Machine Learning Based Estimation and Prediction of Millimeter-Wave Wireless Channels

Emma Olivia
Dept of Communication
Harvard University
Cambridge, Massachusetts, USA
XXXXXXXXXX@gmail.com

Elizabeth Mia
Dept of Communication
Harvard University
Cambridge, Massachusetts, USA
XXXXXXXXX@gmail.com

Abstract*- The massive growth of the wireless communication technology, especially in 5G and beyond, which is used for the efficient and channel modelling for estimation and prediction. The Channel Modeling, works on the statistical approach, it is hard to analyse the dynamic and complex behavior of the channel modelling. Millimeter wave (mmwave) channel modelling operates from the 30-300 GHZ in high bandwidth, highly efficient and high data rate as well as robustness. However, mmWave signals face challenges such as low SNR (signal to noise ratio), Impact on Near and Far field model, robustness issue, Lack of signal detection, inefficient training. To address these challenges, using score based generative model for robustness in channel estimation, MIMO-MRC (Multiple input and multiple output Maximal ratio combining) for High SNR and federated based learning is used for training, using Coordinated multiple point (CoMP) Model for signal detection, Fast alignment algorithm provides far and near field model to execute accurate value. The focus on the channel estimation and prediction, mmwave channel model use of machine learning techniques. Mmwave channel modelling is used to estimate and predict the channel. The propose paper offers some advantages to handling the issues in the communication system such as optimising beamforming, Estimation of CSI (Channel State Information). The proposed model result as Robust, scalability and highly efficient for future wireless network technology when combining these techniques with data. The results shows that our proposed method outperforms significantly mmwave channel prediction and estimation used metrics as BER (Bit Error Rate), Spectral efficiency, Transmission rate, spatial correlation, MSE (Mean Squared Error), SNR.

Index Terms*- *High SNR, Near and Far field Model, MIMO-MRC, Coordinated multiple point (CoMP), Fast alignment algorithm, Channel estimation and prediction.

I. INTRODUCTION

The Mmwave in Multiple-input Multiple-output that led to the fifth-generation mobile communication systems by analysing the massive connectivity high data rate and ultra reliable low-latency communications [1]. The massive mmwave Multiple-input Multiple-output (MIMO) technology can used to increase data transmission rate with higher bandwidth and higher spectral efficiency, emerging as an essential technology for the sixth generation (6G) wireless communication [2]. ML (Machine learning) is used to improve accuracy of detection in physical-layer and new and better waveforms, and reduce complexity of specific parts of receiver algorithms [3]. Reconfigurable intelligent surface (RIS) is an important technique, from Multiple-input Multiple-output (MIMO), mmwave, and provide communications, for future networks (6G) [4]. A Deep unfolding Network is proposed to reduce the computational time in a practical application and used to improve the robustness The constructed deep neural network can be reconstructed that executes faster and more accurate in the interpretable structure of few layers than the iterative algorithm [5]. The probability of transition on Angle-of-Arrival (AoA) and Angle-of-Departure (AoD), a semi-exhaustive search algorithm for beam alignment [6]. mmWave communications are considered a high-potential solution that can improve available bandwidth and spectral efficiency. The frequency band for mmwave ranges from 30 GHz to 300 GHz and supports communication with high transmission rates and ultra-low delay [7]. The simple channel estimation techniques, like LS estimators, which may not execute high-quality channel estimates and low SNR [8]. ML techniques such as Neural Networks (NNs), is used to acquire statistical channel models that tackle the challenges of conventional channel modeling systems [9]. Channel Estimation (CE) is an essential process in modern wireless communication systems, enabling accurate signal detection and robust system optimization [10]. The traditional LLMSE and LS estimation methods are compared to the DNN estimator performance in terms of Bit error rate versus SNR and it can be estimated as channel errors [11]. Channel estimation schemes that accomplish high level accuracy compared to MMSE estimator, while also reducing complexity and enhanced robustness to the imperfect knowledge of channel statistics [12]. The adaptive filtering that provides advancement in mmwave communication technologies and used to enhances the estimation channel performance [13]. The Deep learning-based estimation have been utilizing to reduces the issues and increase the traditional algorithm performance such as signal detection channel estimation and prediction and end-to-end transceiver design [14]. The high directivity is achieved in massive MIMO systems, which often require the need for highlighting the efficient implementation and more complex signal processing techniques [15]. Loss of valuable information about the estimated channel inevitably impacts the estimation performance, particularly at lower SNR regions where the channel noise is usually severe [16]. For Far-field channel estimation, channel sparsity is considered as angle domain, where signals are used to pointed in a particular direction. Whereas, near-field channel estimation considers aperture arrays will experience spherical wavefronts and the channel sparsity is in polar domain [17]. The RIS channel estimation as a maximum likelihood (ML) problem and uses an expectation maximization (EM) algorithm for accurate estimation with reduced training overhead, demonstrating significant improvements over current approaches [18]. The antenna array can be categorized into many sub-arrays, and that can be estimated independently and is used to reduce the complexity. The angles of arrival/departure (AoAs/AoDs) are based on gradient method. The results show off-grid errors. [19]. CSI is important for ensuring reliable signal transmission and reception MIMO systems and is primarily used to design the efficient beamforming techniques [20]. To overcome from these challenges, the proposed model presents MIMO-MRC model, Fast alignment algorithm, federated training, DE-MS QP and CoMP model.

A. *Motivation & Objectives*

The present issues are Low performance of SNR, Impact on Near field and Far field model, Challenging in Robustness, decreasing training, Lack of signal detection

- **Low performance of SNR:** However, the existing method to improve the SNR (signal Noise Ratio). The low SNR will result high noise than the signal.
- **Impact on Near and Far field Model:** Nevertheless, the 3D localization in RIS panel is feasible with near field model and not feasible with far field model. It is essential RIS single panel is effective with Far Field model also.
- **Challenging in Robustness:** However, there is no robustness to estimate the channel in the multi sub band quasi perfect and data embedded (DE-MS QP) techniques. The multi sub band quasi perfect and data embedded model which produces compressed sensing and communication and is used to estimate the channel but need to improve robustness
- **Decreasing Training:** The channel prediction, this research paper needs to train the model in advance. However, the training is low.
- **Lack of Signal Detection:** The signal detection between transmitter and receiver is important for channel estimation and prediction. However, SC attention Network are used to estimate the channel under noise condition but it has to improve the signal detection

The primary objective of this research is mmwave channel estimation and prediction based on machine learning particular objectives of this study are as follows,

- To Develop High SNR where the noise is stronger than the signal.
- To Utilize novel methods, for training the federated learning method are executed
- To develop novel method for near and far field model get feasible with RIS technology.
- To Apply novel methods, provides Robustness for channel estimation and prediction
- To Design the model for signal detection between transmitter and receiver

B. *Research Contribution*

The highlights of this research work are illustrated below;

- To improve High SNR, use the Minimally tuned (MIMO-MRC) method.
- To increase training the federated training method is used
- Near and Far field model executed RIS panel where as, the Fast Alignment Algorithm introduced
- To Improve robustness of channel estimation, Multi sub band quasi perfect and data embedded (DE-MS QP) model is introduced.
- To detect the signal between transmitter and receiver, the Coordinated multiple point (CoMP) model is introduced

C. *Paper organization*

The remainder of this research as follows: Provides an explanation of Section II of a survey of previous work. In section III, states that main issue with the current methods. In section IV, describes the system model. Then in Section V presents study approach for suggested model, appropriate diagrams, mathematical representations and pseudocode. In Section VI, the suggested and current methodology is compared and the experimental results are explained. The proposed method conclusion and future work is explained in Section VII.

II. LITERATURE SURVEY

The Channel estimate is a crucial operation that significantly impacts end-to-end system performance used MIMO digital communication. The MIMO, Score based generative model used for Channel Estimation. To improve estimations given measurements of a signal, to determine the gradient of a distribution logarithm, a model is well trained. However, the Score based model have high inference complexity of posterior sampling with Langevin dynamics [21]. The Research paper examines how wobbling affects the Doppler effect of a mmwave wireless channel connecting ground node and hovering RW UAV. The results show that Amplitude and then Frequency of the ACF oscillation in mmwave link are affected by various RW UAV wobbling patterns. Nevertheless, tangential velocity will produce the Doppler effect and the chance of the Line Of Sight (LoS)link will drop when UAV at low altitude [22]. In this Research, they propose Deep Learning for channel estimation in MIMO system. Convolutional Neural network (CNN) is a Channel estimation process where input and output data which can be referred as H neural network (HNN). HNN which is used to generate the channel information of received signal and it also find connection between channel and received data. The Hopefield Neural network (HNN) algorithm gives channel estimation accurately. However, the large amount of data not collected for channel estimation in deep learning [23]. This research paper proposed orthogonal frequency-division multiplexing (OFDM) systems of compressed-sensing-assisted index modulation, termed as OFDM-CSIM, communicating over mmwave channels. The DNN achieve high throughput than the K-nearest algorithm. The Deep Neural Network (DNN) and Sparse Bayesian learning (SBL) provides better performance and accuracy. However, DNN and SBL decrease the complexity of channel estimation by using received signals as the feature set [24]. The support of the interior channel model simulator, which was developed for the generation of grid-wise channel data (Path Grid Data), a marked Additionally, developing grid-wise channel data (Path Grid Data) with a use of the internal channel model simulator (CP SQDSIM). The 3D model that cannot implement all scattering object during the measurement time [25]. The mmwave in multiple-input multiple output (MIMO), significantly reduces number of Radio frequency (RF) chains by using antenna array. The proposed GM-LAMP results better result on the channel estimation. However, the Gaussian mixture LAMP (GM-LAMP) is applied for improve channel estimation [26]. In MIMO system, the “GPODE” is a channel prediction method. The GPODE is a combination of Genetic programming (GP) with higher order differential equation (HODE). The GPODE method gives high accuracy in channel prediction**.** In GPODE method, which is used for channel prediction. For long term prediction, this research uses online and offline learning. However, Online training gives more accuracy than the offline training [27]. Federated Learning for channel estimation and CNN is trained on the dataset of the users. Federated Learning provides channel estimation performance as well as channel prediction performance. However, to develop compression-based techniques for training the data and provide model parameters to reduce the communication overhead [28]. These Research paper presents Deep Neural Network (DNN) which is used for prediction in Angle of Arrival, Angle of Departure. Dynamic window approach (DWA) which is used to estimate location information of user in User equipment (UE), input parameter is well trained “DNN” to optimise the prediction of “AAOA/AAOD” and “EAOA/EAOD”. There is still chance to improve the channel prediction performance using DNN (Deep Neural Network) with AOA and AOD [29]. The Deep learning compressed sensing (DLCS) is analysed to Estimate the channel. The Deep learning quantized phase (DLQP) hybrid precoder design method is used to develop channel estimation**.** A Deployment hybrid precoding neural network (DHPNN) are presented by changing approximation of ideal phase quantization and output as the DHPNN is analog precoding vector matrix. The proposed DLQP method, there is still chance to develop the Channel estimation and hybrid precoding design for wideband Multi-user mmWave massive MIMO transmission adopting deep learning [30]. DNN-based beam training (DBT) schemes are used to estimate the channel Based on the chances vector, the original DBT (ODBT) use DNN to identify beam combination that best matches mmWave channel longest channel path. However, ODBT and EDBT schemes improve DNN framework and preprocessing method for received signal before DNN processing [31].

This Research paper presents Low complexity Machine learning to reduce reference signal (RS) overhead, latency, and power consumption. The result of the proposed system shows that prediction accuracy and spatial correlational. But there is chance to improve the mmwave beam prediction in machine learning based signal RS overhead [32]. In these research paper, proposed efficient channel estimation for the double-IRS aided Multi-user and the MIMO communication system to solve the cascaded CSI in both single double Reflection link. The performance provided joint training reflection design and channel estimation scheme with double IRS, compared to another benchmark scheme. However, the most generic Multi Assisted Multi-user communication system involves multiple paths of signal reflection, requiring additional intricately designed systems that use multi-IRS deployment, joint passive beamforming, and channel estimation [33]. The optimization of wireless channels and enhance the network as one, Reconfigurable intelligent surface (RIS) and assisted wireless system need precise CSI performance. In this paper, data-driven method that takes beam squint into account is generated for predicting RIS-assisted multi-user mmwave complex MIMO systems use the wideband cascaded channels and the minimal training latency. Nevertheless, the standard Compressed Sensing algorithms still need a significant amount of pilot cost in order to confirm estimation accuracy [34]. RIS-assisted Orthogonal frequency division multiplexing (OFDM) and Multi-user multiple-input multiple-output (MIMO) communication systems determine cascaded channels with high dimensionality and advanced statistical analysis. The proposed Super-resolution convolutional neural network (SRCNN) and Denoising convolutional neural network (DnCNN) results good performance as well as accuracy of channel estimation. The intricate Gaussian distribution is not used by the cascaded channel. The ideal Minimum means square error (MMSE) estimator, have several integrals implementation, cannot be derived in this form [35]. TABLE 1 represents summary of existing work.

