely theorem, is given by $R_k\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right) = \sum_{s \in \mathcal{S}} \rho_k[s] \log_2(1 + \gamma_k[s]) \forall k \in \mathcal{K}$. Under the assumption of puncturing approach nevertheless, the total achievable SE for the eMBB users is partially lost due to coexistence of URLLC requests. For the typical eMBB user k, the lost rate is given by $R_k^{\text{Lost}}\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right) = \sum_{s \in \mathcal{S}} \rho_k[s] \log_2(1 + \gamma_k[s]) \sum_{u \in \mathcal{U}} \sum_{f \in \mathcal{F}} \mathbb{I}(\rho_k[s] = \tilde{\rho}_u^f[s]), \forall k \in \mathcal{K}$, wherein $\mathbb{I}(.)$ is the indicator function [4]. Therefore, the actual achievable SE of the eMBB user k is obtained as $R_k^{\text{Act}}\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right) = R_k\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right) - R_k^{\text{Lost}}\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right)$, which needs to satisfy a minimum threshold as

$$\textbf{C10:}\ R_k^{\text{Act}}\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right) \geq R_k^{\min}, \forall k \in \mathcal{K}. \quad (28)$$

Last but not least, the BS respects a power budget limitation constraint as follows:

$$\textbf{C11:} \sum_{s \in \mathcal{S}} \sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}} \sum_{u \in \mathcal{U}} \left(\rho_k[s] \boldsymbol{w}_k[s] p_k[s] + \tilde{\rho}_u^f[s] \tilde{\boldsymbol{w}}_u^f[s] \tilde{p}_u^f[s]\right) \leq P^{\max}. \quad (29)$$

F. Problem Statement

The binary constraints related to the RB allocation variables for the eMBB and URLLC users can be stated as follows:

$$\textbf{C12:} \quad \rho_k[s] \in \{0, 1\}, \forall k \in \mathcal{K}, s \in \mathcal{S}, \text{ and} \quad (30)$$

$$\textbf{C13:} \quad \tilde{\rho}_u^f[s] \in \{0, 1\}, \forall u \in \mathcal{U}, s \in \mathcal{S}, f \in \mathcal{F}. \quad (31)$$

We aim at maximizing the actual achievable SE of the eMBB users via optimizing the reflection coefficients in RIS elements $\boldsymbol{\Theta}$, the RB allocation $\bar{\boldsymbol{\rho}} = \{\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}\}$ and the BS transmit

power[2] $\boldsymbol{P} = \{\boldsymbol{p}\}$ as the decision variables by respecting the constraints $\textbf{C1} - \textbf{C13}$. Formally, this problem can be stated as (32) at the top of the next page.

In this problem, the objective function is non-convex with respect to $\boldsymbol{\Theta}$, $\bar{\boldsymbol{\rho}}$ and $\boldsymbol{P}$. The constraints $\textbf{C1}$ and $\textbf{C2}$ are linear with respect to $\boldsymbol{\Theta}$ and $\bar{\boldsymbol{\rho}}$, respectively. However, $\textbf{C3}$ and $\textbf{C4}$ are non-linear and non-deterministic, respectively. Moreover, $\textbf{C5}$, $\textbf{C6}$ and $\textbf{C7}$ are a linear function of $\tilde{\boldsymbol{\rho}}$ and $\bar{\boldsymbol{\rho}}$, but a non-linear function of $\boldsymbol{\Theta}$, and are non-convex constraints. With respect to $\bar{\boldsymbol{\rho}}$, the constraints $\textbf{C8}$ and $\textbf{C9}$ are linear, whereas $\textbf{C10}$ is a non-convex constraint with respect to $\boldsymbol{\Theta}$, $\bar{\boldsymbol{\rho}}$ and $\boldsymbol{P}$. Finally, the constraint $\textbf{C11}$ is linear with respect to $\bar{\boldsymbol{\rho}}$ and $\boldsymbol{P}$, while $\textbf{C12}$ and $\textbf{C13}$ induce non-convexity due to their integer form. All in all, the optimization problem (32) is non-convex due to its objective function and the constraints $\textbf{C3}$, $\textbf{C4}$, $\textbf{C5}$, $\textbf{C6}$, $\textbf{C7}$, $\textbf{C10}$, $\textbf{C12}$, and $\textbf{C13}$. In particular, the joint optimization problem (32) is in mixed integer and non-linear programming (MINLP) form and belongs to the NP-hard class.

G. Proposed Solution Methodology

We propose a supervised learning solution policy as depicted in Fig. 1 for the optimization problem (32) by leveraging optimization, deep learning, and ensemble learning. The computations for optimization and deep learning are carried out at the BS. First, an optimization policy is proposed on the basis of alternative decomposition method, so as to address the joint problem (32) in an iterative fashion[3]. Next, leveraging the optimized key system parameters, we train a LSTM by the Levenberg-Marquardt policy and fine-tune its learning rate hyperparameter via the Bayesian optimizer. In this way, the trained LSTM will be able to predict the key system parameters in future in real-time behaviour[4]. Finally, we integrate multiple LSTMs in an ensembling model that leads to a better accuracy of prediction.

III. OPTIMIZATION POLICY

In this section, the joint optimization problem (32) is sub-optimally solved via an alternative decomposition framework. By consecutively solving the sub-problems iteratively, the joint optimization problem (32) converges to a sub-optimal solution.

---

[2]The transmit power of the BS to the URLLC users, i.e., $\tilde{p}$, is set to its maximum value, denoted by $\tilde{p}^{max}$, where $\tilde{p}^{max} \leq P^{\max}$ following [4]. Note that owing to adopting the puncturing approach for multiplexing, there is no inter-service interference and hence, the maximum value of $\tilde{p}$ does not detrimentally contribute to the objective function in (32). Additionally, according to (10), the URLLC users will be better served, with more power budget $\tilde{p}$ and no URLLC intra-service interference. This in turn facilitates to meet the constraint $\textbf{C3}$.

[3]The time-series data is captured by recording the variations of the wireless key system parameters for prior channel realization of the wireless system. The results of the optimization phase are not valid and reliable, before the convergence is reached. Therefore, the time-series data for training the LSTM is obtained from the values at the convergence point of the optimization phase.

[4]Note that the neural network (i.e., the LSTM) is trained only one time in an offline manner. When the wireless network dynamics change, the well-trained LSTM is capable of predicting the key wireless system parameters $(P, \bar{\rho}, \theta)$ in an online manner, without retraining.

$$\max_{\{\boldsymbol{\Theta},\tilde{\boldsymbol{\rho}},\boldsymbol{P}\}} \sum_{k=1}^{K} R_k^{\text{Act}}\Big(\boldsymbol{P},\boldsymbol{\rho},\tilde{\boldsymbol{\rho}},\boldsymbol{\Theta}\Big) \tag{32a}$$

$$\text{s.t.} \quad \text{C1:} \ 0 \leq |\theta_m| \leq 1, \forall m, \tag{32b}$$

$$\text{C2:} \ \sum_{s \in \mathcal{S}} \sum_{f \in \mathcal{F}} \tilde{\rho}_u^f[s] \geq 1, \forall u \in \mathcal{U}, \tag{32c}$$

$$\text{C3:} \ Q\bigg[\sqrt{\frac{\Omega_u^f[s]}{\Gamma_u^f[s]}}\bigg(B_0 \log_2(1+\tilde{\gamma}_u^f[s]) - \frac{\tilde{R}_u^f[s]\ln 2}{\delta}\bigg)\bigg] \leq \bar{\varepsilon}_u^{\max}, \tag{32d}$$

$$\text{C4:} \ \Pr\Big(\sum_{u \in \mathcal{U}} \rho_u^f[s] \leq U\Big) \leq \tilde{\varepsilon}_u^{\max}, \forall f \in \mathcal{F}, s \in \mathcal{S}, \tag{32e}$$

$$\text{C5:} \ D_u^{\text{T}} + D_u^{\text{Q}} \leq D_u^{\max}, \forall u \in \mathcal{U}, \tag{32f}$$

$$\text{C6:} \ D_u^{\text{T}} \leq \frac{C_u}{\tilde{R}_u}, u \in \mathcal{U}, \tag{32g}$$

$$\text{C7:} \ D_u^{\text{Q}} \leq \frac{\ln(1/\tilde{\delta})}{(e^{\tilde{\theta}_u}-1)\tilde{R}_u}, \forall u \in \mathcal{U}, \tag{32h}$$

$$\text{C8:} \ \sum_{k \in \mathcal{K}} \rho_k[s] \leq N^{\max}, \forall s \in \mathcal{S}, \tag{32i}$$

$$\text{C9:} \ \sum_{s \in \mathcal{S}} \rho_k[s] \geq 1, \forall k \in \mathcal{K}, \tag{32j}$$

$$\text{C10:} \ R_k^{\text{Act}}\Big(\boldsymbol{P},\boldsymbol{\rho},\tilde{\boldsymbol{\rho}},\boldsymbol{\Theta}\Big) \geq R_k^{\min}, \forall k \in \mathcal{K}, \tag{32k}$$

$$\text{C11:} \sum_{s \in \mathcal{S}}\sum_{f \in \mathcal{F}}\sum_{k \in \mathcal{K}}\sum_{u \in \mathcal{U}} \Big(\rho_k[s]\boldsymbol{w}_k[s]p_k[s] + \tilde{\rho}_u^f[s]\tilde{\boldsymbol{w}}_u^f[s]\tilde{p}_u^f[s]\Big) \leq P^{\max}, \tag{32l}$$

$$\text{C12:} \ \rho_k[s] \in \{0,1\}, \forall k \in \mathcal{K}, s \in \mathcal{S}, \tag{32m}$$

$$\text{C13:} \ \tilde{\rho}_u^f[s] \in \{0,1\}, \forall k \in \mathcal{K}, s \in \mathcal{S}, f \in \mathcal{F}. \tag{32n}$$

