Preprints.org

Article

Not peer-reviewed version

Channel Estimation using Linear Regression with Bernoulli-Gaussian Noise

Prerna Chaudhary * , B.R. Manoj , Isha Chauhan , Manav Bhatnagar

Posted Date: 27 August 2025

doi: 10.20944/preprints202508.2035.v1

Keywords: Bernoulli-Gaussian noise; channel estimation; gradient descent; linear regression

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Channel Estimation using Linear Regression with Bernoulli-Gaussian Noise

Prerna Chaudhary [1,*], B.R. Manoj [2], Isha Chauhan [3] and Manav Bhatnagar [1]

[1] India Institute of Technology, Delhi
[2] India Institute of Technology, Guwahati
[3] Politecnico di Torino, Italy
* Correspondence: eez218536@iitd.ac.in

Abstract

This study introduces a novel mathematical framework for a machine learning algorithm tailored to address linear regression problems in the presence of non-Gaussian estimation noise. Our focus lies specifically on Bernoulli-Gaussian noise, a prevalent type encountered in various practical scenarios, including wireless communication channels and signal processing systems. We apply our framework within the context of wireless systems, particularly emphasizing its utility in channel estimation tasks. This letter demonstrates the efficacy of linear regression in estimating wireless channel fading coefficients under the influence of additive Bernoulli-Gaussian noise. Through comparative analysis with Gaussian noise scenarios, we underscore the indispensability of our proposed framework. Additionally, we evaluate the performance of the maximum-likelihood estimator using gradient descent, highlighting the superiority of estimators tailored to non-Gaussian noise assumptions over those relying solely on simplified Gaussian models.

Keywords: Bernoulli-Gaussian noise; channel estimation; gradient descent; linear regression

---

1. Introduction

Channel estimation plays a pivotal role in wireless communication systems, facilitating accurate signal reception amidst the challenges posed by various types of noise. Reliable estimation of the channel enables more efficient data transmission by compensating for distortions caused by the communication environment [1–4]. The task becomes particularly challenging when the channel is influenced by non-Gaussian noise models, such as the Bernoulli-Gaussian (BG) noise [5]. This type of noise introduces intermittent and sparse impulses that can severely degrade the performance of traditional channel estimation techniques. In most conventional approaches, such as those assuming Additive White Gaussian Noise (AWGN), linear regression-based estimators have been widely used due to their simplicity and effectiveness in estimating the channel coefficients. However, when the noise deviates from the Gaussian assumption, such as in scenarios with impulsive noise, these methods often fail to provide optimal performance [6,7].

The BG noise model, which combines a sparse noise process with Gaussian perturbations, is a more realistic representation for various wireless environments, including urban and industrial settings where bursty interference is common. Several real-world datasets and measurement campaigns have demonstrated the presence of impulsive noise that follows a BG distribution, particularly in vehicular and cognitive radio networks [8]. Studies like [9] also show that power line noise consists of both Gaussian background noise and impulsive components, which can be modeled accurately using a BG distribution. Moreover, it has been shown that wireless channels often exhibit heavy-tailed noise characteristics beyond Gaussian assumptions, making BG and even Cauchy-based noise models highly relevant for practical deployments [10].

This paper proposes a novel approach to channel estimation tailored for scenarios characterized by BG noise. We explore the limitations of conventional methods and propose an enhanced linear

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

regression approach that incorporates the characteristics of Bernoulli-Gaussian noise to improve the estimation accuracy [11–13]. Additionally, we compare the performance of this method with advanced techniques, such as the log-sum inequality (LSI)-based estimation, which is designed to handle sparse noise distributions. Leveraging the principles of linear regression, our method aims to accurately estimate the channel response, even in the presence of significant noise distortion [14,15]. The proposed technique begins by modeling the received signal as a linear combination of the transmitted symbols convolved with channel impulse response added with a noise, which is BG in nature. We then formulate the channel estimation problem as a linear regression task, where the objective is to learn the parameters of the channel response matrix from the received signal samples. Other regression-based machine learning and deep learning models comprising neural networks have been used in several research articles for channel estimation [12,16–20].

Furthermore, with the rapid evolution of wireless technologies towards 5G, 6G, and beyond, the role of accurate channel estimation has become even more critical. Emerging paradigms such as massive multiple-input multiple-output (MIMO), reconfigurable intelligent surfaces (RIS), integrated sensing and communication (ISAC), and millimeter-wave (mmWave) transmission demand highly reliable estimation methods to ensure low-latency and high-throughput performance [21,22]. In such advanced architectures, non-Gaussian noise, particularly impulsive BG interference, is becoming increasingly common due to spectrum sharing, dense deployments, and hybrid heterogeneous networks [23–25].

Recent studies on deep learning (DL)-based channel estimation [26,27] and unfolded iterative solvers [28] have demonstrated improved robustness to non-linear channel conditions, but these approaches often require large training datasets, suffer from high computational complexity, and lack interpretability [29]. In contrast, our proposed regression-based framework retains computational efficiency while explicitly modeling the impulsive BG noise characteristics, striking an optimal balance between robustness and tractability.

The main contributions of this paper are summarized as follows:

- We employ a robust regression framework to address the challenges posed by BG noise [30], which combines discrete and continuous characteristics. By incorporating suitable regularization techniques and optimization algorithms, our method effectively mitigates the impact of noise on the channel estimation process [31].
- We derive a closed-form approximation of the maximum-likelihood estimator and design an iterative gradient descent approach optimized for sparse-noise distributions.
- We evaluate the performance of the proposed approach through extensive simulations and compare it with existing methods under non-Gaussian noise conditions. The results demonstrate the superior accuracy and robustness of our technique, particularly in scenarios with high noise levels and sparse channel impulse responses.
- We provide insights into the applicability of the proposed framework for emerging wireless paradigms, including RIS-aided systems, massive MIMO, and ISAC-driven 6G communications. Overall, the proposed channel estimation method offers a promising solution for wireless communication systems operating in environments prone to Bernoulli-Gaussian noise, enhancing reliability and performance in practical deployment scenarios.

