Received 16 April 2025; revised 11 May 2025; accepted 29 May 2025. Date of publication 2 June 2025; date of current version 16 June 2025.
Digital Object Identifier 10.1109/OJCOMS.2025.3575947

Channel Estimation for Wideband Multi-RIS-Assisted mmWave Massive-MIMO OFDM System With Beam Squint Effect

THABANG C. RAPUDU [1] AND OLUTAYO O. OYERINDE [1] (Senior Member, IEEE)

[1] School of Electrical and Information Engineering, The University of the Witwatersrand, Johannesburg 2000, South Africa
CORRESPONDING AUTHOR: T. C. RAPUDU (e-mail: 1168108@students.wits.ac.za).

ABSTRACT Reconfigurable intelligent surfaces (RISs) and massive multiple-input multiple-output (massive-MIMO) systems are promising technologies for improving the energy efficiency of millimeter-wave (mmWave) communication. Furthermore, in urban areas, where there is high obscurity, multiple RISs can be deployed to circumvent blockages between communicating nodes. However, deploying both multi-RIS and massive-MIMO systems significantly increases the dimensionality of a wireless communication channel and thus, accurate channel state information (CSI) acquisition by channel estimation (CE) becomes non-trivial mainly due to the passive nature of the RISs. Additionally, existing wideband RIS-assisted CE schemes ignore the beam squint effect despite its severe CE performance degradation. Therefore, in this paper, a beam squint aware machine learning (ML)-based uplink CE scheme for wideband multi-RIS-assisted mmWave massive-MIMO orthogonal frequency division multiplexing (OFDM) system is proposed. Specifically, to reduce the beam squint effect, the bandwidth of the system is divided into subbands, and thereafter, a denoising convolutional neural network bidirectional long-short term memory (DnCNN-Bi-LSTM) scheme is proposed for cascaded uplink CE. For certain parameter settings, the proposed beam squint aware DnCNN-Bi-LSTM CE scheme achieves better normalized minimum mean squared error (NMSE) performance than the state-of-the-art beam squint aware CE methods.

INDEX TERMS Channel estimation, reconfigurable intelligent surface, beam squint, machine learning, MIMO.

I. INTRODUCTION

THE reconfigurable intelligent surfaces (RISs) and massive multiple-input multiple-output (massive-MIMO) technologies are promising technologies to overcome the challenges in millimeter-wave (mmWave) communications [1]. These challenges include but not limited to high path loss and atmospheric absorption. The RISs in a wireless communication system can improve signal coverage by controlling the incident signal towards desired directions [2]. Furthermore, RISs can act as an alternative pathway of the incident signals around the blockage between the transmitter and receiver in non line-of-sight (LoS) cases. On the other hand, massive-MIMO technology improves signal energy and spectral efficiency with very large number of antenna arrays [3]. With massive-MIMO, beamforming becomes possible, which improves mmWave coverage by steering the signal energy from transmitting antenna array of antennas to the desirable direction. There are several proposed beamforming techniques in literature to enhance mmWave performance. In [4], [5], [6], cooperative hybrid transmit beamforming was proposed where hybrid precoders and combiners are designed for cell-free mmWave MIMO networks. Conversely, a deep-learning-based active and passive Beamforming techniques were proposed in [7] for distributed simultaneous transmitting and reflecting (STAR) RIS-assisted systems.

Nevertheless, implementing the RISs and massive-MIMO technologies in a wireless communication system drastically increases the system's dimensionality. As a result, the process of channel estimation (CE) becomes challenging, and conventional CE methods such as least squares (LS) and minimum mean square error (MMSE) become ineffective due to their inability to track non linear data. In [8] and [9], sensing-based CE techniques inspired by the idea of integrated sensing and communication (ISAC) were proposed. However, the proposed schemes were based on the Line-of-sight (LoS) scenarios and they are based on the LS, which performs poorly for highly non linear data. Thus, the schemes in [8] and [9] will deteriorate for CE in non-LoS and high dimensionality scenarios like RIS-assisted systems. In [10], neural network (NN)-based CE scheme was proposed for orthogonal time frequency space (OTFS) modulation system by firt, extracting differential features of the channel to be fed to the NN. However, techniques to hence the NN performance like model pruning were not incorporated in the NN. There, the proposed NN will significantly be slower in high dimensionality scenarios like RIS-assisted systems.

A Bayesian framework-based generalized approximating message passing (GAMP) was proposed in [11] for RIS-assisted CE, assuming a Bernoulli-Gaussian (BG) distribution. However, a MMSE estimator was used for cascaded RIS-assisted CE, which is computationally expensive and requires prior statistical information of the channel for accurate estimates. This statistical information is challenging to obtain in RIS-assisted channels. In [12] and [13], LS-based schemes were proposed for RIS-assisted CE. However, since the LS-based CE techniques cannot optimally exploit the sparsity of the RIS-assisted channel, their CE performance are considerably poor in this case.

Luckily, compared to conventional CE methods as in [11], [12], [13], machine learning (ML) techniques have shown very good performance in estimating RIS-assisted massive-MIMO channels. The long short-term memory convolutional neural network (LSTM-CNN) scheme was proposed for RIS-assisted CE in [14]. The LSTM specialises with sequencial data and is a perfect model for time-varying CE. However, latest normalisation and regularisation methods like batch normalisation (BN) were not utilised to improve the convergence speed of the LSTM-CNN scheme in [14]. Conversely, in [15] a pruned autoencoder was proposed. However, pruning significantly increases the complexity of the autoencoder. In [16], a denoising CNN (DnCNN) scheme was proposed but the user equipments (UEs) were assumed to be static. The DnCNN is the deep-learning model based on the CNN and the residual learning to improve the learning rate of the CNN. However, for mobile UEs, the performance of the DnCNN will deteriorate as is the case in [16] because unlike recurrent neural networks (RNNs), its architecture was not optimised for fast time-varying channels.

In MIMO wideband communication systems however, there is an additional problem called the beam squint effect where the steering vectors of a massive-MIMO communication system are frequency-dependent [17]. This effect negatively affects the performance of CE and need to be accounted for. Ultimately, to obtain an optimal CE performance for a MIMO wideband communication systems, each subcarrier should have a dedicated set of steering vectors that points a beam of that particular subcarrier to a desirable direction. In narrowband communication systems, the effect of the beam squint is negligible and thus, all subcarriers can share a single set of steering vectors [18]. Anyway, just like in narrowband communication systems, conventional CE methods assume a single set of steering vectors for MIMO wideband CE, ignoring the effect of the beam squint, leading to the degraded channel estimates. Thus, it is important to pay special attention to the beam squint effect when performing CE for MIMO wideband communication systems, especially where RISs are deployed.

In [18], [19], [20], [21], [22], the beam squint effect was considered to improve the RIS-assisted CE's accuracy. The single-RIS-assisted channel of a single-input single-output (SISO) system was estimated in [18] by taking advantage of the mutual correlation function between the spatial steering vectors and the downlink cascaded channel from the base station (BS) to the user equipment (UE). And due to the beam squint, this mutual correlation function was found to have two peaks representing a pair of estimated angles for a single propagation path. This pair of estimated angles are the frequency-independent—actual angle' and the frequency-dependent —false angle'. In this case, the beam squint is the result of the interfering false angle. The twin-stage orthogonal matching pursuit (TS-OMP) algorithm was proposed to eliminate the false angle in the first stage by estimating the path angles of the cascaded channel. In the second stage, the TS-OMP algorithm estimates the propagation gains and delays. Similar to [18], a correlation based method was used in [19] and [20] to reduce the beam squint effect. However, this method depends on the number of steering vectors of the channel. That is, the mutual correlation function will have to be determined for every steering vector with the cascaded channel, resulting in high complexity in cases where there are many steering vectors like in a multi-RIS-assisted channel where each RIS has its own steering vector. Furthermore, similar to [18], an OMP algorithm was used for CE in [19]. However, the OMP algorithm is highly sensitive to noise. Thus, for lower signal to noise (SNR) scenarios, the OMP algorithm will significantly deteriorate. In contrast, a data-driven denoising convolutional neural network (DnCNN) algorithm was used for CE in [20]. However, it was assumed that the cascaded single-RIS-assisted channel was quasi-static (i.e., the UE, RIS and BS were assumed to be all static). In more practical scenarios however, UEs are mobile, and the performance of the DnCNN will deteriorate.

In [21], a single-RIS-assisted MIMO system was modeled with an assumption that the uplink channel between the multi-antenna UE and an extremely large-scale RIS was approximated as a near-field channel while the uplink channel between the RIS and the multi-antenna BS was assumed to be a far-field channel. In this case, the beam squint was assumed to be severe for a near-field channel and the beam squint in the far-field channel was assumed to be negligible. To reduce the beam squint effect, a spherical-domain dictionary was constructed by minimizing the coherence of two arbitrary subcarriers. Thereafter, the near-field channel was estimated using a multi-frequency parallelisable subspace recovery (MMPSR) architecture. Similar method was used to reduce the beam squint effect in [22]. However, as in [18], the beam squint reduction method proposed in [21] and [22] will deteriorate for systems with multiple steering vectors. Furthermore, the MMPSR algorithm of [21] performs well only for linear data. For non-linear multi-RIS-assisted massive-MIMO systems, the use of the MMPSR algorithm for CE will be ineffective. On the other hand, as in [18] and [19], the OMP based algorithm was used for CE in [22] and its performance will worsen for high complexity multi-RIS-assisted massive-MIMO systems.

The methods proposed in [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22] dealt with a single-RIS-assisted channels. However, in high obscurity environments like in urban areas, there is a necessity to implement multiple RISs to create a smooth mmWave communication path around many buildings which act as blockages between the communicating nodes [23], [24]. Increasing the number of RISs and/or number of RIS elements was proven to improve both ergodic capacity and outage probability in [23]. The proposed schemes in [25] and [26] considered CE for multi-RIS-assisted system but did not consider the beam squint effect, which therefore, affected CE accuracy. To the best of the authors' knowledge, no work has been reported in literature on CE for massive-MIMO multi-RIS-assisted mmWave system where the beam squint effect is considered. As stated previously, conventional CE methods will not be effective in such systems due to their inability to deal with non linearity of the complex massive-MIMO multi-RIS-assisted mmWave system while ML-based schemes show a much better performance. These serve as a motivation for this paper. Specifically, in this paper, ML-based CE techniques are proposed for a wideband multi-RIS-assisted mmWave massive-MIMO orthogonal frequency division multiplexing (OFDM) system with the consideration of the beam squint effect. The main contributions of this paper are as follows:

1) A novel beam squint aware CE scheme is proposed for wideband multi-RIS-assisted mmWave massive MIMO-OFDM system. That is, unlike all the work in the literature where the beam squint effect is considered for a communication system with no or single RIS implemented, the work in this paper expands to systems with two and more RISs. This is for the applications in high obscurity environments. By adopting the two-timescale channel property [27], the required pilot overhead is minimised in this paper.
2) The proposed methods for the reduction of the beam squint effect in several papers is based on the correlation between the steering vectors and the cascaded channel [18], [19], [20]. This method largely depends on the number of steering vectors of a MIMO wideband system and will be computationally expensive in multi-RIS-assisted systems since a correlation between the cascaded channel and every steering vector of each RIS will have to be computed for the reduction of the beam squint effect. However, in this paper, a beam squint reduction method (called the subdivided bandwidth (SDB) scheme) that is less dependent on the number of steering vectors is proposed. Specifically, the wideband is divided into smaller subbands such that each subband approximates a narrowband. In this case, the subcarriers within a certain subband will share a single set of steering vectors, thereby, reducing the beam squint effect.
3) For CE, the DnCNN bidirectional (Bi)-LSTM scheme is proposed and it is constructed to learn the residual instead of the actual channel coefficients for faster training. This residual learning also improves the accuracy of the estimates. Additionally, a BN regularization is utilised in the DnCNN to further increase the training speed. To cope with the fast time-varying channel between the UEs and the first RIS, the Bi-LSTM is utilised since its architecture was designed to mainly deal with time-varying data than other deep-learning variants.

