Channel Estimation and Pilot Overhead Reduction in OFDM Systems using Compressed Sensing Dynamic Mode Decomposition

Fayad Haddad, *Graduate Student Member, IEEE,*
Carsten Bockelmann, *Member, IEEE* and Armin Dekorsy, *Senior Member, IEEE*

***Abstract*—This work investigates the potential of employing the approach Compressed Sensing Dynamic Mode Decomposition (CS-DMD) in the context of time-varying wireless channels. To the best of the authors' knowledge, this marks the first instance of utilizing CS-DMD for pilot-based channel estimation in Orthogonal Frequency Division Multiplexing (OFDM) systems. The effectiveness of this method is compared with two advanced deep learning-based channel estimation techniques: Interpolation-ResNet and Learned Approximate Message Passing (LAMP). Furthermore, we leverage the advantageous characteristics of DMD in analyzing complex nonlinear dynamic systems to predict the future state of the channel, thereby reducing the required pilot signals. Simulation results show that utilizing CS-DMD can achieve superior channel estimation performance with less pilot overhead.**

***Index Terms*—Channel estimation, compressed sensing, data-driven methods, dynamic mode decomposition.**

I. Introduction

In modern wireless communication systems, Orthogonal Frequency Division Multiplexing (OFDM) has been adopted due to its robustness against frequency selectivity in wireless channels. In OFDM, the total frequency bandwidth is divided into orthogonal subcarriers (SC) and the transmission time into intervals known as OFDM symbols. The fundamental building unit of the OFDM resource grid is the Resource Element (RE), which consists of one subcarrier and one OFDM symbol. Channel estimation in OFDM systems can be accomplished by allocating a portion of the REs to transmit pilots. However, pilots are known symbols and carry no new information. Hence, each RE dedicated to a pilot introduces overhead. Reducing pilot overhead is indeed essential to save more channel resources for user data. Several channel estimation and pilot overhead reduction schemes for OFDM applications have been extensively studied and investigated. Among them, Machine Learning (ML) has gained significant attention due to its success in a wide range of applications, including channel estimation. In [1], authors introduce Interpolation-ResNet, a channel estimation method based on Deep Learning (DL). The method exhibits better performance for channel estimation when compared to two other neural network methods, namely ChannelNet and ReEsNet. Another commonly used channel estimation approach involves exploiting the sparse property of the time-domain Channel Impulse Response (CIR). In this regard, Compressed Sensing (CS) theory is employed due to its ability to recover a high-dimensional sparse signal from a limited number of measurements. In addressing CS for sparse recovery, various methods have been investigated. For instance, in [2], the authors utilize the Orthogonal Matching Pursuit (OMP) algorithm to reconstruct the sparse CIR and thus perform channel estimation. Another deep learning-based method for sparse recovery, known as "Learned Approximate Message Passing" (LAMP), is proposed in [3].

This paper introduces a novel approach for channel estimation. The method is based on Dynamic Mode Decomposition (DMD), originally developed for fluid dynamics analysis in [4]. DMD, a highly versatile matrix decomposition technique, has found applications in various domains. For instance, in [5], DMD is employed to analyze data collected from mechanical equipment, enhancing bearing fault detection. In the field of wireless communication, DMD has been utilized to compress Channel State Information (CSI) and reduce feedback overhead, as demonstrated in [6]. In our investigation, we employ an extended version of DMD known as CS-DMD, where CS is applied to the resulting DMD output. CS-DMD was initially introduced for fluid dynamics in [7]. To the best of the authors' knowledge, this is the first time CS-DMD has been used for pilot-based channel estimation.

Notations: Throughout this paper, we represent matrices by uppercase boldface letters, column vectors by bold lowercase letters, scalars by italic lowercase letters and numbering by italic uppercase letters. Hadamard product and division are denoted with $\odot$ and $\oslash$, respectively. The notation † represents the Moore–Penrose pseudoinverse. $E\{.\}$ denotes the mean.