2. System Model

We consider an end-to-end multiple-input-single-output communication system with additive non-Gaussian noise, as shown in Figure 1. In this system, the goal is to estimate the wireless channel using linear regression techniques while dealing with additive Bernoulli-Gaussian noise. In emerging wireless systems, noise environments deviate significantly from Gaussian assumptions due to impulsive interference caused by device switching, electromagnetic emissions, and spectrum-sharing mechanisms. BG noise captures these dynamics effectively by combining a Gaussian background with

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

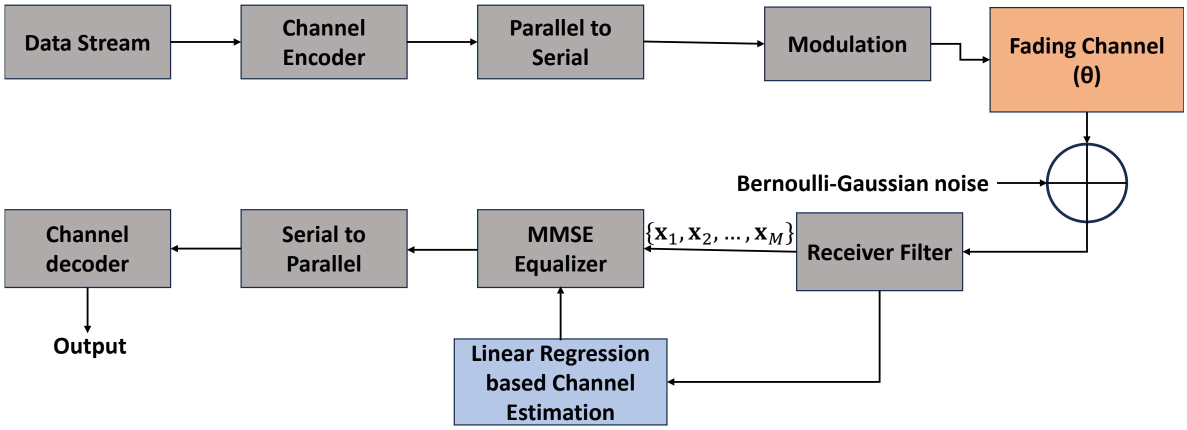

Figure 1. Block diagram for a linear regression-based channel estimation.

sparse impulses, making it particularly suitable for Power Line Communications (PLC), RIS-assisted mmWave, and dense Internet of Things (IoT) deployments [23,24].

We have considered a data stream represented by a real-valued unitary matrix given by $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_M]^T$, where $(\cdot)^T$ is the transpose operator and each column $\mathbf{x}_i^T \in \mathbb{R}^{1\times N}$ is transmitted through an N antenna system. As a result, we receive baseband signal $\mathbf{y} = [y_1, \cdots, y_M]^T$ sampled at M observations, considering each instant of time frame individually. The main focus of this work is to use the parametric models, which entail choosing a function with parameters and figuring out the best parameter values $\boldsymbol{\theta} \in \mathbb{R}^{N\times 1}$ to describe the provided data. In general, the functional dependence of $\mathbf{y}$ over $\mathbf{X}$ is defined as:

$$y_i = \mathbf{x}_i^T \boldsymbol{\theta} + \epsilon_i \implies \epsilon_i = y_i - \mathbf{x}_i^T \boldsymbol{\theta}, i = \{1, \ldots, M\}, \quad (1)$$

where ϵ_i is complex BG noise. Here, $\boldsymbol{\theta}$ indicates the unknown channel-related parameters.

We assume that the i-th sample of an independent variable, $\mathbf{x}_i$, is supplied and that a dependent variable, y_i, is the noisy observation that is produced in parallel with the input. We use signal-to-noise ratio (SNR) to express how strong a signal is in relation to noise levels and is generalized as:

$$\text{SNR} \triangleq \frac{|\mathbf{x}_i^T \boldsymbol{\theta}|^2}{\text{E}\{|\epsilon_i|^2\}}, i = 1, \ldots, M, \quad (2)$$

where $\text{E}\{\cdot\}$ is the expectation operator. In order to evaluate the accuracy of the predictive models, mean squared error (MSE) is used as a performance metric in machine learning [32]. It is averaged mean over channel realization, including number of observations and training samples:

$$\text{MSE} = \text{E}\left\{\sum_{j=1}^{N} (\theta_j - \hat{\theta}_j)^2\right\}, \quad (3)$$

where θ_j is the true value and $\hat{\theta}_j$ is its estimate.

2.1. Approximate Maximum-Likelihood Estimator using Log-Sum Inequality

Let us formulate our problem from (1) by stating the probability density function (PDF) of the Bernoulli-Gaussian noise as [33]:

$$p(\epsilon_i) = w_1 \frac{1}{\sqrt{2\pi N_{o1}}} \exp\left(-\frac{{\epsilon_i}^2}{2N_{o1}}\right) + w_2 \frac{1}{\sqrt{2\pi N_{o2}}} \exp\left(-\frac{{\epsilon_i}^2}{2N_{o2}}\right). \quad (4)$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

Here, $w_1 + w_2 = 1$, w_i, $i \in \{1,2\}$ is the hyperparameter that gives the prior probability of the noise originating with variance N_{oi}. Symbolizing Bernoulli-Gaussian distribution by BG and substituting (1) in (4), we get:

$$p(\mathbf{y}|\mathbf{X},\boldsymbol{\theta}) = \text{BG}(\mathbf{y}|\mathbf{X}\boldsymbol{\theta}, N_{o1}, N_{o2}, w_1, w_2), \tag{5}$$

where $\text{BG}(\mathbf{y}|\mathbf{X}\boldsymbol{\theta}, N_{o1}, N_{o2}, w_1, w_2)$ can be expressed by replacing ϵ_i by $y_i - \mathbf{x}_i^T\boldsymbol{\theta}$ in (4). Let us now move to parameter estimation in the following manner:

$$\begin{aligned} p(\mathbf{y}|\mathbf{X},\boldsymbol{\theta}) &= p(y_1, y_2, \ldots\ldots.., y_M|\mathbf{x}_1, \mathbf{x}_2, \ldots\ldots...\mathbf{x}_M, \boldsymbol{\theta}), \\ &= \prod_{i=1}^{M} \text{BG}(y_i|\mathbf{x}_i^T\boldsymbol{\theta}, N_{o1}, N_{o2}, w_1, w_2). \end{aligned} \tag{6}$$

