

Received 19 February 2024; revised 27 May 2024; accepted 27 May 2024. Date of publication 30 May 2024; date of current version 6 August 2024.

Digital Object Identifier 10.1109/OJAP.2024.3407199

Electromagnetic-Informed Generative Models for Passive RF Sensing and Perception of Body Motions

STEFANO SAVAZZI¹ (Member, IEEE), FEDERICA FIERAMOSCA², SANAZ KIANOUSH¹,
MICHELE D'AMICO² (Senior Member, IEEE), AND VITTORIO RAMPA¹ (Senior Member, IEEE)

¹Consiglio Nazionale delle Ricerche, IEIT Institute, 20133 Milan, Italy

²Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milan, Italy

CORRESPONDING AUTHOR: S. SAVAZZI (e-mail: stefano.savazzi@cnr.it)

Funded by the European Union under Agreement 101099491.

The work has been also partially presented at the IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC), Venice, Italy, October 9–13, 2023 [DOI: 10.1109/APWC57320.2023.10297498].

ABSTRACT Electromagnetic (EM) body models predict the impact of human presence and motions on the Radio-Frequency (RF) field originated from wireless devices nearby. Despite their accuracy, EM models are time-consuming methods which prevent their adoption in strict real-time computational imaging and estimation problems, such as passive localization, RF tomography, and holography. Physics-informed Generative Neural Network (GNN) models have recently attracted a lot of attention thanks to their potential to reproduce a process by incorporating relevant physical laws and constraints. They can be used to simulate or reconstruct missing data or samples, reproduce EM propagation effects, approximated EM fields, and learn a physics-informed data distribution, i.e., the Bayesian prior. Generative machine learning represents a multidisciplinary research area weaving together physical/EM modelling, signal processing, and Artificial Intelligence (AI). The paper discusses two popular techniques, namely Variational Auto-Encoders (VAEs) and Generative Adversarial Networks (GANs), and their adaptations to incorporate relevant EM body diffraction methods. The proposed EM-informed GNN models are verified against classical EM tools driven by diffraction theory, and validated on real data. The paper explores emerging opportunities of GNN tools targeting real-time passive RF sensing in communication systems with dense antenna arrays. Proposed tools are also designed, implemented, and verified on resource constrained wireless devices. Simulated and experimental analysis reveal that GNNs can limit the use of time-consuming and privacy-sensitive training stages as well as intensive EM computations. On the other hand, they require hyper-parameter tuning to achieve a good compromise between accuracy and generalization.

INDEX TERMS EM body models, generative models, variational autoencoders, generative adversarial networks, radio tomography, integrated sensing and communication, localization.

I. INTRODUCTION

PASSIVE radio sensing employs stray ambient radio signals from Radio Frequency (RF) devices to detect, locate, and track people that do not carry any electronic device, namely device-free [1], [2], [3], [4], [5]. In line with the *Communication while Sensing* paradigm [5], these tools provide seamless detection capabilities, while performing wireless communications. Radio signals encode a view of all moving/fixed objects traversed during their propagation:

several data analytic tools, such as Bayesian [6], [7] and machine learning approaches [8], can be used to decode this information, typically by large-scale processing of radio signals exchanged by different devices.

Most of emerging approaches proposed for solving the radio sensing problem require an approximated knowledge of a physical-informed (prior) model to interpret the effects of human subjects on radio propagation. The perturbative effects of the radio signals induced by the presence or

movements of human bodies can be interpreted using electromagnetic (EM) propagation theory considerations [9]. These EM methods have paved the way to several physical and statistical models for passive radio sensing, which exploit full wave approaches [10], ray tracing [11], moving point scattering [12], and diffraction theory [13], [14], [15], [16]. The body-induced perturbations that impair the radio channel, can be thus acquired, measured, and processed using model-based methods to estimate location and track target information. A general EM model for the prediction of body-induced effects on propagation is still under scrutiny [17]. On the other hand, current models are too complex to be of practical use for real-time sensing scenarios [10], [11], although they can be used for off-line applications such as network pre-deployment assessment [18].

Physics-informed generative machine learning is an emerging field in different application contexts ranging from imaging [19], EM field computation [20], to Bayesian estimation for inverse problems [21]. Generative deep Neural Networks (GNN) can be trained to produce observations drawn from a distribution which reflects the complex underlying physics of the environment under study, or rather reproduce approximate fields in an almost negligible time compared with classical numerical methods [22], [23]. For the first time, our paper discusses the adoption of GNN models designed to reproduce the effects of body movements on EM propagation, considering varying size, position and orientation/posture of the body, multiple antenna (e.g., MIMO) setups and different physical and geometrical properties of the radio link(s).

The proposed physics-informed GNN models are trained with samples obtained from EM models based on diffraction theory [15], [16], under different environment configurations. The GNN tools discussed in this paper are based on Variational Auto-Encoders (VAEs) [24] and Generative Adversarial Networks (GANs) [25], [26]. The opportunities and the limitations of each proposed approach are discussed and compared in several case studies targeting the perception of body motions and passive RF sensing applications.

A. RELATED WORKS

Physics-informed GNN models use Machine Learning (ML) methods for computing physical processes. Although still in their infancy, they have been recently proposed to approximate EM fields. A small body of existing works related to this problem does exist. A ML model is proposed in [27] to obtain an approximation of the EM field in a cavity with an arbitrary spatial dielectric permittivity distribution. The model is shown to be one order of magnitude faster than similar finite-difference frequency-domain simulations, suggesting possible applications in inverse problems. In [28], a neural solver for Poisson's equation is proposed using a purely-convolutional neural network structure. An approach for solving Partial Differential Equations (PDEs) using Neural Networks (NNs) has recently emerged [29], where a physics-based loss function is constructed to improve

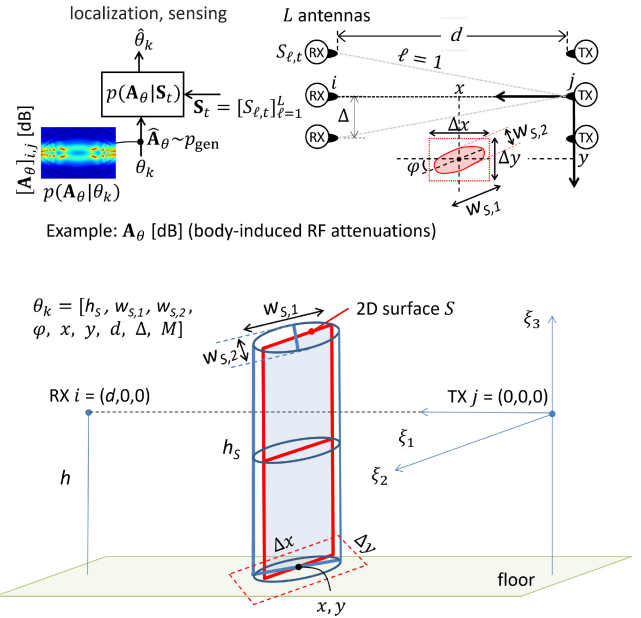


FIGURE 1. EM model geometry (2D sheet-like obstacle, multiple-input multiple-output antenna configuration) and generative approach.

NN training. Compared to traditional EM field computing methods based on numerical integration and/or mesh-based methods, an attractive feature of physics-informed models, based on Deep NN (DNN) implementations, is that they could break the curse of dimensionality [30]. In addition, once trained, DNNs can solve an EM problem in an almost negligible time in comparison to classical numerical methods [22], [23]. Finally, generation accuracy and training time can be improved by incorporating a small amount of *labeled* data or EM field measurements (if available) during the training process.

Applications of GNN tools to communications and localization are also emerging [31]. For example, [32] discussed a convolutional encoder-decoder structure that can be trained to reproduce the results of a ray-tracer, encoding also physics-based information of an indoor environment. A ML-assisted channel modeling approach is proposed in [33] to generate site-specific mmWave channel characteristics. The model is shown to improve the generalization capabilities of conventional physical-statistical models when adopted to reproduce complex network configurations. A Multibranch GAN (MBGAN) has been recently analyzed for radar signal processing to synthesize data that reflect human physical properties and kinematics [34]. The model is shown to provide an increase of 9% in classification accuracy.

B. OBJECTIVES AND CONTRIBUTIONS

The paper discusses for the first time the adoption of EM-informed generative neural network models inspired by VAE and GAN tools [24], [25]. As depicted in Fig. 1, the GNN models are designed and trained to reproduce the human blockage effects on radio propagation as underpinned by scalar diffraction theory. The models consider different body and link configurations relevant in radio sensing:

- 1) varying link geometries: distance (d), height from the ground (h);
- 2) Multiple-Input Multiple-Output (MIMO) antennas (L);
- 3) variable size ($w_{S,1}, w_{S,2}, h_S$), position (x, y) and orientation/posture (φ) of the monitored target.

The proposed GNNs comprise of: *i*) a generator of body-induced RF signal attenuations, and *ii*) a generator of EM field samples which can be used for EM full-wave propagation analysis. The generators are shown to reproduce human body blockage effects under configurations which might be unseen during the training phase, or rather difficult to predict through traditional EM field computing methods. The considered GNN models are designed to reproduce the body blockage effects according to *selectable* configurations: therefore, they generate EM field samples conditioned on input body characteristics, or features, that can be selected at *run-time*, namely during RF sensing deployment.

In line with initial studies [19], [22], [23], the generators are purely based on sequences of convolutional (and deconvolutional) layers whose number varies depending on the physical quantity to be reproduced. Therefore, they are well-suited for real-time localization applications as do not need intensive EM wavefield computations. VAE and GAN [22], [26] generator methods are analyzed in terms of their accuracy in reproducing EM model diffraction effects, implementation complexity, generation times, and model size.

