

Adaptive 2D Pilot Codebook Design for Overhead Minimization in OFDM Systems

Andrew Josephbek

Samueli Electrical and Computer Engineering

University of California, Los Angeles

Los Angeles, CA

andrewjospehbek@ucla.edu

Abstract—Channel estimation in OFDM systems is critical to maintaining system reliability, and is necessary in all systems that leverage the technique [1]. Pilot transmissions in time and frequency improve channel estimation, at the cost of increased overhead. Fixed pilot patterns exacerbate this problem when a channel is quasi static, or when channel coherence time and coherence bandwidth are over-sampled. This paper investigates the design of an $N_f \times N_t$ pilot code book that adapts to instantaneous Delay Spread (τ_{rms}) and Doppler Spread (f_D). Using a Matlab Monte Carlo analysis under Rician Fading, the performance of a Least Squares (LS) estimator enhanced by 2D Interpolation and an adaptive codebook of pilot patterns, was quantified. The Channel Estimation NMSE and system BER were used to establish a codebook performance floor. Results demonstrate that the derived adaptive codebook successfully maps channel parameters to the minimum required pilot spacing, achieving a Goodput increase of 38.49% and 0.143% over a set of delay (1-200 ns) and Doppler (1-1000 Hz) sweeps respectively, compared to a non-adaptive baseline.

I. INTRODUCTION

High speed data transmission in modern wireless communications systems, like 5G LTE [1], rely heavily on Orthogonal Frequency Division Multiplexing (OFDM) to combat the effects of a frequency and time selective channel. Maintaining reliability depends critically on accurate Channel Frequency Response estimation, over multiple successive channel realizations. This estimation requires inserting known pilot symbols into the data stream, which consumes valuable bandwidth and signal power, creating an inherent tradeoff between link reliability and throughput.

Channel distortion is governed by two primary phenomena. Delay Spread (τ_{rms}) causes frequency selective fading and limits Coherence Bandwidth (B_c), the bandwidth along which the channel may be approximated as frequency flat. High delay spread therefore requires dense pilot insertion in frequency for accurate channel estimation. Additionally, Maximum Doppler Spread (f_D), caused by mobility, dictates how fast the channel changes over time and limits the Coherence Time (T_c), the time for which the channel may be approximated as frequency flat, requiring frequent pilot retransmission. Fixed pilot patterns, like the scheme used in WiFi, assume a quasi static channel, and dedicate specific carriers to pilot transmission. In ergodic environments, a fixed pilot pattern must assume the worst fading case, severely limiting throughput via high pilot overhead.

This paper addresses the challenges of maximizing system efficiency by defining an adaptive 2D pilot codebook that dynamically selects the optimal pilot spacing ($N_f \times N_t$) based on instantaneous channel statistics (τ_{rms} and f_D). Monte Carlo simulation was used to quantify the performance of a Least Squares (LS) estimator, and define the performance limits for a set of preselected codebook of pilot configurations under varying Rician fading conditions. The codebook boundaries define the maximum channel sparsity that each pilot pattern can tolerate, allowing for dynamic allocation of pilot resources at low computational complexity. These boundaries are set to maintain acceptable NMSE in channel estimation, allowing for real world systems to tune the boundaries of this codebook before deployment based on BER requirements. The resulting lookup table enables the receiver to minimize pilot overhead for each successive channel realization, optimizing Goodput subject to channel estimation error constraints. The subsequent sections define the LS-based channel estimator, and derivation of the final adaptive codebook.

That said, many pilot patterns and insertion schemes have been devised to address this overhead issue. Early work into this area of study focuses on investigation of comb and block type pilot insertion patterns. As described in Colieri [3], comb pilot patterns transmit pilots continuously across time at select frequencies, while block pilot patterns transmit pilots along all frequencies at regular time intervals. The way in which these comb and block patterns are interpolated to create a complete channel response has a significant impact on system performance. The optimal solution, as described in Sandell [2] is a Linear Minimum Mean Squared Error estimator (LLMSE). Due to its computational complexity, this estimator is generally reserved for use in high-throughput systems. Lower complexity methods, like the ones described in Hsieh [4] and Colieri [3], focus on traditional interpolation approaches (i.e. linear, spline, polynomial).

