RESEARCH ARTICLE

Data Aided Channel Estimation for MIMO-OFDM Wireless Systems Using Reliable Carriers

INAAMULLAH KHAN, MICHAEL CHEFFENA, AND MOHAMMAD MAHMUDUL HASAN

Faculty of Engineering, Norwegian University of Science and Technology (NTNU), 2815 Gjøvik, Norway

Corresponding author: Inaamullah Khan (Inaamullah.khan@ntnu.no)

ABSTRACT Wireless channel estimation is one of the challenging problems in multiple input multiple output orthogonal frequency division multiplexing (MIMO-OFDM) wireless systems. The MIMO-OFDM exploits the spatial resources and increases the reliability and capacity of wireless systems. However, the performance of these systems depends on accurate channel estimation since the receivers require perfect channel state information (CSI) for coherent signal detection. Thus, wireless channel estimation is a necessary component of OFDM systems. In this article, we have proposed an algorithm for MIMO-OFDM systems that combines pilot symbols with reliable data symbols for channel estimation. The reliable data symbols serve as virtual pilots and enhance the spectral efficiency of the system. The proposed data aided channel estimation (DACE) algorithm eliminates the requirement of any additional resources such as excessive number of training sequences to attain the desired performance. Also, it outperforms the traditional least square (LS) and linear minimum mean square error (LMMSE) methods for channel estimation in terms of mean square error (MSE) and bit error rate (BER) performance of the system.

INDEX TERMS Channel estimation, DACE, MIMO-OFDM.

I. INTRODUCTION

Wireless communication is the most effective zone of technological growth where rapid developments have been made in recent times to improve the performance of wireless systems. Different generations of mobile wireless technologies were introduced in past few decades as wireless technology enabled several services including voice to data and to multimedia transmission. Now, we are aiming towards more advancements with vast range of applications in future wireless systems [1], [2], [3], [4], [5]. In this regard, MIMO-OFDM systems meet the requirements of greater bandwidth and capacity, resist noise and multipath effects, and offer high data transmission rate. In wireless communication, CSI is one of the fundamental concepts that provides the possibility to adjust communication according to the current channel condition, which is essential to achieve reliable communication in MIMO-OFDM systems. However, it is highly complicated to obtain accurate CSI due to the time varying nature of the wireless channel [6], [7], [8].

In recent times, different modulation techniques have been devised, but OFDM presents its own advantages. OFDM is a multicarrier modulation (MCM) technique where the whole transmission bandwidth is divided into N number of orthogonal subcarriers and the parallel data streams are multiplexed to offer high data rates. It converts the frequency selective channel into the frequency flat channels and avoids inter symbol interference (ISI) using cyclic prefix. Furthermore, it is easier to implement OFDM as compared to some other techniques like filter bank multicarrier (FBMC) that does not need cyclic prefix but has greater computational complexity [9], [10], [11], [12], [13], [14].

Channel estimation is one of the demanding challenges in OFDM wireless systems [15], [16]. Different channel estimation methods have been discussed in the literature including LS, minimum mean square error (MMSE), LMMSE, singular value decomposition (SVD)-LMMSE etc. but LS and LMMSE are the most common estimation methods because of their simplicity and better performance. Therefore, several

The associate editor coordinating the review of this manuscript and approving it for publication was Usama Mir.

channel estimation techniques have been developed around LS and LMMSE estimators. The LS estimator is modest in comparison with LMMSE because it does not require a priori knowledge of the channel statistics. Also, it offers comparable performance at high signal-to-noise ratio (SNR). Whereas the LMMSE estimator uses second-order channel statistics to minimize the MSE. It is computationally more complex because the multiplication and inversion of higher dimensional matrices are required for each estimate. However, it is robust and offers better performance against fast fading wireless channel. Generally, increasing the number of training sequences (pilots) improve the performance for both LS and LMMSE estimators, but it reduces the spectral efficiency of the wireless systems [17], [18], [19], [20], [21], [22], [23].

One of the important questions that has been addressed in this research is how to improve the MSE and BER performance of MIMO-OFDM systems without increasing the number of pilots? In this regard, different channel estimation techniques have been considered and it is realized that DACE techniques offer better performance using reliable data symbols. The reliable data symbols are those symbols which are decoded correctly by the receiver and can be used as virtual pilots. These virtual pilots are merged with original pilots to update the solutions for conventional channel estimation methods. This improves the quality of channel estimation without compromising the spectral efficiency of the system [24]. However, another key challenge is how to develop an algorithm to select the reliable data symbols out of all possible data symbols transmitted through the wireless channel?

This problem has captured a lot of attention from research community, but only few techniques/algorithms have been proposed to deal with it. One of the most widely used method to select reliable symbols is maximum likelihood (ML). Although it is optimal at high SNR in additive white Gaussian noise (AWGN) scenario, it is not preferable when received symbols are equidistant from a certain constellation point. Also, ML performance deteriorates because it ignores the channel estimation error [25]. In this work, the reliable data symbols are selected using a novel algorithm for OFDM systems. The proposed method calculates the reliability of each observation independently, and judiciously selects the most reliable symbols when they are equidistant from a certain constellation point. It considers both noise and channel estimation error for the detection. Furthermore, the proposed algorithm offers promising performance for both single input single output (SISO) and MIMO-OFDM systems and achieves high accuracy over the existing methods for channel estimation.

Several channel estimation techniques have been proposed and investigated, but each of them has its own limitations. The ultimate goal is to accurately estimate the wireless channel, in order to compensate for its detrimental effects on transmitted signals, and for perfect signal demodulation. The existing channel estimation techniques are discussed below and shown in Fig. 1.

A. BLIND CHANNEL ESTIMATION (BCE)

The statistical characteristics of the received signal are exploited in BCE techniques which require a large amount of data. These techniques are investigated to overcome the wastage of bandwidth in pilot assisted channel estimation (PACE) techniques. The BCE techniques are separated in statistical and deterministic methods. Using deterministic methods, both the received signals and the channel coefficients are considered as the deterministic quantities. The comparison of deterministic and statistical methods indicate that the deterministic methods converge faster than the statistical methods. However, the deterministic methods have greater computational complexity that increases further as the constellation order increases. On the other hand, the performance of statistical methods deteriorate when dealing with exceptionally short sampling sequence. As, no training sequences are required for BCE techniques, thus seem more attractive with effective bandwidth utilization [26], [27]. However, these techniques have greater computational complexity and low convergence rate that cause performance degradation. Also, BCE techniques require time invariant channels for better performance, so these methods are restricted to slow fading channels and not favorable for fast fading channels.

B. NON-BLIND CHANNEL ESTIMATION (NBCE)

In NBCE techniques, information of the previous channel estimates or some portion of the transmitted signal is available to the receiver to be used for channel estimation. These techniques can be divided into the following two types.

1) PILOT ASSISTED CHANNEL ESTIMATION (PACE)

In PACE, also known as training based channel estimation technique, data known to the receiver is multiplexed with the transmitted data symbols at the pre-determined location before transmission. Using optimized space between pilot and data symbols, the PACE techniques offer better performance in terms of the spectral efficiency of the system. However, regardless of several improvements that have been made, the major drawback associated with PACE techniques is the wastage of the bandwidth due to the requirement of large number of pilots for channel estimation. Another setback is that the channel estimates depend only on pilot tones, therefore interpolation techniques are used to estimate the data tones which could not be a perfect estimator all the time [28], [29], [30], [31], [32].