The paper is organized as follows: Section II introduces the passive RF sensing problem and motivations. Section III reviews relevant EM body models that quantify the human body blockage based on diffraction theory considerations. Section IV targets the proposed GNN approaches and discusses VAE and GAN tools. Section V validates the proposed generators against EM diffraction and thus verifies the effectiveness of the models in reproducing body blockage effects on single and multiple antennas wireless receivers. Section VI discusses an experimental case study. The goal is to demonstrate the effectiveness of GNN tools in reproducing real field measurements, supporting the RF sensing and the passive localization processes. Finally, concluding Section VII summarizes the open problems, the opportunities, and the limitations of the study.

II. BACKGROUND AND PROBLEM FORMULATION

The RF sensing goal is to extract the EM human body(ies) blockage effects (\mathbf{E}_θ) from noisy measurements \mathbf{S}_t of the RF radiation observed at time t . The human subject(s) is characterized by an unknown state θ which is recovered from \mathbf{E}_θ . The body effects \mathbf{E}_θ can be evaluated in terms of body-induced *excess* attenuations \mathbf{A}_θ [15], [35], as in the example of Fig. 1. Baseband Channel State Information (CSI) \mathbf{C}_θ [36] can be evaluated as well. The body state θ consists of an ensemble of *features*, e.g., body location, size, height, and orientation (see Section III-A) [6], [15], [16], which depend on the specific sensing application. In what follows, we provide the necessary background on RF sensing and Bayesian methods.

A. RF-SENSING AND BAYESIAN FORMULATION

The objective of the RF sensing inverse problem is to obtain the posterior distribution $p(\mathbf{E}_\theta|\mathbf{S}_t)$:

$$p(\mathbf{E}_\theta|\mathbf{S}_t) = \frac{p(\mathbf{S}_t|\mathbf{E}_\theta) \cdot p(\mathbf{E}_\theta)}{p(\mathbf{S}_t)} \quad (1)$$

of the (unknown) human body blockage effects \mathbf{E}_θ , given the measurements \mathbf{S}_t . Maximum A-Posteriori (MAP) solution to (1) allows to extract the most likely effects:

$$(\widehat{\mathbf{E}}_\theta, \widehat{\theta}) = \arg \max_{\theta} \{p(\mathbf{E}_\theta|\mathbf{S}_t)\} \quad (2)$$

from which it is possible to recover the subject state $\widehat{\theta}$ and any *feature* ($\vartheta \in \widehat{\theta}$) of interest, e.g., body position, size, height, and orientation. Field measurements \mathbf{S}_t can be in the form of received power, Received Signal Strength (RSS), or base-band CSI response [5]. Observations \mathbf{S}_t are perturbed by the body movements according to a *prior* distribution, $p(\mathbf{E}_\theta)$, which predicts the effects of the body (i.e., the target) in the state θ as the result of the propagation of the reflected, scattered, and diffracted EM waves.

The Bayesian approach (1) for solving the radio sensing problem (2) requires the knowledge of the likelihood function $p(\mathbf{S}_t|\mathbf{E}_\theta)$, namely the RF measurement model, and the prior distribution $p(\mathbf{E}_\theta)$. The likelihood term depends on the data collection process as well as on the impairments introduced by the measurement instrument or by the environment. It is typically chosen as log-normal distributed, according to [2], [6]. On the other hand, the prior distribution $p(\mathbf{E}_\theta)$, which models the initial beliefs on \mathbf{E}_θ , is usually hard to represent as it often requires full-wave EM approaches. Approximated solutions, such as diffraction models [13], [16], and several variants [12], are in many cases too time-consuming to be of practical use for real-time sensing scenarios [11]. In addition, when it comes to practice, imperfect knowledge of the scenario, small, involuntary, body movements, or changing configurations of the propagation environment, make the prior even more difficult to obtain with an acceptable level of accuracy [5].

B. BAYESIAN PRIOR MODELLING OF EM BODY EFFECTS

The EM-informed GNN tools discussed in this paper are designed to reproduce the EM effects \mathbf{E}_θ as sampled from the Bayesian *prior* probability distribution $p(\mathbf{E}_\theta)$. The prior $p(\mathbf{E}_\theta)$ quantifies the uncertainties of the body effects \mathbf{E}_θ as caused by imperfect knowledge of the body state θ . It is defined in general as:

$$p(\mathbf{E}_\theta|\theta_k) = \mathbf{E}_{\theta \sim p(\theta|\theta_k)}. \quad (3)$$

In other words, the EM effects \mathbf{E}_θ are obtained for random instances of body features θ that follow a probability function $p(\theta|\theta_k)$. Probability $p(\theta|\theta_k)$ models the uncertainty with respect to the *nominal* body features θ_k .

Some examples are proposed in the following to clarify the approach. First, consider the problem of generating

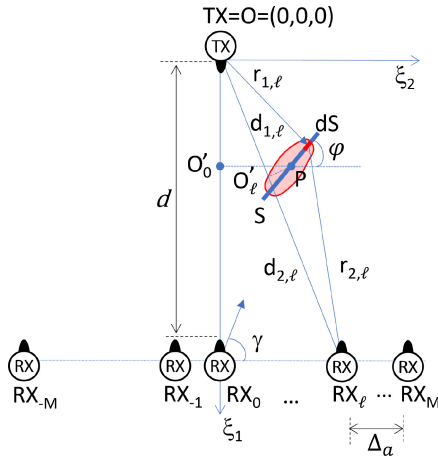


FIGURE 2. 2D layout of the radio link assuming a multiple antenna configuration of Fig. 1. Extension to a MIMO set-up is straightforward. The point O'_l is the projection of the barycenter P of the target S over the l -th LoS path having length d_l .

body-induced RF excess attenuation values (\mathbf{A}_θ) for a subject located at some (nominal) position $\theta_k = (x, y)$ [5]. Involuntary movements, as the result of the complex structure of the human body, strongly affect the RSS [5] and must be adequately taken into account. Body motions can be represented by random movements in an elementary squared area of size Δ around the nominal location (x, y) . Subject movements can be modelled by setting $p(\theta|\theta_k) = \mathcal{U}_{-(\Delta/2), (\Delta/2)}$ with $\Delta = 5 \div 10$ cm [16]. Replacing \mathbf{E}_θ with \mathbf{A}_θ , the prior distribution becomes $p(\mathbf{A}_\theta|\theta_k) = \mathbf{A}_\theta \sim \mathcal{U}_{-(\Delta/2), (\Delta/2)}$.

Likewise, let's consider now the problem of subject activity recognition [36], which requires to real-time track the subject trajectory and orientation w.r.t. the LOS path and its effects on CSI (\mathbf{C}_θ) measurements. Diffraction models [13], [16] can be designed to capture the rotation angle φ of the 2D target (see Fig. 2) by varying the size of the absorbing sheet S that represents the body. However, this operation is often expensive in terms of computational time. Rather than simulating each target rotation angle separately, which is not feasible, the proposed generative model can be set to reproduce the CSI \mathbf{C}_θ for all subject orientations¹ simultaneously, namely $p(\mathbf{C}_\theta|\theta_k) = \mathbf{C}_\theta \sim p(\theta | -\pi/2 \leq \varphi \leq \pi/2)$. Further examples are given in Section V.

III. EM BODY MODELS

The proposed GNN tools are optimized to match the prior distribution $p(\mathbf{E}_\theta)$ in (3) using (few) training examples obtained from scalar diffraction theory [15], [16]. In this section, we discuss relevant diffraction-based EM body models to reproduce the human body blockage effects \mathbf{E}_θ , considering also body-induced RF excess attenuations \mathbf{A}_θ and CSI \mathbf{C}_θ , as special cases. First, we briefly recall the body models proposed in [15] for a single link scenario using scalar diffraction theory considerations. Next, we consider a receiver equipped with an array, i.e., Uniform Linear Array

¹The angular resolution is set to $\pi/200$, which is sufficient for accurate reproduction of complex body trajectories.

(ULA), of L isotropic receiver antennas. The diffraction model represents the body effects relative to each radio link l , namely $\mathbf{E}_\theta = [E_{\ell,\theta}]_{\ell=1}^L$, or equivalently $\mathbf{A}_\theta = [A_{\ell,\theta}]_{\ell=1}^L$ for RSS and $\mathbf{C}_\theta = [C_{\ell,\theta}]_{\ell=1}^L$ for CSI.

In what follows, we will always assume that the monitored target is in the Fraunhofer's region of both transmitting (TX) and receiving (RX) antennas for all the considered links l . Extension to multi-target scenarios can be also inferred according to [16].

A. DIFFRACTION MODELS FOR BODY-INDUCED EXCESS ATTENUATIONS

As depicted in Fig. 1, we assume that the length of the radio link is given by d while h is its height from the floor. The effects of floor, walls, ceiling or other obstacles are not considered. However, with some effort, these obstacles can be included, as shown in [43]. The scalar diffraction theory assumes that the 3D shape of the human body is modeled as a 2D rectangular absorbing sheet S [15] with height h_S and traversal size that changes according to a 3D cylinder view, with max. and min. traversal sizes $w_{S,1}, w_{S,2}$, respectively. The target has nominal position coordinates $\mathbf{p} = [x, y]$, w.r.t. the TX position, which is defined by the projection of its barycenter on the horizontal plane that includes the Line-of-Sight (LOS). The 2D target might be also rotated of an angle φ with respect to the LOS direction. The body/subject state θ is characterized by an ensemble of body features collected into the vector $\theta := \{\mathbf{p}, \varphi, h_S, w_{S,1}, w_{S,2}\}$.