TABLE 1

SUMMARY OF EXISTING WORKS

References	Objectives	Algorithms or methods used	Limitations
[21]	A Score based generative model for posterior sampling and represents a new research direction for MIMO channel estimation.	Score based Generative model for channel estimation	• Limited by the Score based model have "high inference complexity" of posterior sampling with "Langevin dynamics".
[22]	RW UAV wobbling patterns impact the amplitude and frequency of ACF oscillation in the mmwave RW UAV A2G link.	Doppler Effect for the process of RW UAV.	• The Line of Sight is dropped when UAV is low altitude.
[23]	HNN is used to analyse the channel information of received signal and also find the connection between channel and received signal.	Hopefield Neural network (HNN) algorithm for channel estimation accurately.	• The large amount of data is not collected yet for the channel estimation.
[24]	orthogonal frequency-division multiplexing (OFDM) communicating over mmwave and sparse Bayesian learning (SBL) for accurate channel state information.	OFDM and SBL for channel estimation accurately.	• Complexity is increased in OFDM and SBL model.
[25]	Developing grid-wise channel data (Path Grid Data) with a use of the internal channel model simulator.	Q-D channel model framework for channel characteristics in mmwave.	• The 3D model that cannot implement all scattering object during the measurement time.
[26]	Online and Offline training is used to generate the long-term prediction.	GPODE method is used for training	• Limited by Online training is more accurate than the offline training.
[27]	The model is developed for fast channel estimation using CNN	Federated Learning for channel estimation and channel prediction performance.	• Limited by the model reduce the communication overhead.
[28]	Design to reduce the energy consumption and provide high throughput	ES-MPTCP, energy saving scheduling system	• The model parameter is reducing the Communication overhead
[29]	To develop for high robustness channel estimation.	Deep Neural Network (DNN) is used for prediction of channel.	• There is still chance to develop the channel prediction performance
[30]	The model has better channel estimation	Deep learning quantized phase	• There is still chance to develop the

	performance and has high spectral efficiency with low resolution of phase shifters.	(DLQP) is used for better spectral efficiency	channel estimation and hybrid precoding design
[31]	The model proposed to reduce the beam training overhead and improved signal coverage.	ODBT, used to predict the beam combination. EDBT, used for additional beam training.	• The model takes longer time for beam training.
[32]	The model provides low computational complexity and achieves beam prediction accuracy.	Low complexity Machine learning design for reduction in RS overhead.	• There is chance to improve the mmwave beam prediction in machine learning.
[33]	The effectives of the proposed channel estimation scheme and Joint training reflection design.	Double-IRS aided multi-user MIMO system for maximize the training	• The most generic Multi IRS involves multiple paths of signal reflection, requiring additional intricately designed systems.
[34]	The data driven approach for estimated the wide band channel estimation.	Data-driven cascaded channel estimation for denoising neural network.	• The compressed sensing algorithm needs an amount of pilot in order to confirm the estimation accuracy.
[35]	To improve the features and estimates the channel matrix by using pilot locations.	Super-resolution convolutional neural network (SRCNN) used for accuracy of channel estimation.	• The intricate Gaussian distribution is not used by the cascaded channel.

III. PROBLEM STATEMENT

The numerous existing works and their associated responses are arranged in sequence of publication in this section. Furthermore, this study offers the research solutions for the mentioned issues.

Specific research work & Issues: Authors in [36] research paper focuses their affordability and ease of use, Directional scanning sounding (DSS) and Virtual antenna array (VAA) sounding are two widely used models. DSS are mechanically movable omni directional antenna and VAA is a rotatable directional antenna. A new VAA framework based on directional antennas along with the related beamforming algorithm. Unlike the traditional VAA, it is a solution that can be used for frequency bands and polarizations. This research paper accomplishes High angular resolution for mmwave channel measurement without extending time need for measurement. The article [37] presents in RIS, the spherical wavefront propagation in the subTHz systems Near field. A calculated second-order Fresnel approximation of the Near-field channel model propose Near-field channel estimation and localization (NF-JCEL) model. The orthogonal matching pursuit (OMP) model, channel attenuation coefficients, and the simple one-dimensional search can be used to estimate the UE distance. The NF-JCEL method can achieve higher resolution accuracy when compared to the traditional far field approach. Some of the problem detected in these papers are:

- There are still chances to suppress the side lobes and increase the SNR . In Low SNR, the noise is stronger than Signal.

- However, near and far-field models describe the behavior of electromagnetic fields at different distances from a radiating source, like an antenna. However, 3D localization using a single RIS panel is feasible with the near-field model but not with the far-field model.

According to this study [38], Joint radar sensing and communication (JRC) operate in a both function such as Time-domain duplex (TDD) and Multi sub band quasi perfect (MS-QP-TDD) and MS-QP is introduced in target sensing, it will achieve target range and estimation of velocity. To use analog to digital convertor for the detection in sequence. By extended “MS-QP”, data embedded “MS-QP (DE-MS QP)” waveform is created, producing null frequency point on every sub band which is used for the data transmission. This research proposed “DE-MS-QP” the waveform gives interference free sensing and communication. The author proposes [39] Unmanned aerial vehicle (UAV) mmWave provides high data rate transmission in wireless network. The 3D scattering space, includes 3D velocity, 3D antenna array, and 3D rotation. A UAV-to-Vehicle (U2V) and (ML) integrated mm Wave channel model is then proposed. The “back propagation” established neural network and Generative adversarial network (GAN), derived using enormous ray-tracing (RT) simulations to training data set. The U2V mmWave channel is generated under 28 Ghz. The proposed paper presents [40] RIS is a Reconfigurable Intelligent Surface, energy efficient option used in Wireless Communication Networks. Double-RIS aided MIMO have some challenges where large amount of antenna at base station. Skip-connection attention (SC-attention) network that optimize self-attention layer and improve the channel estimation extremely under noisy environment. Normalized mean square error (NMSE) accuracy performance can be successfully increased with SC-attention networks it provides an accurate channel Estimation.

Several issues identified in this research include:

- However, the multi sub band quasi perfect and data embedded "(DE-MS QP)" waveform design is no robust to estimate the channel.
- However, ML networks in the framework need to trained in advance and also need to be pre-processed.
- However Nevertheless, the SC attention network has to improve the Signal detection. Signal detection is important for the channel estimation and channel prediction

Research solution: To overcome from these issues, in this study proposes minimally tuned MIMO-MRC model exhibits asymptotic (high SNR) reductions in both uplink and downlink scenarios. The hybrid Fast alignment algorithm with Sub array partition framework and Hierarchical compressed sensing results Feasible in RIS panel in both near and Far field channel model and novelty as "Hierarchical Compressed Sensing with Spatial-Temporal Sub-Partitioning for Ultra-Fast mmWave Beam Alignment". The hybrid Support vector machine with Score based generative model and bayes optimization to generate robustness in channel estimation and novelty as "SCOVEM: Score-Based Generative Model-Enhanced Support Vector Machine with Bayesian Optimization for mmWave Fault Diagnosis". Federated learning (FL) based framework for hybrid beamforming, where model trained and performed at the Base Station and collected gradients from the users. The Coordinated multiple point (CoMP) transmission is typically used in ultra-dense SCN for better the performance target sensing. Using "CoMP" and SCN to detect the received signal

IV. PROPOSED METHODOLOGY

The proposed methodology is used to estimate channel prediction and estimation using Machine Learning algorithm. The proposed method overall architecture is shown in (Fig. 1). Here the proposed methodology was detailed discussed below,

- System Model
- Data Pre processing
- Near and Far field model
- Federated Based Training
- Machine Learning based Robust channel estimation
- Score based Generative model

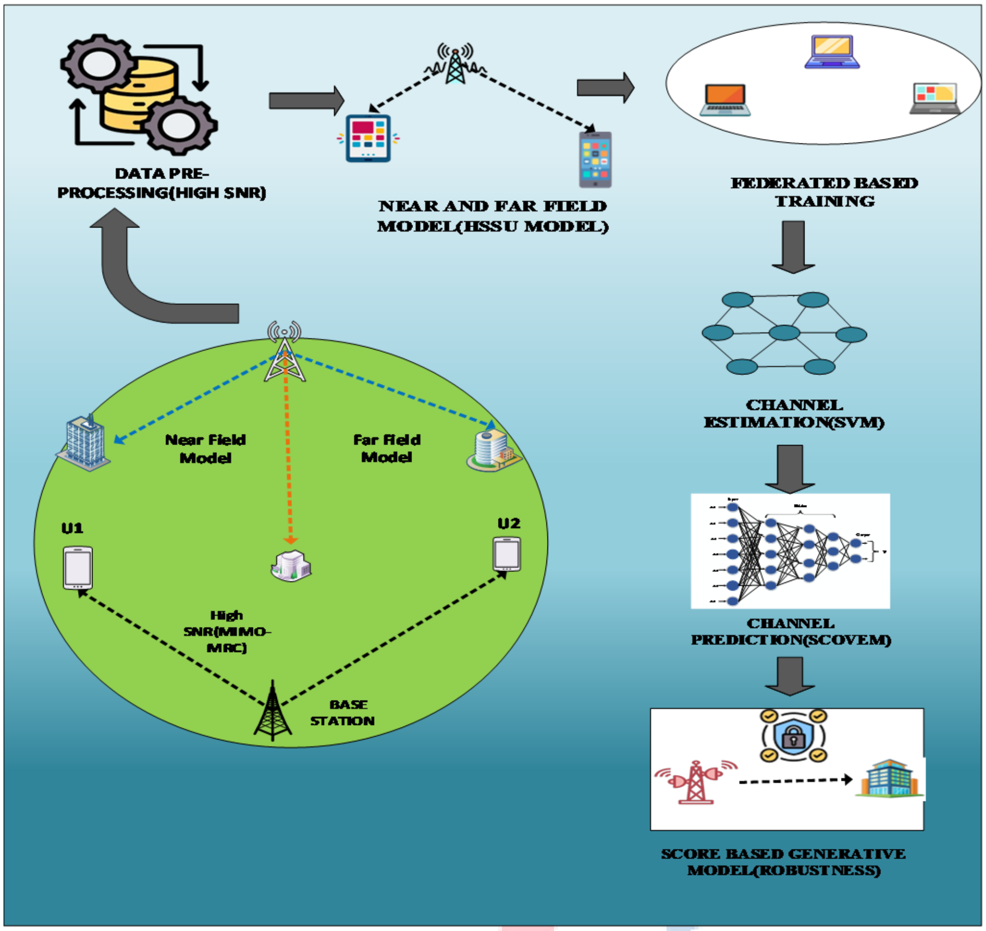

Fig. 1 Overall Proposed Architecture

A. System Model

A 5G-capable wireless communication device can be used by end-user to access the network. UEs are equipped with mmWave transceiver and support beamforming capabilities to communicate with the base stations. It can be static or mobile, and their channel characteristics are influenced by factors such as mobility, location, and the environmental obstacles. A fixed infrastructure node that provides wireless connectivity to UEs within its coverage area. The BS operates in the mmWave frequency band (30–300 GHz) and then uses directional antennas to establish high-data-rate link. It supports only communication and long-range data transmission with low latency. The base station equipped with multiple antennas to enable continuous transmission and reception of multiple data stream. "MIMO-BS" enhances spectral efficiency, and also supports spatial multiplexing, and improves link reliability through diversity gains. It performs a critical role in capturing CSI and provide accurate channel estimation and prediction. The RIS deployed to enhance signal coverage, mitigate blockages, and used to improve channel conditions by intelligently reflecting incident signals toward intended UEs. It supports both near-field and far-field propagation models. In a Channel Environment, the propagation environment such as trees, buildings and user movement. The channel exhibits characteristics like high path loss, multi-path fading, Doppler effects, and spatial-temporal variations, which are analysed and captured through Ray-tracing simulation measurement. In mmwave channel modelling process represents the physical propagation behaviour between the Base station (BS) and user equipment.

$$H = \sqrt{\frac{N_t N_r}{L}} \sum_{i=1}^{L} \alpha_l a_r(\theta_l^r) a_t^H(\theta_l^t) \quad (1)$$

Where H is the "complex channel matrix", N_t and N_r denote the number of transmit and receiver antennas respectively, L is the total number of propagation paths, α_1 represents the complex gain of the l^{th} paths, $a_t^H(\theta_l^t)$ and $a_r(\theta_l^r)$ are the transmit and receive array response vectors corresponding to the AOD and AOA. For a uniform linear array (ULA) configuration, the transmit steering vector is given as

$$\mathbf{a}_t(\theta) = \frac{1}{\sqrt{N_t}}[1, e^{j\frac{2\pi d}{\lambda}\sin(\theta)}, e^{j\frac{2\pi d}{\lambda}2\sin(\theta)}, \dots, e^{j\frac{2\pi d}{\lambda}(N_t-1)\sin(\theta)}]^T \quad (2)$$

$\mathbf{a}_t(\theta)$ is a transmitting array steering vector, Where λ is the carrier wavelength and N_t is the Number of transmit antenna. θ is an angle of departure. $\frac{1}{\sqrt{N_t}}$ is a normalization of power unit and $e^{j(\cdot)}$ is a complex exponential representing phase. d is the antenna spacing and T is a Transpose. The complex gain α_l is represented as

$$\alpha_l = \sqrt{\beta_l} e^{j\phi_l}, \beta_l = \frac{1}{C_0(\frac{d_l}{d_0})^{\gamma}} \quad (3)$$

Where β_l denotes the power gain affected by path loss, C_o is a reference constant at distance $d_o = 1m, \gamma$ is the path loss exponent, d_l is the distance of the l_{th} path, and ϕ_l is the phase uniformly distributed in $[0,2\pi]$. The received signal model is expressed as

$$y = Hx + n \quad (4)$$

Where y denotes the received signal vector, x represents the transmitted signal and $n \sim \mathbb{CN}(0, \sigma^2 I)$ is the additive white Gaussian noise with variance σ^2. H denotes channel effect. The effective post- processing SNR using MIMO-MRC combining is given as

$$SNR_{eff} = \frac{|W^H HF|^2 P_t}{\sigma^2} \quad (5)$$

Where F and w denotes the transmit precoder and receive combiner respectively, and p_t represents the transmitted power. σ^2 is a noise variance. For dynamic environments, the time varying mmwave channel due to mobility is modelled as

$$\mathrm{H}(t) = \sum_{l=1}^{L} \alpha_l e^{j2\pi f_{D,l} t} \mathrm{a}_r(\theta_l^r) \mathrm{a}_t^H(\theta_l^t) \quad (6)$$

Where $f_{D,l} = \frac{v}{\lambda}\cos(\theta_v - \theta_l^r)$ represents the doppler frequency shift for the l^{th}path, with v being the velocity of the UE and θ_v being the direction of motion. Multiple rays are grouped into spatial clusters, and the overall clustered channel is defined as

$$\mathrm{H} = \sqrt{\frac{N_t N_r}{N_c N_p}} \sum_{c=1}^{N_c} \alpha_{c,p} \mathrm{a}_r(\theta_{c,p}^r) \mathrm{a}_t^H(\theta_{c,p}^t) \quad (7)$$