2) DECISION DIRECTED CHANNEL ESTIMATION (DDCE)

The DDCE technique exploits the information of non-pilot symbols to improve the performance of wireless systems. It offers higher reliability than PACE techniques. However, error propagation can occur in successive decisions of DDCE because it utilizes previous channel estimates to decode current OFDM symbols. The newly estimated symbol information is then used to estimate the channel corresponding to the current OFDM symbols. Generally, the estimator of
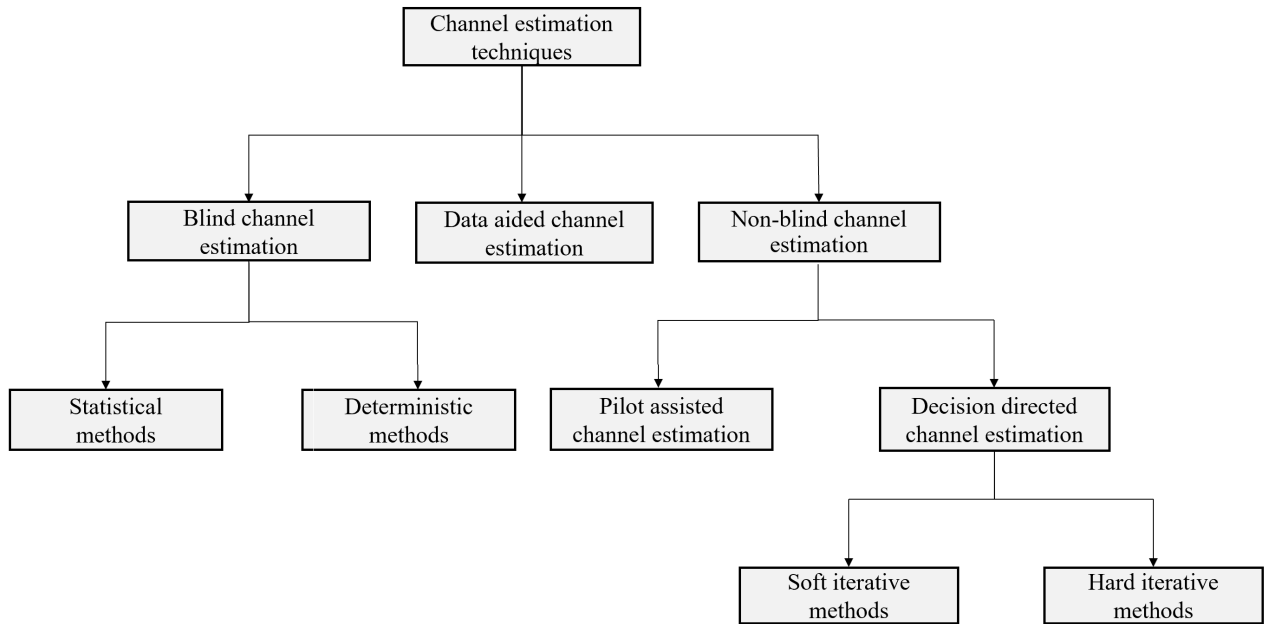
FIGURE 1. Channel estimation techniques.

the channel utilizes soft symbol information or hard symbol information in the iterative form of DDCE techniques. This information is filtered iteratively and estimated by the detector. It is then sent back to the channel estimator to improve the performance of channel estimates with the increasing number of iterations. In this way, the detector takes advantage from the better channel estimates and produces updated symbol information. Hence,the detector works in an iterative manner with the channel estimator. The channel estimator uses hard decision output of the detector in hard iterative channel estimation methods. In contrast, if soft symbol information is employed, the log-likelihood ratios (LLR) are used by the channel estimator. These ratios are estimated by the detector for the channel estimation. Such techniques are known as soft iterative channel estimation methods [33], [34], [35], [36], [37].

C. DATA AIDED CHANNEL ESTIMATION (DACE)
The DACE technique improves the quality of channel estimation using both pilot and decoded data symbols. This is in sheer contrast to the PACE techniques which rely only on pilot symbols. That is why, DACE dominates the PACE techniques, reduces the pilot overhead, and enhances the spectral efficiency of the wireless systems [38], [39], [40], [41], [42], [43]. Different channel estimation techniques have been analyzed with the perspective of the selection of reliable data symbols to improve the MSE and BER performance of MIMO-OFDM systems and it turns out that DACE techniques can give better performance in this regard. The MIMO-OFDM increases the range, offers superior performance using multipath propagation, and improves the data rate combining multiple data streams. Thus, it provides both

the diversity gain as well as the multiplexing gain. The diversity gain increases the reliability of the system as mutiple received signals from different channels are combined to estimate the transmitted signal. Whereas the multiplexing gain increases the capacity as different signals are sent through different paths for this purpose [44], [45], [46], [47], [48], [49].

D. PERFORMANCE MEASURE FOR DACE
The performance measure for DACE depends upon various factors including pilot patterns, estimation methods, and detection methods.

1) PILOT PATTERNS
The estimation of pilot spacing requires extreme care because greater number of pilots improves the quality of channel estimation but reduces the spectral efficiency of the system and vice versa. An optimum pilot pattern for one channel could not be optimum for some other channel because of different fading processes. Another significant feature is the power allocation to the pilot tones in comparison with the data symbols. In numerous cases, equal power is assigned to both pilot and data symbols. However, the accuracy of channel estimation can be enhanced by assigning more power to the pilot tones, but it reduces the SNR.

Pilots can be used in different patterns including block type and comb type pilot patterns as shown in Fig. 2. The block type pilot configuration has pilot tones for all subcarriers in first OFDM block and it is preferable for slow fading channels. Whereas comb type pilot configuration has uniform distribution of pilot tones in all OFDM blocks and it is suitable for fast fading channels [30], [50].
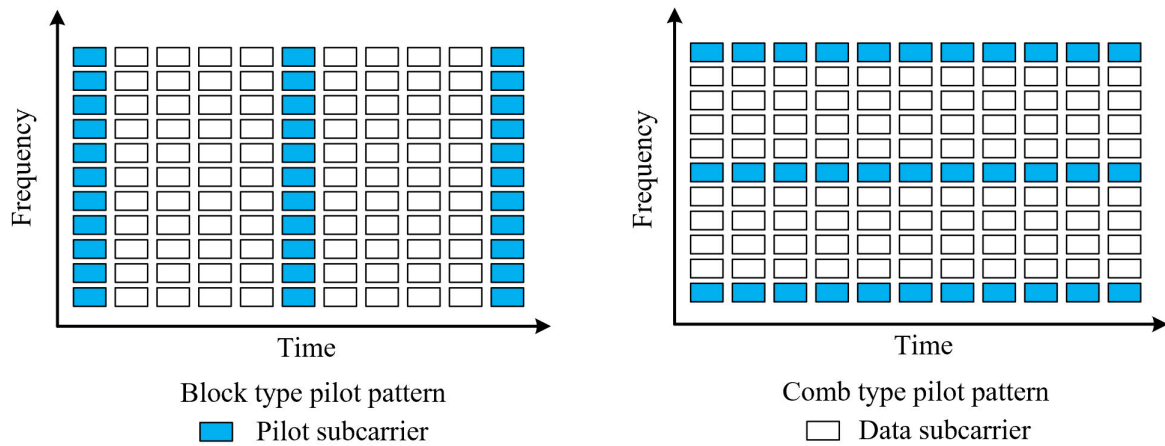
FIGURE 2. Pilot patterns.

2) ESTIMATION METHODS

There are many channel estimation methods like LS, MMSE, LMMSE, SVD-LMMSE, etc. but LS and LMMSE are the most widely used methods for channel estimation. Every method has its own advantages and disadvantages. For instance, LS is simple and direct method, and its performance reaches MMSE at high SNR. It does not require channel and noise statistics but ignoring the noise influence completely, it has greater estimation error particularly at low SNR. Also, LS offers no advantages if the channel length or the channel impulse response (CIR) is unknown. Whereas if channel has L number of taps, an improved performance can be obtained because of the noise reduction.

In contrast to the LS, the MMSE considers the noise influence and gives better performance at low SNR. However, its computational complexity is high due to the extra information incorporated in the estimation technique. Similarly, LMMSE is extensively utilized for the channel estimation of OFDM based systems as it is an optimum method to reduce the MSE while AWGN is present. It requires prior knowledge of pilot symbols or channel statistics for the channel estimation. Therefore, it has greater computational complexity that can be reduced by the appropriate insertion of pilot tones across the OFDM subcarriers [51], [52], [53]. Similarly, SVD also reduces the computational complexity of LMMSE. However, SVD estimation involves complex nonlinear optimization methods, and its estimation is prone to the error of estimated channel matrix [54], [55].

3) DETECTION METHODS

The detection of reliable data symbols is an integral part of DACE techniques, so the detection method plays a vital role in this regard. The most common detection method is ML that determines reliable symbols based on minimum distance from a certain constellation point. However, it is not preferable in equidistant scenario for multiple symbols. Also, intelligent ML decoder needs to update its decisions

iteratively in frequency domain based on the resulting waveforms in time domain [56]. So, one faulty decision in frequency domain can generate error propagation on subsequent decisions.

