NOMA Channel Estimation and Signal Detection using Rotational Invariant Codes and Machine Learning

Ayoob Salari, Student Member, IEEE, Mahyar Shirvanimoghaddam, Senior Member, IEEE, Muhammad Basit Shahab, Member, IEEE, Yonghui Li, Fellow, IEEE, and Sarah Johnson, Member, IEEE

Abstract—This paper studies the joint channel estimation and signal detection for the uplink power domain non-orthogonal multiple access. The proposed technique performs both detection and estimation without the need of pilot symbols by using a clustering technique. To remove the effect of channel fading, we apply rotational invariant coding to assist signal detection at receiver without sending pilots. We utilize Gaussian mixture model (GMM) to cluster received signals without supervision and optimize decision boundaries to enhance the bit error rate (BER) performance.

Index Terms—Clustering, GMM, NOMA, Rotational Invariant codes.

I. INTRODUCTION

The last decade has observed a dramatic growth of mobile services and applications thanks to the introduction of internet of things (IoT) technologies. The fifth-generation (5G) of mobile communication is developed not only to address the demand for capacity growth, but also to provide fast, reliable, and scalable connectivity for a diverse range of IoT scenarios [1]. Accordingly, massive machine-type communications (mMTC) and ultra-reliable low-latency communication (URLLC) have been proposed as two other major use cases of 5G in addition to the enhanced mobile broadband (eMBB) [2]. mMTC is expected to serve a massive number of low complexity devices in a limited spectrum; hence, it should be robust, scalable and energy efficient [3].

As the orthogonal multiple access (OMA) is unable to meet the increasing demand for wireless access resulting from the exponential growth in the number of devices and 5G mMTC services, non-orthogonal multiple access (NOMA) has been proposed to support the massive access by allowing users to share the resources in their transmission and enhance the spectral efficiency [4]. NOMA can be classified into power domain NOMA, and code domain NOMA [5]. Power domain NOMA can be easily implemented in current communication systems. Using the successive interference cancellation (SIC) technique at the receiver, the signals of different users are sequentially decoded according to their signal power. However, in code domain NOMA, distinct codes are assigned to different users, and users are detected based on the spreading codes. Therefore, code domain requires more bandwidth and has higher implementation complexity [6], [7]. In this paper, we focus on power domain NOMA.

Conventional communication systems use large block-length codes to approach the Shannon’s capacity, assuming that channel state information (CSI) is available at the transmitter. Several channel estimation techniques have been proposed to effectively obtain CSI. These techniques can be categorized into training-based, blind, and semi-blind schemes [8]. Using long pilot symbols, training-based channel estimation can estimate the channel accurately. On the other hand, at the cost of higher complexity, blind estimation uses the properties of the transmitted signal to estimate the channel without training symbol. Semi-blind estimation makes use of both blind and training-based techniques to maintain a balance between the pilot length and complexity. In many IoT scenarios, each device usually sends a small packet of data, and using long pilot signals will significantly deteriorate the system throughput. Recently the authors in [9] proposed a clustering-based joint channel estimation and signal detection scheme for NOMA. It has been shown that with only a few pilot symbols, the proposed clustering technique outperforms semi-blind estimation in terms of symbol error rate and achieves the same performance as the conventional maximum-likelihood (ML) detector with full CSI [10].

In this letter, we take a step further to apply phase rotational invariant codes to completely eliminate the necessity of sending even a few pilot symbols. We show that as long as users have a sufficient power difference, which is usually the case for power-domain NOMA, our proposed coding technique, in conjunction with the clustering-based joint channel estimation and detection technique, can achieve a remarkable performance in terms of BER without any pilot signal.

The rest of the paper is organized as follows. Section II presents the system model and provides some preliminary information. The rotational invariant code is detailed in the Section III. Section IV introduces the proposed clustering method based on Gaussian mixture models. Numerical results are provided in Section V. Finally, Section VI concludes the paper.

II. SYSTEM MODEL AND PRELIMINARIES

We consider an uplink NOMA massive IoT system, where users and the BS are equipped with single antenna each.

A. Signal and Channel Model

The $K$ frame-synchronized users transmit their information to the BS using quadrature phase shift keying (QPSK) mod-
Frame synchronicity can be achieved by frequently sending beacon signals from the BS. The channel between the $u$th user and the BS, denoted by $h_u$, is modeled by a zero-mean circular symmetric complex Gaussian random variable, i.e., $h_u \sim \mathcal{CN}(0, 1)$. We also consider block fading, that is the channel between each user and the BS remains unchanged for each transmission frame of length $N$ symbols.

