A Novel Learning-based Estimation Scheme for Communication over Impulsive Noise Channels

Akash Kumar Mandal and Swades De

***Abstract*—This paper proposes a novel statistical hybrid neural network (S-HNN) based estimation of impulse noise infested wireless communication channels. Spatial fading characteristics are found using a convolutional neural network (CNN), while long short-term memory (LSTM) network extracts temporal information over subsequent time horizons. Finite lag samples are employed to extract the channel gain distribution based on multiple recycling of the CNN-LSTM network. The proposed S-HNN framework for orthogonal frequency division multiplexed (OFDM) communication channel with subcarrier spacing of 15 kHz, sampling rate of 15.36 MHz, IFFT size of 1024, and various pilot density deployments is shown to outperform the existing state-of-the-art channel estimation techniques in terms of 50% reduced training length and nearly 49% saving in training time.**

***Index Terms*—Impulse noise, smart grid communication, statistical hybrid neural network, wireless channel estimation**

I. Introduction

With growing wireless infrastructure in system monitoring, control, and automation, impulse noise scenarios have become prevalent in communication setups. Such noise may originate from power lines, motors, high intensity lightning, pulse-type radars, etc. [1]. One such scenario that has gained significance is smart grid monitoring. Due to increased disturbances in the power grids, real-time monitoring of system parameters has gained importance for ensuring reliable operation [2]. Orthogonal frequency division multiplexing (OFDM) provides some robustness against such impulsivity by spreading the noise power over multiple subcarriers. However, its performance degrades sharply when the power or degree of impulsiveness in the noise exceeds a certain threshold [3]. The existing channel estimators are not adaptive to the spatio-temporal dynamics of the impulsive noise environment, and hence are not robust.

A. Literature Review and Motivation

The channel estimation techniques in literature use either mathematical analysis of wireless channel or learning methods for finding the channel coefficients. In the first bracket, least square (LS) and minimum mean square error (MMSE) are the two typical pilot-based methods. LS assumes no additional channel information and provides a simple structure for channel estimation, however with an inferior performance compared to MMSE based measure [4]. The MMSE based estimator uses second order statistics for channel estimation at the cost of complex matrix operations. In [5], a closed-form of impulsive noise channel model is obtained, but it does not provide a channel estimation approach in dynamic impulsive noise environments.

The shortcomings of conventional MMSE were addressed in data-driven MMSE estimators in [6] and using an implementation of LS in obtaining sampled covariance matrix to reduce processing order in [7]. The study in [8] solved the problem of pilot contamination using large number of antennas. Though a decent reliability is assured using such advances in statistical channel estimation, all these methods assume some prior channel knowledge. Such assumptions are hard to meet in practical deployment scenarios. Moreover, the information of noise statistics become scarce as the noise becomes impulsive, which is the case in smart grid communication.

Learning-based estimators relax prior statistical information, thereby gaining widespread popularity. An implementation of convolutional neural network (CNN) and long short term memory (LSTM) based bi-stage offline-online training for extracting channel information is demonstrated in [9], while a multi-layer perceptron based channel estimation is suggested in [10]. However, these techniques suffer from model reliability and implementation complexity. The work in [11] alleviates this issue by using spatial-frequency-temporal CNN for channel estimation, by exploiting correlation among these three aspects of the channel, thus improving estimation accuracy.

However, as prevalent in power grid environments, owing to the dynamic nature of source generation rate, source lifetime, impulse generation rate, impulse duration, and noise floor level in impulse noises, the existing techniques are inadequate to deal with impulse noise, which can significantly affect the performance optimization of such communication systems [12]. The modified convolutional blind denoising network (CBDNet) can tackle conventional communication noise by adjusting its level map [13]. But, its performance degrades when the noise dynamics are impulsive [14]. Moreover, the use of high dimensional dataset renders such implementations unsuitable to the dynamics presented by impulse noise [15].