Where N_c is the number of clusters, N_p is the number of sub paths per cluster and $\alpha_{c,p}$ is the complex gain of the p^{th} sub path within the c^{th} cluster. For near and far field model conditions (Within Rayleigh distance), the wavefront curvature must be considered, and the channel becomes

$$\mathrm{H}_{NF} = \sum_{l=1}^{L} \alpha_l e^{-j\frac{2\pi}{\lambda} r_{mn,l}} \quad (8)$$

Where $r_{m,n,l}$ denotes the propagation distance between the n^{th} transmit and m^{th} receive antennas through l^{th} path, capturing spherical wave propagation. When a Reconfigurable Intelligent Surface (RIS) assists the transmission, the channel can be expressed as

$$\mathrm{H}_{RIS} = \mathrm{H}_{BR} \Phi \mathrm{H}_{RU} \quad (9)$$

Where H_{BR} represents the BS to RIS channel, H_{RU} denotes the RIS-to-UE channel and $\Phi = \mathrm{diag}((e^{j\phi_1}, e^{j\phi_2}, \dots, e^{j\phi_{N_R}})$ is the diagonal matrix of adjustable reflection phase shifts for the N_R RIS elements. The composite effective end-to-end channel is therefore

$$\mathrm{H}_{\mathrm{eff}} = \mathrm{H}_{BU} + \mathrm{H}_{BR} \Phi \mathrm{H}_{RU} \quad (10)$$

Where H_{BU} is the direct BS-UE channel. The degree of spatial correlation between antenna elements is given by

$$\rho_{ij} = \frac{\mathbb{E}[\mathrm{h}_i \mathrm{h}_j^*]}{\sqrt{\mathbb{E}[|\mathrm{h}_i|^2]\mathbb{E}[|\mathrm{h}_j|^2]}}$$

(11)

$[\mathrm{h}_i \mathrm{h}_j^*]$ are the complex channel gains. $\sqrt{\mathbb{E}[|\mathrm{h}_i|^2]\mathbb{E}[|\mathrm{h}_j|^2]}$ normalizes the correlation, so that ρ_{ij} is bounded between the

0 and 1. $\mathbb{E}$ is the average random variable. Where ρ_{ij} measures the correlation between channel vectors corresponding to antenna i and j. The estimation performance of the channel is evaluated through "Mean squared error (MSE)" defined as

$$\mathrm{MSE} = \mathbb{E}[\| \mathrm{H} - \widehat{\mathrm{H}} \|_F^2] \tag{12}$$

Where $\widehat{\mathbf{H}}$ is the estimated channel matrix obtained by the machine learning model. The $F$is a Frobenius norm, which is like the Euclidean norm for matrices. The spectral efficiency is defined by

$$\eta = \log_2 \det(\mathrm{I}_{N_r} + \frac{P_t}{N_t\sigma^2}\mathrm{HH}^H) \tag{13}$$

The η is a spectral efficiency and H^H is a Hermitian transpose of H . I is an identity matrix of size $N_r \times N_t$. P_t is the total transmitting power. $\det(\mathrm{I}_{N_r} + \frac{P_t}{N_t\sigma^2}\mathrm{HH}^H)$ is the determinant of a matrix. The bit error rate (BER) corresponding to modulation order M and SNR γ can be estimated by

$$\mathrm{BER} = Q(\sqrt{2\gamma \sin^2(\frac{\pi}{M})}) \tag{14}$$

Where Q denotes the Gaussian Q-function. $sin^2(\frac{\pi}{M})$ is a minimum distance between the points. Through this comprehensive channel model, the mmwave environment is mathematically, enabling accurate estimation and prediction using the subsequent machine learning based methods such as MIMO-MRC preprocessing, HCSSP channel recovery, federated learning-based training and SCOVEM based prediction.

B. Data Preprocessing

Data pre-processing is used to extract the data and remove the noise. The important performance of pre-processing is Reducing noise, enabling feature extraction, improving efficiency and enhancing accuracy. By using high SNR, it will reduce the noise **MIMO-MRC** (**Multiple input and multiple output Maximal ratio combining**) model exhibits asymptotic (high SNR). Feature Engineering makes the raw data into more interpretable data. To reduce the noise, the clustering algorithm are used to group nearby points that likely belong to the same object. Filtering techniques, such as thresholding or statistical outlier removal, are also applied to remove noise and improve the quality of the point cloud. In Fig. 2 Data preprocessing model is shown Narrowband MIMO baseband receives model (matrix form)

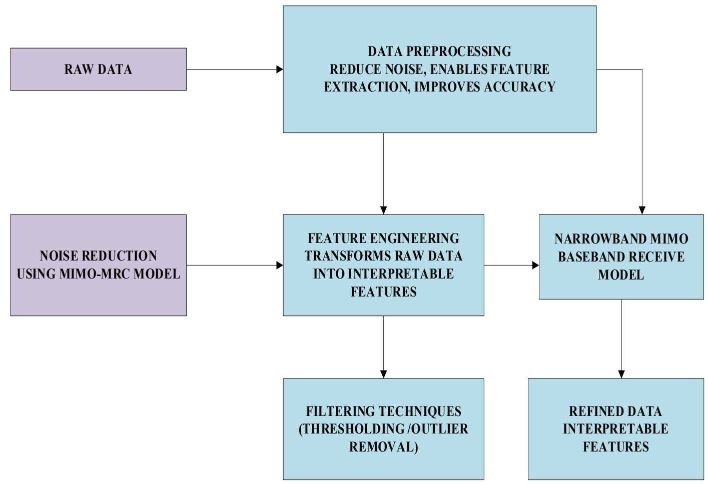

Fig. 2 Data preprocessing

$$y = Hx + n \tag{15}$$

$y \in \mathbb{C}^N$ is the received vector, $H \in \mathbb{C}_p^N$ is the MIMO channel matrix, $x \in \mathbb{C}^N$, the transmitted symbol vector, and $n \sim \mathcal{CN}(0, \sigma_n^2 I)$ AWGN. MIMO-MRC focuses on single stream transmission or combining to maximize the post combiner SNR. The single stream transmits using beam forming vector f.

$$x = fs, \quad ||f||^2 = 1 \tag{16}$$

scalar symbols s is pre coded by unit norm beam former $f \in \mathcal{C}_N$. Linear combiner output (scalar observation)

$$z = w^H y = w^H H f_s + w_H\, n$$
(17)

The w is a combine vector and y is received signal vector. f_s , transmit beamforming and n is the noise vector. $w_H\, n$ is the desired signal after combing with w , affecting SNR and $w^H H f_s$ is the desired signal after combining with w .Receive combiner $w \in C_N$ produces scalar z . MRC (Maximum Ratio combining) chooses w to maximize SNR. MRC optimal combiner for known channel.

$$w_{MRC} = k_{HF} \tag{18}$$

Where k is a scalar normalization. In single stream transmitted on F, the receive MRC weights match the effective receive steering vector HF. The post combiner SNR

$$SNR = \frac{|w^H Hf|^2 P_t}{\sigma_n^2 ||W||_2} \quad (19)$$

SNR after linear combiner w with transmit power p_t. The MRC maximum SNR (choose $w = Hf$)

$$SNR_{MRC} = \frac{Hf\,P_t}{\sigma_n^2}$$

(20)

The numerator becomes squared norm of effective receive vector; this realizes maximal SNR for that f. Optimal transmit beamformer under full CSI. It is a "right singular vector" (v_1) corresponding to largest singular value of H. MRC SNR with SVD beamforming (closed form)

$$SNR_{MRC}^{opt} = \frac{\sigma_1^2\,P_t}{\sigma_n^2} \quad (21)$$

σ_1 is the largest singular value of H. This shows equivalence to strongest spatial eigenmode. Effective array again (diversity+ beamforming)

$$G_{array} = \left||Hf|\right|^2 \sum_{i=1}^{N_r} |h_i^H f|^2$$

(22)

h_i^H is the i th receive antenna row the expression decomposes array gain into per receiver contributions. G_{array} is the array gain and N_r is the number of received antenna. The SNR under independent Rayleigh per antenna channels (MRC with equal power per transmit antenna). If Hf has elements $CN(0, \sigma_g^2)$ then,

$$SNR_{MRC} = \frac{P_t}{\sigma_n^2} \sum_{i=1}^{N_t} |g_i|^2$$

(23)

MRC sums per branch SNR; for Rayleigh fading each $|g_i|^2$ is exponential giving gamma distribution of the sum diversity order. N^t is the number of transmitted antennas.

$$\gamma = p_t \sigma_g^2 / \sigma_n^2$$

(24)

γ represents the ratio of received "SNR" power. σ_g^2 denotes average channel power and σ_n^2 .The outage probability for threshold γ_0

$$P_{out}(\gamma_0) = \mathrm{P_r}\{\gamma < \gamma_0\} = 1 - e \sum_{k=0}^{N_t-1} (2k) \left(\frac{1}{4(1+\gamma)}\right) \quad (25)$$

$\mathrm{P_r}\{\gamma < \gamma_0\}$ denotes formal definition of outage that has chance SNR is less than the required γ_0, minimum SNR required to achieve the target performance.

$\sum_{k=0}^{N_t-1} (2k) \left(\frac{1}{4(1+\gamma)}\right)$ represents the cumulative effective order. The expression may be given via MGF integration; above is representative closed form (can be simplified or expressed via confluent hypergeometric functions). It shows rapid BER improvement with N_r.High SNR ($\mathfrak{h}$) asymptotic SER slope (diversity order)

$$SER(\gamma) \alpha \gamma_n \quad \text{as } \gamma \longrightarrow \mathfrak{h} \quad (26)$$

MRC achieves full diversity N_r error decays with SNR exponent equal to number of combining branches. γ_n denotes normalized SNR term. Effective SNR after preprocessing whitening normalization whiten input covariances $R_y = \mathbb{E}[yy^H]$

$$y = R_y \qquad H = R_y 1\backslash 2\ H \quad (27)$$

R_y is a positive semi definite Hermitian Matrix. Whitening reduces colored noise effects or spatial interreference; subsequent MRC on H yields corrected SNR. Preprocessing increases effective SNR when interference / noise coloring exists. The principal component Analysis (PCA) denoising (rank -r truncation)

$$Y_r = U_r U_r^H Y \quad (28)$$

Where Y is received training snapshot matrix, U_r first r principal eigenvectors; Y_r is the projected or reduced - rank version of the received signal. $U_r U_r^H$ is a "projection matrix" onto subspace spanned by the columns of U_r., improving SNR before MRC. The MMSE prefilter that maximizes SNR in presence of Interference R_i

$$W_{MMSE} = c\left(H_f f^H H^H + R_i + \sigma_n^2 I\right)^{-1} H_f \quad (29)$$

W_{MMSE} filter matrix, to apply at the receiver and c denotes the scaling constant. H_f is the "channel matrix" of the desired signal. Where R_i denotes the interference covariance matrix. I is identity matrix. The MMSE combiner balances signal enhancement and interference/noise suppression. If interference negligible MMSE=MRC. The channel estimation model with estimation error

$$D = H + E, \ \ \mathbb{E}\left[||E||F\right] = \sigma_e^2$$

(30)

D is the estimated channel matrix and $\left[||E||F\right]$ is a frobenius norm squared. E models estimation error (MSE σ_e^2). Use Y for beamforming/combing; imperfect CSI reduces MRC gain. SNR degradation due to channel estimation error

$$SNR_{eff} \approx \frac{||Yf||^2 P_t}{\sigma_n^2 + P_t||EF||^2} \quad (31)$$

The estimation error acts like additional signal dependent noise; the denominator increases by residual beamforming leakage power $P_t||Ef||^2$. The piolet length T_p vs estimation MSE (LS estimator, orthogonal pilots)

$$\sigma_\varepsilon^2 = \frac{\sigma_n^2}{P_p T_P} \quad (32)$$

The pilot power P_p and T_p orthogonal pilot symbols, LS channel MSE decreases inversely with training energy; shows trade-off between piolet overhead and CSI quality (hence MRC effectiveness). The effective SNR gain from averaging (data preprocessing); averaging L independent snapshots

$$\tau_{avg} = L_\tau \quad (33)$$

Temporal averaging of independent noise realizations improves effective SNR by factor L; L_τ is a constant represent the mean delay. used in preprocessing before combining. The spatial correlational model

$$H = R_r h_{id} R_t \quad (34)$$

The R_r represents the received correlation matrix and R_t , transmit correlation Matrix. h_{id} denotes Rayleigh channel fading matrix. The post MRC mean and variance

$$\mathbb{E}[\gamma] = \frac{P_t}{\sigma_n^2} tr\left(R_{eff}\right), Var(\gamma) = \frac{P_t^2}{\sigma_n^4} tr(R_{eff}^2)$$
(35)

The practical quantized combiner/ phase shifter model (finite resolution

$$w_q = Q_B(w), Q_B(e^{jo}) = e_i$$
(36)

w is the original beamforming /precoding vector. w_q is the quantized version applying B bit quantization. The quantization function $Q_B(e^{jo})$ represents bit phase quantizer on analog combining; $e_i$is the quantized phase value and e^{jo} is a complex number on the unit circle. Finite resolution reduces array gain include as multiplicative loss factor in SNR expressions. Pseudocode for MIMO-MRC model is presented in Algorithm 1

Algorithm 1: Multiple input and multiple output Maximal ratio combining (MIMO-MRC) Model

1. **Initialize** system parameters ($T_x, R_x, RIS, \sigma^2, P_t, N_t, N_r$)
2. Acquire received baseband signal vector $y = HX + n$
3. Apply MIMO-MRC combining to maximize post-combiner SNR
4. Compute optimal combiner $w = \frac{H^H}{||H||}$
5. Evaluate post combiner $SNR = (|w^H Hx|^2)/(w^H w\, \sigma^2)$
6. Perform SVD of channel matrix $H = UV^H$ to obtain dominant eigenmode
7. Select transmit beamformer $f = V_1$ corresponding to largest singular value
8. Compute effective array gain $A_{eff} = |h_i|^2$
9. Estimate branch SNRs under Rayleigh fading and compute diversity gain
10. Calculate outage probability $P_{out} = P(SNR < y_{th})$
11. Normalize received signal using whitening matrix $R = n$ to remove the noise.
12. Apply PCA denoising:$Y_r = U_r Y$
13. Compute MMSE filter W_{mmse}
14. Estimate channel $H = H + \Delta H$ and compute estimation error $MSE = ||\Delta H||^2$
15. Evaluate SNR degradation due to estimation error
16. Optimize pilot length τ_p and P^p to minimise channel MSE.
17. Apply temporal averaging over L snapshots to enhance SNR_{eff}
18. Compute spatial correlation metrics $R_{rx} and\ R_{tx}$
19. Quantize beamforming phase: $\theta_q = Q(\theta)$ with bit resolution
20. **Output** pre-processed data Y clean, enhanced channel matrix $\mathcal{H}$ and effective SNR.