A distribution of Huygens' sources of elementary area dS is located on the absorbing sheet S . The electric field E_θ at the receiver [15] is obtained by subtracting the contribution of the obstructed Huygens' sources from the electric field E_0 of the free-space scenario (with no target in the link area):

$$E_\theta = E_0 - \int_S dE \quad (4)$$

where time t is omitted to simplify the reasoning. According to [15], equation (4) can be rewritten in terms of the field ratio $C_\theta = (E_\theta/E_0)$, representing the CSI:

$$C_\theta = 1 - j \frac{d}{\lambda} \int_S \frac{1}{r_1 r_2} \exp \left\{ -j \frac{2\pi}{\lambda} (r_1 + r_2 - d) \right\} d\xi_2 d\xi_3 \quad (5)$$

where λ is the wavelength. Notice that each elementary source $dS = d\xi_2 d\xi_3$ has distance r_1 and r_2 from the TX and the RX, respectively which depends on the relative coordinates \mathbf{p} .

B. MULTIPLE ANTENNA ARRAY CONFIGURATIONS

We now consider an ULA configuration with links ordered as $-M \leq \ell \leq M$ and being RX_ℓ the receiver node for corresponding link l . The central antenna of the array is indicated by the index $\ell = 0$. As shown in Fig. 2, each l -th antenna RX_ℓ of the array is uniformly deployed at mutual distance Δ_a along a segment orthogonal to the LOS at distance d from the TX and horizontally placed w.r.t.

the floor. Ignoring mutual antenna coupling (approximately valid for $\Delta_a > \lambda/4$, see [37]), the CSI observed on the ℓ -th antenna of the array corresponds to the ratio of the electric fields $C_{\ell,\theta} = (E_{\ell,\theta}/E_{\ell,0})$: therefore, using (5) it is:

$$C_{\ell,\theta} = 1 - j \frac{d_\ell}{\lambda} \int_S \frac{1}{r_{1,\ell} r_{2,\ell}} \cdot \exp \left\{ -j \frac{2\pi}{\lambda} (r_{1,\ell} + r_{2,\ell} - d_\ell) \right\} d\xi_2 d\xi_3 \quad (6)$$

where $E_{\ell,0}$ is the EM field received by the same RX_ℓ node in the reference condition, i.e., the free-space scenario. The term d_ℓ indicates the distance of the ℓ -th antenna RX_ℓ of the array from the TX while $d_{1,\ell}$ and $d_{2,\ell}$ are the distances of the projection point O'_ℓ (of the barycenter P of the 2D surface S) from the TX and RX_ℓ nodes. Likewise, $r_{1,\ell}$ and $r_{2,\ell}$ are the distances of the generic elementary area dS of the target S from the TX and RX_ℓ , respectively. Notice that, for $M = 0$, equation (6) reduces to the single-antenna case (5), where RX_0 coincides with the RX antenna at distance $d = d_0$ from the TX.

CSI data ($C_{\ell,\theta}$) and the corresponding excess attenuation values ($A_{\ell,\theta} = -10 \log_{10} |C_{\ell,\theta}|^2$) represent the human body blockage effects for link ℓ . These are organized into the vectors:

$$\begin{aligned} \mathbf{A}_\theta &= \left[A_{\ell,\theta} = -10 \log_{10} |C_{\ell,\theta}|^2 \right]_{\ell=1}^L, \\ \mathbf{C}_\theta &= [C_{\ell,\theta}]_{\ell=1}^L. \end{aligned} \quad (7)$$

In addition to the CSI terms \mathbf{C}_θ , the EM field $E_{\ell,\theta}$, observed on link ℓ can be also re-arranged according to (6) as:

where $E_{0,0}$ is the electric field received by the central antenna of the array of index $\ell = 0$. Using (8), as shown at the bottom of the page, and (6), the EM field for each considered link can be obtained as:

$$\mathbf{E}_\theta = \left[E_{0,0} C_{\ell,\theta} \frac{d}{d_\ell} \exp \left\{ -j \frac{2\pi}{\lambda} (d_\ell - d) \right\} \right]_{\ell=1}^L. \quad (9)$$

C. EFFECTS OF HUMAN BODY BLOCKAGE ON ARRAY RESPONSE

Based on the previous analysis, we now highlight the EM effects of target movements on the *array response* [38] of conventional linear beamforming processing [44]. Using the same ULA configuration, we consider the vector $\mathbf{w}(\gamma) = [w_\ell]_{\ell=-M}^{+M} = [w_{-M} \dots, w_{-1} w_0 w_1 \dots, w_M]^T$ of linear beamforming coefficients designed to *steer* the array in a direction γ . The received baseband signal $r_\theta(\gamma)$ at the output of the beamforming processing is given by:

$$r_\theta(\gamma) = \sum_{\ell=-M}^{+M} w_\ell^*(\gamma) [E_{\ell,\theta} + n_\ell] = \mathbf{w}(\gamma)^H \cdot [\mathbf{E}_\theta + \mathbf{n}], \quad (10)$$

where H indicates conjugate transpose operation, n_ℓ is the ℓ -th element of the Additive White Gaussian Noise (AWGN) complex vector $\mathbf{n} = [n_{-M} \dots, n_{-1} n_0 n_1 \dots, n_M]^T$ of size $2M + 1$, that is assumed to be spatially white with zero mean and covariance $\sigma^2 \mathbf{I}$. Neglecting the AWGN noise² and considering the CSI $C_{\ell,\theta}$ defined in (6), the array response $R_\theta(\gamma)$ as due to a target in state θ is defined as [38]

$$R_\theta(\gamma | \mathbf{C}_\theta) = \sum_{\ell=-M}^{+M} w_\ell^*(\gamma) \frac{E_{\ell,0}}{E_{\ell,\theta}} = \mathbf{w}(\gamma)^H \cdot \mathbf{C}_\theta^{-1}. \quad (11)$$

Notice that conventional ULA scenarios assume planar wavefront propagation. In this case, the steering vector $\mathbf{w}(\gamma)$ for the considered array is given by [44]:

$$\mathbf{w}(\gamma) = \left[\exp \left\{ j \frac{2\pi}{\lambda} \Delta_a \cos \gamma \right\} \right]_{\ell=-M}^{+M} \quad (12)$$

where $\Delta_a = \lambda/2$ is the inter-element antenna distance. According to (12), it is also $|\mathbf{w}|^2 = \mathbf{w}^H \mathbf{w} = (2M + 1)$.

The dominant Direction of Arrival (DoA) γ_{\max} , namely the maximum response of the array, is obtained as:

$$\gamma_{\max} = \arg \max_{\gamma} \{R_\theta(\gamma | \mathbf{C}_\theta)\} \quad (13)$$

and will be considered in the analysis of Section V.

IV. EM-INFORMED GENERATIVE NEURAL NETWORKS TOOLS

The generative models considered in this section reproduce body-induced EM effects \mathbf{E}_θ as sampled from the conditional prior distribution $p(\mathbf{E}_\theta | \theta_k)$ in (3). The prior is thus conditioned on the input body features θ_k . As shown in Fig. 3, the generation process is implemented by a *decoder* (VAE), or a *generator* (GAN), both parameterized by deep Neural Network (NN) parameters \mathbf{W}_D and \mathbf{W}_G , respectively. The neural networks map the input *latent space* $\mathbf{z} \sim p_Z(\mathbf{z})$ of size Z ($\mathbf{z} \in \mathbb{R}^{Z \times 1}$), into the output space:

$$\widehat{\mathbf{E}}_\theta \sim p_{\text{gen}}(\mathbf{E}_\theta | \theta_k) \cong p(\mathbf{E}_\theta | \theta_k). \quad (14)$$

The generated samples $\widehat{\mathbf{E}}_\theta$ are thus set to reproduce the targeted EM model, namely $p_{\text{gen}}(\mathbf{E}_\theta | \theta_k) \cong p(\mathbf{E}_\theta | \theta_k)$, for selectable inputs θ_k . As shown in Fig. 3, the NN parameters \mathbf{W}_D and \mathbf{W}_G constitute the generation models and are trained separately to reproduce body-induced excess attenuations \mathbf{A}_θ or EM field samples \mathbf{E}_θ , respectively.

²In line with the setup described in (1), the generative model is now designed to reproduce the prior effects of body movements on the response of the array; therefore, it appears reasonable to neglect the effect of measurement and AWGN noise, as well as fading.

$$\frac{E_{\ell,\theta}}{E_{0,0}} = \frac{d_0}{d_\ell} \exp \left\{ -j \frac{2\pi}{\lambda} (d_\ell - d_0) \right\} \left[1 - j \frac{d_\ell}{\lambda} \int_S \frac{1}{r_{1,\ell} r_{2,\ell}} \cdot \exp \left\{ -j \frac{2\pi}{\lambda} (r_{1,\ell} + r_{2,\ell} - d_\ell) \right\} d\xi_2 d\xi_3 \right] \quad (8)$$

