

Formal Verification for Deep Learning-based Power Control in Massive MIMO

Thanh Le*, Takeshi Matsumura*, Yusheng Ji[†], John C.S. Lui[‡]

*National Institute of Information and Communications Technology (NICT), Japan

[†]National Institute of Informatics (NII), Japan, [‡]The Chinese University of Hong Kong (CUHK), China

Abstract—Deep learning is a promising approach to optimize wireless communication by simplifying the search for near-optimal solutions. Prior studies on deep learning-based wireless communication optimization have explored supervised learning approaches that map raw user information, such as location or channel state information, to optimal power allocation vectors. While this approach demonstrates competitive performance, it is susceptible to adversarial attacks via input perturbations. Current defense mechanisms primarily rely on empirical methods, which do not provide formal guarantees of robustness. We fill this gap by proposing a formal verification framework to evaluate the robustness of deep learning-based power allocation in multi-cell massive multiple-input multiple-output (MIMO) systems against a wide range of potential adversarial input manipulations. To the best of our knowledge, this is the first attempt to formally verify deep neural networks in a regression setting with non-linear output constraints. We model the adversary’s capabilities using hyper-rectangle constraints on their perturbation, adopt the abstraction-based bound-propagation technique (DeepPoly) to bound the interval of potential allocated powers, and formulate the minimum performance requirements as a constrained program for numerical feasibility analysis. Evaluation on publicly available datasets for power allocation in multi-cell massive MIMO indicates that a well-trained model can guarantee the local robustness under location perturbation by $\pm 1\text{m}$ while retaining a maximum 1% optimality gap.

Index Terms—massive MIMO, DNN, formal verification

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) technology is a key component of next-generation wireless networks, designed to accommodate the growing number of connected user equipments (UE) and their diverse demands for new services and applications [1]. By deploying a large number of antennas at the base station (BS), massive MIMO can leverage multiplexing gains and enable high data rates. In multi-cell massive MIMO networks, optimizing power control is crucial for achieving high performance and reducing interference between UEs both within a cell and across different cells.

Power control in massive MIMO increasingly leverages deep neural network (DNN) to reduce complexity by inferring near-optimal solutions within $O(1)$ time complexity while requiring less engineering effort to design new heuristics if the problem formulation changes. Exciting developments have emerged in this area, *e.g.*, DNN has been supervised to approximate optimal power allocation algorithms in massive MIMO systems [2]. This work utilizes the ability of DNN to approximate arbitrary functions to map the positions of UEs in the multi-cell network to their optimal power allocation.

A few following studies have focused on training distributed architectures using only local information within each cell [3], [4]. Another work has addressed varying numbers of users by utilizing convolutional neural network (CNN) [5], while deep reinforcement learning has been applied to control power in massive MIMO networks with user mobility [6].

However, DNN-based power allocation is susceptible to adversarial attacks, where small perturbations can be added to inputs and manipulated DNNs into allocating non-optimal power to UEs [7], [8]. These perturbations are small enough to evade anomaly detection system but carefully crafted to significantly impair the DNN’s performance, *e.g.*, reducing transmission rates, fairness, and degrade quality of service. Previous work [7] proposed an attack on DNN-based power control with the input being the locations of UEs. By using global navigation satellite system (GNSS) spoofing techniques, attackers can alter the positions of nearby UEs, thereby causing infeasible and non-optimal power allocations.

To the best of our knowledge, there are three main defense approaches: (1) denoising autoencoders, (2) adversarial training, and (3) neural network verification (NNV). Firstly, denoising autoencoder networks can be employed to rectify malicious inputs [9]; however, this incurs additional computational overhead during inference. Secondly, adversarial training mixes sampled adversarial inputs to the training dataset for retraining [10]. However, given the real-valued, high-dimensional input space, the sampled adversarial inputs do not guarantee completeness. Besides, this type of non-certifying defense was found to be evaded by stronger adversarial attack schemes [11]. To mitigate this cat-and-mouse game, there is a gaining traction in defenses with NNV [12], [13], [14], as NNV methods provide absolute robustness guarantee.