II. System and Channel Models

We consider an OFDM transmission where both the transmitter and receiver are equipped with a single antenna. The OFDM resource grid is divided into Resource Segments (RS), each consisting of K subcarriers and M OFDM symbols. Channel estimation is performed for each RS separately. Within the RS, a certain number of REs are designated as pilots for transmission. Let $K' \leq K$ and $M' \leq M$ denote the number of pilots in one OFDM symbol and the number of pilots in one subcarrier, respectively. Thus, the pilot density can be defined as $(K' \times M')/(K \times M)$. The transmitted pilot symbols are organized in a matrix $\mathbf{S} \in \mathbb{C}^{K' \times M'}$. Accounting for the channel effect, the received pilots can be expressed as:

$$\mathbf{Y} = \mathbf{P} \odot \mathbf{S} + \mathbf{Z}, \tag{1}$$

where $\mathbf{P}$ and $\mathbf{Z} \in \mathbb{C}^{K' \times M'}$ are the channel coefficients at pilots positions, and the additive noise with zero mean and variance σ_z^2 per element, respectively.

This work was funded by the German Ministry of Education and Research (BMBF) under grants 16KIS1028 (MOMENTUM) and 16KIS1012 (IRLG).

The authors are with the Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany (e-mail: haddad@ant.uni-bremen.de; bockelmann@ant.uni-bremen.de; dekorsy@ant.uni-bremen.de).

A. Channel Sparsity

The frequency domain channel coefficients within an RS are represented by the matrix $\mathbf{H} \in \mathbb{C}^{K\times M}$. Due to the radio propagation environment, it is accepted that the channel in the time domain exhibits sparsity, which aligns with the 3GPP channel model [8]. We define:

$$\mathbf{H} = \mathbf{F}\mathbf{G}, \tag{2}$$

here, $\mathbf{G} \in \mathbb{C}^{L\times M}$ denotes the channel representation in the time domain, where L signifies the number of channel response taps. $\mathbf{F} \in \mathbb{C}^{K\times L}$ corresponds to the Discrete Fourier Transform (DFT) matrix if $K = L$. In cases where $K < L$, $\mathbf{F}$ is a submatrix of the DFT matrix.

III. Channel Estimation

The process of channel estimation involves transmitting known pilot symbols at predetermined positions on the OFDM resource grid. These positions are determined by the pilot pattern, which can take various forms. In [9], different classical pilot patterns are discussed. However, these conventional pilot assignments are designed to satisfy the Nyquist sampling theorem, resulting in a significant amount of overhead.
At the receiver side, channel estimation involves mainly two steps. First, the estimation of the channel coefficients at pilots positions, denoted as $\mathbf{P}$. To estimate the matrix $\mathbf{P}$, we perform least squares estimation as:

$$\hat{\mathbf{P}} = \mathbf{Y} \oslash \mathbf{S} = \mathbf{P} + \mathbf{Z} \oslash \mathbf{S}. \tag{3}$$

Due to additive noise, the estimation of $\mathbf{P}$ may deviate from actual value. To account for the noisy estimation, we introduce the signal-to-noise ratio (SNR), defined as $\frac{E\{\|\mathbf{S}\|_2^2\}}{\sigma_z^2}$.
Second step is interpolating the channel coefficients between the pilot positions to obtain the estimated channel $\hat{\mathbf{H}}$. This paper introduces and compares various interpolation techniques.

A. Compressed Sensing

Compressed Sensing (CS), as introduced in [10], states that sparse data can be recovered even from a limited number of measurements. This is achieved by adopting a strategy of random measurements, enhancing the chance of capturing the essential signal characteristics. In the context of channel estimation, since the channel exhibits sparsity in the time domain, CS can be applied for channel estimation purpose. In order to capture the variations within the bandwidth, we randomly select K' subcarriers to carry the pilots. To reduce the needed pilots and thus minimize the overhead, we propose considering only a subset of OFDM symbols on each subcarrier, with equal distances between them, as shown in Fig. 1(a).

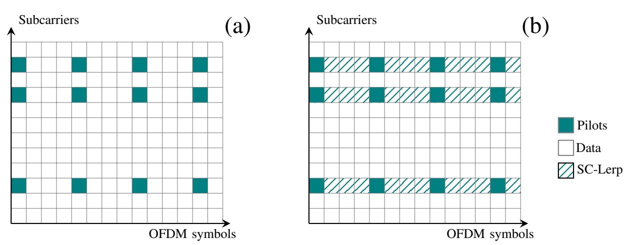

Fig. 1. Randomly selected subcarriers on OFDM resource grid to carry pilots. (a) random subcarriers. (b) SC-Lerp.