A. Power Control Sub-problem

Given the decision variables $\boldsymbol{\Theta}^{(t-1)}$, $\boldsymbol{\rho}^{(t-1)}$ and $\tilde{\boldsymbol{\rho}}^{(t-1)}$ in (32) at $(t-1)$th iteration, we now optimize the transmit power of the BS at tth iteration, i.e., $\boldsymbol{P}^{(t)}$. The BS transmit power control sub-problem for the eMBB users can thus be given as follows:

$$\max_{\{\boldsymbol{P}^{(t)}\}} \sum_{k=1}^{K} R_k^{\text{Act}}\Big(\boldsymbol{P}^{(t)},\boldsymbol{\rho}^{(t-1)},\tilde{\boldsymbol{\rho}}^{(t-1)},\boldsymbol{\Theta}^{(t-1)}\Big) \tag{33a}$$

$$\text{s.t.} \ \text{C10:} \ R_k^{\text{Act}}\Big(\boldsymbol{P}^{(t)},\boldsymbol{\rho}^{(t-1)},\tilde{\boldsymbol{\rho}}^{(t-1)},\boldsymbol{\Theta}^{(t-1)}\Big) \geq R_k^{\min}, \forall k \in \mathcal{K}, \tag{33b}$$

$$\text{C11:} \sum_{s \in \mathcal{S}}\sum_{f \in \mathcal{F}}\sum_{k \in \mathcal{K}}\sum_{u \in \mathcal{U}} \Big(\rho_k[s]\boldsymbol{w}_k[s]p_k[s] + \tilde{\rho}_u^f[s]\tilde{\boldsymbol{w}}_u^f[s]\tilde{p}_u^f[s]\Big) \leq P^{\max}, \tag{33c}$$

which is still non-convex due to the objective function and the constraint **C10**. To solve this sub-problem, let us put (27) in **C10**, which results in

$$(1-\Phi_k)\Big|\big(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\big)\boldsymbol{w}_k[s]\Big|^2 \rho_k[s]p_k[s] \tag{34}$$

$$- \eta \sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) < \Upsilon(i)}} \Big|\big([\boldsymbol{h}_{d,k}[s]]^H + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\big)\boldsymbol{w}_k[s]\Big|^2 \rho_i[s]p_i[s] \geq \omega,$$

with $\eta = 2^{R_k^{\min}/\varpi} - 1$, $\varpi = 1 - \sum_{f \in \mathcal{F}} \mathbb{I}(\rho_k[s] = \tilde{\rho}_u^f[s])$, $\omega = \sigma^2\eta$ and Φ_k is the portion of the lost rate for the eMBB user k, discussed in the **Lemma 1** later. In this sense, the non-linear constraint **C10** can be converted into a linear form and the sub-problem (33) can be reformulated as

$$\max_{\{\boldsymbol{P}^{(t)}\}} \sum_{k=1}^{K} R_k^{\text{Act}}\Big(\boldsymbol{P}^{(t)},\boldsymbol{\rho}^{(t-1)},\tilde{\boldsymbol{\rho}}^{(t-1)},\boldsymbol{\Theta}^{(t-1)}\Big) \tag{35a}$$

$$\text{s.t.} \ \text{C10:} \ (1-\Phi_k)\Big|\big(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\big)\boldsymbol{w}_k[s]\Big|^2 \rho_k[s]p_k[s]$$

$$- \eta \sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) < \Upsilon(i)}} \Big|\big([\boldsymbol{h}_{d,k}[s]]^H + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\big)\boldsymbol{w}_k[s]\Big|^2 \rho_i[s]p_i[s]$$

$$\geq \omega, \forall k \in \mathcal{K}, s \in \mathcal{S}, \tag{35b}$$

$$\text{C11:} \sum_{s \in \mathcal{S}}\sum_{f \in \mathcal{F}}\sum_{k \in \mathcal{K}}\sum_{u \in \mathcal{U}} \Big(\rho_k[s]\boldsymbol{w}_k[s]p_k[s] + \tilde{\rho}_u^f[s]\tilde{\boldsymbol{w}}_u^f[s]\tilde{p}_u^f[s]\Big) \leq P^{\max}, \tag{35c}$$

which is non-convex due to its objective function. By using the extension of Sherman-Morrison-Woodbury formula [29] given as

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(I + CDA^{-1}B)^{-1}CDA^{-1}, \tag{36}$$

we have

$$(1 + \gamma_k[s])^{-1} = 1 - \left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2 p_k[s]$$

$$\left(\left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2 p_k[s] + \xi_k[s] \right)^{-1}, \quad (37)$$

where

$$\xi_k[s] = \sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) < \Upsilon(i)}} \left| \left([\boldsymbol{h}_{d,k}[s]]^H + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2 \rho_i[s] p_i[s] + \sigma^{.2}$$
$$(38)$$

Since (37) is in a MMSE form, we can formulate it as an MMSE detection problem to find $\mathbf{x}_k[s]$ from $\mathbf{y}_k[s]$ as follows:

$$c_k^o[s] = \arg\min_{c_k[s]} e_k[s], \quad (39)$$

where $e_k[s]$ is the mean square error (MSE) and $c_k^o[s]$ is the optimal value of $c_k[s]$ while minimizing the MSE. Besides, the MSE can be stated as $e_k[s] = \mathbb{E}\{|\mathbf{x}_k[s] - c_k[s]y_k[s]|^2\}$, where $c_k[s]$ is the channel equalization coefficient. Obviously, based on MSE, the optimal value of $c_k[s]$ (i.e., $c_k^o[s]$) is obtained such that the detected signal is as close as possible to the transmitted signal.

Lemma 1. *Let us represent by $0 \leq \Phi_k < 1$, the portion of lost rate for the kth eMBB user[5]. The actual achievable rate for the eMBB user k can be approximated as*

$$R_k^{Act} = \max_{c_k[s], a_k[s] > 0} (1 - \Phi_k) \left(-\frac{a_k[s]e_k[s]}{\ln 2} + \log_2 a_k[s] + \frac{1}{\ln 2} \right). \quad (40)$$

Proof. See **Appendix**. □

Using **Lemma 1**, the objective function in (35) turns into a quadratic programming function. Thus, (35) is rewritten as

$$\max_{\boldsymbol{P}^{(t)}} \sum_{s=1}^{S} \sum_{k=1}^{K} \max_{c_k[s], a_k[s] > 0} (1 - \Phi_k) \left(-\frac{a_k[s]e_k[s]}{\ln 2} + \log_2 a_k[s] + \frac{1}{\ln 2} \right)$$
$$\text{s.t.} \quad \mathbf{C10}, \mathbf{C11}. \quad (41a)$$

Now, by optimizing $\{c_k[s]\}$, $\{a_k[s]\}$ and $\{p_k[s]\}$, the problem in (41) can be addressed in an iterative fashion. In particular, given the optimal power control solution $\{p_k^{(t-1)}[s]\}$ at $(t-1)$th iteration, the optimal solution of $\{c_k^{(t)}[s]\}$ (according to **Appendix A**), at the tth iteration, can be obtained as:

$$c_k^{(t)}[s] = \left(\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k^{(t-1)}[s]} \right)^*$$

$$\left(\left| (\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s] \right|^2 \rho_k[s] p_k^{(t-1)}[s] + \xi_k^{(t-1)}[s] \right)^{-1},$$

where

$$\xi_k^{(t-1)}[s] =$$
$$\sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) < \Upsilon(i)}} \left| \left([\boldsymbol{h}_{d,k}[s]]^H + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2 \rho_i[s] p_i^{(t-1)}[s] + \sigma^2.$$

[5]Note that $\Phi_k = 1$ implies that the actual rate of the eMBB user k is sacrificed for serving the URLLC services, which makes the system infeasible according to **C10**.

The corresponding MMSE, defined in **Appendix A**, at the tth iteration, is given by

$$e_k^{o(t)}[s] = \frac{1 - \rho_k[s]p_k^{(t-1)}[s] \left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2}{\left(\left| (\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s] \right|^2 \rho_k[s]p_k^{(t-1)}[s] + \xi_k^{(t-1)}[s] \right)}. \quad (42)$$

Accordingly, $\left\{ a_k^{(t)}[s] \right\}$ at the tth iteration is optimized as $a_k^{(t)}[s] = \frac{1}{e_k^{o(t)}[s]}$. Once we have optimized $\left\{ c_k^{(t)}[s] \right\}$ and $\left\{ a_k^{(t)}[s] \right\}$ at the tth iteration, the sub-problem (41) is reformulated as follows:

$$\min_{\{\boldsymbol{P}^{(t)}\}} \sum_{s=1}^{S} \sum_{k=1}^{K} (1 - \Phi_k)a_k[s]e_k[s] \quad (43a)$$
$$\text{s.t.} \quad \mathbf{C10}, \mathbf{C11}, \quad (43b)$$

where

$$e_k^{(t)}[s] = \left| 1 - c_k[s] \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k^{(t)}[s]} \right|^2$$
$$+ |c_k[s]|^2 \left\| (\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s] \right\|_2^2$$
$$\sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) < \Upsilon(i)}} \rho_i[s]p_i^{(t)}[s] + \left| c_k^{(t)}[s] \right|^2 \sigma^2. \quad (44)$$

Lemma 2. *The sub-problem* (43) *is convex, the Karush-Kuhn-Tucker (KKT) conditions hold and the optimal BS transmit power can be acquired as*

$$p_k^{(t)}[s] = \left(\frac{a_{k,s}^{(t)}\Re\left(c_k^{(t)}[s] \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right)}{\tau} \right), \quad (45)$$

where

$$\tau = \sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) > \Upsilon(i)}} a_k^{(t)}[s] \left| c_k^{(t)}[s] \right| \left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2$$
$$+ \lambda - \mu_k^{(t)}[s] \left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2$$
$$+ \sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) > \Upsilon(i)}} \mu_{u,n}\eta \left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s] \right) \boldsymbol{w}_k[s] \right|^2. \quad (46)$$

Proof. See Appendix. □

B. RB Allocation Sub-problem

So far, the transmit power of the BS at time step t, i.e., $\boldsymbol{P}^{(t)}$ has been optimized. Given $\boldsymbol{\Theta}^{(t-1)}$ from the $(t-1)$th time step, we now aim at optimizing $\bar{\rho}^{(t)}$ at time step t. Hence, the RB allocation sub-problem is written as

$$\max_{\{\bar{\rho}^{(t)}\}} \sum_{k=1}^{K} R_k^{\text{Act}} \left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t-1)} \right) \quad (47a)$$
$$\text{s.t.} \quad \mathbf{C2} - \mathbf{C13}, \quad (47b)$$

which is non-convex due to **C3**, the probabilistic constraint **C4**, as well as the discrete-domain decision variables. In **C3**, $\varepsilon_u^f[s]$ can be approximately represented as [3] in which

$$\varphi_u^f[s] = 2^{\tilde{R}_u^f[s]/B_0} - 1, \quad \kappa_u^f[s] = \left[2\pi\sqrt{2^{2\tilde{R}_u^f[s]/B_0} - 1} \right]^{-1},$$

$$\varepsilon_u^f[s] = Q\left[\frac{\log_2(1+\tilde{\gamma}_u^f[s]) + \tilde{R}_u^f[s]/B_0}{\sqrt{\Gamma_u^f[s](\log_2 e)^2/\Omega_u^f[s]}}\right] = \begin{cases} 1, & \tilde{\gamma}_u^f[s] \leq \chi_u^f[s] \\ \frac{1}{2} - \kappa_u^f[s]\sqrt{\Omega_u^f[s]}(\tilde{\gamma}_u^f[s] - \varphi_u^f[s]), & \chi_u^f[s] \leq \tilde{\gamma}_u^f[s] \leq \xi_u^f[s] \\ 0, & \xi_u^f[s] \leq \tilde{\gamma}_u^f[s] \end{cases}, \quad (48)$$

$\chi_u^f[s] = \varphi_u^f[s] - \left[2\kappa_u^f[s]\sqrt{\Omega_u^f[s]}\right]^{-1}$ and $\xi_u^f[s] = \varphi_u^f[s] + \left[2\kappa_u^f[s]\sqrt{\Omega_u^f[s]}\right]^{-1}$. The non-deterministic constraint **C4** can be efficiently transformed into the convex form [4] as $\sum_{u\in\mathcal{U}} \tilde{\rho}_u^f[s] - F_U^{-1}(1-\tilde{\varepsilon}_u^{\max}) \geq 0, \forall f \in \mathcal{F}, s \in \mathcal{S}$, where F_U is the cumulative distribution function (CDF) of the random variable $U \sim \mathcal{N}(\bar{\mu}, \bar{\sigma}^2)$. By leveraging a time sharing (relaxation) method, the discrete-domain decision variables in $\bar{\rho}$ can be transformed into the continuous-domain form for the ease of solution, expressed as **C12** : $\rho_k[s] \in [0,1], \forall k \in \mathcal{K}, s \in \mathcal{S}$, and **C13** : $\tilde{\rho}_u^f[s] \in [0,1], \forall u \in \mathcal{U}, f \in \mathcal{F}, s \in \mathcal{S}$. However, since the RB allocation is inherently a binary problem, we apply the binary-forcing constraints [31] to the sub-problem (47) to guarantee that the optimized continuous-domain variables are either 0 or 1, as $\sum_{k=1}^{K}\sum_{s=1}^{S}\left[(\rho_k[s]) - (\rho_k[s])^2\right] \leq 0$, and $\sum_{u=1}^{U}\sum_{s=1}^{S}\sum_{f=1}^{F}\left[(\tilde{\rho}_u^f[s]) - (\tilde{\rho}_u^f[s])^2\right] \leq 0$, which incorporate a feasible region over the corner critical points, i.e., $\{0,1\}$ [32]. Due to coexistence of two decision variables, i.e., $\rho_k[s]$ and $\tilde{\rho}_u^f[s]$, the sub-problem (47) is indeed a coupled (joint) problem. Hence, a decomposition process will be used to decouple (47) into two distinct sub-problems over $\rho_k[s]$ and $\tilde{\rho}_u^f[s]$. Finally, by applying the penalty successive upper bound minimization (PSUM) [4], both the sub-problems can be effectively transformed into the convex form, which is straightforward to solve through the existing convex optimization packages such as CVX [33].

