On the Latent Space of mmWave MIMO Channels for NLOS Identification in 5G-Advanced Systems

Bernardo Camajori Tedeschini[ID], *Graduate Student Member, IEEE*, Monica Nicoli[ID], *Senior Member, IEEE*, and Moe Z. Win[ID], *Fellow, IEEE*

Abstract—In mission-critical verticals such as automated driving, 5G-advanced networks must provide centimeter-level dynamic positioning along with ultra-reliable low-latency communication services. Massive Multiple-Input Multiple-Output (mMIMO) and millimeter waves (mmWave) are the key enablers, allowing high accuracy angle and delay estimation. Still, extracting such information from highly-dimensional Channel Impulse Responses (CIRs) results in a complex task, due to channel sparsity and intermittent blockage. In this paper we focus on non-line-of-sight (NLOS) identification from CIR data, proposing a Deep Autoencoding Kernel Density Model (DAKDM) to characterize the statistics of the channel latent features. We formulate the problem as a semi-supervised anomaly detection task in which only LOS samples, i.e., normal data, are adopted for training. DAKDM is a single-stage training model that takes as input the full CIR thanks to an AutoEncoder (AE) structure. The proposed method is able to learn the latent distribution by means of a Kernel Density Estimator (KDE) in combination with a deep learning likelihood network. We validate the proposed solution in a 5G Urban micro (UMi) vehicular scenario. Results show that the proposed model can significantly outperform conventional algorithms and obtain similar performances to variational Bayes algorithms at one tenth of the inference time.

Index Terms—Deep autoencoding kernel density model, anomaly detection, CIR, 5G, deep learning, NLOS identification.

I. INTRODUCTION

THE newest release of the 5th generation (5G) of cellular communication systems, namely the 3rd generation partnership project (3GPP) Release 16, also known as new radio (NR), introduces for the first time high-precision positioning functionalities into cellular networks. location services (LCS) are extended from regulatory services to roaming and commercial capabilities [1], [2], [3], [4]. Higher frequencies, bandwidth improvements and massive-multiple-input multiple-output (MIMO) technologies are the key feature enablers for radio access technology (RAT)-dependent dynamic positioning [2], [3], [4], [5] and location awareness of connected nodes [6], [7], [8], [9], [10], [11]. The major fields of application can be found in target tracking [12], [13], [14], internet-of-things (IoT) [15], [16], [17], crowd sensing [18], [19], smart environments [20] and industrial automation [21]. Strict requirements are foreseen for the most critical services such as automated driving [22], [23]. These include a lateral and longitudinal positioning error of 10 and 50 cm [24], respectively, and a latency down to 5 ms for fully autonomous driving vehicles [25]. Next 5G releases, also known as 5G-Advanced and beyond, will have to meet such challenging localization requirements while coping with complex propagation conditions, due to the extreme path-loss and frequent blockage conditions experienced by millimeter waves (mmWave).

These problems have been widely studied in the field of localization and navigation focusing on fundamental performance limits [26], [27], [28], [29], [30], [31], algorithm design [32], [33], [34], [35], [36], [37], [38], network operation [39], [40], [41], [42], and network experimentation [43], [44], [45], [46], [47], [48]. It is clear that legacy solutions for positioning, based on conventional approaches for multi-lateration/angulation, will struggle to deal with rapidly fading channels and intermittent blockage. Geometrical approaches rely in fact on line-of-sight (LOS) condition for estimating directions and ranges of the positioning reference signals. Real-time detection and prediction of non-LOS (NLOS) links is mandatory to mitigate the false localization due to biased range/angle estimates. Since the environment significantly impacts on the propagation, data-driven techniques have so far produced very encouraging outcomes in NLOS detection [48], [49]. Therefore, machine learning (ML) is expected to play a crucial role in future generation networks [50], [51] and standard compliant solutions are foreseen already from Release 17-18 [52].

Solutions for blockage detection should exploit the whole power-delay-angle profile of the channel impulse response (CIR) as this embeds a wide range of geographical data and propagation characteristics [53]. In 5G industrial use-cases, e.g., automated driving, historical CIR data are largely available in roadside units that receive continuous information from geolocalized vehicles [54], [55], [56], [57]. ML algorithms

Bernardo Camajori Tedeschini is with Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, 20133 Milan, Italy (e-mail: bernardo.camajori@polimi.it).

Monica Nicoli is with Dipartimento di Ingegneria Gestionale (DIG), Politecnico di Milano, 20156 Milan, Italy (e-mail: monica.nicoli@polimi.it).

Moe Z. Win is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).
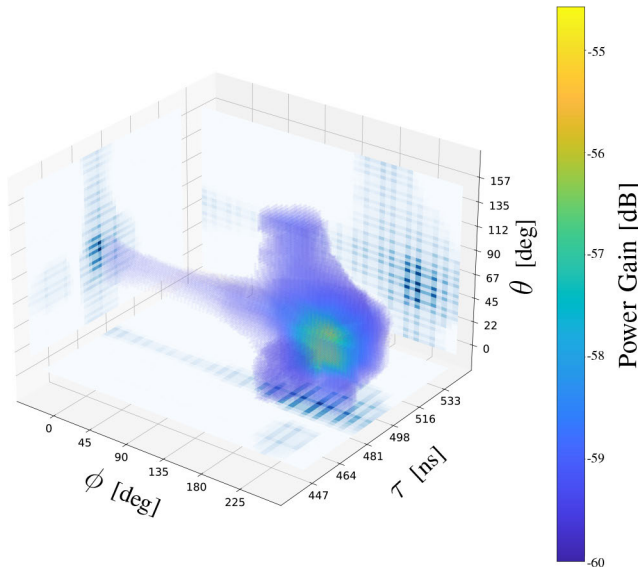
Fig. 1. Sparse channel representation in azimuth ϕ, elevation θ and delay τ domain.

could easily exploit these data for automatic NLOS detection. However, since such signals are highly dependent on the environment, ML approaches for detecting NLOS using CIRs frequently fail to generalize to varied contexts [58]. Moreover, massive MIMO and very high frequencies of advanced-5G networks will produce high dimensional channel responses which may be complex to handle. An example of channel power-delay-angle-profile is shown in Fig. 1, for a 5G urban scenario with carrier frequency 30 GHz, bandwidth 400 MHz and uniform planar antenna array receiver of 64 elements. The sparse power delay-angle profile of the channel is a signature of the user location and should be exploited to infer the visibility conditions of the base station.

In this paper, we propose an innovative strategy to characterize the sparsity of the mmWave MIMO channel and approximate whatever high-dimensional distribution in a fast and compact way. To demonstrate the efficacy of the method, we address the problem of NLOS identification, exclusively employing LOS CIRs for training. This is done because LOS CIRs are easier to extract in training procedures and present more peculiar distributions, i.e., usually the direct path is the dominant factor in a Rician fading channel. In addition, this facilitates the deployment and results in higher generalization compared to other systems that require both classes for training (i.e., LOS and NLOS). Therefore, we treat the problem as an anomaly detection case in which LOS samples are considered as normal samples, while NLOS samples are anomalous.

II. RELATED WORKS

In this section, we first review the literature starting from early works on ultra wide-band (UWB) systems (Sec. II-A) and then we extend the analysis to ML-based algorithms (Sec. II-B). Next, we review the state of the art on anomaly detection focusing on neural network (NN) approaches (Sec. II-C) and we discuss the original contributions provided in the paper (Sec. II-D).

A. NLOS Identification

Existing techniques for NLOS identification/detection problem can be mainly divided into three major categories: based on range estimates, based on position estimates and based on channel statistics. The first group of methods, i.e., based on range estimates, measures the running variance of the ranges and applies a threshold using pre-computed variance statistics [59]. The techniques based on position estimates are mainly map-based, i.e., they observe the user equipment (UE) position in relation to the geometry of the environment [60], [61]. These first two categories are either too oversimplified or require perfect knowledge of the UE's position and of the map geometry.

The third class relies on channel statistics, such as amplitude, mean and root-mean-square delay. In case these statistics are known at-priori, a joint-likelihood ratio test can be adopted for hypothesis selection [62], [63] or as soft information in weighted least squares (WLS) algorithms. The limitations of this last class of existing techniques include experiencing delays while gathering channel statistics to create a database and determining the complex combined probability distributions of necessary features for statistical methods. ML-based approaches overcome these drawbacks by avoiding statistical modeling of the input features.

B. ML for NLOS Identification

ML approaches to NLOS identification can be divided into: supervised, unsupervised and semi-supervised learning. First works (i.e., supervised learning) use hand-crafted channel state information (CSI) features such as energy, maximum rise time, kurtosis, root-mean-square-delay spread, maximum amplitude, time of flight (ToF), Rician-K-factor and mean excess delay [48], [64], [65], [66]. These deterministic features have a solid theoretical basis and capture the differences between LOS and NLOS conditions in terms of power and delay attributes, as well as the strength of the dominant signal component relative to the multipath components. The most popular adopted ML models are support vector machines (SVMs), relevance vector machines (RVMs), random forests (RFs) and Gaussian processes (GP). These methods can also be used to directly mitigate the range bias by applying a regression problem to the ranging-error estimates [49], [67].

Despite achieving good results, these methods highly depend on the pre-selected features which limit their potential. On the other hand, deep learning (DL) approaches can directly learn the most suitable combination of features (typically non-linear) using as input the full CIR and producing as output the desired classification. First works in this direction can be found in [68], [69], [70] using convolutional neural networks (CNN) to perform feature extractions in grid-like data where local patterns and structures are critical. Some recent studies [71], [72] directly exploit the automatic feature extraction of the CNN in order to locate a target by performing a fingerprint training. A main limit is the need of extensive measurement campaigns and time-consuming labeling of data. Moreover, supervised learning approaches require updating the training

database when conditions are changed and need representative samples of all the possible NLOS anomalies.

A solution could be permitting not to have labels at all and manage the problem as an unsupervised one. Authors in [73] fit a Gaussian mixture model (GMM) with two components (one for LOS and one for NLOS) using some key hand-crafted features of the channel and output the classification according to the magnitude of the membership weights. While unsupervised techniques are very promising, unfortunately they do not achieve very high performances, due to lack of knowledge or lack of structured data.

The third class of semi-supervised approaches has the advantage of not needing examples of all the possible anomalies as supervised learning. Moreover, powerful DL semi-supervised methods focus on learning one single distribution which, in many cases, is easier than a separating boundary between two distributions [74]. Works that adopt this strategy can be found in [58], [75] which adopt the Pearson correlation coefficient and one-class SVM to perform NLOS classification, respectively. A very recent work [76] employs variational autoencoder (VAE) to perform feature extraction and imposes a Gaussian distribution to the latent features in order to ease learning of distribution of normal samples. The score adopted to define the probability of NLOS is then used to estimate the bias and variance of time-based measurements. Although the idea of using an autoencoder (AE) to have a compact representation of the channel can give very good results, the usage of sampling-based methods to perform the prediction has the main drawback of not being suitable for real-time applications.

C. Neural Networks for Anomaly Detection

Anomaly detection is frequently employed in problems where we have a large amount of data from normal circumstances but little data from abnormal ones. Here, on the other hand, we consider the setting of semi-supervised learning in which normal training data only are provided. In this case, the problem turns out to be locating those samples that do not conform to the normal ones or a model explaining normal ones. Thus, the objective is to learn in a finer way as possible the distribution related to the normal samples.

To this aim, many works focus on end-to-end models to directly produce the classification using one-class neural networks (OC-NN) [77]. On the other hand, generative models are increasing in popularity with generative adversarial network (GAN) and VAE [78]. However, GANs are problematic to control in the training phase [79] and VAEs have the downside of requiring sampling, which is unfeasible under certain use cases, and furthermore experiments have shown that they tend to perform worse than AE [80], [81].