The goal of maximum-likelihood analysis is to maximize the likelihood function of the received data samples, considering the channel coefficients. The maximum-likelihood estimator (MLE) can be defined using the following equation:

$$\boldsymbol{\theta}_{\text{ML}} = \arg\max_{\boldsymbol{\theta}} \; p(\mathbf{y}|\mathbf{X},\boldsymbol{\theta}) \,. \tag{7}$$

It is evident that maximizing the log-likelihood function is equivalent to minimizing the negative log-likelihood function. In order to minimize the obtained negative log-likelihood function, we need to use the gradient descent method. Hence we take negative log-likelihood of the PDF as follows:

$$\begin{aligned} -\log p(\mathbf{y}|\mathbf{x}_i,\boldsymbol{\theta}) &= -\log \prod_{i=1}^{M} p(y_i|\mathbf{x}_i,\boldsymbol{\theta}) = -\sum_{i=1}^{M} \log p(y_i|\mathbf{x}_i,\boldsymbol{\theta}), \\ &= -\sum_{i=1}^{M} \log\left(\frac{w_1}{\sqrt{2\pi N_{o1}}} \exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o1}} \right) \right. \\ &\quad \left. + \frac{w_2}{\sqrt{2\pi N_{o2}}} \exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o2}} \right) \right), \\ &= -\log\left(\sum_{m=1}^{2} \frac{w_m}{\sqrt{2\pi N_{om}}} \exp\left(\frac{-{\epsilon_i}^2}{2N_{om}} \right) \right), \\ &= \log\left(\frac{1}{\sum_{m=1}^{2} \frac{w_m}{\sqrt{2\pi N_{om}}} \exp\left(\frac{-{\epsilon_i}^2}{2N_{om}} \right)} \right). \end{aligned} \tag{8}$$

Using the LSI: $\sum a_m \log\frac{a_m}{b_m} \geq a\log\frac{a}{b}$, where $a = \sum a_m, b = \sum b_m$, we have:

$$\sum a_m \sum \log\frac{a_m}{b_m} \geq \sum a_m \log\frac{\sum a_m}{\sum b_m}, \text{where} \sum a_m = 1\,. \tag{9}$$

Considering the equi-probability of occurrence of noise with variance N_{om}, we take the value of $a_m = 0.5$ for $m \in \{1,2\}$, as:

$$\sum b_m = \sum_{m=1}^{2} w_m \frac{1}{\sqrt{2\pi N_{om}}} \exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{om}} \right). \tag{10}$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

Substituting (10) into (9), we get:

$$\sum_{m=1}^{2} 0.5 \log\left(\frac{0.5}{\frac{w_m}{\sqrt{2\pi N_{om}}}} \exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{om}}\right)\right) \geq \log\left(\frac{1}{\sum_{m=1}^{2} w_m \frac{1}{\sqrt{2\pi N_{om}}} \exp\left(\frac{-(y_i-\mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{om}}\right)}\right). \tag{11}$$

We observe that the right-hand side of (11) matches that of (8). From LSI, we can write (8) as:

$$\begin{aligned} -\log p(\mathbf{y}|\mathbf{x}_i, \boldsymbol{\theta}) &\leq \sum_{m=1}^{2} 0.5 \log\left(\frac{0.5}{\frac{w_m}{\sqrt{2\pi N_{om}}} \exp\left(\frac{-(y_i-\mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{om}}\right)}\right), \\ &= 0.5 \log\left(\frac{0.5}{\frac{w_1}{\sqrt{2\pi N_{o1}}} \exp\left(\frac{-(y_i-\mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o1}}\right)}\right), \\ &\quad + 0.5 \log\left(\frac{0.5}{\frac{w_2}{\sqrt{2\pi N_{o2}}} \exp\left(-\frac{(y_i-\mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o2}}\right)}\right) \\ &= \log 0.5 - 0.5\left(\log\frac{w_1}{\sqrt{2\pi N_{o1}}} + \log\frac{w_2}{\sqrt{2\pi N_{o2}}}\right) \\ &\quad + (y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2\left(\frac{1}{2N_{o1}} + \frac{1}{2N_{o2}}\right). \end{aligned} \tag{12}$$

Further, (12) can be written as:

$$\begin{aligned} -\log p(\mathbf{y}|\mathbf{x}_i, \boldsymbol{\theta}) &\leq \log 0.5 - 0.5 \log\left(\frac{w_1 w_2}{2\pi\sqrt{N_{o1}N_{o2}}}\right), \\ &\quad + (y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2\left(\frac{1}{2N_{o1}} + \frac{1}{2N_{o2}}\right). \end{aligned} \tag{13}$$

In terms of square distance between $\mathbf{y}$ and $\mathbf{X}\boldsymbol{\theta}$, we can write (13) as:

$$-\log p(\mathbf{y}|\mathbf{x}_i, \boldsymbol{\theta}) \text{ or } L(\boldsymbol{\theta}) = \left(\frac{1}{2N_{o1}} + \frac{1}{2N_{o2}}\right)(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2 + k, \tag{14}$$

$$L(\boldsymbol{\theta}) = \left(\frac{1}{2N_{o1}} + \frac{1}{2N_{o2}}\right)||\mathbf{y} - \mathbf{X}\boldsymbol{\theta}||^2 + k, \tag{15}$$

where k is a relative constant. Computing the partial derivative of $L(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}^T$, we get:

$$\begin{aligned} \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} &= \frac{\partial\left[\left(\frac{1}{2N_{o1}} + \frac{1}{2N_{o2}}\right)(\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\theta})\right]}{\partial \boldsymbol{\theta}^T}, \\ &= \left(\frac{1}{2N_{o1}} + \frac{1}{2N_{o2}}\right)(-\mathbf{X}^T\mathbf{y} + \mathbf{X}^T\mathbf{X}\boldsymbol{\theta}). \end{aligned} \tag{16}$$

The expression for $\boldsymbol{\theta}$ derived under LSI can be further simplified in sub-optimal with closed-form from (16) as:

$$\boldsymbol{\theta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}. \tag{17}$$

We aim to iteratively refine the estimate $\boldsymbol{\theta}$ by utilizing the gradient of the least squares error (LSE) between the received signal and the estimated channel response until the error is minimized. Conse-

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

quently, we iterate the suggested iterative gradient descent algorithm until the computed error falls below a predetermined threshold (α). The variable γ denotes either the step size or the learning rate. Additionally, a residual threshold (res) indicates the convergence point for the estimated channel coefficients, determining when to stop the further iteration.