The rest of the paper is organised as follows. In Section II, the system model and problem formulation is provided followed by the discussion of the proposed scheme in Section III. Thereafter, the simulation results of the proposed scheme are provided in Section IV. Section V concludes the paper.

Notation: The uppercase and lowercase bold letters represent matrices and vectors respectively, while the $||\cdot||_2$ and $||\cdot||_F$ are the $\mathbf{L}_2$ and Frobenius norms respectively. The $\mathbf{I}_X$, $(\cdot)^T$ and $(\cdot)^H$ are the $X \times X$ identity matrix, transpose and Hermitian, respectively. The $\mathrm{diag}(\mathbf{x})$ produces a diagonal matrix with the elements of $\mathbf{x}$ on the main diagonal; vec $(\mathbf{X})$ denotes the vectorization of a matrix $\mathbf{X}$; $\mathbb{C}^{x \times y}$ represents the space of $x \times y$ complex-valued matrices; the $\otimes$ denotes the Kronecker product respectively.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The system model and problem formulation are provided in this section.

A. SYSTEM MODEL

In this paper, the wideband multi-RIS-assisted mmWave massive MIMO-OFDM wireless communication system is considered as depicted in Fig. 1. This system consists of N UEs each with $N_{\text{ant}} = N_{\text{ant}_x} \times N_{\text{ant}_y}$ antennas, K RISs (i.e., RIS_k for $k \in \{1, 2, \ldots, K\}$) and a BS equipped with $M = M_x \times M_y$ uniform planar array (UPA) of antennas. It is assumed that the implemented RISs have equal number of reflecting elements P. Furthermore, the BS and all the RISs are assumed to be fixed at their positions while the UEs are mobile. The distances between the adjacent reflecting elements at every RIS and between the UE's and BS adjacent antennas is d_e. The uplink channel is the focus of this paper. Due to multiple paths, it is assumed that all the channels are frequency-selective and due to blockage, there is no LoS between the UEs and the BS. The OFDM block of the system

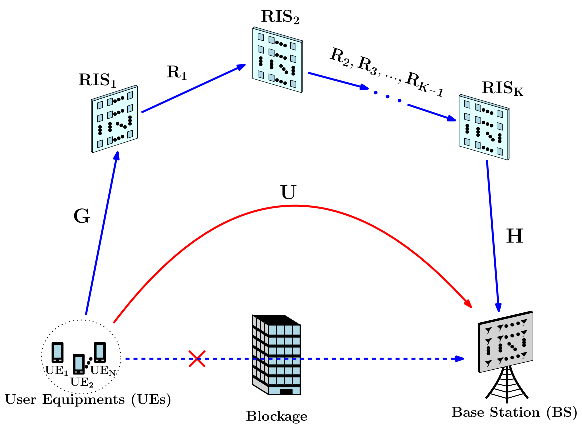

FIGURE 1. A generic system model of K RISs.

has N_{sc} (with $n_{\text{sc}} \in \{1, 2, \ldots, N_{\text{sc}}\}$ as the n_{sc}-th subcarrier) subcarriers and the bandwidth is assumed to be of length B. Therefore, there is $d_{\text{sc}} = \frac{B}{N_{\text{sc}}}$ spacing between adjacent subcarriers [28]. Furthermore, it is assumed that the cyclic prefix (CP) is long enough to overcome the discontinuity between the received signals that may happen due to the multi-path signal delays.

In Fig. 1, $\mathbf{G}$, $\{\mathbf{R}_j\}_{j=1}^{K-1}$ and $\mathbf{H}$ represent the uplink channels between the UEs and RIS_1, between adjacent RISs and between RIS_K and the BS, respectively. Let $\mathbf{G}_{d,n} \in \mathbb{C}^{P\times N_{\text{ant}}}$, $\mathbf{R}_{d,j,n} \in \mathbb{C}^{P\times P}$ and $\mathbf{H}_{d,n} \in \mathbb{C}^{M\times P}$ be the d-th sample of the low-pass filtered discrete-time channels of the n-th UE between the n-th UE and RIS_1, between adjacent RISs and between RIS_K and the BS, respectively, where $j \in \{1, 2, \ldots, K-1\}$ and $n \in \{1, 2, \ldots, N\}$. As a result of the fixed positions of all the RISs and the BS, the channels $\{\mathbf{R}_j\}_{j=1}^{K-1}$ and $\mathbf{H}$ are assumed to be quasi-static. Since all the UEs share the $\mathbf{H}_{d,n}$ and $\{\mathbf{R}_{d,j,n}\}_{j=1}^{K-1}$ channels for all time-domain delay taps, the subscript n can be omitted for these channels. In contrast, the channel $\mathbf{G}$ is time-varying due to the mobility of the UEs and it depends on the n-th UE.

Let parameters $\mathbf{G}_{d,n}$, $\mathbf{R}_{d,j}$, and $\mathbf{H}_d$ be written as

$$\mathbf{G}_{d,n} = \mathcal{G}\sum_{l_g=1}^{L_{g,n}} \alpha_{l_g} p_{rc}\big(dT_s - \tau_{l_g}\big)\mathbf{a}^r\Big(\vartheta_{l_g}^{r,RIS_1}, \psi_{l_g}^{r,RIS_1}\Big) \times \mathbf{a}^{t^H}\Big(\vartheta_{l_g}^{t,n}, \psi_{l_g}^{t,n}\Big)e^{j\frac{v_n f_{sc}}{c}\cos\theta_{l_g} dT_s}, \tag{1}$$

$$\mathbf{R}_{d,j} = \mathcal{R}_{d,j}\sum_{l_{R_{d,j}}=1}^{L_{R_{d,j}}} \alpha_{l_{R_{d,j}}}^{R_{d,j}} p_{rc}\Big(dT_s - \tau_{l_{R_{d,j}}}\Big) \mathbf{a}^r\Big(\vartheta_{l_{R_{d,j}}}^{r,RIS_{k+1}}, \psi_{l_{R_{d,j}}}^{r,RIS_{k+1}}\Big)\mathbf{a}^{t^H}\Big(\vartheta_{l_{R_{d,j}}}^{t,RIS_k}, \psi_{l_{R_{d,j}}}^{t,RIS_k}\Big), \tag{2}$$

and

$$\mathbf{H}_d = \mathcal{H}\sum_{l_H=1}^{L_H} \alpha_{l_H} p_{rc}\big(dT_s - \tau_{l_H}\big) \mathbf{a}^r\Big(\vartheta_{l_H}^{r,BS}, \psi_{l_H}^{r,BS}\Big)\mathbf{a}^{t^H}\Big(\vartheta_{l_H}^{t,RIS_K}, \psi_{l_H}^{t,RIS_K}\Big) \tag{3}$$

respectively.

In (1), (2) and (3), $L_{g,n}$, $\{L_{R_{d,j}}\}_{j=1}^{K-1}$ and L_H are the total number of paths of the uplink channels $\mathbf{G}_{d,n}$, $\{\mathbf{R}_{d,j}\}_{j=1}^{K-1}$ and $\mathbf{H}_d$ respectively; $\mathcal{G} = \sqrt{\frac{P\times N_{\text{ant}}}{L_{g,n}}}$, $\{\mathcal{R}_{d,j} = \sqrt{\frac{P^2}{L_{R_{d,j}}}}\}_{j=1}^{K-1}$ and $\mathcal{H} = \sqrt{\frac{M\times P}{L_H}}$ are the normalisation factors. The D_g, $\{D_{R_{d,j}}\}_{j=1}^{K-1}$ and D_H are the time domain delay tap lengths of the uplink channels $\mathbf{G}_{d,n}$, $\{\mathbf{R}_{d,j}\}_{j=1}^{K-1}$ and $\mathbf{H}_d$ respectively; $\alpha_{l_{g,n}}$, $\{\alpha_{l_{R_{d,j}}}\}_{j=1}^{K-1}$ and α_{l_H} represent the complex gain of the l_i-th path where the subscript $i \in \{\{g,n\}_{n=1}^N, \{R_{d,j}\}_{j=1}^{K-1}, H\}$, $l_{g,n} \in \{1, 2, \ldots, L_{g,n}\}$, $\{l_{R_{d,j}} \in \{1, 2, \ldots, L_{R_{d,j}}\}\}_{k=1}^{K-1}$ and $l_H \in \{1, 2, \ldots, L_H\}$; $p_{rc}(\tau)$ accounts for a low pass filtering and a filter incorporating the pulse-shaping effects evaluated at τ [16]; the $v_n, T_s, f_{n_{\text{sc}}}, c$ and θ_{l_g} are the speed of the n-th user, sampling period, subcarrier frequency, speed of light and the angle between UE direction and the incident direction of the subcarrier frequency for the l_g-th path respectively.

Furthermore, in (1), (2) and (3), $\mathbf{a}^t(\cdot,\cdot)$ and $\mathbf{a}^r(\cdot,\cdot)$ are the normalized UPA steering complex column vectors of length equivalent to the number of antenna/reflecting elements of the transmitting/reflecting and receiving nodes respectively; $\vartheta_{l_\iota}^{t,\cdot}$ and $\psi_{l_\iota}^{t,\cdot}$ are the azimuth and elevation angles of the transmitting/reflecting node of the l_ι-th path respectively, where "ι" is the channel in subject; $\vartheta_{l_\iota}^{r,\cdot}$ and $\psi_{l_\iota}^{r,\cdot}$ are the arrived azimuth and elevation angles of the receiving node of the l_ι-th path respectively.

Generally, the UPA steering complex column vector is given by

$$\mathbf{a}(\vartheta, \psi) = \frac{1}{\sqrt{N_{\text{UPA}}}} e^{-j\frac{d_{e_x}(\vartheta,\psi)}{\lambda_{n_{\text{sc}}}}\mathbf{n}_{\text{UPA}_x}} \otimes e^{-j\frac{d_{e_y}(\psi)}{\lambda_{n_{\text{sc}}}}\mathbf{n}_{\text{UPA}_y}} \tag{4}$$

where $N_{\text{UPA}} = N_{\text{UPA}_x} \times N_{\text{UPA}_y}$ and $\lambda_{n_{\text{sc}}}$ are the number of UPA elements and wavelength of the n_{sc}-th subcarrier respectively; $\mathbf{n}_{\text{UPA}_x} \in \{0, 1, \ldots, N_{\text{UPA}_x} - 1\}$, $\mathbf{n}_{\text{UPA}_y} \in \{0, 1, \ldots, N_{\text{UPA}_y} - 1\}$, $d_{e_x}(\vartheta, \psi) = 2\pi d_e sin(\vartheta)cos(\psi)$ and $d_{e_y}(\psi) = 2\pi d_e sin(\psi)$, where the element of spacing is given as $d_e = \lambda_c/2$, with λ_c representing the carrier wavelength.

The wavelength of the n_{sc}-th subcarrier can be presented in terms of the system's bandwidth B as

$$\lambda_{n_{\text{sc}}} = \frac{c}{f_{n_{\text{sc}}} + f_c} = \frac{\lambda_c}{1 + \frac{f_{n_{\text{sc}}}}{f_c}}, \tag{5}$$

where f_c is the carrier frequency and $f_{n_{\text{sc}}} = \frac{B}{N_{\text{sc}}}(n_{\text{sc}} - 1 - \frac{N_{\text{sc}}-1}{2})$. Note that (4) depends on a subcarrier frequency/wavelength (i.e., it depends on $\lambda_{n_{\text{sc}}}$). This frequency-dependency of the steering vector causes the beam to be squinted to an unintended direction and this phenomenon is called the beam squint effect. For narrowband mmWave systems, $B \ll f_c$ and $\lambda_{n_{\text{sc}}} \approx \lambda_c$, thus, (4) becomes less frequency-dependent. In this case, the beam squint effect is assumed to be negligible and all subcarriers can use the same set of steering vectors of a system. However, in wideband systems, $\lambda_{n_{\text{sc}}} \neq \lambda_c$ and (4) becomes more frequency-dependent, resulting in the beam squint effect.