C. Near and far field model

The novelty "**Hierarchical Compressed Sensing with Spatio-Temporal Sub-Partitioning for Ultra-Fast mmWave Beam Alignment** "is a combine of fast alignment, sub partition framework and hierarchical compressed sensing for providing "near and far field model" accurately. Wireless communication, the "**Near-field** and **Far-field**" models

+91 94448 68310
phdservicesorg@gmail.com

describe different region of an electromagnetic field propagating from an antenna. The far-field is a region where the radiation pattern is stable and the power decays as inverse square of the distance. The near-field, closer to antenna, is more complex, with both reactive and radiative components. The proposed methodology using a **Fast alignment and sub-partition framework** and **hierarchical compressed sensing** is designed to address accuracy in RIS plane of both near and far field model. The Fast alignment algorithm works as coarse alignment which quickly identify the primary **near-field** regions of interest. By doing this, the system avoids spending time on areas with very low signal strength and can access RIS technology for far and near field model. In **Sub partition framework**, the measurement area is divided into smaller, more manageable sub-regions. To combine of this technique the novelty as "Hierarchical Compressed Sensing with Spatio-Temporal Sub-Partitioning for Ultra-Fast mmWave Beam Alignment" The sub-partitions can be dynamically adjusted based on the initial coarse alignment results, capturing the features of both near and far field model in RIS phase shifts. Hierarchical Compresses technique builds on efficiency gained from the sub-partition framework by applying **compressed sensing (CS)** in a hierarchical manner. ". In fig. 3 the "Near and Far field model" is executed. The hybrid of the fast alignment algorithm and sub partition framework and hierarchical compressed sensing is used to get feasible in RIS technology in both "Near and Far field model". The Fresnel (Rayleigh) distance (near and far boundary

$$d_F = \frac{2D^2}{\lambda} \quad (37)$$

Where distance d_F separates near and far field model for an antenna of maximum dimension D at wavelength λ. For $r \ll d_F$ near field effects matter; for $r \gg d_F$ planar approximation holds. Friis path-loss (free space),

$$_r = p_t G_t G_r \left(\frac{\lambda}{4\pi r}\right)^2$$

(38)

p_r is a received power at distance r from transmit power p_t and gains G_t, G_r. Use as a baseline for distance dependent attenuation. Where p_t, p_r is a power and G_t, G_r is an antenna gains, r (distance). General narrowband MIMO baseband channel (Multipath)

$$H = \sum_{l=1}^{L} \alpha_l e^{-j2\pi f_e^{Tl}} \; a_r(\theta_l, r_l) a_t^H(\varphi_l, r_l)$$

(39)

Where, sum of L discrete paths with complex gain α_l, delay T_l, receive/transmit array responses a_r, a_t^H that depend on both direction and range r_l. Far field (planar wave) ULA steering vector (transmit)

$$a_t^{FF}(\varphi) = \frac{1}{\sqrt{N_t}} \quad (40)$$

The a_t^{FF} is a transmit array response for far field (FF) transmission at angle of departure φ .Planar wave steering for ULA with element spacing d. N_t is a number of transmit antennas in the array.

$$a_t^{NF}(\varphi, r) = \frac{1}{\sqrt{N_t}}[e]^T$$

(41)

The $a_t^{NF}(\varphi, r)$ is a "near field" transmit array response vector. Where r is a propagation distance and $[e]^T$ is a phase vector. The account for per element distance dependent phase (spherical curvature). The second order Fresnel approximation (Useful intermediate)

$$\triangle_n \approx r - x_n + \frac{x_n^2}{2r} \quad (42)$$

Taylor expansion of $\triangle_n$ valid in Fresnel region; x_n is a position of the n^{th} antenna element. First two terms give planar plus quadratic correction capturing curvature. Useful in near mid field modelling.

$$H \approx A_r(R, \mathbb{X}) X \, A_t^H(R, \varphi) \quad (43)$$

The $A_t^H(R, \varphi)$ is the transmit "array steering vector" or matrix. $A_r(R, \mathbb{X})$ is the receive array steering vector or matrix. Channel approximated by sparse coefficient matrix X over a polar dictionary parameterized by discrete ranges R and angles $\mathbb{X}$. Compressed sensing forward measurement

$$y = \varphi x + n \quad (44)$$

The measurement vector y collected by beam probing matrix φ from sparse channel vector x with noise n. This is the core equation for CS-based alignment and estimation in each sub-partition. Hierarchical sparse model

$$x = x^1 + x^2, \quad ||x^1||0 \ll ||x^2||0 \quad (45)$$

x^1 models coarse (few strong components, near field spots) and x^2 models fine residuals. Restricted Isometry property (RIP) requirements

$$1 - \delta_k)||v|| \leq ||\varphi v|| \leq (1 + \delta_k)||v||o \leq k \quad (46)$$

The "sensing matrix" φ must satisfy RIP to stably recover k sparse signals, δ_k is RIP constant. v is a k sparse vector. Mutual coherence (design metric)

$$\mu(\varphi) = max \frac{|<\varphi_i,\varphi_j>|}{||\varphi_i||^2 ||\varphi_j||_2} \tag{47}$$

Lower mutual coherence of a matrix φ. $||\varphi_i||^2\, ||\varphi_j||_2$ is a Euclidean norm of the columns and $\langle \phi_i, \phi_j \rangle$ is an inner product between columns i and j and is used when designing probing beams for fast alignment. Beam probing energy allocation across sub partitions

$$\sum_{s=1}^{s} m_s = M \tag{48}$$

The total number of probe measurements M split among S spatial sub partitions, with allocation weights m_s proportional to prior probability of energy in partition s. Coarse to fine success probability (Bayes)

$$P_{succ} = \sum P(select\ s)P(fine\ success) \tag{49}$$

Where p_{succ} is an overall probability of success. $p(select\ s)$ is a probability that a particular strategy and $P(fine\ success \mid s)$ is a conditional probability of success given that s was selected. Aligns coarse selection probability and conditional fine success. Use to optimize probing allocation m_s. MIMO receive signal with beamformer w and precoder f

$$y = w^H Hf\, s + w^H n \tag{50}$$

Baseband scalar observation used in beam alignment design. SNR after MRC combining

$$SNR_{MRC} = \frac{|w^H Hf|^2}{\sigma_n^2} \tag{51}$$

MRC maximum ratio combining SNR: used when evaluating high SNR. MIMO-MRC behaviour,

$$Var(\theta) \geq \frac{1}{2SNR\,(s_H)^2} \tag{52}$$

lower bound on unbiased "AoA" estimate variance; indicates the near field model curvature changes estimation prediction. Range estimation

$$r = cT, T = \frac{r}{c} \tag{53}$$

T is a path delay, c is a speed of light and used to link polar sparse dictionary range bins to physical distances. RIS reflection model

$$\beta_m = p_m^e, \quad 0 \leq p_m \leq 1, \tag{54}$$

Where β_m is the resulting weighted or scaled probability. p_m is a probability value associated with some event $m.e$ is an exponent could represent reliability factor, amplification. Constraint $0 \leq p_m \leq 1$ ensures that p_m is valid probability. and RIS assisted path modification

$$RIS, l \longleftarrow \sum_{m=1}^{M_{RIS}} \beta_m g_{m,l} \tag{55}$$

The RIS is a "Reconfigurable intelligent surface (RIS)". l is the index of a particular element or layer in the RIS. M^{RIS} represents "Number of RIS elements". $g_{m,l}$ is the gain or channel coefficient from RIS element m to layer. where each element applies amplitude p_m and phase. $g_{m,l}$ accounts for geometry and element response. RIS phase optimization

$$\max |w_H(\sum_{m}^{i=1} h_{t,m}^{H} \quad) \qquad | \tag{56}$$

The non convex optimization to choose RIS phases that align reflected components in near field include distance dependent phase in $h_{t,m}$. The phase quantization

$$\varphi_m = \frac{2\pi}{Q}\ round\left(\frac{Q\varphi_m}{2\pi}\right) \tag{57}$$

The φ_m is a phase shift and $\frac{2\pi}{Q}$ is a step size of each quantized phase level. $round\left(\frac{Q\varphi_m}{2\pi}\right)$ denotes to rounds the value to the nearest integer. The Q level quantizer for hardware constrained RIS; include quantization error term in analysis. Polar domain dictionary (Closed form for ULA element n)

$$a_n(\varphi, r) = \exp\left(\left(-j\frac{2\pi}{\lambda}\sqrt{r^2 + x_n^2 - 2rx_n \sin\oint}\right)\right. \tag{58}$$

explicit atom used to assemble $A_t(R,\varphi) in$ (39); supports joint angle range sparsity. Measurement noise model (AWGN+ model mismatch)

$$n \sim cN(0, \sigma^2 I), \qquad y = \varphi x + n + e_{mismstach} \tag{59}$$

The $e_{mismatch}$ to capture dictionary off grid or near field approximation errors; use robust or off- grid recovery to mitigate.

$$a(\varphi + \triangle\varphi, r + \Delta r) \approx a(\varphi, r) + \frac{\partial a}{\partial \varphi}\Delta\varphi + \frac{\partial a}{\partial r}\Delta r$$
(60)

The $a(\varphi, r)$ is an array steering vector. $\Delta\varphi, \Delta r$ is a small change in angle and distance. $\frac{\partial a}{\partial \phi}, \frac{\partial a}{\partial r}$ is a partial derivative of a with respect to φ and r. Then the overall constrained optimization for hierarchical sub partition framework fast alignment

$$\min \frac{1}{2}||Y - \theta(x^1 + x^2)||(\lambda^1 + \lambda^2) \quad (61)$$

Y represents the observed vector and x_1, x_2 represents the components or variables to estimate. $\lambda^1 + \lambda^2$ is a regularization parameter. Joint recovery that enforces coarse support s_{coarse} from fast alignment for the strongest components x_1 and x_2 captures fine residuals. The pseudocode hybrid model of Fast alignment algorithm with Partition framework is presented in Algorithm 2.

Algorithm 2: Fast alignment algorithm with Sub partition Framework

1. **Initialize** antenna parameters (D, λ, Tx, Rx, RIS)
2. Measure distance d between transmitter and reciver.
3. Compute Rayleigh Boundary $r = 2D^2/\lambda$
4. **If** $d < R$ then→ Near- Field Model
5. **Else**→ far-field Model
6. **Initialize** Fast alignment algorithm
7. Perform coarse beam scan→ identify strongest region
8. Divide RIS plane into sub- partitions (spatial cells)
9. For each sub-partition i do
10. Collect measurement vector
11. Apply compressed sensing recovery to estimate h_i
12. Store recovered subchannel H_i
13. **End for**
14. Combine $\{H, i\}$ → From overall channel matrix h_{est}
15. Optimize RIS reflection phases φ to maximize received power
16. Evaluate Mutual coherence $\mu(\varphi)$ to ensure stable recovery.
17. Reconstruct near/far field beam patterns via h_{est}
18. Calculate SNR and spatial correlation for each model.
19. Select Model (Near/ Far) with minimum MSE and higher SNR.
20. Output final aligned channel estimate H_{est} for RIS assisted link.

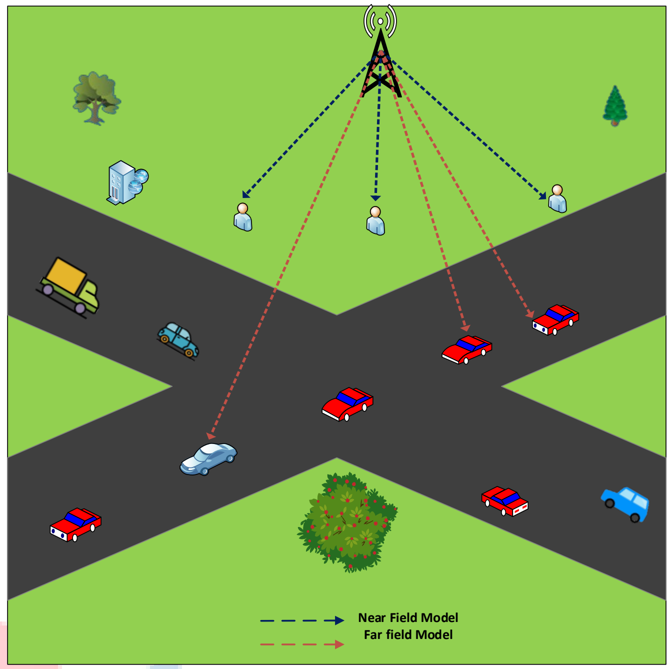

Fig. 3 Near and Far field model

D. Federated based training

"Generative adversarial network "is used to train the dataset from the ray-tracing (RT) simulation. **"Generative adversarial Network"** that uses two competing neural networks to generate new data. To perform better training, the **"Federated Learning"** is used to train the data. The data is collected from the users through the Base station in mmwave channel modelling. Federated learning is a decentralized machine learning technique where model is trained across multiple devices or servers. In Fig.4, the federated based training is shown. Let the central server initialize the model parameter as

