Meta-Learning Based Few Pilots Demodulation and Interference Cancellation For NOMA Uplink

Hebatalla Issa, Mohammad Shehab and Hirley Alves
Centre for Wireless Communications (CWC), University of Oulu, Finland
Email: firstname.lastname@oulu.fi

Abstract—**Non-Orthogonal Multiple Access (NOMA) is at the heart of a paradigm shift towards non-orthogonal communication due to its potential to scale well in massive deployments. Nevertheless, the overhead of channel estimation remains a key challenge in such scenarios. This paper introduces a data-driven, meta-learning-aided NOMA uplink model that minimizes the channel estimation overhead and does not require perfect channel knowledge. Unlike conventional deep learning successive interference cancellation (SICNet), Meta-Learning aided SIC (meta-SICNet) is able to share experience across different devices, facilitating learning for new incoming devices while reducing training overhead. Our results confirm that meta-SICNet outperforms classical SIC and conventional SICNet as it can achieve a lower symbol error rate with fewer pilots.**

Index Terms—**NOMA, SIC, deep learning, SICNet, pilot allocation, meta-learning.**

I. INTRODUCTION

The path to future wireless networks is encompassed by a massive deployment of devices, enabling smart cities, autonomous vehicles, and many unforeseen scenarios. Machine-type Communication (MTC) services have revolutionized the wireless communications industry by shifting the focus toward the Internet of Things (IoT). Besides a large number of devices, these mMTC devices are often battery constraint, with limited computational capability, and have heterogeneous and sporadic traffic patterns. Combined, these characteristics impose many challenges for the design of efficient random-access procedures and radio resource management. Therefore, researchers are increasingly studying multiple access techniques that are able to scale well to this massiveness while coping with the existence of interference and scarce spectrum. In this context, Non-Orthogonal Multiple Access (NOMA) with Successive Interference Cancellation (SIC) has shown potential and has been the focus of academy and industry in recent years [1]–[7]. This is because NOMA schemes allow for multiple-user transmission with superior performance compared to conventional orthogonal schemes. These recent surveys evince the popularity and potential of NOMA and overview key characteristics, techniques, and applications [1], [4], [5].

Despite the recent advances, there are many open challenges concerning NOMA in massive IoT deployments such as latency, complexity of receivers, perfect channel knowledge and effective utilization of radio resources. These challenges prompted researchers to look into data-driven approaches to handle multi-user connectivity [7]–[10]. We should note that data-driven learning-based communication systems aim to co-exist or replace conventional model-based approaches thanks to their low complexity and ability to adapt well to varying channels. Moreover, the learning-based models can be easily customized to any environment, however, at cost of new training instances. For instance, [9] argues that although linear processing has been effective in NOMA systems, non-linear processing is sometimes necessary to maintain good performance. The authors propose a neural network architecture that leverages the benefits of both linear and non-linear processing leading to efficient real-time detection performance. The authors in [8] introduce a deep-learning-based user detection solution (DeepMuD) for the uplink in massive MTC NOMA. The proposed DeepMuD employs an offline-trained Long Short-Term Memory (LSTM)-based network for multi-user detection without the need for perfect channel state information (CSI). Interestingly, authors report that DeepMuD improves error performance compared to conventional detectors and becomes even better as the number of devices increases. Authors in [7] introduce a deep learning approach for estimating the symbols named SICNet in the downlink NOMA. Contrary to conventional SIC, SICNet replaces the interference cancellation blocks with deep neural networks (DNNs) to infer the soft information representing the interfering symbols in a data-driven fashion, yielding robustness against changes in the number of users and power allocation.

More recently, ML-based transceivers have gained lots of interest. For instance, the authors of [11] proposed a meta-SICNet detection scheme for massive MIMO. The results showed that their meta-SICNet outperforms the conventional MMSE detector. Furthermore, the work in [12] investigates learning-based transceivers using joint learning and meta-learning techniques. An important limitation of the autoencoder-based approach to end-to-end training is that training should be generally carried out from scratch for each new channel. Despite the enhanced performance, such models can be trained on an end-to-end basis irrespective of the underlying modulation or multiplexing scheme. However, when relying on conventional training schemes, these models are required to be trained from scratch for each separate transmission condition, which is cumbersome in MTC networks with heterogeneous traffic and radio resources. Therefore, such networks would benefit from a learning scheme that accumulates experience to facilitate learning in new conditions.