Algorithm 1 Iterative Gradient Descent Algorithm

Require:

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{M} \frac{\mathbf{x}_i(\mathbf{x}_i^T\boldsymbol{\theta} - y_i)}{\left(\sqrt{N_{o2}}w_1 e^{\frac{(\mathbf{x}_i^T\boldsymbol{\theta}-y_i)^2}{2N_{o2}}} + \sqrt{N_{o1}}w_2 e^{\frac{(\mathbf{x}_i^T\boldsymbol{\theta}-y_i)^2}{2N_{o1}}}\right)} \times \frac{N_{o2}^{3/2}w_1 e^{\frac{(\mathbf{x}_i^T\boldsymbol{\theta}-y_i)^2}{2N_{o2}}} + N_{o1}^{3/2}w_2 e^{\frac{(\mathbf{x}_i^T\boldsymbol{\theta}-y_i)^2}{2N_{o1}}}}{N_{o1}N_{o2}}$$

▷ Channel to be estimated

1: Initialize $\boldsymbol{\theta} = [1, 1, \ldots, 1]$
2: Calculate initial error $= \|\mathbf{y} - \mathbf{X}^T\boldsymbol{\theta}\|^2$
3: $k = 0$
4: **repeat**
5: $\boldsymbol{\theta}^k = \boldsymbol{\theta}^{k-1} + \gamma \sum_{i=1}^{M} \log\left(\frac{w_1}{\sqrt{2\pi N_{o1}}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta}^{k-1})^2}{2N_{o1}}\right) + \frac{w_2}{\sqrt{2\pi N_{o2}}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta}^{k-1})^2}{2N_{o2}}\right)\right)$
▷ Step Updation
6: Calculate error $= \|\mathbf{y} - \mathbf{X}^T\boldsymbol{\theta}^k\|^2$
7: **until** abs$\left(\|\mathbf{y} - \mathbf{X}^T\boldsymbol{\theta}^{k-1}\|^2 - \|\mathbf{y} - \mathbf{X}^T\boldsymbol{\theta}^k\|^2\right) < \alpha$
▷ Convergence

2.2. *Proposed Maximum-Likelihood Estimator*

To perform channel estimation using the gradient descent method and linear regression, we first need to define the problem and establish the objective function. Given the PDF for the BG noise model in (4), we denote the received signal y_i and transmitted signal $\mathbf{x}_i$ by the i-th antenna.

Our aim is to perform channel estimation for BG noise (impulsive) by substituting the PDF $p(\epsilon_i)$ in (4) into the negative log-likelihood function, we get:

$$J(\boldsymbol{\theta}) = \sum_{i=1}^{M} \log\left(\frac{w_1}{\sqrt{2\pi N_{o1}}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o1}}\right) + \frac{w_2}{\sqrt{2\pi N_{o2}}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o2}}\right)\right). \tag{18}$$

To find the derivative of $J(\boldsymbol{\theta})$ in (18) with respect to $\boldsymbol{\theta}$, we need to differentiate each term with respect to $\boldsymbol{\theta}$. Since (18) involves a logarithm of a sum, we will apply the chain rule, as given by (19) on the top of this page to obtain $\nabla J(\hat{\boldsymbol{\theta}})$.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

$$\begin{aligned}\nabla J(\boldsymbol{\theta}) = \frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= -\sum_{i=1}^{M} \frac{1}{\frac{w_1}{\sqrt{2\pi N_{o1}}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o1}}\right) + \frac{w_2}{\sqrt{2\pi N_{o2}}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o2}}\right)} \\ &\times \left(\frac{w_1}{\sqrt{2\pi N_{o1}}}\frac{\partial}{\partial \boldsymbol{\theta}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o1}}\right) + \frac{w_2}{\sqrt{2\pi N_{o2}}}\frac{\partial}{\partial \boldsymbol{\theta}}\exp\left(\frac{-(y_i - \mathbf{x}_i^T\boldsymbol{\theta})^2}{2N_{o2}}\right)\right), \\ &= \sum_{i=1}^{M} \frac{1}{N_{o1}N_{o2}\left(\sqrt{N_{o2}}w_1 e^{\frac{\left(\mathbf{x}_i^T\boldsymbol{\theta}-y_i\right)^2}{2N_{o2}}} + \sqrt{N_{o1}}w_2 e^{\frac{\left(\mathbf{x}_i^T\boldsymbol{\theta}-y_i\right)^2}{2N_{o1}}}\right)} \\ &\left(\mathbf{x}_i\left(\mathbf{x}_i^T\boldsymbol{\theta} - y_i\right)\left(N_{o2}^{\frac{3}{2}}w_1 e^{\frac{\left(\mathbf{x}_i^T\boldsymbol{\theta}-y_i\right)^2}{2N_{o2}}} + N_{o1}^{\frac{3}{2}}w_2 e^{\frac{\left(\mathbf{x}_i^T\boldsymbol{\theta}-y_i\right)^2}{2N_{o1}}}\right)\right). \end{aligned} \tag{19}$$

Further, we can update the parameter vector $\boldsymbol{\theta}$ iteratively using the gradient descent algorithm:

$$\boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^{k} - \gamma(\nabla J(\boldsymbol{\theta}^{k})), \tag{20}$$

where $\boldsymbol{\theta}^k$ denotes the parameter vector at the k-th iteration, γ is the learning rate, and $\nabla J(\boldsymbol{\theta}^k)$ is the gradient of $J(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}^k$. Similar to Algorithm 1 we use the gradient descent method to update $\boldsymbol{\theta}$ until the algorithm converges.

3. Methodology

In order to evaluate the proposed channel estimation framework under BG noise, we performed extensive Monte Carlo simulations. Each experiment consists of transmitting a sequence of pilot symbols through a wireless channel with additive BG noise. The following parameters were considered for the simulations:

Table 1. Simulation Parameters for Channel Estimation

Parameter	Symbol	Value / Range
Number of trials	M	1000 realizations
Noise variances	N_{o1}, N_{o2}	0.2, 0.6, 0.9
Step size (learning rate)	γ	10^{-4} to 10^{-1}
Residual threshold	α	10^{-4} to 10^{-6}
Signal-to-Noise Ratio (SNR)	—	−10 dB to 30 dB
Performance metric	MSE	Averaged over trials

This setup enables a fair comparison between different channel estimation methods under identical noise environments and fading conditions. The MSE between the estimated channel coefficients and the true channel parameters was used as the primary performance metric.