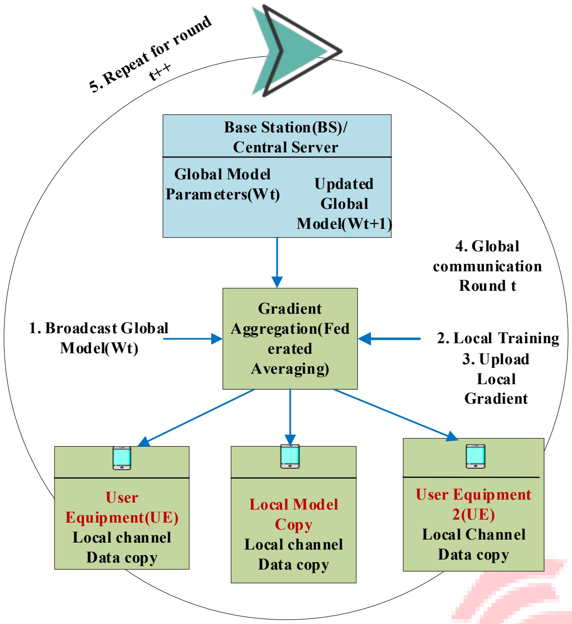

Fig .4 Federated based training

$$\omega_0 = InitializeModel(\theta) \quad (62)$$

Where $\omega_0 \in \mathbb{R}^d$ represents the initial weight vector of the global model, and θ denotes the initial configuration of model parameters such as weights and biases. Each participating device $i \in \{1,2, \ldots .. n_i\}$ possesses its own dataset

$$D_i = \{(x_{i,j}, y_{i,j}) \mid j = 1,2, \ldots ., n_i\} \quad (63)$$

D_i is a dataset belong to client indexed by i. Where $x_{i,j}$ and $y_{i,j}$ represent the input feature vector (eg., received signal or channel coefficient) and its corresponding label(channel state estimated), respectively, and n_i is the number of samples stored locally at device i. j is a sample number within the i^{th} dataset. Each client aims to minimize its local empirical loss function

$$F_i(\omega) = \frac{1}{n_i}\sum_{j=1}^{n_i} l(F_w(x_{i,j}), y_{i,j})) \quad (64)$$

where $F_i(\omega)$ is the prediction model parameterized by ω and $\sum_{j=1}^{n_i} l(F_w(x_{i,j}), y_{i,j}$ denotes the local loss criterion, typically chosen as the mean squared error (MSE) for channel estimation. The global learning objective is expressed as the weighted aggregation of all local objectives:

$$F(w) = \sum_{i=1}^{N} \frac{n_i}{n_{total}} F_i(\omega), \quad n_{total} = \sum_{i=1}^{N} n_i \quad (65)$$

The $F(w)$ denotes the overall function or weighted average. The n_{total} is a total weight or sum of counts.n_i is a count or weight associated with the i_{th} component. Ensuring the devices with larger local datasets contribute proportionally to the global optimization. During each communication round t , each client performs local stochastic gradient descent (SGD) updates on its dataset as

$$\omega_i^{t+1} = \omega^t - \eta \Delta F_i(\omega^t) \quad (65)$$

ω_i^{t+1} is the updated value of the i_{th} element and ω^t is the current value of the variable at i_{th} element. η is a step size or learning rate. Where, $\eta > 0$ is the learning rate controlling the update step size and $\Delta F_i(\omega^t)$ is the Stochastic gradient descent (SGD) updates on its dataset as

$$\Delta F_i(\omega^t) = \frac{1}{n_i} \sum_{j=1}^{n_i} \nabla \omega^l (F_w(x_{i,j}), y_{i,j}) \quad (66)$$

With $\nabla \omega^l$ representing the derivative with respect to model parameters. Upon completing the local update phase, each client transmits the updated model weights to the central server as

$$send(\omega_i^{t+1}) \longrightarrow Server \quad (67)$$

Ensuring that no private data are shared, only parameters updates. The server then aggregates the local models using the federated Averaging (FedAvg) rule,

$$\omega^{t+1} = \sum_{i=1}^{N} \frac{n_i}{n_{total}} \omega_i^{t+1} \quad (68)$$

To form a new global model that captures the statistical diversity across clients. The updated global parameters are then synchronized back to all clients as

$$\omega_i^{t+1} \longleftarrow \omega^{t+1}, \mathbb{A}_i \in \{1,2, \ldots ..., N\} \quad (69)$$

The $\mathbb{A}_i$ is an index set or selected components. Allowing every participant to start the training round from the same global state. The iterative process continues until convergence is achieved according to the stopping criterion

$$\left|\left|\omega^{t+1} - \omega^t\right|\right|_2 < \epsilon \quad (70)$$

Where $||\omega^{t+1} - \omega^t||_2$ denotes the Euclidean norm and $\epsilon > 0$ is a small constant specifying the desired precision threshold. ϵ is a small positive threshold. To handle heterogeneous data distributions and prevent client drift, a proximal regularization term is incorporated into the local objective as

$$F_i^{Prox}(\omega) = F_i(\omega) + \frac{\mu}{2} ||\omega - \omega^t|| \quad (71)$$

Where μ is a positive coefficient that penalizes the deviation between the local and global models, thereby improving stability in non-identically distributed scenarios. $F_i^{Prox}(\omega)$ is the proximal version of $F_i(\omega)$. To further accelerate convergence, adaptive momentum-based optimization can be employed through the Federated Adam variant

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (72)$$

The g_t is the gradient of the loss function at iteration t and m_t is exponentially weighted moving average (EWMA). ϑ_t is a EWMA of past squared gradients and β_1, β_2 are decay rates for the moving averages. where $g_t = \nabla F(\omega^t)$ is the stochastic gradient, $\beta_1, \beta_2 \in [0,1)$ are momentum coefficient and the small constant ϵ measures stability. For devices contributing unequally to the global update due to data imbalance, the aggregation can be refined as

$$\omega^{t+1} = \frac{\sum_{i=1}^{N} p_i \omega_i^{t+1}}{\sum_{i=1}^{N} P_i} \quad (73)$$

Where $p_i = \frac{n_i}{n_{total}}$ denotes the proportional contribution weight for client i. The convergence behaviour of the federated optimization after T rounds can be bounded as

$$\mathbb{E}[F(\omega_T)] - F(\omega^*) \le \frac{C}{T} + O(\eta^2 \sigma^2) \quad (74)$$

$\mathbb{E}[F(\omega_T)]$ is an expected value of the objective function. Where $F(\omega^*)$ is the optimal loss value, c is a constant $F(.)$ and σ^2 represents the variance of stochastic gradients. Finally, to improve communication efficiency, the parameter transmission can be optimized by minimizing the cumulative deviation between local and global updates.

$$min \sum_{i=1}^{N} ||\omega_i^{t+1} - \omega^t||^2 \quad , \left||\omega_i^{t+1} - \omega^t|\right|_2 \le \delta \quad (75)$$

Where δ defines the allowable update bound that controls the communication budget. Through these iterative procedures, the federated- based training mechanism ensures that mmwave channel models are collaboratively trained using distributed observations while maintaining user privacy and minimizing the bandwidth usage. The resulting global model effectively generalizes across diverse propagation environments, providing robust channel estimation and prediction with reduced communication cost and enhanced data security.

E. Machine Learning based Robust channel estimation:

The novelty "**SCOVEM: Score-Based Generative Model-Enhanced Support Vector Machine with Bayesian Optimization for MmWave Fault Diagnosis**" is a combine of SVM, Score based generative model and bayes optimization" The robustness plays an important role, when the channel is estimated or predicted. The **SVM, Score based generative model and the Bayes Optimisation** to make the channel estimation more robustness. The **Support Vector Machine (SVM)** offer a specific kind of **robustness** that makes to estimate the channel in mmWave systems. The robustness of SVM comes from its core principle of maximizing the margin**.** SVM used to directly predict a continuous channel matrix. Then, they are often applied to solve "channel estimation problem" as a classification or "regression task". The Support Vector Machine algorithm is use to classify and predict the channel and it make more robust in channel estimation. For Channel estimation using the novelty method as "Score-Based Generative Model-Enhanced Support Vector Machine with Bayesian Optimization for mmWave channel estimation". The proposed model integrates Score based

The overall process can be formulated as:

$$y = Ah + n \quad (76)$$

Where y denotes the received pilot signal, A represents the known sensing matrix, h is the unknown channel vector to be estimated, and $n \sim \mathcal{CN}(0, \sigma^2 I)$ is additive "Gaussian noise".

$$dx = f(x,t)dt + g(t)dw_t \quad (77)$$

The forward stochastic differential equation (SDE) describes the diffusion process, where $f(x,t)$ is the drift function, $g(t)$ is the noise schedule, and w_t represents "Brownian motion".

$$s_\theta(x,t) = \nabla_x \log P_t(x) \quad (78)$$

Here, $s_\theta(x,t)$ denotes the score function, parameterized by neural network weights θ, which estimates the gradient of log-density at time t.

$$dx = \{f(x,t) - g(t)^2 s_\theta(x,t)dt + g(t)dw^t \quad (79)$$

This reverse time SDE reconstructs samples from the learned data distribution, enabling the generation of high-quality synthetic samples used to augment the SVM training.

$$\nabla_x \log p(x|y) = \nabla_x \log p(x) + \nabla_x \log p(y|x) \quad (80)$$

The conditional posterior score combines prior knowledge $p(x)$ and likelihood information $p(y|x)$ to generate channel estimates consistent with observed data.

$$\nabla_x \log p(y|x) = \frac{1}{\sigma^2} A^H(Ax - y) \quad (81)$$

Assuming a Gaussian likelihood model, this represents the gradient of the measurement log likelihood used during posterior sampling.

$$x_{k+1} = x_k + \eta_k s_\theta(x_k, t_k) + \sqrt{2\eta_k z_k} \quad (82)$$

Discretised Langevin dynamics for generating channel samples, where η_k denotes step size and $z_k \sim \mathcal{N}(0, I)$

$$\mathcal{D} = \{(x_j, y_j)\}_{j=1}^{m} \quad (83)$$

The generated dataset $\mathcal{D}$ augments the original training data, improving generalization in low SNR conditions.

$$\min \frac{1}{2}||\omega||^2 + C \sum_{i=1}^{n} (\vartheta_i + \vartheta_j) \quad (84)$$

This represents the primal optimization of SVM regression using ϑ insensitive loss and regularization constant C.

$$y(x) = \sum_{i=1}^{n} (\propto_i - \propto_i^*) K(x_i, x) + b \quad (85)$$

The prediction function of SVM where $K(x_i, x)$ is the "kernel function defining the feature space" mapping.

$$\mathcal{L}_{aug} = \frac{1}{2}||\omega||^2 + \mathcal{C}[\sum_{i=1}^{n} l_E(y_i, y_l) + \lambda \sum_{j=1}^{m} \omega_j l_E(y_i, y_l)] \quad (86)$$

This augmented objective integrates real and generated samples with weighting factor λ and sample confidence ω_j.

$$\omega_j = \exp(-\gamma ||Ax_j - y||^2) \quad (87)$$

ω_j is the weight assigned to the j^{th} estimate. x_j is the j^{th} candidate vector. A is the channel matrix. $||Ax_j - y||^2$ is the squared Euclidean norm. Weight of each synthetic sample determined by posterior consistency; higher weight implies better alignment with real measurements.

$$\mathcal{R}_{val}(\varphi) = \frac{1}{N_v} \sum_{i=1}^{N_v} ||y_i(\varphi) - \mathrm{y_i}||^2 \quad (88)$$

The $\mathcal{R}_{val}(\varphi)$ is the validation error or residual as a function of φ.$N_v$is the number of validation sample and $y_i(\varphi)$, predicted or estimated measurement for the i^{th} sample. y_i denotes the actual observed measurement for the i^{th} sample. The validation risk function used for hyperparameter tuning.

$$f(\varphi) \sim \mathcal{GP}(0, K_{GP}(\varphi, \varphi_1)) \quad (89)$$

$f(\varphi)$is a function of the variable φ and $\mathcal{GP}$ is Gaussian process. 0 is the mean function of the Gaussian process. $K_{GP}(\varphi, \varphi_1)$ is the covariance (kernel) function between φ and φ_1(input points for the function). Bayesian optimization models the objective function $f(\varphi)$ as a gaussian process with covariance kernel k_{GP}.

$$\mu(\varphi_*) = K_*^T(K + \sigma_n^2 I)^{-1} f,\ \sigma^2(\varphi) = k(\varphi_*, \varphi_*) - k_*^T(K + \sigma_n^2 I)^{-1} k_* \quad (90)$$

$k(\varphi_*, \varphi_*)$ represents the covariance of the scalar. I represents the "identity matrix" and k is the $n \times n$ covariance matrix for training points. The Mean and variance of the GP posterior are computed for candidate parameter φ .

$$EI(\varphi) = (f_{min} - \mu(\varphi))\varphi(z) + \sigma(\varphi)\tau(z) \quad (91)$$

Expected improvement acquisition function where $z = \frac{f_{min} - \mu(\varphi)}{\sigma(\varphi)}$, balancing exploration and exploitation.

$$\varphi_{t+1} = \arg\max EI(\varphi) \quad (92)$$

φ_{t+1} represents next input point to evaluate the function and $EI(\varphi)$ is the expected improvement at input φ. The $\arg max$ is the value of φ that maximizes EI . Next hyperparameter configuration selected by maximizing expected improvement.

$$\mathcal{J}(\theta,w) = \mathcal{M}_{t,x_0,x_t}[||s_\theta(x_t,t) - \nabla_{x_t}\log p_{t|0(x_t|x_0)||^2]+\alpha\mathcal{L}_{aug}}$$
(93)

$\mathcal{J}(\theta,w)$ is the total loss to optimize model parameters θ. $\mathcal{M}_{t,x_0,x_t}$ Expectation over of diffusion time step, original data samples and noisy sample at step $t.\log p_{t|0(x_t|x_0)||^2]}$ is a true score of the perturbation kernel. $\alpha\mathcal{L}_{aug}$ is an optional augmentation loss weighted by α. Joint objective combing SGM score matching loss and augmented SVM regression loss with coupling constant α.

$$h_{final} = \frac{1}{M}\sum_{m=1}^{M} y(x^m) \qquad (94)$$

The h_{final} is a final estimate of the function or quantity of interest. M is the "Number of Monte carlo" samples and x_m is the m^{th} sample from the input distribution. $y(x^m)$ is the function evaluation at sample x^m.

