Deep Learning-Based Pulse-Shaping Filter Estimation for Fine-Grained WiFi Sensing

Han Hu*, Ruiqi Kong*, Ke Xu*, He (Henry) Chen*†
*Department of Information Engineering, The Chinese University of Hong Kong
†Shun Hing Institute of dvanced Engineering, The Chinese University of Hong Kong
Email: {hh022, kr020, xk020, he.chen}@ie.cuhk.edu.hk

***bstract*—In numerous WiFi sensing applications, such as passive human localization, the precision of sensing is often influenced by the estimation accuracy of multipath parameters. Several existing algorithms leverage pulse-shaping filter information to enhance multipath and channel estimation. However, WiFi chips do not disclose this filter information, and no current research has focused on measuring or estimating these pulse-shaping filters. In this paper, we introduce a new deep learning approach for the accurate estimation of pulse-shaping filters using channel state information (SI), which incorporates both multipath channel information and pulse-shaping filter information. Specifically, we construct a convolutional neural network consisting of an encoder-regressor architecture, where the encoder translates the SI into a latent representation, and the regressor subsequently estimates the pulse-shaping filter from this representation. Our proposed model's efficacy is demonstrated through its low normalized root mean squared error (NRMSE) in a variety of channel conditions, highlighting its ability to accurately estimate pulse-shaping filters.**

I. Introduction

In recent years, WiFi has evolved beyond simply interconnecting desktops, laptops, and smartphones, extending its connectivity to a diverse array of compact, intelligent IoT devices [1]. This development has catalyzed an exponential increase in interconnected WiFi devices, resulting in near-ubiquitous coverage with minimal blind spots across various environments [2]. This extensive coverage paves the way for a multitude of wireless sensing applications, including indoor localization, motion speed estimation, and vital sign detection. In the realm of wireless sensing, channel state information (CSI) serves as a critical data source. Derived from the channel frequency response within WiFi OFDM systems, CSI has become increasingly accessible due to tools (e.g., [3]) developed to extract it from commodity WiFi devices . These CSI measurements can reveal intricate details about the physical environment.

In indoor localization tasks, the accuracy of path delay estimation holds paramount importance as it directly determines the relative distance between the target and a WiFi device. Nonetheless, deriving a precise estimation of path delays from CSI presents a substantial challenge due to physical limitations and the implementation of pulse-shaping filters. The disappointing performance of the WiFi Fine Timing Measurement (FTM) in real-world environments [4], a feature introduced in IEEE 802.11-2016 [5], exemplifies these physical constraints, underscoring the difficulty of achieving accurate path delay estimations in practical settings.

In wireless systems, pulse-shaping filters comprise digital filters implemented at the transmitter and corresponding matched filters at the receiver. This combination is designed to ensure the system is band-limited while simultaneously minimizing the occurrence of bit errors at the receiver [6]. Within such band-limited systems, if the delay of a physical path is not an integer multiple of the sampling period, the multipath component in the discrete delay domain appears as multiple taps instead of a singular one—a phenomenon known as *channel leakage* [7]. significant issue arises from this channel leakage effect when two taps within the same pulse are interpreted as two physical paths with distinct delays, leading to a significant degradation of localization accuracy.

Several studies have recognized the leakage effect and addressed channel and delay estimation problems, assuming the pulse-shaping filter is known. Pulse information is used in scenarios employing S GE algorithms for precise channel and multipath parameter estimation [8], [9]. n augmented atomic norm-based approach for pilot-aided channel estimation is proposed in [10] to enhance the low-rank structure of the time-varying narrowband leaked channel. recent study in [11] presents a sparse Bayesian learning method for estimating the number of physical paths and their associated delay parameters, facilitated by known pulse-shaping filters. WiFi sensing tasks can directly benefit from the multipath parameter estimation in these algorithms. However, no existing research delineates a method for measuring or estimating the pulse. In practical systems, pulse-shaping filters are integrated within WiFi network interface cards. This means that we neither have access to the specific filter parameters nor can we infer concrete pulse information, which restricts the implementation of existing algorithms in real-world applications.

