Channel Estimation Methods in Massive MIMO: A Comparative Review of Machine Learning and Traditional Techniques

Amalia Eka Rakhmania, Hudiono, Umi Anis Ro'isatin, and Nurul Hidayati

***Abstract*—Massive Multiple Input Multiple Output (MIMO) has emerged as a crucial technology in 5G and future 6G networks, offering unprecedented improvements in capacity, energy efficiency, and spectral efficiency. A key challenge for Massive MIMO systems is accurate and efficient channel estimation, which significantly impacts system performance. Traditional channel estimation methods such as Least Squares (LS) and Minimum Mean Square Error (MMSE) have been widely employed, but their limitations, particularly in complex and dynamic environments, have led to the exploration of more sophisticated approaches, including machine learning (ML)-based techniques. This review aims to compare traditional channel estimation methods with modern machine learning-based techniques in Massive MIMO systems, providing insights into their performance, computational complexity, and scalability. Furthermore, this paper outlines potential future research directions, emphasizing the integration of machine learning, optimization techniques, and hardware-friendly designs for enhanced performance.**

***Index Terms*—comparative study, machine learning, massive MIMO, traditional methods.**

I. INTRODUCTION

Massive MIMO (Multiple Input Multiple Output) is a revolutionary technology in wireless communication that enhances the capacity and efficiency of networks. It involves the deployment of a large number of antennas at the base station, allowing for the simultaneous transmission and reception of data to multiple users within the same frequency band. This capability significantly improves spectral efficiency and overall network performance, making it a key component in modern wireless systems, especially in the context of 5G and beyond.

The concept of Massive MIMO is built on the principles of spatial multiplexing and beamforming, which enable the base station to serve multiple users by exploiting the spatial dimensions of the wireless channel. By using hundreds of antennas, Massive MIMO can create highly directional beams that focus energy toward specific users, thereby reducing interference and improving signal strength. This technology not only increases capacity but also enhances energy efficiency, as it can adaptively allocate resources based on user demand and channel conditions [1]. However, this benefit comes with challenges, particularly in terms of accurate channel estimation, which is essential for the success of beamforming and resource allocation algorithms [2].

Channel estimation in Massive MIMO is inherently difficult due to the large number of antennas and the complexity of the wireless channel in high-mobility and dense environments. Traditional estimation methods such as LS [3] and MMSE [4] offer basic solutions but fail to cope effectively with increasing system complexity. These limitations have led to the application of machine learning-based techniques, which leverage large datasets and complex models to learn the channel's characteristics and provide more robust solutions.

Recent advancements in deep learning, particularly convolutional neural networks, have shown promise in enhancing channel estimation accuracy by capturing spatial correlations and temporal dynamics more effectively than traditional methods. Moreover, the integration of reinforcement learning approaches has opened new avenues for adaptive channel estimation, allowing systems to dynamically adjust their parameters based on real-time feedback from the environment.

This paper provides a comprehensive review of both traditional and ML-based methods for channel estimation in Massive MIMO, discusses their strengths and weaknesses, and explores possible future developments.

II. TRADITIONAL CHANNEL ESTIMATION METHODS

A. Least Squares (LS)

The Least Squares (LS) method is a widely used approach for channel estimation in wireless communication systems, particularly in OFDM and MIMO configurations, due to its simplicity and computational efficiency [5], [6]. While LS is simple and computationally efficient, it suffers from high mean square error, especially at low signal-to-noise ratios [7]. These approaches aim to improve estimation accuracy while maintaining low complexity.

This work was supported by State Polytechnic of Malang, Indonesia under the DIPA Funding under the contract SP DIPA-023.18.2.77606/2024.
Corresponding author: Amalia Eka Rakhmania.

Amalia Eka Rakhmania is with the Department of Electrical Engineering, State Polytechnic of Malang, Indonesia (e-mail: amaliaeka.rakhmania@polinema.ac.id).

Hudiono is with the Department of Electrical Engineering, State Polytechnic of Malang, Indonesia (e-mail: hudiono@polinema.ac.id).

Umi Anis Ro'isatin is with the Department of Mechanical Engineering, State Polytechnic of Malang, Indonesia (e-mail: umi.anis@polinema.ac.id).

Nurul Hidayati is with the Department of Electrical Engineering, State Polytechnic of Malang, Indonesia (e-mail: nurulhid8@polinema.ac.id).

DOI: 10.36244/ICJ.2025.1.3

Least squares methods are pivotal in channel estimation for MIMO (Multiple Input Multiple Output) and Massive MIMO systems, where accurate channel state information (CSI) is essential for optimizing transmission strategies. These methods focus on estimating the characteristics of communication channels by minimizing the sum of the squares of the differences between observed and estimated values, which is particularly relevant in MIMO systems that utilize multiple antennas at both the transmitter and receiver to enhance communication performance.

In MIMO systems, the least squares method serves as a low-complexity design approach for parameter estimation, allowing for effective retrieval of channel state information even when pilot sequences are limited. The linear least squares problem, often referred to as regression analysis, provides a closed-form solution that is beneficial for estimating parameters in these complex systems. This is crucial because the performance of MIMO systems heavily relies on accurate channel estimation to mitigate the effects of noise and interference, which can significantly degrade transmission quality. Moreover, the application of least squares methods in Massive MIMO systems is particularly advantageous due to the large number of antennas involved. These systems can leverage the additional antennas to compensate for the reduced number of pilot signals, thus maintaining reliable channel estimation. The estimation error analysis is also vital, as it evaluates the accuracy of the channel estimates obtained through least squares methods, helping to improve overall system performance and reliability.

Spatial multiplexing, a technique that allows multiple data streams to be transmitted simultaneously over the same channel, further illustrates the importance of least squares methods in maximizing data rates in MIMO systems. Accurate channel estimation is essential for effectively implementing spatial multiplexing, as it directly impacts the system's ability to handle multiple independent data streams without interference. LS assumes that the channel is static and deterministic. It estimates the channel coefficients by solving a system of linear equations. The solution that minimizes the sum of squared errors between the estimated and actual received signals is chosen as the channel estimate.

In Massive MIMO systems, the LS estimator is commonly employed for uplink channel estimation, although its performance is highly dependent on the choice of training pilots and is sensitive to outlier measurements. To address these challenges, techniques such as sparse Bayesian learning (SBL) have been integrated with LS to improve estimation accuracy by exploiting channel sparsity and separating impulsive noise from the signal of interest [8]. Additionally, combining LS with singular value decomposition (SVD) has been proposed to enhance channel estimation accuracy by using SVD to calculate the initial channel matrix, followed by LS signal detection to refine the channel state information (CSI) [9], [10]. In MIMO-OFDM systems, LS estimation is used alongside adjustable phase shift pilots (APSPs) to reduce pilot overhead and improve the mean square error (MSE) of the channel estimate [11], [12].

Furthermore, LS methods are favored in 5G wireless communications for their practicality and ease of implementation, despite being less accurate than minimum mean square error (MMSE) methods, which require channel statistics [13]. The LS method's performance can be enhanced by increasing the number of base station antennas, which improves the bit error rate (BER) [14].

The Least Squares (LS) method in channel estimation for Massive MIMO systems offers several advantages and disadvantages. One of the primary advantages of LS is its simplicity and ease of implementation, making it resource-friendly and practical for industry applications [13]. LS does not require prior statistical knowledge of the channel, which is beneficial in scenarios where such information is unavailable [15]. Additionally, LS can be computationally efficient, especially when optimized to minimize the relative error between estimated and actual channel coefficients, leading to faster data processing [13].

However, LS has notable disadvantages, particularly its poor performance in low signal-to-noise ratio (SNR) environments, where it provides less accurate channel estimates compared to more sophisticated methods like Minimum Mean Square Error (MMSE) [16], [15]. LS can also introduce significant modeling errors when used to decouple pilot matrices, which can affect the accuracy of channel estimation [17]. Furthermore, LS is less effective in handling pilot contamination and interference, which can degrade the uplink rate in Massive MIMO systems [18]. Despite these limitations, LS remains a widely used method due to its straightforward implementation and low computational complexity, making it suitable for scenarios where computational resources are limited [19].