Final channel estimate obtained as Monte Carlo average over M posterior samples from the SGM.

$$\mathbb{E}\left[||h = h_1||^2\right] \le E_{bias} + o\left(\frac{||w||^2}{n_{eff}}\right) + E_{gen} \qquad (95)$$

h is a True channel and h_1 is the estimated channel. $\left(\frac{||w||^2}{n_{eff}}\right)$ is a variance term, showing decay with effective sample size m_{eff}.Analytical error bound of the proposed model, where E_{bias} is approximation bias, $n_{eff} = n + \lambda \sum_j w_j$ is the "effective number of training samples", and E_{gen} denotes the generative mismatch error. The pseudocode for SCOVEM hybrid model is presented in Algorithm 3.

Algorithm 3

Hybrid model of SCOVEM

1. **Initialize** system parameters $(Y_p, \sigma^2, \varphi, Tx, Rx, RIS)$
2. Load training dataset $D = \{X_i, H_i\}$ from simulation environment
3. **Initialize** SVM parameters (Kernel type, C, ε)
4. **Initialize** Score based generative model weights w_o.
5. **Initialize** Bayesian optimizer with prior $GP(0,k)$.
6. **For each** training H_i in D do
7. Add Gaussian noise σ^2
8. Compute score targets $s^* = \nabla\{H_i^k\}\log p(H_i^k)$
9. Update model weights $w \leftarrow w - \eta\nabla L$ using denoising score matching loss
10. **End for**
11. Generate Synthetic channel samples $\mathcal{H}$ using reverse SDE.
12. Augment dataset $D' = D \cup \{\{\varphi.\mathcal{H}_{gen}, \mathcal{H}_{gen}\}\}$ for training
13. Train support vector regression (SVR) model on D'
14. Obtain prediction function $f(SVM(x)) = X_i, X + b$
15. Perform Bayesian optimization to tune SVM hyper parameters $(C, \varepsilon, kernel)$
16. **For each** received pilot signal Y_p:
17. Predict Channel estimate $\mathcal{H}_{pred} = f\left(SVM(Y_p, \theta)\right)$
18. Refine prediction using posterior correction:
19. $\mathcal{H} = \mathcal{H}_{pred} + \lambda^* s(\mathcal{H}_{pred}, t_0)$
20. Evaluate performance metrics (MSE, BER, spectral efficiency)
21. Output final robust channel estimate $\mathcal{H}$ with optimized parameters θ^*.

F. Score based Generative model

"Score based Generative model" also called as "Score based Diffusion model" or "Score Based Model". The Score Based model then learns the reverse process, which involves iteratively "denoising" random noise to generate new, realistic data and make robustness. In score-based model, it involves Stochasticity as a Regularize and the Agnostic to the Manifold Hypothesis. The Stochasticity as a Regularize is used to The continuous insertion and removal of noise in diffusion process acts as a powerful regularize. It forces the model to learn a smooth, continuous representation of the data distribution, rather than simply memorizing the training data. This makes the model less sensitive to small perturbations or outliers in the input, a key aspect of robustness.

However, SGM can operate even when this assumption does not hold. The diffusion process effectively analyses the data across the entire space, and the model learns to guide samples back to the areas of high data density. This makes SGM robust to data distributions in channel estimation and prediction. The "score based generative model" also known as "Diffusion probabilistic model", learns to estimate the score function of the channel data. It constructs a forward diffusion process that gradually perturbs the channel data with Gaussian noise, followed by a learned reverse process that denoises and reconstructs the original signal distribution.

$$y = Ah + n \qquad (96)$$

Where y represents the received pilot signal, A is the known "measurement matrix", h is the true channel vector, and $n \sim \mathcal{CN}(0, \sigma^2 I)$ is "complex Gaussian noise".

$$q_t(x_t|x_0) = \mathcal{N}(x_t;\ \alpha_t x_0, \sigma_t^2 I) \quad (97)$$

x_o is the original data or signal. x_t is a noisy version of data. The q_t is the conditional "probability distribution" and α_t is the scaling factor. $\mathcal{N}(x_t;\ \alpha_t x_0, \sigma_t^2 I)$ is a gaussian distribution with mean μ. The forward diffusion process gradually corrupts the data x_0(true channel) into noisy latent variables x_t by scaling $\propto_t$ and adding Gaussian noise with variance σ_t^2.

$$dx = f(x,t)dt + g(t)dw_t \quad (98)$$

dx is an "infinitesimal change" in x and dt is the drift term. The forward stochastic differential equation describes the continuous noise adding process, where $f(x,t)$ is drift, $g(t)$ is diffusion strength, and dw_t is "wiener noise".

$$s_\theta(x,t) = \nabla_x \log p_t(x) \quad (99)$$

The "score function" $s_\theta(x,t)$, parameterized by neural weights θ, approximates the gradient of the logarithmic density of the noisy data distribution at time t. The ∇_x is the gradient with respect to x. $p_t(x)$ is the probability density of x at time t.

$$dx = [f(x,t) - g(t)^2 s_\theta(x,t)]dt + g(t)dw_t \quad (100)$$

The reverse time SDE reconstructs clean samples by iteratively denoising, effectively inverting the diffusion process using score function.

$$X_{k+1} = X_k + \eta_k s_\theta(X_k, t_k) + \sqrt{2\eta_k z_k} \quad (101)$$

X_{k+1} represents sample after one update and t_k is a Time or noise level at step K. z_k is a random gaussian noise and $\sqrt{2\eta_k z_k}$ is a Random perturbation ensuring stochasticity. Discretised Langevin dynamics used for sampling; η_k Is the step size and $z_k \sim \mathcal{N}(o, I)$ adds stochasticity to improve coverage of the data distribution.

$$\nabla_x \log p(x|y) = \nabla_x \log p(X) + \nabla_x \log p(y|x) \quad (102)$$

∇_x is a Gradient with respect to x Conditional posterior score decomposition; ∇ $\log p(X)$ is the conditional probability and $\log p(y|x)$ is the likelihood. The generative model integrates both prior $p(x)$ and likelihood $p(y|x)$ terms for channel estimation given pilot signals.

$$\nabla_x \log p(y|x) = \frac{1}{\sigma^2} A^H (Ax - y) \quad (103)$$

Assuming additive Gaussian noise, this provides the analytical expression for the likelihood gradient in the posterior score.

$$x_o = \varsigma[x_0|x_t] = \frac{1}{\alpha_t}(x_t - \sigma_t^2 s_\theta(x_t, t)) \quad (104)$$

x_o is the predicted clean data and x_t noisy data at time t. α_t is a signal scaling coefficient and ς is used to represent function or operator. Posterior mean estimate of the clean signal given its noisy counterpart, forming the core of the reverse reconstruction process.

$$\mathcal{L}_{score}(\theta) = \mathbb{E}_{t,x}[\lambda(t)||s_\theta(x_t,t) - \nabla_x \, log q_t(X_t|X_0)||^2] \quad (105)$$

The $\mathcal{L}_{score}(\theta)$ is a Loss function and t is the diffusion time. X_0 is an original data and X_t is a noisy data. $\mathbb{E}$ is average value or mean. Training loss of the SGM using denoising score matching; $\lambda(t)$ is a time dependent weighting term that stabilizes learning across noise scales.

$$\nabla_x \log q_t(x_t|x_o) = -\frac{x_t - \alpha_t x_o}{\sigma_t^2} \quad (106)$$

The score of the Gaussian noise perturbation process used as the supervised target during training. The Mean squared Error E,

$$h = E_{p(x|y)}[X] \quad (107)$$

The h is the condition mean or estimate and X is a random variable. Final channel estimation obtained as the expectation of the posterior samples generated by the reverse process conditioned on measurements y.

$$P_t(x) = \int p_t(x|x_0)p(x_0)dx_o \quad (108)$$

$p(x_o)$ is the prior distribution of x_o and $p_t(x|x_0)$ is the forward transition "probability" and dx_o integral over all possible x_o. Marginal distribution of noisy data at time t; computed implicit

$$\mathcal{T}(\theta) = \int E[||s_\theta(x_t,t) - \nabla_x, \log q_t(x_t|x_o)||^2]dt \quad (109)$$

$\mathcal{T}(\theta)$ is a Continuous time score matching loss, integrating over diffusion time for stable parameter optimization.

$$x_t - \Delta_t = x_t + \Delta t[f(x,t) - g(t)^2 s_\theta(x_t,t)] + g(t)\sqrt{\Delta t z_k} \quad (110)$$

Discrete time approximation of the reverse SDE used for sample generation, implemented in practical reconstruction.

$$h_{est}^m = A + X_o^m \quad (111)$$

h_{est}^m is the estimated value for the m^{th} sample and X_o^m is an original or random component for m^{th} sample. M is a number of samples and m is a sample index. Each generated clen sampleX_o^m is transformed back to the estimated channel domain using pseudo inverse of the measurement matrix.

$$h = \frac{1}{M}\sum_{m=1}^{M} h_{est}^m \quad (112)$$

Monte Carlo averaging across M reconstructed samples reduces variance in final channel estimate. $Var(h)$ is the sample variance of h.

$$Var(h) = \frac{1}{M-1}\sum_{m=1}^{M} \left|\left|h_{est}^m - h\right|\right|^2 \quad (113)$$

Variance metric quantifying uncertainty in the generative estimation, important for reliability evaluation.

$$E\left[\left|\left|H - h\right|\right|^2\right] = Bias^2 + var \quad (114)$$

Bias-variance decomposition of the channel estimation error; $Bias^2$ is the squared distance between the average estimate and the true value. both components are reduced by SGM due to denoising regularization and posterior averaging.

$$SNR_{eff} = 10\log_{10}\left(\frac{||AH||^2}{||AH-y||^2}\right) \quad (115)$$

SNR_{eff} is the effective signal noise ratio and the $\left|\left|AH\right|\right|^2$is the signal power and the $\left|\left|AH - y\right|\right|^2$ is the noise power. A higher SNR indicates the predicted signal AH closely matches the observed data y mean low error.

V. EXPERIMENTAL RESULT

A. *Simulation setup*

The section proposes channel estimation and prediction using machine learning as well as it is set up for simulation. To simulate the proposed research method, the MATLAB R2023a is used. Table 2 displays the System specification.

Table 2

System specifications

Hardware specifications	Hard disk	512 GB
	RAM	16 GB
Software specifications	Simulation tools	MATLAB R2023a
	OS	Windows 11(64- bit)

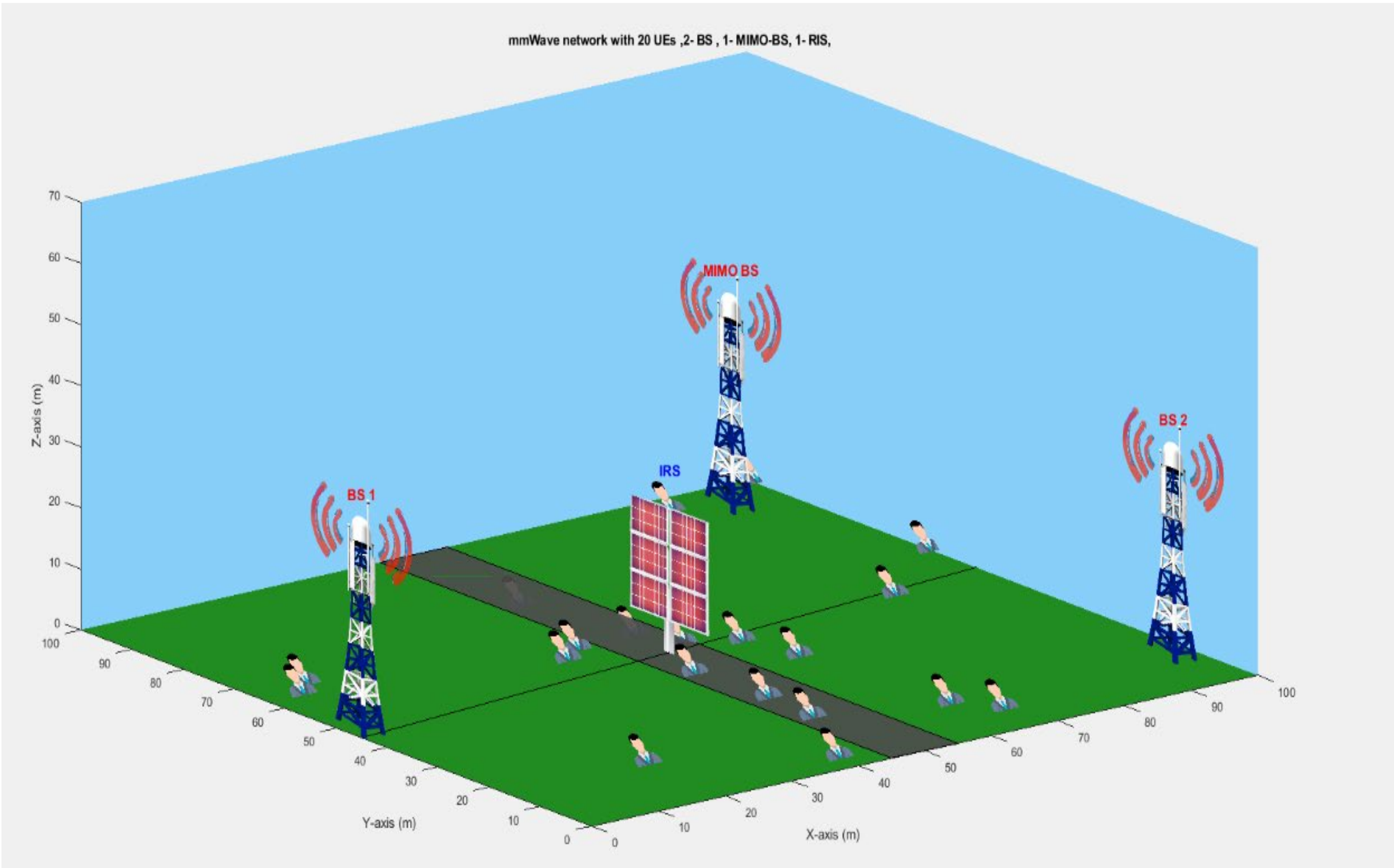

Fig. 5 Mmwave environment

In Fig. 5, The mmwave network which consists of 20 -user equipment (UEs), 2 base station (BS), 1-MIMI-BS, 1-RIS and customized channel environment and parameters.