As noted above, statistical estimators either under-perform or require complex computation with prior channel knowledge. In contrast, learning based estimators ignore the domain information, resulting in unstable performance and high training costs, especially in highly dynamic conditions. The study in [16] considered a mix of MMSE based measure along with learning based channel estimation in a knowledge-driven machine learning (KDML) framework. However, its inability in capturing spatial information in the channel coefficients makes the performance inadequate in impulse noise scenarios.

We note that, on one hand, the dynamic nature of impulsive noise channel does not allow tractable and accurate channel characterization through the standard analytical approach of

The authors are with the Department of Electrical Engineering and Bharti School of Telecommunication, IIT Delhi, New Delhi, 110016, India.

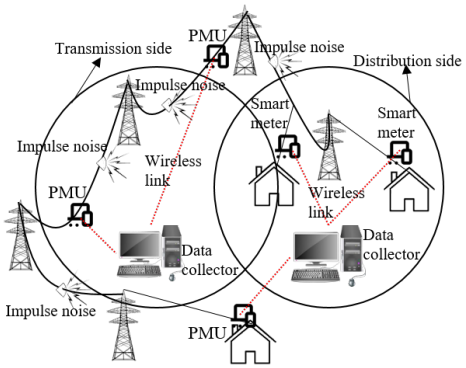


Figure 1: Impulse noise infested smart grid wireless communication.

fading channel modeling. On the other hand, owing to prevalent impulse noise sources in industrial environments, Internet-of-Things (IoT) communications frequently encounter such noise. This letter aims to fill the lacuna of robust channel estimation for communication over impulsive noise channels. A statistical hybrid neural network (S-HNN) based approach to channel estimation is proposed for an OFDM communication system in impulsive noise environments.

B. Contributions and Scope of Application

The key contributions of this work are: 1) A S-HNN based channel estimation framework is proposed for estimation of communication channel over spatio-temporal impulsive noise channel, wherein partial statistics of the channel is used to enhance reliability and reduce the learning input data size. 2) LS-aided MMSE is used as a statistical method to generate approximate channel estimates while addressing time and frequency selective fading in OFDM systems. 3) Deep CNN is used to find spatial estimates of the coarse channel simultaneously across all subcarriers, whereas LSTM is employed to handle the temporal effects of fading. 4) Optimum lag estimates are used to find the channel density function for each subcarrier using multiple recycling of CNN-LSTM network.

Simulation results demonstrate a sufficiently improved performance of the proposed S-HNN estimator over the existing competitive statistical and learning-based estimators in terms of bounded nRMSE, short 30-40 training epochs with 2 retrainings per day, training length saving of 50%, and 48.71% training time saving. This framework can be used for channel estimation in vehicular, power line, underwater acoustic, and plasma channel communication.

In the following, Section II describes the system model; Section III presents S-HNN channel estimation; and results and conclusion are presented in Sections IV and V, respectively.

II. System Model for Wireless Communication

As an example of impulsive noise environment, a smart grid communication scenario is shown in Fig. 1. Impulse noise is originated from ionization around the high voltage lines, represented by solid curved lines. Ambient network monitoring is performed using IoT devices, such as phasor measurement units (PMUs), circuit breakers, and smart meters [17]. It is notable that a monitoring scenario for the whole network implies a federated monitoring of the transmission and distribution systems, with joint data processing at the local or central data collectors. These IoT devices communicate with the local/central data collector over multiple OFDM subcarriers. In an imperfect channel state information (CSI) scenario, the receiver works in coordination with the transmitter to generate channel estimates, which are fine-tuned later using the S-HNN channel estimator. The proposed channel estimation is deployed at the receiver in a data-aided channel estimation approach, for estimation of impulse noise infested wireless channels. The algorithm is implemented on an E3-1285 CPU @4.10 GHz clock frequency for various channel dynamics.

III. S-HNN based Channel Estimation

This section presents the proposed S-HNN based channel estimation framework for an OFDM communication over spatio-temporal impulsive noise channels, as in Fig. 1.