Reconstruction methods, as AE, are the most widely used DL techniques for anomaly detection in images [82]. Usually, they are used in combination with density-based methods, as kernel density estimation (KDE) [83], for score estimation by first performing dimensionality reduction, and then applying density estimation to the latent low-dimensional space. However, these two-steps methods restrict the modification to the dimensionality reduction since fine-tuning is difficult in

well-trained AE. To solve this problem, authors in [84] propose a deep autoencoding Gaussian mixture model (DAGMM) to mutually learn the latent feature representations and their density under the GMM framework by mixture membership estimation. Even though their approach is direct and does not require two step-training, GMM may not be able to fully represent the latent distribution of normal samples and are subject to singularity problems. On the other hand, KDE are perfect to represent complex distributions, but they are very slow in evaluation and require storing the whole dataset for inference.

D. Contributions

In this paper, we address the problem of NLOS identification in 5G-advanced cellular systems using an innovative approach that allows to overcome the aforementioned limitations. The main contributions are the following:

- We propose a feature extraction method that exploits the angle-delay channel power matrix (ADCPM) as input data and allows to characterize the distributions of the latent features of the sparse space-time channel in massive MIMO cellular systems using orthogonal frequency division multiplexing (OFDM).
- We design NLOS identification as a semi-supervised anomaly-detection problem by exploiting a deep autoencoding kernel density model (DAKDM). The DL model allows to identify the few key parameters that describe the sparse space-time channel response and to learn the distributions of such latent features from training data. The proposed approach is able to jointly learn the sparse channel representation and approximate the KDE likelihood in a single training stage without storing the dataset.
- We simulate a realistic 5G-advanced MIMO-OFDM vehicular scenario, according to the standard specifications [85], using a Matlab ray-tracing software [86]. The scenario is composed of multiple vehicular UEs created with simulation of urban mobility (SUMO) software [87].

The paper is organized as follows: Sec. III introduces the channel model for a multi-user MIMO-OFDM system and its extracted fingerprinting. Sec. IV provides the context of anomaly detection applied to the NLOS identification and defines the proposed DAKDM solution. Sec. V is devoted to the description of the simulated 5G scenario and to the comparison with state-of-the-art anomaly detection DL methods. Finally, Sec. VI draws the conclusion.

E. Notation

Random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by x and x; a random vector and its realization are denoted by $\mathbf{x}$ and $\boldsymbol{x}$; a random matrix and its realization are denoted by $\mathbf{X}$ and X, respectively. Random sets and their realizations are denoted by up-right sans serif and calligraphic font, respectively. For example, a random
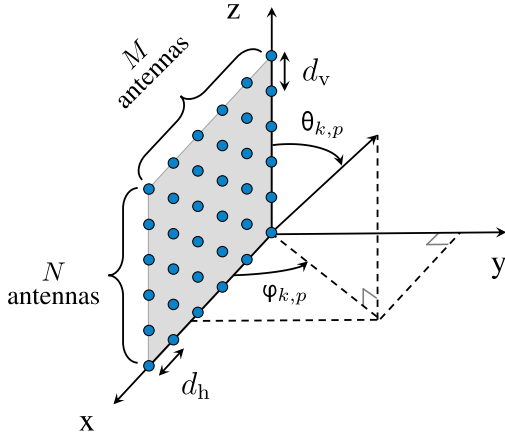
Fig. 2. Uniform planar array with M and N antennas in x and z directions, respectively. The direction of arrival (DoA) is highlighted and decomposed into azimuth $0 \leq \varphi_{k,p} < \pi$ and elevation $0 \leq \theta_{k,p} < \pi$.

set and its realization are denoted by X and $\mathcal{X}$, respectively. The function $p_{\mathsf{x}}(x)$, and simply $p(x)$ when there is no ambiguity, denotes the probability density function (PDF) of x. $j = \sqrt{-1}$ denotes the imaginary unit. The notation $\mathbf{X}^{\mathrm{T}}$, $\mathbf{X}^*$ and $\mathbf{X}^{\mathrm{H}}$ indicate the matrix transposition, conjugation and conjugate transposition. The Kronecker and the Hadamard product between two matrices are denoted with the symbols $\otimes$ and $\odot$, respectively. With the notation $\mathsf{x} \sim \mathcal{CN}(\mu, \sigma^2)$ we indicate a complex Gaussian random variable x with mean μ and standard deviation σ. We use $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$ to denote the expectation and the variance of random variable, respectively. $\mathbb{R}$ and $\mathbb{C}$ stand for the set of real and complex numbers, respectively. $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ are the real and complex part of the complex number x, respectively. $\lfloor x \rfloor$ indicates the largest integer not greater than x, while $\delta(\cdot)$ and $\delta[\cdot]$ are the Dirac delta and Kronecker functions, respectively.

III. SYSTEM MODEL

A. Channel Model

We consider a multi-user mmWave MIMO-OFDM system in which K UEs transmit in uplink direction over a bandwidth B at carrier wavelength λ_{c}. The base station (BS)'s cell panel is equipped with an uniform planar array (UPA) with $N \times M$ isotropic antennas. The antenna spacings are d_{h} and d_{v}, over the horizontal and vertical dimension, respectively. We assume that the UE transmits using only one logical port and a number of physical antennas unknown at the BS. Between the UE $k = 1, \ldots, K$ and the BS, we consider a multipath channel with N_k paths with ToF $\tau_{k,p}$ for path $p = 1, \ldots, N_k$. The DoAs from the k-th UE and of the p-th path are divided into azimuth $0 \leq \varphi_{k,p} < \pi$ and elevation $0 \leq \theta_{k,p} < \pi$. A picture of the panel array can be found in Fig. 2. We restrict the azimuth up to π since we consider an UPA antenna. For tri-sectorial BSs the angular coverage is reduced to $2\pi/3$.

The OFDM modulation is performed over N_{c} sub-carriers, sampling interval T_{s} and symbol duration $T_{\mathrm{c}} = N_{\mathrm{c}}T_{\mathrm{s}}$. Considering a baseband representation of the signal, we define the frequency at the ℓ-th sub-carrier as $f_\ell = \frac{\ell}{T_{\mathrm{c}}}$, $\ell = 0, \ldots, N_{\mathrm{c}}-1$. The cyclic-prefix duration is $T_{\mathrm{g}} = N_{\mathrm{g}}T_{\mathrm{s}}$ and it is assumed

to be larger than the maximum channel delay for all UEs $\tau_{\mathrm{MAX}} = \max_{k,p} \tau_{k,p}$. Consequently, we define with $\mathsf{r}_{k,p} = \lfloor \frac{\tau_{k,p}}{T_{\mathrm{s}}} \rfloor$ the temporally resolvable propagation delay of the p-th path with respect to the k-th UE. Thus, the baseband CIR of user k is modelled as [88]:

$$\mathbf{h}_k(\tau) = \sum_{p=1}^{N_k} \mathsf{a}_{k,p} \beta_{k,p}\, \mathbf{e}(\theta_{k,p}, \varphi_{k,p}) e^{-j2\pi \frac{\mathsf{d}_{k,p}}{\lambda_{\mathrm{c}}}} \delta(\tau - \tau_{k,p}), \quad (1)$$

where the p-th path is characterized by a complex path gain $\alpha_{k,p} = \mathsf{a}_{k,p} e^{-j2\pi \frac{\mathsf{d}_{k,p}}{\lambda_{\mathrm{c}}}} \beta_{k,p}$ with $\beta_{k,p} = e^{j2\pi \mathsf{v}_{k,p} t}$ due to the Doppler frequency shift, a traveled distance $\mathsf{d}_{k,p} = c\tau_{k,p}$, a pulse waveform approximated with a Dirac delta function $\delta(\tau - \tau_{k,p})$ and an array response vector $\mathbf{e}(\theta_{k,p}, \varphi_{k,p}) \in \mathbb{C}^{MN}$. For $p > 1$, the p-th path is $\alpha_{k,p} = \mathsf{a}_{k,p} e^{j\psi_{k,p}}$, with $\psi_{k,p} = 2\pi \mathsf{v}_{k,p} t - 2\pi \frac{\mathsf{d}_{k,p}}{\lambda_{\mathrm{c}}}$ and $\alpha_{k,p} \sim \mathcal{CN}(0, \sigma_{k,p}^2)$. The first path $p = 0$ is $\alpha_{k,0} \sim \mathcal{CN}(s_0 \mathsf{a}_{k,0} e^{j\psi_{k,0}}, \sigma_{k,0}^2)$ where it is $s_0 = 1$ for LOS (i.e., with a deterministic direct path contribution and Rician fading) and $s_0 = 0$ for NLOS (i.e., Rayleigh fading). Additionally, we consider the Doppler-related rotation to be almost constant over time interval τ_{MAX} and that the complex amplitudes $\alpha_{k,p}$ associated to different paths as uncorrelated, according to the wide-sense stationary uncorrelated scattering model. At the BS, the array response vector can be decomposed into [89]:

$$\mathbf{e}(\theta_{k,p}, \varphi_{k,p}) = \mathbf{e}_{\mathrm{v}}(\theta_{k,p}) \otimes \mathbf{e}_{\mathrm{h}}(\theta_{k,p}, \varphi_{k,p}), \quad (2)$$

where the $M \times 1$ response vector to the elevation angle is:

$$\mathbf{e}_{\mathrm{v}}(\theta_{k,p}) = \left[e^{-j2\pi(m-1)\frac{d_{\mathrm{v}}}{\lambda_{\mathrm{c}}}\cos(\theta_{k,p})} \right]_{m=1}^{M} \quad (3)$$

and the $N \times 1$ response vector to the azimuth angle is:

$$\mathbf{e}_{\mathrm{h}}(\theta_{k,p}, \varphi_{k,p}) = \left[e^{-j2\pi(n-1)\frac{d_{\mathrm{h}}}{\lambda_{\mathrm{c}}}\sin(\theta_{k,p})\cos(\varphi_{k,p})} \right]_{n=1}^{N}. \quad (4)$$

Adopting an OFDM modulation with sampling at $t = nT_{\mathrm{s}}$, the channel frequency response (CFR) at the ℓ-th sub-carrier can be written as the discrete Fourier transform (DFT) of the CIR of the different paths [90], [91]:

$$\begin{aligned}
\mathbf{H}_{k,\ell} &\approx \sum_{n=0}^{N_{\mathrm{g}}-1} \sum_{p=1}^{N_k} \alpha_{k,p}\, \mathbf{e}(\theta_{k,p}, \varphi_{k,p})\, \delta[n - \mathsf{r}_{k,p}]\, e^{-j2\pi \tau_{k,p} f_\ell} \\
&= \sum_{p=1}^{N_k} \alpha_{k,p}\, \mathbf{e}(\theta_{k,p}, \varphi_{k,p})\, e^{-j2\pi \frac{\ell \mathsf{r}_{k,p}}{N_{\mathrm{c}}}}, \quad (5)
\end{aligned}$$

where the approximation holds for ToFs multiple of the sampling interval T_{s} or equivalently for $N_{\mathrm{g}} \to \infty$. Finally, the space-frequency channel response matrix (SFCRM) $\mathbf{H}_k \in \mathbb{C}^{MN \times N_{\mathrm{c}}}$ of the k-th UE is obtained as:

$$\mathbf{H}_k = [\mathbf{H}_{k,0}\ \mathbf{H}_{k,1}\ \ldots\ \mathbf{H}_{k,N_{\mathrm{c}}-1}], \quad (6)$$

which will be used in the next section to extract the channel fingerprints.