While LS is straightforward to use and efficient in terms of computation, it does have some drawbacks. Firstly, it is vulnerable to interference, particularly in noisy conditions. Secondly, LS ignores any previous statistical data about the channel, which can affect its precision. Thirdly, LS requires many pilot symbols to obtain reliable channel estimates. Although LS techniques are fundamental in estimating channels for MIMO systems, combining them with other methods and algorithms is essential for overcoming their shortcomings and improving performance in large-scale MIMO scenarios.

B. *Minimum Mean Square Error (MMSE)*

Besides being used as a signal detection technique [20], the Minimum Mean Square Error (MMSE) method addresses the limitations of LS by incorporating noise statistics. It aims to minimize the mean squared error between the estimated and actual channel response. This approach generally provides better accuracy than LS, especially in noisy conditions.

The Minimum Mean Square Error (MMSE) method is a prominent channel estimation technique in Massive MIMO systems, known for its high accuracy in acquiring channel state information (CSI) essential for optimal system performance. MMSE assumes that the channel is a random variable with known statistical properties. It estimates the channel coefficients by minimizing the expected squared error between the estimated and actual channel response. This minimization is achieved by using the channel's prior distribution and the noise statistics. MMSE estimators are particularly effective in environments with spatially correlated Rician fading, where they can achieve improved normalized mean square error (NMSE) as the Rician K-factor decreases, indicating better performance under Rayleigh fading conditions [21]. However, the classical linear MMSE estimator is computationally intensive, especially in Massive MIMO contexts, prompting the development of alternative methods like the rank-1 subspace channel estimator, which offers lower complexity while maintaining high accuracy [22].

The MMSE method's reliance on accurate channel covariance matrices is a critical factor, as imperfections in these matrices can significantly affect estimation accuracy [23]. To address this, techniques such as the generalized eigenvalue decomposition (GEVD) have been proposed to estimate low-rank channel covariance matrices, enhancing MMSE performance in uplink cellular systems [24]. Additionally, model-based approaches and Bayesian estimators have been explored to reduce computational complexity while maintaining estimation quality [25], [23]. Despite these advancements, MMSE estimators still face challenges such as interference in multi-user environments, which can be mitigated by incorporating channel estimation errors into the MMSE detector [26].

One of the primary advantages of MMSE is its ability to provide accurate channel state information (CSI), which is crucial for the performance of Massive MIMO systems, especially in uplink scenarios where interference between user equipments (UEs) can be significant [24], [27]. MMSE estimators are effective in mitigating interference and improving spectral and energy efficiency, particularly when dealing with pilot contamination [28]. Additionally, MMSE can be adapted to various system configurations, such as those involving spatially correlated Rician fading channels, where it shows improved normalized mean square error (NMSE) as the Rician K-factor decreases [21].

However, MMSE channel estimation also has notable disadvantages, including its computational complexity, which can be a significant challenge in systems with large numbers of antennas or when one-bit quantization is used at the receiver [29]. To address this, techniques such as polynomial expansion have been proposed to reduce complexity while maintaining estimation performance [30]. Furthermore, MMSE requires accurate estimates of channel covariance matrices, which can be difficult to obtain, especially in low-rank scenarios [24], [27]. Despite these challenges, MMSE remains a popular choice due to its optimality in estimation theory and its ability to adapt to different channel conditions and system requirements [29], [31].

Both the Minimum Mean Square Error (MMSE) and Least Squares (LS) methods are widely used for channel estimation, each with distinct advantages and limitations. The MMSE estimator is known for its optimality in minimizing the mean square error, making it highly effective in scenarios with high interference and noise, as it requires knowledge of the channel covariance matrix to mitigate interference between user equipments (UEs) in neighboring cells [24], [27]. This method is particularly beneficial in spatially correlated Rician fading channels, where it achieves lower normalized mean square error (NMSE) as the Rician K-factor decreases [21]. However, MMSE's computational complexity and requirement for channel covariance information can be a drawback [29]. On the other hand, the LS method is simpler and does not require prior knowledge of the channel statistics, making it easier to implement [32]. It is effective in scenarios with high signal-to-noise ratio (SNR) and low interference, as demonstrated in MIMO-OFDM systems [32]. Despite its simplicity, LS can suffer from higher estimation errors compared to MMSE, especially in correlated channels [33]. In 5G systems, a modified entropy-based LS (MELS) has been proposed to enhance LS performance, outperforming both LS and MMSE at high SNR values [34].

While MMSE remains a robust method for channel estimation in Massive MIMO systems, ongoing research continues to refine its efficiency and accuracy, addressing computational and interference challenges [34], [35].

C. *Compressed Sensing (CS)*

Compressed Sensing (CS) leverages the sparse nature of wireless channels to reduce the number of required pilot symbols. CS is based on the principle that many signals, including wireless channels, can be represented as sparse vectors in a suitable basis. This means that only a small number of coefficients are non-zero. CS algorithms exploit this sparsity to recover the channel coefficients from a smaller number of measurements than would be required by traditional methods. This technique is particularly effective for Massive MIMO channels, especially in millimeter-wave (mmWave) communications. CS reconstructs sparse channels from fewer measurements, thereby reducing the pilot overhead.

In frequency division duplex (FDD) systems, the pilot overhead is particularly burdensome, and CS offers a solution by leveraging the sparsity of the channel. For instance, structured compressive sensing (SCS) schemes reduce pilot overhead by exploiting spatio-temporal common sparsity in delay-domain MIMO channels, using non-orthogonal pilots and adaptive algorithms like the adaptive structured subspace pursuit (ASSP) to enhance estimation accuracy [36] [37].

In MIMO-OTFS systems, radar sensing information is utilized to aid channel estimation by identifying strong angle-delay-Doppler taps, transforming the problem into a sparse recovery task [38]. Deep learning approaches, such as the two-step orthogonal matching pursuit (OMP) method, integrate CS with neural networks to improve channel state information (CSI) estimation in mmWave systems, even in low SNR conditions [39] [40]. Additionally, algorithms like the zebra optimization-based CoSaMP enhance estimation accuracy by optimizing the atomic matching process [41]. The generalized block adaptive matching pursuit (gBAMP) algorithm further refines channel estimation by optimizing index sets and using adaptive iterative stop conditions [42]. These methods collectively demonstrate that CS-based techniques can significantly reduce pilot and feedback overhead while maintaining high estimation accuracy, thereby enhancing the spectral and energy efficiency of massive MIMO systems [43] [44].

Compared to LS and MMSE, CS offers several advantages. First, it can significantly reduce the pilot overhead, which is especially important in scenarios with limited resources. Second, CS can provide accurate channel estimates even with a small number of measurements. However, CS also has some disadvantages. Its performance may degrade in non-sparse environments, and it can be computationally expensive for sparse recovery.

There are three types of CS methods stated in Table 1.

TABLE I
COMPRESSED SENSING METHODS

Method	Pros	Cons
Orthogonal Matching Pursuit (OMP)	Simple and computationally efficient, making it suitable for real-time applications. It is a greedy algorithm that iteratively selects the atom that has the highest correlation with the residual. This simplicity can be advantageous in scenarios where computational resources are limited. However, OMP can get stuck in local minima, especially when the signal is highly correlated. This can lead to suboptimal performance in some cases.	Can get stuck in local minima, especially when the signal is highly correlated. This can lead to suboptimal performance in some cases.
Basis Pursuit (BP)	Formulates the channel recovery problem as a convex optimization problem, which guarantees a global optimal solution. This makes BP more robust to noise and can provide better performance than OMP in some cases. However, BP can be computationally expensive, especially for large-scale problems.	Can be computationally expensive, especially for large-scale problems.
Compressive Sampling Matching Pursuit (CoSaMP)	Combines the strengths of OMP and BP. It is more robust to noise than OMP and can provide better performance than BP in some cases. However, CoSaMP is more complex than OMP and can be computationally expensive.	More complex than OMP and can be computationally expensive.