Some other detection methods have also been proposed in the literature including symbol selection algorithm (SSA) in which SNR based reliability coefficient (RC) is associated with each received symbol. The symbols with highest RC values are considered reliable, and the rest are discarded [57]. However, SSA is also not an optimal choice for the accurate detection of reliable symbols. Thus, it is key to introduce such an algorithm that detects the reliable symbols with high precision and improves the quality of channel estimation for MIMO-OFDM systems.

E. MACHINE LEARNING BASED METHODS FOR CHANNEL ESTIMATION

Apart from the above-mentioned basic techniques for channel estimation, the machine learning based channel estimation methods have also been developed in recent years. Different from the existing OFDM receivers that first estimate the CSI explicitly and then the transmitted symbols are detected/recovered, using deep learning approach for channel estimation, the CSI is estimated implicitly and the transmitted symbols are recovered directly. The use of deep neural networks (DNN) for channel estimation improves the performance because the various characteristics of wireless channel are learnt and analyzed by DNN. The DNN are basically deeper versions of artificial neural networks (ANN) with greater number of hidden layers. In this regard, the deep learning approach for channel estimation is more effective when the wireless channel has greater interference and distortion [58], [59]. However, it offers various challenges in terms of availability of datasets and training because less datasets have been used in wireless communication. Also, privacy and data security, fluctuations in wireless propagation environment and interference are the various constraints to obtain these datasets. The training of these datasets is

TABLE 1. Channel estimation for MIMO wireless systems.

Method/Modulation	Outcome	Research Gaps	Ref.
MIMO-FBMC	Outperforms LS	Didn't check LMMSE High computational complexity	[12]
MIMO-QPSK	Improved performance than LS and MMSE	Results achieved for limited SNR range (0 dB to 14 dB) Only 2×4 MIMO included	[22]
MIMO-OFDM	Better performance than LS	Didn't check LMMSE High computational complexity	[24]
MIMO-QPSK	Better performance than LMMSE	Didn't check LS Didn't achieve results at low SNR values	[38]
MIMO-OFDM	Outperforms MMSE	Small residual interference required Didn't achieve results at high SNR values	[42]

another crucial problem in deep learning approach because the time complexity becomes too high when all layers are trained simultaneously, and the deviation can transfer from one layer to the other if one layer at a time is trained. Furthermore, the machine learning techniques considered only simplified system models whereas radio environments with high complexity have not been investigated with comprehensive considerations. So, it is paramount to investigate all the aforesaid challenges associated with deep learning approach for wireless channel estimation.

F. RESEARCH PROBLEMS
Looking for different channel estimation techniques, the recent trends indicate that the performance improvement for channel estimation in MIMO-OFDM systems costs us in many ways. Some of the important research gaps which have been addressed in this work are shown in Table. 1. The indicated shortcomings are considered as benchmark to improve the performance of channel estimation. Knowing that the data aided techniques can give better performance and higher spectral efficiency, we propose a novel algorithm for the channel estimation of MIMO-OFDM systems. The proposed algorithm outperforms the existing techniques and implies high quality channel prediction.

II. MIMO-OFDM SYSTEM MODEL
The MIMO-OFDM system model is presented in this section. It offers robust and efficient data communication as demodulation of subcarriers can be performed individually. The inverse fast Fourier transform/fast Fourier transform (IFFT/FFT) implementation of MIMO-OFDM system with DACE is shown in Fig. 3.

The baseband signal for OFDM system is given as:

$$x(t) = \sum_{k=0}^{N-1} X(k)e^{j2\pi k \Delta ft} \tag{1}$$

where N represents the number of subcarriers, $X(k)$ is complex modulation symbol transmitted on the kth subcarrier, and Δf is the subcarrier spacing.

The time domain samples at the output of the IFFT are given by:

$$x(n) = W^H X(k) \tag{2}$$

where W is FFT unitary matrix of order $N \times N$ and its elements are given as:

$$W = e^{-j2\pi kn/N}, 0 \le k, n \le N - 1 \tag{3}$$

The cyclic prefix of length g is added to avoid ISI such as:

$$x'(n) = C_T \times x(n) \tag{4}$$

where C_T is $(N + G) \times N$ order matrix and given as:

$$C_T = \begin{bmatrix} O_{G\times(N-G)} & I_G \\ I_N \end{bmatrix} \tag{5}$$

The frequency selective channel is modelled using L-tap finite impulse response (FIR) filter with taps $[h_0, h_1, \ldots, h_{L-1}]$.

The time domain received signal is given by:

$$y(n) = \sum_{i=0}^{L-1} h_i x'(n-i) + z(n), \quad n = 0, 1, \ldots, N-1 \tag{6}$$

where $z(n) \sim N(0, \sigma_z^2)$ is AWGN.

The convolution is converted into matrix multiplication using following Toeplitz like channel matrix $\widetilde{H}$.

$$\widetilde{H} = \begin{bmatrix} h_0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ h_1 & h_0 & 0 & 0 & \cdots & \vdots & \vdots & 0 \\ \vdots & h_1 & h_0 & 0 & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & \vdots \\ h_{L-1} & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & h_{L-1} & \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & h_1 & h_0 & 0 \\ 0 & \cdots & \cdots & h_{L-1} & \cdots & h_2 & h_1 & h_0 \end{bmatrix} \tag{7}$$

So,

$$y(n) = \widetilde{H}x'(n) + z(n) = \widetilde{H}C_T W^H X(k) + z(n) \tag{8}$$

At the receiver, the removal of the cyclic prefix takes place that can be expressed as multiplication of the received signal with matrix C_R given as:

$$C_R = \begin{bmatrix} O_{N\times G} & I_N \end{bmatrix} \tag{9}$$
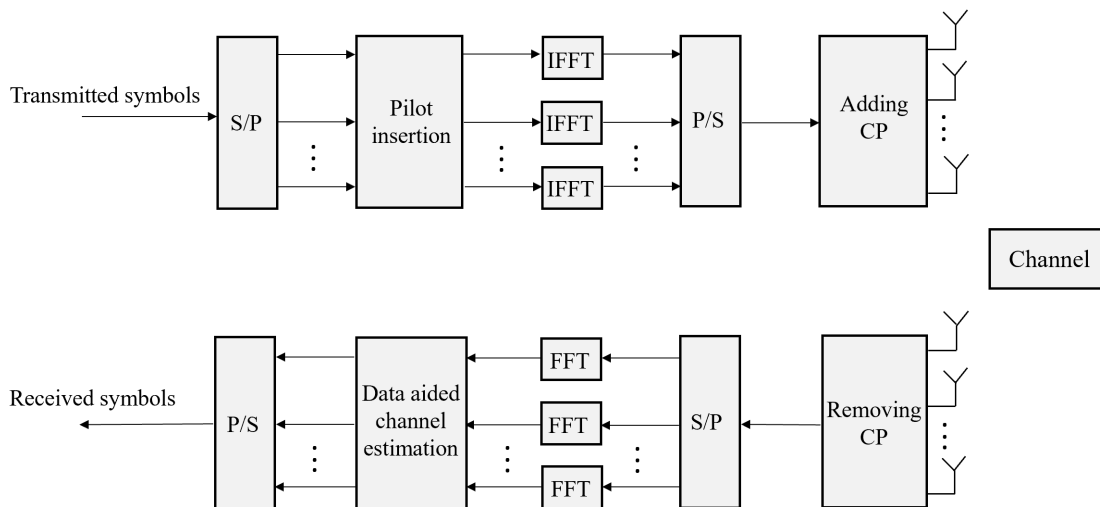
FIGURE 3. MIMO-OFDM system model.

The multiplication converts the Toeplitz like channel matrix into the circulant matrix.

