CNN-Based Automatic Modulation Classification Under Phase Imperfections

Thrassos K. Oikonomou, *Graduate Student Member, IEEE*, Nikos G. Evgenidis, *Student Member, IEEE*, Dimitrios G. Nixarlidis, Dimitrios Tyrovolas, *Graduate Student Member, IEEE*, Sotiris A. Tegos, *Senior Member, IEEE*, Panagiotis D. Diamantoulakis, *Senior Member, IEEE*, Panagiotis G. Sarigiannidis, *Member, IEEE*, and George K. Karagiannidis, *Fellow, IEEE*

***Abstract*—Dynamic spectrum allocation for diverse future applications is anticipated to be supported by sixth-generation (6G) wireless networks. Specifically, automatic modulation classification (AMC) has been highlighted as a technique to enhance spectral utilization. However, its accuracy is influenced not only by additive white Gaussian noise and channel fading but also by phase imperfections (PI) coming from unsynchronized local oscillators and imperfect channel state information (CSI), leading to degraded classification performance. To solve this problem, we propose a convolutional neural network (CNN)-based scheme that transforms the received data to improve the classification accuracy under generalized PI conditions. Moreover, we also modify the kernel dimensions of the CNN layers to further improve the performance based on the geometry of the modulated schemes after the proposed transformation is applied to the received data. Finally, through simulations, we verified the effectiveness of the method in elevating AMC accuracy, even in intense PI conditions.**

***Index Terms*—AMC, CNN, phase noise impairments, spectrum allocation, low-latency.**

I. Introduction

WIRELESS communication networks are designed to effectively manage spectrum resources to support a wide range of future applications. Among the foreseen technologies in spectrum management is automatic modulation classification (AMC), a sophisticated signal processing technique operating at the physical layer. Specifically, AMC's primary function is to automatically identify modulation schemes in wireless transmissions and is expected to play a key role in efficient spectrum utilization, dynamic spectrum allocation, and overhead reduction through pilot symbol omission [1], [2]. However, to provide reliable AMC services, it is imperative to identify and address all factors that could degrade its performance. In this regard, AMC, like any wireless communication-based application, faces a variety of imperfections, including phenomena such as fading and frequency offset, which adversely affect the transmitted symbols at the receiver side. However, although their effects on the magnitude of the symbol are widely recognized and addressed, their impact on its phase is often not considered [3]. In particular, phase imperfections (PI) appear only as time shifts in the received signals, making them more difficult to detect and address. This detection is further complicated by local oscillator (LO) inaccuracies that result in imprecise frequency and phase synchronization, while the lack of accurate channel state information (CSI) can also amplify the negative impact of PI within the system. Consequently, given the pronounced challenges of PI, the improvement of AMC architectures becomes paramount to ensure robust and reliable performance under PI conditions.

Over the last years, various AMC techniques have been investigated to improve classification accuracy under additive white Gaussian noise (AWGN) and fading conditions. Such techniques have mainly focused on higher moments methods for feature extraction or convolutional neural network (CNN) for image recognition approaches to correctly classify the constellation of the received signals [2]. Regarding the image recognition techniques, the potential of AMC as a deep learning application for the physical layer has been explored using numerous CNN-based architectures, as they excel at recognizing spatial patterns in received data [4]. However, the need for CNNs that achieve high accuracy with low time complexity is crucial, particularly in practical scenarios where computational cost is a concern due to their number of convolutional layers and their characteristics such as filters and kernel dimensions [5]. Furthermore, even if CNN-based AMC has demonstrated high accuracy in modulation classification in the presence of AWGN, the adaptation of CNN architecture to mitigate the influence of PI in the AMC process remains unexplored [6], [7]. Specifically, even in cases where fading is considered, as in [6], the effect on the phase of the received symbol is neglected, rendering the framework unsuitable for PI scenarios. Finally, while the authors of [7] illustrated that the presence of PI can significantly degrade classification