B. Channel Fingerprints

To detect the propagation conditions that generated the response (5), classifying them in LOS or NLOS, we propose to analyze the CFR in the angle-delay domain, which eases the recognition of the clustered multipath components associated to the direct (LOS) or secondary (NLOS) macro-paths. We thereby convert the SFCRM (6) into the domain of the angle of arrival (AoA) and the ToF, by introducing the angle-delay channel response matrix (ADCRM). We define with $\boldsymbol{V}_M \in \mathbb{C}^{M \times M}$ and $\boldsymbol{V}_N \in \mathbb{C}^{N \times N}$ the phase-shifted DFT matrices [92] where $[\boldsymbol{V}_M]_{u,v} = \frac{1}{\sqrt{M}} e^{-j2\pi \frac{u(v-\frac{M}{2})}{M}}$ and $[\boldsymbol{V}_N]_{u,v} = \frac{1}{\sqrt{N}} e^{-j2\pi \frac{u(v-\frac{N}{2})}{N}}$. Then, we denote by $\boldsymbol{F} \in \mathbb{C}^{N_c \times N_g}$ the matrix formed by the first N_g columns of N_c dimensional unitary DFT matrix where $[\boldsymbol{F}]_{u,v} = \frac{1}{\sqrt{N_c}} e^{-j2\pi \frac{uv}{N_c}}$. ADCRM is computed as [93]:

$$\mathsf{G}_k = \frac{1}{\sqrt{MNN_c}} (\boldsymbol{V}_M^{\mathrm{H}} \otimes \boldsymbol{V}_N^{\mathrm{H}}) \mathsf{H}_k \boldsymbol{F}^* \in \mathbb{C}^{MN \times N_g}, \quad (7)$$

where $(\boldsymbol{V}_M^{\mathrm{H}} \otimes \boldsymbol{V}_N^{\mathrm{H}})$ and $\boldsymbol{F}^*$ project the SFCRM into the angle and delay domain, respectively.

For NLOS identification, we propose to use the ADCRM to compute the average power of the channel components that are collected into the ADCPM defined as:

$$\mathsf{P}_k = \mathbb{E}\{\mathsf{G}_k \odot \mathsf{G}_k^*\} \in \mathbb{C}^{MN \times N_g}, \quad (8)$$

where $[\mathsf{P}_k]_{i,r} = \mathbb{E}\{|[\mathsf{G}_k]_{i,r}|^2\}$. We recall here that the ADCPM holds some important asymptotic properties, as for N, M and $N_g \to \infty$, it tends to be a sparse matrix with elements $[\mathsf{P}_k]_{i,r}$ matching the i-th AoA and the r-th ToF [93]:

$$\lim_{M,N,N_g \to \infty} [\mathsf{P}_k]_{i,r} = \sum_{p=1}^{N_k} \sigma_{k,p}^2 \delta[i - m_{k,p}N - n_{k,p}]\delta[r - \mathsf{r}_{k,p}], \quad (9)$$

where $m_{k,p}N + n_{k,p}$ denotes the index of the i-th AoA and $\mathsf{r}_{k,p}$ the index of the r-th ToF. Note that the angle and delay indexes $m_{k,p}N + n_{k,p}$ and $\mathsf{r}_{k,p}$, are two distinct and discrete quantities which relates to the physical AoA and ToF in the following way. The ToF can be approximated as $\tau_{k,p} = \mathsf{r}_{k,p}T_s$, while the azimuth $\varphi_{k,p}$ and elevation $\theta_{k,p}$ can be written as $\varphi_{k,p} = \arccos\left(\frac{m_{k,p} - \frac{M}{2}}{M} \frac{\lambda_c}{d_h}\right)$ and $\theta_{k,p} = \arccos\left(\frac{n_{k,p} - \frac{N}{2}}{N} \frac{\lambda_c}{d_v}\right)$, respectively. Consequently, working in the transformed angle-delay domain allows the DL model to learn the location-dependent features and, therefore, the statistics of LOS data to be used for blockage prediction.

Regarding the complexity overhead due to the ADCPM computation, we observe that P_k is obtained from the channel matrix H_k which is always estimated for communications purposes. Therefore the only overhead is the computation of (7), which can be efficiently performed using the 2D-Inverse Fast FT (IFFT), with an overall complexity of $O(MNN_g \cdot \log(MNN_g))$.

IV. BLOCKAGE DETECTION METHODOLOGY

In this section, we first introduce the problem formulation of the semi-supervised setting which serves for the proposed
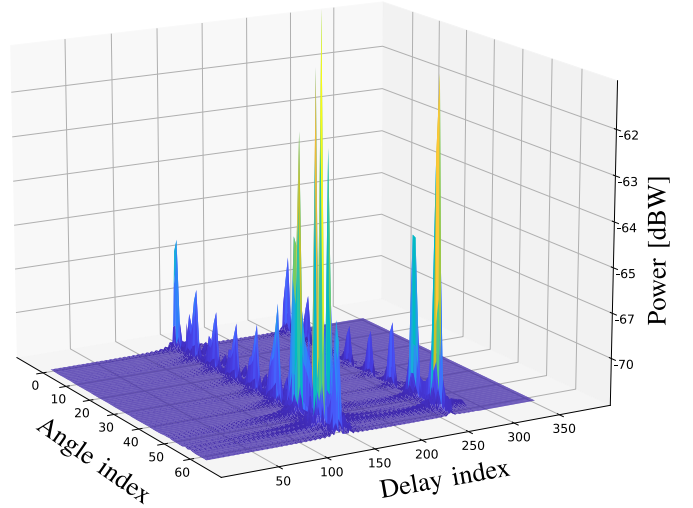

Fig. 3. Example of a sparse ADCPM with $M = N = 8$ antennas at the BS and $N_g = 352$ cyclic prefix duration in terms of sampling intervals T_s, simulated in an urban road environment with ray-tracing software.

DL model's foundation. Then, we describe the network input, i.e., the ADCPM fingerprint, followed by the definition of the DAKDM. Finally, we define the loss function used to train the model.

A. Problem Formulation

We consider a semi-supervised setting in which we are given a training dataset $\mathcal{S}^{\text{train}}$ comprising only normal data, i.e., $\boldsymbol{X}_i$ sampled from p_{X}, and a smaller testing data $\mathcal{S}^{\text{test}}$ comprising normal (label $y_i = 0$) and anomalous data (label $y_i = 1$). Here, we refer to LOS samples as normal, while we consider NLOS samples as anomalous. Nevertheless, the choice of normal/anomalous condition is arbitrary and could be customized to the specific scenario, as the proposed model would still be valid in both cases, i.e., LOS or NLOS samples as normal data. Usually, the high-dimensional distribution of normal samples p_{X} is complex and unknown. Thus, the objective is to first elaborate $\mathcal{S}^{\text{train}}$ such that we can learn its manifold distribution and, subsequently, during inference time, identify the anomalous samples in $\mathcal{S}^{\text{test}}$ as outliers. The mapping of the high-dimensional data is carried out using a DL model $f(\cdot)$ that learns the normal data distribution while also attempting to reduce an anomaly score $A(\boldsymbol{X}_i)$ given as output. The higher the anomaly score of $A(\boldsymbol{X}_i)$ for a test sample $\boldsymbol{X}_i$, the higher probability that $\boldsymbol{X}_i$ is anomalous. For evaluation, a threshold (η) criterion is applied, i.e., $A(\boldsymbol{X}_i) > \eta$ denotes an anomaly, based on a predefined false positive rate (FPR).

B. DL Input

We employ (9) as input to the neural network for NLOS identification, as this matrix represents the clustered multipath structure of the channel and embeds the information on LOS/NLOS propagation conditions that we are interested to extract. Moreover, the sparsity of the matrix helps the CNN in features extraction since the first layers of CNN are usually sparse and they gather the more discriminant features [94]. In Fig. 3 we can see the ADCPM P_k composed by MN angle
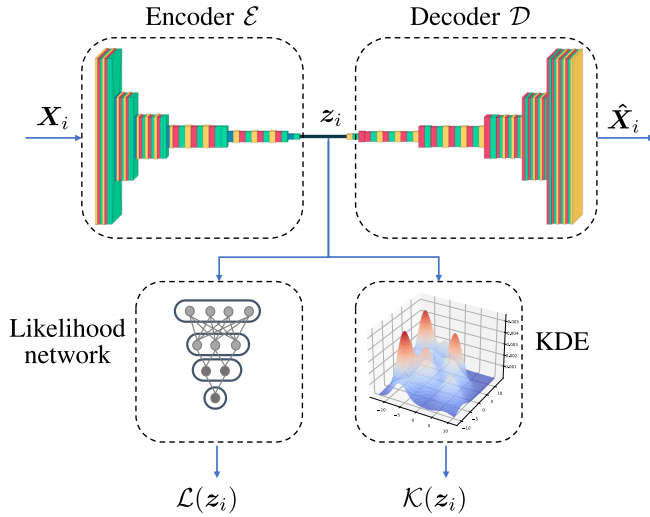
Fig. 4. Structure of the proposed deep autoencoding kernel density model (DAKDM) composed by an autoencoder (AE), a kernel density estimation (KDE) and a likelihood network.

directions and N_g delay samples. The sparsity of the matrix is well-visible even without a huge number of antennas or sample's resolution. From now on, for simplicity of notation, we drop the index k related to user k and we denote the i-th input sample as $\boldsymbol{X}_i = \boldsymbol{P}_i$.

C. Deep Autoencoding Kernel Density Model

The proposed DAKDM system for NLOS identification involves three main elements: an AE, a KDE model and a likelihood network. The model can be seen in Fig. 4. The AE comprises an encoder $\mathcal{E}(\cdot)$, that elaborates the i-th input $\boldsymbol{X}_i \in \mathbb{R}^{MN \times N_\mathrm{g}}$ into a latent representation $\boldsymbol{z}_i \in \mathbb{R}^m$, and a decoder $\mathcal{D}(\cdot)$, that carries out an inverse transformation to return to the original high-dimensional distribution, obtaining $\hat{\boldsymbol{X}}_i$. The latent distribution $p_\mathbf{z}$ may have any form, i.e., it is not constrained to belong to any specific PDF family. This makes the proposed approach general enough to be applied to any channel environment.

The distribution $p_\mathbf{z}$ is automatically learned by the KDE block of the system (see Fig. 4). The KDE is a non-parametric method to estimate any distribution directly from a set of samples drawn from it. Given a set of samples $\{\boldsymbol{z}_j\}_{j=1}^{N_\mathrm{s}}$ from $p_\mathbf{z}$, we define the KDE $\mathcal{K}(\cdot)$ applied to sample $\boldsymbol{z}_i$ as [95]:

$$\mathcal{K}(\boldsymbol{z}_i|\{\boldsymbol{z}_j\}_{j=1}^{N_\mathrm{s}}) = \frac{1}{N_\mathrm{s}} \sum_{j=1}^{N_\mathrm{s}} k_h(\boldsymbol{z}_i - \boldsymbol{z}_j), \qquad (10)$$

where $k_h : \mathbb{R}^m \to \mathbb{R}$ is a kernel function with bandwidth h which regulates the balance between the estimator's variance and bias. The kernel employed in this paper is the widely known Gaussian kernel:

$$k_h(\boldsymbol{x}) = e^{-\frac{|\boldsymbol{x}|^2}{2h^2}}. \qquad (11)$$

The output of a KDE, trained only with normal latent samples $\{\boldsymbol{z}_j\}_{j=1}^{N_\mathrm{s}}$, can be seen as the likelihood of the test sample to belong to the normal distribution. Thus, the derived anomaly score can be obtained as $A_\mathcal{K}(\boldsymbol{z}_i) = -\log(\mathcal{K}(\boldsymbol{z}_i|\{\boldsymbol{z}_j\}_{j=1}^{N_\mathrm{s}}))$. However, the downsides of KDE lie in the fact that it requires

---

Algorithm 1 Mini-Batch Training Procedure

1: **procedure** TRAINING(batch size N_s) ▷ Batch number j
2:     **for** $i = 1, 2, \dots, N_\mathrm{s}$ **do**
3:         Encode incoming signal $\boldsymbol{X}_i$: $\boldsymbol{z}_i = \mathcal{E}(\boldsymbol{X}_i)$.
4:         Compute anomaly score: $A_\mathcal{L}(\boldsymbol{z}_i) = -\log(\mathcal{L}(\boldsymbol{z}_i))$.
5:         Compute KDE prediction:
            $A_\mathcal{K}(\boldsymbol{z}_i) = -\log(\mathcal{K}(\boldsymbol{z}_i^j|\{\boldsymbol{z}_l^{j-1}\}_{l=1}^{N_\mathrm{s}}))$.
6:     **end for**
7:     Compute loss function $\mathrm{L}_\mathrm{tot}^j$.
8:     Perform backward-pass.
9: **end procedure**

---

storing all the training dataset to estimate the density function at inference time.