That said, other less conventional approaches have also been devised. Zhao's DFT/IDFT [5] method uses dynamic low-pass filtering to reduce noise and limit ICI. Yang [6] later expanded on this method with their Windowed DFT method. Still more esoteric approaches have also been developed. Edfors' [7] Singular Value Decomposition (SVD) based approach proposes a low rank approximation of the optimal linear minimum mean squared error estimator (LMMSE).

Apart from interpolation, the problem of 2D time/frequency pilot spacing has also been explored. Simeone's [8] work into the subject proposed an approach that adapts spacing based on estimated channel error. More current work, e.g. Simko [9] and Li [10] focus on adaptive pilot patterns for next generation MIMO systems. That said, by far the most similar research to this work comes from Rao [11], who proposed a codebook of pilot patterns based on instantaneous channel delay and Doppler. Where this work differs from Rao is in its complexity. Crucially, Rao optimizes both pilot power and pilot spacing, where this work only optimizes for pilot spacing. Further, Roa's work focuses on pilot pattern optimization for Carrier Aggregation OFDM (CA-OFDM) systems, where the true spectrum is a fragmented collection of available OFDM carriers.

II. SYSTEM MODEL

The adaptive pilot codebook is designed and validated using a simulated time and frequency selective OFDM system. The model includes the physical phenomena necessary to characterize mobile communication links.

A. Transceiver Architecture

The system utilizes an OFDM frame structure in which pilot patterns may be inserted. Each frame starts as an empty time frequency grid \mathbf{X} ($N_{\text{fft}} \times N_{\text{sym}}$) with OFDM subcarriers along the rows of this matrix, and OFDM symbols along the columns.

- **Transmission Chain:** The transmitter first uses the selected codebook entry to fill the empty OFDM frame matrix \mathbf{X} , with pilot symbols. In this simulation, all pilot symbols were chosen as $\frac{1+1j}{\sqrt{2}}$, though future practical implementations may introduce pilot diversity to improve Peak to Average Power Ratio (PAPR). Carriers close to the edges of each OFDM symbol are left empty as guard bands; a step taken to ease a potential transition to practical implementation. A random bit stream is then mapped to M-QAM symbols and inserted into the spaces between pilot carriers and guard bands. The matrix is then transformed into the time domain via a column-wise Inverse Fast Fourier Transform (IFFT), and protected by a Cyclic Prefix (CP) of length N_{cp} .
- **Reception Chain:** The received signal is mapped to a time domain receive matrix ($N_{\text{fft}} + N_{\text{cp}} \times N_{\text{sym}}$), the CP is removed, and the signal is converted to a time-frequency receive matrix \mathbf{Y}_{raw} ($N_{\text{fft}} \times N_{\text{sym}}$) via a column-wise Fast Fourier Transform (FFT). From there, received pilot symbols estimate the channel response and a simple one-tap-equalizer is used to correct for channel effects through simple division. This step is explored in detail in Section III.

B. Rician Fading Channel Model

The channel environment is modeled using the Matlab **comm.RicianChannel** object to account for high-mobility scenarios where a dominant Line-of-Sight (LOS) path exists.

The channel's behavior is governed by two physical parameters.

Root Mean Square Delay Spread (τ_{rms}): quantifies the time dispersion of the channels multipath components. This dispersion interferes with itself to constructively and destructively combine at the receiver, limiting the channel's Coherence Bandwidth, B_c . High τ_{rms} requires dense pilot spacing in frequency to accurately estimate the channel response:

$$B_c \propto \frac{1}{\tau_{\text{rms}}}$$

The statistical distribution of these scattered multipath components was modeled via an exponential Power-Delay-Profile, such that τ_{rms} is defined as:

$$P(\tau) = P_0 e^{-\tau/\tau_{\text{rms}}}$$

where P_0 is a normalization factor.

Maximum Doppler Spread (f_d): quantifies the frequency dispersion of the channels multipath components. This dispersion quantifies how quickly the channel changes over time, limiting the channel's Coherence Time, T_c . High f_d requires dense pilot spacing in time to accurately estimate the channel response:

$$T_c \propto \frac{1}{f_d}$$

Jake's fading was used to model the spectral distribution of Doppler spread. These parameters form the inputs for the **comm.RicianChannel** object, along with the Rician K factor. This factor introduces a dominant Line-of-Sight (LOS) tap at $\tau = 0$ into the PDP, modeling environments where multipath interference is high, but an LOS path exists. The K factor quantifies the strength of this LOS tap.