A. Impulse Noise Impaired Communication Channel Model

The transmitted data X_s is fed through a serial-to-parallel converter, which is then partitioned into blocks of size N. All the data blocks undergo appropriate modulation, after which orthogonal subcarriers are allocated to them by N-point IFFT operations. Pilots are placed in X_s at equal intervals, which is known at both transmitter and receiver. The OFDM symbols are transmitted through the channel H_s. Using Jake's model with K plane waves arriving uniformly from all directions to capture the time selective fading, we have $g(t) = \frac{E_0}{\sqrt{2K_0+1}}\left(g_I(t) + jg_Q(t)\right)$, where $K_0 = \frac{1}{2}\left(\frac{K}{2}-1\right)$ and E_0 is the average fading channel amplitude. The in-phase and quadrature components are respectively expressed as: $g_I(t) = 2\sum_{m=1}^{K_0}\cos\phi_m\cos\omega_m t + \sqrt{2}\cos\phi_K\cos\omega_D t$ and $g_Q(t) = 2\sum_{m=1}^{K_0}\sin\phi_m\cos\omega_m t + \sqrt{2}\sin\phi_K\cos\omega_D t$, where ϕ_m is the angle of arrival of the m-th incoming wave on the receiver, ω_d is the maximum Doppler shift, and $\omega_m = \omega_d\cos\left(\frac{2\pi m}{K}\right)$ for $m \in 1, \cdots, K_0$. The initial phases are set to yield an uniform phase distribution.

Tapped delay line model is used to capture frequency selective fading. The channel response is expressed as $h(t) = \sum_{i=1}^{l}\sqrt{P_i}g_i(t)\delta(t-\tau_i)$, where l is the number of distinguishable multipaths, P_i is the power of the i-th multipath and τ_i is its delay. Thus, the frequency domain received data is

$$Y_s = H_s X_s + F + E = H_s X_s + W \tag{1}$$

where Y_s is the received data, X_s is the transmitted data, H_s is the frequency domain channel impulse response, F is the receiver thermal noise, and E is the electromagnetic impulse noise peculiar to smart grid communications. The statistics of W is detailed in [5]. Next section explains the proposed S-HNN framework and the time complexity involved in the estimation of an impulse noise infested channel.

B. S-HNN Channel Estimation Framework

Let $\mathbf{X}$, $\mathbf{Y}$, and ($\mathbf{W}$, $\boldsymbol{\theta}$) be the input, output, and parameter space of the S-HNN, respectively (c.f. Fig. 2). From (1), the MMSE estimate of the channel is $\hat{H}_{MMSE} = R_{H_sH_s}\left(R_{H_sH_s}\right.$

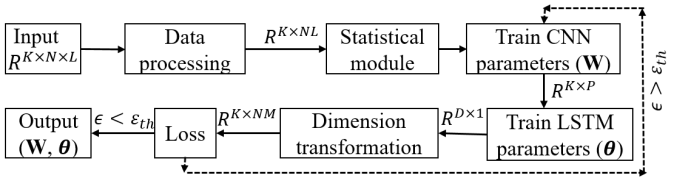


Figure 2: S-HNN based smart grid channel estimation framework.

$+\frac{\sigma_W^2}{\sigma_{X_s}^2}\mathbf{I}\Big)^{-1}\hat{H}_{LS}$, where $\mathbf{I}$ is an identity matrix and $R_{H_sH_s} = \mathbb{E}[H_sH_s^H]$ is the channel auto-correlation matrix, $\sigma_W^2 = \mathbb{E}[W^HW]$ is the noise variance, and $\sigma_{X_s}^2 = \mathbb{E}[X_s^HX_s]$ is the signal power. Under imperfect CSI, we have $R_{H_sH_s} = \mathbb{E}[\hat{H}_{LS}\hat{H}_{LS}^H]$, where $\hat{H}_{LS} = \frac{Y_s}{X_s}$ is the LS channel estimate. We deploy pilots at equal intervals, known at both transmitter and receiver. For non-pilot positions, CSI is interpolated by transforming the estimated frequency domain values into time domain, and setting the samples out of cyclic prefix to null.