Manuscript received 31 January 2024; accepted 15 March 2024. Date of publication 19 March 2024; date of current version 10 May 2024. This work was supported by the Smart Networks and Services Joint Undertaking (SNS JU) through the European Union's Horizon Europe Research and Innovation Programme under Grant 101096456 (NANCY). The work of George K. Karagiannidis was supported by the Framework of H.F.R.I Call Basic Research financing (Horizontal Support of All Sciences) under the National Recovery and Resilience Plan Greece 2.0 funded by the European Union NextGenerationEU (H.F.R.I.) under Project 15642. The associate editor coordinating the review of this article and approving it for publication was X. Liu. *(Corresponding author: Thrassos K. Oikonomou.)*

Thrassos K. Oikonomou, Nikos G. Evgenidis, Dimitrios G. Nixarlidis, Dimitrios Tyrovolas, and Panagiotis D. Diamantoulakis are with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece (e-mail: toikonom@auth.gr; nevgenid@auth.gr; dimnixgeo@auth.gr; tyrovolas@auth.gr; padiaman@auth.gr).

Sotiris A. Tegos is with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, and also with the Department of Electrical and Computer Engineering, University of Western Macedonia, 50100 Kozani, Greece (e-mail: sotiristegos@ieee.org).

Panagiotis G. Sarigiannidis is with the Department of Electrical and Computer Engineering, University of Western Macedonia, 50100 Kozani, Greece (e-mail: psarigiannidis@uowm.gr).

George K. Karagiannidis is with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54636 Thessaloniki, Greece, and also with the Artificial Intelligence and Cyber Systems Research Center, Lebanese American University, Beirut 03797751, Lebanon (e-mail: geokarag@auth.gr).

Digital Object Identifier 10.1109/LWC.2024.3379198

accuracy, they did not tailor their framework to compensate for PI effects. Therefore, to the best of the authors' knowledge, there is currently no research addressing the appropriate modification of CNNs to meet the real-time requirements of AMC applications, while enhancing AMC performance under PI conditions.

In this letter, we present a novel low-complexity method suitable for CNN-based AMC frameworks in the presence of PI, without adding additional time complexity. Specifically, our approach involves the integration of a polar transformation of the received signals and the use of asymmetric kernels in the convolutional layers to capture characteristic modulation patterns in the polar plane. In order to characterize the effectiveness of our method, we applied it to both LeNet (LN) and AlexNet (AN), which are well-established CNN architectures in AMC, demonstrating significant improvements in classification accuracy over traditional IQ mapping, especially when the polar transformation is combined with asymmetric kernels. More specifically, our results show that our method can improve classification accuracy even in the low-SNR regime and under severe PI conditions, without increasing the time complexity of the CNN architectures. Moreover, when our method is applied to LN, its classification accuracy surpasses that of conventional AN, challenging the assumption that more complex CNN architectures inherently lead to better performance. Thus, our method not only improves the classification accuracy of CNN-based AMC under PI conditions, but also paves the way for making existing CNN-based AMC solutions more robust to PI without increasing their time complexity.

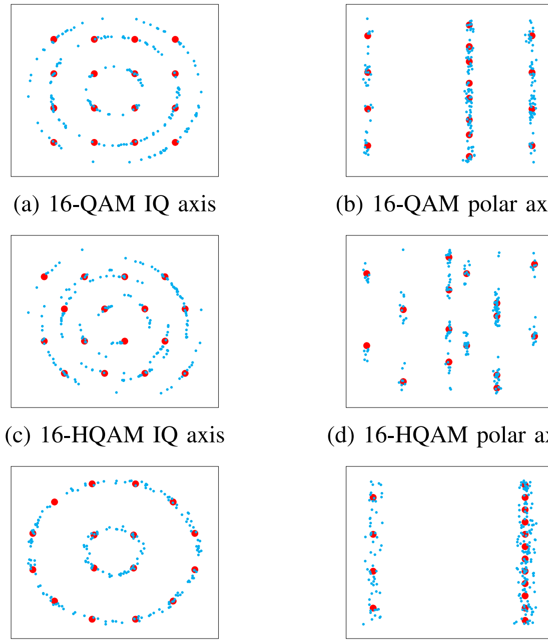
II. System Model

We consider a communication system that transmits a symbol $s \in \mathbb{C}$ belonging to a constellation of M symbols and average symbol energy E_s. Considering an AWGN channel affected by PI, the received signal r can be expressed as