C. Sub-problem for RIS Reflection Coefficient Design

Finally, given the optimized $\bar{\rho}^{(t)}$ and $\boldsymbol{P}^{(t)}$ at the tth iteration, the RIS reflection coefficient design sub-problem can be formulated as follows:

$$\max_{\{\boldsymbol{\Theta}^{(t)}\}} \sum_{k=1}^{K} R_k^{\text{Act}}\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t)}\right) \quad (49a)$$

$$\text{s.t.} \quad \mathbf{C1}, \mathbf{C3}, \mathbf{C6}, \mathbf{C7}, \mathbf{C10}. \quad (49b)$$

Due to the non-convex objective function and constraints, this sub-problem is non-convex. We follow the same procedure applied to the power control sub-problem henceforth. Once we have addressed the distinct (decoupled) sub-problems (33), (47), and (49) iteratively, the joint (coupled) optimization problem (32) is in fact sub-optimally solved. The convergence of the proposed iterative solution approach is proved in **Lemma 3**.

Lemma 3. *The proposed alternative decomposition-based procedure converges to a sub-optimal solution of the joint optimization problem in (32).*

Proof. See **Appendix**. □

To reduce the time complexity of the proposed solution, in what follows, we propose a low-complexity data-driven deep-learning technique that traces the optimized key system parameters thus far and anticipates them in future.

IV. DEEP LEARNING AND ENSEMBLE LEARNING POLICY

In this section, we consider T discrete time steps for capturing the variations in wireless network. In each time step, we form a historical time-series sequence of the optimized key system parameters $\boldsymbol{P}, \bar{\rho}$ and $\boldsymbol{\Theta}$, obtained from the previous section to be predicted at upcoming time steps.

A. Time-series Forecasting

Advantageous of temporal dynamic behavior and time-varying nature, artificial neural networks (ANNs) can accurately process the time-series data (i.e., sequences of data at subsequent time steps). Relaying on this capability, the time-series ANNs are widely employed for tracking the historical information and anticipating their future trend. Due to the dynamic variations in system parameters such as the BS transmit power, they can be modeled in time-series form and efficiently predicted at upcoming time-steps in a real-time manner [32]. Toward this goal, as a powerful time-series ANN, LSTM is considered in this paper.

B. LSTM

A LSTM network includes multiple LSTM units, each of which includes an input gate, a forget gate, a memory cell, a hidden gate, and an output gate. The input gate at state t^6 on one hand, is fed through the historical sequences $\Xi(1), \Xi(2), ..., \Xi(t-1)$, with $\boldsymbol{\Xi} = \{\boldsymbol{P}, \bar{\rho}, \boldsymbol{\Theta}\}$, whereas the output gate on the other hand anticipates $\widehat{\boldsymbol{\Xi}}(t) = \{\widehat{\boldsymbol{P}}(t), \widehat{\bar{\rho}}(t), \widehat{\boldsymbol{\Theta}}(t)\}$. Let us denote by $(\mathcal{I}np, W_{\text{Inp}}^{\Xi}, b_{\text{Inp}})$, $(\mathcal{F}or, W_{\text{For}}^{\Xi}, b_{\text{For}})$, $(\mathcal{C}ell, W_{\text{Cell}}^{\Xi}, b_{\text{Cell}})$, $(\mathcal{H}idd, W_{\varrho}^{\text{Hidd}})$ and $(\mathcal{O}ut, W_{\text{Out}}^{\Xi}, b_{\text{Out}})$, the input gate, the forget gate, the memory cell gate, the hidden gate and the output gate accompanied by their corresponding weights and biases, respectively, with ϱ representing the set of LSTM gates, i.e., $\varrho = \{\mathcal{I}np, \mathcal{F}or, \mathcal{C}ell, \mathcal{O}ut\}$. The mathematical representation of the LSTM structure by using the target system parameters $\Xi(t)$, is expressed as:

$$\mathcal{I}np(t) = f_{\text{sig}}^{\text{act}}\bigg(W_{\text{Inp}}^{\Xi}\, \Xi(t) + W_{\text{Inp}}^{\text{Hidd}}\, \mathcal{H}idd(t-1)$$
$$+ W_{\text{Inp}}^{\text{Cell}}\, \mathcal{C}ell(t-1) + b_{\text{Inp}}\bigg), \quad (50)$$

[6]In order to avoid index conflict, it is emphasized that each state of the deep learning phase corresponds to the convergence point of the optimization phase.

$$\mathcal{F}or(t) = f_{\text{sig}}^{\text{act}}\bigg(W_{\text{For}}^{\Xi}\, \Xi(t) + W_{\text{For}}^{\text{Hidd}}\, \mathcal{H}idd(t-1)$$
$$+ W_{\text{For}}^{\text{Cell}}\, \mathcal{C}ell(t-1) + b_{\text{For}} \bigg), \tag{51}$$

$$\mathcal{C}ell(t) = \mathcal{F}or(t) \odot \mathcal{C}ell(t-1) + \mathcal{I}np(t) \odot f_{\text{hyp}}^{\text{act}}\bigg(W_{\text{Cell}}^{\Xi}\, \Xi(t)$$
$$+ W_{\text{Cell}}^{\text{Hidd}}\, \mathcal{H}idd(t-1) + b_{\text{Cell}} \bigg), \tag{52}$$

$$\mathcal{O}ut(t) = f_{\text{sig}}^{\text{act}}\bigg(W_{\text{Out}}^{\Xi}\, \Xi(t) + W_{\text{Out}}^{\text{Hidd}}\, \mathcal{H}idd(t-1)$$
$$+ W_{\text{Out}}^{\text{Cell}}\, \mathcal{C}ell(t) + b_{\text{Out}} \bigg), \tag{53}$$

and

$$\mathcal{H}idd(t) = \mathcal{O}ut(t) \odot f_{\text{peep}}^{\text{act}}\bigg(\mathcal{C}ell(t) \bigg), \tag{54}$$

where $\odot$ denotes the element-wise multiplication and $f_{\text{sig}}^{\text{act}}, f_{\text{hyp}}^{\text{act}}, f_{\text{peep}}^{\text{act}}$ are the Sigmoid, hyperbolic tangent, and peephole activation functions, respectively. Besides, in (54), $\mathcal{H}idd(t)$ indicates the hidden vector at time step t and updated as $\mathcal{H}idd(t) = \zeta(\mathcal{H}idd(t-1), \Xi(t), \vartheta)$, wherein ϑ is the set of LSTM parameters (i.e., weights and biases), while ζ is a combination of Eqs. (50)-(54). Therefore, the loss function[7] for training the LSTM can be formulated by $f^{\text{Loss}} = \sqrt{\sum_{i=1}^{M} ||\Xi^i(t) - \widehat{\Xi^i}(t)||_2^2}$, with $\widehat{\Xi}(t)$ and $\Xi(t)$, representing the LSTM-forecasted key system parameters and the target key system parameters, both at time step t, respectively. Among the key system parameters, $\boldsymbol{P}, \boldsymbol{\Theta}$ are continuous, whereas $\bar{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{\rho}}$ are discrete. Hence, the LSTM output gate for anticipating $\widehat{\boldsymbol{P}}(t)$ and $\widehat{\boldsymbol{\Theta}}(t)$ employs a regression layer, while a classification layer such as a Softmax is employed for anticipating $\widehat{\boldsymbol{\rho}}(t)$[8]. In the following subsection, we explain how the LSTM is trained by optimizing its weights and biases.

C. Levenberg-Marquardt Policy

By invoking Levenrage-Marquardt policy, we optimize the LSTM parameters (i.e., weights and biases), indicated by $\iota_{\text{LSTM}} = \{(W_{\text{Inp}}^{\Xi}, b_{\text{Inp}}), (W_{\text{For}}^{\Xi}, b_{\text{For}}), (W_{\text{Cell}}^{\Xi}, b_{\text{Cell}}), (W_{\text{Out}}^{\Xi}, b_{\text{Out}}), (W_{\varrho}^{\text{Hidd}})\}$. Due to interpolation between the Gauss-Newton and the gradient descent method, this policy minimizes the squared errors in non-linear problems by means of a generic curve-fitting. By properly initializing in a trust region, the policy is able to approach the LSTM parameters' local minima. The training process aims at optimizing the LSTM parameters indicated by ι_{LSTM}^* and obtaining a curve $\wp\left(\hat{\Xi}(t), \iota_{\text{LSTM}}\right)$, so as to minimize the total deviation from

the target system parameters $\Xi(t)$. In other words, the optimal LSTM parameters are obtained as:

$$\iota_{\text{LSTM}}^* = \arg\min_{\iota_{\text{LSTM}}} Dev(\iota_{\text{LSTM}}) \tag{55}$$
$$= \arg\min_{\iota_{\text{LSTM}}} \sum_{i=1}^{\|\Xi\|} \left[\Xi^i(t) - \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}}\right) \right]^2 .$$

Since it is complicated to achieve ι_{LSTM}^* according to (55), an approximation to the curve can be taken into consideration for the sake of simplicity as follows:

$$\wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right) = \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}}\right) + \text{æ}\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}}, \tag{56}$$

with shifting element æ. Regarding this approximation, the sum of squared deviations in (55) is given by

$$Dev(\iota_{\text{LSTM}}) = \tag{57}$$
$$\sum_{i=1}^{\|\Xi\|} \left[\Xi^i(t) - \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}}\right) - \text{æ}\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right]^2 .$$

Leveraging the Gauss-Newton method, the shifting element can be optimized by letting $\text{æ}^* = \frac{\partial Dev(\iota_{\text{LSTM}})}{\partial \text{æ}} = 0$, calculated as follows:

$$\left[\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right]^T \left[\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right] \text{æ}^* = \tag{58}$$
$$\left[\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right]^T \sum_{i=1}^{\|\Xi\|} \left[\Xi^i(t) - \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}}\right) \right]^2 .$$

By introducing a damped version of (58), the Levenberg-Marquardt policy [17] iteratively achieves the optimal shifting element as follows:

$$\left(\left[\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right]^T \left[\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right] + \text{Æ}\mathbf{I} \right) \text{æ}^* \tag{59}$$
$$= \left[\frac{\partial \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}} + \text{æ}\right)}{\partial \iota_{\text{LSTM}}} \right]^T \sum_{i=1}^{\|\Xi\|} \left[\Xi^i(t) - \wp\left(\hat{\Xi}^i(t), \iota_{\text{LSTM}}\right) \right]^2 ,$$

where Æ indicates the damping factor. Since the objective in training process as mentioned earlier, is to minimize the total deviation from the target system parameters $\Xi(t)$, the Levenberg-Marquardt policy for smaller Æ approaches the Gauss-Newton method, while this policy tends to the gradient descent method for larger values of Æs. Eventually, for a limited squared deviation within a given error range, this policy returns ι_{LSTM}^*. However, the performance of training process can be improved by a proper initialization of the LSTM parameters. This is elaborated in the following subsection.