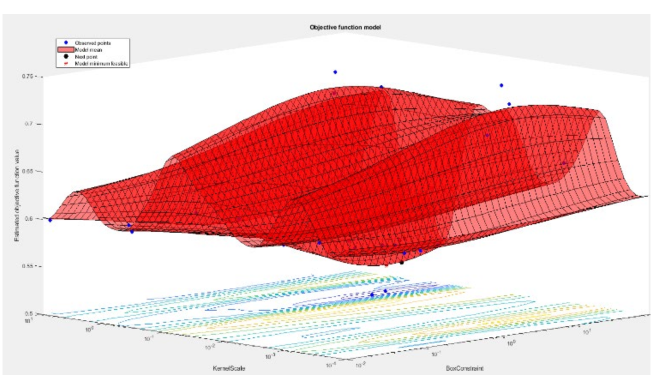

Fig. 6 Objective function Model

In Fig. 6, To estimate the channel by using Score based generative model enhanced support vector machine Bayesian optimization for mmwave fault diagnosis technique. The x axis taken as kernel scale, Box constraint and y axis as estimated objective function value.

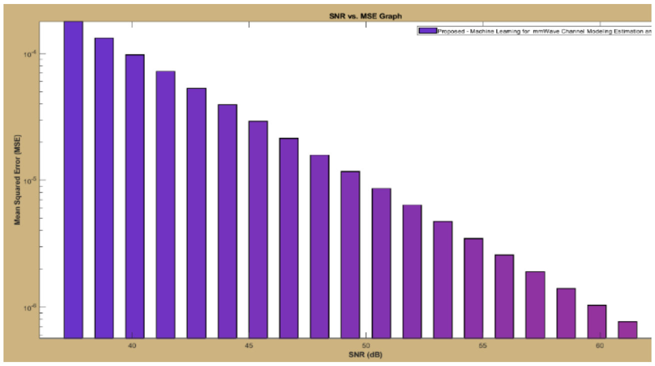

Fig. 7 SNR vs MSE

In Fig. 7, x axis represents SNR and y axis represents Mean squared error (MSE). The SNR increases, the received signal becomes clearer and less corrupted by noise, leading to lower MSE.

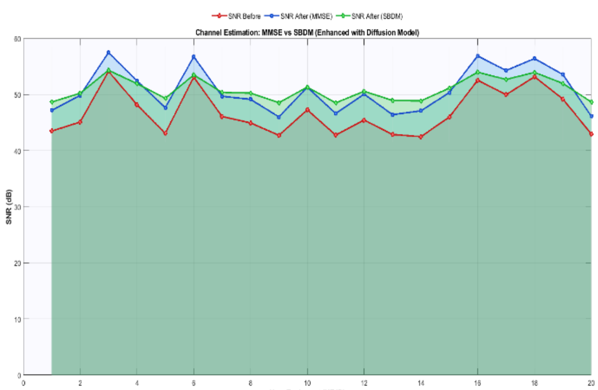

Fig.8 Enhanced with Diffusion Model

In Fig. 8, the channel estimation, User Equipment (UE ID) and SBR (dB) taken as x and y axis respectively. To enhance the data distribution, channel estimation and prediction using score-based diffusion model technique.

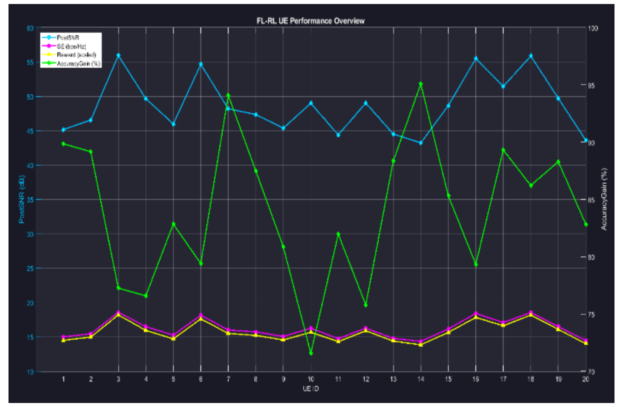

Fig. 9 FL-RL UE performance Overview

In Fig. 9, the UE ID and SNR (dB) are taken as x axis and y axis respectively. The federated learning-based training for hybrid beamforming, where a Generative Adversarial Network (GAN)is trained on ray tracing simulation data.

B. *Comparative analysis*

The proposed method is evaluated by comparing with many existing methods in the following domains: BER vs. SNR (dB), SNR (dB) vs. Spectral Efficiency (bits/s/Hz), Number of RIS elements vs. Transmission Rate (Mbps), Number of RIS elements vs. Spatial Correlation, SNR (dB) vs. MSE (Mean Squared Error). Compared to existing method such as DSSVAA (Directional Scanning Sounding and Virtual Antenna array) [1], DEMSQP (Data embedded multi-sub band quasi-perfect) [3] and DDFFCE (Data-driven frequency-flat cascaded channel estimation) model [13], the proposed model performed good and accurate.

a. BER vs SNR

The “Bit Error Rate (BER)” is a critical metric used to evaluate reliability of communication system. The “SNR” increases, the received signal becomes stronger relative to the noise, which reduces the probability of symbol errors. At low SNR values, noise dominates, resulting in higher BER, while at high SNR helps in selecting optimal modulation schemes and beamforming strategies.

Table 3

BER Vs SNR (dB)

X-axis (BER)	Y-axis (SNR (dB))

	Proposed	DSSVAA	DEMSQP
10^{-16}	10	35	75
10^{-12}	12	39	78
10^{-10}	15	44	81
10^{-5}	25	55	85
10^{-2}	30	65	90

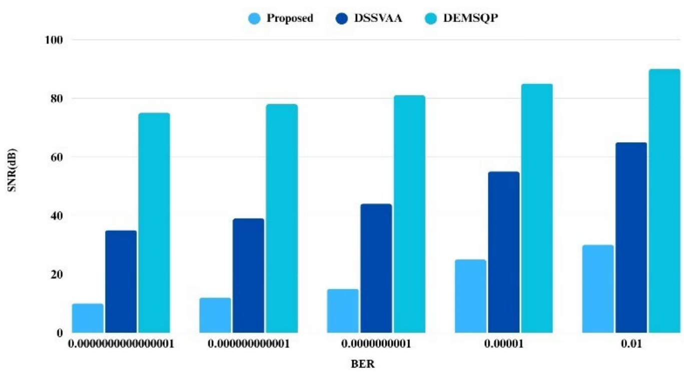

Fig. 10 BER vs SNR (dB)

In table 3 and Fig. 10 represents the variation of "Bit Error Rate (BER)" with "SNR" (dB). At a very low BER of 10^{-16} (0.0000000000000001), the proposed model requires only 10 dB, while ESPRIT and DEMSQP require 35 dB and 75 dB, respectively. As the BER increases to 10^{-2} (0.01), the proposed method reaches 30 dB, whereas DSSVAA and DEMSQP rise to 65 dB and 90 dB. The proposed model achieves "higher efficiency", robustness compared to the existing DSSVAA and DEMSQP methods.

b. SNR Vs Spectral efficiency

The SNR increases, the receiver can better distinguish transmitted signal from noise, resulted as higher achievable data rates and improved spectral efficiency. The relation between SNR and spectral efficiency is typically nonlinear, where efficiency increases logarithmically with SNR

Table 4
SNR (dB) vs Spectral efficiency

X-axis (SNR (dB))	Y-axis Spectral efficiency (mbps)		
	Proposed	DSSVAA	DEMSQP
35	15.5	5.5	1.5
39	16	6	2
45	16.5	7	3
55	18	7.5	4
65	19.5	8.5	4.5

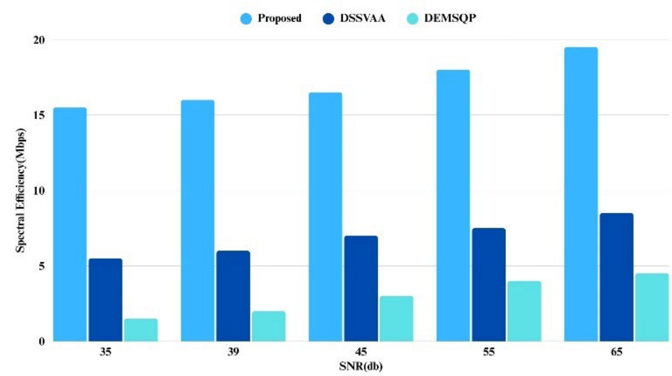

Fig. 11 SNR (dB) vs Spectral efficiency (Mbps)

In table (4) and Fig.11 represents the variation of spectral efficiency (Mbps) with respect to "SNR" in dB. At 35 dB, the proposed model achieves 15.5 Mbps, outperforming DSSVAA (5.5 Mbps) and DEMSQP (1.5 Mbps). Finally, the SNR rises to 65 dB, the proposed method reaches 19.5 Mbps, while DSSVAA and DEMSQP reaches 8.5 Mbps, respectively. The proposed method provides higher spectral efficiency and better performance under improved SNR conditions.

c. No. of RIS elements Vs Transmission rate (Mbps)

The transmission rate determines how much data is successfully transmitted per unit time, measured in megabits per second (Mbps). The number of RIS elements increases, effective channel gain and SNR also improve, leading to a higher transmission rate. The relationship can be expressed as:

$$R = Blog^2(1 + SNR^{eff}(N)) \qquad (116)$$

Where:

- R: transmission rate (bits/s or Mbps)
- B: Bandwidth (HZ)

$SNR^{eff}(N)$ = Effective SNR

Table 5
No. of RIS element vs Transmission rate(mbps)

X-axis (No.of. RIS element)	Y-axis Transmission rate(mbps)		
	Proposed	DSSVAA	DDFFCE
10	1	20	90

20	3	40	92
50	5	50	96
70	7	70	98
100	8	90	100

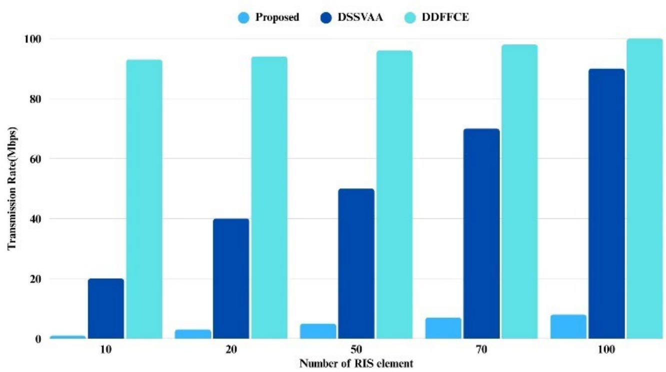

Fig. 12 Number of RIS element vs Transmission Rate (Mbps)

In Table 5 and Fig. 12 illustrate the relationship between the Number of RIS elements and the Transmission Rate (MBPs). At 10 elements, the proposed model achieves 1 Mbps, lower than DSSVAA (20 Mbps) and DDFFCE (90 Mbps). Finally, at 100 elements, the proposed model attains 8 Mbps, while DSSVAA and DDFFCE reach 90 Mbps and 100 Mbps, respectively, shoeing that the proposed model achieves stable growth with efficient use of RIS elements.

d. No. of RIS elements Vs Spatial correlation (Mbps)

The number of RIS elements increases, reflected signals experience more independent paths, resulting in reduced spatial correlation and improved channel diversity. This reduction enhances and handles multipath propagation that increases effective data transmission rate (Mbps).

$$\rho = e^{\frac{d\,N}{\lambda}} \qquad (117)$$

Where:

- ρ: spatial Correlation coefficient
- d: Spacing between RIS elements
- λ: Wavelength of the carrier signal
- N: Number of RIS elements

Table 6
No.of. RIS element vs Spatial correlation

X-axis (No.of. RIS element)	Y-axis spatial correlation (mbps)		
	Proposed	DSSVAA	DDFFCE
10	0.6	4	10
20	0.7	4.5	12
50	0.8	5	14
70	0.9	6	16
100	0.95	9	20

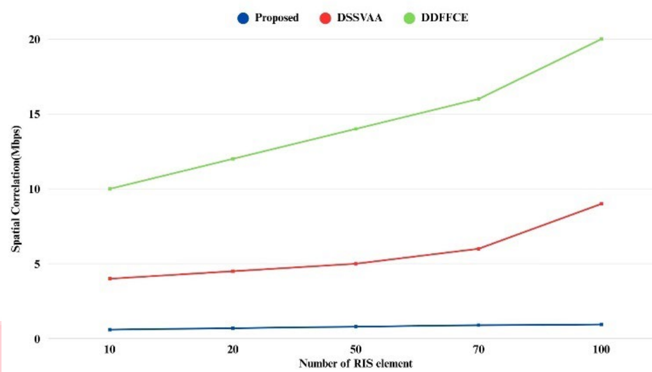

Fig. 13 Number of RIS element vs Spatial correlation (Mbps)

In table (6) and Fig.13 represents the number of RIS elements increase, spatial correlation improves for all models, indicating better signal alignment and system performance. Initially, at 10 elements the proposed model records a correlation of 0.6 Mbps, much lower than DSSVAA (4 Mbps) and DDFFCE (10 Mbps). Finally, the number of increases to 100, the proposed method achieves 0.95 Mbps, showing stable and controlled correlation compared to DSSVAA (9 Mbps) and DDFFCE (20 Mbps). The proposed model maintains the low spatial correlation, ensuring efficient signal reflection and improved communication reliability.

e. SNR vs Mean Squared Error

Mean Squared Error (MSE) measures “the average squared difference between the estimated and the actual signal values. The SNR quantifies the strength of the desired signal compared to background noise. The SNR increases, the received signal becomes clearer and less corrupted by noise, leading to lower MSE.