The received superimposed signal vector at the BS, denoted by $y \in \mathbb{C}^{N \times 1}$, can be expressed as:

$$y = Xh + n, \quad (1)$$

where $X = [x_1, \cdots, x_K]$ represents the transmitted symbols, $x_u \in \mathbb{C}^{N \times 1}$ is the length-$N$ message transmitted from user $u \in \{1, \cdots, K\}$, $h \in \mathbb{C}^{K \times 1}$ is the channel vector for all users, $n \sim \mathcal{CN}(0, I_N)$ is the multivariate additive white Gaussian noise, $I_N$ is the $N \times N$ identity matrix. We further assume that $h_u$, $x_u$, and $n$ are statistically independent of each other. The signal-to-noise ratio (SNR) for user $u$, denoted by $\gamma_u$, is then defined as $\gamma_u = |h_u|^2$.

Each user’s symbols $x_{u,i}$ are drawn from the signal constellation $\mathcal{S}$ with cardinality $|\mathcal{S}|$. For QPSK modulation, we have $x_{u,i} \in \{\exp(-j\pi/4), \exp(j\pi/4), \exp(-j\pi/2), \exp(j\pi/2)\}$ $\forall u \in \{1, \cdots, K\}$ and $i \in \{1, \cdots, N\}$. Assuming that the users’ signals are randomly and uniformly drawn from $\mathcal{S}$, it is easy to show that the entries of $y$, denoted by $y_i$, have the following Gaussian mixture distribution $\mathcal{G}$

$$y_i \sim \frac{1}{|\mathcal{S}|^K} \sum_{s(i) \in \mathcal{S}^K} \mathcal{CN} \left( h^T s(i), \nu \right). \quad (2)$$

We also assume that the CSI is not available at the BS. Therefore, the receiver attempts to jointly estimate the channels and detect the signals. However, we assume that the BS knows the number of users, $K$, and the modulation type of each user.

**B. Rotational Invariant Code**

When the signal constellation has rotational symmetry, e.g. 16QAM, the presence of channel phase rotation combined with the absence of pilot symbols, results in the receiver having no knowledge of which constellation point has been sent. By applying channel coding, one simple technique to find the transmitted symmetry is to decode all possible variations and then based on the syndrome calculation we select the variation with minimum error. However, this method has a very high computational complexity, particularly as the number of users increases. On the other hand, if we send pilot symbols, due to short packet length, we will face the considerable loss of spectral efficiency.

**III. THE PROPOSED CLUSTERING-BASED JOINT CHANNEL ESTIMATION AND SIGNAL DETECTION TECHNIQUE USING RI CODING**

To handle the problem of phase rotation in symmetric constellations when the CSI is not available, rotational invariant (RI) codes have been introduced. Using RI codes, one can guarantee that all rotated versions of transmitted signals are valid sequences and all the rotated versions of the same signal sequence are produced by the same information input and subsequently decode back to the same information. Fig. 1 shows the encoder of a rate 1/2 90-degree RI code for QPSK modulation and Fig. 2 shows its state transition diagram. As can be seen, RI is sub-class of non-linear convolutional codes.

**A. GMM-based Clustering for the SIC receiver**

The Gaussian mixture model (GMM) is a natural choice for our clustering problem, because the noise has a Gaussian distribution, and the received signal can be characterized by a mixture of Gaussian distributions as in $\mathcal{G}$. As part of the SIC approach, the receiver first decodes the strongest user’s signal, then removes it from the overall received signal. The first step in using SIC at the BS is to divide the received signals into four clusters that represent User 1’s signals. Then we subdivided each of those clusters into four additional clusters to represent user 2’s signals, and so on. The joint channel estimation and signal detection problem boils down...
to estimating the unknown latent parameters of the assumed Gaussian mixed distribution in our observed data \([2]\).

Given that we are dealing with complex signals, we represent a two-dimensional multivariate Gaussian probability density function by \(\mathcal{N}(\mathbf{z}; \mu, \Sigma)\), where \(\mu\) and \(\Sigma\) denote the mean vector and covariance matrix, respectively. This method assumes that the data is created by a blend of Gaussian distributions. A GMM parameterizes each Gaussian distribution \(\omega_j \cdot \mathbf{G}\) and then we apply an iterative algorithm to calculate the parameters of the GMM. In the next step of the EM algorithm, we use the calculated \(\omega_j\) and \(\mu_j\) for the parameters of the EM algorithm is possible \([14]\).