Theoretically, measuring the pulse-shaping filter by directly sampling the radio signal using specialized devices with sampling frequencies significantly higher than those of conventional Wi-Fi systems is possible. However, the pulse-shaping filter operates in a pair-matched manner, which means that the pulse is affected by the filters implemented on both the transmitter and receiver. This pair-matched operation produces unique pulse-shaping filters for different transceiver pairs, while directly measuring the radio signal over the air can only

This research was supported in part by project #MMT 79/22 of Shun Hing Institute of dvanced Engineering, The Chinese University of Hong Kong.

ICC 2024 - IEEE International Conference on Communications | 978-1-7281-9054-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICC51166.2024.10622819

obtain information on filters at the transmitter side. Without additional hardware knowledge of the receiver, recovering the pulse-shaping filter from captured radio signals is nearly infeasible. Given that CSI measurements encompass both multipath channel information and pulse-shaping filter information, this paper seeks to answer the question: *can we estimate the pulse-shaping filter directly from CSI measurements*? ccomplishing such an estimation would greatly streamline the practical implementation of pulse shape-augmented channel and multipath estimation algorithms. However, this task is far from straightforward, as constructing an ideal CSI measurement setup with fully controllable multipath parameters poses a significant challenge. For instance, path delays are considerably influenced by uncontrollable factors, such as signal processing delays and sampling time offsets.

s an initial effort to solve the question outlined above, this paper presents a deep learning model, centered around an encoder and a regressor, designed to estimate the time-domain samples of the pulse-shaping filter directly from CSI data, without requiring prior knowledge of multipath parameters. We opted for a deep-learning method over model-based algorithms to handle the high computational complexity that traditional estimation algorithms face when dealing with a large number of parameters. The primary goal of our deep learning model is to accurately extract the characteristics of the pulse-shaping filter and generate precise filter samples using only raw CSI as input. chieving this objective is non-trivial, primarily due to the application of the pulse-shaping filter as an intermediate convolution kernel in tandem with path delays during CSI formulation, and the fact that CSI data is often contaminated by noise. The encoder's design necessitates a careful redesign of contemporary CNN architectures to ensure that the regressor can accurately and robustly estimate the intricate pulse-shaping filter and multipath information from the encoded latent representation of noisy and unseen channels. Our evaluation metrics reveal a significant level of accuracy in estimating pulse-shaping filters using our simulated dataset. This dataset comprises CSI data derived from diverse multipath scenarios and encompasses 16 distinct pulse-shaping filters across four fundamental categories. We assess the performance of our estimation approach for each filter type using the normalized root mean squared error (NRMSE). The achieved NRMSE values, 0.0268 for fixed channel setups and 0.0232 for random channel configurations, indicate a consistently high accuracy in our estimations across a range of multipath conditions.

II. System Model

This section presents the signal models of CSI and pulse-shaping filters, channel leakage effect and problem formulation.

. *Channel State Information*

We consider the physical multipath channel in the time domain as

$$h_p\ \tau) = \sum_{l=}^{L\ \ 1} \ _l\delta\ \tau \quad \tau_l)\ , \tag{1}$$

where L represents the number of paths, while $\ _l$ and τ_l denote the complex amplitude and the delay of the l-th path, respectively.

Taking into account the effect of pulse-shaping, the channel can be expressed as

$$h\ \tau) = h_p\ \tau) \otimes g_t\ \tau) \otimes g_r\ \tau) \tag{2}$$

$$= \sum_{l=}^{L\ \ 1} \ _lg\ \tau \quad \tau_l)\ , \tag{3}$$

where $\otimes$ denotes the convolution operation. Here, $g_t\ \tau)$ and $g_r\ \tau)$ represent the pulse-shaping filter at the transmitter and receiver, respectively, with $g\ \tau) \quad g_t\ \tau) \otimes g_r\ \tau)$.

Upon discretizing the multipath channel in (2), we obtain

$$h_n = h\ nT) = \sum_{l=}^{L\ \ 1} \ _lg\ nT \quad \tau_l)\ ,\ n = 0, \cdots, N \quad 1\ , \tag{4}$$

where N corresponds to the maximum delay spread of the composite channel.

For a WiFi OFDM system with K subcarriers, the frequency response of the composite channel on subcarrier k is

$$H_k = \sum_{n=}^{N\ \ 1} h_n e^{\ j2\pi kn/K}\ ,\ k = 0, \cdots, K \quad 1\ . \tag{5}$$

We note that according to the IEEE 802.11 standards, only a subset of subcarriers is used for data/pilot symbol transmission [12]. In this paper, we concentrate on estimating pulse-shaping filters based on the subcarriers inherent in the practical IEEE 802.11n system designated for data/pilot transmission. Our framework can be seamlessly expanded to accommodate more advanced WiFi protocols.

