

A DEEP REINFORCEMENT LEARNING APPROACH FOR AUTOMATED CHAMBER CONFIGURATION REPLICATING MMWAVE DIRECTIONAL INDUSTRIAL CHANNEL BEHAVIOR

Mohamed Kashef* Sudantha Perera† Carnot Nogueira† Richard Candell* Kate A. Remley†
Matthew T. Simons†

* Communications Technology Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, Maryland, United States

† Communications Technology Laboratory, NIST, Boulder, Colorado, United States

ABSTRACT

Industrial wireless channels have different characteristics than home and office channels due to their reflective nature. Moreover, the millimeter-wave (mmWave) wireless bands can play a big role in improving industrial wireless systems due to their large available bandwidth and the short wavelength that allows a large number of antennas to be located closely to each other. Wireless test chambers are used for over-the-air (OTA) testing and assessment of various protocols and equipment. However, in order to closely characterize a system under test, the chamber should be configured to replicate the environment where the system is deployed. In this work, we present a deep reinforcement learning protocol to configure a test chamber in order to replicate the spatial characteristics of measured mmWave channels in industrial environments. The proposed algorithm is general for any N -dimensional chamber configurations where it can be used to configure various reflectors, absorbers, and paddles inside a wireless test chamber.

Index Terms— Over-the-air test chamber, automatic configuration, Channel modeling, industrial wireless, deep reinforcement learning, wireless systems

I. INTRODUCTION

In future industrial systems, wireless-communication technologies such as 5G and 6G will play a critical role in achieving massive connectivity between various operational components and allowing easier equipment mobility. Industrial physical environments are different than office, home, and even outdoor urban environments which leads to different wireless channel characteristics such as the achievable delay and reliability [1], [2]. Generic models are being studied for indoor industrial channels such as [3] where four different categories of wireless channels in indoor factories are considered. However, various industrial environments differ from each other in their layouts, types of equipment, and the performed industrial activities. Hence, designing and testing of industrial wireless systems require knowledge of the channel characteristics of the corresponding environment [4]. The limited availability of sub-6 GHz wireless spectrum

has motivated the utilization of millimeter-wave (mmWave) bands for many new wireless technologies. Moreover, with many licensed bands, they offer a potential candidate for industrial wireless.

Over-the-air (OTA) testing of wireless devices and systems becomes increasingly important in the technological development and deployment of industrial wireless networks. Both wireless equipment manufacturers and users are interested in assessing system performance and user perceived quality in realistic propagation environments. Industrial wireless electromagnetic environments exhibit rich multipath propagation with strong reflections of the wireless signal over the propagation channel. The reverberation chamber (RC) is a metallic cavity where the signal created by a single source is reflected and diffused to create multipath fading. RCs can be configured as hybrid chambers through adding absorbers and reflectors in different shapes, and changing the position and orientation of different elements including the transmitter, the receiver, the absorbers, and the reflectors. The power delay profile (PDP) of a wireless channel captures the temporal variations of the channel due to multipath components (MPCs) [5]. While the power-angle delay profile (PADP) captures both the temporal and spatial characteristics of a wireless channel. By configuring the RC, various channel parameters can be emulated including the power delay profile and the power spatial pattern [6].

In [7], an RC loading configuration is introduced to emulate realistic indoor environment PDPs through using absorbers placed in the central part of the chamber and in the corners in front of the transmitting and receiving antennas. This configuration resulted in a very steep descendant PDP behavior compared to the conventional central barrier configurations. In [8], a guidance on configuring the absorbing material inside an RC and verifying chamber performance for over-the-air tests is provided. In [9], the accuracy of an RC test method for 5G FR2 bands is verified through designing and manufacturing a small-sized RC. The proposed test method is applied to omni-directional power measurements where a 5G mmWave equipment is measured and its performance is evaluated. Note that refs. [7]–[9]

emulate the delay spread of the channel, rather than its spatial characteristics. Anechoic chambers (ACs) have been described in prior work to emulate spatial characteristics, For example, in [10], realistic multi-path propagation channels in terms of angles of arrivals (AoA) and cross polarization ratio with Rayleigh fading have been emulated with the purpose of diversity measurements inside an anechoic chamber. In [11], a dynamic mmWave channel emulation method for 5G mmWave devices was investigated where the authors reproduced the dominant channel clusters with high accuracy. In [12], spatial channel emulation in an anechoic chamber is used for testing of mmWave radios. Two methods have been investigated for mmWave radios with preliminary experimental results, namely the wireless cable method and the multi-probe anechoic chamber (MPAC) method. In the MPAC method, the spatial pattern at the receiver is generated through having multiple probes inside the chamber.