Circulant matrix for 4-taps is given as:

$$C_R \widetilde{H} C_T = \begin{bmatrix} h_0 & 0 & h_3 & h_2 & h_1 \\ h_1 & h_0 & 0 & h_3 & h_2 \\ h_2 & h_1 & h_0 & 0 & h_3 \\ h_3 & h_2 & h_1 & h_0 & 0 \\ 0 & h_3 & h_2 & h_1 & h_0 \end{bmatrix} \quad (10)$$

Now, the equation for received signal can be written as:

$$Y(k) = W C_R \widetilde{H} C_T W^H X(k) + Z(k) \quad (11)$$

where $Z(k) = W C_R z(n)$ and $H = W C_R \widetilde{H} C_T W^H$

$$\Rightarrow H = \begin{bmatrix} H(0) & 0 & \cdots & \cdots & 0 \\ 0 & H(1) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & H(N-2) & 0 \\ 0 & \cdots & \cdots & 0 & H(N-1) \end{bmatrix} \quad (12)$$

Thus, the frequency domain received signal for OFDM system is given as:

$$Y = HX + Z \quad (13)$$

where H represents the channel frequency response (CFR), X contains the transmitted symbols and Z is AWGN with zero mean and covariance matrix $R_Z = \sigma_Z^2 I_N$ [60].

The matrix vector form of the received signal model is given as:

$$Y = Ah + Z \quad (14)$$

where $A = \sqrt{N} diag(X) F$ and F represents the partial FFT matrix comprising of first L columns of full FFT matrix.

Also, the CFR matrix is given as:

$$H = \sqrt{N} Fh \quad (15)$$

Using N_t number of transmit antennas and N_r number of receive antennas for MIMO-OFDM systems, the intention is to estimate h given the received symbol Y with minimum number of pilots.

III. PROPOSED METHOD

Herein, an algorithm is devised to judicially pick the reliable data symbols to overcome the inadequacies for DACE techniques.

Starting from equation (13) over the kth subcarrier as:

$$Y(k) = H(k)X(k) + Z(k) \quad (16)$$

Finding $\hat{H}(k)$ using LS/LMMSE estimation methods, the estimate of $X(k)$ is given by:

$$\hat{X}(k) = \frac{Y(k)}{\hat{H}(k)} = X(k) + D(k) \quad (17)$$

where $D(k)$ indicates the distortion and it is defined as:

$$D(k) = \hat{X}(k) - \langle \hat{X}(k) \rangle \quad (18)$$

Now, the reliability values are determined by the following expression.

$$R(k) = \log \frac{p_{D(k)}(\hat{X}(k) - \langle \hat{X}(k) \rangle)}{\sum_{i=0, X_i(k) \neq \langle \hat{X}(k) \rangle}^{M-1} p_{D(k)}\left(\hat{X}(k) - X_i(k)\right)} \quad (19)$$

It is a likelihood ratio dependent upon relative posterior probability of $D(k)$ equating $\hat{X}(k) - \langle \hat{X}(k) \rangle$ to the probability of some other vector $\hat{X}(k) - X_i(k)$. Intuitively, $\hat{X}(k)$ is decoded relative to both the nearest constellation point as well as the nearest neighbours. The $p_{D(k)}$ is the probability density function (PDF) of the distortion $D(k)$. As it is highly complex to determine the exact PDF of the distortion, so we consider it circularly symmetric Gaussian with mean zero and variance

$\sigma^2_{D(k)}$. Consequently, the following PDF expression is used to evaluate reliability values in (19).

$$p_{D(k)}(X) = \frac{1}{\pi \sigma^2_{D(k)}} \exp\left(-\frac{1}{\sigma^2_{D(k)}}|X|^2\right) \quad (20)$$

$$p_{D(k)}(\hat{X}(k) - \langle \hat{X}(k) \rangle) = \frac{1}{\pi \sigma^2_{D(k)}} \exp\left(-\frac{1}{\sigma^2_{D(k)}}|\hat{X}(k) - \langle \hat{X}(k) \rangle|^2\right) \quad (21)$$

where the variance of distortion is computed using (16) in (17) as follows:

$$\hat{X}(k) = \frac{H(k)X(k) + Z(k)}{\hat{H}(k)} \quad (22)$$

$$\hat{X}(k) \approx X(k) + \frac{Z(k)}{\hat{H}(k)} \quad (23)$$

Comparing (17) with (23), we get $D(k) = Z(k)/\hat{H}(k)$. Now, the variance of distortion is given as:

$$\sigma^2_{D(k)} = E\left\{|D(k)|^2\right\} = E\left\{\left|\frac{Z(k)}{\hat{H}(k)}\right|^2\right\}$$
$$= (\hat{H}(k))^{-H}(\hat{H}(k))^{-1}\sigma^2_Z \quad (24)$$

Hence, the distortion depends on both AWGN and channel estimates. Now, the important question is how to pick the reliable data carriers from a set of all possible 2^N data carriers? The reliability value for each data carrier is obtained using equation (19), which results in a non-negative number. Therefore, a threshold value is chosen and all data carriers having reliability values greater than the threshold value are considered as reliable data carriers. The threshold value is estimated empirically and a set of reliable data tones Ω_m is selected out of all available data tones. In this regard, the proposed reliability function associates a reliability estimate with each tone and determines the m tones having highest reliability to generate Ω_m. By incorporating the reliable data carriers with pilots, the wireless channel is re-estimated using LS and LMMSE methods.

Consider P as pilot and R as reliable indices, equation (14) is updated as:

$$Y_{rp} = A_{rp}h + Z_{rp} \quad (25)$$

where

$$A_{rp} = diag(X_{rp})F_{rp} \quad (26)$$

and

$$X_{rp} = X(P \cap R) = \begin{bmatrix} X(P) \\ X(R) \end{bmatrix} \quad (27)$$

$$F_{rp} = F(P \cap R) = \begin{bmatrix} F(P) \\ F(R) \end{bmatrix} \quad (28)$$

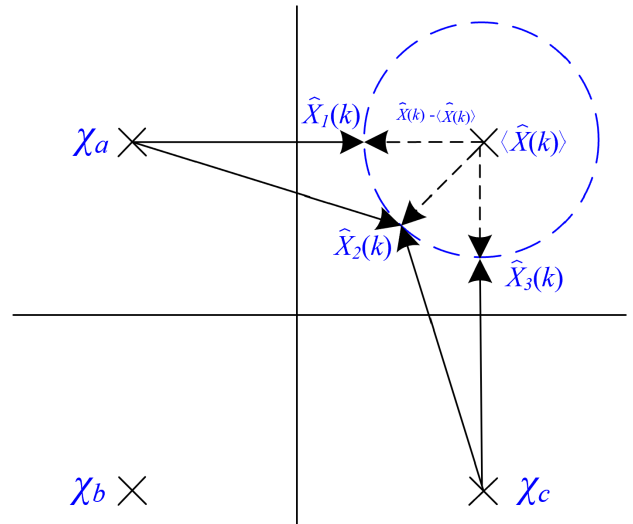$$Y_{rp} = Y(P \cap R) = \begin{bmatrix} Y(P) \\ Y(R) \end{bmatrix} \quad (29)$$

FIGURE 4. The reliability variations of observation.

Finally, LS and LMMSE channel estimates for equation (25) are given by:

$$\hat{h}^{(LS)_{rp}} = \left(A^H_{rp}A_{rp}\right)^{-1}A^H_{rp}Y_{rp} \quad (30)$$

$$\hat{h}^{(LMMSE)_{rp}} = \left[R^{-1}_h + A^H_{rp}R^{-1}_Z A_{rp}\right]^{-1}A^H_{rp}R^{-1}_Z Y_{rp} \quad (31)$$

Summary of the proposed algorithm

1) Get initial channel estimates $\hat{h}$ by using pilots/training symbols (e.g., using LS, LMMSE or any other method).
2) Obtain $\hat{H}$ by frequency transformation of $\hat{h}$ and then equalize the data using (17) to obtain $\hat{X}$.
3) Compute the variance $\sigma^2_{D(k)}$ and the reliability $R(k)$ for each subcarrier, $k = 1, 2, \ldots, N$, using (24) and (19) respectively.
4) Select the index R of the most reliable data carriers by thresholding i.e., the index vector R is computed as follows:
   for $k = 1, 2, \ldots, N, k \in R$ iff $R(k) > threshold$.
5) Re-estimate the channel using both pilots and reliable data carriers (e.g., using LS, LMMSE or any other method).