B. Pulse-Shaping Filter

In multicarrier communication systems such as WiFi OFDM systems, choosing suitable filters is determinative of the correlation between symbols and the robustness of the scheme against dispersive channels [13]. Such filters dictate the pulse shape $g\ \tau)$ employed by the transceiver pair and can be defined by analytical models with control parameters as given in [13, Table I].

While a direct approach to solving the pulse estimation problem might be to estimate the key parameters from its mathematical representation, studies have shown that various filter types with different mathematical definitions can meet frequency response requirements [13]. In real systems, the pulse-shaping filter during the waveform generation process on the baseband can be regarded as the convolution of a series of prototype filters. Let $p_i\ \theta_i)$ denote the ith used prototype

filters, where θ_i is the parameter vector, the pulse-shaping filter used could be written as

$$g = p_1\ \theta_1) \otimes p_2\ \theta_2) \otimes \cdots \otimes p_i\ \theta_i) \otimes \cdots \otimes p_n\ \theta_n)\ .$$

Given the absence of prior knowledge p_i, θ_i and n, it becomes nonviable to estimate the parameters of an unknown mathematical formula. Consequently, we have chosen to estimate the vector that contains the highly precise sampled points of the filter in time domain, using this vector as a representation of the pulse-shaping filter.

C. Channel Leakage Effect and Problem Formulation

The channel leakage effect discussed in [7] stated that in the discrete channel impulse response, the specular scattering with fixed delay does not consist of Dirac-like functions at the delay points of the scatterers. The formulation of the discrete channel impulse response in [7] suggested that the leakage effect is due to the finite transmit bandwidth and can be characterized by a function. Such system limitation is brought by the use of pulse-shaping filters in real systems, and the taps are produced by the time offset in sampling the function of the pulse-shaping filter. n example of the channel impulse response (CIR) with the leakage effect is illustrated in Fig. 1, in which the system applies a truncated raised-cosine filter with a length of 16 and has a sampling period of $T = 50$ ns while one path arrives at $\tau = 25$ ns. Consequently, a pulse with 16 taps is generated, leading to the challenge of determining whether the two taps with the highest amplitude represent two distinct physical paths or are two taps within a single pulse.

The phenomenon observed in Fig. 1 justifies the necessity of estimating the pulse-shaping filter before path delay estimation. Therefore, we define our problem as estimating a vector $\mathbf{g}$ of highly precise sampled points of the pulse-shaping filter $g\ \cdot)$ from raw CSI H_k captured from WiFi OFDM symbols. Based on the mathematical definition, we can infer that estimating the function $g\ \cdot)$ is equivalent to retrieving the intermediate convolution kernel from the formulation of the channel state information, as defined in (4), without any prior knowledge of the channel parameters. To solve this blind convolution kernel estimation problem from CSI, we propose to use deep convolutional networks to estimate the pulse-shaping filter. This is motivated by the image blind deblurring tasks (e.g., [14]) and deep learning CSI feedback systems (e.g., [15]).

III. Proposed Deep-le rning Model

. Overview

The network design is inspired by the encoder-codeword-decoder structure of the CSI feedback task, as described in [15]. However, we use a different representation for the codeword, and the decoder is replaced with a regressor to align with our system model outlined in Sec. II-B. For a detailed design of the network structure, we reference recent machine learning papers on CSI [16], and employ modern CNNs complemented by the latest attention mechanism.

To improve the accuracy of pulse-shaping filter estimation, the encoder's primary goal is to isolate channel features that

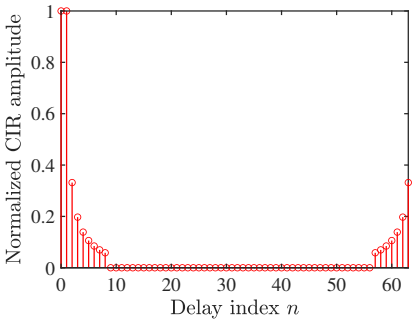

Fig. 1. Illustration of the channel leakage effect with a single path delay of 25 ns and a sampling period of 50 ns, here aligns with in (4).