In this context, meta-learning was introduced [13]–[15], also known as learning to learn. Often communications networks are expected to operate under a variety of system configurations. Using conventional learning requires the training of a separate model for each system configuration, leading to data and training time and processing costs. Owing to efficiency, we want to train a single model that would perform across all configurations. However, for such joint training, there may not exist a model that is able to perform well. On the other hand, meta-learning uses data from multiple configurations to infer a model class and learning procedure, enabling learning on configurations of interest. For example, few-shot classification aims to infer a learning procedure that trains a classifier under limited training data from each class. To do so, the learning procedure is inferred from meta-training data that quickly train a classifier on meta-training tasks, rather than training a single model to classify across all tasks [13], [14]. More specifically, in a wireless communication setting, meta-learning outperforms (with respect to symbol error probability) conventional and joint training in IoT scenarios where devices transmit few pilots while adapting to non-linearities and fading. Learning to communicate on noisy or fading channels requires training to be carried out from scratch for every channel, for example, using pilot symbols. In this case, joint training is equivalent to non-coherent transmission. In [16], authors address the sensing and fusion problem in massive MTC, which aims to collect and process a large amount of information and extract key features representing the observed process. However, sensing and fusion impose communication overheads and data redundancy to perform a given accuracy. The authors then proposed a meta-learning adaptive sensing and reconstruction framework that leverages prediction error and sensing decisions so to reduce the amount of communication overhead while guaranteeing robustness. Besides, they show that meta-learning-based approaches outperform conventional machine-learning algorithms in terms of convergence rate.

A. Contributions

In this context, we build upon [7], [8], [17] and propose a novel few pilot-aided detection mechanism, named meta-SICNet, for NOMA uplink in massive MTC. Herein, as in [7], we infer the interfering symbols via deep neural networks. However, we can reduce training over the network by the use of meta-learning, and in addition, as in [17], we reduce the number of pilots required for detection. Therefore, our contributions are

- We propose a data-driven approach that applies meta-learning for NOMA uplink transceiver design.
- The network model consists of multiple MTC devices transmitting their superimposed signal to the BS. Then, the meta-SICNet framework in the BS learns to recover the transmitted symbols using the meta-training and adaptation approach.
- The meta-SICNet is able to decode symbols with fewer pilots, in the presence of interference, compared to basic SIC and conventional SICNet methods.
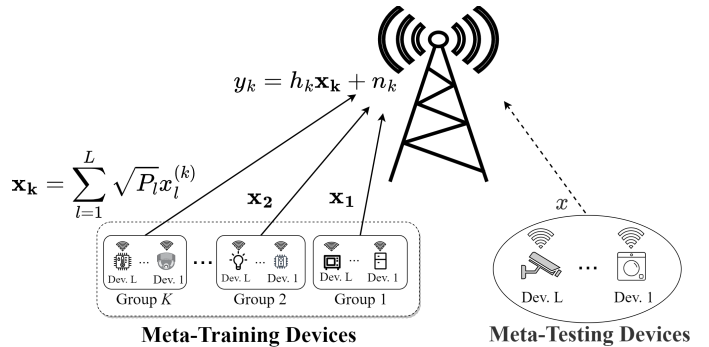

Fig. 1: The system model comprises k sets of meta-training devices and one Meta-testing device. Each meta-training set, k, transmits a superimposed signal, $\mathbf{x_k}$ of L-devices to the BS.

- Extensive simulation results elucidate that the proposed meta-SICNet scheme significantly outperforms both classical SIC and conventional SICNet in terms of outage probability and requires a lower number of pilots (down to 2 pilots).

B. Outline

The rest of the paper is organized as follows: section II depicts the system model and the problem formulation. Conventional SICNet and the proposed meta-learning-based SICNET solutions are presented in Sections III and section IV elucidates the results, and finally, the paper is concluded in Section V.

