

Research Article

Probe Selection and Power Weighting in Multiprobe OTA Testing: A Neural Network-Based Approach

Yong Li ¹, Hao Sun ¹, Xingyu Chen ², Lijian Xin ¹ and Xiang Zhang ³

¹Key Laboratory of Universal Wireless Communications (Ministry of Education),
Beijing University of Posts and Telecommunications, Beijing 100876, China

²Department of Electronic and Information Engineering, College of Electronic Information and Optical Engineering,
Nankai University, Tianjin 300350, China

³China Academy of Information and Communications Technology, Beijing 100191, China

Correspondence should be addressed to Xiang Zhang; zhangxiang1@caict.ac.cn

Received 21 May 2019; Revised 25 July 2019; Accepted 30 July 2019; Published 10 September 2019

Academic Editor: Ahmed Toaha Mobashsher

Copyright © 2019 Yong Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Over-the-air (OTA) radiated testing is an efficient solution to evaluate the performance of multiple-input multiple-output (MIMO) capable devices, which can emulate realistic multipath channel conditions in a controlled manner within lab environment. In a multiprobe anechoic chamber- (MPAC-) based OTA setup, determining the most appropriate probe locations and their power weights is critical to improve the accuracy of channel emulation at reasonable system costs. In this paper, a novel approach based on neural networks (NNs) is proposed to derive suitable angular locations as well as power weights of OTA probe antennas; in particular, by using the regularization technique, active probe locations and their weights can be optimized simultaneously with only one training process of the proposed NN. Simulations demonstrate that compared with the convex optimization-based approach to perform probe selection in the literature, e.g., the well-known multishot algorithm, the proposed NN-based approach can yield similar channel emulation accuracy with significantly reduced computational complexity.

1. Introduction

To efficiently evaluate the performance of multiple-input multiple-output (MIMO) capable devices, the over-the-air (OTA) radiated testing has evolved to be capable of emulating realistic spatial channel characteristics in a controlled manner [1]. Among various candidate methods for MIMO OTA testing, the multiprobe anechoic chamber (MPAC) method is considered as the most promising solution to reproduce the desired multipath channel profiles. In a typical MPAC setup, a number of OTA probe antennas are mounted inside an anechoic chamber. Channel profiles in spectral, polarimetric, temporal, and spatial dimensions can be reconstructed with adequate accuracy in the central area of the anechoic chamber, referred to as the test volume, where the device under test (DUT) is placed. Specifically, the probe antennas are employed to synthesize the cluster-specific power angular spectrum (PAS) shape, which will

essentially determine spatial characteristics among DUT antennas. Therefore, proper power weighting for probe antennas is one important issue for the MPAC setup. In [2], spatial correlation was selected as the criterion to determine probe antenna weights, which aims to minimize the difference between the emulated spatial correlations and the target spatial correlations. Two objective functions were further proposed in [3]; one is to minimize the sum of emulation errors over a set of DUT sampling locations, and the other is to minimize the maximum emulation error within the test volume. Additional constraints on the resulting discrete PAS shape were introduced in [4] to provide a more accurate and realistic spatial characterization; specifically, an optimization toolbox, e.g., the popular convex problem solver CVX [5], can be readily utilized to solve the formulated convex optimization problem.

However, a major challenge for the MPAC setup consists in its high cost because each active probe antenna, allocated

with nonzero power and thus contributing in synthesizing the target channel, will be connected to an output port of an expensive channel emulator. Since not all probe antennas are required to be active to synthesize a particular channel model due to the fact that practical channel models are typically directional, a flexible setup which selects a subset of candidate probe antennas to be active may be more cost-efficient by reducing the number of required hardware resources of the channel emulator, which constitutes a major cost factor for the MPAC setup. In such a flexible MPAC setup, a large number of probe antennas are installed with fixed angular locations and a subset of them will be selected to be driven by a switch box, depending on the target channel model under consideration. Several probe selection algorithms, e.g., the brute force algorithm [6], the one-shot algorithm [6], the multishot algorithm [6, 7], and the successive probe cancellation algorithm [6], are proposed in the literature. All these algorithms rely on the convex optimization framework proposed in [4]. Particularly, the brute force algorithm will perform convex optimization for all possible combinations of active antennas and thus becomes computationally prohibitive when the number of candidate antennas is large; the one-shot algorithm selects a plurality of probes with the largest power weights after convex optimization is performed with all candidate probes, which is the simplest but obviously provides the worse performance; the multishot algorithm and the successive probe cancellation algorithm remove the probe(s) with the least and the most contributions, respectively, and perform convex optimization repeatedly with the remaining probes. It is shown that the multishot algorithm can provide good tradeoff between emulation accuracy and computational complexity [6]. However, as the state-of-the-art MPAC setup has evolved from two-dimensional (2D) probe configuration to three-dimensional (3D) probe configuration to account for channel dispersions in the elevation domain as well as in the azimuth domain, far more probe antennas will be equipped in the anechoic chamber as candidate probes. Particularly, probe antennas operating in high-frequency bands, e.g., 28 GHz, which are promising for the upcoming 5G systems, are inexpensive to manufacture; therefore, it will become economically feasible to install hundreds of probes in the MPAC setup [8], and thus, selecting only a few probes from hundreds of candidates would become a huge challenge. Furthermore, dynamic channel emulations in which the spatial profiles of the target channel may vary constantly have become an emerging demand, which requires quick adjustments of probe parameters. Hence, a method which can achieve the optimal results of probe selection and power weighting more efficiently is urgently required.