are influenced by the filter, while minimizing the impact of multipath effects. This targeted approach is important for identifying the unique distortions caused by different pulse-shaping filters. To achieve this, we carefully group and scale the CSI data during the synthesis process. This step is critical because the network requires clear and comprehensive data representations to accurately estimate pulse-shaping filters. single CSI snapshot is limited to one environment's delay profile, but by grouping similar data, we can highlight relevant features and reduce the impact of 'feature noise' caused by varying path delays in the CSI collection. This approach helps the encoder isolate relevant features and enhances the regressor's ability to accurately map these features to specific pulse-shaping filters.

fter applying the data synthesis, we utilize the system model of CSI along with our problem formulation to guide the design of our encoder network. Our objective is to estimate $\mathbf{g}$ from a group of given H_k in an IEEE 802.11n system. The formulation of H_k for each subcarrier k in the IEEE 802.11n setting results in a small dimension for our input CSI. This motivates us to apply a small 1D convolutional kernel across the subcarriers of a single CSI in the encoder network for feature extraction. Finally, to eliminate the ambiguity caused by the multipath environment, we align the network with the grouping step in data synthesis by employing a 2D convolutional kernel across the grouped CSIs.

During the training process, the pulse-shaping filter has a distinct property where its center value is several orders of magnitude larger than its tail values as shown in Fig. 3. s a result, a loss function that is sensitive to both minuscule and substantial values while remaining insensitive to outliers is required. To address this, we use the smooth L1 loss [17] which is widely recognized for its robustness in regression tasks, but with an adjusted control point.

B. Data Synthesis

In order to efficiently extract the features of the pulse-shaping filters, it is crucial to provide the encoder with a well-organized and informative dataset. We opted to use simulation data for both training and testing instead of real-world data, as it requires precise labeling with ground truth.

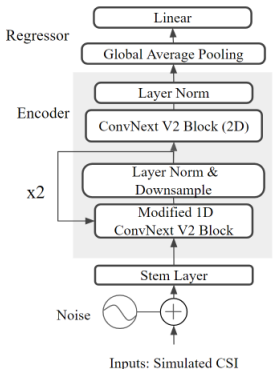

Fig. 2. Network structure of the proposed model.

This approach ensures better accuracy and effectiveness of the training process.

We design finite impulse response (FIR) filters in accordance with the IEEE 802.11n spectral mask [12] and produce the corresponding CSIs using M TL B simulation within a fixed multipath environment. We also account for the time unsynchronization between transceivers to ensure the simulated data closely mirrors experimental data. Specifically, a random sample time offset of $[0, 25]$ ns aligning with the bandwidth of our channel is introduced for all paths. We then group the CSIs using same filter but with random sample time offsets to form a 3D data bundle represented as

$$^{b} = \begin{bmatrix} ^{1}_{1} & \cdots & ^{1}_{n} \\ \vdots & \ddots & \vdots \\ ^{m}_{1} & \cdots & ^{m}_{n} \end{bmatrix},$$

where $^{m}_{n}$ represents the nth CSI using filter m.

This grouping step ensures adequate snapshots of the applied pulse-shaping filter in every group of CSI and enables our proposed deep-learning network to perform 2D convolution to extract features for pulse-shaping filters across CSIs rather than only among subcarriers in a single CSI. This step further highlights the features of the pulse-shaping filters in the CSIs as the equation (4) states that they can be corrupted by the path delay during the formulation of the CSI.

Prior to feeding the CSI into the CNN, we decompose each complex CSI into two real vectors: one representing the amplitude of each subcarrier in the CSI and another for the phase of each subcarrier in the CSI. The decomposition of the data is because general CNNs cannot directly process complex numbers, and we apply a 2-channel training strategy for our model.

C. Network Design

We now delve into the designs of the main components of our proposed model as shown in Figure 2.