II. SYSTEM LAYOUT AND PROBLEM FORMULATION

A. System Layout

We consider a non-orthogonal uplink channel where we have K meta-training device groups and one meta-testing group. For each group k, L IoT devices transmit their data to the BS within the same time and frequency resources as illustrated in Fig.1. The devices transmit $\{x_{l=1,\ldots,L}^{(k)=1,\ldots,K}\}$ symbols to the BS using superposition coding, where the symbol $x_l^{(k)}$ is the lth symbol transmitted from the kth device. Specifically, the symbol $x_l^{(k)}$ is amplified with the transmitted power P_l for $l = 1,\ldots,L$. The channel input is the superimposed signal $\mathbf{x_k}$ given by

$$\mathbf{x_k} = \sum_{l=1}^{L} \sqrt{P_l} x_l^{(k)} \tag{1}$$

The symbols are sampled from an M-point constellation $\mathcal{S}$, and assumed to be mutually independent with unit mean power, i.e., $\mathbb{E}\left[\left|x_l^{(k)}\right|\right] = 1$. For simplicity, we assume all devices have the same modulation order M. The channel output at the BS for each k group denoted as y_k for $k = 1,\ldots,K$ is given by

$$y_k = h_k \mathbf{x_k} + n_k, \tag{2}$$

where $h_k \in \mathbb{C}$ is the channel coefficient between the BS and the device-group k, and $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the complex additive white Gaussian noise (AWGN).
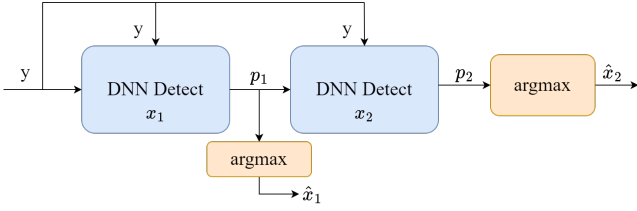
Fig. 2: Architecture of SICNet for DNN-based SIC for $L = 2$ devices.

B. Problem Formulation

We aim to construct a symbol detection demodulator based on a short packet transmission of a few pilot symbols. For each device l in the device group k, the detected symbol, i.e., $\hat{x}_l^{(k)}$, is estimated from the channel output y_k. To enable symbol recovery using few-pilot learning, we use a data-driven approach based on meta-learning and DNNs. Assuming no prior information about the channel model at the receiver, the BS can use the signals received from the previous pilot transmissions of K other IoT device groups, which are referred to as meta-training devices and their data as meta-training data. In particular, the BS has available N pairs of pilots $\mathbf{x_k}$ and received signal y_k for each meta-training group $k = 1, \ldots, K$. The meta-training dataset is denoted as $\mathcal{D} = \{D\}_{k=1,\ldots,K}$, where $D_k = \{(\mathbf{x_k}^{(n)}, y_k^{(n)}) : n = 1, \ldots, N\}$, and $(\mathbf{x_k}^{(n)}, y_k^{(n)})$ are the pilot-received signal pairs for the kth meta-training group of devices.

For the meta-test devices, the BS receives P pilot symbols. It collects the P pilots received from the target device in set $\mathcal{D}_T = \{(x^{(n)}, y^{(n)}) : n = 1, \ldots, P\}$. The demodulator can be trained using meta-training data $\mathcal{D}$ and the pilot symbols $\mathcal{D}_T$ from the meta-test devices. To recover the symbols, the successive interference cancellation algorithm is utilized, such that the power allocations for the superimposed symbols satisfy $P_1 > P_2 \ldots > P_L$. Hence, the meta-learning-based DNN successively detects each symbol from the channel output signal y_k.

III. DEEP LEARNING-BASED SIC

In this section, we describe the data-driven solutions based on DNNs to perform SIC and symbol detection. First, we introduce the conventional DNN based SICNet; then we illustrate the proposed meta-SICNet approach.

A. SICNET

The DNN-based SIC, called SICNet, was introduced in [7] to estimate the transmitted symbols in the downlink NOMA users scheme. For the sake of illustration, we consider an uplink scheme with $L = 2$ devices in each device group. As depicted in Fig. 2, the architecture of SICNet is implemented using sequential DNN blocks. Each block performs symbol recovery, which is considered a classification problem. Therefore, the structure of SICNet consists of L DNN stages, where each stage estimates the transmitted symbol of each device. Each stage uses the received signal y_k and the output vector $\mathbf{p}_l$ for
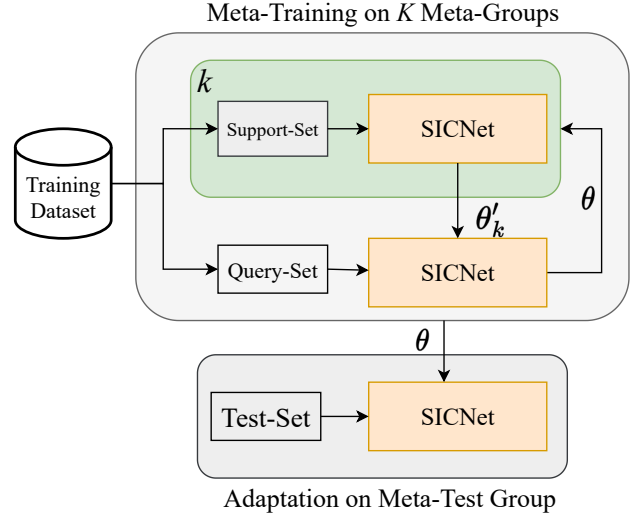