In this work, we propose an automatic hybrid RC/AC configuration approach to reproduce mmWave spatial channel behavior. The approach deploys a Deep Deterministic Policy Gradient (DDPG) reinforcement learning algorithm to tune the configuration of the hybrid chamber. Generally, the proposed approach can work with N -dimensional problems to optimize the variables, and the objective function can include the temporal and spatial behavior of a channel. We only focus on reproducing the spatial behavior through minimizing the root mean squared (RMS) error of the AoAs between the reproduced channel and an exemplar channel that represents measured data from a realistic industrial environment. The configuration parameters are the positions of the reflective bodies inside the chamber. In the results, we validated the proposed approach through testing its optimization performance in a test chamber with four movable reflectors. By providing an automated approach to configuring the multipath components within an OTA test chamber, a wide variety of spatial channels can be created efficiently and repeatably.

II. PROBLEM DESCRIPTION

In this section, an overview of the collection of PADPs from an industrial environment is presented. The data preparation and the format of the resulting data are described. Finally, the automatic chamber configuration problem is stated.

II-A. Environment and Data Collection

The measured data described in this section are the industrial environment data to be used for the validation of the proposed approach. Measurements were performed in the highly reflective Central Utility Plant (CUP) at the Department of Commerce Boulder Laboratories in 2019. This environment consists of large boiler tanks, piping, and numerous racks of control hardware, as shown in Fig. 1. The vector network analyzer (VNA) was placed in a small



Fig. 1. The NIST Central Utility Plant (CUP) measurement environment

rack located between the transmit antenna and the receive array. The measurements were obtained over the band of 26.5 GHz to 40 GHz with a transmit horn antenna and a receive synthetic aperture [13] to scan a 35-by-35 planar grid with 3 mm spacing between the sample points ($\lambda/2$ at 40 GHz).

II-B. Data Preparation and Resulting Data

The data preparation approach in this subsection is deployed in both the original data preparation and later in processing the data obtained from the test chamber. The same receive synthetic aperture is used in both situations and hence the same data processing technique is used. The S_{21} parameters are collected by the synthetic-aperture and are processed using true time delay beamforming to steer the array mainbeam as described in [13]. A low-sidelobe taper is applied across the aperture that is frequency invariant in the boresight direction. Then to steer the array mainbeam towards a desired direction, an additional phase taper is applied across the aperture that varies linearly with frequency.

After coherently combining the product of measured S_{21} values and complex beamforming weights across all the aperture spatial samples, an inverse Fourier transform is utilized to transform the frequency domain data to the temporal domain. The result is known as a directional PDP. The pointing directions specified at the peak of the mainbeam are chosen systematically using the approach described in [14] such that all beams overlap at the 3-dB beamwidth. This algorithm accounts for the fact that the width of the array mainbeam increases in proportion to the product of the cosines of the azimuth and elevation scan angles.

II-C. Exemplar Extraction

In [15], we demonstrated a method to identify and characterize the spatial properties of wireless channels in industrial environments. We introduced an approach for directional

PDP exemplar extraction from measured data. The approach deploys unsupervised machine learning-based clustering for PDP exemplar extraction and uses various types of channel features for the exemplar extraction process. In particular, a cluster of directional PDPs can be represented by few PDPs that capture the most significant characteristics in this cluster. A device under test (DUT) impacted by a signal coming through this exemplar channel from these directions can be tested rigorously and repeatably for this representation of the environment. Building on this, in this work, we focus in reproducing the channel spatial behavior by creating a channel inside an OTA test chamber that has AoAs at the DUT similar to the AoAs of an exemplar channel of the measured CUP environment.

II-D. Problem Statement

We denote the set of the AoAs of the exemplar channel by θ_I and the set of the measured AoAs in the hybrid chamber by θ_C . We have N configuration parameters with normalized ranges $[0, 1]$. The value of the n th configuration parameter at any moment is denoted by x_n . The goal of the Deep Reinforcement Learning algorithm is to achieve the optimal configuration parameter settings and correspondingly each chamber reflecting element to its optimal value \hat{x}_n that minimizes the RMS error of AoAs between the exemplar channel AoAs and the measured AoAs inside the chamber.