B. PROBLEM FORMULATION

Since all RISs are assumed to have equal number of reflecting elements in this paper (i.e., if P_k represents the number of reflecting elements of the k-th RIS, it is assumed that $P_1 = P_2 = P_k = \cdots = P_K = P$), the cascaded channel $\mathbf{U}$ as depicted in Fig. 1 and RISs reflecting coefficients matrices can be represented separately. Specifically, the d-th sample of the low-pass filtered discrete-time cascaded channel between the UE$_n$ and BS can be represented as

$$\mathbf{U}_{d,n} = \mathbf{H}_{d,n}\left(\prod_{j=1,\ldots,K-1} \mathbf{R}_{d,j}\right) \mathbf{G}_{d,n} \in \mathbb{C}^{M\times N_{\mathrm{ant}}}. \tag{6}$$

Note that the assumption that the number of RIS elements is the same for all RIS is just to simplify the computation of (6) and it does not impose any practical limitation. That is, when the RISs have different RIS elements, the multiplication of channel matrices in (6) will require matrix transformations to be resolved to a single matrix $\mathbf{U}_{d,n}$.

Substituting (1), (2) and (3) into (6), the simplified result is

$$\mathbf{U}_{d,n} = \mathcal{U} \sum_{l_{\mathrm{H}}=1}^{L_{\mathrm{H}}} \sum_{l_{R_{d,1}}=1}^{L_{R_{d,1}}} \sum_{l_{R_{d,2}}=1}^{L_{R_{d,2}}} \cdots \sum_{l_{R_{d,K-1}}=1}^{L_{R_{d,K-1}}} \sum_{l_{\mathrm{G}}=1}^{L_{\mathrm{G}}} \alpha_{l_{\mathrm{U}}} p_{rc}\big(dT_s - \tau_{l_{\mathrm{U}}}\big)\mathbf{a}^r\Big(\vartheta_{l_{\mathrm{U}}}^{r,BS}, \psi_{l_{\mathrm{U}}}^{r,BS}\Big) \mathbf{a}^{tH}\Big(\vartheta_{l_{\mathrm{U}}}^{t,n}, \psi_{l_{\mathrm{U}}}^{t,n},\Big) e^{j\frac{v n f s c}{c} cos\theta_{l_g} dT_s} \in \mathbb{C}^{M\times N_{\mathrm{ant}}} \tag{7}$$

where $\mathcal{U} = \mathcal{G} \times \prod_{j=1,\ldots,K-1} \mathcal{R}_{d,j} \times \mathcal{H}$, $\alpha_{l_{\mathrm{U}}} = \alpha_{l_{\mathrm{G}}} \times \prod_{j=1,\ldots,K-1} \alpha_{l_{R_{d,j}}} \times \alpha_{l_{\mathrm{H}}}$, $p_{rc}(dT_s - \tau_{l_{\mathrm{U}}}) = p_{rc}(dT_s - \tau_{l_{\mathrm{H}}}) \times \prod_{j=1,\ldots,K-1} p_{rc}(dT_s - \tau_{l_{\mathrm{R}_{d,j}}}) \times p_{rc}(dT_s - \tau_{l_{\mathrm{G}}})$ and $\tau_{l_{\mathrm{U}}}$ is the total delay of the cascaded channel from UEs to the BS. Furthermore, $\mathbf{a}^r(\vartheta_{l_{\mathrm{U}}}^{r,BS}, \psi_{l_{\mathrm{U}}}^{r,BS}) = \mathbf{a}^r(\vartheta_{l_H}^{r,BS}, \psi_{l_H}^{r,BS}) \times \prod_{j=1,\ldots,K-1} \mathbf{a}^r(\vartheta_{l_{R_{d,j}}}^{r,RIS_{k+1}}, \psi_{l_{R_{d,j}}}^{r,RIS_{k+1}}) \times \mathbf{a}^r(\vartheta_{l_G}^{r,RIS_1}, \psi_{l_G}^{r,RIS_1})$ and $\mathbf{a}^{tH}(\vartheta_{l_{\mathrm{U}}}^{t,n}, \psi_{l_{\mathrm{U}}}^{t,n},) = \mathbf{a}^{tH}(\vartheta_{l_H}^{t,RIS_K}, \psi_{l_H}^{t,RIS_K}) \times \prod_{j=1,\ldots,K-1} \mathbf{a}^{tH}(\vartheta_{l_{R_{d,j}}}^{t,RIS_k}, \psi_{l_{R_{d,j}}}^{t,RIS_k}) \times \mathbf{a}^{tH}(\vartheta_{l_g}^{t,n}, \psi_{l_g}^{t,n})$, with $l_{\mathrm{U}} \in \{1, 2, 3, \ldots, L_{\mathrm{U}}\}$ where $L_{\mathrm{U}} = L_{\mathrm{G}} \times \prod_{j=1,\ldots,K-1} L_{R_{d,j}} \times L_{\mathrm{H}}$ is the total number of cascaded paths from the UEs to the BS and l_{U} is the l_{U}-th path from the UEs to the BS.

The RISs reflecting coefficients vectors of the n-th UE are represented as a combined reflecting coefficients vector as

$$\Theta_n = \prod_{k=1,2,\ldots,K} \Theta_{n,k}, \tag{8}$$

where $\Theta_{n,k} = [\theta_{n,k,1}, \theta_{n,k,2}, \ldots, \theta_{n,k,P}]$ is an $P \times 1$ reflecting vector for the n-th user of the k-th RIS with $\theta_{n,k,p}$ representing the reflecting coefficient for the n-th user of RIS$_k$ at the p-th reflecting element for $p \in \{1, 2, \ldots, P\}$.

By discrete fourier transformation (DFT), the frequency-domain cascaded channel of the n_{sc}-th subcarrier from UE$_n$ to the BS is given by

$$\mathbf{U}_n[n_{\mathrm{sc}}] = \sum_{d=0}^{D-1} \mathbf{U}_{d,n} e^{-j2\pi \frac{n_{\mathrm{sc}}}{N_{\mathrm{sc}}} d}, \tag{9}$$

where

$$D = D_G D_H \left(\prod_{j=1,2,\ldots,K-1} D_{R_j}\right). \tag{10}$$

By singular value decomposition [27], (9) can be represented as

$$\mathbf{U}_{n,q}[n_{\mathrm{sc}}] = \mathbf{A}_{\mathrm{BS}} \Delta_{n,q}[n_{\mathrm{sc}}] \mathbf{A}_n^H, \tag{11}$$

where $\mathbf{A}_{\mathrm{BS}}$ and $\mathbf{A}_n$ represent the $M \times M$ and $N_{\mathrm{ant}} \times N_{\mathrm{ant}}$ complex dictionary matrices of the BS's angle of arrival (AoA) and RIS$_1$'s angle of departure (AoD) respectively; $\Delta_{n,q}[n_{\mathrm{sc}}]$ represents the $M \times N_{\mathrm{ant}}$ sparse matrix, and its nonzero entries are the BS's AoAs and UE$_n$'s AoDs. SVD enables low-rank approximations by removing small or zero singular values especially for the matrix $\Delta_{n,q}[n_{\mathrm{sc}}]$, leading to a sparser representation. In essence, $\mathbf{A}_{\mathrm{BS}}$ and $\mathbf{A}_n$ contain the left and right singular vectors.

Without loss of generality, assuming that $N_{\mathrm{ant}} = P$ for simplicity, at the BS, the received signal of the n-th UE at the q-th time slot is given by

$$\mathbf{y}_{n,q}[n_{\mathrm{sc}}] = \mathbf{U}_{n,q}[n_{\mathrm{sc}}] \Theta_{n,q} s_{n,q}[n_{\mathrm{sc}}] + \mathbf{w}_{n,q}[n_{\mathrm{sc}}], \tag{12}$$

where $q \in \{1, 2, \ldots, Q\}$ and Q is the total number of time slots. The $\Theta_{n,q}$ is the combined reflecting coefficients vector at the q-th time slot for the n-th UE; $s_{n,q}[n_{\mathrm{sc}}]$ and $\mathbf{w}_{n,q}[n_{\mathrm{sc}}] \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_M)$ are the transmitted symbol and noise vector at q-th time slot of the n-th UE respectively, where σ^2 is the noise power at the n_{sc}-th subcarrier and $\mathbf{I}_M$ is an $M \times M$ identity matrix.

Moreover, from (11), (12) can be represented as

$$\mathbf{y}_{n,q}[n_{\mathrm{sc}}] = \mathbf{A}_{\mathrm{BS}} \Delta_{n,q}[n_{\mathrm{sc}}] \mathbf{A}_n^H \Theta_{n,q} s_{n,q}[n_{\mathrm{sc}}] + \mathbf{w}_{n,q}[n_{\mathrm{sc}}] \tag{13}$$

and then reduced to

$$\tilde{\mathbf{y}}_{n,q}[n_{\mathrm{sc}}] = \tilde{\Theta}_{n,q} \Delta_{n,q}[n_{\mathrm{sc}}] + \tilde{\mathbf{w}}_{n,q}[n_{\mathrm{sc}}], \tag{14}$$

where $\tilde{\mathbf{y}}_{n,q}[n_{sc}] = \mathbf{A}_{\mathrm{BS}}^H \mathbf{y}_{n,q}[n_{sc}] \in \mathbb{C}^{M\times 1}$, $\tilde{\Theta} = \mathbf{A}_n^H \Theta_{n,q} s_{n,q}[n_{\mathrm{sc}}] \in \mathbb{C}^{P\times 1}$ and $\tilde{\mathbf{w}}_{n,q}[n_{\mathrm{sc}}] = \mathbf{A}_{\mathrm{BS}}^H \mathbf{w}_{n,q}[n_{\mathrm{sc}}] \in \mathbb{C}^{M\times 1}$ are the effective measurement, sensing and noise vectors for the n_{sc}-th subcarrier respectively.

III. THE BEAM SQUINT EFFECT AND THE PROPOSED SCHEME

In this section, the beam squint effect is discussed followed by the proposed SDB-DnCNN-Bi-LSTM scheme.

A. THE BEAM SQUINT EFFECT

Noted that by (5), the spatial angles $(1+\frac{f_{n_{\mathrm{sc}}}}{f_c})sin(\psi)$ and $(1+\frac{f_{n_{\mathrm{sc}}}}{f_c})sin(\vartheta)cos(\psi)$ in (4) have different angular spread over different subcarriers. This is referred to as the "beam squint", which has negative affects on wideband CE. However, most works on wideband CE (such as works presented in [29] and [30]) ignored the beam squint effect, which result in inaccurate channel estimates.