Fig. 3: Architecture of meta-SICNet

$l = 1, \ldots, L$ that represents the conditional distribution of the corresponding symbol, expressed as

$$\mathbf{p}_l = \begin{bmatrix} \hat{p}(x_l = \gamma_1 | y_k, \mathbf{p}_1, \ldots, \mathbf{p}_{l-1}) \\ \vdots \\ \hat{p}(x_l = \gamma_M | y_k, \mathbf{p}_1, \ldots, \mathbf{p}_{l-1}) \end{bmatrix}, \quad (3)$$

where γ_i is the ith constellation symbol in the constellation space $\mathcal{S}$ for $i = 1, \ldots, M$, and $\hat{p}(x_l = \gamma_i | y_k, \mathbf{p}_1, \ldots, \mathbf{p}_{l-1})$ is a parametric estimation of the probability of x_l given y_k and the previous estimates $\mathbf{p}_1, \ldots, \mathbf{p}_{l-1}$.

The data-driven feature of the SICNet increases its capability to detect symbols reliably without requiring full knowledge of the channel model. Concretely, it works in a model-agnostic manner contrary to the conventional SIC algorithm [7], which requires a restricted channel model [18]. Thus, SICNet operates with arbitrary channel models depending on the classification process of the data-driven DNN. Furthermore, the SICNet architecture only requires knowledge of the modulation order M and the power coefficients of the transmitted signals.

B. Proposed Meta-Learning Approach

In this section, we describe the meta-learning approach for demodulation based on [17]. In our proposed method, the meta-learning-based SICNet allows learning the interference cancellation and signals demodulation using a few-pilot transmission. Moreover, the model can quickly adapt to the change in the channel conditions, in contrast to the classical SICNet.

In our model, the learning is performed in two phases; the meta-learning phase and the test-adaptation phase, as shown in Fig. 3. In the meta-learning phase, we use the meta-training data $\mathcal{D}$ and iterate over the number of K meta-tasks (meta-devices) to learn a general parameter vector θ using the inner task parameter θ'_k. This can be done offline at the BS while collecting data from training devices. For the meta-test adaptation phase, the learned parameter θ is adapted to enable

Algorithm 1: Meta-Learning based SIC

Input: Meta-training data $\mathcal{D} = \{\mathcal{D}_{k=1,\ldots,K}\}$, and meta-testing pilot data $\mathcal{D}_T$; N_{tr}^S and N_{tr}^Q; step size α and β.

Output: Learning parameter vector θ; SER.

1 Randomly initialize the parameter vector θ.

2 START META-TRAINING:

3 **while** *not done* **do**

4   **for** *each meta-training group k* **do**

5   Randomly split $\mathcal{D}_k$ into two sets; support-set $\mathcal{D}_k^S$ of size N_{tr}^S, and query-set $\mathcal{D}_k^Q$ of size N_{tr}^Q.

6   Calculate the gradient $\nabla_\theta L_{\mathcal{D}_T}(\theta)$ from (4) with $\mathcal{D}_T = \mathcal{D}_k^Q$, and $\nabla_\theta^2 L_{\mathcal{D}_T}(\theta)$.