General NNV techniques have primarily focused on verifying the safety of computer vision, natural language processing, and autonomous aviation systems [15], most of them are classification or unconstrained regression tasks. In contrast, our problem involves verifying non-linear performance guarantees, *e.g.*, guarantee minimum product of signal to interference-plus-noise ratio (SINR), which cannot be directly addressed by general NNV techniques and require additional post-processing. In DNN-based for wireless communication domain, a previous work verified robustness of DNN-based antenna selection for massive MIMO systems [16], in which DNNs are trained to classify channel information to optimal antenna configurations. As far as we know, verifying DNN in

regression-based tasks with non-linear throughput and fairness constraints remains an open challenge.

We fill this gap by proposing a verification framework for DNN-based power control in massive MIMO systems. Our proposed framework systematically analyzes whether the trained DNN meets performance requirements under various levels of adversarial perturbations. First, we convert different levels of adversarial perturbation into a wide variety of hyper-rectangle input properties. We apply abstraction-based bound-propagation NNV technique called DeepPoly [12], [14] to compute an over-approximation of the output, which covers all possible power allocations. Then, we propose an adaptation of the existing NNV to work with the optimization goal in the considered problem. Particularly, we check whether the model satisfies a guarantee on product of SINR, *e.g.*, the optimality gap must fall below a certain threshold, by formulating a constrained program for feasibility analysis.

II. PRELIMINARIES

A. System Model

To investigate the robustness of deep learning-based power control in a multi-cell massive MIMO networks, we reproduce the system model and DNN-based power control method which have been made publicly available in [2]. In this system, the massive MIMO network consist of J cells, each cell having a BS of M antennas and covers K UEs. The downlink signal transmitted by the BS in cell l is given by $v_l = \sum_{k=1}^K \mathbf{w}_{lk} u_{lk}$, where $u_{lk} \in \mathcal{N}(0, \rho_{lk})$ is the data signal in the downlink for UE k in BS l , with a precoding vector $\mathbf{w}_{lk} \in \mathbb{C}^M$ that determines the transmission beamforming satisfying $\|\mathbf{w}_{lk}\|^2 = 1$, thus, ρ_{lk} is the transmission power. The best precoding schemes that have been adopted by [2] in generating dataset is multicell-minimum mean square error (M-MMSE). The downlink SINR for UE k in cell j is given as:

$$\gamma_{jk} = \frac{\rho_{jk} a_{jk}}{\sum_{l=1}^J \sum_{i=1}^K \rho_{li} b_{lijk} + \sigma^2}, \quad (1)$$

where σ^2 is the additive white Gaussian noise (AWGN) variance and the average channel gain is $a_{jk} = |\mathbb{E}[\mathbf{w}_{jk}^H \mathbf{h}_{jk}^j]|^2$, with $\mathbf{h}_{jk}^j$ denoting the channel between BS j and UE k in cell j . The average interference gain is:

$$b_{lijk} = \begin{cases} \mathbb{E}[\|\mathbf{w}_{li}^H \mathbf{h}_{jk}^l\|^2] - \mathbb{E}[\|\mathbf{w}_{li}^H \mathbf{h}_{jk}^l\|^2], & \text{if } (l, i) = (j, k), \\ \mathbb{E}[\|\mathbf{w}_{jk}^H \mathbf{h}_{jk}^l\|^2], & \text{otherwise.} \end{cases} \quad (2)$$

B. Power Control Problem

In this study, we attempt to verify the DNN-based optimal power allocation strategy to maximize the max-product SINR [2], which is formulated as:

$$\begin{aligned} & \underset{\rho_{jk} : \forall j, k}{\text{maximize}} && \prod_{j=1}^J \prod_{k=1}^K \gamma_{jk}, \end{aligned} \quad (3a)$$

$$\begin{aligned} & \text{subject to} && \sum_{k=1}^K \rho_{jk} \leq P_{\max}, \quad j \in [1, J]. \end{aligned} \quad (3b)$$

The problem above seeks to maximize objective in Eq. 3a, which is the product of UEs' SINR, while Eq. 3b ensures that the allocated power adhere to the maximum power threshold $P_{\max}$ promoting a trade-off between UE's performance and total throughput. Compared to uniform max power control and max-min spectral efficiency power control, max product SINR offers a balance between fairness and efficiency [2].