D. Kalman Filtering

Kalman filtering is a recursive estimation technique that updates the channel state based on prior knowledge and new measurements. It is well-suited for time-varying channels and is often used in scenarios involving high user mobility. Kalman filtering models the channel as a dynamic system with a state vector that evolves over time. The state vector contains the channel coefficients and their derivatives. Kalman filtering uses a prediction step to forecast the channel state based on the previous state and a measurement update step to correct the prediction based on new measurements.

Kalman filtering methods in channel estimation for massive MIMO systems are pivotal due to their ability to dynamically track and predict channel state information (CSI) in time-varying environments. The Multi-Stage Kalman Filter (MSKF) is a notable approach that leverages a reduced delay-line equalizer and Krylov-space based techniques to achieve fast convergence and reduced channel tracking errors, making it suitable for large-scale MMIMO systems [45]. The Vector Kalman Filter (VKF) is another method that utilizes autoregressive (AR) parameters from spatial channel models (SCM) to predict channels, offering a balance between computational complexity and prediction accuracy compared to machine learning-based methods [46] [47] [48]. In time-varying MIMO-OFDM systems, Kalman filters are used to track regularized zero-forcing (RZF) precoding coefficients, significantly reducing computational complexity by avoiding pseudo-inverse calculations [49].

Adaptive Kalman filters are advantageous in handling channel aging and varying user mobility, providing effective channel coefficient predictions for precoder construction [50]. Additionally, Kalman filters can estimate CSI based on received data without relying heavily on channel statistics, thus reducing the need for frequent pilot transmissions [51]. The use of Kalman filters in TDD massive MIMO systems allows for longer intervals between pilot transmissions, enhancing spectral efficiency by accommodating high Doppler spreads [52]. In STBC MIMO-OFDM systems, Kalman filters improve channel estimation accuracy by utilizing orthogonal pilot sequences and dynamic tracking properties, which are crucial in dynamic multipath environments [53]. Finally, the Kalman filter's ability to adaptively track time-domain changes in channels is enhanced by leveraging space-time reciprocity in antenna arrays, thus improving estimation accuracy in MIMO systems [54].

Compared to LS, MMSE, and CS, Kalman filtering offers several advantages. First, it is well-suited for time-varying channels and can provide accurate channel estimates even in dynamic environments. Second, Kalman filtering can be implemented in a recursive manner, which is efficient for real-time applications. However, Kalman filtering also has some disadvantages. It requires accurate initial state information and can be computationally intensive for large-scale systems.

III. MACHINE LEARNING-BASED CHANNEL ESTIMATION METHODS

Machine learning (ML) methods offer a promising alternative to traditional channel estimation techniques. By leveraging data-driven models, ML can learn complex channel characteristics from historical data, making it well-suited for Massive MIMO systems with their scale and dynamic nature.

Machine learning-based methods for channel estimation in Massive MIMO systems have emerged as powerful tools to address the challenges posed by the complexity and dynamic nature of wireless communication environments. Deep learning models, such as deep belief networks (DBNs) and convolutional neural networks (CNNs), have been effectively utilized to enhance channel estimation accuracy by learning spatial structures and channel statistics, as demonstrated by the DBN-BES technique, which achieves low root mean square error (RMSE) even in low signal-to-noise ratio (SNR) conditions [55].

In high-mobility scenarios, deep learning frameworks like the one proposed for MIMO-OTFS systems leverage CNNs to transform frequency-selective fading channels into quasi-time-invariant channels, significantly improving bit error rate (BER) and normalized mean squared error (NMSE) while reducing computational complexity by 80% [56]. Additionally, CNN-based models have been shown to outperform traditional methods like least squares (LS) and minimum mean square error (MMSE) in low SNR regimes, providing flexibility across various channel conditions without requiring prior statistical knowledge [15]. Other innovative approaches include the use of graph neural networks (GNNs), which incorporate system topology to improve generalization across different antenna configurations [57].

Furthermore, learning-based methods employing non-orthogonal pilots in grant-free multiple access scenarios have demonstrated promising performance in achieving low bit error rates in Massive MIMO systems [58]. Techniques such as the Spatial-Frequency UNet++ exploit spatial and frequency associations to enhance channel estimation accuracy [59].

These advancements highlight the potential of machine learning to not only improve estimation accuracy but also reduce computational complexity, making them suitable for real-world applications in 5G and beyond [60], [16].

A. Deep neural networks (DNNs)

Deep neural networks (DNNs) have been successfully applied to channel estimation tasks. DNNs can approximate the mapping between received pilot signals and the channel response, capturing non-linearities and complex relationships in wireless channels. This enables them to outperform traditional methods, especially in dynamic environments. However, DNNs require large training datasets and can be computationally expensive to train.

Implementing channel estimation in Massive MIMO systems using DNNs involves several innovative approaches that leverage the capabilities of deep learning to enhance performance and efficiency. One such method is a two-stage estimation process that combines pilot-aided and data-aided channel estimation. In the first stage, a two-layer neural network (TNN) and a deep neural network (DNN) are used to jointly design the pilot and the channel estimator, optimizing the pilot length relative to the number of transmit antennas. This is crucial because traditional methods assume the pilot length is equal to or larger than the number of antennas, which is not always feasible in Massive MIMO systems due to resource constraints [61]. The second stage involves further refining the channel estimation accuracy through iterative processes using another DNN, which minimizes the mean square error (MSE) of channel estimation. This iterative approach is shown to converge quickly, typically within five iterations, making it practical for real-time applications [61]. Additionally, deep learning-based methods can be categorized into data-driven and model-driven approaches. Data-driven methods use DNNs to directly map received signals to channel parameters, while model-driven methods, such as those using sparse Bayesian learning (SBL), unfold traditional algorithms into DNNs to capture complex channel sparsity structures effectively [62].

Another approach involves using a Channel State Information Network combined with a gated recurrent unit (CsiNet-GRU) to enhance the recovery quality and balance the trade-off between compression ratio and complexity in Massive MIMO systems. This method also employs dropout techniques to reduce overfitting during the learning process, resulting in significant performance improvements over existing techniques [63]. Furthermore, the use of reinforcement learning (RL) in semi-data-aided channel estimation can reduce communication latency by selecting reliable detected symbol vectors, thus optimizing the channel estimation process in Massive MIMO systems [64]. These methods collectively demonstrate the potential of DNNs to address the challenges of channel estimation in Massive MIMO systems, offering solutions that improve accuracy, reduce latency, and manage computational complexity effectively.

B. Convolutional neural networks (CNNs)

Convolutional neural networks (CNNs) are particularly effective for extracting spatial features from data. In Massive MIMO, CNNs can process channel state information (CSI) and predict the channel based on the spatial correlation between antennas. This approach offers improvements in accuracy and robustness, especially in multi-user MIMO systems. However, CNNs require careful tuning of the network architecture and can be computationally expensive during inference.

CNNs have been increasingly utilized for channel estimation in wireless communication systems due to their ability to handle complex, non-linear problems and their robustness to imperfect channel state information (CSI). CNNs are particularly effective in Massive MIMO systems, where they can refine coarse least squares (LS) estimations by exploiting channel correlations in both frequency and time domains, leading to improved performance and reduced overhead [62].

The CNN-based approach is advantageous because it can process imperfect CSI with strong robustness, which is crucial in practical scenarios where perfect CSI is often unattainable [65]. The architecture typically involves convolutional layers that extract features from the input data, followed by fully connected layers that convert these features into the desired output dimensions, such as combiner weights or refined channel estimates [66]. This structure allows CNNs to learn the statistics of the channel model and acquire sparsity features in the angle domain, which are essential for accurate channel estimation [66]. Moreover, CNNs have been shown to significantly decrease computational complexity compared to traditional algorithms, making them a practical choice for real-time applications [65].