7   Compute adapted parameters θ' using (5) as
$$\theta_k' = \theta - \beta \nabla_\theta L_{\mathcal{D}_k^Q}(\theta)$$

8   **end**

9   update the meta-parameter θ:
$$\theta \leftarrow \theta - \alpha \nabla_\theta \sum_{k=1}^{K} L_{\mathcal{D}_k^S}(\theta_k')$$

10 **end**

11 START META-TESTING ADAPTATION:

12 **for** *testing epochs* **do**

13   Load the learned parameter vector θ.

14   Sample data from $\mathcal{D}_T$.

15   Update θ in the direction of the gradient with step size η by
$$\theta \leftarrow \theta - \eta \nabla_\theta L_{\mathcal{D}_T}(\theta)$$

16 **end**

TABLE I: meta-SICNet Model Parameters

Parameter	Value
Number of device groups for meta-SICNet K	20
Number of devices in each group L	2
Modulation Number M	2
Power factors of the devices P_1, P_2	4, 1
Number of DNN layers for SICNet	4
Neurons for DNN block 1	24-12
Neurons for DNN block 2	32-16
Activation function for hidden layers	ReLU
Activation function for output layers	Softmax
Optimizer	Adam
Outer learning rate α	0.1
Inner learning rate β	0.001
Number of training epochs	300
Number of pilots for training data N_{tr}^S, N_{tr}^Q	4, 4
Size of testing data for target devices	10^6
Training SNR	6 dB
Learning-rate for adaptation on test devices η	0.001
Number of adaptation epochs	1000

it to the performance of classical SIC and conventional SICNet. The simulation parameters for meta-SICNeT are presented in Table I. The model considers $K = 20$ meta-device groups for training, and each group includes $L = 2$ non-orthogonal interference devices for the UL system. The transmitted signal to the BS is modulated using the BPSK modulation scheme of modulation order 2, where the signal is superimposed by the power coefficients $P_1 = 4$ and $P_2 = 1$. The SICNet architecture comprises 2 DNN blocks, and each DNN consists of 4 layers; the input and output layers and two hidden layers. The number of neurons in each hidden layer is shown in Table I. The training loss criteria is based on the *combined loss* from both devices. For the end-to-end training, we used the Adam optimizer with the meta-learning rate $\alpha = 0.1$ and inner learning rate $\beta = 0.001$. For simplicity, we consider a symmetric channel model of ± 1, where half of the training groups have a channel of $+1$ and the other half -1. For adaption on target devices, the channel is chosen randomly as $+1$ or -1. The training is performed on an 11-th Gen Intel Core i5 2.40GHz CPU and 16-GB RAM. The training loss converges after 200 epochs.

Fig. 4 depicts the performance of SER for both meta-SICNet and SICNet for a different number of pilots. The simulation was performed for $N_{te} = 1, 2, \ldots, 8$ with an SNR value of 15 dB. The results show that the meta-learning approach outperforms the SICNet in terms of SER. Moreover, it is clear that increasing the number of pilots improves learning performance. However, for meta-SICNet, the SER relatively saturates faster at 2 or 3 pilots. Note that device 2 always performs better since interference is already removed after decoding the signal from device 1, which captures the SIC effect.

In Fig. 5, we plot the symbol outage probability as a function of the SNR for SNR values $\{0, 2, \ldots, 18\}$ using 4 pilot symbols. It appears that the meta-SICNet can capture the effect of SNR as the symbol error rate significantly improves when increasing the SNR. Again, it is obvious that the meta-SICNet outperforms conventional SICNet and the classical SIC for all

fast adaptation based on the few pilots P transmitted by the target (meta-test) devices. Consequently, the test data-set $\mathcal{D}_T$ is used to train a demodulator $p_\theta(x|y)$ to minimize the cross-entropy loss given by

$$L_{\mathcal{D}_T}(\theta) = - \sum_{(x^{(n)}|y^{(n)}) \in \mathcal{D}_T} \log p_\theta(x^{(n)}|y^{(n)}). \qquad (4)$$

Then, the stochastic gradient descent is used to update the parameter θ iteratively as

$$\theta \leftarrow \theta - \eta \nabla_\theta \log(x^{(n)}|y^{(n)}), \qquad (5)$$

where the pair $(x^{(n)}|y^{(n)}) \in \mathcal{D}_T$, and η is the step size. As discussed in [13], the purpose of the MAML algorithm is to find the initial parameter θ such that, for any device, the loss after one iteration of (5) applied to the received pilots is minimized. The training and testing (adaptation) steps are illustrated in Algorithm 1.