$$r = |h||s|e^{j(\phi+\arg\{s\})} + n, \tag{1}$$

where $|\cdot|$ and $\arg\{\cdot\}$ denote the amplitude and phase of a complex number, respectively, $|h|^2$ denotes the channel gain including both large and small scale fading, and n represents the AWGN with zero mean and variance $\frac{N_0}{2}$. Furthermore, ϕ expresses the PI occurred from the receiver's LO and imperfect CSI, which can be generally modeled as a random variable (RV) following the Von Mises distribution with zero mean and concentration parameter $\kappa \in [0, \infty)$, which is a generalized distribution able to effectively model various PI scenarios [8]. In particular, higher values of κ correspond to more favorable PI conditions (e.g., strong line-of-sight channels, high-quality LO), while in the scenario where $\kappa = 0$ the Von Mises distribution coincides with the circular uniform distribution, indicating severe PI conditions (e.g., non-line-of-sight channels, damaged LO) [8].

Given the significant performance degradation in terms of symbol error rate (SER) under severe PI, there is an urgent need for more resilient modulation schemes to combat PI. Therefore, the evaluation of AMC architectures requires a tailored dataset that considers PI. To this end, the constellations used in this letter include quadrature amplitude modulation (QAM), hexagonal quadrature amplitude modulation (HQAM) [9], and amplitude phase shift keying (APSK) [10], which are chosen due to their widespread applications, energy efficiency, and PI resilience, respectively. It should be highlighted that the diversity among these constellations enriches the provided analysis, as each constellation exhibits unique patterns that imply different levels of accuracy in the AMC due to their different characteristics.



(a) 16-QAM IQ axis (b) 16-QAM polar axis
(c) 16-HQAM IQ axis (d) 16-HQAM polar axis
(e) 16-APSK IQ axis (f) 16-APSK polar axis

Fig. 1. Modulation schemes in IQ and polar form at average received SNR $= 30$ dB and $\kappa = 3$ with constellation points in red and received signals in blue.

III. Architecture and Training

To efficiently perform CNN-based AMC, the selected CNN models must combine high accuracy performance with low time complexity. This requires an appropriate CNN architecture that is as simple as possible without compromising performance. In this direction, we investigate the effectiveness of our proposed technique applied to models such as LN and AN, which are known for their simplicity while maintaining classification accuracy.

A. LeNet and AlexNet Architectures

Given the need for a combination of accuracy and low time complexity in AMC, LN emerges as an appropriate choice for CNN-based AMC [11]. Specifically, LN consists of two sets of convolutional layers, each with a max-pooling layer to reduce the dimension of the waveform representation and obtain more features, followed by a flattening convolutional layer, three fully connected layers, and a softmax classifier. The exact structure of the LN architecture is shown in Table I.

In contrast to LN, AN offers a more complex CNN architecture designed for higher accuracy than LN. It has five convolutional layers, with the first two followed by max-pooling layers, and the next three followed by a max-pooling layer. The output is then fed into a flatten layer and three dense fully connected layers, resulting in a softmax

TABLE I
LN PARAMETERS

Layer (type)	Output dimensions	Parameters
Conv2D	(252,252,6)	156
MaxPooling2D	(126,126,6)	-
Conv2D	(124,124,16)	880
MaxPooling2D	(62,62,16)	-
Flatten	61504	-
Dense	120	7380600
Dense	84	10164
Dense	16	1360

TABLE II
AN PARAMETERS

Layer (type)	Output dimensions	Parameters
Conv2D	(254,248,6)	168
MaxPooling2D	(127,124,6)	-
Conv2D	(125,119,16)	1744
MaxPooling2D	(62,59,16)	-
Conv2D	(60,57,16)	2320
Conv2D	(58,55,16)	2320
Conv2D	(56,53,16)	2320
MaxPooling2D	(28,26,16)	-
Flatten	11648	-
Dense	120	1397880
Dense	84	10164
Dense	16	1360

classifier, similar to LN. As shown in Table II, AN's complex structure allows for more extensive image feature extraction at the cost of greater time complexity [5]. Therefore, the comparison between LN and AN emphasizes that while LN's simpler architecture may suggest inferior performance, its lower time complexity makes it a viable option in real-time AMC scenarios under PI conditions.