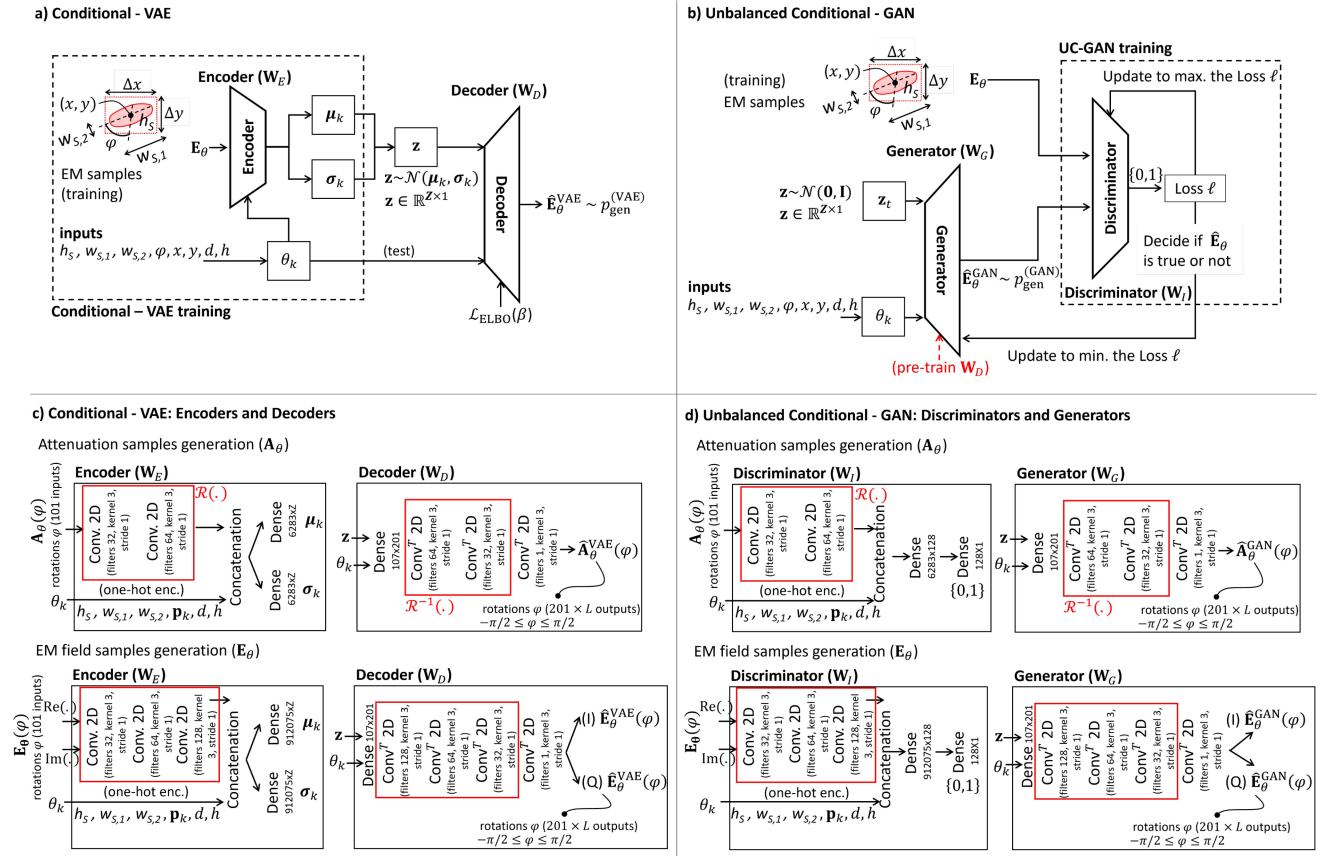


FIGURE 3. a) Conditional VAE (C-VAE) and b) Unbalanced Conditional GAN (UC-GAN) architectures for generating EM body model samples; c) C-VAE encoder and decoder neural network structures for generating excess attenuations A_θ and EM field responses E_θ ; d) corresponding discriminator and generator structures for UC-GAN. Dense, Conv and Conv^T refers to fully connected, convolution and deconvolution layer [48] operators, respectively.

Below we discuss VAE and GAN model architectures [25], [26] referred to as conditional-VAE (C-VAE) and unbalanced conditional-GAN (UC-GAN) [46]. Both models are adapted to generate samples conditioned on input body features θ_k that can be chosen at run-time. In the following, we limit our focus on the body feature set $\theta_k = [\mathbf{p}_k, \varphi_k, h_S, w_{S,1}, w_{S,2}]$ so to generate body effects for varying locations \mathbf{p}_k , orientations $-\pi/2 \leq \varphi_k \leq \pi/2$, and sizes $h_S, w_{S,1}, w_{S,2}$ of the target. Although different approaches are possible, the problem is complex enough to make a full EM simulation unfeasible, thus motivating the use of GNN models.

A. CONDITIONAL VARIATIONAL AUTOENCODER (C-VAE)

As depicted in Figs. 3a) and 3c), the C-VAE model uses an encoder $\mathcal{Q}(\mathbf{z}|\mathbf{E}_\theta, \theta_k; \mathbf{W}_E)$, parameterized by NN parameters \mathbf{W}_E , which learns the latent space $p_{\mathcal{Z}}(\mathbf{z}|\theta_k) \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k^2)$ for inputs θ_k . Latent space is multivariate Gaussian distributed with mean $\boldsymbol{\mu}_k$ and standard deviation $\boldsymbol{\sigma}_k$ parameters (other choices are not investigated here). The encoder is trained using samples \mathbf{E}_θ obtained from the EM model (9) and the corresponding body states θ_k . Model training is further discussed in Section V. The decoder produces a distribution

$$\hat{\mathbf{E}}_\theta^{\text{VAE}} \sim p_{\text{gen}}^{\text{VAE}}(\mathbf{E}_\theta|\theta_k):$$

$$p_{\text{gen}}^{\text{VAE}}(\mathbf{E}_\theta|\theta_k) = \int_{\mathcal{Z}} p_{\text{gen}}^{\text{VAE}}(\mathbf{E}_\theta|\mathbf{z}, \theta_k; \mathbf{W}_D) p_{\mathcal{Z}}(\mathbf{z}|\theta_k) d\mathbf{z} \quad (15)$$

which is the marginalization of the conditional probability $p_{\text{gen}}^{\text{VAE}}(\mathbf{E}_\theta|\mathbf{z}, \theta_k; \mathbf{W}_D)$ function of the NN parameters \mathbf{W}_D . The goal is to maximize the likelihood bound called Evidence Lower Bound (ELBO) $\mathcal{L}_{\text{ELBO}}$ described in [45]. Omitting dependency on parameters \mathbf{W}_E and \mathbf{W}_D , it is:

$$\mathcal{L}_{\text{ELBO}} = \ell_k - \beta \cdot \text{D}_{\text{KL}}[\mathcal{Q}(\mathbf{z}|\mathbf{E}_\theta, \theta_k) \| p_{\mathcal{Z}}(\mathbf{z}|\theta_k)]. \quad (16)$$

The first term $\ell_k = \mathbb{E}_{\mathbf{z} \sim \mathcal{Q}(\cdot|\cdot)} \log[p_{\text{gen}}^{\text{VAE}}(\mathbf{E}_\theta|\mathbf{z}, \theta_k)]$ is the log-likelihood function, while the second one is the Kullback-Leibler (KL) divergence D_{KL} [47] between the encoder output and the input latent space.

Maximization of the likelihood ℓ_k in (16) makes the generated samples $\hat{\mathbf{E}}_\theta^{\text{VAE}}$ more correlated to the latent variables \mathbf{z} , which typically cause the model to be more deterministic. On the other hand, the number of latent variables Z as well as the ELBO weight term $\beta > 0$ can be tuned to increase the contribution of the KL divergence between the posterior and the prior to the total ELBO and thus increase the randomness of generated samples. Targeting passive localization applications, in Section V, we will show

that both terms (Z, β) can be optimized to improve the generation process, also to account for measurements \mathbf{S}_l affected by noise and multipath interference.

B. UNBALANCED CONDITIONAL GAN (UC-GAN)

GAN training, depicted in Figs. 3b) and 3d), is formulated as a min-max problem that can be interpreted as an adversarial game with two players: the *discriminator* $\mathcal{D}(\hat{\mathbf{E}}_\theta, \mathbf{E}_\theta; \mathbf{W}_D) \in [0, 1]$, namely a binary classifier which tries to improve the detection of *fake* EM field samples, and the generator which is designed to fool the discriminator. The generator produces $\hat{\mathbf{E}}_\theta^{\text{GAN}} \sim p_{\text{gen}}^{\text{GAN}}(\mathbf{E}_\theta|\theta_k; \mathbf{W}_G)$ with:

$$p_{\text{gen}}^{\text{GAN}}(\mathbf{E}_\theta|\theta_k) = \int_{\mathbf{Z}} p_{\mathcal{Z}}(\mathbf{z}) p_{\text{gen}}^{\text{GAN}}(\mathbf{E}_\theta|\mathbf{z}, \theta_k; \mathbf{W}_G) d\mathbf{z} \quad (17)$$

now driven by $p_{\mathcal{Z}}(\mathbf{z}) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The goal is now to minimize the statistical distance (Jensen-Shannon divergence, based on [47]) between $p(\mathbf{E}_\theta|\theta_k)$ and $p_{\text{gen}}^{\text{GAN}}$. This corresponds to maximizing the discriminator loss while minimizing the generator one. More details can be found in [26]. Physics-informed GAN based models were proposed in several works [23], [32]. For the considered problem, we adopted an unbalanced implementation [46] which pre-trains the generator using the parameters \mathbf{W}_D of the C-VAE decoder (15). This prevents the faster convergence of the discriminator at early epochs which could limit the generator reproduction accuracy.

C. MODEL TRAINING AND IMPLEMENTATION CONSIDERATIONS

C-VAE and UC-GAN pre-trained models shown in Fig. 3 are available on-line [41] together with example codes for training on new samples, and testing, namely generating body-induced excess attenuations ($\hat{\mathbf{A}}_\theta$) and CSI ($\hat{\mathbf{C}}_\theta$), according to specific body configurations. In the following sections, we discuss few critical implementation constraints and related general considerations useful for the case study of Section VI.