III. AUTOMATIC CHAMBER CONFIGURATION

In this section, we explain the need for automatic configuration and the use of deep learning for this purpose. Then, we discuss how we modified a typical deep reinforcement learning approach in order to make the algorithm feasible in light of the long synthetic-aperture channel-measurement times.

III-A. Deep Reinforcement Learning Approach

We describe the conventional way of using deep reinforcement learning for chamber configuration as shown in Fig. 2. In this approach, a channel measurement is taken after each configuration change inside the chamber. This can be very time consuming in many scenarios such as the high dimensional problems in the case of a large number of configuration variables with multiple degrees of freedom or when each channel measurements take a long period of time. With four reflectors and a 35 x 35 synthetic aperture array, our measurements present both of these difficulties.

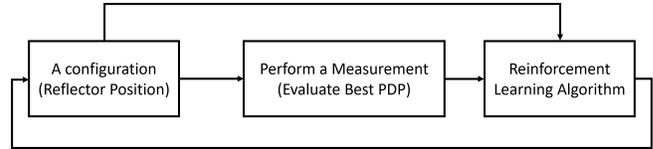


Fig. 2. The conventional reinforcement learning applied to automatic chamber configuration.

In a deep reinforcement learning approach, the following needs to be defined: a) the objective function, b) the tuning parameters, and c) the deep learning algorithm and its parameters. We start by formally defining the objective function and describe the approach for evaluating its various variables. The objective function to be minimized through the reinforcement learning approach is the RMS error of the AoAs as follows

$$J = \sqrt{\sum_{m \in \mathcal{M}} (\theta_{I,m} - \theta_{C,m})^2} \quad (1)$$

where \mathcal{M} is the set of indices of the AoAs to be replicated and $\theta_{I,m}, \theta_{C,m}$ are the m th elements of the sets θ_I, θ_C , respectively. In order to evaluate the objective function, the AoAs have to be measured. In the present work, the 2D spatial pattern of the received power is measured and it is followed by a peak detection algorithm to evaluate the AoAs at the receiver.

The tuning parameters, generally, include all the configurable elements of the test chamber. In this specific scenario, we use the linear positions of spherical reflectors as the tuning parameters. In each step of the reinforcement learning, algorithm, the vector of the tuning parameters is optimized to change the reflectors' positions and get a new measurement from which the AoAs may be estimated.

III-B. Deep Deterministic Policy Gradient (DDPG)

We next briefly overview the used deep reinforcement learning algorithm and the reason for deploying this algorithm in our work. In configuring the chamber, we are dealing with a multi-dimensional continuous output optimization problem to minimize the AoA RMS error. Modeling the impact of the reflectors' positions on the received signal AoAs is a complex problem that contains multiple non-linear effects. As a result, the use of deep learning is proposed to solve this modeling problem while the optimization problem can be solved either through supervised learning or reinforcement learning. The use of supervised deep learning in a multi-dimensional continuous output optimization problem requires a large training set of labeled data, here consisting of specific reflectors' positions and the associated resulting AoAs. Solving such a complex problem is not feasible for the case of long chamber measurement time, such as we have with the VNA-based synthetic aperture. As a result,