B. Polar Transformation and Asymmetric Kernels

As mentioned above, our primary goal is to accurately classify constellations, especially in the presence of PI. Specifically, a set of symbols is received, processed by the CNN to classify the constellation scheme that is used, and then visualized on the IQ coordinate system as shown in Figs. 1(a), 1(c), 1(e). However, PI introduce deviations in the expected symbol phase, which complicates the constellation recognition task. To this end, based on the previously presented constellations, we exploit the transformation of the received symbols into the polar coordinate system in order to achieve a visualization of the signal that is more robust to PI [10]. As it can be seen in Figs. 1(b), 1(d), 1(f), from an imaging point of view, this transformation provides received symbols with a more consistent structure, allowing easier recognition by feature extraction techniques. Specifically, since PI affects the phase of the symbols, it acts as a rotating factor in the IQ system, while in the polar system, it acts only as a shift factor on the phase component, meaning that the transformed constellations form vertical lines, as illustrated in Figs. 1(b), 1(d), 1(f). Moreover, since the magnitude of the received symbols is expected to be on discrete levels, affected only by AWGN, and only the phase has a larger deviation from the ideal symbols, the classification task is expected to have better accuracy, due to the aforementioned line patterns that are easier to recognize.

Observing the structures of the constellations in Figs. 1(a), 1(c), 1(e) and 1(b), 1(d), 1(f), it is of paramount importance to adapt the CNN architecture to the classification task. Therefore, inspired by the constellation structures in the polar coordinate system, we introduce asymmetric kernels into the convolutional layers of CNN, to improve feature extraction along vertical and horizontal directions. Specifically, unlike symmetric kernels, which may unintentionally correlate different vertical magnitude zones as shown in Figs. 1(b), 1(d), 1(f), asymmetric kernels improve the ability to detect lines and the distances between them within the images. Furthermore, determining the optimal kernel dimensions requires fine-tuning, but given the geometric characteristics of our classified constellations, asymmetric kernels are particularly suitable in this context, suggesting that other CNN-based architectures can also benefit from our method. Finally, regarding the complexity of the proposed method, it is important to emphasize that, compared to the original CNN models on which it is applied, no additional time complexity is added. In particular, since the type and number of layers and filters remain the same, the only difference in terms of architecture is derived from the asymmetric kernel dimensions, which does not change the overall complexity. Moreover, as shown in [5], more complex CNN architectures typically increase their complexity due to the addition of extra layers and filters, thus it is important to keep the CNN structure as simple as possible to achieve real-time decisions for the AMC task.

C. Dataset, Training, and Testing

The discussed LN and AN models have been trained with the Adam optimizer for learning rate 10^{-4} and for the categorical cross-entropy loss function, as it is the most suitable for the classification task of AMC. In addition, no padding or stride was implemented for the convolutional layers of the DNNs. Regarding the activation functions, both convolutional and dense layers implement the Rectified Linear Unit (ReLU) function, while the last layer implements the Softmax activation function, as it is the most suitable for the classification task of AMC.

Regarding the utilized data,[1] this letter employed a dataset comprising 16 digital modulation schemes, with the majority of them sourced from the RadioML2018.01 dataset [7], i.e., 4-ASK, 8-ASK, BPSK, QPSK, 4-HQAM, 16-HQAM, 64-HQAM, 16-QAM, 32-QAM, 64-QAM, 128-QAM, 256-QAM, 16-APSK, 32-APSK, 64-APSK, 128-APSK. In more detail, the training dataset consists of 7168 images with dimensions of $256 \times 256 \times 1$, each containing 1000 symbols, while the validation and testing datasets consist of 3584 and 17920 images with the same dimensions, respectively. It should be mentioned that the chosen resolution of 256×256 effectively balances symbol distinction across different SNR values with manageable CNN model complexity, in line with the original AN model's input size [12]. Furthermore, since a wireless communication system is expected to operate under varying SNR conditions, the LN was trained with a dataset consisting of multiple SNR values simultaneously. Specifically, the average received SNR of the dataset covers the range from 0 dB to 30 dB in steps of 5 dB, for a total of 7 SNR values. Hence, the accuracy of the CNN models is lower than what could be achieved by training them

[1] In the spirit of reproducible research, the code used for the simulations is available at: https://github.com/wcipAUTH/Automatic-Modulation-Classification-under-Phase-Imperfections.