The last component of the channel model is AWGN receiver noise, modeled using the MATLAB **awgn()** function. This noise process is Additive White Gaussian Noise (AWGN), which possesses a power spectral density that is uniform across the entire bandwidth and is scaled according to the specified Signal-to-Noise Ratio (SNR).

C. Pilot Resource Allocation and Overhead

The goal of this adaptive codebook is to minimize pilot overhead by tracking these two variables (τ_{rms} , f_d) and adjusting pilot spacing to track B_c and T_c . The codebook is defined by two design variables:

- **Frequency Spacing (N_f):** The spacing between two adjacent pilot symbols in frequency. A spacing $N_f = 2$ would signify a pattern with pilots separated by 1 data "bin" across the N_{fft} dimension.
- **Time Spacing (N_t):** The spacing between two adjacent pilot symbols in time. A spacing $N_t = 3$ would signify a pattern with pilots separated by 2 OFDM data "bins" across the N_{sym} dimension.

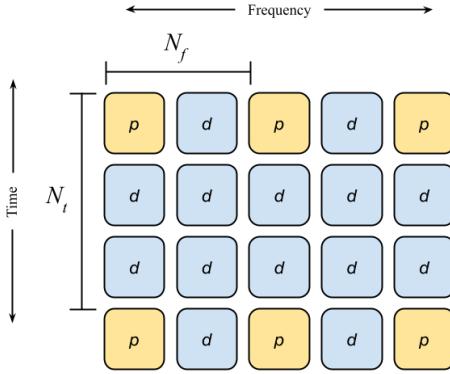


Fig. 1. Pilot time/frequency resource allocation for $N_f = 2$ and $N_t = 3$. p bins denote pilot symbols and d bins denote data symbols

III. METHODOLOGY AND CONTRIBUTION

This section details the receiver channel estimation and equalization protocol, and the procedure used to develop the codebook pattern limits. The main objective of this work is to define the minimum pilot resources required to maintain channel estimation error limits.

A. Least Squares (LS) Channel Estimation

The receiver architecture leverages a least squares estimator because of its low computational complexity, a highly desirable quality in low complexity embedded devices [12]. The estimation process begins with the $\mathbf{Y}_{\text{pilots}}$ matrix, extracted from \mathbf{Y}_{raw} based on the chosen codebook pattern. With knowledge of the original pilot pattern, $\mathbf{X}_{\text{pilots}}$, the estimated channel impulse response may be found through simple element-wise division:

$$\hat{\mathbf{H}}_{\text{raw}} = \mathbf{Y}_{\text{pilot}} \oslash \mathbf{X}_{\text{pilot}}$$

With channel estimation on pilot symbols complete, 2D linear interpolation fills in the rest of the time frequency grid, forming the full channel estimate, $\hat{\mathbf{H}}$. This estimation works under the assumption that fading may be approximated as linear between symbols. Though more complex interpolation algorithms may be adopted, this work attempts to find the limits of low complexity linear interpolation.

B. One Tap Equalizer

Having estimated the complete channel response, the raw data contained in \mathbf{Y}_{raw} is fed into the One Tap Equalizer. This equalizer implements simple element-wise division of the raw data frame with the channel response. The equalized symbol matrix, \mathbf{Y}_{eq} , is the input to the QAM demodulator:

$$\mathbf{Y}_{\text{eq}} = \mathbf{Y}_{\text{raw}} \oslash \hat{\mathbf{H}}$$

Accurate channel estimation is crucial to this step, as poor equalization will affect demodulation.

C. Quantifying Performance

The M-QAM symbol matrix \mathbf{Y} is formed from \mathbf{Y}_{eq} through nearest neighbor estimation. SER and BER are then calculated by comparing the input data in \mathbf{X} to the output data in \mathbf{Y} . With BER, goodput (G) (the number of good bits sent per data symbol transmitted) could be calculated as:

$$G = \frac{N_{\text{data}}}{N_{\text{sym}}} (1 - \text{BER}) \quad [\text{bits/symbols}]$$

where N_{data} is the total number of data symbols sent per QAM frame. While SER and BER are useful in assessing real-world system performance, they introduce significant variance across multiple simulations take significant time to converge. Additionally, BER and SER are dependent on system conditions, like modulation order and coding scheme, that make them unsuitable for comparisons across systems. To these ends, codebook performance was assessed via Normalized Mean Squared Error (NMSE) of the channel estimate $\hat{\mathbf{H}}$:

$$\text{NMSE} = \frac{\mathbb{E}[|\mathbf{H} - \hat{\mathbf{H}}|^2]}{\mathbb{E}[|\mathbf{H}|^2]}$$