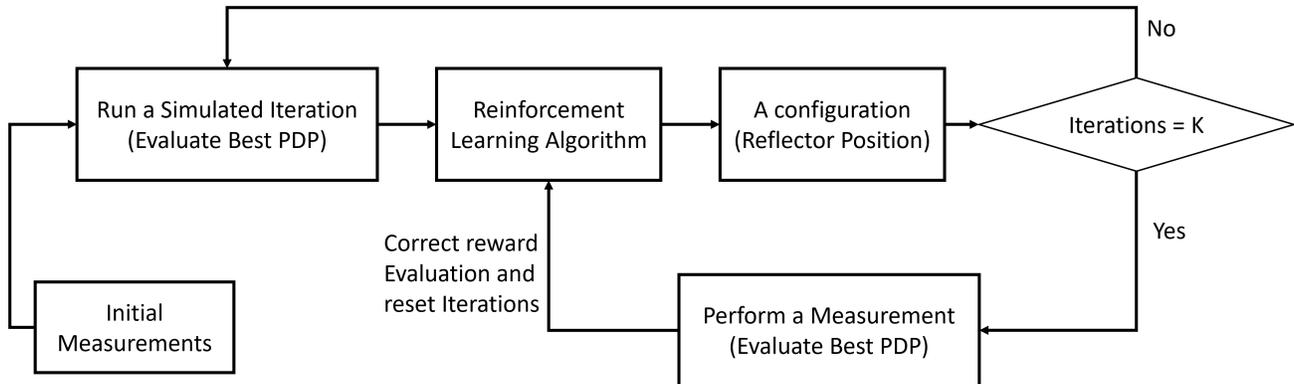


Fig. 3. The proposed hybrid reinforcement learning approach that combines numerical and measurement-based inputs to the deep learning algorithm.

the use of reinforcement learning is more suitable to the stated problem.

Reinforcement Learning (RL) is the type of learning guided by a specific objective. An agent learns by interacting with an unknown environment, typically by trying certain actions. The agent receives feedback in terms of a reward or cost from the environment; then, it trains itself and collects knowledge about the environment. RL algorithms may be policy-based or value-based or a combination of both such as the actor-critic method. The RL algorithms can also be classified as model free methods such as Q-learning or model-based algorithms such as dynamic programming. In a model-based algorithm, an agent does not rely on trials, instead it exploits an already learned model. In model-based RL, an agent can make predictions about different states and corresponding rewards after learning [16].

Deep reinforcement learning has a main advantage in its ability to learn from the actions it experiences and to be able to work in continuously changing dynamic environments [17]. Typical algorithms include Deep Q Network (DQN) algorithm and the Deep Deterministic Policy Gradient (DDPG) algorithm [18]. DQN is the simplest algorithm to implement nonlinear function approximation. When used in conjunction with the Q-Learning mechanism, DQN can enhance training stability by breaking reinforcement learning difficulties into manageable supervised learning tasks. The DDPG combines the advantages of DQN and the actor-critic framework to generate a deterministic strategy. The deep neural network parameterizes this method, making DDPG perform well in continuous control tasks [19].

In this work, we used the DDPG algorithm because of the need to have a continuous values for the reflectors' positions and because of the non-linear behavior of the AoAs with respect to the reflectors. Specifically, we deployed the implementation of the DDPG in [20]. The actions and the state of the problems are defined to be similar and to be the normalized reflectors' positions in the range of [0,1] where the position value in a specific direction is normalized with

respect to the whole linear range of this direction. The reward is the inverse of the defined objective function in eq. (1).

III-C. Hybrid Deep Reinforcement Learning Approach

We next describe our modified approach that allows the deep reinforcement learning algorithm to run its iterations using both numerical and measured data. This proposed algorithm is necessary because, for our measurement system, obtaining an AoA measurement is a lengthy process and hence obtaining the typically many measurements needed for the RL to converge is infeasible. The algorithm starts by measuring the channel at the corners of the optimization space by measuring all combinations of the interval edges of all the tuning parameters (here, linear positioners). Then, the AoA of a simulated iteration is obtained through a multi-dimensional linear interpolation. Every K simulated iterations, we perform a measured iteration to add a point to the measured data and hence improve the interpolation results in the subsequent simulated iterations. The modified algorithm block diagram is shown in Fig. 3.

IV. RESULTS

In this section, we present the results of a realistic example of the automated chamber configuration approach on the OTA test chamber to reproduce the spatial behavior of a mmWave exemplar channel representing the measured utility plant data. We start this section by describing the chamber measurement procedure and the data processing. We then show the results for running the procedure to reproduce a physical channel that provide angles of arrival from the four desired spatial directions.

IV-A. OTA Test Chamber and Data Processing

The hybrid chamber can take measurements in any frequency bandwidth within 26.5 GHz to 40 GHz range. In this paper we took S-parameter measurements from 26.5 GHz to 29.5 GHz with 100 MHz frequency steps, 100 Hz IF bandwidth, and -15 dBm power level settings of the VNA. The synthetic aperture beam-forming was based on

S_{21} parameters from 1225 measurements in 35×35 virtual array for each polarization.