C. Deep Learning-based Power Control

In [2], DNNs are employed to compute an approximated optimal power allocation that maximizes the SINR product of all UEs. A publicly available dataset¹ is provided as $\{x^{(n)}, \rho^{(n)}\}_{n=1}^{N_T}$, where $x^{(n)} \in \mathbb{R}^{J \times K \times 2}$ represents an input feature contains location of UEs, the output label $\rho^{(n)} \in [0, \rho_{\max}]^{J \times K}$ is the optimal power, and N_T denotes the size of the dataset. We denote a regression-based DNN model as $N(\cdot; \theta)$, with x as the input to the model and the predicted output as $N(x; \theta) = \rho$, where θ constitutes the set of parameters of the model N . The corresponding loss function is represented as $\mathcal{L}(x, \rho; \theta)$. The goal of the DNN model is to learn the mapping from the UEs' geographical positions, x , to the optimal power allocation solution, ρ produced by mathematical optimization tools. A feed-forward neural network architecture and mean square error (MSE) loss function that has been adopted to approximate the optimal power allocation. In the standard setting outlined in [2], the input x is a vector representing the geographical locations of all UEs in the system, while ρ is a vector containing the optimal power solution obtained by solving the power allocation problem (refer to Eq. 3)

D. NNV - An Overview

Given a DNN N and a property ϕ , the general NNV problem asks if ϕ is a valid property of N . Typically, ϕ is a formula of the form $\phi_{in} \implies \phi_{out}$, where ϕ_{in} is a property over the inputs of N and ϕ_{out} is a property over the outputs of N . This property has been used to encode safety and security requirements of DNN [17]. A DNN verifier attempts to find a *counterexample* to N that satisfies ϕ_{in} but violates ϕ_{out} . If no such counterexample exists, ϕ is an unsatisfiable property of N (UNSAT or proven robust); otherwise, satisfiable (SAT).

III. PROPOSED VERIFICATION FRAMEWORK

Fig. 1 illustrates our proposed robustness guarantees framework for DNN-based power control in massive MIMO through a three-phase pipeline. First, we formulate local robustness input properties encapsulating possible adversarial noise around system inputs. Then, an abstraction-based bound-propagation DNN verifier method [12], [14] is adopted to compute bounds on possible power allocations provided by a trained DNN. Finally, given the power allocation bounds, we formulate a constrained program to prove the non-existence of a power allocation that violates the maximum optimality gap guarantee, *e.g.*, reducing the optimization objective by more than ζ .

¹<https://data.ieeeemlc.org/Ds2Detail>

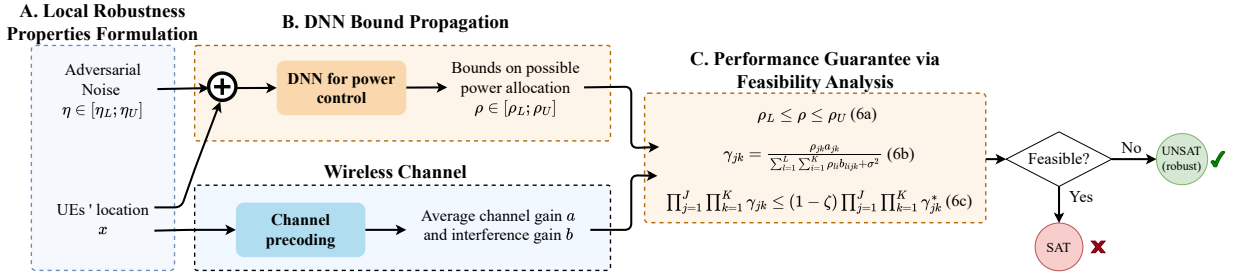


Fig. 1: Overview of the proposed NNV framework to verify DNN-based power control in multi-cell massive MIMO networks.

A. Local Robustness Properties Formulation

We first define the input properties encoding the local robustness requirements of DNN-based power allocation. Following the threat model in [7], [10], we assume that an attacker can introduce controlled perturbations to the input fed into the DNN. For instance, the attacker may manipulate the reported positions of UEs using spoofing techniques, in which global navigation satellite system receivers are deployed near the users' actual locations. Furthermore, the perturbation remains minimal relative to the actual input $x^{(n)}$. The adversary's objective is to compute input perturbations along the gradient direction to maximize the loss function (e.g., maximize MSE loss to optimal power allocation), thereby degrading the performance of the DNN-based power allocation system.