Table 7

SNR (dB) vs Mean Squared Error (MSE)

X-axis (SNR (dB))	Y-axis Mean Squared Error (MSE)		
	Proposed	DSSVAA	DDFFCE
40	1.75×10^{-6}	2.00×10^{-6}	3.15×10^{-6}
45	1.85×10^{-6}	2.75×10^{-6}	3.20×10^{-6}
50	1.89×10^{-6}	2.85×10^{-6}	3.25×10^{-6}
55	1.90×10^{-6}	2.89×10^{-6}	3.30×10^{-6}
60	1.98×10^{-6}	2.95×10^{-6}	4.00×10^{-6}

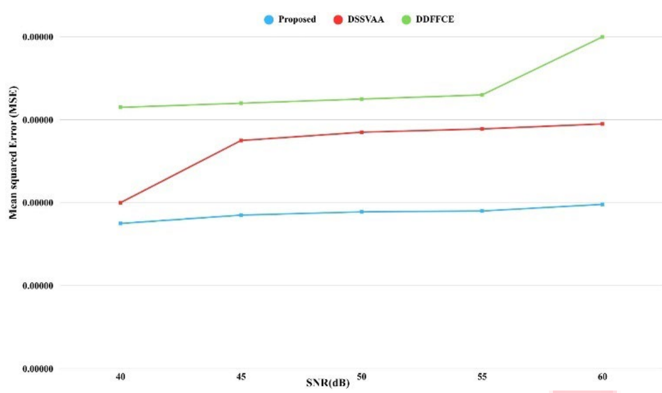

Fig. 14 SNR (dB)vs Mean Squared Error (MSE)

In table (7) and Fig. 14 represent the variation of Mean Squared Error (MSE) with respect to SNR (dB). In 40 dB, the model attains an MSE of 1.75×10^{-6} (0.00000175), outperforming DSSVAA as 1×10^{-4} (0.00000200) and DDFFCE as 1×10^{-3} (0.00000315). As the SNR increases to 60 dB, the method maintains a low MSE of 1.98×10^{-4} (0.00000198), while DSSVAA and DDFFCE exhibit higher errors of 1×10^{-3} (0.00000295) and 1×10^{-1} (0.00000400), respectively. This demonstrates that the proposed model provides superior estimation accuracy, better noise resistance, and more stable performance under varying SNR conditions compared to the existing methods

C. Research Summary

Initially we design the mmWave network which consists of 20- User Equipment (UEs), 2- Base stations (BS), 1- MIMO-BS, 1- RIS and customized channel environment and parameters. Next, we perform data preprocessing to reduce noise and enhance accuracy by applying high-SNR MIMO-MRC modeling technique, clustering algorithms, and filtering techniques to enable efficient feature extraction and improved data quality. Then, we perform Fast Alignment to quickly identify strong near-field regions, then apply HCSSP (Hierarchical Compressed Sensing with Sub-Partitioning) technique on those sub-regions to recover sparse channel paths. We perform (FL) based training for “hybrid beamforming”, where a Generative Adversarial Network (GAN) is trained on raytracing (RT) simulation data, and the base station (BS) aggregates user gradients in a decentralized manner without accessing raw data. Next, we estimate the channel by using SCOVEM (Score-Based Generative Model-Enhanced Support Vector Machine with Bayesian Optimization for mmWave Fault Diagnosis) technique. Then, we enhance the data distribution, channel estimation and prediction by using the Score based Diffusion model technique. Finally, we plot performance metrics such as BER vs. SNR (dB), SNR (dB) vs. Spectral Efficiency (bits/s/Hz), Number of RIS elements vs. Transmission Rate (Mbps), Number of RIS elements vs. Spatial Correlation and SNR (dB) vs. MSE (Mean Squared Error.

VI. CONCLUSION

The proposed paper presents an efficient machine learning based approach for mmwave channel estimation and prediction. The proposed model integrates MIMO-MRC preprocessing, federated learning, and score based generative model for robustness. The use of MIMO-MRC resulted in high signal reliability under noisy condition. By incorporating the Coordinated Multiple Point (CoMP) model, signal detection between transmitter and receiver. The federated learning framework improved training efficiency while preserving data security. The hierarchical compressed sensing with spatio- temporal Sub partitioning ensured accurate near and far filed modelling. The Score based generative model enhanced noise resilience and channel estimation accuracy. The model achieved higher transmission rates and lower spatial correlation through optimized RIS integration. Overall, the proposed hybrid system achieved superior robustness and scalability for mmwave communication. In future, the work can be extended to real time mmwave environments with dynamic user mobility. Further optimization using deep reinforcement learning can enhance adaptive beamforming and prediction accuracy. Integration with 6G intelligent reflecting surfaces and edge computing will enable ultra-reliable, low latency communication.

Reference

1. RIS-Aided mmWave, M. I. M. O. (2022). using Deep Learning and Compressive Sensing.
2. Zhang, R., Tan, W., Nie, W., Wu, X., & Liu, T. (2022). Deep learning-based channel estimation for mmwave massive MIMO systems in mixed-ADC architecture. *Sensors*, *22*(10), 3938.
3. Huttunen, J. M., Korpi, D., & Honkala, M. (2022). DeepTx: Deep learning beamforming with channel prediction. *IEEE Transactions on Wireless Communications*, *22*(3), 1855-1867.
4. Nguyen, C., Hoang, T. M., & Cheema, A. A. (2023). Channel estimation using CNN-LSTM in RIS-NOMA assisted 6G network. *IEEE Transactions on Machine Learning in Communications and Networking*, *1*, 43-60.
5. Wu, P., Cheng, J., Eldar, Y. C., & Cioffi, J. M. (2024). Learned Trimmed-Ridge Regression for Channel Estimation in Millimeter-Wave Massive MIMO. *IEEE Transactions on Communications*.
6. Raj, V., Nayak, N., & Kalyani, S. (2022). Deep reinforcement learning based blind mmWave MIMO beam alignment. *IEEE Transactions on Wireless Communications*, *21*(10), 8772-8785.
7. Qiao, Y., Niu, Y., Su, L., Mao, S., Wang, N., Zhong, Z., & Ai, B. (2024). Deep Reinforcement Learning based MmWave Beam Alignment for V2I Communications. *IEEE Transactions on Machine Learning in Communications and Networking*.
8. Sattari, M., Guo, H., Gündüz, D., Panahi, A., & Svensson, T. (2024). Full-duplex millimeter wave MIMO channel estimation: a neural network approach. *IEEE Transactions on Machine Learning in Communications and Networking*.
9. Bano, S., Cassarà, P., Tonellotto, N., & Gotta, A. (2023, June). A federated channel modeling system using generative neural networks. In *2023 IEEE 97th Vehicular Technology (VTC2023-Spring)* (pp. 1-5). IEEE.
10. Hu, Z., Chen, Y., & Han, C. (2023). PRINCE: A pruned AMP integrated deep CNN method for efficient channel estimation of millimeter-wave and terahertz ultra-massive MIMO systems. *IEEE Transactions on Wireless Communications*, *22*(11), 8066-8079.
11. Kim, D., Castellanos, M. R., & Heath, R. W. (2024). Joint Band Assignment and Beam Management Using Hierarchical Reinforcement Learning for Multi-Band Communication. *IEEE Transactions on Vehicular Technology*, *73*(9), 13451-13465.
12. Bacci, G., D'Amico, A. A., & Sanguinetti, L. (2024). MMSE channel estimation in large-scale MIMO: Improved robustness with reduced complexity. *IEEE Transactions on Wireless Communications*.
13. RIS-Aided mmWave, M. I. M. O. (2022). using Deep Learning and Compressive Sensing.
14. Guo, J., Chen, T., Jin, S., Li, G. Y., Wang, X., & Hou, X. (2024). Deep learning for joint channel estimation and feedback in massive MIMO systems. *Digital Communications and Networks*, *10*(1), 83-93.
15. Jawarneh, A., Kadoch, M., & Albataineh, Z. (2022). Decoupling energy efficient approach for hybrid precoding-based mmWave massive MIMO-NOMA with SWIPT. *IEEE Access*, *10*, 28868-28884.
16. Fola, E., Luo, Y., & Luo, C. (2024, December). AE-DENet: Enhancement for Deep Learning-based Channel Estimation in OFDM Systems. In *GLOBECOM 2024-2024 IEEE Global Communications* (pp. 1449-1454). IEEE.
17. Shen, H., Chen, Y., Han, C., & Yuan, J. (2025). Hybrid Beamforming with Widely-spaced-array for Multi-userCross-Near-and-Far-Field Communications. *IEEE Transactions on Communications*.
18. Shukla, V. B., Krejcar, O., Choi, K., Mishra, A. K., & Bhatia, V. (2024). Adaptive sparse channel estimator for irs-assisted mmwave hybrid mimo system. *IEEE Transactions on Cognitive Communications and Networking*, *10*(6), 2224-2235.
19. Zhu, X., Liu, Y., & Wang, C. X. (2023). Sub-array-based millimeter wave massive MIMO channel estimation. *IEEE Wireless Communications Letters*, *12*(9), 1608-1612.
20. Xu, H., Zhou, G., Wong, K. K., New, W. K., Wang, C., Chae, C. B., ... & Zhang, Y. (2023). Channel estimation for FAS-assisted multiuser mmWave systems. *IEEE Communications Letters*, *28*(3), 632-636.
21. Arvinte, M., & Tamir, J. I. (2022). MIMO channel estimation using score-based generative models. *IEEE Transactions on Wireless Communications*, *22*(6), 3698-3713.
22. Yang, S., Zhang, Z., Zhang, J., & Zhang, J. (2022). Impact of rotary-wing UAV wobbling on millimeter-wave air-to-ground wireless channel. *IEEE Transactions on Vehicular Technology*, *71*(9), 9174-9185.
23. Lyu, S., Li, X., Fan, T., Liu, J., & Shi, M. (2022). Deep learning for fast channel estimation in millimeter-wave MIMO systems. *Journal of systems engineering and electronics*, *33*(6), 1088-1095.

24. Liu, H., Zhang, Y., Zhang, X., El-Hajjar, M., & Yang, L. L. (2022). Deep learning assisted adaptive index modulation for mmWave communications with channel estimation. *IEEE Transactions on Vehicular Technology*, *71*(9), 9186-9201.
25. Tsukada, H., Suzuki, N., Bag, B., Takahashi, R., & Kim, M. (2024). Millimeter-wave urban celluar channel characterization and recipe for high-precision site-specific channel simulation. *IEEE Transactions on Vehicular Technology*.
26. Wei, X., Hu, C., & Dai, L. (2022). Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems. *IEEE Transactions on Communications*, *69*(1), 182-193.
27. Wang, L., Liu, G., Xue, J., & Wong, K. K. (2022). Channel prediction using ordinary differential equations for MIMO systems. *IEEE Transactions on Vehicular Technology*, *72*(2), 2111-2119.
28. Elbir, A. M., & Coleri, S. (2023). Federated learning for channel estimation in conventional and RIS-assisted massive MIMO. *IEEE transactions on wireless communications*, *21*(6), 4255-4268.
29. Wang, R., Klaine, P. V., Onireti, O., Sun, Y., Imran, M. A., & Zhang, L. (2022). Deep learning enabled beam tracking for non-line of sight millimeter wave communications. *IEEE open journal of the communications society*, *2*, 1710-1720.
30. Ma, W., Qi, C., Zhang, Z., & Cheng, J. (2022). Sparse channel estimation and hybrid precoding using deep learning for millimeter wave massive MIMO. *IEEE Transactions on Communications*, *68*(5), 2838-2849.
31. Qi, C., Wang, Y., & Li, G. Y. (2022). Deep learning for beam training in millimeter wave massive MIMO systems. *IEEE Transactions on Wireless Communications*.
32. Khan, M. Q., Gaber, A., Parvini, M., Schulz, P., & Fettweis, G. (2024). A low-complexity machine learning design for mmWave beam prediction. *IEEE Wireless Communications Letters*, *13*(6), 1551-1555.
33. Zheng, B., You, C., & Zhang, R. (2023). Efficient channel estimation for double-IRS aided multi-user MIMO system. *IEEE Transactions on Communications*, *69*(6), 3818-3832.
34. Abdallah, A., Celik, A., Mansour, M. M., & Eltawil, A. M. (2022, May). Deep-learning based channel estimation for RIS-aided mmWave systems with beam squint. In *ICC 2022-IEEE International on Communications* (pp. 1269-1275). IEEE.
35. Shen, W., Qin, Z., & Nallanathan, A. (2023). Deep learning for super-resolution channel estimation in reconfigurable intelligent surface aided systems. *IEEE Transactions on Communications*, *71*(3), 1491-1503.
36. Li, M., Zhang, F., Ji, Y., & Fan, W. (2022). Virtual antenna array with directional antennas for millimeter-wave channel characterization. *IEEE Transactions on Antennas and Propagation*, *70*(8), 6992-7003.
37. Pan, Y., Pan, C., Jin, S., & Wang, J. (2023). RIS-aided near-field localization and channel estimation for the terahertz system. *IEEE Journal of Selected Topics in Signal Processing*, *17*(4), 878-892.
38. Mao, T., Chen, J., Wang, Q., Han, C., Wang, Z., & Karagiannidis, G. K. (2022). Waveform design for joint sensing and communications in millimeter-wave and low terahertz bands. *IEEE Transactions on Communications*, *70*(10), 7023-7039.
39. Mao, K., Zhu, Q., Song, M., Li, H., Ning, B., Pedersen, G. F., & Fan, W. (2022). Machine-learning-based 3-D channel modeling for U2V mmWave communications. *IEEE Internet of Things Journal*, *9*(18), 17592-17607.
40. Liu, M., Li, X., Ning, B., Huang, C., Sun, S., & Yuen, C. (2022). Deep learning-based channel estimation for double-RIS aided massive MIMO system. *IEEE Wireless Communications Letters*, *12*(1), 70-74.