The LSI based approximation is employed to handle the Bernoulli-Gaussian distribution, which consists of a mixture of Gaussian components with different variances. By applying LSI, the negative log-likelihood function can be bounded and simplified, allowing for tractable optimization.

The gradient descent algorithm is then applied to iteratively minimize the resulting loss function. Compared to closed-form MLE, gradient descent provides:

- Lower computational complexity in high-dimensional settings,
- Robustness to impulsive noise samples,
- Flexibility in adapting to varying noise variances.

Recently, machine learning and DL methods have been applied to channel estimation, showing promising results under complex fading and non-linear channel conditions [26,27,29]. However, these methods typically suffer from three limitations:

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

- Requirement of large labeled training datasets,
- High computational complexity during both training and inference,
- Lack of interpretability, which restricts practical deployment.

In contrast, the proposed regression-based framework explicitly models the impulsive characteristics of BG noise while maintaining computational efficiency and interpretability. This ensures robustness in non-Gaussian noise environments without relying on data-intensive training procedures.

4. Results and Discussions

In this section, we illustrate the impact of Bernoulli-Gaussian noise in channel estimation for a single antenna wireless communication system under Rayleigh fading. Figure 2 illustrates the relationship between MSE and SNR for four different parameter settings of the gradient descent algorithm in the context of Bernoulli-Gaussian noise using (19) for $N_{o1} = 0.6$ and $N_{o2} = 0.2$ (as in Algorithm 1). As the SNR increases, the MSE decreases, indicating improved estimation accuracy at higher SNR levels. The curve corresponding to $\gamma = 5 \times 10^{-4}$ and $res = 10^{-4}$ demonstrates a relatively unstable and high MSE across the entire SNR range, suggesting this parameter setting fails to provide a good balance between convergence speed and stability. The curve for $\gamma = 10^{-4}$ and $res = 10^{-5}$ suggests moderate performance. The parameter setting $\gamma = 10^{-3}$ and $res = 10^{-5}$ indicates high accuracy but potentially slower convergence. The parameter setting $\gamma = 0.1$ and $res = 10^{-6}$ results in significantly lower MSE values. This indicates that a smaller step size can lead to poor estimation accuracy and instability. Thus, Figure 2 highlights the importance of parameter tuning to achieve optimal performance, demonstrating that smaller step sizes and residual thresholds generally yield more accurate and stable estimations.

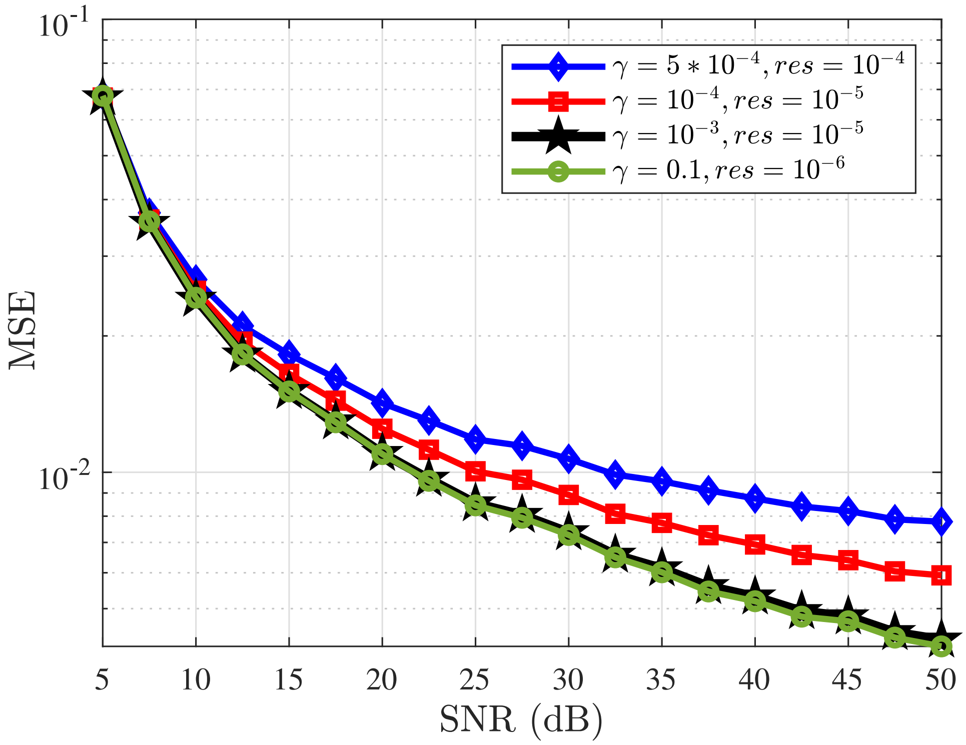

Figure 2. MSE versus SNR curves for the conventional linear regression-based channel estimation with varying values of step size and residual threshold over Bernoulli-Gaussian Noise.

Figure 3 illustrates the performance of MSE versus SNR for Gaussian noise and Bernoulli-Gaussian noise with two different set of variances. In order to use our conventional estimator given in (19) for the Gaussian noise we have ignored the Bernoulli term by setting its probability of impulses to

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

zero. The curve corresponding to a variance of $N_0 = 0.9$ consistently shows a higher MSE compared to the curve with a variance of $N_0 = 0.6$ across the entire SNR range. This indicates that Gaussian noise with a smaller variance yields better (lower MSE) at higher SNR values. The MSE for Gaussian noise follows a predictable pattern. Similarly, the MSE for Bernoulli-Gaussian noise decreases with increasing SNR but at a rate different from that of Gaussian noise. The curves for Bernoulli-Gaussian noise with a variance of $N_{o1} = 0.9$ and $N_{o2} = 0.5$ show a higher MSE than that with a variance of $N_{o1} = 0.6$ and $N_{o2} = 0.2$ at higher SNR values, indicating better performance for higher variance. This figure highlights the impact of Bernoulli-Gaussian noise on MSE performance, showing how the Bernoulli-distributed impulses affect the overall error compared to Gaussian noise.