To formally verify the robustness of the considered DNN-based power allocation scheme, we define input specifications $\phi_{in} := [\eta_L, \eta_U]$, such that η_L and η_U are two vectors having the same size as input $x^{(n)}$, to encode permitted perturbation levels. The adversary seeks to generate an adversarial example x' within a constrained neighborhood of x such that the DNN output becomes erroneous. This adversarial example is constructed by adding a bounded perturbation $x' = x^{(n)} + \eta$ to the original input $x^{(n)}$, where $\eta \in [\eta_L, \eta_U]$. Similarly, output specifications ϕ_{out} are introduced to capture minimum performance guarantees, requiring that the achieved performance remains at least $1 - \zeta$ times that of the optimal solution under controlled conditions. This property, referred to as local robustness, ensures stability against small adversarial perturbations around given inputs.

B. DNN Bound Propagation

In this phase, we aim to bound the upper and lower limits of power allocations given the input property and the trained DNN. Bounding the output range of a DNN under constrained input perturbations has emerged as a critical topic in NNV and certified adversarial defense [12], [14]. For a neural network $N(x, \theta)$, we analyze its behavior at a nominal input $x^{(n)}$ subject to a bounded perturbation η , where $x = x^{(n)} + \eta$ and x lies within an l_p -norm ball $\mathcal{B}_p(x^{(n)}, r)$. Computing exact output bounds of a DNN is generally intractable [12].

To address this challenge, we employ DeepPoly [12] abstraction, which provides provable linear relaxation of bounds on neuron outputs given specified input perturbations. This method constructs two linear functions that serve as guaranteed

lower and upper bounds for each neuron's output with respect to its perturbed input. The procedure begins by relaxing non-linear activation functions into linear constraints. These relaxations of bounds on output of non-linear activations are then propagated through the network architecture, layer by layer, from the input to the final output. DeepPoly accepts input specifications consisting of bounds $[x^{(n)} + \eta_L, x^{(n)} + \eta_U]$, along with the trained DNN in Verification of Neural Networks Competitions (VNN-COMPs) format [15]. The output is a set of tight bounds $[\rho_L, \rho_U]$ on the DNNs' outputs.

Note that NNV has been generalized to general computational graphs [14], enabling the computation of linear output bounds not restricted to feed-forward neural networks but also able to find bounds of modern architectures, e.g., CNNs, ResNets, or Transformers. However, we will adhere to the feed-forward neural network architecture used in publicly available implementations to ensure reproducibility and consistency with previous works [2], [7].

C. Performance Guarantee via Feasibility Analysis

The performance guarantee of DNN-based power control in multi-cell massive MIMO networks cannot be captured by existing NNV properties, which primarily target classification tasks or generic regression objectives with linear constraints [15]. We propose a constrained program to encode the non-linear performance guarantee as follow:

$$\rho_L \leq \rho \leq \rho_U, \quad (4a)$$

$$\gamma_{jk} = \frac{\rho_{jk} a_{jk}}{\sum_{l=1}^J \sum_{i=1}^K \rho_{li} b_{lik} + \sigma^2}, \quad (4b)$$

$$\prod_{j=1}^J \prod_{k=1}^K \gamma_{jk} \leq (1 - \zeta) \prod_{j=1}^J \prod_{k=1}^K \gamma_{jk}^*, \quad (4c)$$

where Eq. 4a constrains the power allocation variables ρ to lie within the bounds $[\rho_L, \rho_U]$ obtained from the DNN bound propagation, Eq. 4b defines the SINR γ_{jk} for each UE k in cell j based on the power allocation and channel conditions, and Eq. 4c encodes the performance guarantee by ensuring that the product of all SINR values does not fall below $(1 - \zeta)$ times the optimal product γ_{jk}^* . Feasibility of this program is verified using numerical optimization tools. If there exists no feasible solution then the verification problem is deemed to be unsatisfiable or UNSAT. Consequently, the DNN can guarantee to achieve a relative optimality gap of ζ for all adversarial

inputs within $[\eta_L, \eta_U]$. Otherwise, the verification problem is satisfiable (SAT) and there may exist an exploitation.