Upon receiving, the pilots are organized into matrix $\mathbf{S}$, then the matrix $\hat{\mathbf{P}}$ is estimated based on (3). Subsequently, using the estimated elements in $\hat{\mathbf{P}}$, Linear interpolation is applied to the subcarriers with pilots as in Fig.,1(b). We refer to this interpolation as SC-Lerp. As a result, we obtain the matrix $\hat{\mathbf{B}} \in \mathbb{C}^{K'\times M}$, which can be thought as projection of $\hat{\mathbf{H}}$ through a pilot selecting matrix $\mathbf{C}$:

$$\hat{\mathbf{B}} = \mathbf{C}\hat{\mathbf{H}}, \tag{4}$$

here, $\mathbf{C} \in \mathbb{B}^{K'\times K}$ comprises elements from Boolean domain $\mathbb{B} = \{0, 1\}$ and required to be incoherent with respect to $\mathbf{F}$. Each row of $\mathbf{C}$ contains only one 1, indexing the selected subcarrier of the resource grid to carry the pilots. The rest of the row is set to zeros. By substituting $\hat{\mathbf{H}}$ in (4) analogous to (2) we obtain:

$$\hat{\mathbf{B}} = \mathbf{C}\mathbf{F}\hat{\mathbf{G}} = \mathbf{\Psi}\hat{\mathbf{G}}. \tag{5}$$

The estimated channel response $\hat{\mathbf{G}}$ can be determined by solving (5). But it is an underdetermined equation. However, since $\hat{\mathbf{G}}$ is sparse, a sparse recovery algorithm can be applied to detect the sparse solution. Once $\hat{\mathbf{G}}$ is obtained, it can be substituted into (2) to find the estimated channel $\hat{\mathbf{H}}$.

Orthogonal Matching Pursuit (OMP) [2] is an effective approach for sparse recovery in compressed sensing applications. OMP performs an iterative greedy search to identify the most significant nonzero elements and their corresponding locations.

Learned Approximate Message Passing (LAMP) [3] is built by unfolding the iterations of AMP as a feedforward neural network, where the parameters in the AMP can be learned. AMP is a sparse recovery algorithm. It employs a message-passing framework, iteratively updating estimates to converge to an efficient solution.

B. Compressed Sensing - Dynamic Mode Decomposition

Dynamic Mode Decomposition (DMD) [4] is a data-driven method for decomposing dynamical systems into spatiotemporal coherent structures that exhibit oscillations at fixed frequencies which either grow or decay at fixed rates. The method relies on collecting snapshots from a dynamical system. In the context of wireless channels, the matrix $\mathbf{H}$ comprises M channel snapshots. Specifically, $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2\ ...\ \mathbf{h}_M]$, with each $\mathbf{h}_m \in \mathbb{C}^{K\times 1}$ representing the channel vector at all subcarriers over the OFDM symbol m, with $m = 1, 2, ..., M$. To use DMD, the channel vectors need to be arranged into two data matrices:

$$\begin{aligned}\mathbf{H}' &= [\mathbf{h}_1\ \mathbf{h}_2\ ...\ \mathbf{h}_{M-1}] \in \mathbb{C}^{K\times M-1},\\ \mathbf{H}'' &= [\mathbf{h}_2\ \mathbf{h}_3\ ...\ \mathbf{h}_M] \in \mathbb{C}^{K\times M-1}.\end{aligned} \tag{6}$$

DMD defines a linear approximation, expressing how $\mathbf{H}''$ evolves from $\mathbf{H}'$ as:

$$\mathbf{H}'' \approx \mathbf{A_h}\mathbf{H}', \tag{7}$$

where $\mathbf{A_h} \in \mathbb{C}^{K\times K}$ is an approximating linear operator, determined as: $\mathbf{A_h} = \mathbf{H}''\mathbf{H}'^{\dagger}$. This solution minimizes the Frobenius norm $\|\mathbf{H}''-\mathbf{A_h}\mathbf{H}'\|_F$ functioning as a linear regression of data onto the dynamics represented by $\mathbf{A_h}$. In practice, direct analysis of the matrix $\mathbf{A_h}$ may be intractable, especially when the number of subcarriers is extensive. However, the rank of $\mathbf{A_h}$ is at most $M-1$, since it is constructed as a linear combination of the $M-1$ columns of $\mathbf{H}$. Therefore, instead of solving for $\mathbf{A_h}$, DMD projects the data onto a low-rank subspace defined by at most $M-1$ Proper Orthogonal Decomposition (POD) modes. It then solves for a low-dimensional solution evolving on these POD mode coefficients. The DMD then uses this low-dimensional solution to find the leading r eigenvectors $\mathbf{\Phi_h} \in \mathbb{C}^{K\times r}$ and eigenvalues $\mathbf{\Lambda_h} \in \mathbb{C}^{r\times 1}$, which are called DMD modes and dynamics, respectively. It has been demonstrated in [4] that the snapshots are recomposed as:

$$\mathbf{h}_m \approx \mathbf{\Phi_h}\mathbf{\Lambda_h^m}. \tag{8}$$

Here r denotes the DMD rank truncation. It indicates the number of used eigendecompositions. Formula (8) implies that the higher the r, the better the resolution of recomposed $\mathbf{h}_m$. However, it is important to mention that the generated eigendecompositions are sorted in descending order of significance. This implies that a few eigendecompositions contain most of the channel power. Accordingly, it may be sufficient to take just a few modes and dynamics to ensure an adequate resolution of the recomposed $\mathbf{h}_m$.

One important feature of DMD is its capability for future state prediction. This can be achieved by extending the application of formula (8) by growing the index m beyond M, such as $m = M+1, M+2, \ldots$.

CS-DMD: According to [7], it is possible to determine the dynamical model of a high-dimensional data when limited number of measurements are available. This capability is particularly applicable when the full high-dimensional data exhibit sparsity in some basis. By exploiting the principles of compressed sensing in conjunction with dynamic mode decomposition, it becomes possible to effectively reconstruct the full data from their measurements.

In this context, we adopt the approach proposed by [7] to conduct channel estimation for an OFDM system operating in a time-varying environment. Since the matrix $\hat{\mathbf{B}}$ of the pilots' subcarriers is available. Here, CS-DMD can be employed. By arranging the columns of $\hat{\mathbf{B}}$ in two matrices $\hat{\mathbf{B}}'$ and $\hat{\mathbf{B}}''$, analogous to (6), DMD can be utilized to determine the modes $\mathbf{\Phi_b}$ and dynamics $\mathbf{\Lambda_b}$ of the operator $\mathbf{A_b}$, with $\hat{\mathbf{B}}'' \approx \mathbf{A_b}\hat{\mathbf{B}}'$. It has been mathematically demonstrated in [7] that the dynamics of the operators $\mathbf{A_b}$ and $\mathbf{A_h}$ are the identical:

$$\mathbf{\Lambda_b} = \mathbf{\Lambda_h}. \tag{9}$$

Also, the relationship between the operator's modes is that, $\mathbf{\Phi_b}$ is the projection of $\mathbf{\Phi_h}$ through $\mathbf{C}$:

$$\mathbf{\Phi_b} = \mathbf{C}\mathbf{\Phi_h}. \tag{10}$$

It is worth mentioning that, since the channel $\mathbf{H}$ is sparse in time domain, DMD can identify the most dominant modes of $\mathbf{\Phi_h}$ that capture these sparse features. Accordingly, we define:

$$\mathbf{\Phi_h} = \mathbf{F}\mathbf{\Theta_h}, \tag{11}$$

where $\mathbf{\Theta_h} \in \mathbb{C}^{L\times r}$ is the sparse time-domain representation of the modes $\mathbf{\Phi_h}$. By substituting (11) in (10), we obtain:

$$\mathbf{\Phi_b} = \mathbf{C}\mathbf{F}\mathbf{\Theta_h} = \mathbf{\Psi}\mathbf{\Theta_h}. \tag{12}$$

Since (12) is underdetermined, it can be solved to find the sparse $\mathbf{\Theta_h}$ by employing CS technique, as discussed in III-A. Then, by substituting $\mathbf{\Theta_h}$ in (11), we can calculate $\mathbf{\Phi_h}$.

Now that we have obtained the modes $\mathbf{\Phi_h}$ and dynamics $\mathbf{\Lambda_h}$ for the dynamical model of the channel evolution, we can use equation (8) to recover the entire channel coefficients and also perform channel prediction.

Complexity: DMD is built on POD, which is based on Singular Value Decomposition (SVD), making SVD the most computationally demanding part. In CS-DMD, since DMD is applied only once on $\hat{\mathbf{B}}$ for every channel estimation iteration, the SVD complexity is determined by $\mathcal{O}(K'M\min(K',M))$. The sequential execution of DMD and CS sums their complexities for the overall computational load.