The graphical illustration of the reliability variations of observation for the proposed algorithm is shown in Fig. 4. Consider $\hat{X}_1(k)$ and $\hat{X}_2(k)$ such that:

$$\frac{|\hat{X}_1(k) - \langle \hat{X}(k) \rangle|}{|\hat{X}_1(k) - X_a|} > \frac{|\hat{X}_2(k) - \langle \hat{X}(k) \rangle|}{|\hat{X}_2(k) - X_a|} \quad (32)$$

It is obvious from Fig. 4 that X_a and X_c are the nearest neighbours and X_b is the next nearest neighbour for $\langle \hat{X}(k) \rangle$. The received symbols $\hat{X}_1(k)$, $\hat{X}_2(k)$ and $\hat{X}_3(k)$ are equidistant form $\langle \hat{X}(k) \rangle$ and exist at the circumference of the circle of radius r. In this scenario, ML is not an optimum method for detection because the received symbols are equidistant from $\langle \hat{X}(k) \rangle$. Now, consider X_a as the nearest neighbour to compare the reliability of $\hat{X}_1(k)$ and $\hat{X}_2(k)$. It can be observed that
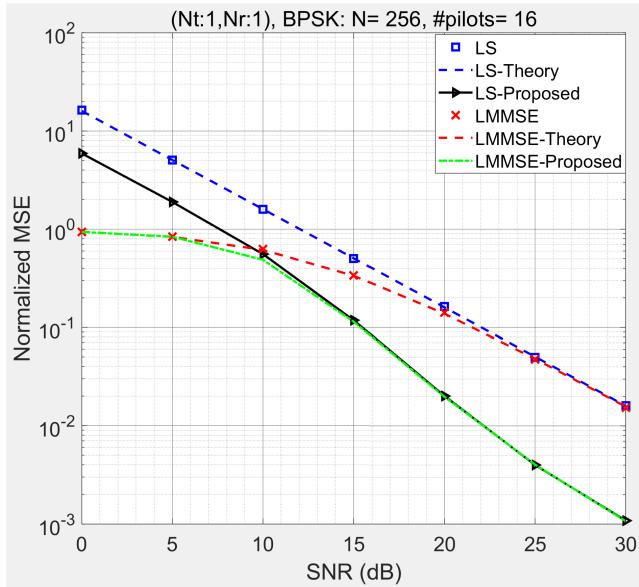
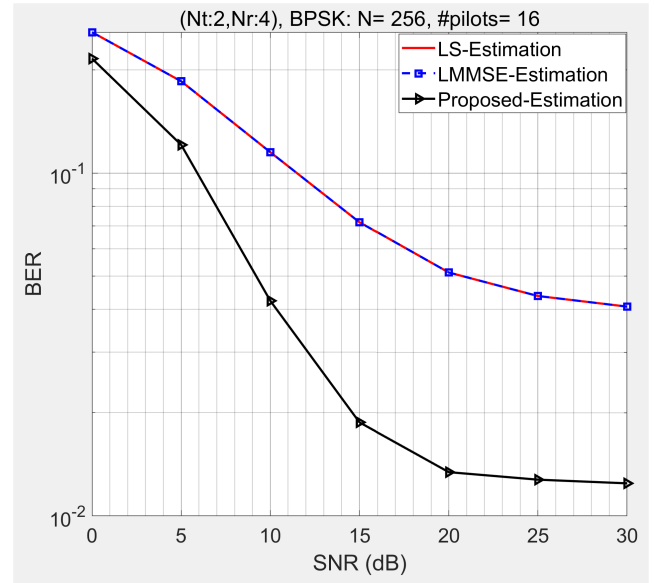**FIGURE 5.** MSE vs. SNR for BPSK, $N_t = 1$, $N_r = 1$.

FIGURE 7. BER vs. SNR for BPSK, $N_t = 2$, $N_r = 4$.
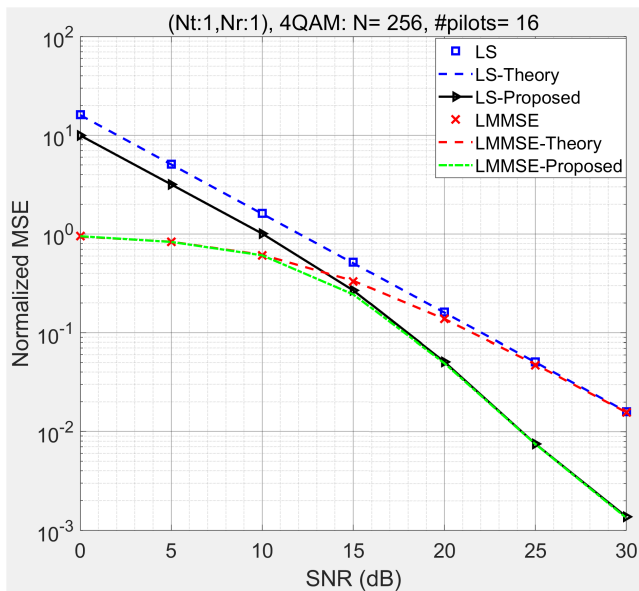

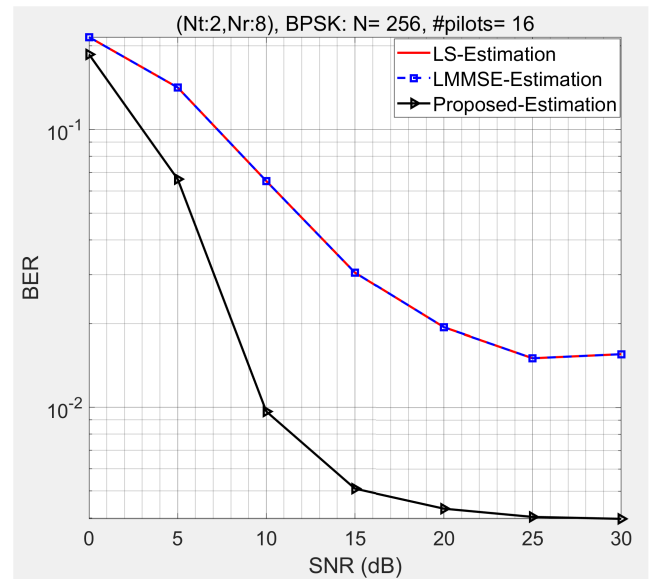**FIGURE 6.** MSE vs. SNR for 4-QAM, $N_t = 1$, $N_r = 1$.

FIGURE 8. BER vs. SNR for BPSK, $N_t = 2$, $N_r = 8$.

$\hat{X}_2(k)$ is farther from X_a than $\hat{X}_1(k)$, so $\hat{X}_2(k)$ is more reliable than $\hat{X}_1(k)$. Alternatively, the reliability to assume $\hat{X}_2(k)$-$\langle \hat{X}(k) \rangle$ is higher than $\hat{X}_1(k)$-$\langle \hat{X}(k) \rangle$. Similarly, if we consider X_c as the nearest neighbour for the comparison of $\hat{X}_2(k)$ and $\hat{X}_3(k)$, the distance between $\hat{X}_2(k)$ and X_c is greater than the distance between $\hat{X}_3(k)$ and X_c. So, the reliability to assume $\hat{X}_2(k)$-$\langle \hat{X}(k) \rangle$ is also higher than $\hat{X}_3(k)$-$\langle \hat{X}(k) \rangle$. Thus, $\hat{X}_2(k)$ is more reliable than $\hat{X}_1(k)$ and $\hat{X}_3(k)$.

IV. RESULTS AND DISCUSSION
In this section, the simulated results are achieved using N=256 number of OFDM subcarriers and $N_p = 16$ number