IV. NUMERICAL ANALYSIS

In this section, we evaluate the performance of the proposed meta-SICNet in terms of symbol error rate (SER) and compare
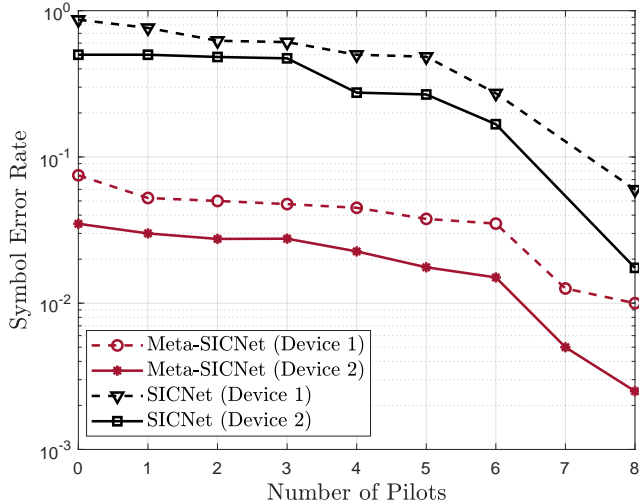
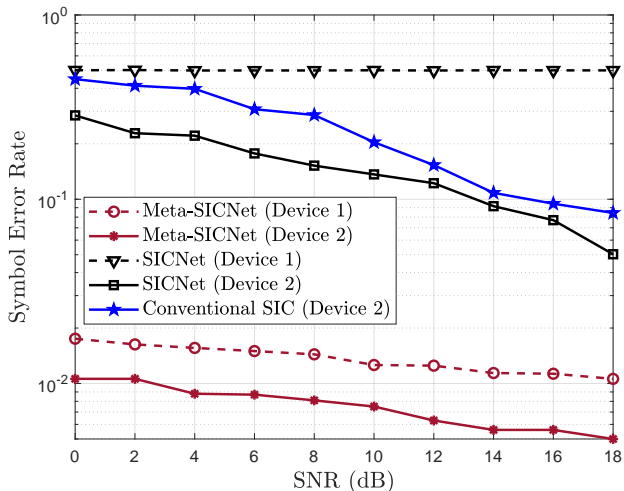Fig. 4: Symbol error rate vs. number of pilots for Meta-testing devices and SICNet (SNR = 15 dB).

Fig. 5: Symbol error rate vs. SNR for Meta-testing devices, SICNet, and conventional SIC (number of pilots = 4).
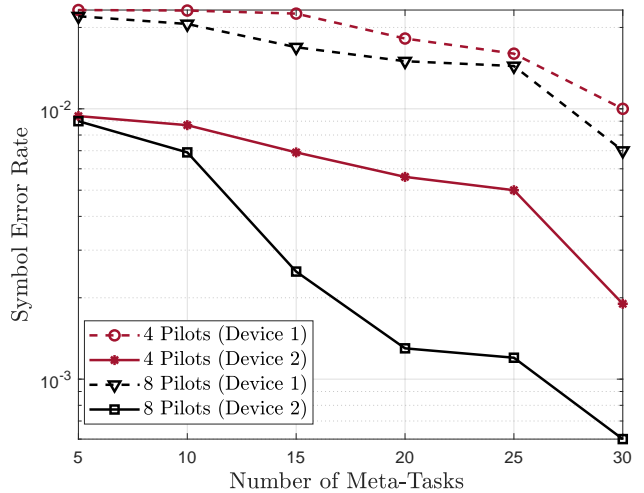

Fig. 6: Symbol error rate vs. number of meta-training tasks (number of device groups) for SNR = 15 dB, and number of pilots 4, 8.

TABLE II: Complexity analysis of meta-SICNet vs. SICNet, 8 pilots

Model	# Parameters	Training time	Test-time
Meta-SICNet	2240	50.3 ms	4.01 ms
SICNet	1120	10.9 ms	41.01 ms

the meta-SICNet approach achieves higher computational complexity, the training can be done offline among the meta-training device groups. Then, the online adaptation process is performed to continuously update the model parameters without having to train the model again. On contrary, the SICNet model requires performing the whole training process online whenever the channel condition alters. Therefore, the testing (adaptation) time for the meta-based approach is much less than SICNet for the same number of pilots. Hence, meta-SICNet can achieve lower SER with relatively low online complexity (i.e, only 4.01 ms of adaptation time).