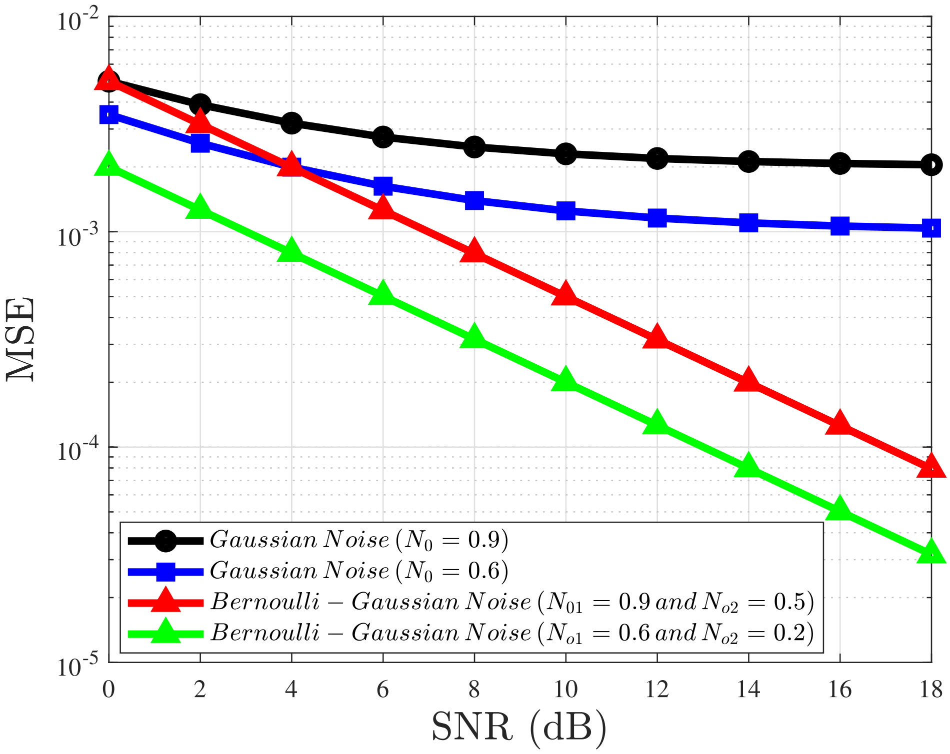

Figure 3. Comparison of MSE versus SNR for Gaussian Noise and Bernoulli-Gaussian Noise for different variances.

Figure 4 illustrates the performance of two channel estimation methods, approximate MLE using LSI and conventional MLE using iterative gradient descent method by seeking derivatives from (16) and (19) for Bernoulli-Gaussian noise, for two set of noise variances. The MSE is plotted against SNR for each estimation technique. For both variance sets $(N_{o1} = 0.9, N_{o2} = 0.5)$ and $(N_{o1} = 0.6, N_{o2} = 0.2)$, conventional MLE consistently outperforms the LSI method across the entire range of SNR values. Notably, the MSE for the conventional MLE decreases more rapidly with increasing SNR, particularly for negative SNR values, demonstrating better robustness under low SNR conditions. This aligns with the expectation that the conventional MLE method offers a more reliable channel estimation across all data points. In contrast, the approximate MLE using LSI curves shows a higher MSE at low SNR values, indicating a slower convergence rate as the noise becomes more pronounced. However, with the SNR $>$ 10 dB, the performance of both estimators begins to converge though the proposed MLE still maintains a lower MSE. This suggests that while LSI methods might be suitable in higher SNR regimes, they are less effective under noisier conditions, as evidenced by the flatter decay in MSE for lower SNR values. The results indicate that the conventional MLE can better adapt to varying noise conditions, making it a preferred choice for practical communication systems where maintaining accurate channel estimates across a wide range of SNR values is crucial.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

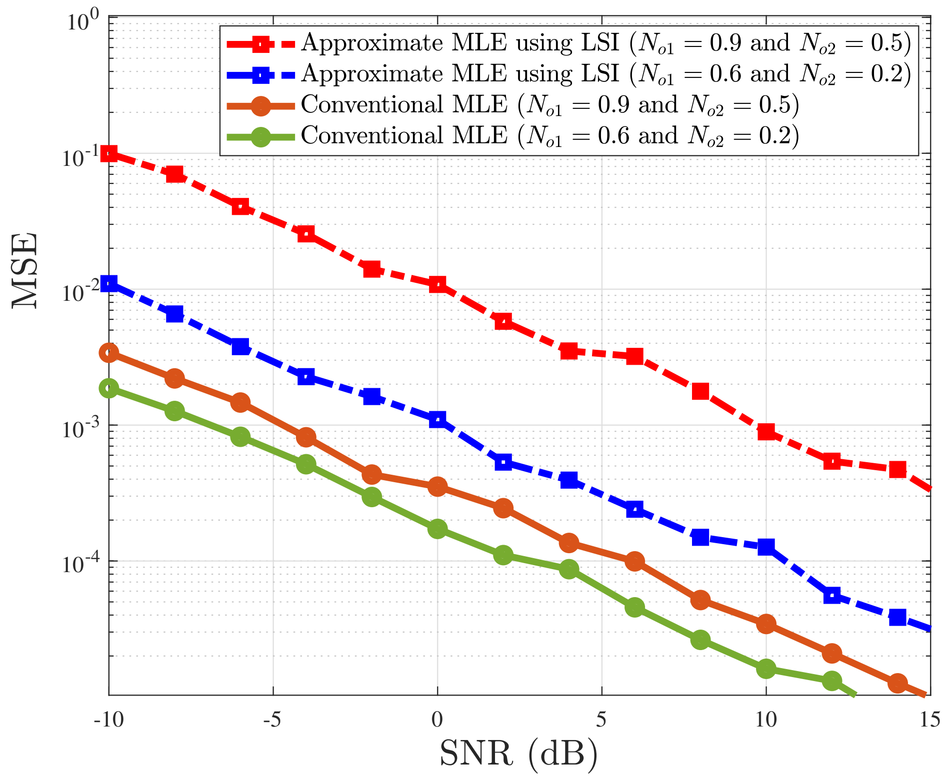

Figure 4. Comparison of MSE versus SNR by using iterative gradient descent method by seeking derivatives from (16) and (19) for Bernoulli-Gaussian noise as a function of variances.