On the other hand, as in [18] and [19] several works used a correlation-based method to eliminate the influence

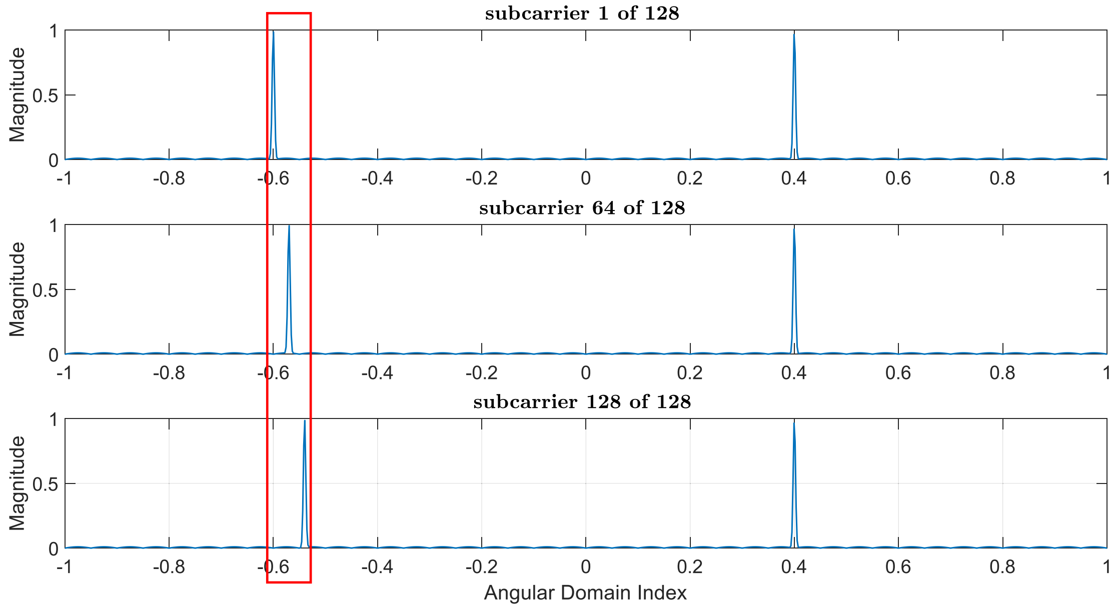

FIGURE 2. The beam squint effect.

of the beam squint effect by taking advantage of the angular orthogonality property of steering vectors. In this case, by finding the peaks of mutual correlation function

$$\Upsilon(x, y) = \mathbf{a}(x, y)^H \mathbf{U}_n[n_{sc}], \tag{15}$$

the equivalent angles of arrival (AoAs) and/or angles of departure (AoDs) can be estimated. The function $\Upsilon(x, y)$ will have two peaks/angles, one called the actual peak/angle at $(x = \eta_1, y = \eta_2)$ and the other at $(x = \eta_1 + \frac{f_c}{f_{n_{sc}}+f_c}, y = \eta_2 + \frac{f_c}{f_{n_{sc}}+f_c})$ which is referred to as a false angle. Note that the false peak/angle is frequency-dependent. Therefore, the actual peaks will remain the same for all subcarrier frequencies while the false peaks change as demonstrated in Fig. 2 for $B = 500$MHz, $f_c = 10$GHz and $N_{sc} = 128$. The peaks on the right are the actual peaks while the peaks on the left (inside the red rectangle) represent the false angles. To reduce the beam squint effect, the works in [18], [19], [20] proposed to accumulate the mutual correlation function $\Upsilon(x, y)$ over all subcarriers since for such accumulation, the total peak of the false angles will decrease as the total peak of the actual angles increase.

However, this method largely depends on the number of steering vectors in a system. For a multi-RIS-assisted system which has a high number of steering vectors, this method will be computationally expensive since $\Upsilon(x, y)$ will have to be computed for every steering vector and thereafter, accumulated for all subcarriers. In this paper, the beam squint elimination method which is less dependent on the number steering vectors is proposed as it will be discussed later.

B. THE PROPOSED TWO-TIMESCALE SDB-DNCNN-BI-LSTM SCHEME

Due to the deployed RISs in a system, high pilot overhead will be required for CE. This is because an RIS-assisted channel has significantly higher number of channel coefficients (especially if the transceivers are mobile [31]) which have to be estimated than in a non-RIS system. Therefore, it is necessary to make efforts to reduce the required pilot overhead especially for multi-RIS-assisted CE. Hence, in this paper, the two-timescale channel scheme is adopted to significantly reduce the required pilot overhead for multi-RIS-assisted CE. That is, the small-timescale and large-timescale coherence times are assumed for a time-varying and quasi-static channels respectively. However, unlike in [32], [33], [34], [35] where the large-timescale coherence time was assumed for a single-link/channel between the static RIS and a BS, the proposed large-timescale coherence time (T_{RB}) is assumed for the cascaded uplink channel (i.e., multi-link/channel) from RIS$_1$ all the way to the BS, and a single-link/channel between the UEs and RIS$_1$ adopts a small-timescale coherence time (T_{UE}).

Specifically, the CE in the proposed scheme is performed in two stages. Since the RISs and the BS are assumed to be at fixed positions at all times, the channels between them (i.e., the cascaded channel RIS$_1$-BS) are assumed to be quasi-static, and therefore, they can be estimated only once during the large-timescale coherence time (T_{RB}) in the first stage. In this case, pilots are transmitted from the UEs through the RISs all the way to the BS. The channel $\mathbf{G}$ can be assumed to be quasi-static and given an arbitrarily known constant since doing so will not affect the estimation of the RIS$_1$-BS cascaded channel in this stage. Therefore, the RIS$_1$-BS cascaded channel is estimated without the ambiguity of channel $\mathbf{G}$. Thereafter, in the second stage, the time-varying channel $\mathbf{G}$ is estimated during the small-timescale coherence time (T_{UE}) using the estimated cascaded channel RIS$_1$-BS. This two-timescale CE methods significantly reduces the

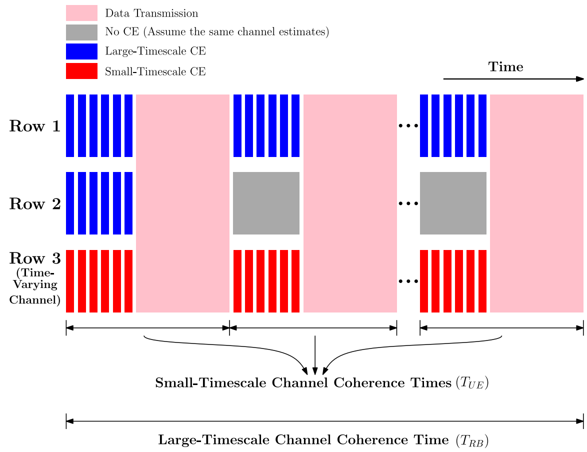

FIGURE 3. Difference between conventional and two-timescale CE.

required pilot overhead for CE. Unlike in conventional CE where the quasi-static channel is estimated together with the time-varying channel during every small-timescale coherence time, the proposed two-timescale method estimates the quasi-static channel altogether only once.

Fig. 3 demonstrates the difference between the conventional and the two-timescale CE where Rows 1 and 2 are the quasi-static uplink RIS_1-BS CE processes for a conventional and proposed two-timescale schemes respectively, and Row 3 is an uplink time-varying channel UEs-RIS_1. Therefore, the pair consisting of Row 3 and 1 represents the conventional CE scheme while the Row 3 and 2 pair is the two-timescale CE scheme. The red and blue bars represent the small-timescale and large-timescale CE processes respectively. Note that the small-timescale coherence time is a fraction of the large-timescale coherence time. At the first small-timescale coherence time period, all the channels are estimated and thereafter, data is transmitted during the pink region. However, from the second small-timescale coherence time period onwards, there is no CE for the two-timescale quasi-static channel (Row 2) indicated by a gray region (i.e., it assumes the same channel coefficients estimated from the previous small-timescale coherence time period) while the conventional quasi-static channel (Row 1) is estimated every time when the time-varying channel (Row 3) is estimated. Thus, the conventional CE unnecessarily uses higher pilot overhead than the two-timescale CE.

The two-stage CE in this paper is estimated using the proposed DnCNN and Bi-LSTM schemes for the cascaded uplink quasi-static channel RIS_1-BS and time-varying channel $\mathbf{G}$ respectively, as discussed below. But first, the proposed beam squint reduction method is discussed.

1) THE BEAM SQUINT EFFECT REDUCTION SCHEME

Basically, since (4) (i.e., steering vectors) depends on the subcarrier frequency, and for narrowband, subcarriers are not widely spread (i.e., the entire signal bandwidth fits within a small frequency range), the beam squint has less effect and a single frequency can be used to compute the steering vectors of the system. However, for wideband systems, the subcarriers are widely spread and the steering vectors becomes more dependent on the individual subcarrier frequencies. Thus, using a single frequency to compute the steering vectors of the system will significantly affect the CE, since the channels (1), (2) and (3) becomes inaccurate. Fig. 2 shows the correlation of the steering vector with the cascaded channel as (15). In this paper, the SDB method is proposed where the wideband is divided into a number of subbands where each subband is an approximation of the narrowband, thus, reducing the beam squint effect. A center subcarrier frequency of a subband is used to compute the steering vectors of that particular subband. Each subband will therefore have a unique set of steering vectors.

The bandwidth is divided into N_{sb} subbands with the assumption that each subband can be approximated as a narrowband. Thus, a set of steering vectors for the n_{sb}-th subband is computed using a center frequency

$$f_{n_{\text{sb}}} = f_c + \frac{B}{N_{\text{sb}}}\left(n_{\text{sb}} - \frac{N_{\text{sb}}-1}{2}\right) \tag{16}$$

of that particular subband, where $n_{\text{sb}} \in \{0, 1, \ldots, N_{\text{sb}}-1\}$ is the n_{sb}-th subband and each subband consists of $N_{\text{bs/sc}} = \frac{N_{\text{sc}}}{N_{\text{sb}}}$ subcarriers. Thus, a steering vector for an n_{sb}-th subband can be computed as

$$\mathbf{a}_{n_{\text{sb}}}(\vartheta, \psi) = \frac{1}{\sqrt{N_{\text{UPA}}}} e^{-j\frac{d_{e_x}(\vartheta,\psi)}{\lambda_{n_{\text{sb}}}}\mathbf{n}_{\text{UPA}_x}} \otimes e^{-j\frac{d_{e_y}(\psi)}{\lambda_{n_{\text{sb}}}}\mathbf{n}_{\text{UPA}_y}}, \tag{17}$$

where $\lambda_{n_{\text{sb}}} = (c/f_{n_{\text{sb}}}) + f_c = \lambda_c/1 + \frac{f_{n_{\text{sb}}}}{f_c}$,is the wavelength of the n_{sb}-th subband.

Fig. 4 demonstrates how the function $\Upsilon(x, y)$ responds as a result of using $\mathbf{a}_{n_{\text{sb}}}(x, y)$ instead of $\mathbf{a}(x, y)$ for $N_{\text{sb}} = 1$, $N_{\text{sb}} = 3$, $N_{\text{sb}} = 6$ and $N_{\text{sb}} = 10$. The total number of subcarriers is set to $N_{\text{sc}} = 128$ and the bandwidth $B =$ 500MHz. When the bandwidth is not divide (i.e., $N_{\text{sb}} =$ 1), the proposed SDB scheme uses the carrier frequency f_c as a frequency of the whole bandwidth (i.e., $f_{n_{\text{sb}}} = f_c$). Thus, the angular domain index of the false angle for all subcarriers varies from the subcarrier 1's angular domain index to the subcarrier 128's angular domain index. It can be concluded that as the number of subbands increases, the angular domain index spread of the false angles for the n_{sb}-th subband reduces. That is, the angular domain index of the false angle of the last subcarrier within the n_{sb}-th subband gets closer and closer to the reference (red dashed line) as the value of N_{sb} increases. For example, when $N_{\text{sb}} = 3$ ($n_{\text{sb}} \in \{0, 1, 2\}$), the number of subcarriers within the n_{sb}-th subband will be

$$N_{\text{bs/sc}} = \frac{N_{\text{sc}}}{N_{\text{sb}}} = \frac{128}{3} \approx 42. \tag{18}$$

In this case, the angular domain index spread of the false angles for all subcarriers within the n_{sb}-th subband is shorter than that of $N_{\text{sb}} = 1$. Similar explanation can be given for the $N_{\text{sb}} = 6$, $N_{\text{sb}} = 10$ and $N_{\text{sb}} > 10$ cases.