The idea to solve this issue is to first reduce the number of samples N_s used to estimate the distribution, and then approximate the output of the KDE with a NN that is much faster in the prediction. We call the NN to estimate the output of the KDE as likelihood network and denote it with $\mathcal{L}(\cdot)$. The logical steps for the training procedure with a batch of N_s samples are described in Algorithm 1. First, we encode the input with the encoder. Then, we extract the anomaly score as $A(\boldsymbol{X}_i) \triangleq A_\mathcal{L}(\boldsymbol{z}_i) = -\log(\mathcal{L}(\boldsymbol{z}_i))$ and we compute the KDE prediction $A_\mathcal{K}(\boldsymbol{z}_i)$. Finally, we compute the loss function which is described in Sec. IV-D and perform the backward pass. The key aspect here is that the KDE output is computed using the previous mini-batch, i.e., $\mathcal{K}(\boldsymbol{z}_i^j|\{\boldsymbol{z}_l^{j-1}\}_{l=1}^{N_\mathrm{s}})$. This permits to avoid storing all the training dataset to estimate the density function. The underlying assumption is that the batch size N_s is able to give a good representation of the likelihood through the KDE. Formally:

$$\mathrm{KL}\left(\mathcal{K}\left(\boldsymbol{z}_i|\{\boldsymbol{z}_j\}_{j=1}^{N_\mathrm{s}}\right) \,\|\, p(\boldsymbol{z}_i)\right) \simeq 0, \qquad (12)$$

where $\mathrm{KL}(\cdot\|\cdot)$ is the Kullback-Leibler divergence. On the contrary, at inference time, we just check if the anomaly score $A_\mathcal{L}(\boldsymbol{z}_i) > \eta$. This implies that, during deployment, we can completely discard both the decoder $\mathcal{D}(\cdot)$ and the KDE $\mathcal{K}(\cdot)$, just relying on the faster prediction of the encoder $\mathcal{E}(\cdot)$ and likelihood network $\mathcal{L}(\cdot)$.

D. Loss Function

The objective of the loss function is to first induce the DAKDM to learn the latent representation of normal data and then to approximate $A_\mathcal{K}(\boldsymbol{z}_i)$ with $A_\mathcal{L}(\boldsymbol{z}_i)$. To this aim, we consider the training dataset $\mathcal{S}^\mathrm{train} = \{\boldsymbol{B}_j\}_{j=1}^{N_\mathrm{b}}$, where N_b is the number of batches in the training dataset and $\boldsymbol{B}_j = \{\boldsymbol{X}_i^j\}_{i=1}^{N_\mathrm{s}}$ is the j-th mini-batch with N_s samples. We define the total loss related to mini-batch j as follows:

$$\begin{aligned}
\mathrm{L}_\mathrm{tot}^j = &\frac{1}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} \mathrm{L}_\mathrm{rec}(\boldsymbol{X}_i^j, \hat{\boldsymbol{X}}_i^j) \\
&+ \frac{w_\mathrm{KL}}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} \mathrm{KL}_\mathrm{point}\left(\mathcal{L}(\boldsymbol{z}_i^j) \,\|\, \mathcal{K}\left(\boldsymbol{z}_i^j|\{\boldsymbol{z}_l^{j-1}\}_{l=1}^{N_\mathrm{s}}\right)\right) \\
&+ \frac{w_\mathrm{lik}}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} \left(-\log \mathcal{K}\left(\boldsymbol{z}_i^j|\{\boldsymbol{z}_l^{j-1}\}_{l=1}^{N_\mathrm{s}}\right)\right), \qquad (13)
\end{aligned}$$

where $\mathrm{L}_{\mathrm{rec}}(\boldsymbol{X}_i^j, \hat{\boldsymbol{X}}_i^j) = \|\boldsymbol{X}_i^j - \hat{\boldsymbol{X}}_i^j\|_2$ is the loss function that describes the reconstruction error given by the AE, w_{KL} and w_{lik} are the weighting parameters that control how much the single losses affect the objective function as a whole and $\mathrm{KL}_{\mathrm{point}}$ is the pointwise KL-divergence defined as:

$$\mathrm{KL}_{\mathrm{point}}(x\|y) = x\log\left(\frac{x}{y}\right). \tag{14}$$

With the second right-hand side of (13), we exploit the power of the likelihood network to learn the KDE output trained with the previous mini-batch. The choice of the loss function is motivated by the fact that, if assumption (12) holds, then we can write the contribution of $\boldsymbol{z}_i$ to the anomaly score with the following [96]:

$$\begin{aligned}
-\log p(\boldsymbol{z}_i) \lesssim &-\log\mathcal{K}\left(\boldsymbol{z}_i|\{\boldsymbol{z}_j\}_{j=1}^{N_{\mathrm{s}}}\right) \\
&+ \mathrm{KL}\left(\mathcal{L}(p(\boldsymbol{z}_i)|\boldsymbol{z}_i)\,\big\|\,p\left(\mathcal{K}\left(\boldsymbol{z}_i|\{\boldsymbol{z}_j\}_{j=1}^{N_{\mathrm{s}}}\right)|\boldsymbol{z}_i\right)\right),
\end{aligned} \tag{15}$$

where $\mathcal{L}(p(\boldsymbol{z}_i)|\boldsymbol{z}_i)$ is the likelihood network that provides the probability of $\boldsymbol{z}_i$ given $\boldsymbol{z}_i$. For the proof of (15), please refer to Appendix A. We directly insert the first right-hand side of (15) in the loss function to induce the AE to decrease the anomaly score, thus increasing the likelihood. On the other hand, we do not have a KDE that provides the probability of its predictions, therefore we consider the $p(\mathcal{K}(\boldsymbol{z}_i|\{\boldsymbol{z}_j\}_{j=1}^{N_{\mathrm{s}}})|\boldsymbol{z}_i)$ as a single deterministic value that we approximate through the likelihood network.

V. SIMULATION EXPERIMENTS

A. 5G NR Network Simulation

To evaluate the proposed DAKDM method for NLOS identification, we simulate realistic CSI data based on the 5G NR clustered delay line (CDL) channel model [97] which is characterized by a maximum bandwidth of 2 GHz over the whole frequency range of 0.5 GHz to 100 GHz. We simulate the wave propagation using a ray-tracing method [98], [99], [100] provided by Antenna Toolbox Matlab package where the propagation pathways from the UE to the BSs are computed based on the surface geometry from a map file. Ray-tracing uses the shooting and bouncing ray (SBR) method [101], accounting for up to 10 path reflections. The method does not take into account buildings' windows and possible foliage, which would require a high-definition 3D mapping of the environment or a complex simulation with artificially created maps. The channel model is then produced by coupling all the paths taking into account the small-scale fading due to the UE's movement, angle spread and cluster properties. This permits to achieve spatial consistency, meaning that two adjacent positions present similar channel characteristics due to comparable scattered environments.

B. Urban Mobility Scenario

For the experiments, we simulate a 3GPP urban micro (UMi) scenario in an area of 1000 m × 1000 m, near the Leonardo Campus of Politecnico di Milano, with specific parameters described in [85]. As shown in Fig. 5, the scenario comprises 19 urban sites, placed in an hexagonal layout with Inter-Site Distance (ISD) of 200 m, each equipped with 3 cells. The BS antennas are characterized by an UPA configuration with $M = N = 8$ elements and a downtilt of 15°. The transmission power is 44 dBm and each antenna element was defined using the specifications in [102], providing a front-to-back ratio of about 30 dB and a maximum gain of 8 dBi.

The vehicular UEs move in the area traveling along different trajectories generated with SUMO software which replicates actual traffic patterns on a particular route network. We generate up to 50 trajectories sampled every second, for a total simulation time of 170 s. Each UE is equipped with an omnidirectional antenna and transmits the 5G standard compliant sounding reference signals (SRSs) to all the BSs in the area using a carrier frequency $f_{\mathrm{c}} = 30$ GHz and a transmission bandwidth $B = 400$ MHz. The BSs, which can be in either LOS or NLOS condition due to occlusions and reflections, demodulate the OFDM signal and estimate the channel using a least squares (LS) estimator. Subsequently, they obtain the channel fingerprint using the estimated channel response to compute the angle-delay channel structure (7) and then the associated power structure (8).

For the experiments, we do not consider the multi-user interference (MUI), but it is worth mentioning few considerations for possible real implementations of the method. In practice, the BSs can adopt various techniques to manage the inaccuracy of channel estimation due to factors such as the MUI. One common technique is to use channel estimation algorithms that are robust to MUI, such as linear minimum mean-square error (LMMSE) [103] which obtains sub-optimal performance (sub-optimal as it does not use the knowledge of the full CSI) with moderate computational complexity. Additionally, other techniques may rely on non-linear pre-coding schemes which have been found to provide near-optimal performance [104], [105]. In the standard of 5G-NR, codebook-based MIMO precoding techniques have been proposed and they are described in the 3GPP technical specifications (TS) 38-214 [106] and 38-211 [107]. With latest releases, i.e., Rel 16 and 17, MU-MIMO codebook (type II) has been improved with the reduction of the feedback overhead. By implementing these techniques, the MUI is highly reduced and the residual interference resembles to background noise. Moreover, in case the model has been trained in a channel in presence of non-null interference, we would have an even-broader LOS distribution characterization, which would be beneficial in case of single-user transmission.

C. CSI Dataset

In the offline phase, each BS is assured to gather LOS only realizations of the channel, composing a training dataset for the blockage detection. In the online phase, on the other hand, we create the test dataset adopting unobserved positions of the UEs and collecting a balanced number of samples between LOS and NLOS conditions. We saved more than $7.5 \cdot 10^4$ and $8.6 \cdot 10^4$ samples in the training and testing set, respectively. Before the training, all the samples are standardized (i.e., transformed such that the mean intensity is 0 with standard deviation of 1) and shuffled at each epoch.
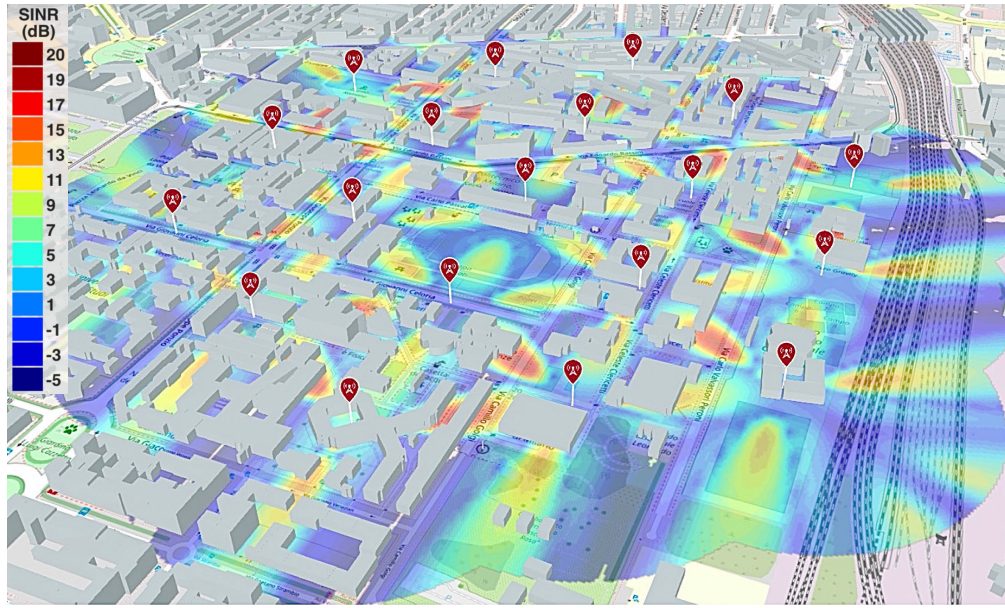
Fig. 5. UMi scenario composed of 19 sites in the area of Politecnico di Milano, Leonardo campus, Italy. The signal-to-interference-plus-noise ratio (SINR) is shown in downlink as a function of the UE position when BSs use trisectorial cells with broadside transmission.