Further, the effect of pilot contamination can be mitigated using known techniques such as, sparse pilot design, pilot clustering, and compressed sensing [4]. The coefficients mined by the statistical module for N subcarriers at time t are converted into time domain, $\hat{h}_t \in \mathbb{C}^{K\times N}$, before feeding to HNN. Each element of both the input and output vectors is a complex vector of dimension $\mathbb{C}^{K\times 1}$, owing to spatial randomness, defining its density function. $\hat{h}_t$ is transformed into a dataset of supervised learning $\hat{h}_{K\times N(L+M)}$, and divided into a training dataset with L lag samples, $x := [\hat{h}_{t-NL+1}, \cdots, \hat{h}_{t-1}, \hat{h}_t]^T \in \mathbb{C}^{K\times NL}$ and a validation set with M samples $y := [\hat{h}_{t+1}, \hat{h}_{t+2}, \cdots, \hat{h}_{t+NM}]^T \in \mathbb{C}^{K\times NM}$ for N subcarriers, where $\mathbb{C}$ is the set of complex numbers.

CNN is employed in spatial channel estimation because of its ability in harnessing the spatial correlations in the data. CNN takes the real and imaginary parts separately as inputs into different channels. Each channel is processed using CNN, and these outputs are combined to produce the complex-valued results. A layer j can be found from layer $(j-1)$ as

$$\begin{aligned} x_j &= \rho_j(W_j x_{j-1}) = \rho_j\left(\sum_k \sum_v x_{j-1}(v,k) w_{j,k_j}(u-v,k)\right) \\ &= \rho_j\left(\left[\sum_k x_{j-1}(:,k) * w_{j,k_j}(:,k)\right](u)\right) \end{aligned} \tag{2}$$

where $\rho(\cdot)$ is the non-linear mapping for the output of the convolutional layer achieving contraction in data size as a virtue of spatial correlation. In this work, a rectified linear unit function is employed as a non-linear transform. For J convolutional layers, the output of the final layer is given by $x_{t,o}^J = \rho_J(\cdots(\rho_1(W_1x_0))) = \alpha(W_1x_0)$, where $x_{t,o}^J \in \mathbb{R}^{K\times P}$, and $x_0 = x \in \mathbb{R}^{K\times NL}$ is the input to the CNN. This output is then fed to the recurrent neural network (RNN).

LSTM is employed as a variant of RNN because of its reliability in handling time series data. LSTM performs 3 gated operations as listed in (3), namely, forget gate f_t, input gate i_t, and output gate o_t, thus avoiding gradient vanishing.

$$\begin{aligned} f_t &= \delta\left(\omega_f^T\left[\left(x_t^J\right)^T, z_{t-1}^T\right]^T + b_f\right);\ i_t = \delta\left(\omega_i^T\left[\left(x_t^J\right)^T, z_{t-1}^T\right]^T + b_i\right) \\ c_t^{'} &= \tanh\left(\omega_c^T\left[\left(x_t^J\right)^T, z_{t-1}^T\right]^T + b_c\right);\ c_t = f_t c_{t-1} + i_t c_t^{'} \\ O_t &= \delta\left(\omega_o^T\left[\left(x_t^J\right)^T, z_{t-1}^T\right]^T + b_o\right);\ z_t = O_t \tanh(c_t) \end{aligned} \tag{3}$$

where $\delta(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function, and $\tanh(\cdot)$ is the hyperbolic tangent function. $x_{t,o}^J$ is transformed to $x_t^J \in \mathbb{R}^{KP\times 1}$ and fed as the input of LSTM, whereas $z_t \in \mathbb{R}^{D\times 1}$, and $c_t \in \mathbb{R}^{1\times D}$ denote the output and cell state of the LSTM at time t, respectively. Further, $\omega \in \mathbb{R}^{(KP+D)\times D}$ and $b \in \mathbb{R}^{D\times 1}$ capture the parameters of the LSTM, s.t. $\omega, b \in \theta$, which are optimally determined in the training process.