We start by initializing the mean, covariance and weight \(\omega_j\) and \(\mu_j\) of each Gaussian mixture, and then we apply an iterative algorithm to calculate the parameters of the GMM. In the first iteration of the EM algorithm, we evaluate the so-called responsibility variable \(\gamma_{i,j}\) for every observation \(i\) and each cluster \(j\), and the underlying Gaussian mixture distribution is a weighted sum of the \(M\) Gaussian distributions (each representing a cluster), i.e.,

\[
p(\mathbf{z}; \mu_1, \ldots, \mu_M; \Sigma_1, \ldots, \Sigma_M) = \sum_{j=1}^{M} \omega_j \mathcal{N}(\mathbf{z}; \mu_j, \Sigma_j)
\]

where \(\sum_{j=1}^{M} \omega_j = 1\). We are interested in estimating \(\mu_j\), \(\Sigma_j\), and \(\omega_j\), \(j = 1, \ldots, M\), from the observed data. This can be done by maximizing the likelihood function \([4]\) for all received signals. To this end, we utilize the expectation maximization (EM) algorithm \([13]\), which is suitable for solving maximum likelihood problems with unobserved latent variables. Having a finite number of Gaussian mixtures, a closed-form expression for the parameters of the EM algorithm is possible \([14]\).

We start by initializing the mean, covariance and weight of each Gaussian distribution, and then we apply an iterative algorithm to calculate the parameters of the GMM. In the \(t\)-th iteration of the EM algorithm, we evaluate the so-called responsibility variable of each cluster \(j\) for every observation \(i\) defined as

\[
\hat{\gamma}_{i,j}^{(t)} = \frac{\omega_j^{(t-1)} \mathcal{N}_j(\mathbf{z}_i; \mu_j^{(t-1)}, \Sigma_j^{(t-1)})}{\sum_{k=1}^{M} \omega_k^{(t-1)} \mathcal{N}_k(\mathbf{z}_i; \mu_k^{(t-1)}, \Sigma_k^{(t-1)})}.
\]

We then assign each data point to its corresponding cluster. In the next step of the EM algorithm, we use the calculated responsibilities to update the mean, variance, and weight of each cluster as

\[
\hat{\omega}_j^{(t)} = \frac{\sum_{i=1}^{N} \hat{\gamma}_{i,j}^{(t)}}{\sum_{i=1}^{N} \sum_{k=1}^{M} \hat{\gamma}_{i,k}^{(t)}}.
\]

The EM algorithm is guaranteed to converge to a local optimum \([14]\). When the signals of the users are uniformly drawn from the same QPSK constellation, the weights of the Gaussian distributions are the same, i.e., \(\omega_j = \frac{1}{4}, j = 1, \ldots, 4\). Moreover, using the QPSK modulation and a SIC receiver, at each stage of the SIC, we need to estimate only four Gaussian distributions. This helps limit the computational complexity.

The proposed algorithm for joint channel estimation and signal detection using RI coding is summarized in Algorithm 1. Based on the modulation order \(m\), we fix the weights of each cluster as \(\omega_j = \frac{1}{m}\). Next, starting from the strongest user, we initialize the mean and covariance of GMM clustering. To do that, we run a few iterations (less than five) of the \(k\)-mean algorithm. Once we initialize the system, we continue by finding the responsibility and log-likelihood function based on \([4]\) and \([5]\), respectively. After that, we calculate the centroids of clusters and covariance matrices (lines 8 to 11 in the Algorithm 1). We continue by evaluating the phase of each cluster centroid. Considering that we are using QPSK modulation, the phase difference between any two adjacent clusters is \(\frac{\pi}{2}\). However, due to noise, the phase of each centroid might diverge from \(\frac{\pi}{2}, \frac{3\pi}{4}, \frac{5\pi}{4}, \text{ or } \frac{7\pi}{4}\). To minimize the effect of phase rotation, we average the phase difference between the centroids of each cluster and their expected values (Step 14), and update the decision boundaries according to the average phase rotation. We then use the state transition diagram (Fig. 2) and Viterbi algorithm to decode the clustered signals of the user (Step 17). On the one hand, by averaging the means of all clusters, the absolute value of channel gain is calculated. On the other hand, after decoding the signal of this user, the
Algorithm 1: Applying GMM for joint channel estimation and signal detection

Input: Received data at BS, number of Gaussian distributions $M$, modulation order $m$, and convergence threshold $\epsilon$
Output: Decoded Information.