Neural networks (NNs), capable of capturing with a high degree of accuracy the underlying complicated relationship between input and output without a priori assumption of the knowledge between them, regained massive amounts of attention in the past couple of years in a wide range of technological disciplines. Recently, NNs have been introduced into the domain of wireless communications to solve a variety of challenging problems [9, 10]. For example,

NNs were employed as a powerful tool in the design of antenna array in [11, 12]. In this paper, a NN-based approach is proposed to jointly solve the problems of probe selection and power weighting, which can be casted into a simple linear NN model with ℓ_0 norm regularization which tends to produce sparse solutions for NN weights. A remarkable advantage of the proposed NN-based approach lies in that once the training of the NN is finished, the subset of the selected active probes as well as their power weights can be known; that is, the structure and the weights of the considered NN model can be optimized in one go, which implies significant improvement in computational efficiency. Note that, on the contrary, power weights have to be recomputed repeatedly by using convex optimization after the candidate set of probes are updated in the aforementioned multishot algorithm. Therefore, the proposed NN-based approach can be more efficient than the convex optimization-based multishot algorithm, which will be verified by our simulations presented later. To our best knowledge, this work is the first reported effort to introduce the NN technique into multiprobe OTA testing.

2. Probe Selection and Power Weighting

A state-of-the-art 3D probe configuration is assumed for the MPAC setup under consideration. A pair of virtual antenna elements u and v with isotropic radiation patterns is used to sample the test volume. The same sampling method as used in [3] is adopted such that u and v are located at the opposite positions of the ellipsoid surface enclosing the test volume. Suppose that M sampling pairs are obtained in the test volume. Given a realistic spherical power spectrum (SPS) known as $P(\theta, \phi) = P(\theta)P(\phi)$, which should satisfy $\int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} P(\theta)P(\phi)\cos(\theta)d\phi d\theta = 1$, the target spatial correlation between the m th sampling pair can be written as

$$\rho_m = \oint \exp\left(j\frac{2\pi}{\lambda}(\mathbf{r}_{u,m} - \mathbf{r}_{v,m}) \cdot \Omega\right)P(\Omega)d\Omega, \quad (1)$$

where λ denotes the wavelength, $\mathbf{r}_{u,m}$ and $\mathbf{r}_{v,m}$ denote, respectively, the location vectors of antenna elements u and v in the m th sampling pair, Ω denotes the solid angle vector, and (\cdot) denotes the dot product operator.

Let N denote the number of candidate probe antennas, and let $\Phi \triangleq [\Phi_1, \dots, \Phi_N]^T$ denote the vector indicating the known angular locations of the N probe antennas. The contribution from the n th probe in the spatial correlation at the m th antenna pair can be given by $\hat{\rho}_{m,n} = w_n\beta_{m,n}$, where w_n is the real-valued power weight of the n th probe and

$$\beta_{m,n} \triangleq \exp\left(j\frac{2\pi}{\lambda}(\mathbf{r}_{u,m} - \mathbf{r}_{v,m}) \cdot \Phi_n\right). \quad (2)$$

Finally, the emulated spatial correlation of the m th antenna pair can be given by

$$\hat{\rho}_m = \mathbf{w}^T \boldsymbol{\beta}_m, \quad (3)$$

where $\mathbf{w} = [w_1, \dots, w_N]^T$ and $\boldsymbol{\beta}_m = [\beta_{m,1}, \dots, \beta_{m,N}]^T$.

Note that the ℓ_1 norm and the nonnegative constraints should apply for the weight vector \mathbf{w} , therefore, we have

$$\|\mathbf{w}\|_1 = 1, \quad 0 \leq w \leq 1, \quad (4)$$

where $\|\cdot\|_1$ denotes the ℓ_1 norm. Furthermore, the aim of probe selection is to find the optimal subset of probes whose number should not be larger than the number of available output ports of the channel emulator denoted by K ; thus, we further have

$$\|\mathbf{w}\|_0 \leq K, \quad (5)$$

where the ℓ_0 norm operation $\|\cdot\|_0$ denotes the number of nonzero entries in the vector.