D. Bayesian Optimizer

In order to further optimize the training behavior of the LSTM, we propose to initialize its hyperparameters before the training process starts [20]. One of the most important hyperparameters in training the LSTM with respect to its memory-based structure, is the learning rate that can be efficiently optimized by minimizing its total degree of error per training

---

[7]Consider the root mean square error (RMSE) for instance as the loss function.

[8]In fact, due to the existence of four variables with different dimensions and domains (i.e., continuous and integer) in Ξ, we train four LSTMs independently, each on a separate variable in Ξ.

sample. Let us denote by $œ_l^i$, the degree of error associated with the ith sample and the lth neuron in the final layer of the LSTM. Therefore, the total degree of error for the ith sample can be given by $œ^i = \frac{1}{2} \sum_{l \in \mathcal{L}_{N_{\text{Lay}}}} \left[\Xi^i(t) - \hat{\Xi}^i(t) \right]_l$, where $\mathcal{L}_{N_{\text{Lay}}}$ is the neuron index set in the last layer of the LSTM. During the LSTM training process, the LSTM parameters (i.e., its weights and biases) related to the lth neuron and the ith sample are updated as follows:

$$\Delta \dddot{\lambda}_{\text{LSTM}^i} = \dddot{\lambda}_{\text{LSTM}^i}(t) - \dddot{\lambda}_{\text{LSTM}^i}(t-1)$$
$$= -\eta_{\text{LSTM}} \frac{\partial œ^i}{\partial \zeta_{\text{LSTM}_l^i}} \left[\Xi^i(t) \right]_l, \quad (60)$$

where η_{LSTM} is the LSTM learning rate and $\zeta_{\text{LSTM}_l^i}$ represents the induced local field corresponding to the ith sample and lth neuron in the LSTM final layer. In Bayesian optimizer, the previous states of the hyperparameter play an important role in updating to its current state. This optimizer requires considerably fewer number of iterations for seeking the optimal initial value of the hyperparameter in comparison with the grid search counterpart. Therefore, by using the Bayesian optimizer, it is anticipated for the LSTM learning process to converge quickly. This optimizer assumes a probability distribution for η_{LSTM} and evaluates its prior states through an assessment function. This function minimizes the training error of the LSTM, which is a regression operation. We introduce the assessment function at state t as: $f^{\text{Loss}}(\eta_{\text{LSTM}}(t-1), \eta_{\text{LSTM}}(t-2), ..., \eta_{\text{LSTM}}(1)) = \tilde{\eta}_{\text{LSTM}}$, that can be considered as a temporal benchmarking, while analyzing the next state of η_{LSTM}. Thus, we can obtain the optimal initial η_{LSTM}, through minimizing the loss function at state t as follows:

$$\eta_{\text{LSTM}}(t) = \arg \min_{\eta_{\text{LSTM}}} f^{\text{Loss}}(\eta_{\text{LSTM}}(t-1), ..., \eta_{\text{LSTM}}(1)), \quad (61)$$

meaning that the previous states should be analyzed one by one towards minimizing the loss function. For real-time prediction, the duration of states should be short. Within a large space of previous states for the hyperparameter η_{LSTM}, this assessment can incur a significantly high computational cost. To facilitate the computations, a surrogate approximator for f^{Loss} is commonly used regarding the prior states of η_{LSTM}. Using the Bayes theorem [20], the surrogate approximator function as a probability distribution at instant t is expressed as:

$$\Pr(\tilde{\eta}_{\text{LSTM}} | \eta_{\text{LSTM}}(t)) = \frac{\Pr(\eta_{\text{LSTM}}(t) | \tilde{\eta}_{\text{LSTM}}) \Pr(\tilde{\eta}_{\text{LSTM}})}{\Pr(\eta_{\text{LSTM}}(t))}, \quad (62)$$

which is achieved through the tree parzen estimator (TPE) [34]. In fact, TPE is a probabilistic evaluator on f^{Loss} to get $\tilde{\eta}_{\text{LSTM}}$ from $\eta_{\text{LSTM}}(t)$ in a computationally efficient way. This is commonly done by expected improvement (EI) measure on the basis of "exploration and exploitation" trade-off rule [35]. After all, by optimizing the LSTM learning rate $\eta_{\text{LSTM}}(t)$, at instant t, the LSTM parameters can be efficiently updated in (60).

E. Ensemble Learning Policy

Ensemble learning is an efficient technique to develop a strong learner by using multiple weak learners[9]. The strong learner predicts the future of the key system parameters as a combination of the predictions made by its incorporating weak learners. As an efficient method for realizing ensemble learning, bootstrap aggregation is popular for various applications [21]. According to Fig. 2, the training samples are categorized and bootstrapped in replicas (subsets), where each of the weak learners adopts a distinct subset for training. The weak learners are trained in an independent and simultaneous manner. The predictions from weak learners are integrated in a combination phase, over which a regression/classification is applied, forming the prediction of the strong learner. The predicted key system parameters in $\hat{\Xi}^{(t)}$ made by each LSTM as a weak learner are categorized either into discrete-domain (i.e., $\hat{\boldsymbol{\rho}}(t)$) or continuous-domain (i.e., $\hat{\boldsymbol{P}}(t)$ and $\hat{\boldsymbol{\Theta}}(t)$). Depending on which domain the key system parameters belong to (i.e., continuous or discrete), the final prediction made by the strong learner is calculated according to either of the following principles:

- *Regression*: For continuous parameters $\hat{\boldsymbol{P}}(t)$ and $\hat{\boldsymbol{\Theta}}(t)$, the bootstrap aggregation (or so-called bagging algorithm) involves a regression problem. Suppose that we have M_1 trained weak learners (i.e., LSTM regressors) by M_1 randomly subsets, obtained from the alternative decomposition technique. As observable from Fig. 2, any of the trained LSTMs as a regressor is capable of forecasting the next time step for the key system parameters independent of others. A strong learner afterwards is responsible for aggregating the LSTM predictions, upon which a universal regression is performed to obtain the strong learner prediction. For this, the bagging mechanism [21] exerts a weighted averaging policy over the predictions, whereby the LSTM regressors with a higher prediction accuracy[10] contribute more to the final prediction made by the strong learner.

- *Classification*: The future prediction of the discrete parameters $\hat{\boldsymbol{\rho}}(t)$, induces a classification problem for the bagging mechanism. As before, as a weak learner, each of the M_1 LSTM classifiers is trained beforehand on a distinct subset of the optimized system parameters. The trained LSTM classifiers (in Fig. 2) are responsible for predicting $\hat{\boldsymbol{\rho}}(t)$ simultaneously and independently. A strong learner accordingly aggregates the predictions made by the LSTM classifiers so as to make a universal classification as the final prediction. With respect to the binary nature of the RB allocation, the predicted $\hat{\boldsymbol{\rho}}(t)$ by each LSTM in fact belongs to either of the two binary classes $\{0, 1\}$. In this regard, the well-known maximum voting approach is applied over the predictions, whereby a voting counter specifies the total vote for each class (i.e., the total number of LSTM classifiers, adopted that

[9]Note that the "weak learner" in this paper is a general predictor (e.g., LSTM).

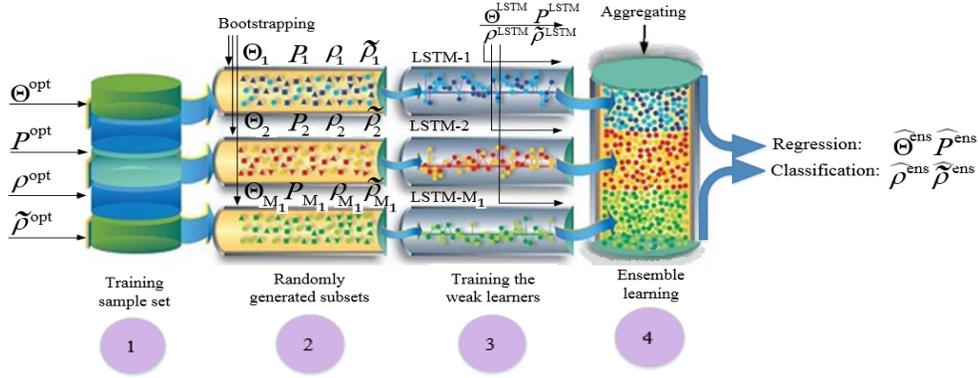[10]The accuracy of weak learners is assessed beforehand through validation and testing over the existing samples.

Fig. 2: Ensemble learning schematic.

Algorithm 1 Alternative Decomposition Policy

Initialization: Initialize all the parameters, such that the system feasibility holds (i.e., the constraints **C1** − **C13** are initially satisfied).

1: **repeat**
2:     Solve the power control sub-problem (33) as follows.
3:     **repeat**
4:         $t=1$.
5:         Obtain the optimal $a_k^{(t)}[s]$.
6:         Obtain the optimal $c_k^{(t)}[s]$.
7:         Obtain the optimal $p_k^{(t)}[s]$.
8:         $t=t+1$.
9:     **until** $t \leq T_{\max}$
10:     Solve the RIS reflection coefficient design sub-problem (49) accordingly.
11:     Solve the RB allocation sub-problem (47) according to [4].
12:     Calculate $\sum_{k=1}^{K} R_k^{\text{Act}(t)}$.
13:     **if** $\left| \sum_{k=1}^{K} R_k^{\text{Act}(t)} - \sum_{k=1}^{K} R_k^{\text{Act}(t-1)} \right| < \tilde{\varepsilon}$ **then**
14:         final convergence = **true**.
15:     **else**
16:         $T = T + 1$.
17:     **end if**
18: **until** final convergence

---

class as their prediction). Eventually, the class with more voting counter is selected as the prediction of the strong learner.