In some implementations, CNNs are combined with other neural network architectures, such as long short-term memory (LSTM) networks, to enhance their ability to handle fast time-varying channels and further improve estimation accuracy [66]. The use of CNNs in channel estimation is part of a broader trend of applying deep learning techniques to various physical layer problems in wireless communications, demonstrating their versatility and effectiveness in addressing the challenges posed by Massive MIMO systems [62].

CNNs offer a promising solution for channel estimation by providing a balance between performance, robustness, and computational efficiency, which are critical for the next generation of wireless communication systems. Moreover, the integration of attention mechanisms within these architectures can further refine the model's focus on relevant features, leading to even greater improvements in performance and adaptability in dynamic environments.

C. Recurrent neural networks (RNNs)

Recurrent neural networks (RNNs) and their variant LSTM are designed to handle sequential data and capture time dependencies in channel matrices, making them ideal for time-varying channel estimation in mobile environments. LSTMs can capture long-term dependencies in time-series data, which is particularly useful for tracking slow-fading channels in Massive MIMO systems. However, RNNs and LSTMs can be computationally expensive and require large-scale training.

The RNN-based approach leverages the inherent time and frequency correlations in wireless channels, which allows for more accurate channel estimation without the need for extensive channel-state-information (CSI) feedback or pilot assignment [67]. The LSTM networks are trained to predict the current channel matrix using a series of past channel matrices, optimizing the number of time steps considered to balance between capturing time correlation and avoiding excessive randomness [67]. This method is particularly beneficial in scenarios with long time coherence and channel hardening, where it outperforms traditional blind detection and least squares (LS) estimators [67].

Additionally, RNNs, including LSTM and Gated Recurrent Unit (GRU) architectures, have been proposed for doubly-selective channel estimation, addressing challenges posed by multi-path propagation and Doppler effects in dynamic environments. These RNN-based schemes demonstrate superior performance in terms of bit error rate and throughput across various mobility scenarios and modulation orders, while also reducing computational complexity and execution time compared to conventional methods [68].

Furthermore, the integration of deep learning techniques, such as combining LSTM with deep neural networks (DNNs), enhances the generalization capabilities of channel estimation models, allowing them to adapt more efficiently to non-stationary environments and reduce pilot overhead [69], [70]. These advancements highlight the potential of RNNs in improving the accuracy and efficiency of channel estimation in modern wireless communication systems, making them a promising tool for future developments in this field.

D. Autoencoders

Autoencoders are used for dimensionality reduction in high-dimensional systems like Massive MIMO. They compress the channel state information into a lower-dimensional space while maintaining key features for accurate channel reconstruction. This can reduce the computational complexity of channel estimation and improve efficiency.

Autoencoders, particularly variational autoencoders (VAEs) and convolutional neural network (CNN) autoencoders, have been explored for channel estimation in various contexts. The use of VAEs in channel estimation is highlighted in the context of underdetermined systems, where they are employed to parameterize an approximation to the mean squared error (MSE)-optimal estimator. This approach is advantageous as it does not require perfect channel state information (CSI) during the offline training phase, which is a significant improvement over other deep learning-based methods that typically demand such data [71].

Additionally, CNN autoencoders have been utilized in differential encoding networks for CSI estimation, where they have shown to outperform traditional compressed sensing approaches. These autoencoders are used to encode and feedback estimation errors, leveraging their ability to compress error terms effectively. This method combines unrolled optimization networks with autoencoders, demonstrating superior performance compared to previous autoencoder-based approaches [72].

Furthermore, the integration of autoencoders in channel estimation frameworks allows for the exploitation of sparsity in channel realizations, particularly in the angular domain, which enhances the network's ability to handle interference and improve estimation accuracy [73]. These applications underscore the versatility and effectiveness of autoencoders in addressing the challenges of channel estimation, such as pilot contamination and feedback compression, in modern wireless communication systems. The use of autoencoders, therefore, represents a promising direction for improving the efficiency and accuracy of channel estimation processes in various MIMO system configurations. However, autoencoders require large datasets and their performance is limited by the network design.

E. Reinforcement learning (RL)

Reinforcement learning (RL) is another promising approach to channel estimation. RL agents can learn policies that optimize the estimation process over time by interacting with the environment. This allows RL to adapt to different channel conditions without explicit training data, making it robust to changing environments.

The use of RL for channel estimation is explored in the context of optimizing the selection of detected symbols for a semi-data-aided channel estimator. This approach involves formulating an optimization problem that adaptively selects these symbols, which is then solved using an efficient RL algorithm. The RL-based channel estimator demonstrates superior performance in terms of normalized mean square error (NMSE) and block error rate (BLER) compared to conventional pilot-aided methods by leveraging detected symbol vectors as additional pilot signals [64]. This method is particularly advantageous as it can be universally applied to any soft-output data detection method that computes log-likelihood ratios (LLRs) of transmitted data bits, thus enhancing its applicability across various data detection scenarios [64].

The RL algorithm's ability to utilize a priori probabilities (APPs) obtained from maximum a posteriori (MAP) data detection methods further underscores its versatility and effectiveness in channel estimation tasks. This approach not only improves the accuracy of channel estimation but also reduces the reliance on traditional pilot signals, thereby optimizing the overall communication system performance. The integration of RL in channel estimation represents a significant advancement in wireless communication, offering a robust framework for enhancing signal processing capabilities in complex and dynamic environments. However, RL requires exploration and may converge slowly.

TABLE II
ML-BASED METHODS

Method	Advantages	Disadvantages
DNNs	- Highly flexible models capable of learning complex non-linear relationships between received signals and channel coefficients. - Can capture both spatial and temporal correlations in the channel, making them suitable for a wide range of wireless environments. - Relatively easy to train and deploy compared to other ML methods.	- Require large training datasets, which can be challenging to obtain and label. - Computationally expensive, especially for large-scale models, requiring significant computational resources. - Sensitive to overfitting, which can lead to poor generalization performance.
CNNs	- Efficiently extract spatial features from CSI, which is crucial for channel estimation in Massive MIMO systems. - Computationally efficient compared to DNNs, making them suitable for real-time applications. - Relatively easy to train and deploy, especially when using pre-trained models.	- May struggle to capture temporal dependencies in time-varying channels, limiting their effectiveness in certain environments. - Require careful tuning of the network architecture and hyperparameters to achieve optimal performance.
RNNs/LSTMs	- Handle sequential data effectively, making them well-suited for time-varying channels in mobile environments. - Can capture long-term dependencies in the channel, which is important for predicting future channel states. - Relatively easy to train and deploy compared to other ML methods.	- Computationally expensive, especially for large-scale models and long sequences. - Can suffer from the vanishing gradient problem, which can make training difficult. - Sensitive to noise and outliers in the data.
Autoencoders	- Reduce the dimensionality of CSI, which can improve computational efficiency and reduce storage requirements. - Can be trained unsupervised, which can be advantageous when labeled data is limited. - Can capture important features of the channel while reducing noise and redundancy.	- May not capture all important features of the channel, leading to suboptimal performance. - Sensitive to noise and outliers in the data. - Difficult to train, especially for complex architectures.
RL	- Adapts to changing channel conditions without requiring explicit training data. - Can optimize channel estimation performance over time, improving accuracy and efficiency. - Can be used in environments with limited or no prior knowledge of the channel.	- Computationally expensive, especially for complex environments and large state spaces. - Can converge slowly, especially in challenging environments. - Difficult to tune and evaluate RL agents.

In addition to these methods, hybrid approaches that combine ML with traditional techniques have also been explored. For example, hybrid methods can use ML to learn the parameters of a traditional channel model or to improve the accuracy of traditional estimation algorithms.

One of the key challenges in ML-based channel estimation is the need for large training datasets. These datasets must be representative of the diverse channel conditions that the system will encounter in practice. Collecting and labeling such datasets can be time-consuming and expensive. To address this challenge, researchers have explored techniques such as data augmentation, transfer learning, and generative models.