For M sampling locations, the goal is to minimize the sum of squared emulation errors. Therefore, the problem can be formulated as

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{E}(\mathbf{w}) \triangleq \|\hat{\boldsymbol{\rho}}(\mathbf{w}) - \boldsymbol{\rho}\|_2^2, \\ \text{s.t.} \quad & (4) \text{ and } (5), \end{aligned} \quad (6)$$

where $\hat{\boldsymbol{\rho}} = [\hat{\rho}_1, \dots, \hat{\rho}_M]^T$, $\boldsymbol{\rho} = [\rho_1, \dots, \rho_M]^T$, and $\|\cdot\|_2$ denotes the ℓ_2 norm operation. Note that (6) is a combinatorial nonconvex optimization problem which is NP-hard due to the ℓ_0 norm constraint in (5). The brute force method is obviously computationally prohibitive when N is large.

3. NN-Based Approach

3.1. Proposed NN Model. Although (6) has already been treated in a convex optimization framework as in [4], in this work, we will propose a NN-based approach to solve this problem. Towards this purpose, a simple (however, the ℓ_0 norm constraint given by (5) is nontrivial and needs to be carefully addressed) linear NN model (A linear NN model without hidden layers is adopted in this work due to the following reasons. Firstly, the relationship between the spatial correlation ρ and probe weights \mathbf{w} is a straightforward linear function; thus, our proposed NN model is capable of capturing this relationship. Secondly, from the point of computational complexity, the computational complexity increases as the structure of the NN model becomes more sophisticated. This is because the gradient descent method is based on the so-called ‘‘back propagation’’ algorithm, where the gradients should be calculated across all the layers. Last but not the least, the weight of each neuron at the input layer has its practical meaning, which can be immediately translated to the power weight of the corresponding probe. This advantage would not be available if additional hidden layers were introduced into the NN model.) can be used, as illustrated in Figure 1. The proposed NN model has one input layer consisting of N neurons, which have the same number as the candidate probe antennas in the OTA setup under consideration, and one output layer consisting of one neuron. The overall input is $\boldsymbol{\beta}$, which is a N -dimensional vector. The input on the n th neuron is first multiplied with a weight w_n , and then, all weighed inputs are summed up to generate the output $\hat{\rho}$, which is expected to approximate ρ . Therefore, the training data set for the proposed NN model can be defined as $\mathcal{M} = \{(\boldsymbol{\beta}_m, \rho_m) | m = 1, 2, \dots, M\}$, which has the same size as the sampling locations within the test volume. When the NN model is converged after the training

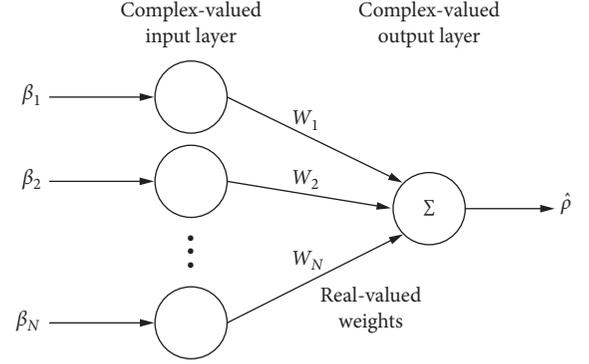


FIGURE 1: Proposed NN model for probe selection and power weighting.

process is finished, the set of weights from the input layer to the output layer, i.e., $\{\mathbf{w}_k | k = 1, 2, \dots, N\}$, is the proper power weighting result for the N probe antennas. Furthermore, due to the regularization applied in the training process, there should be no more than K nonzero weights among the resulting N weight values.

Note that compared with a usual NN model, the non-linear activation function is omitted in Figure 1 due to the linear nature indicated by (3). Also note that both the input and the output of the proposed NN model are complex-valued while the weights between them are real-valued.

3.2. Training of the Proposed NN Model. In the following, we will provide a training method suitable for the proposed NN model.

3.2.1. Regularization. It is well known that ℓ_p norm regularization is a pruning method to simultaneously optimize the structure and weights of NNs [13]. By appending a regularization term to the error function $\mathbf{E}(\mathbf{w})$ in (6), the modified error function can be given by $\mathbf{E}(\mathbf{w}) + \lambda \|\mathbf{w}\|_p^p$, where $\lambda > 0$ is the regularization parameter to balance the tradeoff between training accuracy and structural complexity of the NN and $\|\cdot\|_p$ is the ℓ_p norm. It has been shown that, when $0 \leq p \leq 1$, the ℓ_p regularization method can produce sparse solution; particularly, the smaller the p , the sparser the solution [14]. However, it is challenging to solve the ℓ_0 norm due to its NP-hard nature; instead, the $\ell_{1/2}$ norm regularizer has proved to possess the representativeness among all ℓ_p regularizers for $0 < q < 1$ and has attracted intensive research interests on sparse modeling recently [15]. Thus, by appending the $\ell_{1/2}$ norm regularizer, the error function in (6) can be rewritten as

$$\boldsymbol{\varepsilon}(\mathbf{w}) = \mathbf{E}(\mathbf{w}) + \lambda \|\mathbf{w}\|_{1/2}^{1/2}. \quad (7)$$