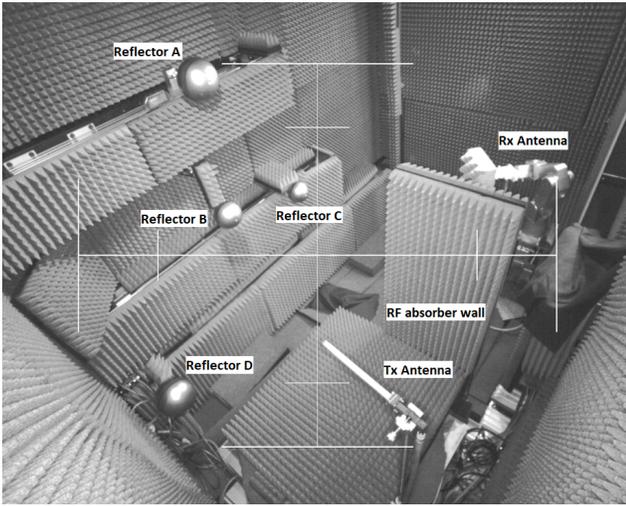


Fig. 4. A photo of the hybrid measurement chamber showing various equipment and components.

Fig. 4 shows the main components of the measurement setup in the chamber. The transmit (Tx) antenna is stationary and connected to port 1 of the VNA. Port 2 of the VNA is connected to the receive (Rx) antenna which is mounted on the robotic arm. Both antennas are WR28 open-ended waveguides with the same model number. There is an RF absorber wall (a metal plate sandwiched between two RF absorbers) to prevent Tx-Rx direct coupling. Four reflectors (A, B, C, and D) mounted on four linear positioners, which can reconfigure the chamber to emulate different exemplar channels.

IV-B. Measured AoAs

In this subsection, we show the obtained 2D spatial pattern of the received power at the beginning and the end of the RL algorithm. We compare the received power pattern to the AoAs of the desired exemplar channel. In Fig. 5, we present the received power pattern for the first measurement inside the chamber after the initial measurement phase and running the RL algorithm for the initial K simulated iterations. The black dots in the figure represent the AoAs of the exemplar channel that should be replicated. It can be seen that the received power pattern peaks are getting closer to the exemplar AoAs but have not converged yet.

Each complete cycle of the hybrid RL approach includes a chamber measurement and K simulated iterations. After running 7 complete cycles, the RL algorithm converged. The result in Fig. 6 shows the resulted received power pattern. In this figure, we see that the peaks of the received power pattern after running our hybrid deep learning algorithm almost overlapped with the exemplar AoAs.

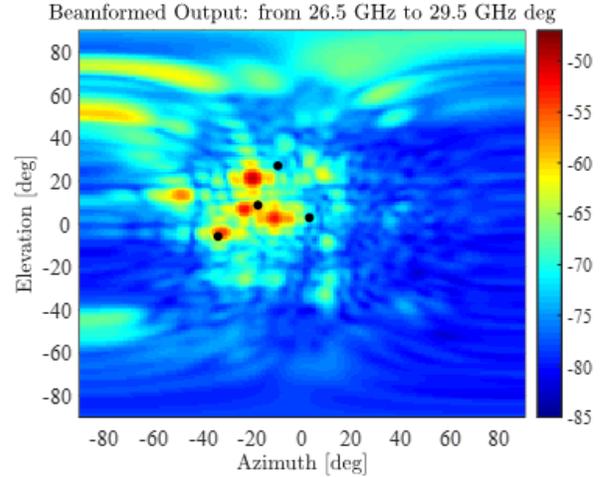


Fig. 5. The 2D spatial received power pattern resulting from the set of first measurements inside the test chamber.

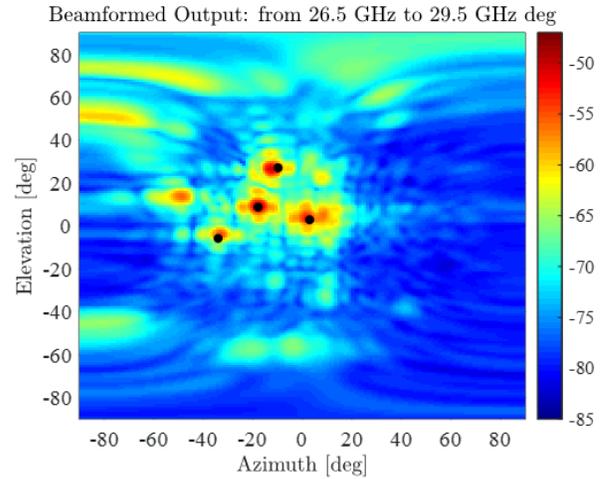


Fig. 6. The power pattern resulting from taking the final synthetic-aperture measurement inside the test chamber after the approach converged.