GNN trainable parameters and models. Considering C-VAE in Figs. 3a) and 3c), the encoder model \mathbf{W}_E includes two key components: a sequence of convolutional layers and a feed-forward network. The encoder takes as input training samples obtained from the diffraction model as well as the conditional inputs θ_k . Excess attenuations \mathbf{A}_θ and EM field generation \mathbf{E}_θ require a different number of convolutional layer subcomponents, to reflect the dimension of the data: 2 layers are chosen for reproducing excess attenuations $\hat{\mathbf{A}}_\theta$, while 3 layers are required to generate EM field samples $\hat{\mathbf{E}}_\theta$. This choice is conservative since it is critical to limit the size of trainable model parameters, while no apparent performance improvement is observable beyond this limit. The decoder \mathbf{W}_D reproduces samples of body effects as a function of customized inputs θ_k that are one-hot encoded before being used as input to the neural network. It uses transposed convolution layers, also referred to as fractionally-strided convolutions, to increase (upsample) the

TABLE 1. C-VAE and UC-GAN vs diffraction model: body-induced excess attenuation and EM field generation time analysis.

Model	Configuration (L links)	Generation time [s/sample/link]
C-VAE/UC-GAN ($\hat{\mathbf{A}}_\theta$ generation)	$Z = 16, L = 1$	3.5×10^{-5}
	$Z = 32, L = 1$	5.2×10^{-5}
	$Z = 16, L = 81$	4.3×10^{-2}
	$Z = 32, L = 81$	5.2×10^{-2}
C-VAE/UC-GAN ($\hat{\mathbf{E}}_\theta$ generation)	$Z = 16, L = 1$	4.6×10^{-5}
	$Z = 32, L = 1$	6.3×10^{-5}
	$Z = 16, L = 81$	5.3×10^{-2}
	$Z = 32, L = 81$	6.1×10^{-2}
EM Diffraction (omnidirectional antennas)	Err.= $10^{-3}, L = 1$	5.4×10^{-3}
	Err.= $10^{-6}, L = 1$	3.81×10^{-2}
	Err.= $10^{-3}, L = 81$	0.78
	Err.= $10^{-6}, L = 81$	3.1
EM Diffraction (directional antennas)	Err.= $10^{-3}, L = 1$	0.64
	Err.= $10^{-6}, L = 1$	1.56
	Err.= $10^{-3}, L = 81$	51.8
	Err.= $10^{-6}, L = 81$	121

spatial dimensions of intermediate features, so that generated outputs respect the desired dimensions.

Considering now the UC-GAN model in Figs. 3b) and 3d), the discriminator (\mathbf{W}_D) and the generation (\mathbf{W}_G) model structures include similar components. Following the unbalanced GAN implementation, the C-VAE decoder model parameters \mathbf{W}_D are transferred to discriminator \mathbf{W}_D at the beginning of the training stage. To simplify comparison, the outputs of both the models have the same dimension. For an assigned input $\theta_k = [\mathbf{p}_k, h_S, w_{S,1}, w_{S,2}]$ VAE and GAN generate 201 different subject orientations φ_k in the interval $-\pi/2 \leq \varphi_k \leq \pi/2$, for all the configured physical links³ L .

Generation times. Generation times of C-VAE, UC-GAN and EM diffraction models are compared in Tab. 1 considering single antenna TX and RX and a MIMO setup with $L = 81$ links. Reproduction of excess attenuations $\hat{\mathbf{A}}_\theta$ and EM field samples $\hat{\mathbf{E}}_\theta$ are analyzed separately. For each case, time measurements are obtained using a Jetson Nano single-board computer equipped with a quad-core ARM-Cortex-A57 SoC, 4 GB RAM, 128 Compute Unified Device Architecture (CUDA) cores, and a Maxwell GPU architecture. This is representative of a typical resource-constrained wireless device. Note that, on average, the VAE/GAN-based generation of the EM body effects is about $\times 60 \div 100$ times faster than EM model computation. The latter also depends on the chosen numerical integration configurations, i.e., tiled integration method, and absolute error tolerance, target size, and antenna configuration (omnidirectional [15] vs directional [42] antenna radiation patterns). The generative model can be therefore used to reproduce the desired prior distribution in real-time, with sufficiently high randomness of samples. Generation rate is in the order of $50 \div 100$

³Pre-trained models [41] generate EM samples for $L = 81$ links where TX and RX have 9 antennas with spacing $\Delta_a = \lambda/2$.

TABLE 2. C-VAE and UC-GAN comparative analysis of model footprints for body-induced $\hat{\mathbf{A}}_\theta$, RF excess attenuations and EM field $\hat{\mathbf{E}}_\theta$ generation.

Model	Configuration (L links)	Model footprint [Mbyte]
C-VAE/UC-GAN ($\hat{\mathbf{A}}_\theta$ generation)	$Z = 16, L = 1$	0.891
	$Z = 32, L = 1$	1.687
	$Z = 16, L = 81$	62.8
	$Z = 32, L = 81$	~ 123
C-VAE/UC-GAN ($\hat{\mathbf{E}}_\theta$ generation)	$Z = 16, L = 81$	~ 117
	$Z = 32, L = 81$	~ 235

samples per second, which is reasonable considering typical body movement speeds (max. 1 m/s).

GNN model footprints. For the proposed implementations, Tab. 2 analyzes the size of the trainable parameters, namely the model footprint, of the decoder (C-VAE) and the generator (UC-GAN). Footprints range from 1 MB to 240 MB, being the EM field generation $\hat{\mathbf{E}}_\theta$ more demanding in terms of memory occupation than excess attenuation $\hat{\mathbf{A}}_\theta$ generation. Although out of the scope of the current paper, accurate model pruning is desirable to minimize the memory footprint on resource-constrained devices [5].

V. ANALYSIS OF EM FIELD GENERATION ACCURACY

In this section, we assess the ability of C-VAE and UC-GAN approaches to reproduce the EM diffraction effects (Section III) and, more generally, the effectiveness of the models in sampling from the prior $p(\mathbf{E}_\theta|\theta_k)$ in (3). We consider the problem of reproducing attenuations (\mathbf{A}_θ) and CSI (\mathbf{C}_θ), separately, for varying input features and scenarios.

The setup consists of TX and RX nodes equipped with ULAs having 9 omni-directional antennas (corresponding to $L = 81$ total links) that are spaced at $\Delta_a = \lambda/2$. The length of the central link of the array is equal to $d = 4$ m while all the links of the array are horizontally placed at height $h = 0.99$ m from the ground. The human target has also variable height h_S , traversal max. and min. sizes equal to $w_{S,1}$ and $w_{S,2}$, respectively. Generative models C-VAE and UC-GAN are trained using samples of EM body diffraction at carrier frequency $f_c = 2.4$ GHz. The training samples correspond to the following settings of the body configurations $\theta_k = [\mathbf{p}_k, \varphi_k, h_S, w_{S,1}, w_{S,2}]$:

- 1) 75 marked positions $\mathbf{p}_k, k = 1, \dots, K = 75$, located inside the Fresnel's area of the considered link, so that the spacing between marked positions is 0.25 m along and across the LOS (for $d = 4$ m);
- 2) 4 subject rotations $\varphi_k = [-\pi/2, -\pi/3, -\pi/6, 0]$;
- 3) 21 different dimensions of the targets ranging from $h_S = 1.2$ m to $h_S = 2.2$ m and $w_{S,1} = 0.25$ m to $w_{S,1} = 0.55$ m. Note that, in the reported tests, the dimension $w_{S,2}$ is fixed, by setting $w_{S,2} = 0.25$ m.

Notice that these limitations are reasonable as far as the goal is to represent a human body [16].

C-VAE method requires parameter optimization, namely optimization of the number of latent variables Z and ELBO weights β , which is the goal of the first part of the analysis in Section V-A. Next, in Section V-B C-VAE and UC-GAN reproduction accuracy is compared considering a multiple antenna array configuration. Finally, in Section V-C we evaluate the accuracy of the generated EM field samples for reproducing the array response.

A. C-VAE LATENT VARIABLE OPTIMIZATION

Fig. 4 shows an example of C-VAE generation of diffraction model samples, namely body-induced excess attenuations, using varying latent variables ranging from $Z = 8$ to $Z = 32$ and ELBO weight $\beta = 0.05$. The subject is moving along and across the (single) radio link in specific marked locations \mathbf{p}_k , as well as changing its orientation $0 \leq \varphi \leq -\pi/2$. TX and RX are equipped with a single antenna ($L = 1$). Here, we are interested in generating body excess attenuations $\hat{\mathbf{A}}_\theta^{\text{VAE}} = [\hat{A}_{1,\theta}^{\text{VAE}}]$. To account for the uncertainties introduced by different body postures and small, i.e., involuntary, movements in the assigned location \mathbf{p}_k , we report the average excess attenuations w.r.t. 50 generated samples from $\hat{A}_{1,\theta}^{\text{VAE}} \sim p_{\text{gen}}^{\text{VAE}}(\mathbf{A}_\theta|\theta_k)$.

In Fig. 4a), the C-VAE model is used to reproduce the average excess attenuation samples corresponding to a subject that is moving along the LOS ($0.25 \text{ m} \leq x \leq 3.75 \text{ m}$, $y = 0$) with a step of 0.25 m, namely occupying 15 marked locations, from $\mathbf{p}_1 = [0.25 \text{ m}, 0]$ to $\mathbf{p}_{15} = [3.75 \text{ m}, 0]$. The target has different dimensions, namely $h_S = 1.4$ m, $w_{S,1} = 0.35$ m (black), $h_S = 2.0$ m, $w_{S,1} = 0.65$ m (red), $h_S = 1.65$ m, $w_{S,1} = 0.65$ m (green) and $w_{S,2} = 0.25$ m. Also, it is changing its orientation while standing in each marked location. Fig. 4b) shows the corresponding generated samples now featuring a subject moving across the LOS ($-0.5 \text{ m} \leq y \leq 0.5 \text{ m}$, $x = 1 \text{ m}$) and with same dimensions. Finally, in Fig. 4c), the target is now fixed in position $\mathbf{p} = [0.5 \text{ m}, 0]$ but uniformly changing its orientation φ from $\varphi = -\pi/2$ to $\varphi = 0$. Generated samples are compared with the average EM body excess attenuations $A_{1,\theta}$ obtained from (6) and (7) via numerical methods for the same link and corresponding positions (dashed lines). The excess attenuations are averaged over 50 random target movements in an elementary squared area of size $\Delta = 0.1$ m surrounding the corresponding marked positions \mathbf{p}_k . Considering all the tests, we found that using $Z = 16$ latent variables constitutes a good compromise between complexity and accuracy.