3.2.2. Gradient. Let $\beta_{m,n}$, $\hat{\rho}_m$, and $\varepsilon_m \triangleq \hat{\rho}_m - \rho_m$ denote the n th input, the output, and the complex-valued error for the m th sample in the training data set, respectively; defining $E_m \triangleq \varepsilon_m \varepsilon_m^*$ is the contribution of the m th sample in $\mathbf{E}(\mathbf{w})$.

The gradient of E_m with respect to \mathbf{w} , i.e., $(\partial E_m(\mathbf{w}))/\partial \mathbf{w}$, can be given by $[(\partial E_m(\mathbf{w}))/(\partial w_1), \dots, (\partial E_m(\mathbf{w}))/(\partial w_N)]^T$.

According to the chain rule, we have

$$\begin{aligned} \frac{\partial E_m(\mathbf{w})}{\partial w_n} &= \frac{\partial \{\Re^2[\varepsilon_m] + \Im^2[\varepsilon_m]\}}{\partial w_n} \\ &= 2\Re[\varepsilon_m] \frac{\partial \Re[\varepsilon_m]}{\partial w_n} + 2\Im[\varepsilon_m] \frac{\partial \Im[\varepsilon_m]}{\partial w_n} \\ &= 2\Re[\varepsilon_m] \Re[\beta_{m,n}] + 2\Im[\varepsilon_m] \Im[\beta_{m,n}] \\ &= \varepsilon_m^* \beta_{m,n} + \varepsilon_m \beta_{m,n}^* \end{aligned} \quad (8)$$

where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real part and the imaginary part of a complex number, respectively.

Depending on the subset of training data used for each iteration in the NN training process, denoted as $\mathcal{M}' \subseteq \mathcal{M}$, the gradient of (7) with respect to \mathbf{w} can be calculated as

$$\frac{\partial E(\mathbf{w})}{\partial w_k} = \sum_{m \in \mathcal{M}'} \frac{\partial E_m(\mathbf{w})}{\partial w_k} + \frac{\lambda}{2\sqrt{w_k}}. \quad (9)$$

Three well-known learning methods to choose \mathcal{M}' include stochastic gradient descent (SGD), batch gradient descent (BGD), and mini-batch gradient descent (MBGD). The SGD method uses only one sample from the training data set in each iteration; the BGD method uses all samples in each iteration; finally, the MBGD method uses a subset of samples, which can provide good tradeoff between training efficiency and convergence and is adopted here in our study.

3.2.3. Weight Update. Compared with the standard gradient descent, gradient descent with momentum, which computes an exponentially weighted average of gradients to update weights, is shown to work faster. The momentum is updated as

$$\mathbf{V}(t) = \gamma \mathbf{V}(t-1) + \eta \frac{\partial E(\mathbf{w})}{\partial \mathbf{w}}, \quad (10)$$

where η is the momentum parameter to control the exponentially weighted average and η is the learning rate. Finally, the weights can be updated as

$$\mathbf{w}(t) = \mathbf{w}(t-1) - \mathbf{V}(t). \quad (11)$$

Furthermore, as the regularization parameter λ is used to control the sparsity of weights, its appropriate value is unknown at the beginning of training and has to be adapted in the training process. The training process for the proposed NN model can be summarized as in 1.

In Algorithm 1, T is the total number of iterations for training the NN, Δ_T is the interval in the number of iterations to update λ , Δ_λ is the step to increase λ in the case that more nonzero weights than K are obtained. It is worth noting that the *slow start* strategy is adopted to gradually increase λ to avoid the phenomenon of over-regularization in the case of more than K remaining nonzero weights, whereas the *exponential backoff* strategy is used to decrease λ to help recover from over-regularization quickly in the case

of less than K remaining nonzero weights. Furthermore, the operator $[\cdot]^+$ denotes the projection onto the nonnegative orthant, which is used to guarantee the nonnegative constraint for \mathbf{w} in (4), equally importantly, since a smaller (and therefore less significant) w_n tends to yield a larger gradient component in its direction, as indicated by the second term on the right-hand side of (9), due to the existence of the regularization term, and thus is more likely to become less than zero after being updated by (11), forcing such weight to be equal to zero can avoid its oscillations near zero and lead to quick convergence.

3.3. Complexity Analysis. The computational complexity will determine the potential applicability of the proposed method, especially for dynamic channel emulation. In this subsection, we will choose the well-known multishot algorithm as the baseline and compare the computational complexity of both methods.