D. Building the Adaptive Codebook

The adaptive codebook was derived from two Monte Carlo sweeps to establish the exact boundaries at which the LS estimator failed for a given pilot spacing and channel conditions. For each tested spacing tested, $N_{\text{trials}} = 1000$ Monte Carlo trials were averaged to ensure statistical stability.

Frequency Selectivity Sweep (N_f Optimization):

- Fixed Parameters: Maximum Doppler Spread (f_D) was set low (e.g., 1 Hz) and Time Spacing (N_t) was set dense (e.g., $N_t = 1$).
- Variable: Frequency Spacing (N_f) was swept against increasing RMS Delay Spread (τ_{rms}). This identified the maximum permissible N_f before frequency interpolation failed.

Time Selectivity Sweep (N_t Optimization):

- Fixed Parameters: RMS Delay Spread (τ_{rms}) was set low (e.g., 10 ns) and Frequency Spacing (N_f) was set dense (e.g., $N_f = 2$).
- Variable: Time Spacing (N_t) was swept against increasing Maximum Doppler Spread (f_D). This identified the maximum permissible N_t before the channel tracking failed.

In both sweep scenarios, the channel parameters were fixed at a Rician K-factor of 15 dB and an AWGN SNR of 30 dB. The SNR was set high to deliberately minimize the impact of thermal noise on the SER and ensure faster simulation convergence, allowing the experiment to isolate the dominant effects of channel fading (τ_{rms} and f_D). Additional OFDM numerology parameters (Modulation Order, Number of Subcarriers N_{fft} , Cyclic Prefix Length N_{cp} , OFDM symbols per frame N_{sym} , Dead Band Length, Sampling Frequency) are summarized in Table I.

The performance limits for each spacing scheme were established by setting an channel estimate NMSE threshold

TABLE I
FREQUENCY AND TIME SELECTIVITY SWEEP PARAMETERS

Parameter	Frequency Sel. Sweep	Time Sel. Sweep
N_f	[2, 3, 4, 5]	1
N_t	1	[30, 40, 50, 60]
τ_{rms}	1-100 ns	1 ns
f_D	1 Hz	1-1000 Hz
K-factor	15 dB	15 dB
AWGN SNR	30 dB	30 dB
Modulation	QPSK	QPSK
N_{fft}	65	65
N_{cp}	16	16
N_{sym}	16	16
Guard Bands	2 carriers/side	2 carriers/side
Sampling Freq.	20 MHz	20 MHz

(6% error in both cases), and selecting the highest spacing that met said threshold at each value of the parameter sweep. The 6% NMSE error threshold was selected by mapping the point on the parameter sweep curves where the Bit Error Rate (BER) dropped below the target reliability floor of $\sim 5 \times 10^{-3}$ to the corresponding NMSE value. This target error floor is arbitrary, but was set in an attempt to demonstrate the codebook's utility practical systems.

IV. RESULTS AND EVALUATION

This section presents the results from the two Monte Carlo sweeps described in Section III. The performance boundary for all analysis was set at an NMSE threshold of 6%, corresponding to the $BER < 5 \times 10^{-3}$.

A. Frequency Selectivity Analysis

Figure 2 shows NMSE performance plotted against τ_{rms} . As delay spread increases, the estimation performance of the $N_f = 2$ spacing remains relatively constant while the $N_f = 8$ spacing increases rapidly. As N_f approaches the coherence bandwidth at a given delay spread, its estimation performance degrades rapidly. These trends are also reflected in Figure 3, as BER grows rapidly for higher N_f spacing.