1. Set $\omega_j = \frac{1}{N}$
2. for user $i$ to $K$
   3. Initialize $\hat{\mu}^{(0)}$ and $\Sigma^{(0)}$ by running a few iterations of K-means clustering
   4. Calculate $\hat{\gamma}_{i,j}^{(0)}$ according to (4)
   5. Calculate log-likelihood according to (8)
   6. Set $t = 1$
   7. while $t(t) - t(t-1) \geq \epsilon$
      8. Update $\hat{\mu}^{(t)}$ and $\Sigma^{(t)}$ using (6) and (7)
      9. Update $\hat{\gamma}_{i,j}^{(t)}$ according to (4)
     10. Update log-likelihood according to (8)
   11. end
   12. Return optimal $\hat{\mu}$ and $\Sigma$
   13. Calculate the phase of each cluster centroid ($\phi_i$)
   14. Calculate the mean phase rotation as $\theta = \sum_{i=1}^{M} \phi_i - M \pi$
   15. Update the QPSK decision boundaries based on the overall mean phase rotation due to noise
   16. Use the updated decision boundaries to demodulate the received symbols for this user into bits
   17. User Transition State (Fig. 2) to decode the received bits using the Viterbi algorithm
   18. Re-encode this user, modulate and multiply by estimated channel and subtract from superimposed received signal
   19. end

channel phase rotation is estimated. Finally, we re-encode this user and modulate and multiply by the estimated channel and subtract it from the superimposed received signal (Step 18) and repeat the algorithm for the next user.

IV. NUMERICAL RESULTS AND DISCUSSION

In this section, we evaluate the performance of the proposed clustering-based joint channel estimation and detection algorithm utilizing RI codes. For all simulations, we assume that each user has a packet of length 50 symbols to send to the BS. We also assume that the QPSK modulation is employed by the users.

Fig. 4 shows the bit error rate (BER) performance of the proposed GMM-based clustering approach using RI coding. As shown in this figure, for point-to-point communication, the proposed GMM clustering algorithm with RI coding performs extremely close to the ideal maximum-likelihood (ML) receiver with full CSI. Fig. 5 illustrates the proposed GMM clustering approach with RI coding in a two-user NOMA scenario with a 9dB power difference between the users. As demonstrated in this figure, the proposed GMM-clustering technique with RI coding is capable of accurately determining clusters and performing symbol detection with a BER that is very close to that of optimal ML detection with full CSI.

To further investigate the effectiveness of the proposed technique, in Fig. 6 we provide the results for the two-user NOMA scenario when both of users apply the GMM clustering while implementing different coding techniques with rate 1/2. Since the intention is to transmit without any pilot signals, we make use of syndromes for LDPC detection. When QPSK modulation is used, there are four possible constellation mappings; by calculating the syndrome for each of these mappings, the proper mapping may be determined by picking the mapping with the most syndrome checked. As illustrated in this Fig. 6 RI coding outperforms LDPC coding in terms of BER. For the strong user, the gap between the proposed RI coding technique and LDPC coding with GMM is more than 2dB at the BER of $10^{-3}$.

Thus far, we have compared our proposed technique with the optimal ML detection with full CSI. However, attaining full CSI with a limited number of symbols is not achievable in real-world settings. Fig. 7 compares the performance of our proposed approach with ML-based detection with 2 training symbols for channel estimation. As can be noticed, when ML is used to estimate the channel using two training symbols, its performance is much worse than our proposed technique of using clustering in conjunction with RI coding. For ML to have a BER performance similar to our proposed scheme,
error rate (BER). Results showed that the proposed scheme





Gaussian mixture model to cluster incoming signals without
communicate without the use of pilot signals. We employ the
we employ rotational invariant coding, which allows us to
mulation. In order to counteract the effects of channel rotation,
and signal detection for the uplink of non-orthogonal multiple
outperforms the K-mean clustering approach.

as can be seen the GMM-based clustering
approach: it estimates the appropriate covariance matrices
for each cluster. As can be seen the GMM-based clustering





RI coding for GMM clustering versus K-means clustering.