The encoder converts the preprocessed CSI samples into a latent representation. For this purpose, we propose a variant of `ConvNext V2` [18] as our encoder. Given that ConvNext V2 was primarily structured for pixel-level feature extraction in 2D RGB images, and considering that the CSI is in a 1D format, we have made major changes to the blocks of ConvNext V2. Specifically, we eliminate the downsampling step in the stem layer following the initial input and modify the convolutional layers and downsample layers. The convolutional layers' size is adjusted from 7 in 2D to 3 in 1D to enable the network to extract features from neighboring subcarriers in the CSI of a single OFDM symbol. We preserve the progressive sizes of the 'feature' dimension in the original `ConvNext V2` model but with less depth used during the progression of each 'feature' dimension as the size of our dataset is significantly smaller than the image dataset for the original `ConvNext V2` model. The downsampling is conducted over the subcarrier dimension. This setup enables the network to transition from discerning local features derived from each subcarrier to understanding global ones sourced from groups of adjacent subcarriers. Finally, we consider the grouping step in data synthesis to retain the last block of `ConvNext V2` unchanged. This medium-sized 2D convolution block allows the network to grasp the commonalities of the CSI data with the same pulse-shaping filter applied after each being thoroughly processed and downsampled by the 1D convolutional layers.

The regressor translates the representations into the estimated pulse shape with the smooth L1 loss applied. Considering the use of real-world pulse-shaping filters, as discussed in Sec. II-B, we employ a global average pooling layer with w neurons to derive the estimation of each individual sampled point of the pulse-shaping filter. Here, w is contingent on the number of sample points the application aims to estimate.

D. Loss Function

We employ the smooth L1 loss [17] for supervised learning. The smooth L1 loss uses the L2 loss for targets within the range $[0, \beta]$, and the L1 loss beyond β to avoid excessively penalizing outliers. Specifically, it is defined as

$$f\ x) = \begin{cases} 0.5^{\underline{x}}, & \text{if } |x| < \beta, \\ |x| \quad 0.5\beta, & \text{otherwise.} \end{cases} \tag{6}$$

IV. Perform nce Ev lu tion

. Simulation Setup

We crafted 16 finite impulse response (FIR) filters aligned with the IEEE 802.11n spectral mask, each possessing distinct configurations across four primary types: raised cosine filters, FIR filters with Kaiser window, FIR filters with Tukey window, and FIR filters with Hann window. Each filter spans a fixed length of 16 samples. To ensure precise model training, we oversampled these filters by a factor of 50, producing a total of 801 sampled points with unit power normalization. The added 1 point is automatically generated by M TL B to maintain symmetry. Example filters are given in Fig. 3.

To generate various delay profiles, we used `comm.RayleighChannel` with a configuration in compliance with the IEEE 802.11n standard, operating in a 20MHz channel. By default, the M TL B Communications Toolbox generates CSI by taking path delays and loss gains as parameters

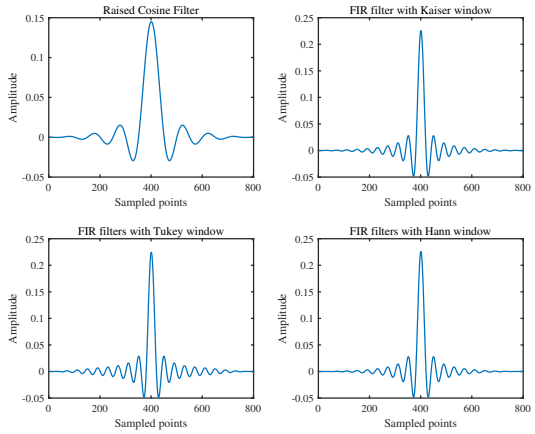

Fig. 3. Example of crafted filters

and using a general pulse-shaping filter. We replaced this generic filter with our specifically designed ones to produce the corresponding CSIs.

We implemented a fixed physical 4-path environment for the training dataset, enforcing consistent path delay intervals and path gain for every path. The random sampling time offset is indiscriminately selected from the range $[0, 50]$ ns. For the test dataset, we employed an additional randomized multipath environment with different path delays and path gain. The parameters of the environment are limited in the maximum path gain being 35dB in SNR and the maximum path delay being 150 ns. These settings ensure our objective to gauge the generalizability of the encoder and regressor.

fter deriving the CSIs from the simulation, we apply data synthesis as explained in Sec. III-B. For the training dataset, we pick $m = 2000$ and $n = 100$ with 16 crafted filters being randomly chosen. We then create the test dataset by choosing different primary types of the filter separately, with parameters that differ from the ones in the training set. This resulted in a total of 320 groups of CSIs, each containing 100 CSIs collected in both randomized and fixed multipath environments. Specifically, for each scenario, we selected 50 groups each for raised cosine filters, FIR filters with Kaiser window, and FIR filters with Tukey window filters as they contain more than one control parameter. dditionally, we picked 10 groups for FIR filters with Hann window.