of equally spaced pilots in comb type pilot configuration. The MSE and BER are considered as the yardstick for the performance of MIMO-OFDM systems. The results are obtained for both LS and LMMSE channel estimation methods. The Rayleigh fading channel is modelled using $L = 16$ taps FIR filter, where each tap has a complex Gaussian distribution with zero mean and unit variance. The performance of proposed DACE algorithm is evaluated for both BPSK and 4-QAM modulation schemes. The MSE and BER curves are plotted over a diverse range of SNR values from 0 dB to 30 dB. The MSE curves for SISO-OFDM system using BPSK and 4-QAM constellations are shown in Figs. 5 and 6, respectively. These curves suggest that the proposed LS and
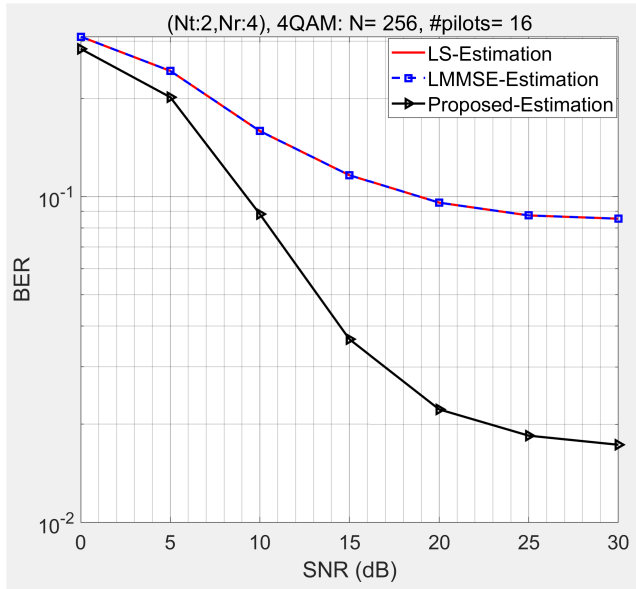
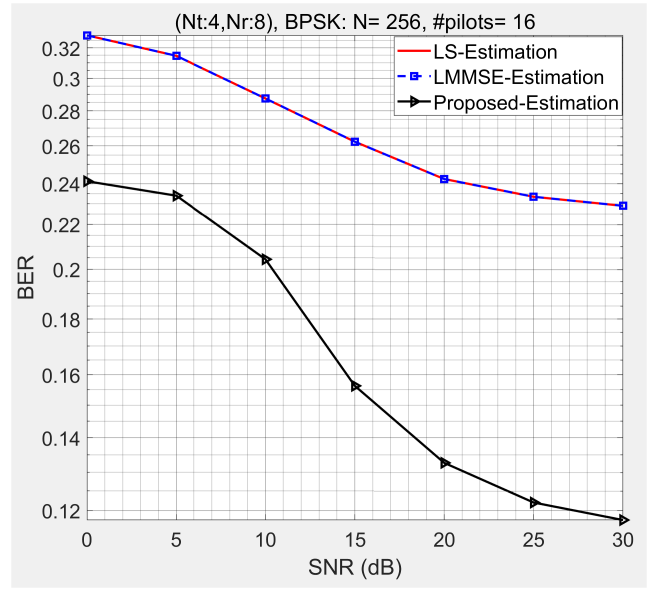**FIGURE 9.** BER vs. SNR for 4-QAM, $N_t = 2, N_r = 4$.

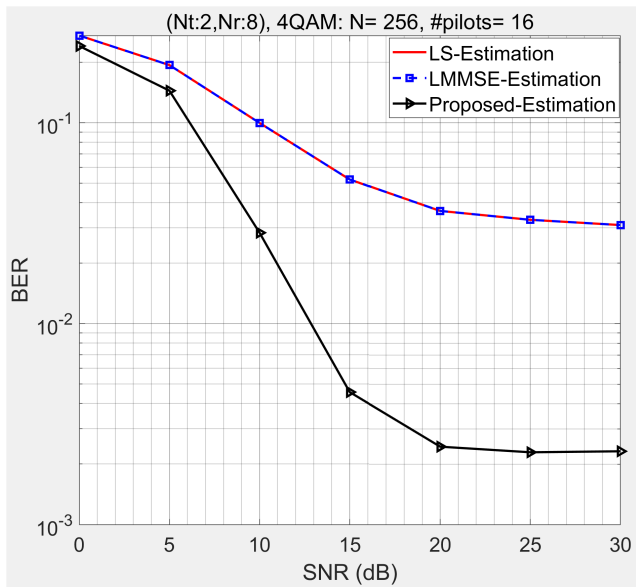**FIGURE 11.** BER vs. SNR for BPSK, $N_t = 4, N_r = 8$.

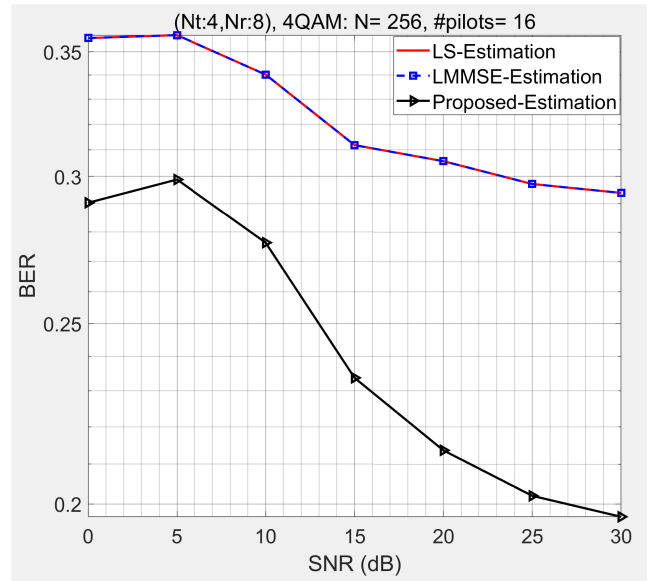**FIGURE 10.** BER vs. SNR for 4-QAM, $N_t = 2, N_r = 8$.

FIGURE 12. BER vs. SNR for 4-QAM, $N_t = 4, N_r = 8$.

LMMSE methods using virtual pilots outperform the conventional LS and LMMSE methods for channel estimation. The LS curves reach LMMSE curves at high SNR and give comparable performance. The performance improvement seems better at high SNR as data equalization/channel estimation has less distortion at high SNR and vice versa.

The BER curves for different MIMO configurations including 2×4, 2×8, and 4×8 are shown in Figs. 7 to 12. These curves indicate that unlike many other methods/algorithms applied for channel estimation, the BER performance for the proposed algorithm improves for different MIMO configurations, from low SNR values to high SNR

values i.e., from 0 dB to 30 dB. It is clear from Figs. 7 and 8 that BPSK attains better BER performance for a greater number of receive antennas due to the diversity gain. Similarly, Figs. 9 and 10 show that the 4-QAM constellation also offers superior performance as the number of receive antennas increases. However, Figs. 11 and 12 imply that the BER performance deteriorates with an increasing number of transmit antennas for BPSK and 4-QAM constellations, respectively. The reason is that it is challenging to pick reliable symbols as the size of the data increases. Furthermore, the MSE and BER performance for BPSK appears slightly better than that of 4-QAM because it is difficult to detect reliable symbols as the constellation order increases.
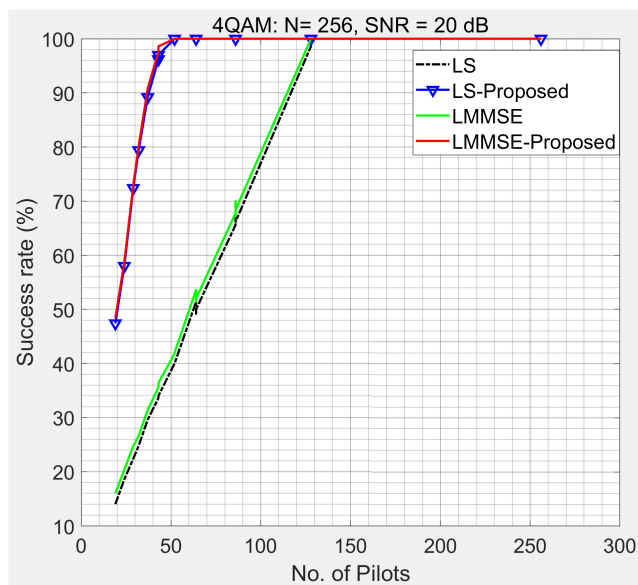
FIGURE 13. Success rate vs. number of pilots.

Finally, it is investigated how much pilot overhead reduction/spectral efficiency improvement has been achieved using the proposed DACE algorithm? To ascertain this, the performance curves are obtained for both LS and LMMSE channel estimation methods at a fixed SNR. The targeted MSE value is chosen as 0.02. The success rate is defined as the ratio of number of symbols for which the MSE is less than the targeted MSE to the total number of transmitted symbols. The SNR value is fixed at 20 dB for 4-QAM modulation scheme. The success rate is plotted against number of pilots varying from L to N as shown in Fig. 13. It is observed that the traditional LS and LMMSE channel estimaton methods require 128 number of pilots to achieve a 100% success rate. However, using the proposed DACE algorithm, the LS and LMMSE methods reach a 100% success rate with only 52 number of pilots. Thus, the proposed method reduces approximately 60% pilot overhead as compared to the conventional methods for channel estimation, which means a significant improvement in terms of the spectral efficiency of the system.