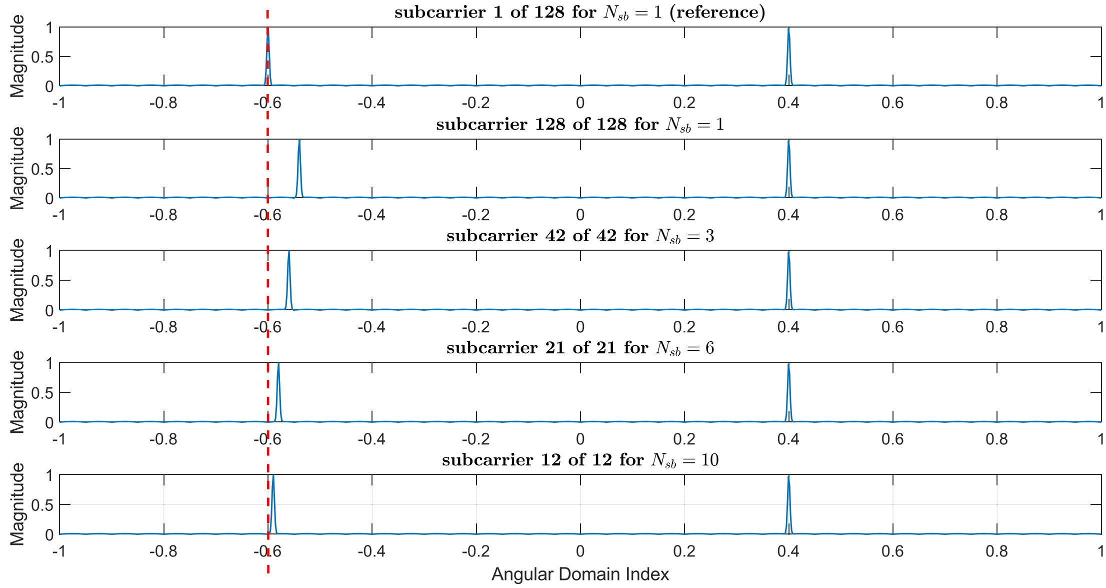

FIGURE 4. **The effect of the proposed SDB scheme on the mutual correlation function $\Upsilon(x, y)$ for various values of N_{sb} at the n_{sb}-th subband.**

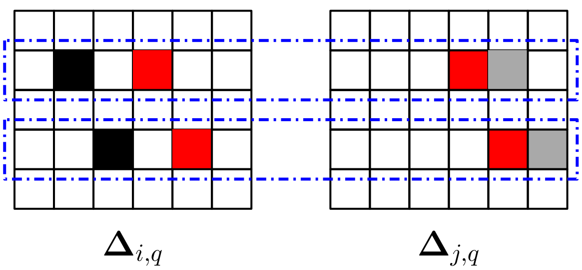

FIGURE 5. **The double-structured sparsity of the cascaded channels for UE$_i$ and UE$_j$ at the q-th time slot.**

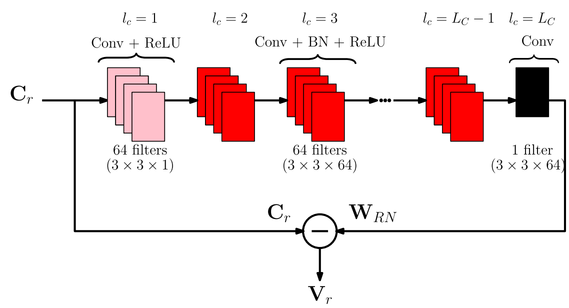

FIGURE 6. **The proposed DnCNN architecture.**

In essence, when the bandwidth is divided, the factor $\frac{f_{n_{sc}}}{f_c}$ in the spatial angles $(1 + \frac{f_{n_{sc}}}{f_c})sin(\psi)$ and $(1 + \frac{f_{n_{sc}}}{f_c})sin(\vartheta)cos(\psi)$ reduces. And if the value of N_{sb} becomes large enough, each n_{sb}-th subband can be assumed as a narrowband and the beam squint effect will not have effect.

2) STAGE 1: CASCADED QUASI-STATIC CE

The multi-RIS-assisted channel model in (1) exhibits a double-structured sparsity property as demonstrated in Fig. 5. That is, the cascaded channel has a full sparsity in the RIS$_1$-BS cascaded channel and a partial sparsity in the $\mathbf{G}$ (UEs-RIS$_1$) channel [36]. In Fig. 5, the columns of channel matrices $\Delta_{i,q}$ and $\Delta_{j,q}$ for the i-th and j-th UEs represent the paths of the time varying-channel $\mathbf{g}_n$ while rows are the paths for the cascaded channel RIS$_1$-BS, and the non-empty entries are shown by the shaded squares. Note that all the non-entries of $\Delta_{i,q}$ share rows with all the non-entries of $\Delta_{j,q}$. This is a demonstration of the full sparsity of the cascaded quasi-static channel RIS$_1$-BS between any i-th and j-th UEs. For the partial sparsity between any i-th and j-th UEs, observe that some non-entries of $\Delta_{i,q}$ and $\Delta_{j,q}$ share columns (indicated by the red squares) and some do not (indicated by the black and gray squares). The time-varying channel between UEs and RIS$_1$ exhibits this partial sparsity property.

Therefore, the quasi-static cascaded channel RIS$_1$-BS is estimated by taking advantage of the full sparsity property. For this purpose, the problem is viewed as a denoising problem and the DnCNN architecture is proposed as shown in Fig. 6. The proposed DnCNN scheme learns the error residual than the actual channel coefficient. This approach improves the accuracy of the channel estimates and increases the training speed [37]. To further increase the training speed of the proposed DnCNN, the BN for regularization is incorporated. And since the DnCNN is based on the CNN, to reduce the CE complexity, it detects smaller, meaningful features of the cascaded Δ_n for processing [38].

As shown in Fig. 6, the proposed DnCNN scheme has L_C convolutional layers with each layer comprising of c_{l_c} different filters of size $D_{l_c,x} \times D_{l_c,y} \times D_{l_c,z}$, where $l_c \in \{1, 2, \ldots, L_C\}$ is the l_c-th convolutional layer. The first convolutional layer $l_c = 1$ (the stacks of pink rectangles) has $c_1 = 64$ filters of size $(3 \times 3 \times 1)$ and it is followed by a rectified linear unit (ReLU) while the last convolutional layer $l_c = L_C$ (black rectangle) consists of a single filter ($c_{L_C} = 1$) of size $(3 \times 3 \times 64)$ for constructing the output. The convolutional layers between the first and the last convolutional layers are identical (the stacks of red rectangles) and each consists of 64 filters of size $(3 \times 3 \times 64)$ followed by the BN and ReLu operations.

Therefore, since the cascaded uplink quasi-static channel RIS_1-BS exhibits the full sparsity property as previously discussed, the DnCNN takes in, as an input, the $M \times 1$ correlation vector

$$\mathbf{c}_r = \sum_{n_{sc}=1}^{N_{sc}} \sum_{n=1}^{N} \mathbf{c}_n[n_{sc}]^H. \tag{19}$$

In this case, $\mathbf{c}_n[n_{sc}] = |\mathbf{s}_n \tilde{\mathbf{Y}}_n[n_{sc}]| \in \mathbb{C}^{1\times M}$, where the collection of the received effective signals for Q time slots of the n_{sc}-th subcarrier is presented as a matrix $\tilde{\mathbf{Y}}_n[n_{sc}] \in \mathbb{C}^{Q\times M}$ and the vector $\mathbf{s}_n = [s_{n,1}, s_{n,2}, \ldots, s_{n,Q}]^H$ consists of symbols transmitted by the n-th UE for Q time slots. To actually use (19) as the input to the DnCNN, it must be formulated as a matrix as

$$\mathbf{C}_r = vec2math(\mathbf{c}_r, [M_x, M_y]) \in \mathbb{C}^{M_x \times M_y} \tag{20}$$

since CNNs can only process a grid-like data [38].

Note that the output of the proposed DnCNN is the residual noise $\mathbf{W}_{RN} \in \mathbb{C}^{M_x\times M_y}$ of the input $\mathbf{C}_r$. Therefore, the matrix associated with the cascaded channel can be computed as

$$\mathbf{V}_r = \mathbf{C}_r - \mathbf{W}_{RN} \in \mathbb{C}^{M_x\times M_y} \tag{21}$$

and its vector form can be presented as

$$\mathbf{v}_r = \sum_{n_{sc}=1}^{N_{sc}} \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{p=1}^{P} |[\Delta_n[n_{sc}]]_{:,n}| \in \mathbb{C}^{1\times M}, \tag{22}$$

and contains all the accumulated cascaded channel amplitudes coming from all the UEs. In essence, the proposed DnCNN estimates the clean received signal from a noisy signal by learning a mapping function $\mathcal{F}(\mathbf{C}_r) = \mathbf{V}_r$.

The proposed DnCNN learns a set of trainable parameters η_1 through a loss function

$$\ell_1(\eta_1) = \frac{1}{2I} \sum_{i=1}^{I} ||\mathcal{W}(\mathbf{C}_{r,i}; \eta_1) - (\mathbf{C}_{r,i} - \mathbf{V}_{r,i})||_F^2, \tag{23}$$

where $\{\mathbf{C}_{r,i}, \mathbf{V}_{r,i}\}_{i=1}^{I}$ are the I training patch pairs.

This proposed DnCNN will perform considerably well in multi-RIS-assisted CE practical scenarios since it is only used for quasi-static CE (where the channel does not rapidly change). That is, in a practical scenario where both the BS and the RISs are stationery, the uplink channels from RIS_1 to the BS do not change rapidly. This means that less channel coefficients are required to be estimated as compared to dynamic channel scenarios due to, for example, high speed movement of UEs (as in the UEs-RIS_1 channel). Thus, with the incorporated convolutional layers which captures spatial correlation of the channels, the CE performance of the proposed DnCNN will perform significantly well in practical quasi-static channels.

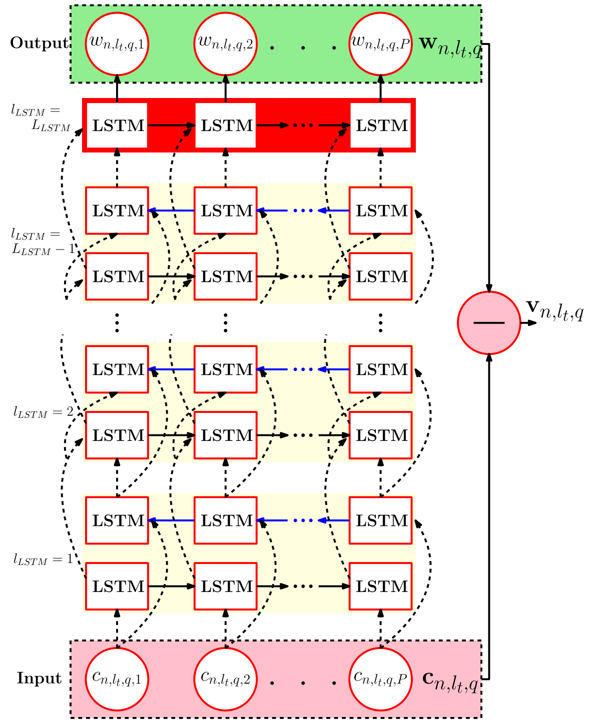

FIGURE 7. The proposed Bi-LSTM architecture.

However, the uplink time-varying CE between the UEs and RIS_1 will be poor in practical scenarios if the DnCNN is used. Therefore, there is a need for a CE scheme which is more robust to time-varying data to estimate the time-varying UEs-RIS_1 channel. Thus, in this paper, the Bi-LSTM scheme is proposed to estimate time-varying UEs-RIS_1 channel as discussed next (stage 2: time-varying CE).

3) STAGE 2: TIME-VARYING CE

Since the channel $\mathbf{G}$ is time-varying, the DnCNN architecture used to estimate the cascaded quasi-static channel RIS_1-BS may not optimally track the fast changing coefficients of channel $\mathbf{G}$. Thus, for this purpose, the Bi-LSTM scheme shown in Fig. 7 is proposed. The proposed Bi-LSTM scheme, unlike the DnCNN scheme, can capture time series data dependencies more effectively [39]. And unlike the LSTM schemes in [14], [40] and [41], which track series data in only forward direction, the proposed Bi-LSTM scheme tracks time series data dependencies in both forward and backward directions, thereby, significantly improving the channel estimates.