In addition to average excess attenuation terms, Fig. 5 analyzes the distribution of the generated excess attenuation samples $p_{\text{gen}}^{\text{VAE}}(\mathbf{A}_\theta|\theta_k)$ compared with those obtained from the EM diffraction model. Prior distributions are obtained for varying latent variables Z and two choices of ELBO weight β , namely $\beta = 0.05$, for Figs. 5a) and 5c); and $\beta = 1e-09$, for Figs. 5b) and 5d). Three target configurations are considered, namely $h_S = 1.4$ m, $w_{S,1} = 0.35$ m in Figs. 5a), 5b), 5c), and 5d); $h_S = 1.65$ m, $w_{S,1} =$

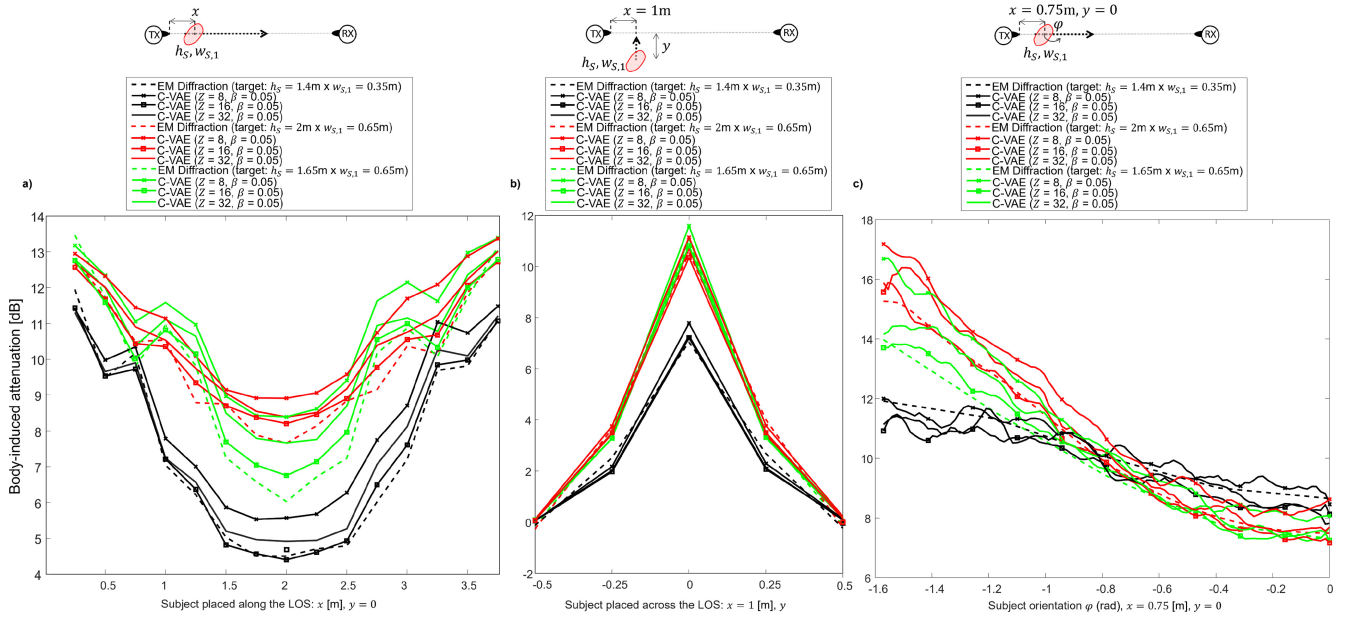


FIGURE 4. C-VAE generation of body-induced excess attenuation values A_{θ} for different target movements (along, across the LOS and varying orientations) ($h_s, w_{S,1}$), with $w_{S,2} = 0.25$ m. C-VAE results are shown for varying latent samples Z and $\beta = 0.05$. From left to right: a) the subject is moving along the LOS ($0.25 \text{ m} \leq x \leq 3.75 \text{ m}$, $y = 0$). The generated EM body excess attenuation values obtained via numerical methods are represented in dashed lines by averaging over random target orientations $-\pi/2 \leq \varphi \leq \pi/2$ and random movements in an elementary squared area of size $\Delta = 0.1$ m. b) The subject is moving across the LOS ($x = 1 \text{ m}$, $-0.5 \text{ m} \leq y \leq 0.5 \text{ m}$). Dashed lines show the corresponding ground-truth diffraction model samples obtained similarly as in a). c) The target is in position $x = 0.75 \text{ m}$, $y = 0$ and changing orientation $-\pi/2 \leq \varphi \leq \pi/2$ while performing small movements in the same elementary area. Dashed lines shows the EM body model excess attenuations for a subset of the subject orientations. Figs. a) and c) are also presented in [24]

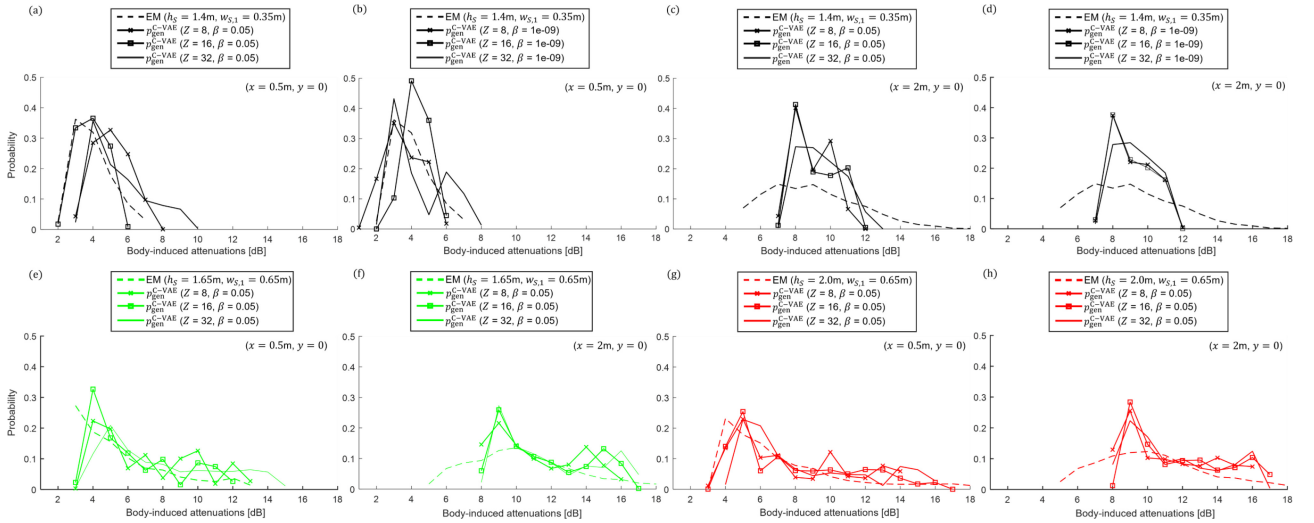


FIGURE 5. C-VAE generated sample probabilities $p_{\text{gen}}^{\text{VAE}}(A_{\theta}|\theta_k)$ of body-induced RF excess attenuations for varying latent dimensions ($Z = 8, 16, 32$) and ELBO weights β , compared with samples obtained from EM body model (dashed lines). In (a) and b) the target stands at $x = 0.5 \text{ m}$ from the TX with size $h_s = 1.4 \text{ m}$, $w_{S,1} = 0.35 \text{ m}$. Generation exploits $\beta = 0.05$ in a) and $\beta = 1e-09$ in b). In c) and d) the target stands at $x = 2 \text{ m}$ from the TX with size $h_s = 1.4 \text{ m}$, $w_{S,1} = 0.35 \text{ m}$ while β is set to $\beta = 0.05$ in c) and $\beta = 1e-09$ in d), respectively. The target, with size $h_s = 1.65 \text{ m}$, $w_{S,1} = 0.65 \text{ m}$, stands at distance from the TX equal to $x = 0.5 \text{ m}$ in e) and at $x = 2 \text{ m}$ in f). The target with size $h_s = 2.0 \text{ m}$, $w_{S,1} = 0.65 \text{ m}$ stands at distance from the TX equal to $x = 0.5 \text{ m}$ in g) and at $x = 2 \text{ m}$ in h). For the cases in e), f), g) and h), it is $\beta = 0.05$.

0.65 m in Figs. 5e), and 5f); and $h_s = 2.0 \text{ m}$, $w_{S,1} = 0.65 \text{ m}$ in Figs. 5g) and 5h). The generated distributions $p_{\text{gen}}^{\text{VAE}}$ reproduce the excess attenuations observed with a target standing at distance from the TX equal to 0.5 m in Figs. 5a), 5b), 5e), and 5g); and equal to 2 m in Figs. 5c), 5d), 5f), and 5h), but changing its posture in the same squared elementary area of size $\Delta = 0.1 \text{ m}$ previously considered.