V. CONCLUSIONS

In this work, we introduced a data-driven, meta-learning-aided NOMA uplink model. Unlike classical SIC and conventional deep learning SIC, the proposed meta-SICNet can accumulate experience across different devices, facilitating the learning process for newly introduced devices and reducing the training overhead. Our results confirm that meta-SICNet outperforms conventional SICNet as it can achieve a lower symbol outage probability. Moreover, meta-SICNet can converge faster and renders significantly good performance for a few pilots (only 2 or 3 pilots). Although meta-SICNet consumes a longer time for training, the training phase is performed offline, and the online adaptation phase consumes a very short time (≈ 4 ms) compared to conventional SIC, which reduces the online complexity. There are plenty of possible extensions to the meta-SICNet approach proposed here, among which full-duplex NOMA with self-interference cancellation, adaptation to

SNR values, which proves the superiority of the proposed meta-learning solution for in NOMA uplink. Note that the classical SIC is the worst performer here, so we just added one curve for illustration.

In Fig. 6, we address the performance of meta-SICNet for a different number of meta-training tasks (device groups) with 15dB SNR and $N_{te} = \{4, 8\}$ pilots. The simulation shows that the capacity of learning for the meta-SICNet model improves when increasing the meta-training device groups. Therefore, the model can achieve lower SER if the training experience is shared between more device groups.

Table II illustrates the training complexity for both meta-SICNet and SICNet in terms of the number of parameters, training time per epoch, and adaptation (testing) time. Although

a higher number of devices, and meta-learning aided massive MIMO.

References

[1] Z. Ding *et al.*, "A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, 2017.

[2] Y. Liu *et al.*, "Nonorthogonal multiple access for 5g and beyond," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2347–2381, 2017.

[3] Y. Yuan, Z. Yuan, and L. Tian, "5G Non-Orthogonal Multiple Access Study in 3GPP," *IEEE Communications Magazine*, vol. 58, no. 7, pp. 90–96, 2020.

[4] 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.

[5] M. Elbayoumi, M. Kamel, W. Hamouda, and A. Youssef, "NOMA-Assisted Machine-Type Communications in UDN: State-of-the-Art and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1276–1304, 2020.

[6] E. Dosti, M. Shehab, H. Alves, and M. Latva-aho, "On the performance of non-orthogonal multiple access in the finite blocklength regime," *Ad Hoc Networks*, vol. 84, pp. 148–157, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S157087051830708X

[7] T. Van Luong *et al.*, "Deep learning based successive interference cancellation for the non-orthogonal downlink," *IEEE transactions on vehicular technology*, vol. 71, no. 11, pp. 11 876–11 888, 2022.

[8] A. Emir, F. Kara, H. Kaya, and H. Yanikomeroglu, "DeepMuD: Multi-User Detection for Uplink Grant-Free NOMA IoT Networks via Deep Learning," *IEEE Wireless Communications Letters*, vol. 10, no. 5, pp. 1133–1137, 2021.

[9] D. Schaufele *et al.*, "GPU-Accelerated Machine Learning in Non-Orthogonal Multiple Access," in *2022 30th European Signal Processing Conference (EUSIPCO)*, 2022, pp. 667–671.

[10] X. Zhang, P. Fan, J. Liu, and L. Hao, "Bayesian learning-based multiuser detection for grant-free noma systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 6317–6328, 2022.

[11] M. Khani, M. Alizadeh, J. Hoydis, and P. Fleming, "Adaptive neural signal detection for massive mimo," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5635–5648, 2020.

[12] S. Park, O. Simeone, and J. Kang, "Meta-learning to communicate: Fast end-to-end training for fading channels," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 5075–5079.

[13] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International conference on machine learning*. PMLR, 2017, pp. 1126–1135.

[14] L. Chen *et al.*, "Learning with limited samples: Meta-learning and applications to communication systems," *Foundations and Trends in Signal Processing*, vol. 17, no. 2, pp. 79–208, 2023. [Online]. Available: http://dx.doi.org/10.1561/2000000115

[15] J. Beck *et al.*, "A Survey of Meta-Reinforcement Learning," Jan. 2023. [Online]. Available: http://arxiv.org/abs/2301.08028

[16] H. Wu *et al.*, "Learn to sense: A meta-learning-based sensing and fusion framework for wireless sensor networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8215–8227, 2019.

[17] S. Park, H. Jang, O. Simeone, and J. Kang, "Learning how to demodulate from few pilots via meta-learning," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2019, pp. 1–5.

[18] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 19–29, 2005.