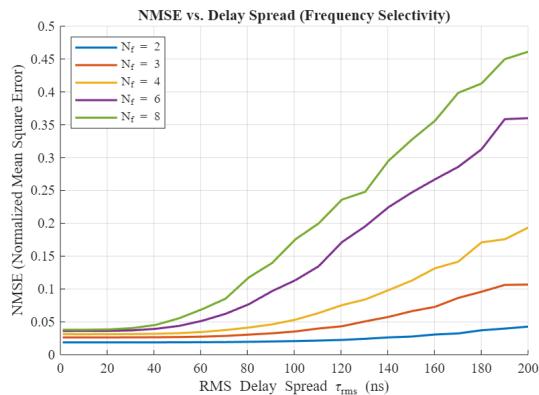


Fig. 2. Channel Estimation NMSE vs. RMS Delay Spread

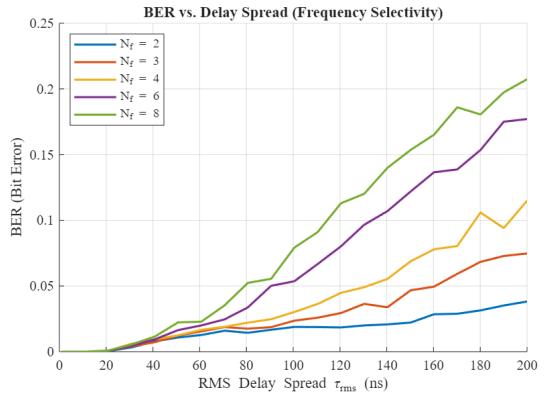


Fig. 3. BER vs. RMS Delay spread

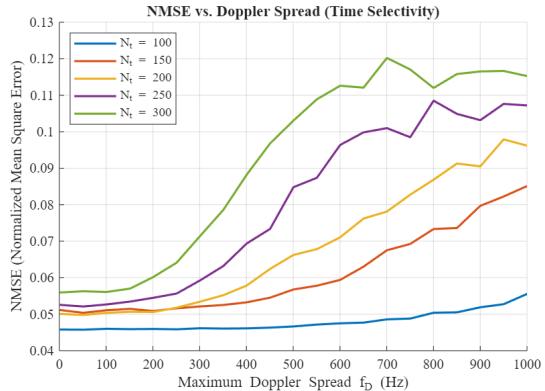


Fig. 4. Channel Estimation NMSE vs. Max Doppler

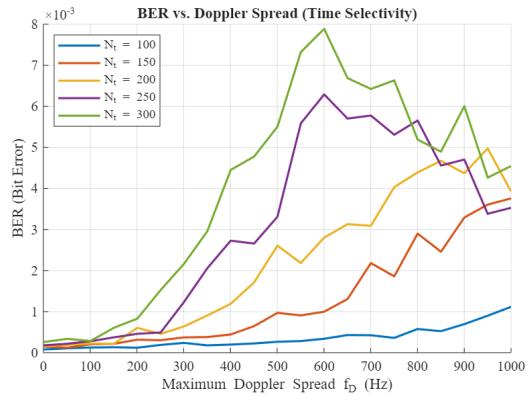


Fig. 5. BER vs. Max Doppler

B. Time Selectivity Analysis

Figure 4 shows NMSE performance plotted against f_D . As Doppler spread increases, the estimation performance of the $N_t = 100$ spacing remains relatively constant while the $N_t = 300$ spacing increases rapidly. As N_t approaches the coherence time at a given Doppler spread, its estimation performance degrades rapidly, and flattens off. That said, although the NMSE flattens off at high f_D , Figure 5 shows that BER

continues to decrease. This is may be due to symbol errors becoming temporal de-correlated at high Doppler. The channel begins to change quickly enough in time that sustained deep fading across multiple symbols is prevented, improving BER.

C. Adaptive Pilot Codebook

From the data contained in Figures 2 - 5 the procedure described in III was used to construct the codebook for using the NMSE < 0.06 threshold.

TABLE II
DELAY SPREAD CODEBOOK

N_f	τ_{rms} min	τ_{rms} max	N_t	f_D min	f_D max
8	0 ns	50 ns	300	0 Hz	150 Hz
6	50 ns	60 ns	250	150 Hz	300 Hz
4	60 ns	100 ns	200	300 Hz	400 Hz
3	100 ns	140 ns	150	400 Hz	600 Hz
2	140 ns	200 ns	100	600 Hz	1000 Hz

D. Goodput Gain and Performance Validation

Figures 6 and 7 show the goodput gains achieved from the use of the codebook described in Table II. The plots show goodput over the τ_{rms} and f_D sweeps described previously for the dynamic pilot spacing codebook and compared to fixed dense pilot spacing. The characteristic stair-step shaped curve in both plots shows the discrete switching between pilot patterns. Figure 7 is significantly more noisy since simulation hardware proved insufficient to run larger Monte Carlo simulations in a reasonable amount of time.