B. Training details

We train the network using the damW optimizer, starting with an initial learning rate of 1×10^{-4}. We employ the `Cosine nnealingLR` learning rate scheduler to adjust the learning rate throughout the training process dynamically. We set the batch sizes to 16 and the boundary β used for clipping to 0.1. The proposed network is trained on a system furnished with an NVIDI RTX 6000.

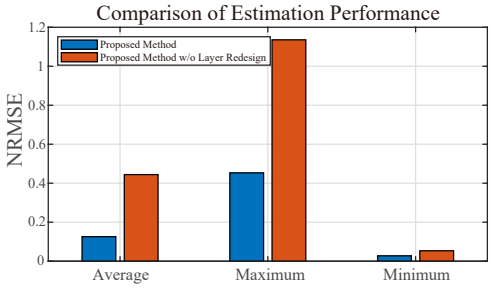

Fig. 4. Performance comparison of proposed method with and without layer redesign.

C. Performance Evaluation

To assess the efficacy of our proposed model, we carried out a performance evaluation using NRMSE to compare the estimated values against the actual ones across 16 different types of filters. The NRMSE is calculated by dividing the RMSE by the mean of the estimation.

We computed the NRMSE over all 801 estimated points for each single filter within the test dataset and obtained the average, maximum, and minimum NRMSE over different filters used in our test set. This includes both randomized and fixed multipath environments. The results are presented in Table I.

T BLE I
ESTIM TION PERFORM NCE FOR LL S MPLED POINTS

Evaluation \ Environment	Random	Fixed
vg. NRMSE	0.0232	0.0268
Min. NRMSE	0.0014	0.0009
Max. NRMSE	0.0781	0.1044

D. blation Experiment

The ablation experiment aims to investigate the effectiveness of our proposed method by removing certain components in our deep network to understand their contributions. We performed two ablation experiments to gauge the effects of the network designs and data synthesis techniques on the accuracy of pulse-shaping filter estimation.

First, we aim to show the effectiveness of the dimensional redesign of the convolution layers in our proposed model. This redesign of the convolutional layers aims to extract features from neighboring subcarriers from the CSI in a single OFDM symbol. To do this, we compare the estimation performance in random multipath environments between our proposed method and our proposed method without layer redesign.

The performance of our proposed method without layer redesign is notably poor, with significant estimation accuracy degradation compared to our proposed method as shown in Fig. 4. The results indicate that reducing the dimensions of input data aggressively and using large convolution kernels for image tasks are unsuitable for CSI data as they discard information from neighboring subcarriers. The poor estimation

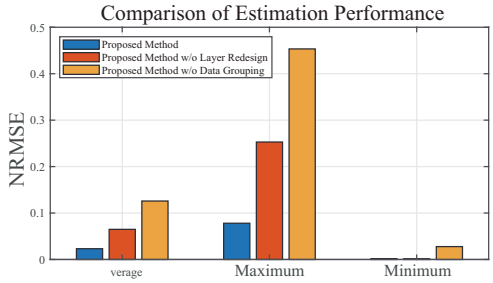

Fig. 5. Performance comparison of proposed method with and without data grouping.

performance of our proposed method without layer redesign indicates that there is significant feature loss during the training process. Our design can prevent this feature loss, resulting in more accurate estimation results.

Next, we investigate the effectiveness of our proposed data synthesis step by performance evaluation among two control groups and our proposed model in Sec. III-C. We first use the original `ConvNext V2` model with 2D convolution kernels to serve as the control group representing the proposed method without layer redesign. The second control group is chosen as the model with our proposed redesign in the first experiment, serving as the proposed method without data grouping. Different estimation accuracy degradation can be found in Fig. 5, with our proposed model significantly outperforming two control group models. The reason for the performance differences is that the resolution of the CSI from a single OFDM symbol is much smaller than the impulse response of the pulse-shaping filter.

This result suggests the importance of grouping the CSIs with the same applied filter but with random sampling time offsets and guiding the training process to extract 'features' among CSIs with different sampled points of the pulse-shaping filter. The 'features' are then combined in the final linear layer to estimate a high-resolution vector representation of the pulse-shaping filter. By doing so, our proposed model is able to effectively learn and extract distortions introduced by the pulse-shaping filter rather than multipath environments.