Another challenge is the computational cost associated with training and deploying ML models. DNNs, CNNs, and RNNs can be computationally intensive, especially for large-scale models. To address this challenge, researchers have explored techniques such as model compression, quantization, and hardware acceleration.

Despite these challenges, ML-based channel estimation methods offer a promising avenue for addressing the challenges of channel estimation in Massive MIMO systems. By leveraging the power of data-driven models, ML can learn complex channel characteristics and provide accurate channel estimates in dynamic and challenging environments. However, further research is needed to address the computational and data requirements of ML-based methods and to explore new hybrid approaches that combine the strengths of ML and traditional techniques.

IV. DISCUSSION

Machine learning (ML) methods offer a promising alternative to traditional channel estimation techniques. By leveraging data-driven models, ML can learn complex channel characteristics from historical data, making it well-suited for Massive MIMO systems with their scale and dynamic nature.

TABLE III
COMPARISON OF ML AND TRADITIONAL BASED METHODS

Method	Complexity	Accuracy	Adaptability	Pilot Overhead	Data Requirements	Hardware Requirements
LS	$O(n^2)$	High MSE at low SNR	Static channels	Requires many pilot symbols	No prior channel statistics required	Can run on CPUs
MMSE	$O(n^3)$	Low MSE, especially at low SNR	Requires channel statistics	Requires many pilot symbols, but fewer than LS	Requires channel covariance matrix	May require GPUs for large-scale system
Compressed Sensing	Sparse recovery algorithms can be computationally intensive	Good for sparse channels, degrades in non-sparse environment	Works best in sparse environment	Significantly reduces pilot overhead	Requires sparse channel representation	May require GPUs for real-time applications
Kalman Filtering	Recursive updates, $O(n^2)$	Good for time-varying channels	Excellent for dynamic environments	Requires periodic pilot updates	Requires initial state information	Can run on CPUs, but GPUs may speed up processing
DNN	Training: $O(n^4)$ Inference: $O(n^3)$	Captures non-linearities, low MSE	Adapts well to dynamic environment	Learns channel structure, hence reducing pilot overhead	Requires large labeled datasets	Requires GPUs for training and inference
CNN	Training: $O(n^4)$ Inference: $O(n^3)$	captures spatial correlations, low MSE	Adapts well to multi-user MIMO	Learns spatial features, reduces pilot overhead	Requires large labeled datasets	Requires GPUs for training and inference
RNN/LSTM	Training: $O(n^4)$ Inference: $O(n^3)$	Captures temporal dependencies, low MSE	Excellent for time-varying and mobile environments	Learns temporal features, reduces pilot overhead	Requires large labeled datasets with temporal sequences	Requires GPUs for training and inference
Autoencoder	Training: $O(n^4)$ Inference: $O(n^3)$	Good for dimensionality reduction, moderate MSE	Works well for high-dimensional systems	Compresses CSI, reduces pilot overhead	Requires large datasets for training	Requires GPUs for training and inference
Reinforcement Learning	Exploration and policy optimization can be computationally intensive	Adapts to changing environments, low MSE	Excellent for dynamic and non-stationary environments	Reduces reliance on pilot signals	Requires interaction with the environment for training	Requires GPUs for training and inference

A. Accuracy

ML-based methods generally outperform traditional methods, especially in complex and dynamic environments. They can capture intricate relationships in the channel that traditional methods may overlook, such as non-linear dependencies, spatial correlations, and temporal variations. This is particularly beneficial in scenarios with rapidly changing channel conditions or when the channel is highly correlated. For example, in Massive MIMO systems with a large number of antennas, ML methods can effectively exploit the spatial correlation between antennas to improve channel estimation accuracy.

B. Complexity

While ML methods often require more computational resources due to their complexity and the need for large training datasets, their improved accuracy and adaptability can justify the increased computational cost. In many cases, the benefits of ML-based methods outweigh the additional computational overhead, especially in applications where high accuracy and adaptability are critical. For instance, in autonomous vehicles or critical infrastructure, accurate channel estimation is essential for reliable communication, and the increased computational cost of ML methods may be acceptable in exchange for improved performance.

C. Pilot Overhead

ML methods can significantly reduce pilot overhead by learning the structure of the channel more efficiently. This is particularly beneficial in scenarios with limited resources or bandwidth constraints, such as in mobile communication systems or IoT networks. By reducing the number of pilot symbols required for channel estimation, ML methods can improve spectral efficiency and increase data throughput. For example, in IoT networks where devices have limited power and bandwidth, ML-based channel estimation can help reduce the overhead associated with transmitting pilot symbols, enabling more efficient communication.

D. Adaptability

ML methods like LSTMs and RL excel in environments with high mobility or time-varying channels. They can adapt to changing conditions and provide accurate channel estimates in real-time, which is essential for applications like mobile communication, vehicular networks, and wireless sensor networks. For instance, in mobile communication systems where users are constantly moving and the channel conditions are changing rapidly, ML-based methods can continuously learn and adapt to the channel, ensuring reliable communication even in challenging environments.

E. Additional Considerations

To provide a clearer and more structured overview of the additional considerations in machine learning-based channel estimation, Table IV summarizes key aspects such as data quality, model selection, interpretability, privacy, hardware acceleration, and hybrid approaches.

TABLE IV
ADDITIONAL CONSIDERATIONS

Aspects	Descriptions
Data Quality and Quantity Considerations	
Data Collection	Collect diverse datasets covering various environments (indoor, outdoor, urban, rural), frequency bands, and propagation conditions.
Data Cleaning	Remove outliers, inconsistencies, and errors using techniques like outlier detection, imputation, and normalization.
Data Augmentation	Generate additional training data through noise injection, rotation, and scaling to improve model robustness.
Data Labeling	Accurately label channel estimates in training data to provide correct supervision. This task may require domain expertise.
Model Selection and Hyperparameter Tuning	
Model Architecture	Choose a model suitable for the task (e.g., DNNs for non-linear relationships, CNNs for spatial features).
Hyperparameter Tuning	Experiment with learning rate, batch size, number of layers, and activation functions using techniques like grid search or Bayesian optimization.
Interpretability and Explainability	
Visualize Decisions	Use feature importance plots or decision trees to understand how the model makes predictions.
Identify Biases	Detect and mitigate biases arising from training data or model architecture.
Explain Behavior	Provide human-understandable explanations for model decisions to build trust and improve transparency.
Privacy and Security	
Data Encryption	Protect data from unauthorized access during transmission and storage.
Data Anonymization	Remove or disguise personal information to protect user privacy.
Model Security	Protect models from adversarial examples (misleading inputs) and model theft (stealing parameters).
Hardware Acceleration	
GPU-Based Training	Use GPUs to accelerate training, especially for large-scale models requiring parallel computations.

Aspects	Descriptions
TPU-Based Inference	Use TPUs to accelerate inference, making ML models more suitable for real-time applications.
Hybrid Approaches	
ML-Enhanced Traditional Methods	Use ML to learn parameters of traditional models or improve their accuracy (e.g., predicting channel coefficients).
Hybrid Architectures	Combine traditional and ML components (e.g., traditional estimator for initial estimation, ML for fine-tuning).

V. Future Research Directions

Channel estimation in Massive MIMO systems is a rapidly evolving field with numerous opportunities for future research. One promising avenue is the development of hybrid methods that combine the strengths of traditional and machine learning (ML) techniques.

1) **Hybrid Methods**

Hybrid methods can leverage the complementary advantages of traditional and ML-based approaches. For example, traditional methods can be used for initial channel estimation, providing a baseline estimate that can be refined by ML models. This can reduce the computational cost of pure ML models while improving their performance in dynamic environments. Additionally, hybrid methods can incorporate domain knowledge into the ML models, enhancing their interpretability and robustness.

One potential hybrid approach is to use a traditional channel estimator to provide an initial estimate of the channel, and then use an ML model to refine the estimate based on additional information, such as the received signal or the channel statistics. This can help to improve the accuracy of the channel estimate, especially in challenging environments. Another approach is to use ML models to learn the parameters of a traditional channel model, making it more adaptable to different channel conditions.