V. COMPLEXITY ANALYSIS

The computational complexity of the power control and sub-problem depends on λ and μ, which is linear to the number of eMBB users, i.e., $\mathcal{O}(K)$. Note that, the KKT dual variables λ and μ can be obtained via a bisection method. This makes the total complexity of the power control sub-problem of order $\mathcal{O}_{\text{PC}}(T_{\max}K^2\log_2(\psi_{\text{PC}}))$, where $T_{\max}$ and ψ_{PC} are the number of iterations required to converge and the required accuracy of the convergence in power control, respectively. Since the RIS reflection coefficient design sub-problem follows a similar solution procedure, it has the computational complexity of

$\mathcal{O}_{\text{RPSD}}(T_{\max}K^2\log_2(\psi_{\text{RPSD}}))$, with ψ_{RPSD} being the required accuracy of the convergence. The RB allocation sub-problem, however, is solved according to the PSUM method presented in [4]. This method follows a penalty-based approach, wherein the constraints as penalty terms are incorporated into the objective function and the problem is solved by determining a sequence of approximation of the objective functions. To this end, the CVX is applied with the computational complexity of

$$\mathcal{O}_{\text{PC}}\left(\frac{\log_2(N_{\text{Cons}})/q_0\psi_{\text{acc}}}{\log_2(\psi_{\text{step}})} \right), \tag{63}$$

where N_{Cons} represents the total number of constraints, $0 \leq \psi_{\text{acc}} \leq 1$ denotes the desired accuracy level of the interior point method for convergence with the initial point q_0, and ψ_{step} denots the gradient step size. The computational complexity of training a LSTM is [16]:

$$\mathcal{O}_{\text{LSTM}}\left(\left((N_{\text{node}}^{\max})^2 + \mathcal{O}(f_{\text{sig/hyp/peep}}^{\text{act}}) \right) N_{\text{LSTM}}^{\text{hor}} N_{\text{LSTM}}^{\text{ver}} \right), \tag{64}$$

where $N_{\text{node}}^{\max}$ denotes the maximum number of neurons within the LSTM layers, $\mathcal{O}(f_{\text{sig/hyp/peep}}^{\text{act}})$ is the highest computational complexity among the activation functions and $N_{\text{LSTM}}^{\text{hor}}(N_{\text{LSTM}}^{\text{ver}})$ indicates the number of LSTM units in horizontal (vertical) rows. Finally, the computational complexity of the ensembling network is given by

$$\mathcal{O}_{\text{Ens}}\left(\mathcal{O}_{\text{LSTM}_1} + \mathcal{O}_{\text{LSTM}_2} + ... + \mathcal{O}_{\text{LSTM}_{M_1}} \right)$$
$$= \max\{\mathcal{O}_{\text{LSTM}_j}\}, \; \forall j \in \{1, 2, ..., M_1\}. \tag{65}$$

VI. SIMULATION RESULTS

A. Simulation Parameters

We consider a THz BS equipped with $N_t = 256$ antennas with a coverage radius of 200 m. We consider the same number of eMBB and URLLC users (i.e., $K = U = 5$ users) and the users are randomly distributed within the cell coverage area. The BS's power budget is $P^{\max}=15$ dBm. The maximum number of multiplexed eMBB users over a specific RB is constrained to $N^{\max} = 3$. The duration of a time slot is set to 1 ms and each time slot is further divided into 7 equi-spaced mini-slots. A zero-forcing (ZF) beamforming is also
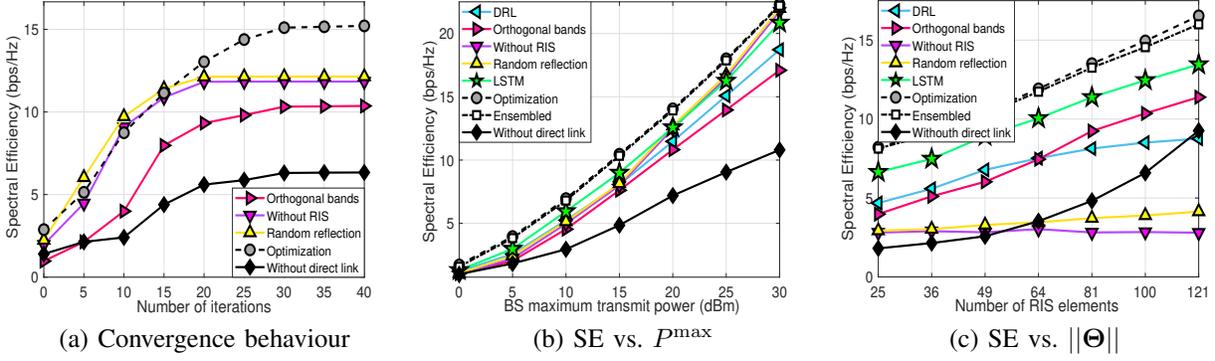
(a) Convergence behaviour (b) SE vs. $P^{\max}$ (c) SE vs. $||\boldsymbol{\Theta}||$

Fig. 3: Key system parameter optimization.

(a) prediction (phase) (b) prediction (power) (c) prediction (RB)
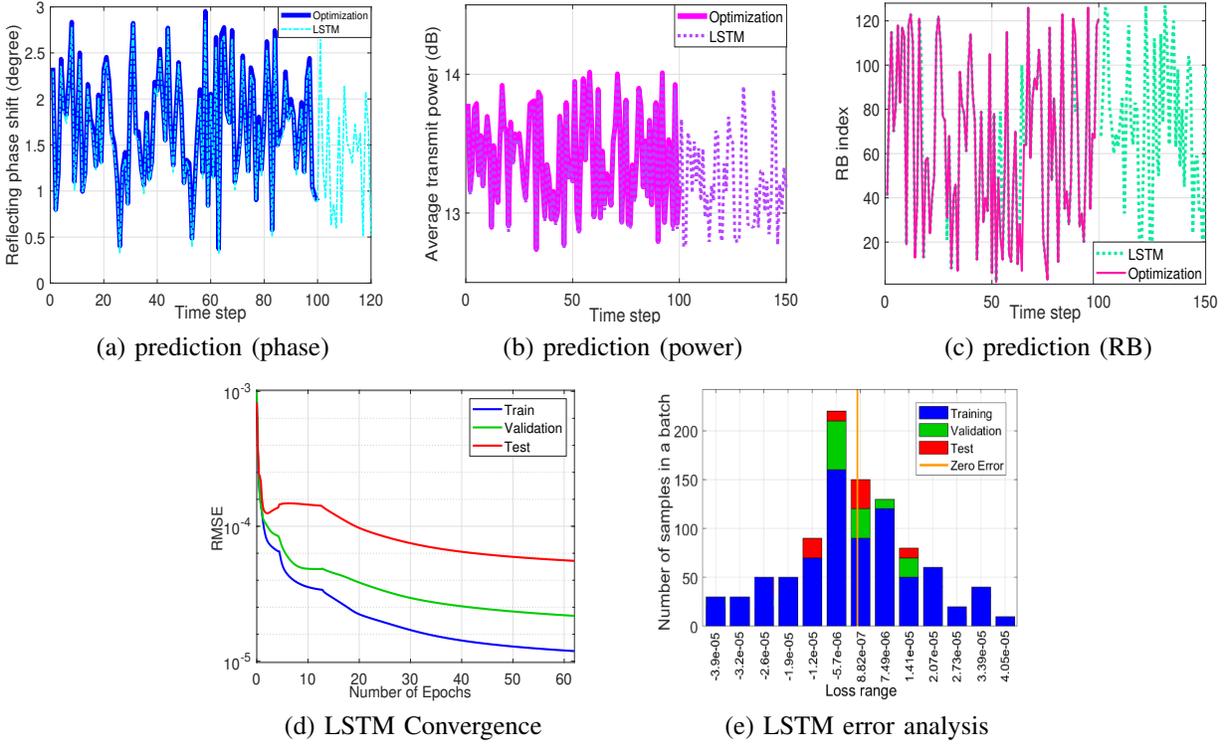

(d) LSTM Convergence (e) LSTM error analysis

Fig. 4: Analyzing the LSTM performance in terms of training, validation, and test accuracy.

adopted at the BS. The RIS is equipped with $M = 10 \times 10$ elements. The conventional modeling of the RIS is used in simulations, where the amplitude of the RIS elements is not a function of their phase shift, unless otherwise stated. However, we will also present some results based on a more practical RIS model [26]. The URLLC service traffic model follows a Poisson process, where the packet size is assumed to be 50 bytes with an arrival rate of 5 packets per mini-slot [14].

The frequency of the sth RB in THz spectrum satisfies $f_s = f_c + \frac{B}{S}(s - 1 - \frac{S-1}{2})$, with a central frequency of $f_c = 100$ GHz. The transmission bandwidth of each RB is $B = 10$ GHz and we consider $S = 128$ RBs [39]. The numbers of clusters accounting for the quasi-optical characteristic in THz signals, are assumed to be $N_{\mathrm{cl}}^1 = N_{\mathrm{cl}}^2 = N_{\mathrm{cl}}^3 = 2$, with $N_{\mathrm{ray}}^1 = N_{\mathrm{ray}}^2 = N_{\mathrm{ray}}^3 = 2$ rays in each cluster. The

delays of clusters and the delay offsets for the rays follow a uniform distribution within $[0, 20]$(ns) and $[-0.1, 0.1]$(ns), respectively [39].

To evaluate the training performance of the LSTM, we use the RMSE metric, i.e., RMSE $= \sqrt{\sum_{i=1}^{M} ||\boldsymbol{\Xi}^i - \widehat{\boldsymbol{\Xi}^i}||_2^2}$. We observe 100 time slots of tracking with 0.15 sec interval in a LSTM with 7 hidden layers. The training, validation and test phases of the LSTM use 70%, 15%, and 15% of the total samples, respectively. Finally, to enable the ensemble learning, we consider $M_1 = 15$ LSTMs as the weak regressors/classifiers.

B. Simulation Baselines

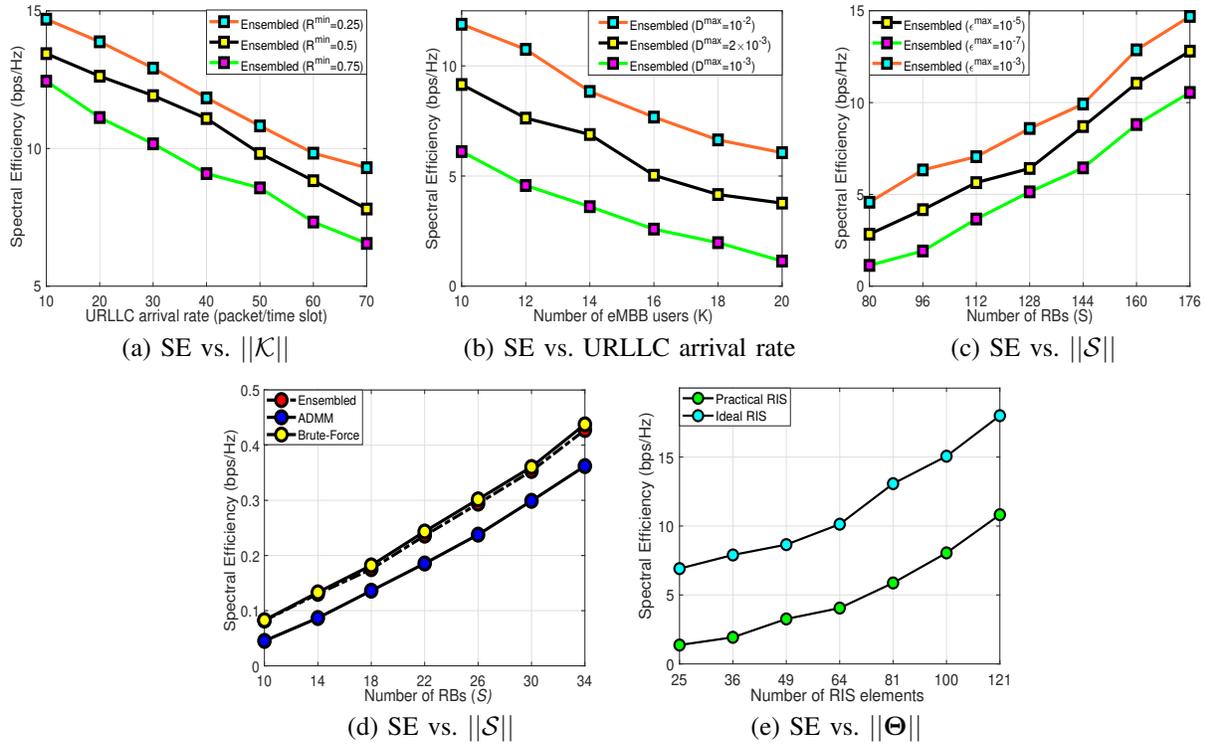In Figs. 3, we evaluate the achievable eMBB SE for the following schemes:

Fig. 5: Analyzing the performance of the proposed ensemble model for coexisting URLLC and eMBB services.