Overall, the codebook achieved a goodput gain of 38.49% over the τ_{rms} sweep and a 0.143% gain over the f_D sweep. Notably, gains made from adaptive pilot spacing in frequency are significantly higher than those made from adaptive pilot spacing in time. That said, if adaptive time spacing gains were compared at significantly higher Doppler, and with significantly denser pilot patterns, gains would likely become comparable.

Crucially, the codebook achieves this performance while maintaining an arbitrarily low NMSE threshold, and by extension BER threshold, low enough to sustain link reliability under FEC (Forward Error Correction).

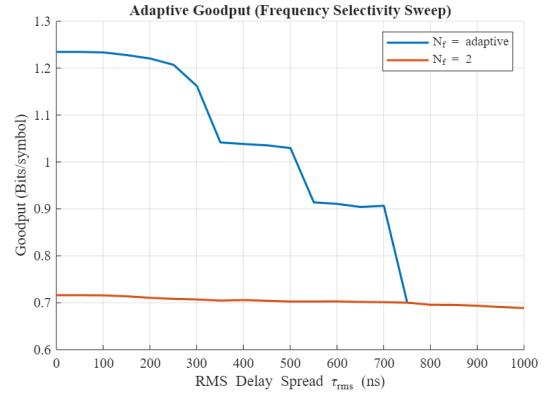


Fig. 6. Adaptive pilot codebook and fixed pilot pattern ($N_f = 2$) plotted over RMS Delay Spread sweep

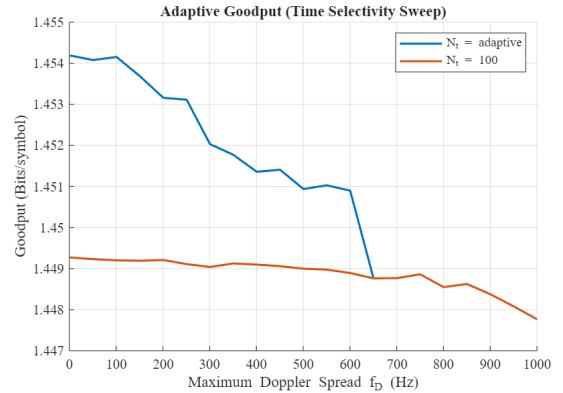


Fig. 7. Adaptive pilot codebook and fixed pilot pattern ($N_t = 100$) plotted over Max Doppler sweep

V. CONCLUSIONS

To conclude, this work successfully designed and validated an Adaptive 2D Pilot Codebook that improves spectral efficiency in a low-complexity OFDM receiver under varying Rician Fading Conditions. Through Monte Carlo analysis, the performance limits of a set of time and frequency pilot spacings were assessed using a Least Squares (LS) channel estimator. Performance limits were quantified under channel statistics τ_{rms} and f_D .

The key contribution is a functional look-up-table II that may be used as-is in channels with highly variable delay and Doppler spread, and that may be tailored to different Rician channel conditions by following the provided methodology. This lookup table achieves a channel estimation NMSE below 6%.

The resulting scheme demonstrated a significant achieved a Goodput increase of 38.49% and 0.143% over a set of delay (1-200 ns) and Doppler (1-1000 Hz) sweeps respectively, compared to a non-adaptive baseline.

Future work may explore the effects of different non-uniform pilot patterns, like the optimal diamond pattern designed analytically by Cho [14]. Additionally, while the LS estimator served to define the efficiency limits, replacing it

with the Linear Minimum Mean Square Error (LMMSE) estimator is the logical next step. A practical implementation that maintains the low computational complexity of the current implementation is described in Lowe [13]. Alternatively, the new low-complexity SVD based LMMSE approximation method in Park [15] might be used. Doing this would improve BER to usable levels at higher modulation order, making this system viable in real deployments.

VI. SUPPLEMENTAL MATERIALS

Code associated with this work (OFDM simulation and parameter sweeps) may be accessed at: [GitHub Repository Link](#).