The input is combined with the previous output of the LSTM, as $[x_t^J, z_{t-1}]$, and $f_t, i_t = [0,1]$ is calculated based on the respective activation functions. This decides the percentage of the previous cell state information to be retained at the current moment, and the current cell state c_t based on the combination of previous cell state c_{t-1} with the candidate cell state $c_t^{'}$. c_t is generated by mixing $c_t^{'}$ and c_{t-1}. Finally, the output z_t of the LSTM is computed based on the weights assigned through $O_t \in [0,1]$ for c_t. The final dimension transformation from $z_t \in \mathbb{R}^{D\times 1}$ to $\tilde{z}_t \in \mathbb{R}^{NM\times 1}$ is achieved by a fully connected layer, $\tilde{z}_t = \omega_{\tilde{z}} z_t + b_{\tilde{z}}$, where $\omega_{\tilde{z}} \in \mathbb{R}^{NM\times D}$ is the feature transformation matrix. The network parameters are optimized based on the loss function $\mathcal{L}\left(\tilde{z}_{t+i}^p, y_{t+i}\right) = \frac{1}{M}\sum_{i=1}^M ||\tilde{z}_{t+i}^p - y_{t+i}||_2^2$.

Proof of convergence: $\mathcal{L}(\cdot)$ is a composition of the Euclidean norm and the square function, both of which are smooth. Therefore, their composition results in a smooth function as well. Further, since Euclidean norm is always non-negative, applying reverse triangle inequality we have: $0 \le \mathcal{L}\left(\tilde{z}_{t+i}^p, y_{t+i}\right) \le \|\tilde{z}_{t+i}^p\|_2^2 + \|y_{t+i}\|_2^2$. Thus, $\mathcal{L}\left(\tilde{z}_{t+i}^p, y_{t+i}\right)$ is bounded on both sides. Since the loss function is convex, smooth, and bounded, it must converge to its global minimum, which guarantees convergence of the S-HNN framework [18].

C. Time Complexity Analysis

The time complexity of a neural network translates to the number of floating point operations (FLOPs) used in generating the output. From (2), we note that the CNN processes the inputs with a constant feature map of size equal to the input, i.e., $K \times NL$. The time complexity is given by $\mathcal{O}(K \times NL)$. From (3), we conclude that the FLOPs depend on the total matrix operation and biases. The matrix-bias couples given by $\{\omega_f, b_f\}$, $\{\omega_i, b_i\}$, $\{\omega_c, b_c\}$, and $\{\omega_0, b_0\}$ offer an equal complexity of order $\mathcal{O}((KP+D)D)$ as $\omega \in \mathbb{R}^{(KP+D)\times D}$, while the feature transformation renders a time complexity resulting from $\omega_{\tilde{z}} \in \mathbb{R}^{NM\times D}$. For a training length l_t and establishing $l_t \ll D$, $KP \ll D$, $NM \ll D$, we evaluate the total time complexity of the LSTM network as $\mathcal{C}_{LSTM} \approx \mathcal{O}\left(l_t D^2\right)$. Further, it was noted during simulation that the training length l_t is $\ge K$, the spatial inputs used in extracting the density function of channel gain.

Thus, time complexity presented by the S-HNN framework is $\mathcal{C}_{S-HNN} \approx \mathcal{O}(l_t(NL + D^2)) = \mathcal{O}\left((1+\eta)l_t D^2\right)$, where $\eta D = \frac{NL}{D}$ is the compression provided by S-HNN.

Time complexity of the two competitive schemes, KDML and MLP, are [16]: $\mathcal{C}_{KDML} = \mathcal{O}(l_{t,KDML} D^2)$ and $\mathcal{C}_{MLP} = \mathcal{O}\left(l_{t,MLP} \sum_{i=1}^{n_l-1} h_{s,i} h_{s,i+1}\right)$. $l_{t,KDML}$, $l_{t,MLP}$ are the training lengths for KDML and MLP schemes, respectively, and $h_{s,i}$ is the size of the i-th layer of MLP. Thus, it can be concluded that learning-based estimators provide a linear complexity. Whereas, MMSE estimator operates under a cubic complexity, $\mathcal{O}\left(l_{t,MMSE}^3\right)$, where $l_{t,MMSE}$ is the dimension of the input matrix. Further comparison between the learning-based counterparts of S-HNN is done in the next section.