2) **Federated Learning**

Federated learning is another promising area of research for channel estimation in Massive MIMO systems. This technique allows multiple devices to train a shared ML model without sharing their raw data, preserving privacy and reducing communication overhead. Federated learning can be particularly useful in distributed Massive MIMO systems where channel data is collected from a large number of devices.

By using federated learning, channel estimation can be performed in a decentralized manner, reducing the reliance on a central server and improving privacy. Additionally, federated learning can enable the training of ML models on large-scale datasets that would be difficult or impossible to collect and process centrally.

3) **Real-Time ML Inference**

While training ML models can be computationally expensive, future work could focus on optimizing inference time to make real-time deployment feasible in Massive MIMO systems. Techniques like model pruning and quantization can help reduce the model size and speed up the estimation process. Additionally, hardware acceleration using specialized hardware like GPUs or TPUs can further improve the inference performance.

By optimizing inference time, ML-based channel estimators can be deployed in real-time applications, such as mobile communication systems and autonomous vehicles. This will enable more accurate and responsive channel estimation, leading to improved system performance.

4) **Cross-Layer Optimization**

There is growing interest in cross-layer optimization, where channel estimation is integrated with higher-layer functions like resource allocation and power control. By jointly optimizing these processes using ML models, system performance could be significantly enhanced. For example, ML models could be used to predict the channel conditions and allocate resources accordingly, or to optimize power control to maximize data throughput while minimizing interference.

Cross-layer optimization can help to achieve more efficient and reliable wireless communication by taking into account the interactions between different layers of the system. By jointly optimizing these layers, it is possible to achieve better overall system performance than by optimizing each layer in isolation.

5) **Explainable Machine Learning**

As ML models become more complex, their interpretability decreases. Future research should focus on developing explainable ML methods for channel estimation to increase transparency and trust in ML-based wireless systems. Explainable ML techniques can help to understand how the model makes decisions, identify potential biases, and improve the model's reliability.

Explainable ML is particularly important in critical applications where it is essential to understand how the model works and why it makes certain decisions. By making ML models more explainable, we can increase trust in their predictions and ensure that they are not biased or unfair.

In conclusion, channel estimation in Massive MIMO systems is a rapidly evolving field with numerous opportunities for future research. By developing hybrid methods, leveraging federated learning, optimizing real-time inference, exploring cross-layer optimization, and improving the explainability of ML models, we can continue to advance the state of the art in this critical area of wireless communication.

V. Conclusion

Massive MIMO systems are pivotal for modern wireless communication, offering significant improvements in capacity, spectral efficiency, and energy efficiency. However, accurate channel estimation remains a critical challenge, especially in dynamic and complex environments. Traditional methods like Least Squares (LS) and Minimum Mean Square Error (MMSE) are widely used due to their simplicity and computational efficiency, but they struggle in low SNR and high-mobility scenarios. In contrast, machine learning (ML)-based methods, such as Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs), have demonstrated superior performance by capturing complex spatial and temporal correlations in the channel. These methods reduce pilot overhead, improve accuracy, and adapt well to dynamic environments, though they require large datasets and significant computational resources.

Future research should focus on hybrid approaches that combine the strengths of traditional and ML-based methods, leveraging the simplicity of traditional techniques for initial estimates and the adaptability of ML for refinement. Additionally, advancements in federated learning, real-time ML inference, and cross-layer optimization can further enhance the efficiency and robustness of channel estimation in Massive MIMO systems. By addressing challenges such as data requirements, computational complexity, and model interpretability, ML-based methods hold great promise for advancing wireless communication in the era of 5G and beyond.

REFERENCES

[1] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications*, Nov. 2010, DOI: 10.1109/TWC.2010.092810.091092.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, Feb. 2014, DOI: 10.1109/MCOM.2014.6736761.

[3] R. Hidayat, A. F. Isnawati, and B. Setiyanto, "Channel estimation in MIMO-OFDM spatial multiplexing using Least Square method," in *International Symposium on Intelligent Signal Processing and Communication Systems*, Dec. 2011. DOI: 10.1109/ISPACS.2011.6146157.

[4] Z. Luo and D. Huang, "General MMSE Channel Estimation for MIMO-OFDM Systems," in *Vehicular Technology Conference*, Oct. 2008. DOI: 10.1109/VETECF.2008.151.

[5] N. Dubey and A. Pandit, "A Comprehensive Review on Channel Estimation in OFDM System," Mar. 2019, DOI: 10.24113/IJOSCIENCE.V5I3.201.

[6] Z. Fang and J. Shi, "Least Square Channel Estimation for Two-Way Relay MIMO OFDM Systems," *Etri Journal*, Oct. 2011, DOI: 10.4218/ETRIJ.11.0210.0424.

[7] B. Zhang, "A LS channel estimation algorithm based on adaptive filtering for OFDM system," *Journal of Xi'an University of Posts and Telecommunications*, Jan. 2011.

[8] G. Zhu, J. Dai, C. Chang, and W. Xu, "Robust uplink channel estimation for massive MIMO systems with general array geometry," *Digital signal processing*, Oct. 2022, DOI: 10.1016/j.dsp.2022.103793.

[9] S. Li, J. Li, H. Dong, and Z. Li, "A Massive MIMO Channel Estimation Algorithm Design Combined the SVD Method with LS Signal Detection," Aug. 2022, DOI: 10.1109/IWS55252.2022.9977532.

[10] "A Massive MIMO Channel Estimation Algorithm Design Combined the SVD Method with LS Signal Detection," *2022 IEEE MTT-S International Wireless Symposium (IWS)*, Aug. 2022, DOI: 10.1109/iws55252.2022.9977532.

[11] S. S. Devi, B. S. Kumar, S. Narayanan, and G. Vimal, "Channel Estimation of OFDM – MIMO," Dec. 2022, DOI: 10.1109/ICMNWC56175.2022.10032004.

[12] "Channel Estimation of OFDM – MIMO," *2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, Dec. 2022, DOI: 10.1109/icmnwc56175.2022.10032004.

[13] A. Hasan, S. M. A. Motakabber, F. Anwar, M. H. Habaebi, and M. I. Ibrahimy, "A Computationally Efficient Least Squares Channel Estimation Method for MIMO-OFDM Systems," *International Conference on Computer and Communication Engineering*, Jun. 2021, DOI: 10.1109/ICCCE50029.2021.9467142.

[14] A. Riadi, M. Boulouird, and M. M. Hassani, "Least Squares Channel Estimation of an OFDM Massive MIMO System for 5G Wireless Communications," Dec. 2018, DOI: 10.1007/978-3-030-21009-0_43.

[15] L. Carro-Calvo, A. de la Fuente, A. Melgar, and E. Morgado, "Massive MIMO Channel Estimation With Convolutional Neural Network Structures," *IEEE Transactions on Cognitive Communications and Networking*, Jan. 2024, DOI: 10.1109/tccn.2024.3435478.

[16] R. Thakur and V. N. Saxena, "A Survey on Learning-Based Channel Estimation Methods Used for 5G Massive Mimo System," Jul. 2023, DOI: 10.1109/icccnt56998.2023.10308216.

[17] G. Zhu, Y. Xia, and S. Li, "VBI-based Uplink Channel Estimation for Massive MIMO Systems," *IEEE International Conference on Electronic Information and Communication Technology*, Aug. 2022, DOI: 10.1109/ICEICT55736.2022.9909288.

[18] B. Sissokho, J.-P. Cances, and A. D. Kora, "Uplink Rate Based on Massive MIMO Channel Estimation Approach," *International Conference Frontiers Signal Processing*, Sep. 2018, DOI: 10.1109/ICFSP.2018.8552055.

[19] G. Sklivanitis, K. Tountas, D. A. Pados, and S. N. Batalama, "Small-Sample-Support Channel Estimation for Massive Mimo Systems," *International Conference on Acoustics, Speech, and Signal Processing*, Apr. 2018, DOI: 10.1109/ICASSP.2018.8461599.