While K-means assumes that the covariance matrix for all
clusters is the same for all clusters, GMM uses a different
approach: it estimates the appropriate covariance matrices
for each cluster. As can be seen the GMM-based clustering
outperforms the K-mean clustering approach.

V. Conclusion

By combining clustering technique with rotational invariant
coding, this paper investigated the joint channel estimation
and signal detection for the uplink of non-orthogonal multiple
access without the use of any pilot symbols for channel esti-
mation. In order to counteract the effects of channel rotation,
we employ rotational invariant coding, which allows us to
communicate without the use of pilot signals. We employ the
Gaussian mixture model to cluster incoming signals without
supervision, and then optimise decision boundaries in accor-
dance with the clustering results in order to improve the bit
error rate (BER). Results showed that the proposed scheme

at least eight training symbols should be used. However, this
results in about 16% loss in throughput. Moreover, even with
eight training symbols, our proposed GMM-based approach
has better performance for the weak user.

Fig. 8 shows the BER performance of two-user NOMA with
RI coding for GMM clustering versus K-means clustering.

without any pilot symbols can achieve the same performance
as the maximum-likelihood detector that needs to obtain full
CSI to operate well.

REFERENCES

[1] H. Viswanathan and P. E. Mogensen, “White paper: Communications in
the 6G Era,” Nokia Bell Labs, 2020.
[2] J. Peisa, P. Persson, S. Parkvall, E. Dahlman, A. Grovlen, C. Hoymann,
and D. Gerstenberger, “5G New Radio Evolution,” Ericsson Technology
Review, no. 2, March 2020.
[3] N. H. Mahmood, S. Böcker, A. Munari, F. Clazzer, I. Moerman,
K. Mikhailov, O. Lopez, O.-S. Park, E. Mercier, H. Bartz et al., “White
paper on critical and massive machine type communication towards 6G,”
arXiv preprint arXiv:2004.14146, 2020.
[4] N. Iswarya and L. Jayashree, “A survey on successive interference
cancellation schemes in non-orthogonal multiple access for future radio
access,” Wireless Personal Communications, pp. 1–22, 2021.
[5] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson,
“Grant-free non-orthogonal multiple access for IoT: A survey,” IEEE
Communications Surveys & Tutorials, vol. 22, no. 3, pp. 1805–1838,
2020.
[6] F. Ghanami, G. A. Hodtani, B. Vucetic, and M. Shirvanimoghaddam,
“Performance analysis and optimization of NOMA with HARQ for short
packet communications in massive IoT,” IEEE Internet of Things
Journal, 2020.
[7] Z. Ding, R. Schober, and H. V. Poor, “Unveiling the importance of SIC
in NOMA systems—part I: State of the art and recent findings,” IEEE
Communications Letters, vol. 24, no. 11, pp. 2373–2377, 2020.
[8] F. Nadeem, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, “Non-
orthogonal HARQ for URLLC: Design and analysis,” IEEE Internet
of Things Journal, 2021.
[9] A. Salari, M. Shirvanimoghaddam, M. B. Shahab, R. Arablouei, and
S. Johnson, “Clustering-based joint channel estimation and signal detec-
tion for grant-free NOMA,” in 2020 IEEE Globeocom Workshops (GC
Wkshps), IEEE, 2020, pp. 1–6.
[10] ———, “Clustering-based joint channel estimation and signal detection
for NOMA,” 2022.
[11] S. S. Pietrobon, G. Ungerboeck, L. C. Pérez, and D. Costello, “Rotation-
ally invariant nonlinear trellis codes for two-dimensional modulation,”
IEEE Transactions on Information Theory, vol. 40, no. 6, pp. 1775–
1791, 1994.
[12] R. Singh, B. C. Pal, and R. A. Jabr, “Statistical representation of distrib-
ution system loads using Gaussian mixture model,” IEEE Transactions
on Power Systems, vol. 25, no. 1, pp. 29–37, 2009.
[13] T. Hastie, R. Tibshirani, and J. Friedman, The elements of statistical
learning: data mining, inference, and prediction. Springer Science &
Business Media, 2009.
[14] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood
from incomplete data via the EM algorithm,” Journal of the Royal
Statistical Society: Series B (Methodological), vol. 39, no. 1, pp. 1–22,
1977.