- *Optimization*: The SE is plotted using the optimized key system parameters $\Xi(t)$ via the proposed alternative decomposition approach.
- *LSTM*: The SE is plotted using the predicted key system parameters $\widehat{\Xi}(t)$ via the trained LSTM.
- *Ensemble LSTM*: The SE is plotted using the predicted key system parameters $\widehat{\Xi}^{\text{ens}}(t)$ in the ensemble model.
- *Orthogonal RBs*: The SE employs the alternative decomposition optimization procedure only (i.e., neither the LSTM nor the ensemble model), where each RB is orthogonally allocated only to one eMBB user.
- *DRL*: The resource management is performed according to the DRL procedure using the configurations in [14].
- *Without Direct Link*: All direct links between the BS and the users are blocked, which enables RIS-only transmission.
- *Random reflection*: The reflection coefficients of the RIS elements are randomly adopted from the feasible set $\mathcal{Q}$.
- *Without RIS*: The conventional cellular network without RIS, wherein the communications completely rely on direct links between the BS and the users.

For the DRL model to address the problem in (32), the state, action and the reward of the agent can be defined as follows:

- **State Space:** At instant t, the state of the agent is defined as the channel gains at instant $t-1$ i.e., $\mathcal{H}(t-1)$.
- **Action Space:** The action adopted by the agent at instant t indicates the decision variables of (32) i.e., $a(t) = \left[\Theta(t), \bar{\rho}(t), P(t) \right]$.
- **Reward:** The value returned by the reward function of the agent at instant t is categorized in manifold cases as

follows:

1) **C1**-**C13** hold and $\Lambda(t) \geq \Lambda^*$; therefore, $\mathcal{R}(t) = \Lambda(t)$ and $\Lambda^* = \Lambda(t)$.

2) **C1**-**C13** hold, $\Lambda(t) \leq \Lambda^*$ and $\Lambda(t) \geq \Lambda(t-1)$; therefore, $\mathcal{R}(t) = \Lambda(t) - |\Lambda(t)|$.

3) **C1**-**C13** hold, $\Lambda(t) \leq \Lambda^*$, $\Lambda(t) \leq \Lambda(t-1)$; therefore, $\mathcal{R}(t) = \Lambda(t) - 2|\Lambda(t)|$.

4) **C1**-**C13** do not hold; therefore. $\mathcal{R}(t) = \Lambda(t) - 3|\Lambda(t)|$.

In all the cases, $\Lambda(t) = \sum_{k=1}^{K} R_k^{\text{Act}}\left(P(t), \rho(t), \tilde{\rho}(t), \Theta(t) \right)$, and Λ^* stands for the highest Λ achieved by the agent thus far. By doing so, the decision variables $\{P, \rho, \tilde{\rho}, \Theta\}$ are optimized, such that the most gain on Λ is achieved, while the constraints **C1**-**C13** are satisfied.

Moreover, in Fig. 5, we consider the brute-force method and the alternating direction method of multipliers (ADMM) method [32], respectively as globally optimal and sub-optimal baseline schemes, for the RB allocation, to be compared with our proposed method. In Fig. 5, as well, the performance of our proposed resource allocation framework is already investigated relying on both the ideal RIS architecture, as well as the practical RIS architecture proposed in [26].

C. Optimization Results

In this subsection, we show the convergence behaviour and achievable SE for eMBB service of the proposed alternative decomposition method.

1) Convergence of the Optimization Algorithm: Fig. 3(a) demonstrates the convergence of the proposed optimization algorithm, wherein five optimization-driven baselines, i.e., "Optimization", "Orthogonal bands", "Without direct link",

"Random reflection" and "Without RIS" are included for comparison[11]. It is clearly observed that within the first 30 iterations, all schemes converge to their corresponding sub-optimal values. However, the baselines take more iterations to converge due to addressing the four aforementioned sub-problems in section III. In comparison, it takes only 15 iterations for "Random reflection" and "Without RIS" baselines since there is no RIS reflection coefficient optimization therein.

Another observation from Fig. 3(a) is that the "Optimization" baseline is superior compared to the "Orthogonal bands" baseline, because a non-orthogonal RB allocation achieves better SE than its orthogonal counterpart [30]. Since the baseline "Without direct link" merely benefits from the RIS-only transmissions, it achieves lower eMBB SE due to neglecting the direct links between the BS and the users. Eventually, "Random reflection" and "Without RIS" baselines exhibit the lowest achieved eMBB SE due to the absence of optimized RIS-aided transmissions.

2) BS transmit power threshold: Fig. 3(b) analyzes the achievable eMBB SE of the baselines as a function of the BS transmit power budget. The SE of all the baselines increase with the increasing BS transmit power budget. The leading "Optimization" baseline is followed by the "Ensemble" baseline accompanied by a negligible performance loss (lower than $\leq 1\%$ more precisely). This achievement is brought by the precise regression/classification power of the ensemble model. Next, the LSTM baseline experiences some performance loss (at most 7% in comparison with the "Optimization" baseline) yet superior than others.

Also, Fig. 3(b) demonstrates the significance of RIS when the BS transmit power budget is within a moderate range [36]. For small BS transmit power budget (e.g., 0 dBm), the direct links between the BS and users are inactive and the signals reflected by the RIS are weak to overcome the severe attenuation in THz band. Consequently, all baselines experience limited achievable eMBB SE in this case. In contrast, with high transmit power budget (e.g., 30dBm), the direct links become more dominant and that is why the "Without direct link" baseline has a limited contribution to the eMBB SE improvement when $P^{\max} > 20$ dBm.

3) Number of RIS elements: According to Fig. 3(c), the SE of all baselines except "Without RIS" increases with increasing number of RIS elements. Evidently, the "Random reflection" has the mildest improvement due to non-optimized reflection coefficients. More importantly, owing to the large scale state-action space of "DRL" [15], the SE performance deteriorates with increasing number of RIS elements. For instance, in comparison with this baseline for an RIS of 11×11, up to 49%, 48%, and 34% higher SE is achieved by the "optimization", "ensemble model", and the "LSTM" baselines, respectively. For a small RIS, the baseline "Without direct link" offers no significant improvement since the communicating signals in this baseline benefit neither from direct links, nor from RIS-aided transmission links. For a large RIS, this baseline however improves due to efficient RIS-aided transmissions. We also

[11]Note that other baselines named in Subsection VII-B are excluded from this figure since they do not include any optimization operation.

note that, when the number of RIS elements in Fig. 3(c) grows, the gap between the "Optimization", "Ensemble", and "LSTM" baselines becomes more evident. For instance, given the 8×8 RIS, additional 17% SE is achieved for the "Ensemble" scheme compared to "LSTM" scheme, whereas for 11×11 RIS, the difference increases up to 21%.

D. LSTM Results

The training/prediction performance of the LSTM is analyzed in Fig. 4. Specifically in Fig. 4(a), the LSTM precisely tracks the RIS reflecting elements in first 100 time steps and predicts the RIS reflection coefficients, for the next 20 time steps. In Fig. 4(b) and Fig. 4(c), respectively, the LSTM meticulously tracks the average transmit power of the BS, as well as the wideband THz RB allocation within the first 100 time steps to predict them in the upcoming 50 time steps. Fig. 4(d) depicts that the LSTM converges quickly within limited number of epochs with negligible RMSE values. Note that the number of epochs is derived from optimizing the learning rate in LSTM via the Bayesian optimizer, as mentioned in Subsection IV.D. Evidently, the convergence rate of our proposed LSTM in Fig. 4(d) outperforms the one in [37] (shown in Fig. 7 of [37]). In addition, training the LSTM through the Levenberg-Marquardt policy leads to lower prediction variance, thanks to the optimized parameters.

Fig. 4(e) portrays the deviation error for training, validation and test phases when different number of training samples are taken within a batch. Note that the deviation error is the difference between the predicted values of the LSTM and the desired target values. The x-axis indicates $\left[\Xi(t) - \hat{\Xi}(t)\right]$ and there can be negative values related to the deviation errors. Since training the neural networks with a single batch of training data leads to inefficiencies, the training process is done in several episodes, each episode with a batch of the training data. Clearly, the more the number of training samples within a batch, the smaller is the error. Another observation from Fig. 4(e) is that the errors in validation and test phases are only limited to a small range, adjacent to zero error line and thus negligible. This indicates that the larger errors occur only during the training stage. The absence of large errors during validation and testing reveals that the future trend is well-predicted by the fine-tuned LSTM.

E. Results for the Ensemble Model

In this subsection, we specifically analyze the results derived from the ensemble model as the final state in the proposed supervised learning resource management framework. In particular, we evaluate the actual achievable SE for eMBB against varying URLLC traffic, the number of eMBB users, and the number of RBs in Figs. 5(a)-(c), respectively.

Fig. 5(a) shows that the actual achievable SE degrades by increasing the URLLC traffic load, owing to the priorities associated with the URLLC service over the eMBB service. It is also observed that the higher the SE target requirement $R_k^{min} \; \forall k$ for eMBB users, the lower is the number of users that can be served. Fig. 5(b) shows the actual achievable SE as a function of increasing number of eMBB users. Clearly,

SE degrades due to increasing NOMA interference in (27). Another observation is the detrimental impact of URLLC service latency threshold $D_u^{\max} \forall u$ on the actual achievable eMBB SE. Fig. 5(c) shows that a higher SE is achieved with increasing number of RBs. Also, the URLLC reliability threshold (according to **C4** in (32)) is a limiting factor for the actual achievable SE, since it exerts more restrictions on the solution boundary of (32). In Fig. 5(d), the proposed framework is evaluated for a small number of users and RBs using a brute-force search (optimal), and also using the ADMM method [32]. While former provides a globally optimal RB allocation strategy along with the ensemble method for reflection coefficient design and BS power allocation, the ADMM baseline scheme provides a locally optimal solution for the joint problem (32), where all the sub-problems are addressed by the ADMM method [32]. According to Fig. 5(d), for example, with 30 RBs, the proposed ensemble method achieves up to 17% performance gain, compared to the sub-optimal ADMM baseline, at the expense of a trivial performance loss (less than 1%), in comparison with the brute-force method. Fig. 5(e) reveals that the proposed resource allocation framework can be also applied on a practical RIS architecture such as the one proposed in [26]. For $\beta_{\min} = 0.2$, $k = 1.6$ and $\phi = 0.43\pi$ [26], more than 80% performance loss is observed in comparison with the ideal model, for a deployed RIS with 10×10 elements.

VII. CONCLUSION

We have investigated the problem of coexistence of eMBB and URLLC services over RIS-aided THz communications. We have proposed a supervised learning-based resource management framework for optimizing the transmit power of the BS, the reflection coefficients for reflections by the RIS elements, as well as RB allocation. The proposed approach leverages an optimization policy, a deep learning method, and an ensemble learning method together to fulfill the eMBB and URLLC services requirements. We have shown that, in presence of an RIS of size 11×11, the proposed optimization, deep learning and ensemble learning techniques, respectively, achieve up to 49%, 34%, and 48% eMBB SE gain over the DRL scheme [15].