Since the output bounds $[\rho_L, \rho_U]$ represent an over-approximation, and the feasibility analysis in Eq. 4 check for violation in the superset of all possible values of ρ , the proof of robustness is a sound proof. Soundness ensures that the proposed verification framework will never produce any false positive. If a DNN is claimed to be robust against an input bound $[\eta_L, \eta_U]$, then the claim is guaranteed to be true.

IV. EVALUATION RESULTS

A. Evaluation Setups

1) *Dataset*: We verify the robustness of the DNN-based power control models for multi-cell massive MIMO networks², which are trained using the publicly available dataset [2]. We then extract a training set of 329,000 pairs of UE locations and their optimal power allocations, and another 500 samples form the test dataset, which is independent of the training dataset [7]. The dataset is generated with $J = 4$ cells, with each cell covering a square area of 250×250 m. A wrap-around topology is used to better represent interference for UEs in the edges of the network. Within each cell, there are $K = 5$ UEs at a randomly and uniformly distributed location, and at a distance that is larger than 35m from the BS. The bandwidth is $B = 20$ MHz, with the total receiver noise power $\sigma = -94$ dBm. The pilot reuse factor $\tau_p = 1$, and the maximum transmit power per UE is 20 dBm.

2) *DNN Hyperparameters, input/output properties*: Similar to previous work [7], we tested on two DNN architectures: (a) `fc.small` contain fully connected layers with $[64, 32, 32, 32, 5]$ neurons and (b) `fc.medium` contain fully connected layers with $[512, 256, 128, 128, 5]$ neurons. Two networks have 6,981 and 202,373 trainable parameters, respectively, and employ ReLU as activation units. Both models are trained on 329,000 data pairs for 50 epochs, with a batch size of 1024 and a learning rate of $3e^{-4}$ using the Adam optimizer.

The input properties for verification are constructed based on the perturbation constraint from previous adversarial attackers [7], [8]. We assume the input is in the L_∞ -norm, *e.g.*, the adversarial perturbation constraint is given as $[\eta_L, \eta_U] := [-\eta_0, \eta_0]^{J \times K \times 2}$. The level of perturbation η_0 is a scalar with values from $\{0.01, 0.1, 1, 10, 100\}$ (meter(s)) which correspond to distance perturbation of 1 cm to 100 m. The output properties, which are the maximum optimality gap ζ , are a scalar and take values from $\{0.001, 0.005, 0.01, 0.05, 0.1\}$, which guarantee to retain 90% to 99.9% of the optimally trained model. For each property, we test the local robustness of 30 randomly drawn set of UEs input locations from the test dataset. Also, we need to test $J = 4$ models for $J = 4$ cells. Thus, the total number of verification properties for the general scenario is 3,000 properties.

²The Python implementation of our proposed framework is available at https://anonymous.4open.science/r/verify_dnn_power_control

	Perturbation level η_0				
	$\pm 1\text{cm}$	$\pm 10\text{cm}$	$\pm 1\text{m}$	$\pm 10\text{m}$	$\pm 100\text{m}$
<code>fc.small</code>	0.002	0.015	0.153	1.560	22.610
<code>fc.medium</code>	0.007	0.068	0.688	7.664	341.999

TABLE I: The average L_1 -distance between ρ_L and ρ_U (mW).

3) *Evaluation Metrics: L_1 -distance between lower and upper output bound*: We measure the L_1 -based distance between the lower bound ρ_L and upper bound ρ_U computed by abstraction-based NNV [12], [14]. A smaller L_1 -distance indicates tighter bounds and more predictable model behavior.

Unsatisfiable percentage: We report the percentage of UNSAT properties, averaged across multiple test inputs $x^{(n)}$. An UNSAT result indicates that the verification tool successfully proves that the DNN model maintains the required optimality gap ζ despite input perturbations within the specified bounds $[\eta_L, \eta_U]$.

B. Robustness Analysis

We trained two architectures, `fc.medium` and `fc.small`, on a publicly available dataset to reproduce previous work [2], [7], [10], which map UE locations to maximum product SINR power allocation. Under normal operational conditions without adversarial perturbations, both models achieve competitive performance with average throughput per UE of approximately 30 Mbit/s or about 1.5bit/s/Hz, which corresponds to data from Fig. 2a in [2].