MATLAB 2022a is used to create the channel fingerprints of the data points, while the DL model for training and testing is implemented using Pytorch [108] (v1.12 with Python 3.7.11). We run our simulations on a workstation with an Intel(R) Xeon(R) Silver 4210R CPU @ 2.40 GHz, 96 GB of RAM and a Quadro RTX 6000 24 GB GPU. The testing times, described in Sec. V-F1, only apply to the run time on Pytorch 1.12. Unless otherwise specified, we train the model for a number of epochs $E = 30$ with a batch size $N_s = 64$. w_{KL} and w_{lik} are both empirically set to 0.1. We adopted the Adam [109] with an initial learning rate $lr = 10^{-4}$, and momentum $\beta_1 = 0.9$, $\beta_2 = 0.999$.

D. DL Model Characteristics

For the AE part, we adopt the Segnet architecture [110] with one single channel encoder and decoder. The upsampling layers employ the encoder pool indices to create a sparse feature mapping which is ideal for reproducing the sparse ADCPM input. The AE is the most complex part of the model, however, at testing time, we use only the encoder part, thus halving the inference time if compared with VAE models or in general solutions that adopt the reconstruction error as a monitoring feature.

On the other hand, we develop the likelihood network using a simpler multi-layer perceptron (MLP) which is able to learn whatever non-linear function. The network can be found in Table I. To cope with the overfitting we adopt the dropout technique [111] after each activation function. Furthermore, we insert a single batch normalization layer [112] right before the ReLU function. This is done to avoid that the output of the network will converge to a unique value after a long training.

E. Baselines

To evaluate the performances of the proposed model, we compare it against a number of DL approaches proposed in the literature to solve anomaly detection problems:

TABLE I
LIKELIHOOD NETWORK LAYER COMPOSITION

Layer Num.	Type	Output Size
0	Input	8×1
1	Linear + Tanh + Dropout	8×1
2	Linear + Tanh + Dropout	5×1
3	Linear	1×1
4	BatchNorm1d	1×1
5	ReLU	1×1

- DAGMM [84]. Single-stage training model composed by an AE and a GMM used for learning the latent feature distribution. The membership weights, which represent the probability that a given data point belongs to each component, are usually computed with the expectation (E)-step of the expectation-maximization (EM) algorithm used for the GMM fitting. However, in this case, the membership weights are produced by an estimation DL network.
- AE-KDE [83]. Double-stage training model in which first the AE is trained and then a KDE is used to learn the distribution of latent features from all the training dataset. The bandwidth and the kernel are the same of DAKDM.
- VAE [76]. Auto-encoding variational Bayes applied to NLOS identification. Here, the sampling mechanism is mandatory since we need to sample new latent variables from the learned probability distribution, i.e., in this case a Gaussian distribution. The anomaly score $A(\boldsymbol{X}_i)$ is computed as $A(\boldsymbol{X}_i) = \frac{1}{N_m} \sum_{j=1}^{N_m} L_{rec}(\boldsymbol{X}_i, \hat{\boldsymbol{X}}_i^j)$, where N_m is the number of samples. As suggested by the authors, we draw 10 samples from the latent space representation to derive the anomaly score.

- GANomaly [113]. Deep-generative model composed by an AE, a second encoder and a discriminator. The model minimizes simultaneously the reconstruction error, the encoder loss given by the second encoder and the adversarial loss yielded by the discriminator.

For a fair comparison, we give the same input to each model and we adopt the unchanged architecture for the encoders and decoders with respect to DAKDM. Therefore, we adopt the same number of latent features for all architectures.

In addition to DL model baselines, we compare our method with classical ML and statistical algorithms. In particular, we implement:

- JLRT [63]. Joint-likelihood ratio test considering the statistics of the kurtosis, mean excess delay and root-mean-square delay spread. The PDFs of the statistics are approximated as log-normal distributions and they are considered independent of each other.
- RF [66]. Random forest model with 100 trees and, as input features, the Rician K-factor, root-mean-square delay spread, mean excess delay and dominant channel tap.
- CORR [75]. Pearson correlation coefficient computed using a reference set of LOS ADCPM sliced in the direction of arrival with higher received power. We gathered 100 LOS reference signals and we considered only the samples comprising 10 points before and 100 points after the first peak. The likelihood of a test input is obtained by averaging the correlation coefficient with the reference LOS signals.
- OC-SVM [58]. One-class support vector machine which computes the smallest hyper-sphere containing normal, i.e., LOS, samples. We use the score function as a likelihood measure. As regards the feature selection, we adopt both static channel characteristics as the maximum received power, kurtosis, skewness, rising time, root-mean-square delay spread, Rician K-factor, angular spread of arrival and both time-varying features [114] like the angular variant of arrival.

Note that, while CORR and OC-SVM are semi-supervised learning algorithms, JLRT and RF are supervised learning methods since they require statistics/samples of both classes. The models and algorithms are run independently by each BS, after the UE uplink transmission. The training, if required, is performed before the validation procedure at each BS using the locally collected input samples.

F. Results

1) Inference Timings: In this assessment, we want to measure the time required by each DL model to predict the output of a sample. This is of particular relevance in real-time applications where the inference time must be as low as possible. An example is the vehicular applications where the end-to-end latency must be contained within 100 ms or less [115]. In Fig. 6, we show the boxplot of the inference time for each sample over the whole testing dataset. We notice that the proposed DAKDM is able to predict the anomaly score in half of the time with respect to DAGMM as it does not require
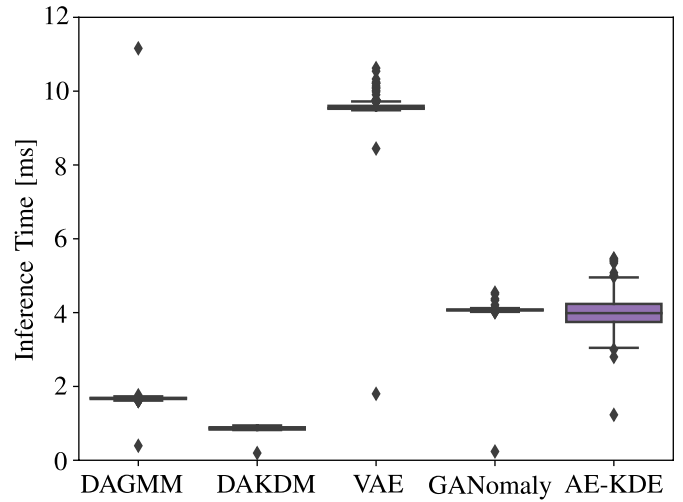

Fig. 6. Boxplot of the distribution of the inference time per sample [ms], varying the adopted DL models.

the decoder part prediction. Moreover, GANomaly and AE-KDE models require up to 4 ms for a single prediction. This is because GANomaly holds a more complex model, while AE-KDE has to pass through the whole training dataset for a single prediction. Finally, VAE takes about ten times more than DAKDM due to the sampling strategy.

2) Batch Size: This assessment has the goal of verifying how the batch size N_s affects the performances of the proposed DAKDM. Theoretically, N_s should be large enough to generate a good representation of the latent features' distribution. To verify this behaviour, in Fig. 7 we analyze the anomaly score $A_{\mathcal{L}}(z_i)$ of normal (Fig. 7a) and anomalous (Fig. 7b) samples in the testing dataset after 30 epochs for $N_s = \{8, 16, 32, 64, 128\}$. To avoid singular issues due to possible zeros values given as input to the logarithm, we shift the likelihood distribution as $A_{\mathcal{L}}(z_i) = -\log(\mathcal{L}(z_i)+1)$. The first thing to notice is that the anomalous score of normal data is lower than the abnormal data and this is because the likelihood network outputs higher values for samples with normal distributions. Second, we observe that decreasing N_s for normal data, will produce lower mean and variances distributions, thus enhancing the NLOS identification capabilities. This is due to the fact that with a large batch size, the model struggles to learn the pointwise KL-divergence since in the loss function we have the contributions of many points. On the contrary, with lower batch sizes, the likelihood network learns exactly which value assign to each latent representation. Reducing N_s has the benefit of being suitable for simpler devices with low computational capabilities, in exchange for higher training times. As a trade-off between performances and training times, we choose $N_s = 16$.

3) Hyper-Parameters Tuning: This experiment aims of tuning the main hyper-parameters related to the density models, i.e., the bandwidth h of the KDE for DAKDM and AE-KDE and the number of GMM components, denoted with g, for DAGMM. In Fig. 8, we report the area under the curve (AUC) obtained in the testing set after 30 training epochs varying the bandwidth $h \in \{0.05, 0.1, 0.2, 0.4, 0.8, 1.6, 3, 6, 12\}$ (Fig. 8a)
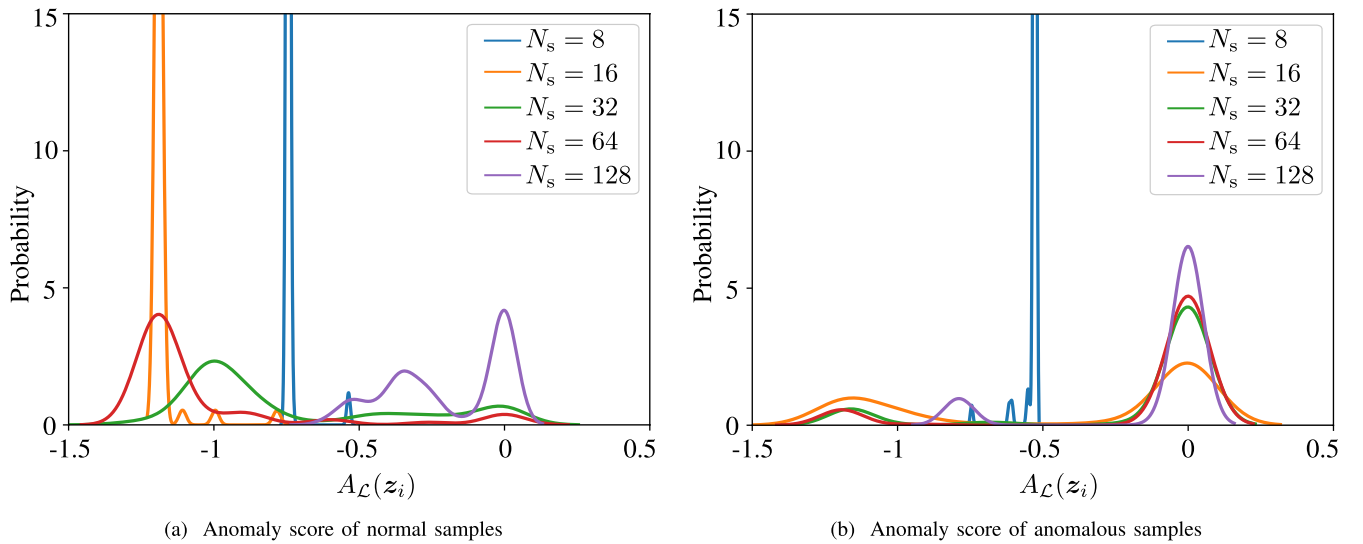
(a) Anomaly score of normal samples

(b) Anomaly score of anomalous samples

Fig. 7.   Comparison of the anomaly score $A_{\mathcal{L}}(\boldsymbol{z}_i)$ after 30 epochs between normal and anomalous data, for varying batch size N_{s}.

for DAKDM and the number of GMM components $g \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ (Fig. 8b) for DAGMM. Starting from Fig. 8a, we notice that the higher AUC is reached by $h = 0.2$ and that for not optimal values, the AUC can differ significantly. This is somehow due to the range values of $\boldsymbol{z}_i$ and to the number of points that we have. Since, in practice, we have few anomalous samples, for tuning the bandwidth we can simply rely on maximizing the likelihood of normal samples varying the bandwidth. Comparing the results with Fig. 8b, we see that DAGMM is not able to achieve high peaks of performances in average, i.e., above 90% of AUC. This means that the latent distribution cannot be well approximated with less than 10 Gaussian distributions. Clearly, increasing g will improve the performance but at the cost of a higher complexity of the DAGMM estimation network.