The computational complexity of the proposed estimator is significantly lower than that of conventional MLE approaches. The closed-form MLE requires matrix inversion, which has complexity $\mathcal{O}(N^3)$, where N is the number of channel coefficients. In contrast, the proposed gradient descent estimator has complexity $\mathcal{O}(kN^2)$ per iteration, where k is the number of iterations required for convergence. This makes the approach suitable for practical real-time implementations.

5. Conclusions

This work presented a robust channel estimation framework for wireless communication systems operating under BG noise environments. We developed an enhanced linear regression-based approach that explicitly incorporates the impulsive characteristics of BG noise and evaluated its performance against conventional Gaussian-noise-based estimators and advanced maximum-likelihood methods. Through detailed analysis and extensive simulations, we demonstrated that the proposed framework achieves superior estimation accuracy and robustness, particularly in low-SNR regimes and scenarios characterized by sparse channel impulse responses. The results confirmed that the gradient descent-based estimator, derived from the log-sum inequality formulation, provides a favorable trade-off between accuracy and computational complexity. Specifically, the proposed approach achieves significant MSE improvements compared to Gaussian-based estimators, while maintaining lower complexity than conventional closed-form MLE solutions. This makes the framework suitable for real-time and resource-constrained wireless systems.

Beyond its methodological contributions, this study highlights the broader significance of BG noise modeling in modern communication environments. In applications such as power line communications, vehicular ad-hoc networks, and dense IoT deployments, impulsive interference frequently dominates, making Gaussian assumptions unrealistic. The proposed approach offers a practical solution that can be extended to advanced wireless paradigms including massive MIMO, RIS, mmWave transmission, and ISAC architectures envisioned for 6G systems. While the current work focuses on

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

simulation-based evaluation, future research should extend this framework to experimental validation using real-world datasets. Moreover, hybrid approaches that combine machine learning models with regression-based estimation could further enhance adaptability under highly dynamic non-Gaussian environments. Another promising direction lies in extending the proposed methodology to multi-antenna and multi-user scenarios, as well as integrating it with adaptive pilot design strategies.

In summary, the proposed linear regression-based estimator provides an analytically tractable, computationally efficient, and practically relevant solution for robust channel estimation under Bernoulli-Gaussian noise. By bridging the gap between traditional statistical signal processing and emerging requirements of 5G/6G networks, this work contributes a valuable step towards reliable communication in non-Gaussian wireless environments.

Author Contributions: Conceptualization, P.C., B.M. and M.B.; methodology, P.C., I.C. and M.B.; software, P.C.; validation, I.C.; formal analysis, P.C.; investigation, P.C., B.M., I.C. and M.B.; resources, B.M. and M.B.; data curation, M.B.; writing—original draft preparation, P.C.; writing—review and editing, I.C.; visualization, B.M.; supervision, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, Z.; Liu, L.; Cui, S. Channel Estimation for Intelligent Reflecting Surface Assisted Multiuser Communications: Framework, Algorithms, and Analysis. *IEEE Trans. Wireless Commun.* **2020**, *19*, 6607–6620. https://doi.org/10.1109/TWC.2020.3004330.
2. Chen, J.; Liang, Y.C.; Cheng, H.V.; Yu, W. Channel Estimation for Reconfigurable Intelligent Surface Aided Multi-User mmWave MIMO Systems. *IEEE Trans. Wireless Commun.* **2023**, *22*, 6853–6869. https://doi.org/10.1109/TWC.2023.3246264.
3. Chatelier, B.; Magoaroul, L.L.; Redieteab, G. Efficient Deep Unfolding for SISO-OFDM Channel Estimation. In Proceedings of the ICC 2023 - IEEE International Conference on Communications, Rome, Italy, May 2023; pp. 3450–3455. https://doi.org/10.1109/ICC45041.2023.10278825.
4. Cuc, A.M.; Morgoș, F.L.; Grava, A.M.; Grava, C. Estimation of the Impulse Response of the AWGN Channel with ISI within an Iterative Equalization and Decoding System That Uses LDPC Codes. *Entropy* **2024**, *26*. https://doi.org/10.3390/e26090720.
5. Yu, Y.; Yang, L.; Shen, Y. Bernoulli–Gaussian Model with Model Parameter Estimation. *Journal of Surveying Engineering* **2024**, *150*. https://doi.org/https://doi.org/10.1061/JSUED2.SUENG-1510.
6. Wilson, A.M.; Panigrahi, T. Robust SIMO Channel Estimation Under Impulsive Noise with "Fair" Cost Function. In Proceedings of the Recent Trends in Intelligent Systems and Next Generation Wireless Communication (IIWCS 2024), 2025, Vol. 1329, *Lecture Notes in Networks and Systems*, pp. 265–275. https://doi.org/10.1007/978-981-96-4741-5_23.
7. Zhang, M.; Zhou, X.; Wang, C. A Novel Noise Suppression Channel Estimation Method Based on Adaptive Weighted Averaging for OFDM Systems. *Symmetry* **2019**, *11*. https://doi.org/10.3390/sym11080997.
8. Zhao, X.; Li, F. Sparse Bayesian Compressed Spectrum Sensing Under Gaussian Mixture Noise. *IEEE Transactions on Vehicular Technology* **2018**, *67*, 6087–6097. https://doi.org/10.1109/TVT.2018.2810283.
9. Mohan, V.; Mathur, A. Pulse Jamming in PLC Over Log-Normal Channel Gain With Bernoulli-Gaussian Additive Noise. *IEEE Communications Letters* **2023**, *27*, 2603–2607. https://doi.org/10.1109/LCOMM.2023.3301063.
10. Gülgün, Z.; Larsson, E.G. Massive MIMO with Cauchy Noise: Channel Estimation, Achievable Rate and Data Decoding. *arXiv preprint* **2023**. Available: arXiv:2307.02724.
11. Mathur, A.; Bhatnagar, M.R.; Panigrahi, B.K. Maximum likelihood decoding of QPSK signal in power line communications over Nakagami-m additive noise. In Proceedings of the Proc. IEEE International Symposium on Power Line Communications and Its Applications (ISPLC), Austin, TX, USA, Mar. 2015; pp. 7–12. https://doi.org/10.1109/ISPLC.2015.7147581.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