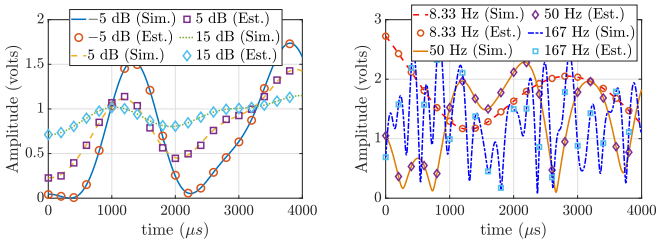
IV. Results and Discussions

This section presents the performance of the proposed S-HNN channel estimator with the parameters in Tables I. The considered OFDM system subcarrier spacing is 15 kHz, sampling rate is 15.36 MHz, IFFT size is 1024, and the numbers of pilot spacing (NPS) are 2, 4, and 8. Communication scenarios under different noise impulsiveness are simulated, and the performance of the proposed framework is compared with MLP [10] and KDML [16] for learning-based estimators. Further, it is compared with the conventional statistical estimators, LS and MMSE+DFT. Channel mobility effect is studied by considering Doppler spread, which can arise due to mobility of the users and/or the surrounding environment.

A. S-HNN based Impulsive Noise Channel Estimation

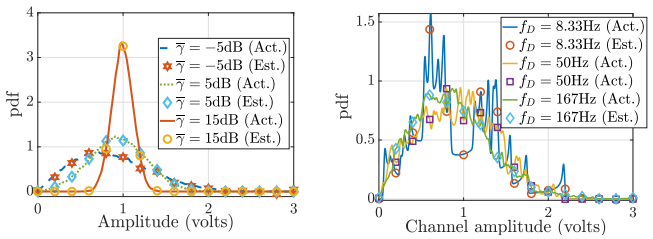
Figs. 3 and 4 present the temporal estimation of the channel with NPS $= 2$. Fig. 3(a) shows the channel estimation for different $\overline{\gamma}$ over a window of 500μs. The actual and S-HNN estimated plots closely match with nRMSE $\approx 10^{-3}$ to 10^{-4}, $\forall\ \overline{\gamma}$. Further, from Fig. 3($b$), we observe that, even for varying mobility environments the proposed framework estimates the temporal variations of the channel with a similar accuracy.

Fig. 4(a) shows the estimation of channel density function at various SNR. The proposed channel estimator is observed to provide a highly accurate channel statistics, with an nRMSE $\approx 10^{-4}$. Fig. 4(b) estimates the density function for different Doppler shifts, which indicate that the estimator works with similar accuracy for varying Doppler shifts as well. Thus, the proposed S-HNN based estimator provides spatio-temporal estimation of the channel and its density function under various conditions with high accuracy.



(a) S-HNN estimate at different SNR (b) S-HNN in mobile environment

Figure 3: (a) Static; (b) at various Doppler frequencies, f_D.

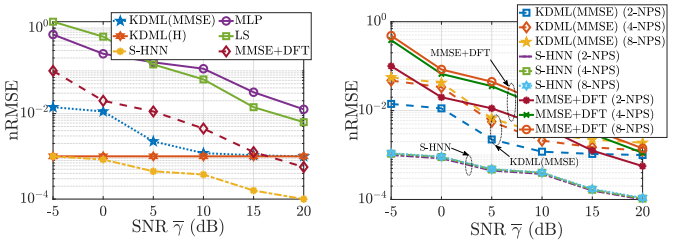


(a) (b)

Figure 4: Channel pdf estimation using S-HNN under (a) different SNR at $f_D = 0$ Hz, and (b) varying Doppler shift, $\overline{\gamma} = 5$ dB.