4) Performance Comparison: In this last assessment, we compare the performances of the proposed DAKDM with the models described in Sec. V-E. In Table II, we report the average AUC after 10 runs and the F-score, Accuracy, Precision and Recall using a threshold on the anomaly score related to 20% of FPR. We notice that the performances of the AE-KDE (highlighted in green) are superior with respect to all the others. The reason behind this is that the AE-KDE represents the perfect unfeasible upper-bound, i.e., it requires storing the whole training dataset for inference and thus it can perfectly reconstruct the latent features distribution. On the other hand, the lower-bound is represented by the statistical JLRT method (highlighted in red), which obtains an AUC of 63%. This method assumes the independence of few hand-crafted features, which may not hold in any situation. The second non ML-based method CORR reaches an AUC of 64%, meaning that the LOS reference signals are not a good representation of the LOS distribution. The OC-SVM, a classical ML method, achieves a slightly higher AUC due to its capabilities of projecting the original features in a higher hyper-space (kernel trick). However, its main limitations lay in the features-choice which, for sparse and

high-dimensional spaces, constitutes a non-trivial task. Moreover, we can notice that the precision (94%) is much higher with respect to the recall (60%). This means that OC-SVM tends to classify all test samples as LOS, learning a rough, i.e., too general, LOS distribution. Finally, among classical ML methods, the RF achieves the highest performances by reaching an AUC of 79%. However, we remark that this method requires the knowledge of NLOS samples, thus it holds an advantage with respect to semi-supervised learning methods.

Focusing now on the DL models, numerical results show that they highly outperform the classical ML and statistical algorithms. Indeed, while deterministic feature extraction might be more suitable for low-dimensional or simple channels, using the raw ADCPM as input to the CNN structure allows the DL models to utilize the full potential of automatic feature extraction, which contributes to the superior performance of the DL methods in comparison to classical ML and statistical algorithms. However, this does not exclude the possibility of incorporating deterministic features in future work to further improve the performance of the proposed model. Among the DL methods, DAKDM (highlighted in bold) and VAE hold the highest AUCs if compared with DAGMM and GANomaly. In particular, DAKDM and VAE outperform DAGMM and GANomaly of 7% and 16%, respectively. The reasons behind this are that GANomaly is a very complex network that requires a non-negligible effort in hyper-parameter tuning and optimization, with additional issues in training stability [79]. On the other hand, DAGMM is not able to accurately learn the LOS latent feature representation due to its limited number of Gaussian components. Both DAKDM and VAE achieves the highest performances, i.e., 95% and 96% of AUC, but with two different methods. While VAE imposes a simple latent distribution, DAKDM automatically learns the low-dimensional LOS distribution thanks to the KDE in training phase. However, the main advantage of DAKDM is that it does not require sampling procedures
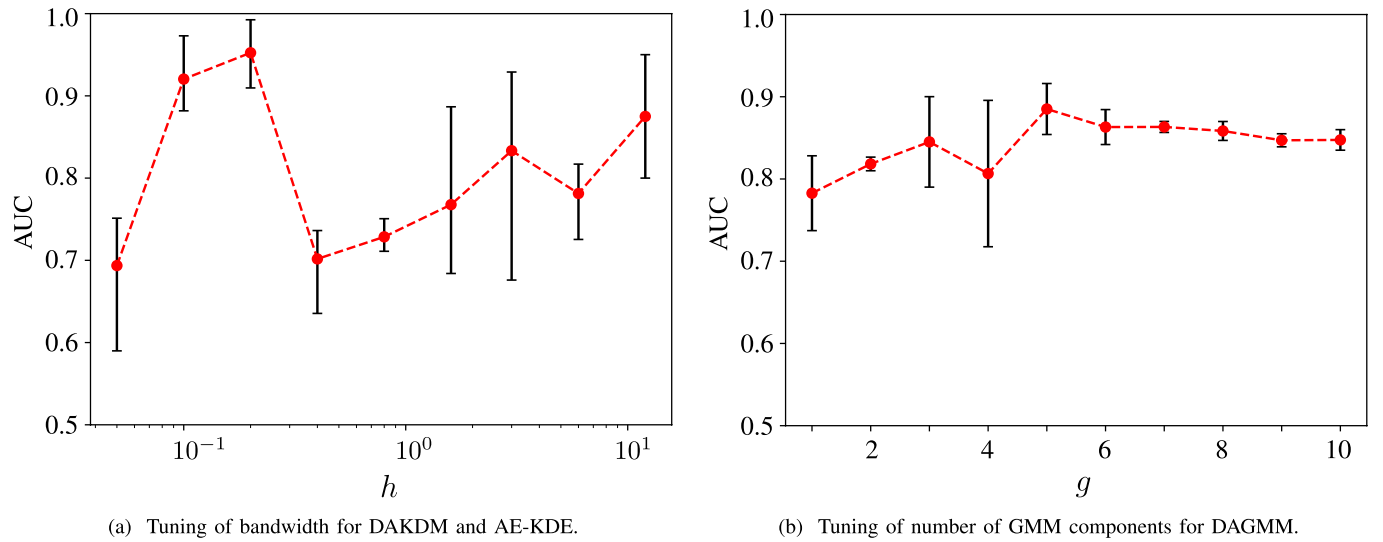
(a) Tuning of bandwidth for DAKDM and AE-KDE.

(b) Tuning of number of GMM components for DAGMM.

Fig. 8. Comparison of the AUC reached after 30 epochs using (a) DAKDM and (b) DAGMM, for varying bandwidth h and number of GMM components g, respectively. The mean value (red dot) is plotted together with the associated uncertainty (error bar) computed using the maximum and minimum values of AUC as boundaries.

TABLE II

COMPARISON ON MEAN PERFORMANCE INDICATORS AFTER 10 DIFFERENT RUNS BETWEEN THE PROPOSED DAKDM AND THE BASELINES

	AUC	F-score	Accuracy	Precision	Recall
AE-KDE	0.9981	0.9376	0.9196	0.8845	0.9973
VAE	0.9620	0.8654	0.8491	0.9408	0.8013
DAKDM	**0.9535**	**0.8625**	**0.8456**	**0.9406**	**0.7972**
DAGMM	0.8919	0.8131	0.7767	0.8305	0.7964
GANomaly	0.8214	0.8337	0.8068	0.8697	0.8006
RF	0.7979	0.7753	0.7506	0.8495	0.7131
OC-SVM	0.7735	0.7419	0.7435	0.9477	0.6096
CORR	0.6449	0.7153	0.6152	0.6474	0.7992
JLRT	0.6395	0.5636	0.5962	0.7658	0.4459

and necessitates only 10% of the inference time needed by VAE (see Sec. V-F1). This makes it suitable for low-latency and mission-critical applications such as V2X networks for automated driving.

VI. CONCLUSION

This paper addressed the problem of high-dimensional channel distribution characterization for next generation cellular networks. In order to demonstrate the method, we tackle the problem of NLOS identification in a mm-wave MIMO system with sparse space-time channel responses. We model the problem within the semi-supervised anomaly detection framework where LOS samples correspond to normal data with peculiar characteristics and distributions. We propose a deep autoencoding kernel density model (DAKDM) where the manifold distribution of normal data is elaborated with an AE that takes as input the sparse ADCPM which univocally map the position-dependent features of the channel. The AE is jointly learned together with a likelihood network which is trained to learn the output of a KDE that directly estimates the distribution of the latent features.

The DAKDM has the main advantage of learning whatever latent distribution without storing the whole training dataset.

We validated the model in a 5G standard compliant UMi scenario simulated with Matlab ray-tracing package, permitting spatial channel consistency between adjacent positions. The UEs are vehicles which move in the area according to dynamics simulated by the SUMO software. Compared with DL state-of-the-art models, results showed that the proposed DAKDM is much more efficient, both in terms of inference time and computational requirements. In particular, DAKDM holds a prediction time per sample which is up to one fourth and one tenth of GANomaly and VAE, respectively. This makes it appropriate for edge devices with strong latency requirements for mission-critical applications. From a performance point-of-view, DAKDM is able to achieve similar performances of the top-performer VAE, outperforming GMM-based method such as DAGMM of about 7%.

In the next years, ML and in particular DL methods are expected to play a crucial role in next generation cellular networks. Communication systems, but also localization techniques, are required to increase performance capabilities and types of services to accomplish increasingly high standards. Thus, DL-based methods as the proposed DAKDM become essential to push further the performances. A natural extension of our work would be to integrate NLOS mitigation into the system in order to compensate the induced error given by lack of visibility or directly integrate DL techniques into positioning algorithms suited for high complexity environments. A further direction of research could be the extension to a cooperative inference framework where BSs exchange mutual-soft information for accuracy enhancement. Moreover, more realistic environments with simulated foliage and dynamic obstructions should be explored. Challenges are represented by NLOS situations, changes in the environment and lack of possible representative samples for each feasible location.
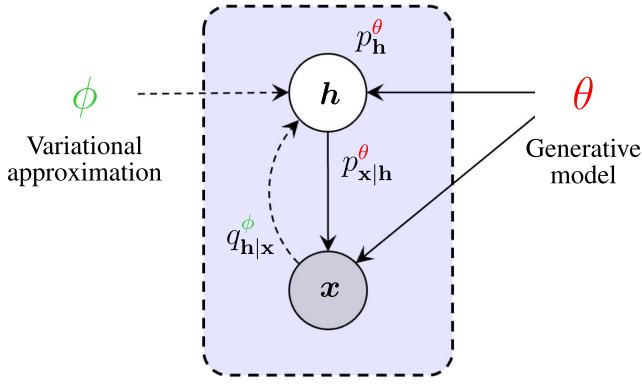
Fig. 9. Variational inference problem as described in [116]. The generative model is indicated by solid lines while variational approximation is reported with dashed lines.

APPENDIX A
PROOF OF (15)

In this Appendix we provide a proof of the anomaly score contribution given by z_i. From the variational inference approach [96], we can note that the likelihood network performs the same objective of latent variable inference. To see this parallelism, we recall the variational inference context where we are given an observation h (i.e., latent variable) from the prior distribution p_{h}^{θ} with parameters θ. Subsequently, a datapoint x is generated from $p_{\mathsf{x}|\mathsf{h}}^{\theta}$, which is considered intractable. The objective is to estimate the exact posterior $p_{\mathsf{h}|\mathsf{x}}^{\theta}$, also intractable, with a simpler variational posterior $q_{\mathsf{h}|\mathsf{x}}^{\phi}$ with parameters ϕ. For a graphical representation of the problem, please refer to Fig. 9 [116].

We can view the datapoint x as a compact representation of the channel z_i and the latent variable h as the distribution $p(z_i)$. Let us denote the probability as $\mathsf{s}_i = p_{\mathsf{z}}(z_i)$, which can be approximated with the KDE $\hat{\mathsf{s}}_i = \mathcal{K}\big(z_i | \{z_j\}_{j=1}^{N_{\mathsf{s}}}\big)$, the likelihood network $\mathcal{L}^{\phi}(\cdot)$ (with parameters ϕ) will be acting as the variational distribution $q^{\phi}(\cdot)$. Following this parallelism, the contribution of z_i to the anomaly score, i.e., negative log-likelihood, can be written as follows:

$$-\log p_{\mathsf{z}}(z_i) = -\log \int_s p_{\mathsf{z},\mathsf{s}_i}\big(z_i, s\big)\, ds \qquad (16)$$

$$= -\log \int_s \mathcal{L}^{\phi}\big(s|z_i\big) \frac{p_{\mathsf{z},\mathsf{s}_i}\big(z_i, s\big)}{\mathcal{L}^{\phi}\big(s|z_i\big)}\, ds$$

$$\leq -\int_s \mathcal{L}^{\phi}\big(s|z_i\big) \log \frac{p_{\mathsf{z},\mathsf{s}_i}\big(z_i, s\big)}{\mathcal{L}^{\phi}\big(s|z_i\big)}\, ds \qquad (17)$$

where (17) is the consequence of Jensen's inequality. It follows that

$$-\log p_{\mathsf{z}}(z_i) \leq -\mathbb{E}_{\mathcal{L}^{\phi}}\Big\{ \log p_{\mathsf{z},\mathsf{s}_i}\big(z_i, s\big) - \log \mathcal{L}^{\phi}\big(s|z_i\big) \Big\}$$

$$= -\log p_{\mathsf{z}}(z_i) + \mathrm{KL}\big(\mathcal{L}^{\phi}\big(s|z_i\big) \,\big\|\, p_{\mathsf{s}_i|\mathsf{z}}(s|z_i)\big)$$

$$\simeq -\log \mathcal{K}\big(z_i | \{z_j\}_{j=1}^{N_{\mathsf{s}}}\big)$$

$$+ \mathrm{KL}\big(\mathcal{L}^{\phi}\big(s|z_i\big) \,\big\|\, p_{\hat{\mathsf{s}}_i|\mathsf{z}}(\hat{s}|z_i)\big), \qquad (18)$$

where the approximation in (18) is due to (12), concluding the proof.