The proposed Bi-LSTM has L_{LSTM} layers of LSTM cells where the first $L_{\text{LSTM}} - 1$ layers are the Bi-LSTM layers. That is, the first $L_{\text{LSTM}} - 1$ layers consists of pairs of LSTM networks where one LSTM network propagates data in the forward direction and the other LSTM network propagates data in the backward direction. The last LSTM layer (the red filled rectangle) has only the forward direction LSTM network to compute the output. Note that, just like the proposed DnCNN scheme, the proposed Bi-LSTM scheme learns the error residual. Therefore, to estimate the $\mathbf{g}_n$ channel amplitudes, the pilots are transmitted from UE$_n$ through the multi-RIS network and the signal is received by the multi-antenna BS. The estimation of $\mathbf{g}_n$ by the proposed Bi-LSTM scheme utilises the already estimated quasi-static channel (RIS$_1$-BS) coefficients by the DnCNN. That is, the time-varying channel $\mathbf{g}_n$ is estimated without the ambiguity of the cascaded channel RIS$_1$-BS.

If $l_t \in \{1, 2, \ldots, L_T\}$ is the l_t-th cascaded path of channel RIS$_1$-BS where L_T represents the total number of paths of channel RIS$_1$-BS, the time-varying channel $\mathbf{g}_n$ can be estimated by the proposed Bi-LSTM scheme by taking in as an input, a $P \times 1$ correlation vector

$$\mathbf{c}_{n,l_t,q} = \sum_{n_{\mathrm{sc}}=1}^{N_{\mathrm{sc}}} |\tilde{\Theta}_{n,q}[\tilde{\mathbf{Y}}[n_{\mathrm{sc}}]]_{:,\hat{\Omega}_r(l_t)}|. \tag{24}$$

If the estimation of vector $\mathbf{v}_r$ by the DnCNN is represented as a vector $\hat{\mathbf{v}}_r$, then $\hat{\Omega}_r$ in (24) is a set containing the nonzero entries of $\hat{\mathbf{v}}_r$. Thus, the amplitude of the l_t-th path is presented by $\hat{\Omega}_r(l_t)$ which is one of the nonzero entries of $\hat{\Omega}_r$. As shown in Fig. 7, the output of the proposed Bi-LSTM scheme is the residual noise vector $\mathbf{w}_{n,l_t,q} \in \mathbb{C}^{P\times 1}$ and the vector corresponding to the channel amplitudes of the channel $\mathbf{g}_n$ can be computed as

$$\begin{aligned}\mathbf{v}_{n,l_t,q} &= \mathbf{c}_{n,l_t,q} - \mathbf{w}_{n,l_t,q}\\ &= \sum_{n_{\mathrm{sc}}=1}^{N_{\mathrm{sc}}} |\Delta_{n,q}[n_{\mathrm{sc}}]_{:,\hat{\Omega}_r(l_t)}| \in \mathbb{C}^{P\times 1}.\end{aligned} \tag{25}$$

The proposed Bi-LSTM learns a set of trainable parameters η_2 through a loss function

$$\ell_2(\eta_2) = \frac{1}{2I}\sum_{i=1}^{I} ||\mathcal{W}(\mathbf{c}_{n,l_t,q,i};\eta_2) - (\mathbf{c}_{n,l_t,q,i} - \mathbf{v}_{n,l_t,q,i})||_F^2 \tag{26}$$

where $\{\mathbf{c}_{n,l_t,q,i}, \mathbf{v}_{n,l_t,q,i}\}_{i=1}^{I}$ are the I training patch pairs.

Each LSTM cell has the forget (gray region), input (pink region) and output (yellow region) gates as shown in Fig. 8. The forget gate discards less important information from the previous cell state by computing

$$z_p = \Psi(\varrho_z c_{n,l_t,q,p} + \varpi_z w_{n,l_t,q,(p-1)} + b_z), \tag{27}$$

while the input and the output gate are computing

$$r_p = \Psi(\varrho_r c_{n,l_t,q,p} + \varpi_r w_{n,l_t,q,(p-1)} + b_r) \tag{28}$$

and

$$o_p = \Psi(\varrho_o c_{n,l_t,q,p} + \varpi_o w_{n,l_t,q,(p-1)} + b_o) \tag{29}$$

respectively. The input gates adds new information to the memory $cs_{n,l_t,q,(p-1)}$ while the output gate decides which information should be propagated as a hidden state $w_{n,l_t,q,p}$.

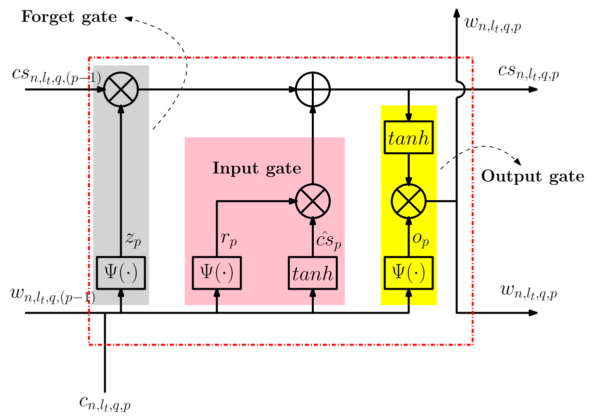

FIGURE 8. Internal structure of an LSTM cell.

TABLE 1. System parameters.

Parameter	Value	Parameter	Value
N	16	N_{sc}	2048
P	256	N_{ant}	4
M	64	B	500 MHz
K	2	f_c	60 GHz

In (27), (28) and (29), ϱ_i and ϖ_i are the weights while b_i is a bias for $i \in \{z, r, o\}$. The subscripts z, r and o correspond to the forget, input and output gates respectively. Therefore, the outputs of an LSTM cell are

$$cs_{n,l_t,q,p} = (z_p \otimes cs_{n,l_t,q,(p-1)}) + (r_p \otimes \hat{cs}_p) \tag{30}$$

and

$$w_{n,l_t,q,p} = tanh(cs_{n,l_t,q,p}) \otimes o_p \tag{31}$$

where $\hat{cs}_p = tanh(\varrho_{cs} c_{n,l_t,q,p} + \varpi_{cs} w_{n,l_t,q,(p-1)} + b_{cs})$ ($\{\varrho_{cs}, \varpi_{cs}\}$ and b_{cs} are weights and bias respectively). An LSTM cell learns all the weights and biases.

Algorithm 1 shows the complete process of the proposed SDB-DnCNN-Bi-LSTM scheme.

IV. SIMULATION RESULTS

In this section, simulation results are provided for the convergence of the proposed scheme and thereafter, a comparative performance of the proposed scheme with schemes in [18] and [20] is provided. For training, the DeepMIMO dataset in [42] is used and an "O1_60" scenario is adopted. It is an outdoor scenario operating at 60GHz. The simulation results are provided for a double-RIS-assisted system without loss of generality and similar procedure can be followed for systems with more than three RISs. Table 1 summarises the system's parameters for simulations.

The total generated data for both training (70%) and validation (30%) is 20 000 input-output dataset pairs.

A. THE CONVERGENCE OF THE PROPOSED DNCNN AND BI-LSTM SCHEMES

The values of L_C and L_{LSTM} in Figs. 6 and 7 respectively are determined by simulating the DnCNN and Bi-LSTM

Algorithm 1 The proposed SDB-DnCNN-Bi-LSTM

Initialisation: N_{sb}, $\tilde{\mathbf{Y}}_n[n_{\mathrm{sc}}]$ $\forall(n_{\mathrm{sc}}, n)$, $\Theta_{n,q}$ $\forall(n, q)$, L_G, L_H, L_{R_j} $\forall(R_j)$, $\hat{\Omega}_r$, $\hat{\Omega}_{n,l_t,q}$ $\forall(n, n_{\mathrm{sc}}, q, l_t)$
$\Delta_{n,q}[n_{sc}] \leftarrow \mathbf{0}_{M\times P}$ $\forall(n_{\mathrm{sc}}, n, q)$
procedure REDUCE THE BEAM SQUINT EFFECT
$\mathbf{a}_{n_{\mathrm{sb}}}(\vartheta, \psi) \xleftarrow[\text{SDB}]{\text{Online}} N_{\mathrm{sb}}$
Return $\mathbf{a}_{n_{\mathrm{sb}}}(\vartheta, \psi)$ $\forall$(RISs, BS)
procedure ESTIMATE Ω_r
$\mathbf{C}_r \leftarrow vec2math(\mathbf{c}_r, [M_x, M_y])$.
$\hat{\mathbf{V}}_r \xleftarrow[\text{DnCNN}]{\text{Online}} \mathbf{C}_r$
$\hat{\mathbf{v}}_r \leftarrow vec(\hat{\mathbf{V}}_r)$
$\hat{\Omega}_r \leftarrow$ INDEXSORTDESCEND $(\hat{\mathbf{v}}_r, L_T)$
Return $\hat{\Omega}_r$
end procedure
procedure ESTIMATE $\Omega_{n,l_t,q}$
$\hat{\mathbf{v}}_{n,l_t,q} \xleftarrow[\text{GRU}]{\text{Online}} \mathbf{c}_{n,l_t,q}$
$\hat{\Omega}_{n,l_t,q} \leftarrow$ INDEXSORTDESCEND $(\hat{\mathbf{v}}_{n,l_t,q}, L_G)$
Return $\hat{\Omega}_{n,l_t,q}$
end procedure
procedure RECONSTRUCT THE CASCADED CHANNEL $\mathbf{U}_{n,q}[n_{\mathrm{sc}}]$
  for $q = 1 : Q$ **do**
    for $n = 1 : N$ **do**
      for $n_{\mathrm{sc}} = 1 : N_{\mathrm{sc}}$ **do**
        for $l_t = 1 : L_T$ **do**
          $\hat{\mathbf{\Delta}}_{n,q}[n_{\mathrm{sc}}]_{(\hat{\Omega}_r, \hat{\Omega}_{n,l_t,q})}$
        end for
        $\hat{\mathbf{U}}_{n,q}[n_{\mathrm{sc}}] = \mathbf{A}_{BS}\hat{\mathbf{\Delta}}_{n,q}[n_{\mathrm{sc}}]\mathbf{A}^H_{RIS_1}$
      end for
    end for
  end for
return $\hat{\mathbf{U}}_{n,q}[n_{\mathrm{sc}}]$ $\forall(n, n_{\mathrm{sc}}, q)$
end procedure

schemes with various L_C and L_{LSTM} values. It can be seen in Figs. 9 and 10 that the DnCNN and Bi-LSTM schemes converge faster for $L_C = 3$ and $L_{\mathrm{LSTM}} = 3$ respectively. The reason for the behavior is that the gradients during learning become very small (vanish) during backpropagation for both the DnCNN and Bi-LSTM schemes with a larger number of layers (in this case larger than $L_C = L_{\mathrm{LSTM}} = 3$) for a given learning rate. This in turn makes it difficult for the network to learn effectively. Similarly, in Fig. 10 when the number of layers becomes smaller, the Bi-LSTM schemes performance degrades. This could be attributed to insufficient representational capacity. Note that for DnCNN, L_C cannot be less than 3. Therefore, for double-RIS-assisted CE, the DnCNN and Bi-LSTM schemes must have three number of convolutional layers and three LSTM layers for optimal performance respectively. This schemes were both trained with a 10^{-6} learning rate.

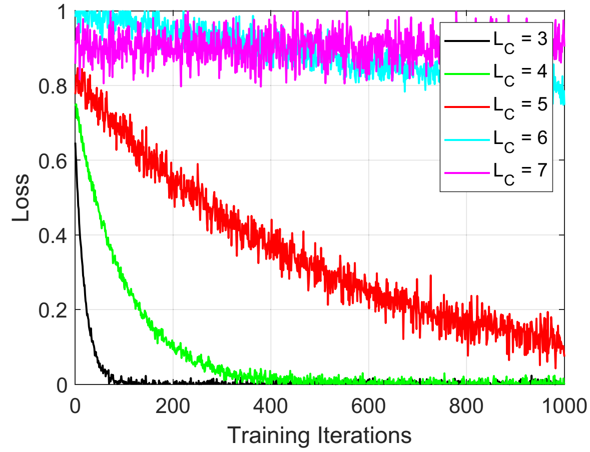

FIGURE 9. The DnCNN Convergence in terms of training loss versus number of training iterations for various number L_C layers.