As evident from the corresponding cases, the number of latent variables Z substantially affects the generated samples, while the ELBO weight β seems to have less evident effects. The C-VAE tool configured for $Z = 16$ (and $Z = 32$) provides a good representation of the distribution of the excess attenuations when compared with the EM diffraction model one. On the other hand, the C-VAE model seems to better reproduce the excess attenuation samples

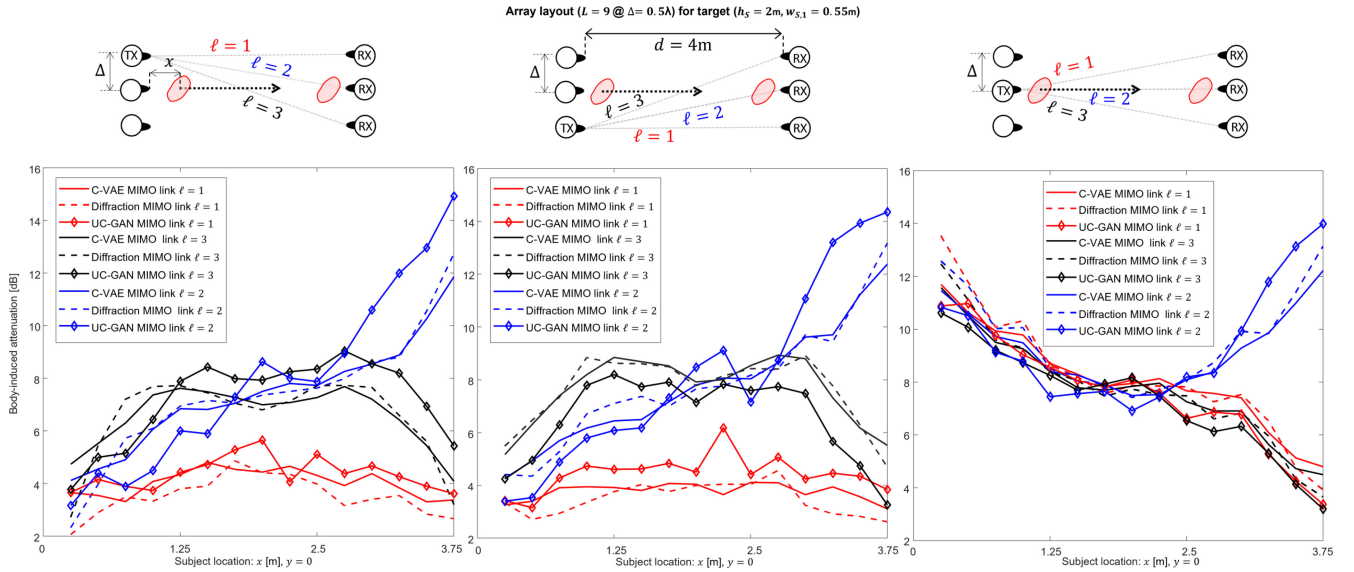


FIGURE 6. C-VAE vs UC-GAN generation of EM body model for a MIMO array consisting of 3 antennas at the TX and RX, respectively, and $L = 9$ radio links. The subject has dimensions ($h_S = 2$ m, $w_{S,1} = 0.55$ m, $w_{S,2} = 0.25$ m) and is moving along the LOS path of link $\ell = 5$ (0.25 m $\leq x \leq 3.75$ m, $y = 0$). The EM body average excess attenuation values $\mathbf{A}_\theta = [A_{\ell,\theta}]_{\ell=1}^{L=9}$ obtained through C-VAE (solid lines) and UC-GAN (diamond markers) methods are compared with the corresponding diffraction model samples (dashed lines). C-VAE latent variable dimension is $Z = 16$ with $\beta = 0.05$, as optimized as in Fig. 4. UC-GAN is pre-trained using C-VAE decoder model.

corresponding to targets placed at some distance, i.e., 2 m from the TX, rather than close to TX, i.e., 0.5 m. The trend is particularly evident when the C-VAE model is set to reproduce sample distribution with high variance and few training samples, as in the case for small target size ($h_S = 1.4$ m, $w_{S,1} = 0.35$ m). Finally, the choice for $\beta = 0.05$ stands as a good compromise between the average reproduction accuracy and the reconstruction of the entire probability function which require to increase the randomness of generated samples [24].

B. C-VAE AND UC-GAN MODEL COMPARISON

In Fig. 6, we compare the C-VAE generation tool using the optimized parameters $Z = 16$, $\beta = 0.05$ shown previously with the UC-GAN implementation described in Section IV-B. The following analysis is of interest as it shows the behavior of two different generative systems and compares their ability to reproduce the EM body diffraction effects. With respect to previous section, we now consider a MIMO ULA setup consisting of 3 antennas at the TX and RX, respectively, $L = 9$ radio links and distance $d = 4$ m. The samples $\hat{\mathbf{A}}_\theta^{\text{VAE}} = [\hat{A}_{\ell,\theta}^{\text{VAE}}]_{\ell=1}^{L=9}$ obtained through C-VAE (solid lines) and with UC-GAN $\hat{\mathbf{A}}_\theta^{\text{GAN}} = [\hat{A}_{\ell,\theta}^{\text{GAN}}]_{\ell=1}^{L=9}$ (diamond markers) tools are compared with the corresponding EM body-induced excess attenuations $\mathbf{A}_\theta = [A_{\ell,\theta}]_{\ell=1}^{L=9}$ obtained from diffraction theory (dashed lines). The subject has fixed dimensions $h_S = 2$ m, $w_{S,1} = 0.55$ m, $w_{S,2} = 0.25$ m, and it is moving along the LOS path of the link $\ell = 5$ (0.25 m $\leq x \leq 3.75$ m, $y = 0$).

Tab. 3 reports a comparative analysis of C-VAE and UC-GAN generation for the same MIMO setup, in terms of Mean Squared Error (MSE) and Kullback-Leibler (KL) divergence D_{KL} [47]. The latter compares the distance (divergence)

TABLE 3. C-VAE and UC-GAN comparative analysis in terms of MSE and KL divergence D_{KL} .

	Target (h_S, x, y) [m]	MSE [dB]	D_{KL}
C-VAE	$h_S = 1.4 \div 1.6, x \leq 0.5, y = 0$	0.28	0.87
	$h_S = 1.4 \div 1.6, x = 0.5 \div 2, y = 0$	0.19	0.46
	$h_S = 1.6 \div 1.8, x \leq 0.5, y = 0$	0.44	0.37
	$h_S = 1.6 \div 1.8, x = 0.5 \div 2, y = 0$	0.37	0.41
	$h_S = 1.8 \div 2, x \leq 0.5, y = 0$	0.14	0.38
	$h_S = 1.8 \div 2, x = 0.5 \div 2, y = 0$	0.31	0.37
UC-GAN	$h_S = 1.4 \div 1.6, x \leq 0.5, y = 0$	1.74	>3
	$h_S = 1.4 \div 1.6, x = 0.5 \div 2, y = 0$	1.07	>3
	$h_S = 1.6 \div 1.8, x \leq 0.5, y = 0$	1.43	>3
	$h_S = 1.6 \div 1.8, x = 0.5 \div 2, y = 0$	0.74	>3
	$h_S = 1.8 \div 2, x \leq 0.5, y = 0$	1.12	2.21
	$h_S = 1.8 \div 2, x = 0.5 \div 2, y = 0$	0.41	2.3

among the generated sample probability functions, $p_{\text{gen}}^{\text{VAE}}$ or $p_{\text{gen}}^{\text{GAN}}$ and the theoretical ones obtained from EM diffraction. The MSE and the KL divergence terms are computed for link $\ell = 5$ and varying target dimensions. The MSE values remain below 0.5dB for C-VAE, on the other hand they are about 1dB higher for UC-GAN. Similarly, by observing the KL divergence, the C-VAE model is able to better reproduce the true distribution of the samples (KL divergence < 1) compared with UC-GAN, which features in some cases large deviations (> 3). As previously noticed, the C-VAE model seems to better reproduce the diffraction samples for targets placed at distance $x > 0.5$ m from the TX. Also, large targets surfaces S ($h_S \geq 1.8$ m) are better represented by C-VAE model than small surfaces ($h_S \leq 1.6$ m). Even if not considered in this paper, a possible solution could