REFERENCES

[1] M. Win et al., "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011. [Online]. Available: http://ieeexplore.ieee.org/document/5762798/

[2] A. Conti et al., "Location awareness in beyond 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 22–27, Nov. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9665420/

[3] S. Fischer, *5G and Beyond: Fundamentals and Standards*, X. Lin and N. Lee, Eds. Cham, Switzerland: Springer, 2021, doi: 10.1007/978-3-030-58197-8.

[4] F. Mogyorósi et al., "Positioning in 5G and 6G networks—A survey," *Sensors*, vol. 22, no. 13, p. 4757, Jun. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/13/4757

[5] O. Kanhere and T. S. Rappaport, "Position location for futuristic cellular communications: 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 70–75, Jan. 2021. [Online]. Available: http://ieeexplore.ieee.org/document/9356512/

[6] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1458275/

[7] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1458287/

[8] M. Z. Win, Z. Wang, Z. Liu, Y. Shen, and A. Conti, "Location awareness via intelligent surfaces: A path toward holographic NLN," *IEEE Veh. Technol. Mag.*, vol. 17, no. 2, pp. 37–45, Jun. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9781656/

[9] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, "Location awareness in beyond 5G networks via reconfigurable intelligent surfaces," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2011–2025, Jul. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9729782/

[10] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8240645/

[11] R. Karlsson and F. Gustafsson, "The future of automotive localization algorithms: Available, reliable, and scalable localization: Anywhere and anytime," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 60–69, Mar. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7872535/

[12] X. Wang, T. Li, S. Sun, and J. Corchado, "A survey of recent advances in particle filters and remaining challenges for multitarget tracking," *Sensors*, vol. 17, no. 12, p. 2707, Nov. 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/12/2707

[13] M. Chiani, A. Giorgetti, and E. Paolini, "Sensor radar for object tracking," *Proc. IEEE*, vol. 106, no. 6, pp. 1022–1041, Jun. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8365918/

[14] N. Wang and C. K. Ahn, "Coordinated trajectory-tracking control of a marine aerial-surface heterogeneous system," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 6, pp. 3198–3210, Dec. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9339888/

[15] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, "Efficient multisensor localization for the Internet of Things: Exploring a new class of scalable localization algorithms," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 153–167, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8454389/

[16] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014. [Online]. Available: https://ieeexplore.ieee.org/document/6740844/

[17] A. A. Saucan and M. Z. Win, "Information-seeking sensor selection for Ocean-of-Things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10072–10088, Oct. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9086780/

[18] B. Guo et al., "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Comput. Surv.*, vol. 48, no. 1, pp. 1–31, Sep. 2015. [Online]. Available: https://dl.acm.org/doi/10.1145/2794400

[19] D. E. Boubiche, M. Imran, A. Maqsood, and M. Shoaib, "Mobile crowd sensing—Taxonomy, applications, challenges, and solutions," *Comput. Hum. Behav.*, vol. 101, pp. 352–370, Dec. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0747563218305181

[20] G. Pasolini, P. Toppan, F. Zabini, C. D. Castro, and O. Andrisano, "Design, deployment and evolution of heterogeneous smart public lighting systems," *Appl. Sci.*, vol. 9, no. 16, p. 3281, Aug. 2019. [Online]. Available: https://www.mdpi.com/2076-3417/9/16/3281

[21] A. Eskandarian, C. Wu, and C. Sun, "Research advances and challenges of autonomous and connected ground vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 683–711, Feb. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/8936542/

[22] B. C. Tedeschini, M. Brambilla, L. Barbieri, and M. Nicoli, "Addressing data association by message passing over graph neural networks," in *Proc. 25th Int. Conf. Inf. Fusion (FUSION)*, Linköping, Sweden, Jul. 2022, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/document/9841233/

[23] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, Dec. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/8246850/

[24] S. Ko, H. Chae, K. Han, S. Lee, D. Seo, and K. Huang, "V2X-based vehicular positioning: Opportunities, challenges, and future directions," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 144–151, Apr. 2021. [Online]. Available: http://ieeexplore.ieee.org/document/9382021/

[25] *Study on Enhancement of 3GPP Support for 5G V2X Services*, document TR 22.886, 3GPP, Dec. 2018, version 16.2.0.

[26] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8421288/

[27] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Communication-efficient distributed learning over networks—Part I: Sufficient conditions for accuracy," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1081–1101, Apr. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10043782

[28] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Communication-efficient distributed learning over networks—Part II: Necessary conditions for accuracy," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1102–1119, Apr. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10043777

[29] A. A. D'Amico, A. D. J. Torres, L. Sanguinetti, and M. Win, "Cramér–Rao bounds for holographic positioning," *IEEE Trans. Signal Process.*, vol. 70, pp. 5518–5532, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9950340

[30] S. Mazuelas, Y. Shen, and M. Z. Win, "Spatiotemporal information coupling in network navigation," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7759–7779, Dec. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8418783/

[31] Y. Shen, S. Mazuelas, and M. Z. Win, "Network navigation: Theory and interpretation," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1823–1834, Oct. 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6311242/

[32] A. T. Ihler, J. W. Fisher, R. L. Moses, and A. S. Willsky, "Nonparametric belief propagation for self-localization of sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 809–819, Apr. 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1413473/

[33] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for Localization-of-Things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8827486/

[34] F. Meyer et al., "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8290605/

[35] S. Bartoletti, A. Giorgetti, M. Z. Win, and A. Conti, "Blind selection of representative observations for sensor radar networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1388–1400, Apr. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7024174/

[36] U. A. Khan, S. Kar, and J. M. F. Moura, "DILAND: An algorithm for distributed sensor localization with noisy distance measurements," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1940–1947, Mar. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5352246/

[37] S. Bartoletti, W. Dai, A. Conti, and M. Z. Win, "A mathematical model for wideband ranging," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 2, pp. 216–228, Mar. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/6955781/

[38] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, Mar. 2009. [Online]. Available: http://ieeexplore.ieee.org/document/4802193/

[39] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. V. Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8421291/

[40] U. A. Khan, S. Kar, and J. M. F. Moura, "Distributed sensor localization in random environments using minimal number of anchor nodes," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 2000–2016, May 2009. [Online]. Available: http://ieeexplore.ieee.org/document/4776469/

[41] T. Wang, Y. Shen, A. Conti, and M. Z. Win, "Network navigation with scheduling: Error evolution," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7509–7534, Nov. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7953679/

[42] J. Chen, W. Dai, Y. Shen, V. K. N. Lau, and M. Z. Win, "Resource management games for distributed network localization," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 2, pp. 317–329, Feb. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7835141/

[43] A. Conti, M. Guerra, D. Dardari, N. Decarli, and M. Z. Win, "Network experimentation for cooperative localization," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 467–475, Feb. 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6136832/

[44] A. Conti, D. Dardari, M. Guerra, L. Mucchi, and M. Z. Win, "Experimental characterization of diversity navigation," *IEEE Syst. J.*, vol. 8, no. 1, pp. 115–124, Mar. 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6581875/

[45] Z. Liu, W. Dai, and M. Z. Win, "Mercury: An infrastructure-free system for network localization and navigation," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, pp. 1119–1133, May 2018. [Online]. Available: https://ieeexplore.ieee.org/document/7983453/

[46] B. Teague, Z. Liu, F. Meyer, A. Conti, and M. Z. Win, "Network localization and navigation with scalable inference and efficient operation," *IEEE Trans. Mobile Comput.*, vol. 21, no. 6, pp. 2072–2087, Jun. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9247273/

[47] L. Barbieri, M. Brambilla, A. Trabattoni, S. Mervic, and M. Nicoli, "UWB localization in a smart factory: Augmentation methods and experimental assessment," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–18, Apr. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9409138/

[48] S. Marano, W. Gifford, H. Wymeersch, and M. Win, "NLOS identification and mitigation for localization based on UWB experimental data," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1026–1035, Sep. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5555901/

[49] H. Wymeersch, S. Marano, W. M. Gifford, and M. Z. Win, "A machine learning approach to ranging error mitigation for UWB localization," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1719–1728, Jun. 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6192275/

[50] A. Bourdoux et al., "6G white paper on localization and sensing," Jun. 2020, *arXiv:2006.01779*.

[51] T. Wild, V. Braun, and H. Viswanathan, "Joint design of communication and sensing for beyond 5G and 6G systems," *IEEE Access*, vol. 9, pp. 30845–30857, 2021. [Online]. Available: http://ieeexplore.ieee.org/document/9354629/

[52] *Moderator's Summary for Discussion [RAN93e-R18Prep-10] Expanded and Improved Positioning*, document TSG-RAN Meeting (RP) 0.0.4, 3GPP, Version 0.0.4, Sep. 2021. [Online]. Available: https://www.3gpp.org/ftp/tsgran/TSGRAN/TSGR93e/Docs/

[53] A. Nessa, B. Adhikari, F. Hussain, and X. N. Fernando, "A survey of machine learning for indoor positioning," *IEEE Access*, vol. 8, pp. 214945–214965, 2020. [Online]. Available: http://ieeexplore.ieee.org/document/9264122/

[54] L. Cazzella et al., "Deep learning of transferable MIMO channel modes for 6G V2X communications," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4127–4139, Jun. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9765739/

[55] K. Wang et al., "Task offloading with multi-tier computing resources in next generation wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 2, pp. 306–319, Feb. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9978919/

[56] M. Brambilla, D. Pardo, and M. Nicoli, "Location-assisted subspace-based beam alignment in LOS/NLOS mm-wave V2X communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9148587/

[57] M. Brambilla, L. Combi, A. Matera, D. Tagliaferri, M. Nicoli, and U. Spagnolini, "Sensor-aided V2X beam tracking for connected automated driving: Distributed architecture and processing algorithms," *Sensors*, vol. 20, no. 12, p. 3573, Jun. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/12/3573

[58] Z. Miao, L. Zhao, W. Yuan, and F. Jin, "Application of one-class classification in NLOS identification of UWB positioning," in *Proc. Int. Conf. Inf. Syst. Artif. Intell. (ISAI)*, Jun. 2016, pp. 318–322. [Online]. Available: https://ieeexplore.ieee.org/document/7816727/

[59] J. Borras, P. Hatrack, and N. Mandayam, "Decision theoretic framework for NLOS identification," in *Proc. 48th IEEE Veh. Technol. Conf. Pathway Global Wireless Revolution (VTC)*, vol. 2, May 1998, pp. 1583–1587. [Online]. Available: http://ieeexplore.ieee.org/document/686556/

[60] Y.-H. Jo, J.-Y. Lee, D.-H. Ha, and S.-H. Kang, "Accuracy enhancement for UWB indoor positioning using ray tracing," in *Proc. IEEE/ION PLANS*, Apr. 2006, pp. 565–568. [Online]. Available: http://ieeexplore.ieee.org/document/1650645/

[61] S. Seidel and T. Rappaport, "A ray tracing technique to predict path loss and delay spread inside buildings," in *Proc. Commun. Global Users (GLOBECOM)*, vol. 2, Dec. 1992, pp. 649–653. [Online]. Available: http://ieeexplore.ieee.org/document/276436/

[62] I. Guvenc, C.-C. Chong, and F. Watanabe, "NLOS identification and mitigation for UWB localization systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2007, pp. 1571–1576. [Online]. Available: http://ieeexplore.ieee.org/document/4224541/

[63] İ. Güvenç, C.-C. Chong, F. Watanabe, and H. Inamura, "NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, Aug. 2007, Art. no. 271984.