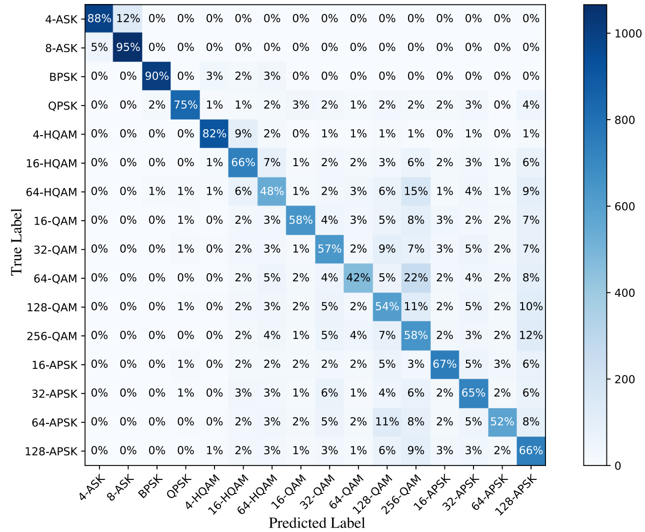
separately for each SNR. However, training the CNN models for each specific SNR value, while potentially more accurate, poses challenges in that it increases the training time due to the large number of SNR values and complicates memory management since different weight configurations must be stored.

IV. Simulation Results

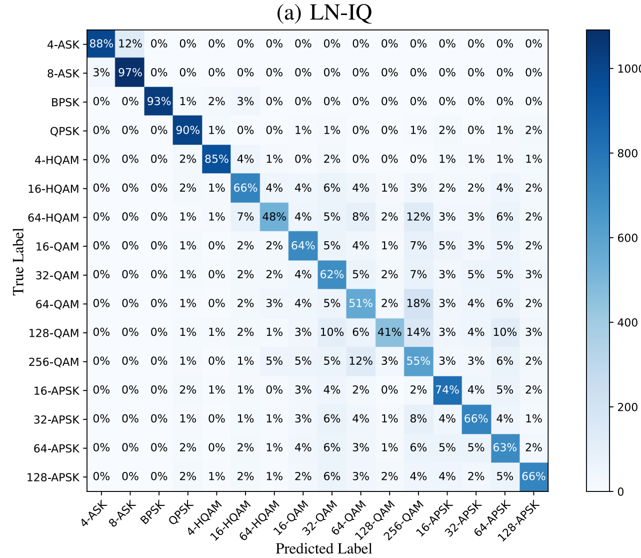
In this section, we assess the performance of various CNN architectures under different PI conditions to evaluate our method's effectiveness. Specifically, we examine the standard LN and AN architectures trained on IQ image along with their polar domain-trained versions, LN-P and AN-P, which, while structurally similar to LN-IQ and AN-IQ, process data in the polar plane. The layer structures and outputs for both LN and AN in their symmetric forms and their polar domain adaptations are presented in Tables I and II. Additionally, we explored LN and AN variants incorporating asymmetric kernels, labeled LN-IQAK and AN-IQAK, as well as LN-PAK and AN-PAK, where both polar transformation and asymmetric kernels are applied. In contrast to their symmetric counterparts, the asymmetric LN model features convolutional layers of 5×5 and 3×6, whereas the asymmetric AN model comprises layers of 3×9 followed by 3×6 and three consecutive layers of 3×3. These dimensions were selected through a process of experimentation, testing various kernel sizes to optimize the CNN performance.

Figs. 2(a) and 2(b) provide a detailed examination of the accuracy of our method within the dataset through confusion matrices showing the performance of LN-IQ and LN-PAK when $\kappa = 2$, which corresponds to a typical PI scenario. These matrices are useful not only for classifying the accuracy of each constellation, but also for highlighting the frequency with which different constellations are mistakenly identified as each other. As can be seen, both LN-PAK show an increase in the majority of the diagonal elements compared to the LN-IQ variant, thus confirming the improved accuracy achieved by our method. Moreover, while certain constellations such as HQAM and QAM are more prone to confusion in LN-IQ, our method significantly reduces such confusion in LN-PAK, as evidenced by the notable increase in the majority of diagonal elements in the confusion matrix of LN-PAK.