V. CONCLUSION

In this work, we have proposed a data aided channel estimation method for both SISO and MIMO-OFDM systems. A novel algorithm for the selection of reliable data symbols/virtual pilots is developed for optimal channel estimation. The reliable data symbols enable us to attain accurate channel estimates with minimum number of pilots. Alternatively, the reduction of pilot overhead allows us to enhance the spectral efficiency of the system. The simulation results endorse our theoretical analysis and performance comparison against existing methods for channel estimation. The proposed algorithm is simple, precise, and efficient. Also, it

selects the reliable data symbols intelligently and improves the MSE and BER performance of the system.

REFERENCES

[1] J. Jiao, X. Sun, L. Fang, and J. Lyu, "An overview of wireless communication technology using deep learning," *China Commun.*, vol. 18, no. 12, pp. 1–36, Dec. 2021.

[2] S. S. Shree and J. F. G. Poovathy, "Communication technologies in IoT and related concepts: A review," in *Proc. 3rd Int. Conf. Intell. Comput. Instrum. Control Technol. (ICICICT)*, Aug. 2022, pp. 310–314.

[3] S. M. Periannasamy, C. Thangavel, S. Latha, G. V. Reddy, S. Ramani, P. V. Phad, S. R. Chandline, and S. Gopalakrishnan, "Analysis of artificial intelligence enabled intelligent sixth generation (6G) wireless communication networks," in *Proc. IEEE Int. Conf. Data Sci. Inf. Syst. (ICDSIS)*, Jul. 2022, pp. 1–8.

[4] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine learning for resource management in cellular and IoT networks: Potentials, current solutions, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1251–1275, 2nd Quart., 2020.

[5] O. O. Oyerinde and S. H. Mneney, "Review of channel estimation for wireless communication systems," *IETE Tech. Rev.*, vol. 29, no. 4, pp. 282–298, Jul. 2012.

[6] Z. Shang, T. Zhang, G. Hu, Y. Cai, and W. Yang, "Secure transmission for NOMA-based cognitive radio networks with imperfect CSI," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2517–2521, Aug. 2021.

[7] A. Naikoti and A. Chockalingam, "Performance analysis of OTFS with imperfect delay-Doppler channel state information," in *Proc. IEEE 95th Veh. Technol. Conf.*, Jun. 2022, pp. 1–6.

[8] S. Jazir, P. A. Haris, and K. Suresh, "Efficient channel estimation of massive MIMO systems using artificial bee colony algorithm," in *Proc. IEEE Recent Adv. Intell. Comput. Syst. (RAICS)*, Dec. 2020, pp. 190–194.

[9] Y. Xiang, K. Xu, B. Xia, and X. Cheng, "Bayesian joint channel-and-data estimation for quantized OFDM over doubly selective channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1523–1536, Mar. 2023.

[10] T. Fathi, S. Ashraf, and A. A. Rhebi, "Deep learning for channel estimation and signal detection," in *Proc. Int. Conf. Eng. MIS (ICEMIS)*, Jul. 2022, pp. 1–6.

[11] W. Gao, M. Yang, W. Zhang, and L. Liu, "Efficient OFDM channel estimation with RRDBNet," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2022, pp. 1–5.

[12] P. Singh, H. B. Mishra, A. K. Jagannatham, and K. Vasudevan, "Semi-blind, training, and data-aided channel estimation schemes for MIMO-FBMC-OQAM systems," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4668–4682, Sep. 2019.

[13] W. Yuan, S. Li, Z. Wei, J. Yuan, and D. W. K. Ng, "Data-aided channel estimation for OTFS systems with a superimposed pilot and data transmission scheme," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1954–1958, Sep. 2021.

[14] A. Zaib and S. Khattak, "Structure-based low complexity MMSE channel estimator for OFDM wireless systems," *Wireless Pers. Commun.*, vol. 97, no. 4, pp. 5657–5674, Dec. 2017.

[15] Y. Zhao, W. Zhao, G. Wang, B. Ai, H. H. Putra, and B. Juliyanto, "AoA-based channel estimation for massive MIMO OFDM communication systems on high speed rails," *China Commun.*, vol. 17, no. 3, pp. 90–100, Mar. 2020.

[16] Y. Kabalci, "5G mobile communication systems: Fundamentals, challenges, and key technologies," in *Smart Grids and Their Communication Systems*. Cham, Switzerland: Springer, 2018, pp. 329–359.

[17] T. A. Kumar and L. Anjaneyulu, "Channel estimation techniques for multicarrier OFDM 5G wireless communication systems," in *Proc. IEEE 10th Int. Conf. Syst. Eng. Technol. (ICSET)*, Nov. 2020, pp. 98–101.

[18] T. Taoliu and H. Zhang, "Improved channel estimation method jointing channel coding," in *Proc. 3rd Int. Conf. Electron. Inf. Technol. Comput. Eng. (EITCE)*, Oct. 2019, pp. 1624–1628.

[19] M. B. Sutar and V. S. Patil, "LS and MMSE estimation with different fading channels for OFDM system," in *Proc. Int. Conf. Electron., Commun., Aerosp. Technol. (ICECA)*, vol. 1, Apr. 2017, pp. 740–745.

[20] J. A. Sam and A. K. Nair, "Analysis and implementation of channel estimation in OFDM system using pilot symbols," in *Proc. Int. Conf. Control, Instrum., Commun. Comput. Technol. (ICCICCT)*, Dec. 2016, pp. 725–728.

[21] T. Hesketh, R. C. de Lamare, and S. Wales, "Adaptive MMSE channel estimation algorithms for MIMO systems," in *Proc. 18th Eur. Wireless Conf.*, 2012, pp. 1–5.

[22] B. Mthethwa and H. Xu, "Deep learning-based wireless channel estimation for MIMO uncoded space-time labeling diversity," *IEEE Access*, vol. 8, pp. 224608–224620, 2020.

[23] A. Farzamnia, N. W. Hlaing, M. K. Haldar, and J. Rahebi, "Channel estimation for sparse channel OFDM systems using least square and minimum mean square error techniques," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–5.

[24] K. C. Jomon and S. Prasanth, "Joint channel estimation and data detection in MIMO-OFDM using distributed compressive sensing," *Radioelectron. Commun. Syst.*, vol. 60, no. 2, pp. 80–87, Feb. 2017.

[25] E. B. Al-Safadi, T. Y. Al-Naffouri, M. Masood, and A. Ali, "Nonlinear distortion reduction in OFDM from reliable perturbations in data carriers," 2015, *arXiv:1506.09060*.

[26] B. Liu, C. Gong, J. Cheng, Z. Xu, and J. Liu, "Blind and semi-blind channel estimation/equalization for Poisson channels in optical wireless scattering communication systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 5930–5946, Aug. 2022.

[27] M. Liu, H. Wang, Y. Li, and P. Li, "Research on pilot-based channel estimation algorithms," in *Proc. IEEE 5th Int. Conf. Comput. Commun. (ICCC)*, Apr. 2019, pp. 758–761.

[28] M. H. Rahman, M. Shahjalal, M. O. Ali, S. Yoon, and Y. M. Jang, "Deep learning based pilot assisted channel estimation for Rician fading massive MIMO uplink communication system," in *Proc. 12th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Aug. 2021, pp. 470–472.

[29] H. Dogan, O. Sayli, and E. Panayirci, "Pilot assisted channel estimation for asymmetrically clipped optical OFDM over visible light channels," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, Jun. 2016, pp. 1–4.

[30] Q. Yu, W. Liu, S. Feng, Y. Yan, and Z. Shao, "Channel estimation based on diagonal pilot in OFDM systems," in *Proc. IEEE Int. Conf. Inf. Autom.*, Aug. 2015, pp. 245–249.

[31] E. Manasseh, S. Ohno, and M. Nakamoto, "Pilot symbol assisted channel estimation for OFDM-based cognitive radio systems," *EURASIP J. Adv. Signal Process.*, vol. 2013, no. 1, pp. 1–11, Dec. 2013.

[32] N. Ito, D. Shimbo, N. Maeda, H. Mishima, and H. Nakayama, "Pilot-assisted channel estimation for OFDM systems in time-varying multipath channels," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2014, pp. 1–5.