12. Chaudhary, P.; Chauhan, I.; Manoj, B.R.; Bhatnagar, M.R. Linear Regression-Based Channel Estimation for Non-Gaussian Noise. In Proceedings of the Proc. IEEE 99th Vehicular Technology Conference (VTC2024-Spring), Singapore, Jun. 2024; pp. 1–6. https://doi.org/10.1109/VTC2024-Spring62846.2024.10683532.
13. Rong, J. Robust sparse Bayesian learning based on the Bernoulli–Gaussian model. *Signal Processing* **2023**. Uses Bernoulli–Gaussian to model impulsive noise.
14. Ohtsuki, T. Machine Learning in 6G Wireless Communications. *IEICE Trans. Commun.* **2023**, *E106-B*, 75–83. https://doi.org/10.1587/transcom.2022CEI0002.
15. Weththasinghe, K.; Jayawickrama, B.; He, Y. Machine Learning-Based Channel Estimation for 5G New Radio. *IEEE Wireless Commun. Lett.* **2024**, *13*, 1133–1137. https://doi.org/10.1109/LWC.2024.3362963.
16. Balevi, E.; Doshi, A.; Andrews, J.G. Massive MIMO Channel Estimation With an Untrained Deep Neural Network. *IEEE Trans. Wireless Commun.* **2020**, *19*, 2079–2090. https://doi.org/10.1109/TWC.2019.2962474.
17. Gao, J.; Zhong, C.; Li, G.Y.; Soriaga, J.B.; Behboodi, A. Deep Learning-Based Channel Estimation for Wideband Hybrid mmWave Massive MIMO. *IEEE Trans. Commun.* **2023**, *71*, 3679–3693. https://doi.org/10.1109/TCOMM.2023.3258484.
18. Kim, W.; Ahn, Y.; Kim, J.; Shim, B. Towards deep learning-aided wireless channel estimation and channel state information feedback for 6G. *Journal of Commun. and Networks* **2023**, *25*, 61–75. https://doi.org/10.23919/JCN.2022.000037.
19. Manoj, B.R.; Sadeghi, M.; Larsson, E.G. Adversarial Attacks on Deep Learning Based Power Allocation in a Massive MIMO Network. In Proceedings of the Proc. IEEE International Conference on Communications, Montreal, QC, Canada, Jun. 2021; pp. 1–6. https://doi.org/10.1109/ICC42927.2021.9500424.
20. Hu, Z. Robust channel estimation based on the maximum entropy principle. *Science China Information Sciences* **2023**. https://doi.org/10.1007/s11432-023-3822-9.
21. Wang, Z.; Liu, L.; Cui, S. Channel Estimation for Intelligent Reflecting Surface Assisted Multiuser Communications: Framework, Algorithms, and Analysis. *IEEE Transactions on Wireless Communications* **2020**, *19*, 6607–6620. https://doi.org/10.1109/TWC.2020.3004330.
22. Chen, J.; Liang, Y.C.; Cheng, H.V.; Yu, W. Channel Estimation for Reconfigurable Intelligent Surface Aided Multi-User mmWave MIMO Systems. *IEEE Transactions on Wireless Communications* **2023**, *22*, 6853–6869. https://doi.org/10.1109/TWC.2023.3246264.
23. Yu, Y.; Yang, L.; Shen, Y. Bernoulli–Gaussian Model with Model Parameter Estimation. *Journal of Surveying Engineering* **2024**, *150*. https://doi.org/10.1061/JSUED2.SUENG-1510.
24. Mohan, V.; Mathur, A. Pulse Jamming in PLC Over Log-Normal Channel Gain With Bernoulli-Gaussian Additive Noise. *IEEE Communications Letters* **2023**, *27*, 2603–2607. https://doi.org/10.1109/LCOMM.2023.3301063.
25. Park, H.; Choi, J.W. Binary Hypothesis Testing-Based Low-Complexity Beamspace Channel Estimation for mmWave Massive MIMO Systems. *arXiv preprint* **2025**. Available: arXiv:2508.01007.
26. Gao, J.; Zhong, C.; Li, G.Y.; Soriaga, J.B.; Behboodi, A. Deep Learning-Based Channel Estimation for Wideband Hybrid mmWave Massive MIMO. *IEEE Transactions on Communications* **2023**, *71*, 3679–3693. https://doi.org/10.1109/TCOMM.2023.3258484.
27. Kim, W.; Ahn, Y.; Kim, J.; Shim, B. Towards Deep Learning-Aided Wireless Channel Estimation and Channel State Information Feedback for 6G. *Journal of Communications and Networks* **2023**, *25*, 61–75. https://doi.org/10.23919/JCN.2022.000037.
28. Weththasinghe, K.; Jayawickrama, B.; He, Y. Machine Learning-Based Channel Estimation for 5G New Radio. *IEEE Wireless Communications Letters* **2024**, *13*, 1133–1137. https://doi.org/10.1109/LWC.2024.3362963.
29. Luan, D.e.a. Achieving Robust Channel Estimation Neural Networks. *arXiv preprint* **2025**. Includes SVM-based robust estimation under non-Gaussian noise.
30. Finamore, W.A.; da Silva Pinho, M.; Sharma, M.; Ribeiro, M.V. Modeling Noise as a Bernoulli-Gaussian Process. *Journal of Commun. and Information Systems* **2023**, *38*, 174–186. https://doi.org/10.14209/jcis.2023.20.
31. Le Saux, B.; Helard, M. Iterative Channel Estimation based on Linear Regression for a MIMO-OFDM system. In Proceedings of the Proc. IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, Montreal, QC, Canada, Jun. 2006; pp. 356–361. https://doi.org/10.1109/WIMOB.2006.1696362.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.2035.v1

32. Hodson, T.O.; Over, T.M.; Foks, S.S. Mean squared error, deconstructed. *Journal of Advances in Modeling Earth Systems* **2021**, *13*. https://doi.org/https://doi.org/10.1029/2021MS002681.
33. Deisenroth, M.P.; Faisal, A.A.; Ong, C.S. *Mathematics for Machine Learning*; Cambridge University Press, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.