Tab. I presents the output bound analysis for both DNN architectures across different perturbation levels. The results demonstrate the relationship between input uncertainty and the bounds on output power allocation. As the perturbation level η_0 increases, the L_1 -distance between lower and upper bounds grows exponentially for both models. This behavior aligns with our expectations, where larger input perturbations lead to wider output ranges. However, despite having identical network depth, the `fc.medium` model consistently produces looser bounds compared to `fc.small`. This occurs because the wider architecture, with more neurons per layer, introduces additional complexity and nonlinearity, making abstraction-based verification more conservative in its bound estimates.

Tab. II and Tab. III present the verification success rates for both model architectures across different combinations of perturbation levels η_0 and optimality gap requirements ζ . The general trend reveals that 100% UNSAT rates are consistently achieved when both the perturbation level η_0 is sufficiently small and the tolerable optimality gap ζ is sufficiently large. The `fc.small` model consistently exhibits higher verification success rates (P_{UNSAT}) compared to `fc.medium` across most experimental configurations. Particularly for the `fc.small` model, our verification results guarantee that input perturbations below 1m can maintain 99% of optimal performance ($\zeta \leq 0.01$) with 100% confidence.

These guarantees have significant implications for practical network deployment. BS can implement anomaly-detection

	Perturbation level η_0				
	$\pm 1\text{cm}$	$\pm 10\text{cm}$	$\pm 1\text{m}$	$\pm 10\text{m}$	$\pm 100\text{m}$
$\zeta = 0.001$	100.0	100.0	6.7	0.0	0.0
$\zeta = 0.005$	100.0	100.0	80.0	0.0	0.0
$\zeta = 0.01$	100.0	100.0	100.0	0.0	0.0
$\zeta = 0.05$	100.0	100.0	100.0	70.0	0.0
$\zeta = 0.1$	100.0	100.0	100.0	100.0	0.0

TABLE II: The average percentage of UNSAT properties for model `fc.small`.

mechanisms to identify potential adversarial attacks by monitoring location reports for movement speeds exceeding specifications derived from our verification bounds. For instance, network operators can leverage the evaluation results when designing safe DNN-based power allocation for massive MIMO networks, ensuring an optimality gap of no more than 10% by constraining UE movement to less than 10m between location update intervals.

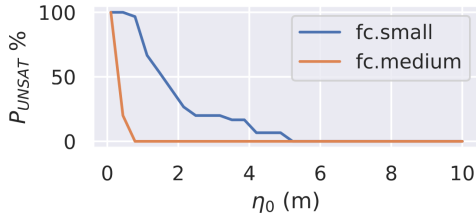


Fig. 2: Compare two models `fc.small` and `fc.medium` in terms of the average percentage of UNSAT verification properties per each perturbation level η_0 and $\zeta = 0.5\%$.

Fig. 2 provides a more fine-grained analysis of the verification results presented in the previous tables, examining verification success rates across a continuous range of perturbation levels, $\eta_0 \in [0.1, 10]$ meters, with a fixed optimality gap tolerance of $\zeta = 0.5\%$. The results demonstrate that `fc.small` maintains a high UNSAT rate for all perturbation levels up to 1m, indicating strong provable robustness within this range. In contrast, `fc.medium` exhibits a sharp decline in UNSAT percentage, degrading rapidly as perturbation levels increase beyond 0.5m. Critically, both models lose all provable robustness guarantees when perturbation levels exceed 5m, with percentage of provable properties drop to 0%. This establishes a clear operational boundary for safe deployment of these DNN-based power control systems.

V. CONCLUSION

This work proposed a formal verification framework for DNN-based power allocation in massive MIMO, providing optimality gap guarantee for DNN-based power control against adversarial location perturbations via abstraction-based bounds and feasibility checking. Evaluation results show that the smaller network architecture (`fc.small`) attains 100% verification success up to 1m while preserving 99% performance. Future works will be extended to verifying different quality of

	Perturbation level η_0				
	$\pm 1\text{cm}$	$\pm 10\text{cm}$	$\pm 1\text{m}$	$\pm 10\text{m}$	$\pm 100\text{m}$
$\zeta = 0.001$	100.0	13.3	0.0	0.0	0.0
$\zeta = 0.005$	100.0	100.0	0.0	0.0	0.0
$\zeta = 0.01$	100.0	100.0	10.0	0.0	0.0
$\zeta = 0.05$	100.0	100.0	100.0	0.0	0.0
$\zeta = 0.1$	100.0	100.0	100.0	3.3	0.0

TABLE III: The average percentage of UNSAT properties for model `fc.medium`.

service and scale verification to larger deep-learning models and more complex network settings.