C. *Interpolation-ResNet*

Interpolation-ResNet is introduced in [1] as an improved residual convolutional neural network structure. As illustrated in Fig. 2, interpolation is carried out following the application of ResNet to the channel matrix at the pilots' positions $\hat{\mathbf{P}}$.

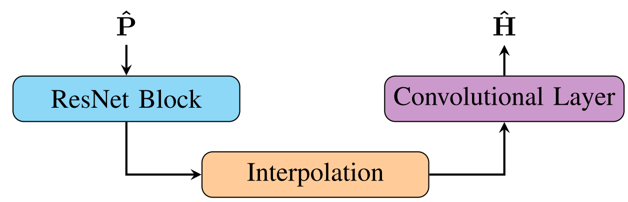

Fig. 2. Interpolation-ResNet architecture.

IV. Simulation Results

In this section, we conduct numerical simulations to evaluate the performance of CS-DMD in comparison to other approaches. CS-based channel estimation employs two different sparse recovery methods: OMP and LAMP. For applying CS-DMD, OMP is utilized, so we refer to it as OMP-DMD. The DMD rank $r = 3$. For ML-based channel estimation, we apply the interpolation-ResNet with Linear interpolation, so we refer it as Lerp-ResNet. The parameters of the ML-based methods remain the same as specified in their original publications. Additionally, we include Linear interpolation (Lerp)-based channel estimation as a baseline technique. For a fair comparison, we evenly space the pilots' subcarriers in Lerp and Lerp-ResNet, otherwise they are selected randomly. The pilot density for all methods is set to be 3.5% with $M' = 5$ and $K' = 102$. We employ Heterogenous Radio Mobile Simulator (HermesPy) [11] to generate the channel coefficients. System parameters are listed in Table I.

TABLE I
Simulation Parameters

System Parameters	Value
Channel model	COST 259 [12]
Carrier frequency	2 GHz
Receiver velocity	50 Km/h
RS size K, M	1024, 14
Channel taps in time domain L	1024
Subcarrier spacing	15 kHz
Channel sparsity S	8

To assess the performance, we utilize the normalized mean square error (NMSE), defined as $\frac{E\{\|\mathbf{H}-\hat{\mathbf{H}}\|_2^2\}}{E\{\|\mathbf{H}\|_2^2\}}$.

In Fig. 3, the simulation shows that OMP-DMD outperforms other methods. Notably, combining OMP with DMD enhances performance compared to using OMP alone. The improvement is attributed to the noise reduction achieved through the truncation of DMD modes. This explains the convergence of OMP-DMD and OMP performances at high SNR values.

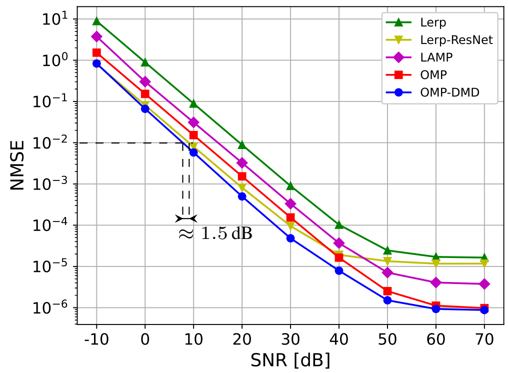

Fig. 3. Channel estimation NMSE in terms of SNR.

At very low SNR value, the performance of Lerp-ResNet is comparable to OMP-DMD. However, at very high SNR, Lerp-ResNet converge to applying Lerp. This can be explained from Fig. 2 since interpolation is applied after ResNet block. For example, to meet a 10^{-2} NMSE requirement, OMP-DMD needs less SNR by about $1.5\,$dB compared Lerp-ResNet.

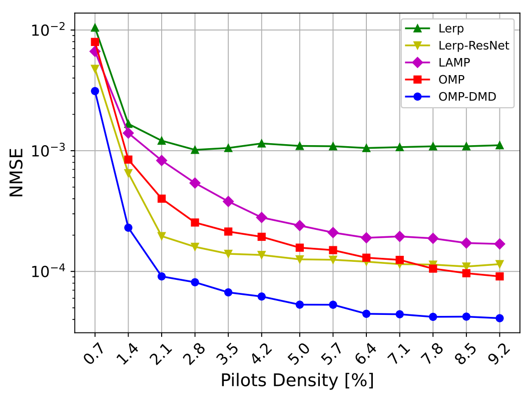

Fig. 4. Channel estimation NMSE in terms of pilots density, SNR= 30 dB.