[64] C. Huang et al., "Artificial intelligence enabled radio propagation for communications—Part II: Scenario identification and channel modeling," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3955–3969, Jun. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9713743/

[65] T. Van Nguyen, Y. Jeong, H. Shin, and M. Z. Win, "Machine learning for wideband localization," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1357–1380, Jul. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7102989/

[66] E. Kurniawan, L. Zhiwei, and S. Sun, "Machine learning-based channel classification and its application to IEEE 802.11ad communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/8254052/

[67] X. Yang, "NLOS mitigation for UWB localization based on sparse pseudo-input Gaussian process," *IEEE Sensors J.*, vol. 18, no. 10, pp. 4311–4316, May 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8322223/

[68] Q. Zheng et al., "Channel non-line-of-sight identification based on convolutional neural networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1500–1504, Sep. 2020.

[69] C. Jiang, J. Shen, S. Chen, Y. Chen, D. Liu, and Y. Bo, "UWB NLOS/LOS classification using deep learning method," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2226–2230, Oct. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9108193/

[70] Z. Cui, Y. Gao, J. Hu, S. Tian, and J. Cheng, "LOS/NLOS identification for indoor UWB positioning based on Morlet wavelet transform and convolutional neural networks," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 879–882, Mar. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9264213/

[71] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8027020/

[72] Y. Jing, J. Hao, and P. Li, "Learning spatiotemporal features of CSI for indoor localization with dual-stream 3D convolutional neural networks," *IEEE Access*, vol. 7, pp. 147571–147585, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8864050/

[73] J. Fan and A. S. Awan, "Non-line-of-sight identification based on unsupervised machine learning in ultra wideband systems," *IEEE Access*, vol. 7, pp. 32464–32471, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8666972/

[74] J. E. van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Mach. Learn.*, vol. 109, no. 2, pp. 373–440, Feb. 2020.

[75] Z. Zeng, R. Bai, L. Wang, and S. Liu, "NLOS identification and mitigation based on CIR with particle filter," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8886002/

[76] M. Stahlke, S. Kram, F. Ott, T. Feigl, and C. Mutschler, "Estimating TOA reliability with variational autoencoders," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5133–5140, Mar. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9503384/

[77] P. Oza and V. M. Patel, "One-class convolutional neural network," *IEEE Signal Process. Lett.*, vol. 26, no. 2, pp. 277–281, Feb. 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8586962/

[78] L. Ruff et al., "A unifying review of deep and shallow anomaly detection," *Proc. IEEE*, vol. 109, no. 5, pp. 756–795, May 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9347460/

[79] M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," Jan. 2017, *arXiv:1701.04862*.

[80] H. Xu et al., "Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications," in *Proc. World Wide Web Conf. (WWW)*, 2018, pp. 187–196, doi: 10.1145/3178876.3185996.

[81] E. Nalisnick, A. Matsukawa, Y. W. Teh, D. Gorur, and B. Lakshminarayanan, "Do deep generative models know what they don't know?" Oct. 2018, *arXiv:1810.09136*.

[82] C. Zhou and R. C. Paffenroth, "Anomaly detection with robust deep autoencoders," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*. New York, NY, USA: Association for Computing Machinery, Aug. 2017, pp. 665–674, doi: 10.1145/3097983.3098052.

[83] V. L. Cao, M. Nicolau, and J. McDermott, "A hybrid autoencoder and density estimation model for anomaly detection," in *Parallel Problem Solving From Nature—PPSN XIV*, J. Handl, E. Hart, P. R. Lewis, M. López-Ibáñez, G. Ochoa, and B. Paechter, Eds. Cham, Switzerland: Springer, Aug. 2016, pp. 717–726.

[84] B. Zong et al., "Deep autoencoding Gaussian mixture model for unsupervised anomaly detection," in *Proc. Int. Conf. Learn. Represent.*, Feb. 2018, pp. 1–19. [Online]. Available: https://openreview.net/forum?id=BJJLHbb0-

[85] *Study on NR Positioning Support*, document TR 38.855, Version 16.0.0, 3GPP, Sep. 2019. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3501

[86] *MATLAB, R2022a*. Natick, MA, USA: MathWorks, 2022. [Online]. Available: https://it.mathworks.com/help/antenna/ref/rfprop.raytracing.html

[87] P. A. Lopez et al., "Microscopic traffic simulation using SUMO," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2575–2582. [Online]. Available: https://elib.dlr.de/124092/

[88] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[89] H. L. Van Trees and E. Detection, *Modulation Theory, Optimum Array Processing (Part IV)*. New York, NY, USA: Wiley, 2002.

[90] C. Sun, X. Gao, S. Jin, M. Matthaiou, Z. Ding, and C. Xiao, "Beam division multiple access transmission for massive MIMO communications," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2170–2184, Jun. 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7093168/

[91] X. Sun, X. Gao, G. Y. Li, and W. Han, "Single-site localization based on a new type of fingerprint for massive MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6134–6145, Jul. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8307353/

[92] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep., Jan. 2007, pp. 10–15, vol. 5.

[93] C. Wu et al., "Learning to localize: A 3D CNN approach to user positioning in massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4556–4570, Jul. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9149427/

[94] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Computer Vision—ECCV*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham, Switzerland: Springer, 2014, pp. 818–833.

[95] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962. [Online]. Available: http://www.jstor.org/stable/2237880

[96] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, "An introduction to variational methods for graphical models," *Mach. Learn.*, vol. 37, no. 2, pp. 183–233, Nov. 1999.

[97] *Study on Channel Model for Frequencies From 0.5 to 100 GHz (Rel-16)*, document TR 38.901, Version 16.1.0, 3GPP, Nov. 2020. [Online]. Available: https://www.etsi.org/deliver/etsitr/138900138999/138901/16.01.0060/tr138901v160100p.pdf

[98] C.-F. Yang and C.-J. Ko, "A ray tracing method for modeling indoor wave propagation and penetration," in *IEEE Antennas Propag. Soc. Int. Symp. Dig.*, vol. 1, Jul. 1996, pp. 441–444. [Online]. Available: https://ieeexplore.ieee.org/document/686780/

[99] H.-J. Li, C.-C. Chen, T.-Y. Liu, and H.-C. Lin, "Applicability of ray-tracing technique for the prediction of outdoor channel characteristics," *IEEE Trans. Veh. Technol.*, vol. 49, no. 6, pp. 2336–2349, Nov. 2000. [Online]. Available: https://ieeexplore.ieee.org/document/901902/

[100] A. Hsiao, C. Yang, T. Wang, I. Lin, and W. Liao, "Ray tracing simulations for millimeter wave propagation in 5G wireless communications," in *Proc. IEEE Int. Symp. Antennas Propag., USNC/URSI Nat. Radio Sci. Meeting*, Oct. 2017, pp. 1901–1902. [Online]. Available: https://ieeexplore.ieee.org/document/8072993/

[101] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7152831/

[102] *Guidelines for Evaluation of Radio Interface Technologies for IMT-2020*, document, Report International Telecommunication Union (ITU), 2017. [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2412-2017-PDF-E.pdf

[103] X. He, Q. Guo, J. Tong, J. Xi, and Y. Yu, "Low-complexity approximate iterative LMMSE detection for large-scale MIMO systems," *Digit. Signal Process.*, vol. 60, pp. 134–139, Jan. 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1051200416301452

[104] L. Gopal, Y. Rong, and Z. Zang, "Tomlinson–Harashima precoding based transceiver design for MIMO relay systems with channel covariance information," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5513–5525, Oct. 2015. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7115189/

[105] A. Hindy and A. Nosratinia, "Ergodic fading MIMO dirty paper and broadcast channels: Capacity bounds and lattice strategies," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5525–5536, Aug. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7942023/

[106] *5G; NR; Physical Channels and Modulation*, document TS 38.214, Version 16.12.0 Release 16, 3GPP, Jan. 2023. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3216

[107] *5G; NR; Physical Channels and Modulation*, document TS 38.211, Version 16.10.0 Release 16, 3GPP, Jun. 2022. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213

[108] A. Paszke et al., "Automatic differentiation in PyTorch," Tech. Rep., 2017. [Online]. Available: https://openreview.net/forum?id=BJJsrmfCZ

[109] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Jan. 2014, *arXiv:1412.6980*.

[110] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7803544/

[111] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[112] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, in Proceedings of Machine Learning Research, vol. 37, F. Bach and D. Blei, Eds. Lille, France: PMLR, Jul. 2015, pp. 448–456. https://proceedings.mlr.press/v37/ioffe15.html

[113] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "GANomaly: Semi-supervised anomaly detection via adversarial training," in *Computer Vision—ACCV*, C. V. Jawahar, H. Li, G. Mori, and K. Schindler, Eds. Cham, Switzerland: Springer, May 2019, pp. 622–637.

[114] C. Huang et al., "Machine learning-enabled LOS/NLOS identification for MIMO systems in dynamic environments," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3643–3657, Jun. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8968748/

[115] S. Sun, J. Hu, Y. Peng, X. Pan, L. Zhao, and J. Fang, "Support for vehicle-to-everything services based on LTE," *IEEE Wireless Commun.*, vol. 23, no. 3, pp. 4–8, Jun. 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7498068/

[116] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," Dec. 2013, *arXiv:1312.6114*.
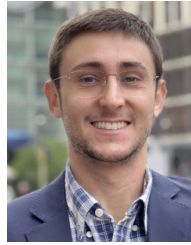
Bernardo Camajori Tedeschini (Graduate Student Member, IEEE) received the B.Sc. degree (Hons.) in computer science and M.Sc. degree (Hons.) in telecommunications engineering from the Politecnico di Milano, Italy, in 2019 and 2021, respectively. From November 2021, he started as Ph.D. fellow in Information Technology at Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano. He is currently a Visiting Researcher with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology (MIT), Cambridge, MA.

His research interests include federated learning, machine learning and localization methods. He was a recipient of the Ph.D. grant from the ministry of the Italian government Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) and the Roberto Rocca Doctoral Fellowship granted by MIT and Politecnico di Milano.

Monica Nicoli (Senior Member, IEEE) received the M.Sc. (Hons.) and Ph.D. degrees in communication engineering from Politecnico di Milano, Milan, Italy, in 1998 and 2002, respectively. She was a Visiting Researcher with ENI Agip, from 1998 to 1999, and Uppsala University, in 2001. In 2002, she joined Politecnico di Milano as a Faculty Member. She is currently an Associate Professor in telecommunications with the Department of Management, Economics and Industrial Engineering.

Her research interests include signal processing, machine learning, and wireless communications, with emphasis on smart mobility and Internet of Things (IoT) applications. She was a recipient of the Marisa Bellisario Award, in 1999, and a co-recipient of the best paper awards of the IEEE Symposium on Joint Communications and Sensing, in 2021, the IEEE Statistical Signal Processing Workshop, in 2018, and the IET *Intelligent Transport Systems journal*, in 2014. She is an Associate Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. She has also served as an Associate Editor for the EURASIP *Journal on Wireless Communications and Networking*, from 2010 to 2017, and a Lead Guest Editor for the Special Issue on Localization in Mobile Wireless and Sensor Networks, in 2011.

Moe Z. Win (Fellow, IEEE) is a Professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and with NASA Jet Propulsion Laboratory.

His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include ultra-wideband systems, network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial positions for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, coauthored with students and colleagues, have received several awards. Other recognitions include the MIT Everett Moore Baker Award (2022), the IEEE Vehicular Technology Society James Evans Avant Garde Award (2022), the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* (2008) from the Università degli Studi di Ferrara, and the U.S. Presidential Early Career Award for Scientists and Engineers (2004). He is an ISI Highly Cited Researcher.