[20] Jyoti P. Patra, Bibhuti Bhusan Pradhan, and M. Rajendra Prasad, "An Ordered QR Decomposition based Signal Detection Technique for Uplink Massive MIMO System", *Infocommunications Journal*, Vol. XVI, No 1, March 2024, pp. 20–25., DOI: 10.36244/ICJ.2024.1.3

[21] J. F. Arellano, C. D. Altamirano, H. R. C. Mora, N. O. Garzón, and F. D. A. García, "On the Performance of MMSE Channel Estimation in Massive MIMO Systems over Spatially Correlated Rician Fading Channels," *Wireless Communications and Mobile Computing*, Apr. 2024, DOI: 10.1155/2024/5445725.

[22] B. Li *et al*., "Beyond MMSE: Rank-1 Subspace Channel Estimator for Massive MIMO Systems," Apr. 2024, DOI: 10.48550/arxiv.2404.13603.

[23] A. H. Shatti and E. A. Hussein, "Low-complex Bayesian estimator for imperfect channels in massive muti-input multi-output system," *International Journal of Electrical and Computer Engineering*, Dec. 2022, DOI: 10.11591/ijece.v12i6.pp6261-6271.

[24] "GEVD-based Low-Rank Channel Covariance Matrix Estimation and MMSE Channel Estimation for Uplink Cellular Massive MIMO Systems," *2022 30th European Signal Processing Conference (EUSIPCO)*, Aug. 2022, DOI: 10.23919/eusipco55093.2022.9909626.

[25] X. Li, Z. Song, L. Liu, X. Sha, and Y. Li, "Model-Based Structured Covariance-Aided Channel Estimation for Massive MIMO Systems," *International Conference on Speech Technology and Human-Computer Dialogue*, Nov. 2022, DOI: 10.1109/ICCT56141.2022.10072431.

[26] N. Zhao *et al*., "The MMSE MIMO Detector Under a New LMMSE Channel Estimation Error Analysis Method," Jul. 2023, DOI: 10.1109/iceict57916.2023.10245661.

[27] R. V. Rompaey and M. Moonen, "GEVD-based Low-Rank Channel Covariance Matrix Estimation and MMSE Channel Estimation for Uplink Cellular Massive MIMO Systems," arXiv: Signal Processing, Nov. 2021, DOI: 10.48550/arXiv.2111.11902

[28] E. Mukubwa and O. Sokoya, "Comparison of Improved MMSE and the Semi-Blind Channel Estimation Methods," *International Conference on Artificial Intelligence*, Aug. 2020, DOI: 10.1109/ICABCD49160.2020.9183807.

[29] M. Ding, I. Atzeni, A. Tolli, and A. L. Swindlehurst, "On the Optimal MMSE Channel Estimation for One-Bit Quantized MIMO Systems," arXiv.org, Apr. 2024, DOI: 10.48550/arxiv.2404.05536.

[30] Y. Wang and P. Fortier, "Polynomial Expansion-Based MMSE Channel Estimation for Massive MIMO-GFDM Systems," *Vehicular Technology Conference*, Nov. 2020, DOI: 10.1109/VTC2020-FALL49728.2020.9348717.

[31] J. Mirzaei, F. Sohrabi, R. S. Adve, and S. Shahbazpanahi, "MMSE-Based Channel Estimation for Hybrid Beamforming Massive MIMO with Correlated Channels," *International Conference on Acoustics, Speech, and Signal Processing*, May 2020, DOI: 10.1109/ICASSP40776.2020.9053980.

[32] 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," *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Jun. 2022, DOI: 10.1109/hora55278.2022.9799887.

[33] M. H. Sadraei, M. S. Fazel, and A. M. Doost-Hoseini, "Ergodic spectral efficiency of massive MIMO with correlated Rician channel and MRC detection based on LS and MMSE channel estimation," *Iet Communications*, Oct. 2020, DOI: 10.1049/IET-COM.2019.0905.

[34] J. Kelvin, "Analytical Analysis on LS, MMSE and Modified Entropy-Based LS Channel Estimation Techniques for 5G Massive MIMO Systems," Jan. 2023, DOI: 10.1007/978-981-19-9512-5_40.

[35] Q. Liu, Y. Li, and J. Sun, "A Model-Driven Channel Estimation Method for Millimeter-Wave Massive MIMO Systems," *Sensors*, Feb. 2023, DOI: 10.3390/s23052638.

[36] Z. Gao, Y. Mei, and L. Qiao, "Compressive Sensing Sparse Channel Estimation in FDD Massive MIMO Systems," Oct. 2023, DOI: 10.1007/978-981-99-5394-3_3.

[37] "High Spectrum and Efficiency Improved Structured Compressive Sensing-Based Channel Estimation Scheme for Massive MIMO Systems," Jan. 2022, DOI: 10.1007/978-981-16-7610-9_19.

[38] S. Jiang and A. Alkhateeb, "Sensing Aided OTFS Massive MIMO Systems: Compressive Channel Estimation," May 2023, DOI: 10.1109/iccworkshops57953.2023.10283647.

[39] W. Tong, W. Xu, F. Wang, J. Shang, M. Pan, and J. Lin, "Deep Learning Compressed Sensing-Based Beamspace Channel Estimation in mmWave Massive MIMO Systems," *IEEE Wireless Communications Letters*, Sep. 2022, DOI: 10.1109/LWC.2022.3188530.

[40] "Deep Learning Compressed Sensing-Based Beamspace Channel Estimation in mmWave Massive MIMO Systems," *IEEE Wireless Communications Letters*, Sep. 2022, DOI: 10.1109/lwc.2022.3188530.

[41] S. Li and Y. Yang, "Improved compressed sensing channel estimation algorithm for MIMO-OFDM systems," May 2024, DOI: 10.1109/ccdc62350.2024.10587886.

[42] Y. Huang, Y. He, W. He, L. Shi, T. Cheng, and Y. Sui, "Channel Estimation in Massive MIMO Systems Based on Generalized Block Adaptive Matching Pursuit Algorithm," *IEEE Wireless Communications Letters,* Aug. 2020, DOI: 10.1109/LWC.2020.3013689.

[43] "A Compressive Sensing and Deep Learning-Based Time-Varying Channel Estimation for FDD Massive MIMO Systems," *IEEE Transactions on Vehicular Technology*, Aug. 2022, DOI: 10.1109/tvt.2022.3176290.

[44] N. Nouri, M. J. Azizipour, and K. Mohamed-pour, "A Compressed CSI Estimation Approach for FDD Massive MIMO Systems," *Iranian Conference on Electrical Engineering*, Aug. 2020, DOI: 10.1109/ICEE50131.2020.9260725.

[45] "Multi Stage Kalman Filter (MSKF) Based Time-Varying Sparse Channel Estimation With Fast Convergence," *IEEE open journal of signal processing*, Jan. 2022, DOI: 10.1109/ojsp.2021.3132583.

[46] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO Channel Prediction: Kalman Filtering Vs. Machine Learning," *IEEE Transactions on Communications*, Jan. 2021, DOI: 10.1109/TCOMM.2020.3027882.

[47] H. Kim, S. Kim, H. Lee, and J. Choi, "Massive MIMO Channel Prediction: Machine Learning Versus Kalman Filtering," *Global Communications Conference*, Dec. 2020, DOI: 10.1109/GCWKSHPS50303.2020.9367471.

[48] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO Channel Prediction: Kalman Filtering vs. Machine Learning," arXiv: Information Theory, Sep. 2020, DOI: 10.1109/TCOMM.2020.3027882

[49] S. Wu, X. Tao, D. Mishra, Y. Chen, and J. Xu, "Efficient Kalman Filter-Based Precoder Tracking for Time-Varying Massive MIMO-OFDM Systems," *IEEE Communications Letters*, Mar. 2020, DOI: 10.1109/LCOMM.2020.2981937.

[50] V. Arya and K. Appaiah, "Kalman Filter based Tracking for Channel Aging in Massive MIMO Systems," *International Conference on Signal Processing*, Jul. 2018, DOI: 10.1109/SPCOM.2018.8724442.