In Fig. 3, we present the classification accuracy of the algorithms for average received SNR values, which is defined as $\text{SNR} = \frac{\mathbb{E}[|h|^2]E_s}{N_0}$ and varies from 0 dB to 30 dB, and PI described by $\kappa = 0$ and $\kappa = 10$. The chosen SNR values are typically higher than those simulated in AMC, but due to the inclusion of PI, the performance of the constellations is already severely compromised, as explained in Section II. Furthermore, the selected PI levels in Fig. 3 represent two extreme cases of phase impairments. Specifically, $\kappa = 0$ represents the most challenging situation where the phases are uniformly distributed (e.g., Rayleigh distributed channel), while $\kappa = 10$ represents the presence of small phase distortions (e.g., Ricean distributed channel). As can be observed, the application of our method to both LN and AN enhances the resilience against PI across all SNR levels and significantly improves the classification accuracy even under severe PI conditions. More specifically, LN-IQAK and AN-IQAK consistently showed superior classification accuracy compared to conventional LN-IQ and AN-IQ, respectively, while the same was also observed for both LN-P and AN-P. However, the



(a) LN-IQ



(b) LN-PAK

Fig. 2. LN confusion matrices for PI $\kappa = 2$.

most significant improvement was achieved by LN-PAK and AN-PAK, which confirms the effectiveness of our method regardless of the SNR. Therefore, by combining the polar transformation of the received signals and incorporating asymmetric kernels in the convolutional layers, the classification accuracy can be significantly improved even in the case of severe PI.

Finally, in Fig. 4, we investigate the effect of the proposed polar transformation and CNN architecture modifications on the average accuracy of AMC for different PI levels, where the average is in terms of SNR (0–30 dB). As it can be seen, the application of our method resulted in increased resilience in the presence of PI for both LN and AN, leading to a significant improvement in classification accuracy across all PI conditions. In more detail, for the LN model, LN-IQAK consistently demonstrated superior classification accuracy compared to the conventional LN-IQ, while the same was also observed with LN-P. However, the most substantial improvement was achieved by LN-PAK. Similar behavior is observable for AN, with AN-PAK achieving the greater accuracy gain, thus, verifying the robustness of our method even

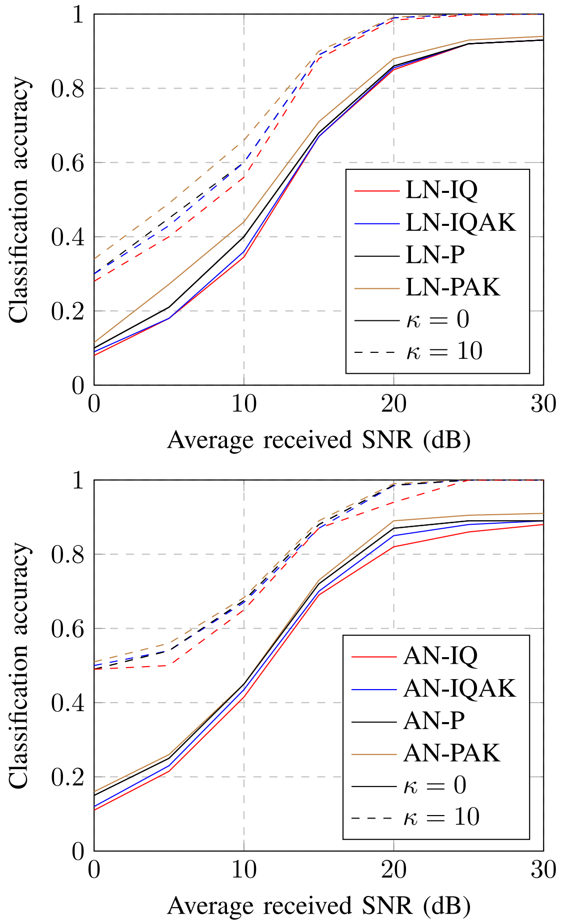


Fig. 3. Classification accuracy vs SNR for a) LN variants b) AN variants.