REFERENCES

- [1] 3GPP, "Release 20," 3GPP, <https://www.3gpp.org/specifications-technologies/releases/release-20> (accessed Dec. 12, 2025).
- [2] M. Sandell and O. Edfors, "A comparative study of pilot-based channel estimators for wireless OFDM," Luleå Univ. of Technology, Luleå, Sweden, Research Rep. TULEA 1996:19, 1996. [Online]. Available: <http://www.sm.luth.se/csee/sp/research/report/sae96r.pdf>.
- [3] S. Colieri, M. Ergen, A. Puri and A. Bahai, "A study of channel estimation in OFDM systems," Proceedings IEEE 56th Vehicular Technology Conference, Vancouver, BC, Canada, 2002, pp. 894-898 vol.2, doi: 10.1109/VETECF.2002.1040729.
- [4] M.-H. Hsieh and C.-H. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," IEEE Transactions on Consumer Electronics, vol. 44, no. 1, pp. 217-225, Feb. 1998. doi: 10.1109/30.663750.
- [5] Y. Zhao and A. Huang, "A novel channel estimation method for OFDM mobile communication systems based on pilot signals and transform-domain processing," in Proc. 1997 IEEE 47th Vehicular Technology Conference. Technology in Motion, Phoenix, AZ, USA, 1997, pp. 2089-2093. doi: 10.1109/VETEC.1997.605966.
- [6] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Windowed DFT based pilot-symbol-aided channel estimation for OFDM systems in multipath fading channels," in Proc. VTC2000-Spring. 2000 IEEE 51st Vehicular Technology Conference, Tokyo, Japan, 2000, pp. 1480-1484. doi: 10.1109/VETECS.2000.851372.
- [7] O. Edfors, M. Sandell, J. . -J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," IEEE Transactions on Communications, vol. 46, no. 7, pp. 931-939, July 1998. doi: 10.1109/26.701321.
- [8] O. Simeone and U. Spagnolini, "Adaptive pilot pattern for OFDM systems," in Proc. 2004 IEEE International Conference on Communications, Paris, France, 2004, pp. 978-982. doi: 10.1109/ICC.2004.1312647.
- [9] M. Simko, P. S. R. Diniz, Q. Wang, and M. Rupp, "Adaptive Pilot-Symbol Patterns for MIMO OFDM Systems," IEEE Transactions on Wireless Communications, vol. 12, no. 9, pp. 4705-4715, Sep. 2013. doi: 10.1109/TWC.2013.081413.121998.
- [10] G. Li and G. Liao, "A Pilot-Pattern Based Algorithm for MIMO-OFDM Channel Estimation," Algorithms, vol. 10, no. 1, p. 3, 2017. doi: 10.3390/a10010003.
- [11] R. M. Rao, V. Marojevic, and J. H. Reed, "Adaptive Pilot Patterns for CA-OFDM Systems in Nonstationary Wireless Channels," IEEE Transactions on Vehicular Technology, vol. 67, no. 2, pp. 1231-1244, Feb. 2018. doi: 10.1109/TVT.2017.2751548.
- [12] K. Liu and K. Xing, "Research of MMSE and LS channel estimation in OFDM systems," in Proc. The 2nd International Conference on Information Science and Engineering, Hangzhou, China, 2010, pp. 2308-2311. doi: 10.1109/ICISE.2010.5688562.
- [13] D. Lowe and X. Huang, "Adaptive Low-Complexity MMSE Channel Estimation for OFDM," in Proc. 2006 International Symposium on Communications and Information Technologies, Bangkok, Thailand, 2006, pp. 638-643. doi: 10.1109/ISCIT.2006.339825.
- [14] J.-W. Cho and Y.-H. Lee, "Design of the optimum pilot pattern for channel estimation in OFDM systems," in Proc. IEEE Global Telecommunications Conference, 2004. GLOBECOM '04., Dallas, TX, USA, 2004, pp. 3661-3665. doi: 10.1109/GLOCOM.2004.1379052.
- [15] H.-R. Park, "A Low-Complexity Channel Estimation for OFDM Systems Based on CIR Length Adaptation," IEEE Access, vol. 10, pp. 85941-85951, 2022. doi: 10.1109/ACCESS.2022.3198962.