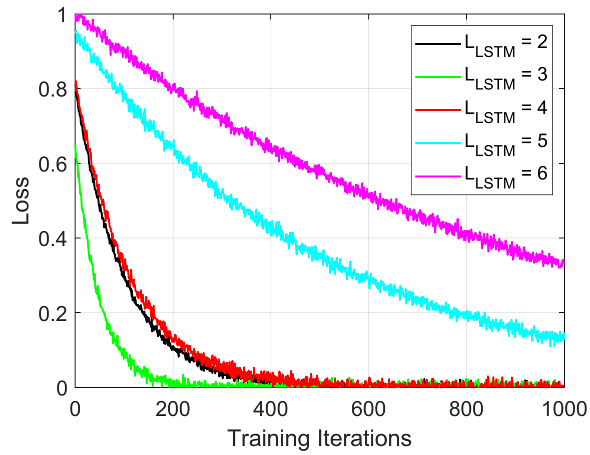

FIGURE 10. The Bi-LSTM Convergence in terms of training loss versus number of training iterations for various number L_{LSTM} layers.

B. THE NMSE COMPARISON

The proposed SDB-DnCNN-Bi-LSTM scheme is benchmarked in terms of NMSE performance with the following beam squint aware CE schemes:

1) *OMP-based Scheme in [18]*. This scheme used a correlation-based method to eliminate the beam squint effect. For CE, the TS-OMP algorithm is used.
2) *Double-structured DnCNN scheme in [20]*. Similar to [18], a correlation-based beam squint elimination method was adopted. However, the DnCNN was used for cascaded CE from UEs to the BS.
3) *The Oracle LS*. This scheme assumes a perfect elimination of the beam squint effect, and provides the lowest bound of CE. The CE is performed with an assumption that the supports are perfectly known.

The NMSE is computed as

$$\mathrm{NMSE} = \frac{1}{N}\sum_{n=1}^{N}\frac{\sum_{n_{\mathrm{sc}}=1}^{N_{\mathrm{sc}}}||\hat{\mathbf{U}}_n[n_{\mathrm{sc}}] - \mathbf{U}_n[n_{\mathrm{sc}}]||_F^2}{\sum_{n_{\mathrm{sc}}=1}^{N_{\mathrm{sc}}}||\mathbf{U}_n[n_{\mathrm{sc}}]||_F^2} \tag{32}$$

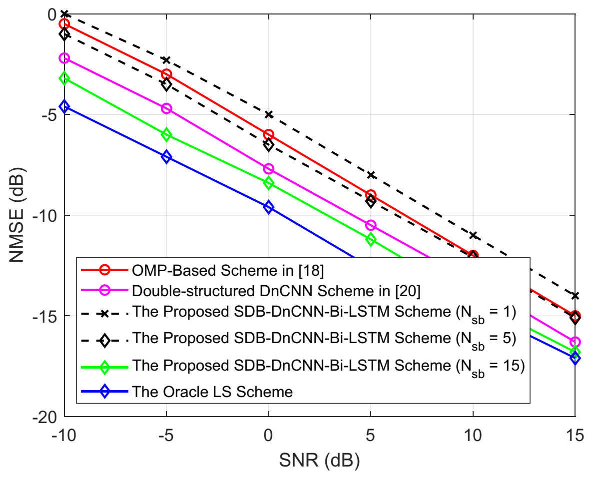

FIGURE 11. NMSE comparison of the proposed SDB-DnCNN-Bi-LSTM scheme (for various values of N_{sb}) with schemes in [18] and [20].

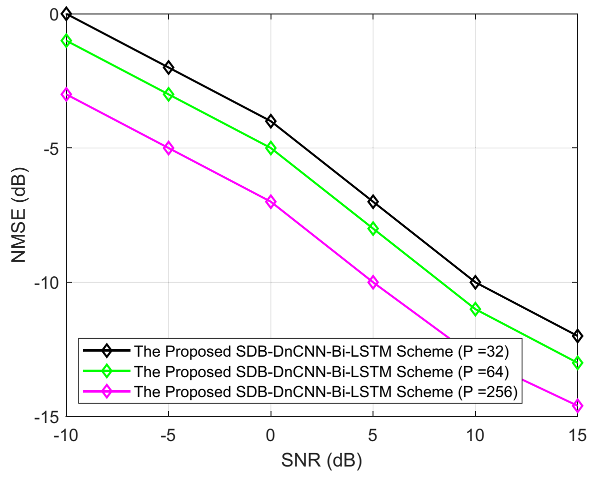

FIGURE 12. NMSE comparison of the proposed SDB-DnCNN-Bi-LSTM scheme with $N_{sb} = 15$, for various values of P.

As it can be observed in Fig. 11 that the NMSE performance of the proposed SDB-DnCNN-Bi-LSTM scheme improves as the value of N_{sb} increases. This is because the steering vectors of the system becomes less dependent on the subcarrier frequencies as the subbands become narrower and narrower. Note that for $N_{\text{sb}} = 1$, the proposed scheme performs worser than both the OMP-based and Double-structured DnCNN schemes mainly because for $N_{\text{sb}} = 1$, the bandwidth is simply not divided and the beam squint effect is severe, affecting the performance of the proposed scheme, but the OMP-based and Double-structured schemes are not affected (i.e., the beam squint elimination scheme for both schemes took effect regardless). For $N_{\text{sb}} = 5$, the proposed scheme is starting to perform better than the OMP-based scheme (only at low SNR up to about SNR of 7 dB) even though the SDB scheme may not be performing better than the correlation-based for eliminating the beam squint. This is because the proposed scheme has a better CE scheme (DnCNN-Bi-LSTM), an ML-based scheme which can handle the sparsity of the cascaded multi-RIS-assisted system than the OMP-based scheme. For about $N_{\text{sb}} = 15$, the proposed scheme performs better than the schemes in [18] and [20] by 3.5 dB and 1.5 dB respectively. Unlike the scheme in [20] which implement the DnCNN also for time-varying channel between the UEs and RIS_1, the proposed scheme utilise the Bi-LSTM scheme which is optimised to deal with time-varying data than the DnCNN. On the other hand, the proposed scheme with $N_{\text{sb}} = 15$ has about 1.5 dB performance gap from the Oracle LS scheme. This gap can be reduced by increasing the value of N_{sb}. Thus, the proposed scheme demonstrates a promising solution for multi-RIS-assisted CE.

Fig. 12 illustrates how the NMSE of CE improves as the number of RIS elements of the system increases for the proposed scheme with $N_{\text{sb}} = 15$ while keeping all parameters in TABLE 1 constant but varying the value of P. When P increases from $P = 32$ to $P = 64$, the NMSE improves by about 1.2 dB. Similarly, the NMSE improves by about 1.8 dB when P increases from $P = 64$ to $P = 256$.

Although the simulation results of the proposed scheme are provided for the double-RIS-assisted channel, as discussed in [23], it is expected that both the complexity and resource consumption will be relatively higher as the number of RISs increases. That is, for the proposed DnCNN scheme, the value of L_C (i.e., the number of convolutional layers) will become relatively higher as the number of RISs increases since more channel coefficients will need to be estimated. On the other hand, the value of the L_{LSTM} only depends on the channel coefficient estimation of the UEs-RIS_1 (i.e., channel $\mathbf{G}$) channel, which will be the same for any number of RISs. Therefore, $L_{\text{LSTM}} = 3$ for any number of RISs in the system.

C. COMPLEXITY ANALYSIS

In this section, the online complexity analysis of the proposed SDB-DnCNN-Bi-LSTM scheme as compared with the schemes in [18] and [20] is provided.

The complexity of each LSTM cell in the proposed Bi-LSTM scheme depends on the forget, input and output gates by (27), (28) and (29) respectively, and candidate state (cell state). Each of these variables (the gates and the cell state) involves two matrix-vector multiplications. Thus, the complexity cost of each LSTM is of order $\mathcal{O}(4(N_{inp}N_{Hidd} + N_{hidd}^2))$, where N_{inp} and N_{hidd} are the cell input size and number of hidden units. And since there are P LSTM cells per layer ($2P$ LSTM cells within each Bi-LSTM layer), the total complexity cost of the proposed Bi-LSTM scheme is $\mathcal{O}(L_{LSTM} \times 2P \times 4(N_{inp}N_{Hidd} + N_{hidd}^2)) = \mathcal{O}(8L_{LSTM}P(N_{inp}N_{Hidd} + N_{hidd}^2))$.

On the other hand, for the DnCNN with L_C number of convolutional layers, the complexity cost order is given by $\mathcal{O}(\sum_{l_c=1}^{L_C} D_{l_c,x}D_{l_c,y}D_{l_c,z}c_{l_c})$. Therefore, for the Q time slots and N_{sb} subbands, the overall complexity of the proposed SDB-DnCNN-Bi-LSTM scheme is $\mathcal{O}(QN_{\text{sb}}(8L_{LSTM}P(N_{inp}N_{Hidd} + N_{hidd}^2) + \sum_{l_c=1}^{L_C} D_{l_c,x}D_{l_c,y}D_{l_c,z}c_{l_c}))$.

As in [43], the complexity cost of the OMP-based scheme in [18] is $\mathcal{O}(2MN_{ant}N_{atom})$, where N_{atom} is the number of atoms of the cascaded channel matrix, and complexity cost of the scheme in [20] is $\mathcal{O}(2\sum_{l_c=1}^{L_C} D_{l_c,x}D_{l_c,y}D_{l_c,z}c_{l_c})$.

It therefore, becomes clear that our proposed scheme traded off the computational complexity for high CE accuracy as both [18] and [20] have relatively lower complexity.

V. CONCLUSION

This paper considered a wideband multi-RIS-assisted mmWave massive-MIMO OFDM system and an SDB-DnCNN-Bi-LSTM scheme was proposed for uplink CE. To reduce the required pilot overhead of CE, a two-timescale scheme was adopted, and the beam squint effect was combated by dividing the system's bandwidth into N_{sb} subbands. It was proven that the beam squint effect decreases with an increase of N_{sb}. For the uplink cascaded quasi-static channel between RIS_1 and BS, the DnCNN scheme was used for CE. The BN and residual learning were incorporated in the proposed DnCNN scheme to increasing the learning speed. Thereafter, without the ambiguity of the RIS_1-BS channel, the uplink time-varying channel between the UEs and the RIS_1 was estimated using the Bi-LSTM scheme, which is capable of tracking time-varying data in both forward and backward directions, resulting in more accurate channel estimates than a basic LSTM scheme which only tracks time-varying data in only forward direction.

As illustrated in Fig. 11, it can be inferred that the proposed SDB scheme is effective in reducing the beam squint effect, thereby, improving the CE accuracy. However, if CE is performed without deploying the SDB scheme (i.e., SDB scheme with $N_{sb} = 1$, thereby ignoring the beam squint effect), the proposed DnCNN-Bi-LSTM scheme still exhibits enhanced performance which is slightly lower than the case where beam squint is considered as documented in [18] and [20], but with lower performance gain of about 0.5 dB and 2.5 dB respectively. In essence, the proposed SDB-DnCNN-Bi-LSTM scheme shows improved performance in compassion with other schemes considered in this paper. For future work, the proposed scheme can be tested for other system models and parameter settings.

REFERENCES

[1] B. G. Kashyap et al., "A reconfigurable intelligent surface for 5G wireless communication applications," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, Singapore, 2021, pp. 111–112, doi: 10.1109/APS/URSI47566.2021.9704068.