Table I: S-HNN estimator parameters

Parameters		Values
CNN	Activation function	Sigmoid and tanh
	Hidden layers	4
	Feature maps	64
	Kernel size	3×3
RNN	Activation function	ReLU
	Hidden layers	64
Learning rate		0.001
Batch size		500
Training length		15000
Training epochs		30-40 for $\overline{\gamma}$ = 5-15 dB
Optimum lag samples, L		100



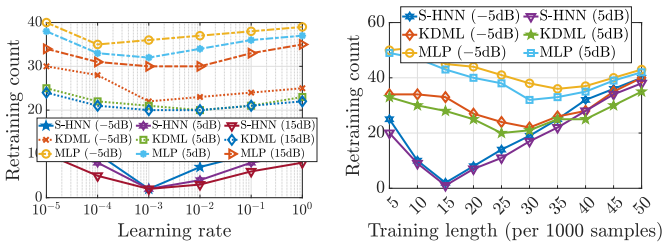
(a) (b)

Figure 5: Performance comparison of S-HNN with state-of-the-art (a) at NPS $= 2$ and (b) varying NPS $= \{2, 4, 8\}$.

Remark 1. *As noted above, under practical consideration of dynamic impulsive noise, while analytical characterization of channel pdf is non-tractable, S-HNN based channel estimation can be effective for efficient channel utilization.*

B. Error Performance Comparison

Fig. 5(a) compares the performance of S-HNN against various existing channel estimation techniques in literature. It is observed that the proposed algorithm outperforms the existing techniques in estimating an impulse noise infested wireless channel. Even at average SNR levels of -5 dB, the error remains below $\approx 10^{-3}$, which decreases as the SNR improves. The effectiveness of the proposed S-HNN over the baseline algorithms results from the advantage of accounting the temporal and spatial correlation in the channel data, thereby increasing the accuracy of channel estimation. From Fig. 5(b) it is inferred that the decrease in pilot density does not affect the performance of the estimator significantly, while other channel estimators digress with NPS varying from 2 to 8, the S-HNN estimator stays highly stable. Thus it can



(a) (b)

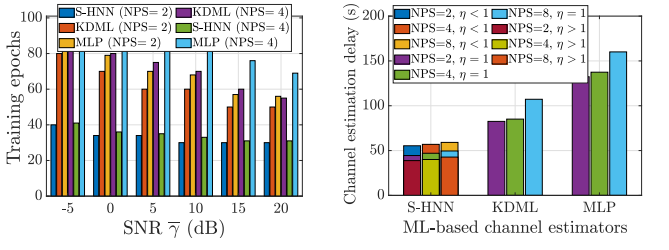
Figure 6: Retraining count of various ML channel estimators with (a) learning rate, and (b) training length; over 24 hours horizon.

be inferred that, the recycling of CNN-LSTM network makes the proposed estimator robust to the varying number of pilots.

C. Complexity Comparison

Fig. 6 shows the retraining instances required in various learning-based estimation schemes over a day's time window. Fig. 6(a) reveals that the number of retrainings required by the proposed S-HNN is 2, which is fairly less than the most competitive KDML framework. This proves the robustness and reliability of the proposed framework. As expected, the retraining instances vary with learning rate and average SNR. However, at the optimal learning rate 0.001, the proposed S-HNN framework presents the same number of retrainings at all $\overline{\gamma}$, unlike the other competitive frameworks. From Fig. 6(b) we observe that the training length at optimality is the least for the proposed S-HNN framework, which manifests to a lesser estimation overhead. Furthermore, the S-HNN framework requires 50% less training data than the KDML approach.

Fig. 7 presents the time complexity for various learning-based channel estimators. From Fig. 7(a) it is noted that the number of training epochs required for S-HNN estimator is less compared to the best learning-based estimator (KDML) in literature, while MLP estimator requires even higher training epochs. Moreover, the training epochs for the proposed S-HNN framework does not vary considerably over a wide range of SNR. As compared to KDML, S-HNN estimator achieves a lower nRMSE within 30-40 training epochs. From Fig. 7(b), we note that the S-HNN estimator requires 48.71% less modeling time, compared to the most competitive KDML channel estimator, which is corresponding to $\eta = 1$, leading to least channel estimation delay. This time saving increases when $\eta > 1$, and decreases marginally with $\eta < 1$. Therefore, for all η, the proposed algorithm presents an improved performance.