[33] G. Gaspar, A. Ferreira, P. Montezuma, M. Marques da Silva, and R. Dinis, "An iterative decision directed channel estimation for multi-user MIMO systems," in *Proc. Prog. Electromagn. Res. Symp.-Spring (PIERS)*, 2017, pp. 3699–3706.

[34] W. Zhang, H. Li, and B. Li, "Iterative decision-directed channel estimation for MIMO-OFDM system," in *Proc. 2nd IEEE Int. Conf. Comput. Commun. (ICCC)*, Oct. 2016, pp. 1678–1682.

[35] A. Idris, N. I. A. Razak, and M. Sabry, "Performance evaluation of decision-directed (DD) channel estimation using diversity techniques," in *Proc. IEEE 3rd Int. Conf. Syst. Eng. Technol.*, Aug. 2013, pp. 352–355.

[36] J. Ketonen, M. Juntti, J. Ylioinas, and J. R. Cavallaro, "Decision-directed channel estimation implementation for spectral efficiency improvement in mobile MIMO-OFDM," *J. Signal Process. Syst.*, vol. 79, no. 3, pp. 233–245, Jun. 2015.

[37] O. Ogundile, O. Oyerinde, and D. Versfeld, "Decision directed iterative channel estimation and Reed–Solomon decoding over flat fading channels," *IET Commun.*, vol. 9, no. 17, pp. 2077–2084, Nov. 2015.

[38] X. Li, Q. Wang, H. Yang, and X. Ma, "Data-aided MIMO channel estimation by clustering and reinforcement-learning," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2022, pp. 584–589.

[39] B. Liu, C. Gong, J. Cheng, and Z. Xu, "Data-aided channel estimation for Poisson channels with inter-symbol interference," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[40] Y. Zhang, J. Kang, J. Wang, Y. Xiao, and S. Hu, "Data-aided channel estimation for OFDM-IDMA systems," in *Proc. IEEE 23rd Int. Conf. Digit. Signal Process. (DSP)*, Nov. 2018, pp. 1–4.

[41] P. Li, "Data aided channel estimation in massive MIMO systems," in *Proc. Int. Workshop High Mobility Wireless Commun.*, Nov. 2014, p. 3.

[42] S. Park, B. Shim, and J. W. Choi, "Iterative channel estimation using virtual pilot signals for MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3032–3045, Jun. 2015.

[43] V. A. Irtuga and D. V. Sarana, "Frequency domain data-aided channel estimation for OFDM signals," in *Proc. 6th Int. Congr. Ultra Mod. Telecommun. Control Syst. Workshops (ICUMT)*, Oct. 2014, pp. 29–32.

[44] L. Wang, G. Liu, J. Xue, and K.-K. Wong, "Channel prediction using ordinary differential equations for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 2111–2119, Feb. 2023.

[45] S. E. Khamy, N. Korany, and H. Hassan, "Channel estimation techniques for wideband MIMO-OFDM communication systems using complementary codes two-sided sequences," in *Proc. 37th Nat. Radio Sci. Conf. (NRSC)*, Sep. 2020, pp. 74–84.

[46] A. Thomas and A. K. Nair, "Null-subcarrier based channel estimation and mutual interference reduction in MIMO OFDM systems," in *Proc. Int. Conf. Wireless Commun. Signal Process. Netw. (WiSPNET)*, Mar. 2019, pp. 275–278.

[47] T. Yang, R. Zhang, X. Cheng, and L. Yang, "Secure massive MIMO under imperfect CSI: Performance analysis and channel prediction," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 6, pp. 1610–1623, Jun. 2019.

[48] Q. Qin, L. Gui, P. Cheng, and B. Gong, "Time-varying channel estimation for millimeter wave multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9435–9448, Oct. 2018.

[49] X.-F. Kang, Z.-H. Liu, and M. Yao, "Deep learning for joint pilot design and channel estimation in MIMO-OFDM systems," *Sensors*, vol. 22, no. 11, p. 4188, May 2022.

[50] A. S. Ahmed, M. M. Hamdi, M. S. Abood, A. M. Khaleel, M. Fathy, and S. H. Khaleefah, "Channel estimation using LS and MMSE channel estimation techniques for MIMO-OFDM systems," in *Proc. Int. Congr. Hum.-Comput. Interact., Optim. Robot. Appl. (HORA)*, Jun. 2022, pp. 1–6.

[51] A. K. Shrivas, "A comparative analysis of LS and MMSE channel estimation techniques for MIMO-OFDM system," *Int. J. Sci. Res. Develop.*, vol. 1, no. 8, pp. 44–48, Jan. 2013.

[52] S. A. Ghauri, S. Alam, M. F. Sohail, A. Ali, and F. Saleem, "Implementation of OFDM and channel estimation using LS and MMSE estimators," *Int. J. Comput. Electron. Res.*, vol. 2, no. 1, pp. 41–46, 2013.

[53] D. Pandey and K. N. Dewangan, "MSE and SER compare for LS, MMSE and SMMSE for BCE technique in OFDM," *J. Commun. Eng. Syst.*, vol. 3, no. 3, pp. 14–21, Jul. 2014.

[54] R. Tang, X. Zhou, and C. Wang, "Singular value decomposition channel estimation in STBC MIMO-OFDM system," *Appl. Sci.*, vol. 9, no. 15, p. 3067, Jul. 2019.

[55] K. Bangash, I. Khan, J. Lloret, and A. Leon, "A joint approach for low-complexity channel estimation in 5G massive MIMO systems," *Electronics*, vol. 7, no. 10, p. 218, Sep. 2018.

[56] E. B. Al-Safadi and T. Y. Al-Naffouri, "Pilotless recovery of clipped OFDM signals by compressive sensing over reliable data carriers," in *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 580–584.

[57] M. Siti, A. Assalini, E. Dall'Anese, and S. Pupolin, "Low complexity decision-directed channel estimation based on a reliable-symbol selection strategy for OFDM systems," in *Proc. IEEE Int. Conf. Commun. Workshops*, May 2010, pp. 1–5.

[58] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[59] H. A. Le, T. Van Chien, T. H. Nguyen, H. Choo, and V. D. Nguyen, "Machine learning-based 5G-and-beyond channel estimation for MIMO-OFDM communication systems," *Sensors*, vol. 21, no. 14, p. 4861, Jul. 2021.

[60] F. Khan, *LTE for 4G Mobile Broadband: Air Interface Technologies and Performance*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

INAAMULLAH KHAN received the B.Sc. degree in electrical engineering from the University of Engineering and Technology (UET), Peshawar, Pakistan, in 2012, the M.S. degree in electrical engineering from the COMSATS University Islamabad (CUI), Pakistan, in 2019, and the B.Ed. degree from the Abbottabad University of Science and Technology (AUST), Pakistan, in 2022. He is currently pursuing the Ph.D. degree in information and communications technology with the Norwegian University of Science and Technology (NTNU), Norway. He was a Lecturer with different institutions. His research interests include wireless communications, modeling and estimation of wireless propagation channels, MIMO-OFDM systems, and signal processing.

MOHAMMAD MAHMUDUL HASAN received the B.Tech. and M.Tech. degrees in electronics and telecommunication engineering from KIIT University, India. He is currently pursuing the Ph.D. degree in information and communication technology with the Norwegian University of Science and Technology (NTNU), Norway. In 2010, he was an Assistant Professor with KIIT University. From 2011 to 2022, he was an Assistant Professor with the Department of Electronics and Communication Engineering, University of Information Technology and Sciences (UITS), Bangladesh. He has written and/or edited several journals and conference articles. His research interests include wireless communications, sensor networks, array signal processing, intelligent systems, machine learning, antenna engineering, massive MIMO, and millimeter wave communications.

• • •

MICHAEL CHEFFENA received the M.Sc. degree in electronics and computer technology from the University of Oslo, Norway, in 2005, and the Ph.D. degree from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2008. In 2007, he was a Visiting Researcher with the Communications Research Centre Canada, Ottawa, ON, Canada. From 2009 to 2010, he was a Postdoctoral Researcher with the University Graduate Center, Kjeller, Norway, and the French Space Agency, Toulouse, France. He is currently a Full Professor with NTNU, Gjøvik, Norway. His research interests include modeling and prediction of propagation radio channels, signal processing, medium access control protocol design, antenna sensors, and sensor systems.