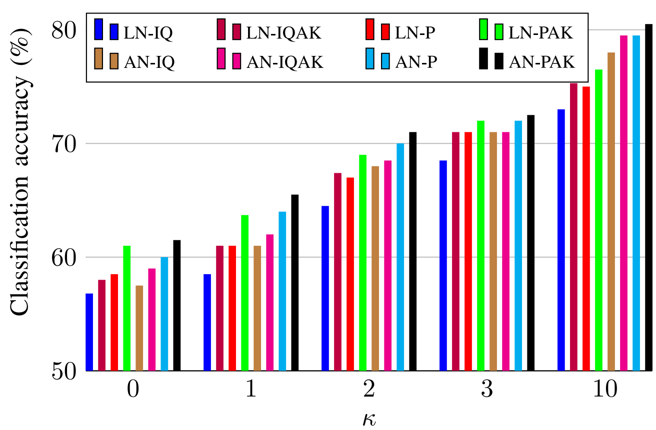


Fig. 4. Overall classification accuracy for different κ.

in more complicated CNN architectures. Furthermore, Fig. 4 provides a significant insight, where LN-PAK outperforms AN-IQ across strong PI conditions, a finding that is pivotal as it challenges the prevailing notion that more complex CNN architectures inherently lead to better performance. Finally, under challenging PI conditions, denoted by $\kappa = 0$, both LN-IQ and AN-IQ exhibit similar classification accuracy. Consequently, the augmentation of CNN complexity, achieved by increasing either its depth through additional layers or its width through extra filters, does not necessarily result in an improvement in classification accuracy. This underscores the impact of PI in diminishing the performance gains associated with a more complex CNN. However, our proposed method proves to be significantly effective even in the extreme scenario of $\kappa = 0$, providing a substantial enhancement in performance.

V. Conclusion

In this letter, we addressed the challenge of AMC under PI conditions, focusing on enhancing AMC accuracy while maintaining low time complexity. Specifically, we proposed a scheme incorporating a polar transformation of the received data and an asymmetric kernel dimension approach that can applied to every CNN-based AMC framework. The provided simulation results verified the PI resilience of the proposed AMC framework, illustrating the gain in classification accuracy of the proposed scheme, especially as the effect of PI increases. Interestingly, our results illustrate that the proposed method can lead simpler CNN architectures such as LN to achieve better accuracy than more complex ones such as AN, thus challenging the popular notion that more complex architectures inherently lead to better performance. This highlights the importance of our approach in AMC applications, as it can improve classification accuracy in scenarios with PI without increasing system complexity.

References

[1] A. Hermawan, R. Ginanjar, D.-S. Kim, and J. M. Lee, "CNN-based automatic modulation classification for beyond 5G communications," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1038–1041, May 2020.

[2] T. Huynh-The et al., "Automatic modulation classification: A deep architecture survey," *IEEE Access*, vol. 9, pp. 142950–142971, 2021.

[3] L. M.-M. Suárez and A. G. Armada, "Compensation of phase noise in 5G NR with machine learning," in *Proc. IEEE Globecom Workshops*, 2022, pp. 717–722.

[4] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.*, 2016, pp. 213–226.

[5] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Los Alamitos, CA, USA, 2015, pp. 5353–5360.

[6] P. Ghasemzadeh, S. Banerjee, M. Hempel, and H. Sharif, "A novel deep learning and polar transformation framework for an adaptive automatic modulation classification," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13243–13258, Sep. 2020.

[7] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.

[8] K. V. Mardia and P. E. Jupp, *Directional Statistics*. Hoboken, NJ, USA: Wiley, 2009.

[9] T. K. Oikonomou, S. A. Tegos, D. Tyrovolas, P. D. Diamantoulakis, and G. K. Karagiannidis, "On the error analysis of hexagonal-QAM constellations," *IEEE Commun. Lett.*, vol. 26, no. 8, pp. 1764–1768, Jun. 2022.

[10] S. Bicaïs and J.-B. Doré, "Design of digital communications for strong phase noise channels," *IEEE Open J. Veh. Technol.*, vol. 1, pp. 227–243, May 2020, doi: 10.1109/OJVT.2020.2994626.

[11] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017.