(a) (b)

Figure 7: Comparison of (a) training epochs, and (b) channel estimation delay for different learning-based channel estimation schemes.

V. Concluding Remarks

This paper proposed a novel S-HNN framework for estimation in impulsive noise channels, which is not possible with classical methods. The approach aimed at estimating spatio-temporal channel coefficients and gain distribution in OFDM systems. Estimations were done under varying impulse noise amplitudes leading to different average SNR. It was concluded that the CNN-LSTM recycling led to a robust channel estimation. The statistical aid along with HNN resulted in considerable decrease of retraining instances. Furthermore, the use of spatio-temporal correlation reduced the training data length, leading to reduced channel estimation delay.

References

[1] R. Barazideh, *Impulsive noise detection and mitigation in communication systems*. Kansas State University, 2019.

[2] A. K. Mandal, A. Malkhandi, S. De, N. Senroy, and S. Mishra, "A multi-path model for disturbance propagation in electrical power networks," *IEEE Trans. Circuits Syst. II: Express Briefs*, 2022.

[3] T. Bai, H. Zhang, J. Zhang, C. Xu, A. F. Al Rawi, and L. Hanzo, "Impulsive noise mitigation in digital subscriber lines: The state-of-the-art and research opportunities," *IEEE Comm. Mag.*, vol. 57, no. 5, pp. 145–151, 2019.

[4] T. E. Bogale, L. B. Le, X. Wang, and L. Vandendorpe, "Pilot contamination mitigation for wideband massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7889–7906, 2019.

[5] A. K. Mandal and S. De, "Analysis of wireless communication over electromagnetic impulse noise channel," *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2022.

[6] M. Ju, L. Xu, L. Jin, and D. Defeng Huang, "Data aided channel estimation for massive MIMO with pilot contamination," in *Proc. IEEE Int. Conf. Commun.*, 2017, pp. 1–6.

[7] F. A. Pereira de Figueiredo, D. Mendes Lemes, C. Ferreira Dias, and G. Fraidenraich, "Massive MIMO channel estimation considering pilot contamination and spatially correlated channels," *Electron. Lett.*, vol. 56, no. 8, pp. 410–413, 2020.

[8] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, pp. 3111–3124, 2014.

[9] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, "Channel state information prediction for 5G wireless communications: A deep learning approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 227–236, 2020.

[10] M. Soltani, V. Pourahmadi, and H. Sheikhzadeh, "Pilot pattern design for deep learning-based channel estimation in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 12, pp. 2173–2176, 2020.

[11] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 989–1000, 2019.

[12] S. Sadrizadeh, H. Otroshi-Shahreza, and F. Marvasti, "Impulsive noise removal via a blind CNN enhanced by an iterative post-processing," *Signal Process.*, vol. 192, p. 108378, 2022.

[13] Y. Jin, J. Zhang, B. Ai, and X. Zhang, "Channel estimation for mmWave massive MIMO with convolutional blind denoising network," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 95–98, 2020.

[14] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, "Toward convolutional blind denoising of real photographs," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, 2019, pp. 1712–1722.

[15] M. T. Islam, S. M. Rahman, M. O. Ahmad, and M. Swamy, "Mixed Gaussian-impulse noise reduction from images using convolutional neural network," *Sig. Proc.: Image Commun.*, vol. 68, pp. 26–41, 2018.

[16] D. Li, Y. Xu, M. Zhao, J. Zhu, and S. Zhang, "Knowledge-driven machine learning and applications in wireless communications," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 2, pp. 454–467, 2022.

[17] R. Gupta, V. Gupta, A. K. Mandal, and S. De, "Learning-based multivariate real-time data pruning for smart PMU communication," in *Proc. IEEE Consum. Commun. Netw. Conf.*, 2022, pp. 326–331.

[18] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.