The multishot algorithm performs convex optimization based on the objective function defined in (6) and removes the probe with the least power weight at each shot, until the number of the remaining probes reaches the target. It has proved that the optimization problem in (6) is essentially a second-order cone programming (SOCP) problem [16]. In the equivalent SOCP problem corresponding to the i th shot, the second-order cone will have the dimension of $M+1$ and one equality constraint with the dimension of $N-i+1$. Thus, the worst-case computational complexity in the i th shot of the multishot algorithm can be given by [17]

$$\mathcal{O}\left((N-i+1)^3 + (N-i+1)^2(M+1) + (M+1)^2\right). \quad (12)$$

It is obvious that the computational complexity of the multishot algorithm, in proportion to the cube power of the probe number, is rather high and will become unaffordable for 3D probe configurations with hundreds of candidate probes.

As a comparison, in our proposed method, the main computational overhead is the gradient calculation at each iteration. Thus, the complexity of the proposed method can be given by $\mathcal{O}(SN)$, where S denotes the size of the mini-batch. Therefore, the proposed method has a much lower computational complexity compared with the multishot algorithm and is more suitable for 3D probe selection application.

4. Simulations and Discussions

In this section, simulation results are provided to verify the performances of our proposed method in terms of both emulation accuracy and computational efficiency. The multi-shot algorithm is used for comparison as the representative of convex optimization based methods.

4.1. Simulation Configurations. Given the directional and specular channel characteristics of the millimeter-wave (mmWave) frequency bands intended for the usage of the fifth generation (5G) mobile communication systems, a sectorized MPAC setup as proposed in [18] is considered in

```

Input: Training data set  $\mathcal{M}$ 
Output: Weight vector  $\mathbf{w}$ 
(1) Initialize: the iteration index  $t = 1$ ,  $\mathbf{w}(1)$ ,  $\mathbf{V}(1)$ ,  $\lambda$ ,  $\eta$ ,  $\gamma$ .
(2) while  $t < T$  do
(3)   Randomly select a mini-batch of size  $S$  from  $\mathcal{M}$ ;
(4)   Calculate the gradient and the momentum according to (9) and (10), respectively;
(5)   Update  $\mathbf{w}$  according to (11);
(6)   Adjust  $\mathbf{w}$ :  $\mathbf{w} = [\mathbf{w}]^+$ , and  $\mathbf{w} = \mathbf{w}/\|\mathbf{w}\|_1$ ;
(7)   if  $t \bmod \Delta_T = 0$  then
(8)     if  $\|\mathbf{w}(t)\|_0 < K$  then
(9)        $\lambda = \lambda/2$ 
(10)    else if  $\|\mathbf{w}(t)\|_0 > K$  then
(11)       $\lambda = \lambda + \Delta_\lambda$ 
(12)    else if
(13)  end if
(14)  Update  $t = t + 1$ .
(15) end while

```

ALGORITHM 1: Training process for the proposed NN model.

this work, which is illustrated in Figure 2. Unlike the conventional MPAC OTA setup of the 2D or 3D probe ring structure, the DUT, positioned within the test volume, is placed at the edge of the anechoic chamber. The distance from each point on the sectorized probe wall to the center of the test volume is equal. Apparently, the sectorized MPAC setup can fully utilize the space of the chamber and reduce required hardware resources such as probe antenna to allow a more cost-efficient solution for DUTs with large dimension, e.g., 5G massive MIMO devices.

Two different probe configurations are assumed. In the first probe configuration, the probe wall covers 60° in elevation direction and 120° in horizontal direction and the spacing between adjacent probes is 10° ; thus, $N = 91$ candidate probes are available. In the second probe configuration, the coverage range of the probe wall is the same, but the probe spacing is reduced to 5° ; thus, $N = 325$ candidate probes are available. The radius of the test volume is set to be 2.5λ , and there are $M = 2592$ sampling antenna pairs in the test volume. A uniform initial weight vector is chosen such that $\mathbf{w}_n(1) = (1/N)$ for any $n \in \{1, \dots, N\}$. The parameters used in the NN-based approach are listed in Table 1, unless otherwise stated.

4.2. Execution Speed. In this subsection, practical execution speeds for the proposed method and the multishot algorithm are compared. The practical execution times are tracked by using the Matlab™ built-in Profiler command. To concentrate on the execution time of the two methods, the initialization stage for obtaining the training data set is excluded from the recorded execution time. Two cases are investigated. In Case A, the target channel composed of a single cluster following truncated Laplacian distribution is to be reconstructed under the first probe configuration and the target probe number is $K = 8$; in Case B, the CDL-A channel model defined in [19] is chosen as the target channel for a more complicated channel environment under the second probe configuration and the target probe number is $K = 16$.