IV-C. Resulting Metrics

In this subsection, we present the evolution of relevant metrics versus the seven chamber-configuration measurements that were performed between each set of 7 complete cycles of the hybrid approach. In Fig. 7, the AoA RMS error value in degrees from eq. 1 is presented. The RMS error has dropped significantly through the seven measurements using the modified RL approach. The error has saturated to a value higher than 0 because each reflector has a single degree of freedom and can only move horizontally.

In Fig. 8, we present another metric that assesses the

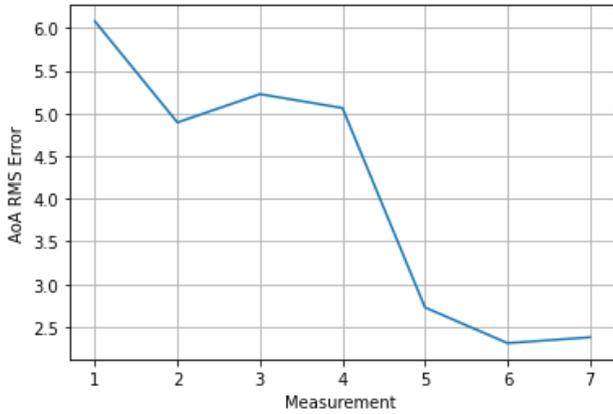


Fig. 7. The AoA RMS error values at various measured data points.

performance of the algorithm: the average received power over the exemplar AoAs. The higher the average received power, the better the system has achieved the exemplar AoAs.

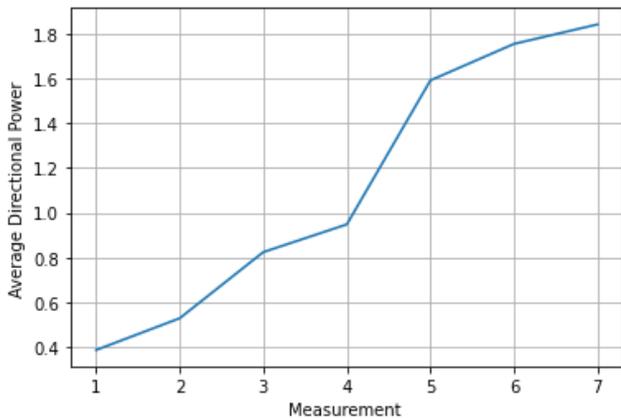


Fig. 8. The average power averaged over the exemplar AoAs at various measured data points.

V. CONCLUSIONS

We introduced an automatic OTA test chamber configuration approach that deploys deep reinforcement learning. The approach is generic that can be used to optimize any N -dimensional tuning parameters to optimize an objective function. In this work, we validated the proposed approach to configure the position of four reflectors to minimize the RMS error of the AoAs between the channel created inside the test chamber and an exemplar channel of a realistic industrial environment.

The ability to physically recreate industrial channels will allow the testing and performance evaluation of wireless

equipment in realistic environments. The flexible hybrid chamber presented here can create these channels in a controllable fashion with the ability to test various directional stress scenarios by configuring specific directional channel patterns. The modified deep reinforcement learning approach allows the automatic configuration approach to converge with a smaller number of measurements.