[2] S. Kisseleff, W. A. Martins, H. Al-Hraishawi, S. Chatzinotas, and B. Ottersten, "Reconfigurable intelligent surfaces for smart cities: Research challenges and opportunities," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1781–1797, 2020, doi: 10.1109/OJCOMS.2020.3036839.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013, doi: 10.1109/TCOMM.2013.020413.110848.

[4] M. Jafri, S. Srivastava, P. Kumar, A. K. Jagannatham, and L. Hanzo, "Cooperative hybrid beamforming for the mitigation of realistic asynchronous interference in cell-free mmWave MIMO networks," *IEEE Trans. Commun.*, vol. 72, no. 11, pp. 6737–6751, Nov. 2024, doi: 10.1109/TCOMM.2024.3403474.

[5] M. Jafri, S. Srivastava, N. K. D. Venkategowda, A. K. Jagannatham, and L. Hanzo, "Cooperative hybrid transmit beamforming in cell-free mmWave MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 5, pp. 6023–6038, May 2023, doi: 10.1109/TVT.2022.3229832.

[6] G. Sun, X. Yi, W. Wang, W. Xu, and S. Jin, "Hybrid beamforming for millimeter-wave massive grant-free transmission," *IEEE Trans. Commun.*, vol. 73, no. 3, pp. 2061–2076, Mar. 2025, doi: 10.1109/TCOMM.2024.3455241.

[7] H. A. Le, T. V. Chien, and W. Choi, "Graph neural network based active and passive beamforming for distributed STAR-RIS-assisted multi-user MISO systems," *IEEE Trans. Commun.*, early access, Apr. 7, 2025, doi: 10.1109/TCOMM.2025.3558504.

[8] C. Qing, W. Hu, Z. Liu, G. Ling, X. Cai, and P. Du, "Sensing-aided channel estimation in OFDM systems by leveraging communication echoes," *IEEE Internet Things J.*, vol. 11, no. 23, pp. 38023–38039, Dec. 2024, doi: 10.1109/JIOT.2024.3443640.

[9] C. Qing, Z. Liu, W. Hu, Y. Zhang, X. Cai, and P. Du, "LoS sensing-based channel estimation in UAV-assisted OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 5, pp. 1320–1324, May 2024, doi: 10.1109/LWC.2024.3368652.

[10] C. Qing, Z. Liu, G. Ling, W. Hu, and P. Du, "Channel estimation in OTFS systems by leveraging differential modulation," *IEEE Trans. Veh. Technol.*, vol. 74, no. 5, pp. 6907–6918, May 2025, doi: 10.1109/TVT.2024.3522940.

[11] X. Su, R. He, P. Zhang, and B. Ai, "Generalized approximating message passing based channel estimation for RIS-aided THz communications with beam split," in *Proc. IEEE 99th Veh. Technol. Conf.*, Singapore, 2024, pp. 1–5, doi: 10.1109/VTC2024-Spring62846.2024.10683198.

[12] J. Wu, S. Kim, and B. Shim, "Parametric sparse channel estimation for RIS-assisted terahertz systems," *IEEE Trans. Commun.*, vol. 71, no. 9, pp. 5503–5518, Sep. 2023, doi: 10.1109/TCOMM.2023.328575.

[13] W. Yang, R. Liu, W. Cai, M. Li, and Q. Liu, "Channel estimation and pilot allocation for practical RIS-aided wideband OFDMA systems," in *Proc. IEEE Int. Conf. Commun.*, Rome, Italy, 2023, pp. 4347–4352, doi: 10.1109/ICC45041.2023.10279798.

[14] C. Nguyen, T. M. Hoang, and A. A. Cheema, "Channel estimation using CNN-LSTM in RIS-NOMA assisted 6G network," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 1, pp. 43–60, 2023, doi: 10.1109/TMLCN.2023.3278232.

[15] K. Kim and C. S. Hong, "Pruned autoencoder based mmWave channel estimation in RIS-assisted wireless networks," in *Proc. 23rd Asia-Pacific Netw. Oper. Manag. Symp. (APNOMS)*, Takamatsu, Japan, 2022, pp. 1–4, doi: 10.23919/APNOMS56106.2022.9919942.

[16] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "RIS-aided mmWave MIMO channel estimation using deep learning and compressive sensing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3503–3521, May 2023, doi: 10.1109/TWC.2022.3219140.

[17] B. Wang et al., "Spatial-wideband effect in massive MIMO with application in mmWave systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 134–141, Dec. 2018, doi: 10.1109/MCOM.2018.1701051.

[18] S. Ma, W. Shen, J. An, and L. Hanzo, "Wideband channel estimation for IRS-aided systems in the face of beam squint," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6240–6253, Oct. 2021, doi: 10.1109/TWC.2021.3072694.

[19] H. Ozen, O. Yilmaz, and G. M. Guvensen, "Beam-squint-aware channel estimation for dual-wideband UPA-type RIS-aided massive MIMO," *IEEE Commun. Lett.*, vol. 28, no. 10, pp. 2367–2371, Oct. 2024, doi: 10.1109/LCOMM.2024.3446320.

[20] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Deep-learning based channel estimation for RIS-aided mmWave systems with beam squint," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2022, pp. 1269–1275, doi: 10.1109/ICC45855.2022.9839142.

[21] S. Yang, C. Xie, W. Lyu, B. Ning, Z. Zhang, and C. Yuen, "Near-field channel estimation for extremely large-scale reconfigurable intelligent surface (XL-RIS)-aided wideband mmWave systems," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 6, pp. 1567–1582, Jun. 2024, doi: 10.1109/JSAC.2024.3389120.

[22] Z. Li, Z. Gao, and T. Li, "Sensing user's channel and location with terahertz extra-large reconfigurable intelligent surface under hybrid-field beam squint effect," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 4, pp. 893–911, Jul. 2023, doi: 10.1109/JSTSP.2023.3278942.

[23] Y. Wang, W. Zhang, Y. Chen, C.-X. Wang, and J. Sun, "Novel multiple RIS-assisted communications for 6G networks," *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1413–1417, Jun. 2022, doi: 10.1109/LCOMM.2022.3163587.

[24] M. Liu, X. Li, B. Ning, C. Huang, S. Sun, and C. Yuen, "Deep learning-based channel estimation for double-RIS aided massive MIMO system," *IEEE Commun. Lett.*, vol. 12, no. 1, pp. 70–74, Jan. 2023, doi: 10.1109/LWC.2022.3217294.

[25] T. C. Rapudu and O. O. Oyerinde, "Machine learning-based channel estimation for multi-RIS-assisted mmWave massive-MIMO OFDM system in a dynamic environment," *IEEE Trans. Wireless Commun.*, early access, Mar. 7, 2025, doi: 10.1109/TWC.2025.3546671.

[26] T. C. Rapudu and O. O. Oyerinde, "Joint CFO and channel estimation for multi-RIS-assisted mmWave massive-MIMO OFDM system in a dynamic environment," *IEEE Trans. Veh. Technol.*, to be published.

[27] S. Yang, W. Lyu, Y. Xiu, Z. Zhang, and C. Yuen, "Active 3D double-RIS-aided multi-user communications: Two-timescale-based separate channel estimation via Bayesian learning," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3605–3620, Jun. 2023, doi: 10.1109/TCOMM.2023.3265115.

[28] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6853–6869, Oct. 2023, doi: 10.1109/TWC.2023.3246264.

[29] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Design and optimization on training sequence for mmWave communications: A new approach for sparse channel estimation in massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1486–1497, Jul. 2017, doi: 10.1109/JSAC.2017.2698978.

[30] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, Jul. 2017, doi: 10.1109/JSAC.2017.2699338.

[31] B. P. S. Sahoo, S. Narang, P. Jain, D. Kumar, and S. K. Vankayala, "Mobility-aware hybrid channel estimation scheme for RIS-aided millimeter-wave systems," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol.*, 2022, pp. 1–5, doi: 10.1109/CONECCT55679.2022.9865721.

[32] Z. Huang, C. Liu, Y. Song, H. Wang, H. Zhou, and S. Shen, "Manifold optimization empowered two-timescale channel estimation for RIS-assisted systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15034–15048, Oct. 2024, doi: 10.1109/TVT.2024.3405955.

[33] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021, doi: 10.1109/TCOMM.2021.3072729.

[34] J. Sun, L. Zhao, C. Liu, and Y. Gao, "Sensing-based two-timescale channel estimation for ris-assisted hybrid millimeter wave systems," in *Proc. Int. Symp. Wireless Commun. Syst.*, Hangzhou, China, 2022, pp. 1–6, doi: 10.1109/ISWCS56560.2022.9940404.

[35] K. Zhi et al., "Two-timescale design for reconfigurable intelligent surface-aided massive MIMO systems with imperfect CSI," *IEEE Trans. Inf. Theory*, vol. 69, no. 5, pp. 3001–3033, May 2023, doi: 10.1109/TIT.2022.3227538.

[36] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications—Part II: An improved solution based on double-structured sparsity," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1403–1407, May 2021, doi: 10.1109/LCOMM.2021.3052787.

[37] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, pp. 3142–3155, 2017, doi: 10.1109/TIP.2017.2662206.

[38] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. Accessed: Nov. 15, 2023. [Online]. Available: https://www.deeplearningbook.org

[39] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [review article]," *IEEE Comput. Intell. Mag.*, vol. 13, no. 3, pp. 55–75, Aug. 2018, doi: 10.1109/MCI.2018.2840738.

[40] Z. Mai and Y. Chen, "A CSI entropy-based estimation algorithm and modified recurrent neural networks for time-varying non-stationary channels," *IEEE Trans. Veh. Technol.*, vol. 73, no. 5, pp. 6614–6625, May 2024, doi: 10.1109/TVT.2023.3339706.

[41] X. Han, Z. Jiao, P. Liang, and J. Fan, "A joint time-varying channel estimation based on compressive sensing and LSTM," in *Proc. IEEE 95th Veh. Technol. Conf.*, Helsinki, Finland, 2022, pp. 1–5, doi: 10.1109/VTC2022-Spring54318.2022.9860883.

[42] A. Alkhateeb, "DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications," in *Proc. Inf. Theory Appl. Workshop (ITA)*, 2019, pp. 1–8.

[43] Z. He, L. Zhou, Y. Yang, Y. Chen, X. Ling, and C. Liu, "Compressive sensing-based channel estimation for FBMC-OQAM system under doubly selective channels," *IEEE Access*, vol. 7, pp. 51150–51158, 2019, doi: 10.1109/ACCESS.2019.2898896.

THABANG C. RAPUDU received the B.Eng. and M.Eng. degrees in electrical and electronic engineering from the University of Johannesburg, South Africa, in 2022 and 2023, respectively. He is currently pursuing the Ph.D. degree in electrical and information engineering with the University of the Witwatersrand, South Africa, where he is currently a Sessional Lecturer for courses on Microprocessors and Signals and Systems. His current research interests are, but not limited to signal processing, information theory, 5G/6G technologies, and the application of machine learning techniques for wireless communication systems. He has been a member of the Golden Key International Honour Society since 2023. He has done reviews for IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE Wireless Africa Conference 2025.

OLUTAYO O. OYERINDE (Senior Member, IEEE) received the Ph.D. degree in electronic engineering from the University of Kwazulu-Natal, Durban, South Africa, in 2011. He has been with the School of Electrical and Information Engineering, the University of the Witwatersrand, Johannesburg, South Africa, since 2013, where he is currently an Associate Professor. His current research interests include wireless communications with specific interests in 5G and beyond 5G technologies, and signal processing techniques for wireless communication systems. He is a National Research Foundation C2-Rated Scientist, a Registered Professional Engineer with the Engineering Council of South Africa, a Registered Engineer with COREN, a Senior Member of the Institute of Electrical and Electronics Engineer, and a Corporate Member of NSE amongst others. He is an Associate Editor for IEEE ACCESS, IEEE COMMUNICATIONS LETTERS, the *Journal of the Franklin Institute*, and the *Franklin Open*. He is also an Editorial Member of the Editorial Board of *Wireless Communications and Control*.