In Fig. 4, the NMSE performance varies with pilot density. As the pilot density increases, the NMSE improves rapidly up to a pilot density of about $2.5\,\%$. Beyond this point, NMSE tends to stabilize for Lerp and Lerp-ResNet, where pilots' subcarriers are evenly spaced. While, the NMSE for LAMP, OMP and OMP-DMD show a gradual improvement, due to the lower likelihood of obtaining substantial information from added pilots, given their randomly selected positions.

In the CS-DMD, we take advantage of the prediction capability of DMD. As channel estimation is conducted on one RS, the prediction is performed on the subsequent RSs. Fig. 5 illustrates a comparison of NMSE between applying OMP and OMP-DMD. First, the RS $\#0$ is estimated, the next RSs are predicted for OMP-DMD and repeated for OMP. The plot reveals a less steep slope in the growing NMSE for the first predicted RS. As an example, if the system requires the NMSE to be less than 10^{-2}, the predicted channel for the RS $\#1$ can satisfy this requirement, while simply repeating the channel values from RS $\#0$ does not meet the desired NMSE. Thus, no additional pilots are needed for RS $\#1$, resulting in further reduction in pilot overhead and computational complexity.

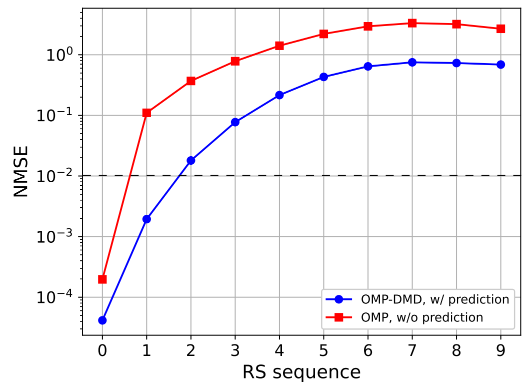

Fig. 5. Channel estimation NMSE in terms of RS sequence, SNR= 30 dB.

V. Conclusion

In this work, we proposed utilizing a method that incorporates two techniques, CS and DMD. CS exploits the inherent sparsity features, while DMD leverages temporal correlation present in time-varying channels. The results showed that applying OMP-DMD provides improved channel estimation performance compared to using OMP alone, albeit with the added complexity of DMD. We also compared OMP-DMD to ML-based channel estimation, namely Interpolation-ResNet and LAMP, and observed superior performance. Additionally, the approach OMP-DMD allows for predicting the future RSs, leading to avoid any potential pilot overhead in the predicted RS, as long as the prediction meets the system requirements.

References

[1] Dianxin Luan, and John Thompson. "Low complexity channel estimation with neural network solutions." in *Proc. WSA 2021; 25th International ITG Workshop on Smart Antennas*. VDE. Nov 2021.

[2] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[3] Mark Borgerding, Philip Schniter, and Sundeep Rangan. "AMP-inspired deep networks for sparse linear inverse problems." *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4293-4308, 2017.

[4] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

[5] Z. Cai, Z. Dang, M. Wen, Y. Lv, and H. Duan, "Application of compressed sensing based on adaptive dynamic mode decomposition in signal transmission and fault extraction of bearing signal," *Machines*, vol. 10, no. 5, p. 353, 2022.

[6] F. Haddad, C. Bockelmann, and A. Dekorsy, "A dynamical model for csi feedback in mobile mimo systems using dynamic mode decomposition," in *Proc. IEEE International Conference on Communications*, Rome, Italy, May 2023, pp. 5327–5333.

[7] S. L. Brunton, J. L. Proctor, J. H. Tu, and J. N. Kutz, "Compressed sensing and dynamic mode decomposition," *Journal of computational dynamics*, vol. 2, no. 2, pp. 165–191, 2016.

[8] "ETSI TR 138 901".[Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/16.01.00_60/tr_138901v160100p.pdf.

[9] Lee, Yong Chan, et al. "The Pilot Alignment Pattern Design in OFDM Systems." *arXiv preprint arXiv:1601.07811* (2016).

[10] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[11] "HermesPy". [Online]. Available: https://www.barkhauseninstitut.org/en/results/hermespy.

[12] "ETSI TR 125 943". [Online]. Available: https://www.etsi.org/deliver/etsi_tr/125900_125999/125943/06.00.00_60/tr_125943v060000p.pdf.