APPENDIX

A. Proof of Lemma 1

By substituting (27), the MMSE detection problem (39) can be represented as (66) whereby the optimal equalization coefficient $c_k^o[s]$ in (39) is obtained as:

$$\frac{\partial e_k[s]}{\partial c_k[s]}\Big|_{c_k^o[s]} = 0, \tag{67}$$

$$\implies -\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]} + (c_k^o[s])^*$$
$$\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2 \rho_k[s]p_k[s] + \xi_k[s]\right) = 0, \tag{68}$$

$$\implies c_k^o[s] = \left(\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right)^*$$
$$\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2 \rho_k[s]p_k[s] + \xi_k[s]\right)^{-1}. \tag{69}$$

Besides, by replacing (67) into (66), the objective of the MMSE problem can be rewritten as (70), which implies

$$\left(1 + \gamma_k[s]\right)^{-1} = \min_{c_k[s]} e_k[s]. \tag{71}$$

The actual achievable data rate for the eMBB user k over the RB s therefore, can be given by [29]

$$R_k^{\text{Act}} = \sum_{s \in \mathcal{S}} \left[\rho_k[s]\log\left(1 + \gamma_k[s]\right) - \tag{72}\right.$$
$$\left. \rho_k[s]\log\left(1 + \gamma_k[s]\right) \sum_{u \in \mathcal{U}} \sum_{f \in \mathcal{F}} \mathbb{I}(\rho_k[s] = \tilde{\rho}_u^f[s]) \right] = \sum_{s \in \mathcal{S}} (1 - \Phi_k)$$
$$\rho_k[s]\log\left(1 + \gamma_k[s]\right) = \left(\max_{c_k[s]}(1 - \Phi_k)\left(-\log e_k[s]\right) \forall k \in \mathcal{K}.$$

Note that in (72), the logarithmic function raises a challenge, for which the following theorem is invoked.

Theorem 1. *Suppose that* $f(a) = -\frac{ab}{\ln 2} + \log_2 a + \frac{1}{\ln 2}$, *where a is a positive real number, it can be concluded that* $\max f(a) = -\log_2 b$, *where the optimal value of a is $a^o = \frac{1}{b}$.*

Proof. Due to the logarithmic function, $f(a)$ is concave and its maximum value can be obtained as $\frac{\partial f(a)}{\partial a}\big|_{a=a^o} = 0$, where substituting $a = a^o = \frac{1}{b}$, results $-\log_2 b$ to be the maximum value of $f(a)$. Relying on **Theorem 1**, (72) can be given as (40) and the proof is complete. $\square$

B. Proof of Lemma 2

With respect to the linear objective function and the constraints, the reformulated sub-problem (43) is convex. The Lagrange function to this sub-problem is formulated as

$$L(p, \lambda, \mu) = \sum_{s=1}^{S} \sum_{k=1}^{K} (1 - \Phi_k) a_k^{(t)}[s] e_k^{(t)}[s]$$
$$+ \lambda \left(\sum_{s=1}^{S} \sum_{k=1}^{K} p_k^{(t)}[s] - P \right) + \sum_{s=1}^{S} \sum_{k=1}^{K} \mu_k[s]\theta_k[s], \tag{73}$$

and

$$\theta_k[s] = \eta \sum_{\substack{i \in \mathcal{K}, i \neq k, \\ \Upsilon(k) < \Upsilon(i)}} \left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_i[s] \right|^2 \rho_i[s]p_i^{(t)}[s] +$$
$$\left| \left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s] \right|^2 \rho_k[s]p_k^{(t)}[s] + \omega, \tag{74}$$

in which $\lambda \geq 0$ and $\mu_k[s] \geq 0, \forall k \in \{1, 2, ..., K\}, \forall s \in \{1, 2, ..., S\}$. The KKT conditions of (43) as well, hold as

$$
\begin{aligned}
e_k[s] &= 1 - 2\Re\left(c_k[s](\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right) + |c_k[s]|^2\left(\left|(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\right|^2\rho_k[s]p_k[s] + \xi_k[s]\right) \\
&= \left|1 - c_k[s](\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right|^2 + |c_k[s]|^2\,||(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]||_2^2 \sum_{\substack{i\in\mathcal{K}, i\neq k, \\ \Upsilon(k)<\Upsilon(i)}}\rho_i[s]p_i[s] + |c_k[s]|^2\,\sigma^2. \quad (66)
\end{aligned}
$$

$$
\begin{aligned}
e_k^o[s] &= 1 - 2\Re\left(c_k^o[s]\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right) + |c_k^o[s]|^2\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2\rho_k[s]p_k[s] + \xi_k[s]\right) \\
&= 1 - 2\Re\left(\left(\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right)\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2\rho_k[s]p_k[s] + \xi_k[s]\right)^{-1}\right. \\
&\quad \left.\left(\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right) + \left|\left(\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\rho_k[s]\sqrt{p_k[s]}\right)\right.\right. \\
&\quad \left.\left.\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2\rho_k[s]p_k[s] + \xi_k[s]\right)^{-1}\right|^2\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2\rho_k[s]p_k[s] + \xi_k[s]\right)\right) \\
&= 1 - \rho_k[s]p_k[s]\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2\left(\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2\rho_k[s]p_k[s] + \xi_k[s]\right)^{-1}, \quad (70)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial p_k[s]} &= a_k^{(t)}[s]\left(\left|c_k^{(t)}[s]\right|\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2 - \Re\left(c_k^{(t)}[s](\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\right)\left(p_k^{(t)}[s]\right)^{-1/2}\right) \\
&\quad + a_k^{(t)}[s]|c_k[s]|^2\left|(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\right|^2\sum_{i\in\mathcal{K}, i\neq k, \Upsilon(k)>\Upsilon(i)}\rho_i[s] + \lambda - \mu_k[s]\left|\left(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s]\right)\boldsymbol{w}_k[s]\right|^2 \\
&\quad + \mu_k[s]\eta\left|(\boldsymbol{h}_{d,k}[s] + \mathbf{G}[s]\boldsymbol{\Theta}\boldsymbol{h}_{r,k}[s])\boldsymbol{w}_k[s]\right|^2\sum_{\substack{i\in\mathcal{K}, i\neq k, \\ \Upsilon(k)>\Upsilon(i)}}\rho_i[s] = 0, \quad (75)
\end{aligned}
$$

$$
\lambda\left(\sum_{k=1}^{K}\sum_{s=1}^{S}p_k^{(t)}[s] - P\right) = 0, \quad (76)
$$

and

$$
\mu_k^{(t)}[s]\theta_k^{(t)}[s] = 0. \quad (77)
$$

In accordance with (75), the optimal transmit power of the BS at iteration t can be obtained as (45).

C. Proof of Lemma 3

The primary joint optimization problem (32) has been decoupled into the sub-problems (33), (47) and (49) up to now and addressed in a joint iterative procedure on the basis of the alternative decomposition approach. The overall iterative procedure is outlined as follows:

$$
\begin{aligned}
&[\boldsymbol{P}^{(0)}, \boldsymbol{\rho}^{(0)}, \tilde{\boldsymbol{\rho}}^{(0)}, \boldsymbol{\Theta}^{(0)}] \to [\boldsymbol{P}^{(1)}, \boldsymbol{\rho}^{(1)}, \tilde{\boldsymbol{\rho}}^{(1)}, \boldsymbol{\Theta}^{(0)}] \to \dots \to \\
&[\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t)}] \to \dots \to [\boldsymbol{P}^{(\text{opt})}, \boldsymbol{\rho}^{(\text{opt})}, \tilde{\boldsymbol{\rho}}^{(\text{opt})}, \boldsymbol{\Theta}^{(\text{opt})}].
\end{aligned}
$$
(78)

Until the convergence is achieved, the objective function increases or remains unchanged [30]. For instance, at the iteration t, first the power control sub-problem is addressed, where the objective function is accordingly updated as:

$$
\begin{aligned}
&U\left(\boldsymbol{P}^{(t-1)}, \boldsymbol{\rho}^{(t-1)}, \tilde{\boldsymbol{\rho}}^{(t-1)}, \boldsymbol{\Theta}^{(t-1)}\right) \\
&\leq U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t-1)}, \tilde{\boldsymbol{\rho}}^{(t-1)}, \boldsymbol{\Theta}^{(t-1)}\right),
\end{aligned}
$$
(79)

where

$$
U\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right) = \sum_{k=1}^{K} R_k^{\text{Act}}\left(\boldsymbol{P}, \boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}, \boldsymbol{\Theta}\right). \quad (80)
$$

The RB allocation sub-problem for the eMBB service is then solved, where the objective function is accordingly updated as:

$$
U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t-1)}, \tilde{\boldsymbol{\rho}}^{(t-1)}, \boldsymbol{\Theta}^{(t-1)}\right) \leq U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t-1)}, \boldsymbol{\Theta}^{(t-1)}\right).
$$
(81)

Relying on the results obtained formerly, the RB allocation sub-problem for the URLLC service is addressed, where the objective function is updated as:

$$
U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t-1)}, \boldsymbol{\Theta}^{(t-1)}\right) \leq U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t-1)}\right).
$$
(82)

We then solve the phase shift design sub-problem by updating the objective function as:

$$U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t-1)}\right) \leq U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t)}\right). \tag{83}$$

The convergence is eventually achieved, subject to meeting the following condition:

$$\left| U\left(\boldsymbol{P}^{(t)}, \boldsymbol{\rho}^{(t)}, \tilde{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\Theta}^{(t)}\right) \right.$$
$$\left. - U\left(\boldsymbol{P}^{(t-1)}, \boldsymbol{\rho}^{(t-1)}, \tilde{\boldsymbol{\rho}}^{(t-1)}, \boldsymbol{\Theta}^{(t-1)}\right) \right| \leq \nu. \tag{84}$$

Therefore, the alternative decomposition approach is convergent into the sub-optimal solution for the joint optimization problem (32) and the proof is complete.

REFERENCES

[1] S. R. Pokhrel, J. Ding, J. Park, O.-S. Park, and J. Choi, "Towards enabling critical MMTC: A review of URLLC within MMTC," *IEEE Access*, vol. 8, pp. 131796-131813, 2020.

[2] J. Zhang, X. Xu, K. Zhang, B. Zhang, X. Tao, and P. Zhang, "Machine learning based flexible transmission time interval scheduling for eMBB and uRLLC coexistence scenario," *IEEE Access*, vol. 7, pp. 65811–65820, 2019.

[3] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for eMBB and URLLC coexistence in 5G and beyond: A deep reinforcement learning based approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4585-4600, July 2021.

[4] A. K. Bairagi, Md. S. Munir, M. Alsenwi, N. H. Tran, S. S. Alshamrani, M. Masud, and Z. Han, "Coexistence mechanism between eMBB and URLLC in 5G wireless networks," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1736-1749, March 2021.

[5] A. Anand, G. De Veciana, and S. Shakkottai, "Joint scheduling of URLLC and eMBB traffic in 5G wireless networks," *IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, pp. 1970–1978, Ap. 2018.

[6] K. Pedersen, F. Frederiksen, G. Berardinelli, and P. Mogensen, "A flexible frame structure for 5G wide area," *IEEE 82nd Vehc. Technl. Conf. (VTC-Fall)*, pp. 1–5, 2015.

[7] W. Sui, X. Chen, S. Zhang, Z. Jiang, and S. Xu, "Energy-efficient resource allocation with flexible frame structure for heterogeneous services," *IEEE Green Comput. Commun. (GreenCom), IEEE Cyb. Phys. Soc. Comput.*, pp. 749–755, 2019.

[8] M. Alsenwi and H. Choong Seon, "Resource scheduling of URLLC/eMBB traffics in 5G new radio: A punctured scheduling approach", *IEEE Global Commun. Conf. (IEEE GLOBECOM)*, pp. 1271–1277, 2020.

[9] M. Series, "IMT vision—framework and overall objectives of the future development of IMT for 2020 and beyond." *Recomm. ITU 2083*, 2015.