ACKNOWLEDGEMENT

The authors thank the NICT AI R&D Promotion Unit for providing cloud GPU credits.

REFERENCES

- [1] E. Björnson, J. Hoydis, L. Sanguinetti, et al., “Massive MIMO networks: Spectral, energy, and hardware efficiency,” *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [2] L. Sanguinetti, A. Zappone, and M. Debbah, “Deep learning power allocation in massive MIMO,” in *2018 52nd Asilomar conference on signals, systems, and computers*, IEEE, 2018, pp. 1257–1261.
- [3] S. Chakraborty, E. Björnson, and L. Sanguinetti, “Centralized and distributed power allocation for max-min fairness in cell-free massive MIMO,” in *2019 53rd Asilomar conference on signals, systems, and computers*, IEEE, 2019, pp. 576–580.
- [4] M. Zaher, Ö. T. Demir, E. Björnson, and M. Petrova, “Learning-based downlink power allocation in cell-free massive MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 174–188, 2023.
- [5] T. Van Chien, T. N. Canh, E. Björnson, and E. G. Larsson, “Power control in cellular massive MIMO with varying user activity: A deep learning solution,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5732–5748, 2020.
- [6] X. Zhang, V. A. Le, M. Kaneko, J. C. Lui, and Y. Ji, “Multi-Agent Deep Reinforcement Learning-based Uplink Power Control in Cell-Free Massive MIMO with Mobile Users,” *IEEE Transactions on Vehicular Technology*, 2025.
- [7] B. Manoj, M. Sadeghi, and E. G. Larsson, “Adversarial attacks on deep learning based power allocation in a massive MIMO network,” in *ICC 2021-IEEE International Conference on Communications*, IEEE, 2021, pp. 1–6.
- [8] P. M. Santos, B. Manoj, M. Sadeghi, and E. G. Larsson, “Universal adversarial attacks on neural networks for power allocation in a massive MIMO system,” *IEEE Wireless Communications Letters*, vol. 11, no. 1, pp. 67–71, 2021.
- [9] R. Sahay, M. Zhang, D. J. Love, and C. G. Brinton, “Defending adversarial attacks on deep learning-based power allocation in massive MIMO using denoising autoencoders,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 4, pp. 913–926, 2023.
- [10] B. Manoj, M. Sadeghi, and E. G. Larsson, “Downlink power allocation in massive MIMO via deep learning: Adversarial attacks and training,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 707–719, 2022.
- [11] A. Athalye, N. Carlini, and D. Wagner, “Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples,” in *International conference on machine learning*, PMLR, 2018, pp. 274–283.
- [12] G. Singh, T. Gehr, M. Püschel, and M. Vechev, “An abstract domain for certifying neural networks,” *Proceedings of the ACM on Programming Languages*, vol. 3, no. POPL, pp. 1–30, 2019.
- [13] J. Cohen, E. Rosenfeld, and Z. Kolter, “Certified adversarial robustness via randomized smoothing,” in *International conference on machine learning*, PMLR, 2019, pp. 1310–1320.
- [14] K. Xu et al., “Automatic perturbation analysis for scalable certified robustness and beyond,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 1129–1141, 2020.
- [15] C. Brix, S. Bak, T. T. Johnson, and H. Wu, “The Fifth International Verification of Neural Networks Competition (VNN-COMP 2024): Summary and Results,” *arXiv preprint arXiv:2412.19985*, 2024.
- [16] J. Kim, H.-S. Lim, and K. Choi, “Certified Robustness of Antenna Selecting Neural Networks for Massive MIMO Wireless Communications,” *IEEE Access*, 2025.
- [17] H. Duong, T. Nguyen, and M. B. Dwyer, “NeuralSAT: A High-Performance Verification Tool for Deep Neural Networks,” in *Computer Aided Verification (CAV 2025)*, R. Piskac and Z. Rakamarić, Eds., ser. Lecture Notes in Computer Science, vol. 15932, Springer, Cham, 2025, pp. 409–423. DOI: 10.1007/978-3-031-98679-6_19