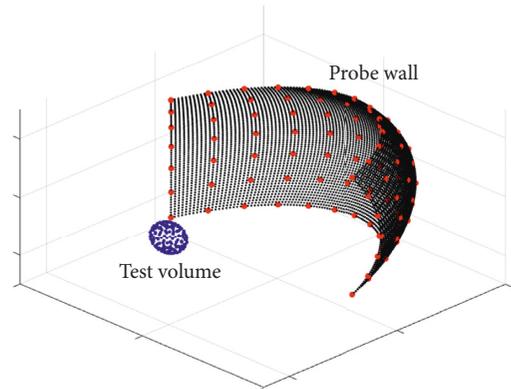


FIGURE 2: Sectorized MPAC OTA setup. The black spherical surface denotes the probe wall, and each red dot on the probe wall denotes a probe antenna. The volume enclosed by blue dots denotes the test volume. Note that the test volume is enlarged for better visualization.

It can be expected that due to the randomness in the MBGD method, the results from different runs of the NN training processes may vary in the proposed method. Thus, five independent runs are performed for both methods in both cases, whose execution time as well as accuracy performance in terms of root-mean-square deviation (RMSD) is shown in Table 2. It can be observed that there are variations in the RMSD results from the proposed NN-based approach due to the nature of the MBGD method used in the training process, which randomly selects a subset from the training data set in each iteration.

We then concentrate on the comparisons of the two methods. It can be observed that the emulation accuracy of the proposed NN-based approach can be very close to or even better than that of the multishot algorithm, but its computational complexity in terms of execution time is several magnitude orders less than that of the counterpart. In particular, note that the execution times of the multishot algorithm increase more sharply than the NN-based

TABLE 1: Parameters used in the NN-based approach.

Symbol	Meaning	Value
T	Iteration number	10,000
S	Batch size	128
H	Learning rate	0.00001
Γ	Momentum parameter	0.99
$\lambda(1)$	Initial regularization parameter	0.8
Δ_T	Update interval for λ	400
Δ_λ	Step for adjusting λ	0.05

TABLE 2: Execution times and RMSD results for different methods.

		Run	1	2	3	4	5
Case A	Multishot	RMSD	0.0338	0.0338	0.0338	0.0338	0.0338
		Execution time (s)	300.7	299.4	299.1	301.5	310.4
	Proposed	RMSD	0.0344	0.0266	0.0334	0.0339	0.0277
		Execution time (s)	5.1	4.9	5.2	5.5	5.0
Case B	Multishot	RMSD	0.0262	0.0262	0.0262	0.0262	0.0262
		Execution time (s)	8673.3	9311.5	9168.2	9038.6	8937.1
	Proposed	RMSD	0.0223	0.0292	0.0261	0.0389	0.0258
		Execution time (s)	37.4	36.4	36.3	35.6	34.0

approach when the scale of the problem is increased from Case A to Case B. This is because the additional computational complexity for the multishot algorithm lies in both the increased number of iterations and the increased scale of optimization problem in each iteration, the latter typically involving an increase of computational complexity in a superlinear order; on the contrary, the additional complexity for the NN-based approach only comes from the linearly increased number of weights in the training process. It is worth noting that the comparison in practical execution times is consistent with our complexity analysis in Section 3.3.

4.3. Robustness. In this subsection, multiple target channel models with different SPS shapes are used to verify the robustness of the proposed method; in particular, target channel models with a single cluster and more practical multiple clusters are both considered. The second probe configuration with denser probe spacings is assumed.

For the single-cluster cases, the target clusters are randomly generated with different shape parameters to investigate the applicability of the proposed method in arbitrary channel environments. As shown in Table 3, six different single-cluster channel models with different SPS shapes are generated, all of which follow the truncated Laplacian distribution with various angles of arrival and angle spreads. The target probe number is $K = 8$ in the single cluster cases. The RMSD results in Table 3 indicate that the proposed method can achieve similar accuracy when compared with the multishot algorithm in the single-cluster cases. For the multicluster cases, the clustered delay line (CDL) channel models under non-line-of-sight (NLOS) scenarios defined in [19] are assumed. It is noted that the clusters whose arrival angles are beyond the probe wall coverage are ignored and the target probe number is $K = 16$.

The simulation results in Table 4 suggest that quite high accuracy can be obtained by the proposed method when compared to the multishot algorithm. Thus, it can be concluded that the proposed approach is suitable for general channel environments.