[10] ——, "User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone (Release 15," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.101-1, 12 2018, version 15.4.0. [Online]. Available: http://www.3gpp.org/ftp//Specs/archive/38_series/38.101-1/

[11] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, "A flexible 5G frame structure design for frequency-division duplex cases," *IEEE Commun. Magazine*, vol. 54, no. 3, pp. 53–59, March 2016.

[12] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," *IEEE INFOCOM*, pp. 270–278, 2018.

[13] Y. Huang, S. Li, C. Li, Y. T. Hou, and W. Lou, "A deep-reinforcement-learning-based approach to dynamic eMBB/URLLC multiplexing in 5G NR," *IEEE Internet of Things Journal,* vol. 7, no. 7, pp. 6439-6456, July 2020.

[14] J. Li and X. Zhang, "Deep reinforcement learning-based joint scheduling of eMBB and URLLC in 5G networks," *IEEE Wireless Commun. Letters*, vol. 9, no. 9, pp. 1543-1546, Sept. 2020.

[15] Y. Li, C. Hu, J. Wang, and M. Xu, "Optimization of URLLC and eMBB multiplexing via deep reinforcement learning," *IEEE ICCC Workshops*, Changchun, China, pp. 245-250, 2019.

[16] X. Gao, Y. Liu, X. Liu, and Z. Qin, "Resource allocation in RISs aided MISO-NOMA networks: A machine learning approach," *IEEE Globecom*, pp. 1–6, 2020.

[17] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters", *J. App. Math.*, pp. 431-441, 1963.

[18] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications: Part II - an improved method based on double-structured sparsity," *IEEE Commun. Letters*, vol. 25, no. 5, pp. 1403-1407, May 2021.

[19] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications: Part I - fundamentals, solutions, and future opportunities," *IEEE Commun. Letters*, vol. 25, no. 5, pp. 1398-1402, May 2021.

[20] H. Cho, Y. Kim, E. Lee, D. Choi, Y. Lee, and W. Rhee, "Basic enhancement strategies when using Bayesian optimization for hyperparameter tuning of deep neural networks," *IEEE Access*, vol. 8, pp. 52588-52608, 2020.

[21] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123-140, 1996.

[22] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, 1987.

[23] S. Zargari, A. Hakimi, C. Tellambura, and S. Herath, "Multiuser MISO PS-SWIPT Systems: Active or Passive RIS?," *IEEE Wireless Commun. Letters*, vol. 11, no. 9, pp. 1920-1924, Sept. 2022.

[24] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. Passive RIS: Which Will Prevail in 6G?", [Online]. Available: https://arxiv.org/abs/2103.15154v5, Mar.

[25] K. Feng, Q. Wang, X. Li, and C.-K. Wen, "Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems," *IEEE Wireless Commun. Letters*, vol. 9, no. 5, pp. 745–749, 2020.

[26] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent Reflecting Surface: Practical Phase Shift Model and Beamforming Optimization," *Proc. IEEE Int. Conf. Commun. (ICC),* Dublin, Ireland, pp. 1-6, 2020.

[27] K. D. Katsanos, N. Shlezinger, M. F. Imani, and G. C. Alexandropoulos, "Wideband multi-user MIMO communications with frequency selective RISs: Element response modeling and sum-rate maximization," *Proc. IEEE ICC,* Seoul, South Korea, May 2022.

[28] C. She and C. Yang, "Ensuring the quality-of-service of tactile internet," in *Proc. IEEE Veh. Techn. Conf. (VTC Spring)*, pp. 1–5, May 2016.

[29] S. Boyd and L. Vandenberghe, "*Convex Optimization.*" Cambridge, U.K.: Cambridge Univ. Press, 2004.

[30] M. Moltafet, R. Joda, N. Mokari, M. R. Sabagh, and M. Zorzi, "Joint access and fronthaul radio resource allocation in PD-NOMA-based 5G networks enabling dual connectivity and CoMP," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6463-6477, Dec. 2018.

[31] A. Khalili, M. R. Mili, M. Rasti, S. Parsaeefard, and D. W. K. Ng, "Antenna selection strategy for energy efficiency maximization in uplink OFDMA networks: A multi-objective approach," *IEEE Trans. Wireless Commun.*, pp. 595—609, Oct. 2019.

[32] H. Zarini, A. Khalili, H. Tabassum, M. Rasti, and W. Saad, "AlexNet classifier and support vector regressor for scheduling and power control in multimedia heterogeneous networks," *IEEE Trans. Mobile Comput.*, Early Access.

[33] S. P. Boyd and M. C. Grant. *The CVX users' guide release 2.2.* Accessed: Jan. 2020. [Online]. Available: http://cvxr.com/cvx/doc/CVX.pdf/

[34] J. Li and D. He, "A Bayesian optimization AdaBN-DCNN method with self-optimized structure and hyperparameters for domain adaptation remaining useful life prediction," *IEEE Access*, vol. 8, pp. 41482-41501, 2020.

[35] J. G. Stoddard, J. S. Welsh, and H. Hjalmarsson, "EM-based hyperparameter optimization for regularized Volterra kernel estimation," *IEEE Cont. Sys. Lett.*, vol. 1, no. 2, pp. 388-393, 2017.

[36] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *IEEE Trans. Signal Processing*, vol. 69, pp. 4085-4101, 2021.

[37] K. Hamedani, L. Liu, S. Hu, J. Ashdown, J. Wu, and Y. Yi, "Detecting dynamic attacks in smart grids using reservoir computing: A spiking delayed feedback reservoir based approach," *IEEE Trans. Emrg. Top. Comput. Intell.*, vol. 4, no. 3, pp. 253-264, June 2020.

[38] Z. Zhang and L. Dai, "Capacity improvement in wideband reconfigurable intelligent surface-aided cell-free network," *IEEE 21st SPAWC*, pp. 1-5, Atlanta, GA, USA, 2020.

[39] J. Tan and L. Dai, "Wideband beam tracking in THz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1693-1710, June 2021.

Hosein Zarini is currently pursuing his Ph.D. degree with the department of computer engineering at Sharif University of Technology, Tehran, Iran. Priorly, he obtained the M.Sc. degree under the supervision of Dr. Mehdi Rasti from Amirkabir University of Technology, Tehran, Iran. His research interests include Beyond 5G and 6G wireless network, non-orthogonal multiple access (NOMA), ultra-reliable and low-latency communication (URLLC) service, enhanced mobile broadband (eMBB) service, massive machine type communication (mMTC) service, Massive MIMO communication, mMWave and THz communication, digital twin network, intelligent reflecting surface (IRS) communication, unmanned aerial vehicle (UAV) communication, green communication, mobile edge computing, optimization theory and machine learning.

Narges Gholipoor received the M.Sc. degree in telecommunication engineering from the Amirkabir University of Technology, Tehran, Iran, in 2016. She completed her Ph.D. from Tarbiat Modares University, Tehran, Iran, by working on radio and network resource allocation for the Tactile Internet service, an ultra-reliable low latency communication (URLLC) service, in February 2021. In addition, she was also involved in a number of large-scale network design and consulting projects in the telecom industry. She has been a member of ITU-R WP5D since September 2020, and she is pursuing her research goals by participating in ITU meetings. Her general research areas include Tactile Internet services, URLLC, non-orthogonal multiple access (NOMA), multiplexing of different services, dynamic framing, network function virtualization (NFV), THz communication, digital twin network, and radio resource allocation in 5G and beyond.

Mohammad Robat Mili received the Ph.D. degree in electrical and electronic engineering from the University of Manchester, U.K., in 2012. He held postdoctoral research positions at the Department of Telecommunications and Information Processing, Ghent University, Belgium and the Department of Electrical Engineering, Sharif University of Technology, Iran. His main research interests are in the area of design and analysis of wireless communication networks with particular focus on 5G and 6G cellular networks using mathematical methods such as optimization theory, game theory, and machine learning."

Mehdi Rasti (S'08-M'11-SM'21) is currently an Associate Professor at the Centre for Wireless Communications, University of Oulu, Finland. From 2012 to 2022 he was with the Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran. From Feb. 2021 to Jan. 2022, he was a visiting researcher at the Lappeenranta-Lahti University of Technology (LUT), Lappeenranta, Finland. From November 2007 to November 2008, he was a visiting researcher at the Wireless@KTH, Royal Institute of Technology, Stockholm, Sweden. From September 2010 to July 2012 he was with Shiraz University of Technology, Shiraz, Iran. From June 2013 to August 2013, and from July 2014 to August 2014 he was a visiting researcher in the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada. He received his B.Sc. degree from Shiraz University, Shiraz, Iran, and the M.Sc. and Ph.D. degrees both from Tarbiat Modares University, Tehran, Iran, all in Electrical Engineering in 2001, 2003 and 2009, respectively. His current research interests include radio resource allocation in IoT, Beyond 5G and 6G wireless networks.

Hina Tabassum (SM'17) is an Assistant Professor in the Department of Electrical Engineering & Computer Science, Lassonde School of Engineering, York University, Canada. Prior to that, she was a postdoctoral research associate at the Department of Electrical and Computer Engineering, University of Manitoba, Canada. She received her PhD degree from King Abdullah University of Science and Technology (KAUST) in 2013. She has been listed in the Stanford's list of the world's top 2% researchers in 2021 and has been recognized as N2Women: Rising Stars in Computer Networking and Communications in 2022. She is the founding chair of a special interest group on THz communications in IEEE Communications Society (ComSoc) - Radio Communications Committee (RCC). She is a Senior member of IEEE and registered Professional Engineer in the province of Ontario, Canada. Currently, she is serving as an Associate Editor for the *IEEE Transactions on Communications*, *IEEE Communications Letters*, *IEEE Transactions on Green Communications*, *IEEE Communications Surveys and Tutorials*, and *IEEE Open Journal of Communications Society*. Her current research interests include stochastic modeling and machine learning driven optimization of vehicular and aerial wireless networks enabled with millimeter (5G) and terahertz (6G) communication networks.

Ekram Hossain (F'15) is a Professor and the Associate Head (Graduate Studies) in the Department of Electrical and Computer Engineering at University of Manitoba (home.cc.umanitoba.ca/~hossaina), Canada. He is a Member (Class of 2016) of the College of the Royal Society of Canada. Also, he is a Fellow of the Canadian Academy of Engineering and a Fellow of the Engineering Institute of Canada. Dr. Hossain's current research interests include design, analysis, and optimization of 5G-Advanced and 6G cellular wireless networks. He was elevated to an IEEE Fellow "for contributions to spectrum management and resource allocation in cognitive and cellular radio networks". He received the 2017 IEEE ComSoc TCGCC (Technical Committee on Green Communications & Computing) Distinguished Technical Achievement Recognition Award "for outstanding technical leadership and achievement in green wireless communications and networking". Dr. Hossain has won several research awards including the "2017 IEEE Communications Society Best Survey Paper Award and the 2011 IEEE Communications Society Fred Ellersick Prize Paper Award. He was listed as a Clarivate Analytics Highly Cited Researcher in Computer Science in 2017, 2018, 2019, 2020, 2021, and 2022. Currently he serves as an Editor for the *IEEE Transactions on Mobile Computing* and the Director of Online Content for the IEEE Communications Society (ComSoc). Previously, he served as the Editor-in-Chief (EiC) of the IEEE Press (2018-2021) and the EiC of the *IEEE Communications Surveys and Tutorials* (2012-2016). He was a Distinguished Lecturer of the IEEE ComSoc and the IEEE Vehicular Technology Society. He served as the Director of Magazines for the IEEE ComSoc (2020-2021). Also, he was an elected member of the Board of Governors of the IEEE ComSoc for the term 2018-2020