[51] A. Almamori and S. Mohan, "Estimation of channel state information for massive MIMO based on received data using Kalman filter," *IEEE Annual Computing and Communication Workshop and Conference*, Jan. 2018, DOI: 10.1109/CCWC.2018.8301698.

[52] S. Kashyap, C. Mollen, E. Björnson, and E. G. Larsson, "Performance analysis of (TDD) massive MIMO with Kalman channel prediction," *International Conference on Acoustics, Speech, and Signal Processing*, Mar. 2017, DOI: 10.1109/ICASSP.2017.7952818.

[53] R. Tang, X. Zhou, and C. Wang, "Kalman Filter Channel Estimation in 2 × 2 and 4 × 4 STBC MIMO-OFDM Systems," *IEEE Access*, Sep. 2020, DOI: 10.1109/ACCESS.2020.3027377.

[54] R. Tang, X. Zhou and C. Wang, "Kalman Filter Channel Estimation in 2 × 2 and 4 × 4 STBC MIMO-OFDM Systems," in *IEEE Access*, vol. 8, pp. 189 089–189 105, 2020, DOI: 10.1109/ACCESS.2020.3027377.

[55] O. Pabbati and R. Joshi, "An optimized deep learning model for a highly accurate DOA and channel estimation for massive MIMO systems," *International Journal of Communication Systems*, Jul. 2024, DOI: 10.1002/dac.5902.

[56] M. Payami and S. D. Blostein, "Deep Learning-based Channel Estimation for Massive MIMO-OTFS Communication Systems," Apr. 2024, DOI: 10.1109/wts60164.2024.10536672.

[57] M. Ye, X. Liang, C. Pan, Y. Xu, M. Jiang, and C. Li, "Channel Estimation for mmWave Massive MIMO Systems Using Graph Neural Networks," Aug. 2023, DOI: 10.1109/iccc57788.2023.10233621.

[58] S. Park, Y. Kim, J. Jang, and H.-J. Yang, "Learning-Based Channel Estimation Method with Non-Orthogonal Pilots for Grant-Free Multiple Access in Massive MIMO Systems," *The Journal of Korean Institute of Communications and Information Sciences*, Nov. 2023, DOI: 10.7840/kics.2023.48.11.1464.

[59] K. Long, H. Liu, D. Qiu, and J. Yang, "A Channel Estimator of Millimeter-Wave Massive MIMO Systems Through Deep Learning," Nov. 2023, DOI: 10.1109/apcap59480.2023.10470212.

[60] J. Meng, Z. Wei, Y. Zhang, B. Li, and C. Zhao, "Machine learning based low-complexity channel state information estimation," *EURASIP Journal on Advances in Signal Processing*, Oct. 2023, DOI: 10.1186/s13634-023-00994-4.

[61] C. J. Chun, J. M. Kang, and I. M. Kim, "Deep learning-based channel estimation for massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1228–1231, 2019. DOI: 10.1109/lwc.2019.2912378

[62] J. Gao, C. Zhong, G. Y. Li, J. B. Soriaga, and A. Behboodi, "Deep Learning-Based Channel Estimation for Wideband Hybrid MmWave Massive MIMO," *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3679–3693, Jun. 2023, DOI: 10.1109/TCOMM.2023.3258484.

[63] H. M. N. Helmy, S. El Daysti, H. Shatila, and M. Aboul-Dahab, "Performance Enhancement of Massive MIMO Using Deep Learning-Based Channel Estimation," vol. 1051, no. 1, p. 012 029, Feb. 2021, DOI: 10.1088/1757-899X/1051/1/012029.

[64] T.-K. Kim, Y.-S. Jeon, J. Li, N. Tavangaran, and H. V. Poor, "Semi-Data-Aided Channel Estimation for MIMO Systems via Reinforcement Learning," *IEEE Transactions on Wireless Communications*, p. 1, Apr. 2022, DOI: 10.1109/twc.2022.3227312.

[65] X. Bao, W. Feng, J. Zheng, and J. Li, "Deep CNN and Equivalent Channel Based Hybrid Precoding for mmWave Massive MIMO Systems," *IEEE Access*, vol. 8, pp. 19 327–19 335, Jan. 2020, DOI: 10.1109/ACCESS.2020.2967402.

[66] H. Hirose, T. Ohtsuki and G. Gui, "Deep Learning-Based Channel Estimation for Massive MIMO Systems With Pilot Contamination," in *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 67–77, 2021, DOI: 10.1109/OJVT.2020.3045470.

[67] T. Faghani, A. Shojaeifard, K.-K. Wong, and A. Aghvami, "Recurrent Neural Network Channel Estimation Using Measured Massive MIMO Data" *34th IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC 2023*, Toronto, ON, Canada, September 5-8, 2023," Sep. 2023, DOI: 10.1109/PIMRC48278.2020.9217192

[68] A. K. Gizzini and M. Chafii, "RNN Based Channel Estimation in Doubly Selective Environments," *IEEE transactions on machine learning in communications and networking*, vol. 2, pp. 1–18, Jan. 2024, DOI: 10.1109/tmlcn.2023.3332021.

[69] J. Abouei et al., "RNN-FL-based Channel Estimation Approach in mmWave Massive MIMO Systems" *Authorea Preprints*, 2023.

[70] L. Ge, C. Shi, S. Niu, and G. Chen, "Mixed RNN-DNN based channel prediction for massive MIMO-OFDM systems," *Iet Communications*, Oct. 2023, DOI: 10.1049/cmu2.12685.

[71] M. Baur, N. Turan, B. Fesl, and W. Utschick, "Channel Estimation in Underdetermined Systems Utilizing Variational Autoencoders," Apr. 2024, DOI: 10.1109/icassp48485.2024.10447622.

[72] M. Rosario and Z. Ding, "Learning-Based MIMO Channel Estimation under Practical Pilot Sparsity and Feedback Compression", *IEEE Transactions on Wireless Communications* 22.2 (2022): 1161–1174, DOI: 10.1109/TWC.2022.3202750

[73] A. Kasibovic, B. Fesl, M. Baur, and W. Utschick, "Addressing Pilot Contamination in Channel Estimation with Variational Autoencoders", *arXiv preprint arXiv:*2409.07071, 2024, DOI: 10.48550/arXiv.2409.07071.

Amalia Eka Rakhmania received the Bachelor of Engineering degree from the University of Brawijaya, Malang, Indonesia, in 2012, Master of Science from National Central University, Taiwan, and Master of Engineering degree from the University of Brawijaya, Malang, Indonesia in 2015 all in electrical engineering, majoring in Telecommunication. She is currently a lecturer at the Department of Electrical Engineering, State Polytechnic of Malang, Indonesia. Her research interests include wireless and optical communication, interference mitigation, and the internet of things for smart cities.

Hudiono Hudiono received the Bachelor of Engineering and Master of Engineering. in electrical engineering, majoring in Telecommunication degree from the Tenth November Institute of Technology, Surabaya, Indonesia, in 1993 and 1998, respectively. He is currently a lecturer at the Department of Electrical Engineering, State Polytechnic of Malang, Indonesia. His research interests include radio communication, antenna propagation, wireless sensor network, and the internet of things for smart cities.

Umi Anis Ro'isatin received her Bachelor and Master of English Education from State University of Malang, Malang, Indonesia in 1988 and 2009, respectively. She is currently a lecturer at the Department of Mechanical Engineering, State Polytechnic of Malang, Indonesia. Her research interests include English for specific purpose, pedagogy, and English for second language.

Nurul Hidayati received the Bachelor of Engineering and Master of Engineering. in electrical engineering, majoring in Telecommunication degree from the Tenth November Institute of Technology, Surabaya, Indonesia, in 2016 and 2018, respectively. She is currently a lecturer at the Department of Electrical Engineering, State Polytechnic of Malang, Indonesia. Her research interests include internet of things, cooperative wireless network, and algorithms for wireless sensor network.