4.4. Training Process. In order to obtain a deeper insight of how our proposed method works, one example of the training process for the proposed NN model is illustrated in Figures 3–5. Figure 3 illustrates the adjustment of the regularization parameter λ in the training process, and Figure 4 demonstrates the variation of the number of nonzero weights. It seems that the initial λ is too large such that far less active probes than available ($K = 16$ in this case) are selected in the initial stage since the regularizer under large λ would impose a superfluous constraint on sparsity. The training process will then reduce λ to its half, which will in turn allow more weights to become nonzero to approach the available number of probes. It is shown that the regularization parameter λ and thus the number of nonzero probes can achieve their convergence after only a few adjustments of λ , so does the emulation error which finally converges to 0.083 in this example training process. Once a training process is finished, a set of nonzero weights, each of which corresponds to an active probe and its power weight, can be obtained as a local optimum. A plural of training processes can be run to further improve the accuracy performance. However, as shown in Table 2, even one training process can provide performance on the same order as the multishot algorithm with much reduced computational complexity in almost all cases.

Finally, as demonstrated in Figures 4 and 5, the emulation error fluctuates more violently when the number of nonzero probes changes and becomes stable as the number of nonzero probes goes still, due to fixed learning parameters

TABLE 3: RMSD results under single-cluster cases.

Cluster	1	2	3	4	5	6
AOA	2.7	22.4	3.2	-5.5	22.0	26.5
ASA	3	10	15	20	24	30
ZOA	106.9	112.9	105.2	78.33	84.6	107.4
ZSA	3	8	7	8	10	15
Multishot	0.0184	0.0923	0.0401	0.0620	0.0671	0.1215
Proposed	0.0227	0.0795	0.0642	0.0455	0.0810	0.1084

TABLE 4: RMSD results under multicluster cases.

Channel	CDL-A	CDL-B	CDL-C
Multishot	0.0366	0.0760	0.0301
Proposed	0.0347	0.0669	0.0251

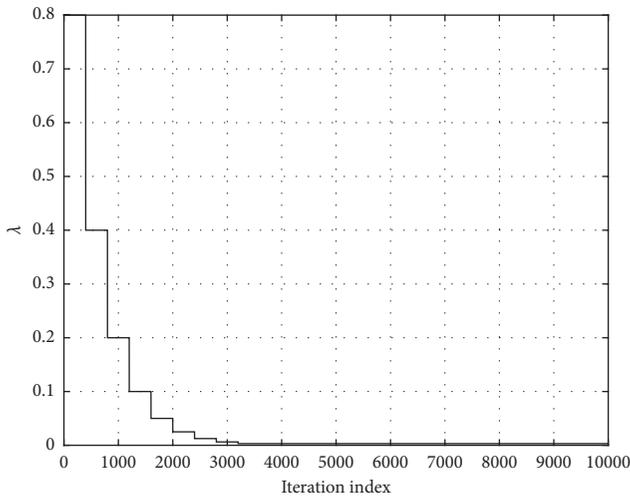
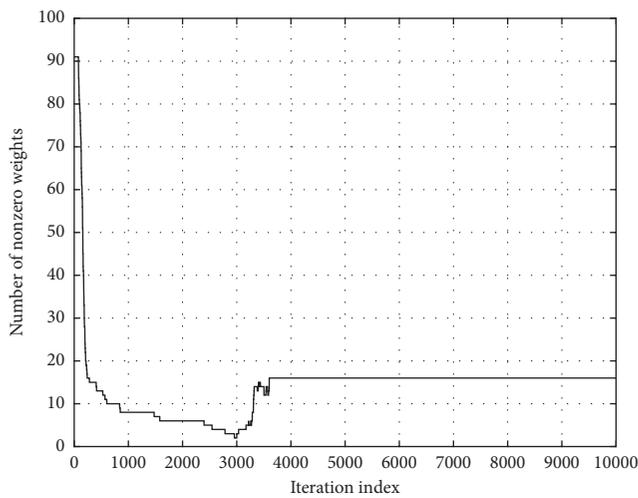
FIGURE 3: Regularization parameter λ vs. iteration number.

FIGURE 4: Number of nonzero weights vs. iteration number.

in the training process. If the learning parameters, e.g., the learning rate η , can be adjusted in an adaptive manner, the convergence of the proposed NN model may be more

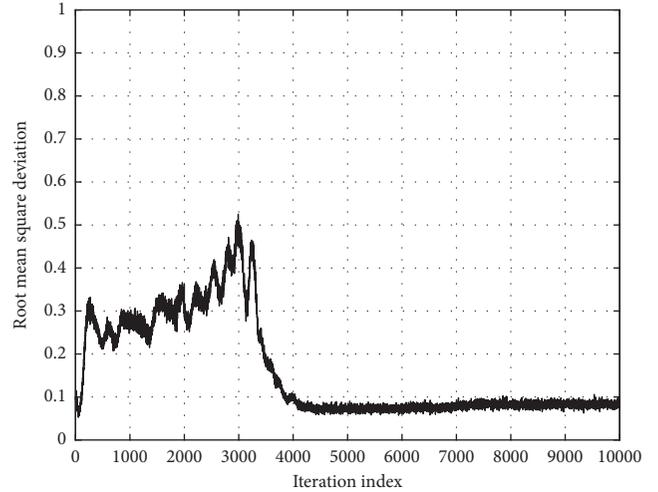


FIGURE 5: Emulation error vs. iteration number.

smooth and quick. Fine tuning of the training parameters in the NN model will be a topic in our future work.