Disclaimer Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

VI. REFERENCES

- [1] K. S. Low, N. Win, and M. J. Er, "Wireless Sensor Networks for Industrial Environments," 2005.
- [2] M. Cheffena, "Propagation channel characteristics of industrial wireless sensor networks [wireless corner]," *IEEE Antennas and Propagation Magazine*, vol. 58, no. 1, pp. 66–73, Feb. 2016.
- [3] 3GPP, "Addition of indoor industrial channel model to Doc. 38.901 - Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Tech. Rep. R1-1909807, 09 2019. [Online]. Available: <https://portal.3gpp.org/ngppapp/CreateTdoc.aspx?mode=view&contributionId=1045855>
- [4] R. Candell, M. Kashef, Y. Liu, K. B. Lee, and S. Fofou, "Industrial wireless systems guidelines: Practical considerations and deployment life cycle," *IEEE Industrial Electronics Magazine*, vol. 12, no. 4, pp. 6–17, Dec. 2018.
- [5] B. Soret, M. C. Aguayo-torres, and J. T. Entrambasaguas, "Capacity with explicit delay guarantees for generic sources over correlated Rayleigh channel," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 1901–1911, June 2010.
- [6] V. M. Primiani, M. Barazzetta, L. Bastianelli, D. Micheli, E. Moglie, R. Diamanti, and G. Gradoni, "Reverberation chambers for testing wireless devices and systems," *IEEE Electromagnetic Compatibility Magazine*, vol. 9, no. 2, pp. 45–55, 2020.
- [7] A. Sorrentino, F. Nunziata, S. Cappa, S. Gargiulo, and M. Migliaccio, "A semi-reverberation chamber configuration to emulate second-order descriptors of real-life indoor wireless propagation channels," *IEEE Transactions on Electromagnetic Compatibility*, vol. 63, no. 1, pp. 3–10, 2021.
- [8] K. A. Remley, J. Dortmans, C. Weldon, R. D. Horansky, T. B. Meurs, C.-M. Wang, D. F. Williams, C. L. Holloway, and P. F. Wilson, "Configuring and

- verifying reverberation chambers for testing cellular wireless devices,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 3, pp. 661–672, 2016.
- [9] Z. Wang, Y. Ren, Y. Zhang, C. Pan, and X. Zhang, “A novel mode-stirred reverberation chamber design for 5g millimeter wave bands,” *IEEE Access*, vol. 9, pp. 38 826–38 832, 2021.
- [10] A. Choumane, A. E. S. Ahmad, and K. Khoder, “Emulation of realistic multi-path propagation channels inside an anechoic chamber for antenna diversity measurements,” *Wirel. Eng. Technol.*, vol. 11, no. 01, pp. 1–12, 2020.
- [11] X. Cai, Y. Miao, J. Li, F. Tufvesson, G. F. Pedersen, and W. Fan, “Dynamic mmwave channel emulation in a cost-effective MPAC with dominant-cluster concept,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 6, pp. 4691–4704, 2022.
- [12] W. Fan, L. Hentilä, and P. Kyösti, “Spatial fading channel emulation for over-the-air testing of millimeter-wave radios: concepts and experimental validations,” *Frontiers of Information Technology & Electronic Engineering*, vol. 22, no. 4, pp. 548–559, Apr 2021. [Online]. Available: <https://doi.org/10.1631/FITEE.2000484>
- [13] P. Vouras, B. Jamroz, J. Quimby, A. Weiss, D. Williams, R. Leonhardt, D. Guven, R. Jones, J. Kast, and K. Remley, “Wideband synthetic aperture millimeter-wave spatial channel measurements with uncertainties,” *in review*, 2020.
- [14] G. V. T. W. M. Waters, D. P. Patel, “Optimum number of faces of a volume-scanning active array radar,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, July 1998.
- [15] M. Kashef, P. Vouras, R. D. Jones, R. Candell, and K. A. Remley, “A machine-learning approach for the exemplar extraction of mmwave industrial wireless channels,” *IEEE Open Journal of Instrumentation and Measurement*, vol. 1, pp. 1–15, 2022.
- [16] M. Naeem, S. T. H. Rizvi, and A. Coronato, “A gentle introduction to reinforcement learning and its application in different fields,” *IEEE Access*, vol. 8, pp. 209 320–209 344, 2020.
- [17] S. S. Mousavi, M. Schukat, and E. Howley, “Deep reinforcement learning: An overview,” in *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016*, Y. Bi, S. Kapoor, and R. Bhatia, Eds. Cham: Springer International Publishing, 2018, pp. 426–440.
- [18] Q. Zhou, “A novel movies recommendation algorithm based on reinforcement learning with DDPG policy,” *International Journal of Intelligent Computing and Cybernetics*, vol. 13, no. 1, pp. 67–79, Mar. 2020. [Online]. Available: <https://doi.org/10.1108/ijicc-09-2019-0103>
- [19] W. Zhang, Q. Chen, J. Yan, S. Zhang, and J. Xu, “A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting,” *Energy*, vol. 236, p. 121492, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0360544221017400>
- [20] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, “Continuous control with deep reinforcement learning,” 2015. [Online]. Available: <https://arxiv.org/abs/1509.02971>