V. Conclusions

The use of pulse-shaping filters in wireless communication systems can cause channel leakage effect, which can significantly impact the accuracy of delay estimation for a multipath environment. In this paper, we propose a new deep-learning model that can estimate high-precision samples of pulse-shaping filters from raw CSI without prior knowledge of multipath parameters. Our proposed model has been evaluated based on simulation data that shows its proficiency to estimate pulse-shaping filters under various unseen multipath environments, with the NRMSE being 0.0268 and 0.0232 in fixed and random multipath environments, respectively. Our proposed network structure and data synthesis techniques greatly enhance model accuracy by carefully incorporating the relationships between neighboring subcarriers and the similarities under different multipath settings caused by the pulse shaping filters in the formation of the CSI, as demonstrated by additional ablation studies.

References

[1] Y. Ma, G. Zhou, and S. Wang, "WiFi sensing with channel state information: survey," *CM Computing Surveys (CSUR)*, vol. 52, no. 3, pp. 1–36, 2019.

[2] S. Tan, Y. Ren, J. Yang, and Y. Chen, "Commodity WiFi sensing in ten years: Status, challenges, and opportunities," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17 832–17 843, 2022.

[3] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your CSI: channel state information extraction platform for modern Wi-Fi chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, 2019, pp. 21–28.

[4] S. ggarwal, R. K. Sheshadri, K. Sundaresan, and D. Koutsonikolas, "Is WiFi 802.11 mc fine time measurement ready for prime-time localization?" in *Proceedings of the 16th CM Workshop on Wireless Network Testbeds, Experimental evaluation & CHaracterization*, 2022, pp. 1–8.

[5] "IEEE standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements - part 11: Wireless L N medium access control (M C) and physical layer (PHY) specifications," *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pp. 1–3534, 2016.

[6] K. Gentile, "Digital pulse-shaping filter basics," *pplication Note N-922. nalog Devices, Inc.*, 2007.

[7] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE Journal of selected topics in signal processing*, vol. 4, no. 2, pp. 255–271, 2010.

[8] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the S GE algorithm," *IEEE Journal on selected areas in communications*, vol. 17, no. 3, pp. 434–450, 1999.

[9] D. Shutin and B. H. Fleury, "Sparse variational bayesian S GE algorithm with application to the estimation of multipath wireless channels," *IEEE Transactions on signal processing*, vol. 59, no. 8, pp. 3609–3623, 2011.

[10] J. Li and U. Mitra, "Improved atomic norm based time-varying multipath channel estimation," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 6225–6235, 2021.

[11] K. Xu, H. Chen, and C. Wu, "Pulse shape-aided multipath delay estimation for fine-grained WiFi sensing," *arXiv preprint arXiv:2306.15320*, 2023.

[12] "IEEE standard for information technology– local and metropolitan area networks– specific requirements– part 11: Wireless L N medium access control (M C) and physical layer (PHY) specifications amendment 5: Enhancements for higher throughput," *IEEE Std 802.11n-2009 (mendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, and IEEE Std 802.11w-2009)*, pp. 1–565, 2009.

[13] . Sahin, I. Guvenc, and H. rslan, " survey on multicarrier communications: Prototype filters, lattice structures, and implementation aspects," *IEEE communications surveys & tutorials*, vol. 16, no. 3, pp. 1312–1338, 2013.

[14] K. Zhang, W. Ren, W. Luo, W.-S. Lai, B. Stenger, M.-H. Yang, and H. Li, "Deep image deblurring: survey," *International Journal of Computer Vision*, vol. 130, no. 9, pp. 2103–2130, 2022.

[15] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Overview of deep learning-based CSI feedback in massive MIMO systems," *IEEE Transactions on Communications*, vol. 70, no. 12, pp. 8017–8045, 2022.

[16] B. Zhang, H. Sifaou, and G. Y. Li, "CSI-fingerprinting indoor localization via attention-augmented residual convolutional neural network," *IEEE Transactions on Wireless Communications*, 2023.

[17] R. Girshick, "Fast R-CNN," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[18] S. Woo, S. Debnath, R. Hu, X. Chen, Z. Liu, I. S. Kweon, and S. Xie, "ConvNeXt V2: Co-designing and scaling convnets with masked autoencoders," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 16 133–16 142.