5. Conclusions

A NN-based approach is proposed to address the joint problems of probe selection and power weighting in multiprobe OTA testing, which was conventionally treated in a convex optimization-based framework. A NN model and its training method are detailed; in particular, the probe selection problem is considered as a sparsity constraint and solved by introducing an $\ell_{1/2}$ norm regularizer. When the training process of the proposed NN is completed, active probe locations and their weights can be obtained simultaneously. Compared with the classic multishot algorithm which iteratively makes use of convex optimization, the proposed NN-based approach can yield similar channel emulation accuracy but at significantly reduced computational complexity.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61971067, in part by the State Major Science and Technology Special Projects under Grant 2018ZX03001028-003, and in part by Beijing Municipal Science and Technology Project under Grant Z191100007519002. The work of L. Xin was supported by the BUPT Excellent Ph.D Students Foundation under Grant CX2018103.

References

- [1] M. Rumney, R. Pirkel, M. H. Landmann, and D. A. Sanchez-Hernandez, "MIMO over-the-air research, development, and testing," *International Journal of Antennas and Propagation*, vol. 2012, Article ID 467695, 8 pages, 2012.
- [2] P. Kyösti, T. Jämsä, and J.-P. Nuutinen, "Channel modelling for multiprobe over-the-air MIMO testing," *International Journal of Antennas and Propagation*, vol. 2012, Article ID 615954, 11 pages, 2012.
- [3] W. Fan, X. Carreño, J. Nielsen et al., "3D channel emulation in multi-probe setup," *Electronics Letters*, vol. 49, no. 9, pp. 623–625, 2013.
- [4] W. Fan, X. Carreno Bautista de Lisbona, F. Sun, J. O. Nielsen, M. B. Knudsen, and G. F. Pedersen, "Emulating spatial characteristics of MIMO channels for OTA testing," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 8, pp. 4306–4314, 2013.
- [5] CVX Research, "Inc. CVX: Matlab software for disciplined convex programming, version 2.0," April 2011, <http://cvxr.com/cvx>.
- [6] W. Fan, F. Sun, J. O. Nielsen et al., "Probe selection in multiprobe OTA setups," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 4, pp. 2109–2120, 2014.
- [7] W. Fan, I. Szini, J. O. Nielsen, and G. F. Pedersen, "Channel spatial correlation reconstruction in flexible multiprobe setups," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, no. 4, pp. 1724–1727, 2013.
- [8] P. Kyösti, L. Henttilä, W. Fan, J. Lehtomäki, and M. Latva-aho, "On radiated performance evaluation of massive MIMO devices in multiprobe anechoic chamber OTA setups," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 10, pp. 5485–5497, 2018.
- [9] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2017.
- [10] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2595–2621, 2018.
- [11] J. Dudczyk and A. Kawalec, "Adaptive forming of the beam pattern of microstrip antenna with the use of an artificial neural network," *International Journal of Antennas and Propagation*, vol. 2012, Article ID 935073, 13 pages, 2012.
- [12] R. G. Ayestaran, "Fast near-field multifocusing of antenna arrays including element coupling using neural networks," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 7, pp. 1233–1237, July 2018.
- [13] H. Zhang, W. Wu, F. Liu, and M. Yao, "Boundedness and convergence of online gradient method with penalty for feedforward neural networks," *IEEE Transactions on Neural Networks*, vol. 20, no. 6, pp. 1050–1054, 2009.
- [14] G. Gnecco and M. Sanguineti, "Regularization techniques and suboptimal solutions to optimization problems in learning from data," *Neural Computation*, vol. 22, no. 3, pp. 793–829, 2010.
- [15] Z. Xu, X. Chang, F. Xu, and H. Zhang, " $L_{1/2}$ regularization: a thresholding representation theory and a fast solver," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 7, pp. 1013–1027, 2012.
- [16] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and Its Applications*, vol. 284, no. 1–3, pp. 193–228, 1998.
- [17] S. Tomic, M. Beko, and R. Dinis, "RSS-based localization in wireless sensor networks using convex relaxation: non-cooperative and cooperative schemes," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2037–2050, 2015.
- [18] P. Kyösti, W. Fan, G. F. Pedersen, and M. Latva-aho, "On dimensions of OTA setups for massive MIMO base stations radiated testing," *IEEE Access*, vol. 4, pp. 5971–5981, Sept. 2016.
- [19] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," TR 38.901 V14.2.0, Springer, Berlin, Germany, 2017, <http://www.3gpp.org>.