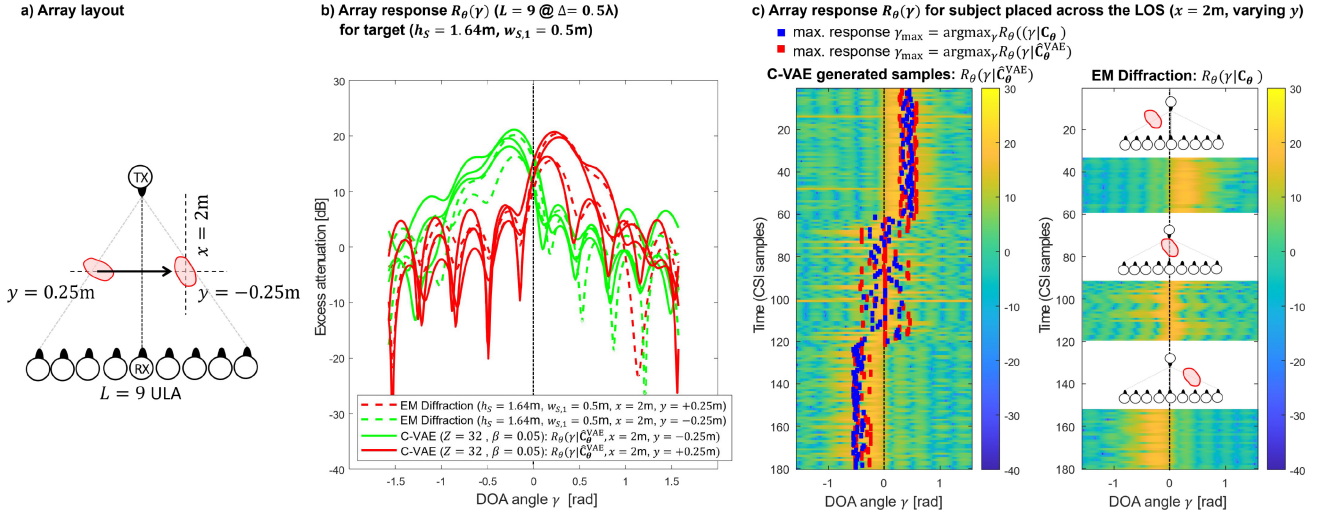


FIGURE 7. a) Array layout setup; b) Array response $R_\theta(\gamma)$ for $\gamma = 0$ and a target with dimensions $h_S = 1.65$ m, $w_{S,1} = 0.55$ m, $w_{S,2} = 0.25$ m and 2 positions, namely $x = 2$ m, $y = -0.25$ m and $x = 2$ m, $y = 0.25$ m. Response is obtained using C-VAE EM field generated samples $\hat{\mathbf{C}}_\theta^{\text{VAE}}$ ($Z = 16$, $\beta = 0.05$) - green/red solid lines - and compared with the array response obtained with EM Diffraction field samples (dashed lines); c) Direction of Arrival (DoA) analysis for a subject moving across the LOS ($-0.25 \text{ m} \leq y \leq 0.25 \text{ m}$, $x = 2 \text{ m}$), and changing orientations randomly in the interval $-\pi/2 \leq \varphi \leq \pi/2$.

be to augment the size of training data for these more disadvantageous cases.

C. ANALYSIS OF GENERATED BEAMFORMING ARRAY RESPONSES

Based on the analysis in Section III-C, we now verify the ability of the proposed C-VAE model to reproduce the array response of a conventional linear beamforming processing. We thus compare the reproduced response $R_\theta(\gamma|\hat{\mathbf{C}}_\theta^{\text{VAE}})$ using the generated EM field samples $\hat{\mathbf{C}}_\theta^{\text{VAE}}$ with the *true* response $R_\theta(\gamma|\mathbf{C}_\theta)$ obtained by diffraction as in (11).

Fig. 7 considers an RX-side UL array layout now consisting of $L = 9$ antennas (see Fig. 7a) and shows the body-induced array response $20 \log_{10} |R_\theta(\gamma)|$ as a function of the DoA γ and for different values of the y displacement of the target (w.r.t. the central LOS) and $x = 2$ m. The array signal processing is set to extract the response for varying DoA γ and is based on Fast Fourier Transform (FFT) with $N_{\text{FFT}} = 257$ points. Fig. 7b) shows two responses (red and green lines) for corresponding target locations $y = -0.25$ m and $y = 0.25$ m, respectively. The theoretical responses $R_\theta(\gamma|\mathbf{C}_\theta)$ using EM diffraction are in dashed lines while solid lines represent the reproduced responses $R_\theta(\gamma|\hat{\mathbf{C}}_\theta^{\text{VAE}})$ using C-VAE generated CSI samples $\hat{\mathbf{C}}_\theta^{\text{VAE}}$.

Fig. 7c) compares the maximum response γ_{\max} of the array as defined in (13). Blue dots are the dominant DoA obtained by maximizing the array response $R_\theta(\gamma|\mathbf{C}_\theta)$. Red dots refer to the DoA produced by C-VAE generated response $R_\theta(\gamma|\hat{\mathbf{C}}_\theta^{\text{VAE}})$. Both cases simulate a target moving across the LOS ($-0.25 \text{ m} \leq y \leq 0.25 \text{ m}$, $x = 2 \text{ m}$), with speed 0.5 m/s and changing orientation randomly in the interval $0 \leq \varphi \leq \pi/2$. It can be noticed that the maximum response γ_{\max} is perturbed by the presence of the subject and such alteration is well reproduced by the C-VAE model.

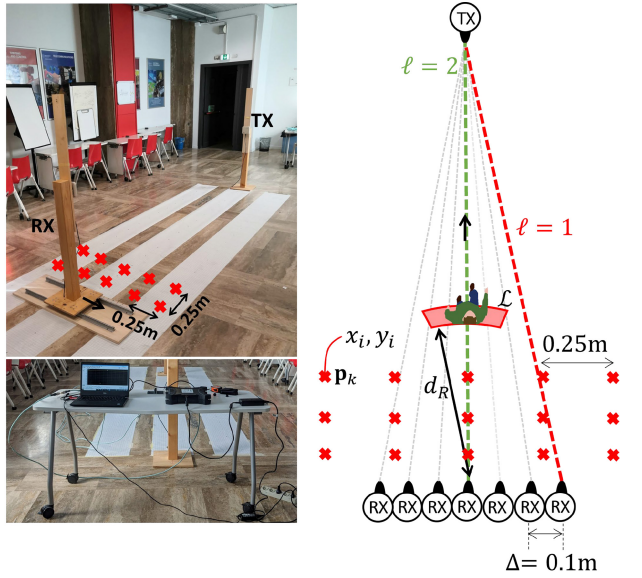
VI. CASE STUDIES IN PASSIVE RADIO LOCALIZATION

A specific case study is considered in this section. The goal is to demonstrate the effectiveness of the proposed GNN tools to reproduce the prior $p(\mathbf{E}_\theta|\theta_k)$ in (3) and to support a real-time localization process. The EM-informed C-VAE tool has been thus validated with measurements taken in a hall of size $6.1 \text{ m} \times 14.4 \text{ m}$ as shown in Fig. 8. Both TX and RX are equipped with directional antennas with parameters summarized in the table embedded in Fig. 8. A mechanical handling mechanism, shown in the top left of Fig. 8, is used to move the RX antenna at specified positions where RF measurements on multiple links ℓ are collected. The target is located in $K = 75$ marked positions \mathbf{p}_k , $k = 1, \dots, K$, which belong to a regular 2D grid as shown in the same figure. A tracking generator enabled spectrum analyzer [49] is used to measure the RSS \mathbf{S}_t in the $2.4 \div 2.5$ GHz band and over 81 frequencies with 1.25 MHz spacing. For each frequency and target position under test, 500 consecutive time samples are acquired in 1 min. (120 ms sampling time). The human target (one of the authors who volunteered) has height $h_S = 1.80$ m and traversal max. and min. sizes approximated as $w_{S,1} = 0.55$ m and $w_{S,2} = 0.25$ m, respectively.

A. GENERATION OF RECEIVED SIGNAL STRENGTH SAMPLES

So far we considered generative models to reproduce human-body blockage effects \mathbf{E}_θ . In what follows, we highlight a dual use of the tool to reproduce the raw RSS measurements $\mathbf{S}_t = [S_{\ell,t}]_{\ell=1}^L$ for an assigned body state θ_k and link set L as in the scenario of Fig. 8. The RSS measurements $\hat{S}_{\ell,t}$ are generated via Monte Carlo sampling method and using the C-VAE tool as follows:

$$\hat{S}_{\ell,t} = \begin{cases} P_{\ell,0} - \hat{A}_{\ell,\theta}^{\text{VAE}} + w_{\ell,0} & \mathbf{p}_k \notin \mathcal{F}_\ell \\ P_{\ell,0} - \hat{A}_{\ell,\theta}^{\text{VAE}} + w_{\ell,T} & \mathbf{p}_k \in \mathcal{F}_\ell, \end{cases} \quad (18)$$



Start/stop frequency: 2.4GHz – 2.5GHz	Power: 0 dBm
Frequency spacing: 1.25MHz	Antenna HPBW: H (60 deg.), V (76 deg.)
Resolution: 100 kHz	Antenna gain: 9 dBi

FIGURE 8. Measurement setup, TX/RX antennas, linear guide system for RX antenna positioning. Explored target positions around the Fresnel's area.

where $\hat{A}_{\ell,\theta}^{\text{VAE}} \sim p_{\text{gen}}^{\text{VAE}}(A_{\ell,\theta}|\theta_k)$, $P_{\ell,0}$ is the free-space received power, known or measured during calibration, while $w_{\ell,0}$ and $w_{\ell,T}$ model the log-normal multipath fading and the other noise sources. Disturbances $w_{\ell,0}$ and $w_{\ell,T}$ differ depending on whether the target is located inside ($\mathbf{p}_k \in \mathcal{F}_\ell$) or outside ($\mathbf{p}_k \notin \mathcal{F}_\ell$) the first Fresnel zone \mathcal{F}_ℓ of the considered link ℓ . Noise $w_{\ell,0} \sim \mathcal{N}(0, \sigma_0^2)$ refers to the *free-space case* with $\mathbf{p}_k \notin \mathcal{F}_\ell$, while noise $w_{\ell,T} \sim \mathcal{N}(\mu_T, \sigma_T^2)$ refers to the case with $\mathbf{p}_k \in \mathcal{F}_\ell$. In the followings, we set $\mu_T = 1.5\text{dB}$, $\sigma_T = 1\text{dB}$ and $\sigma_0 = 1\text{dB}$ [16].

Fig. 9 compares the generated RSS samples $\hat{S}_{\ell,t}$ with RSS measurements $S_{\ell,t}$ at 2.4GHz for the scenario in Fig. 8 and link $\ell = 2$. The subject is standing while performing small movements around 4 nominal positions $\mathbf{p}_k = (x, y)$, namely $x = 0.25\text{ m}$, $y = 0$ (blue color), $x = 0.5\text{ m}$, $y = 0$ (violet color), $x = 0.75\text{ m}$, $y = 0$ (yellow color). Position outside the Fresnel area is $x = 2\text{ m}$, $y = -0.5$ (black color). Considering the same scenario, Fig. 10 evaluates the C-VAE generation of RSS samples obtained from two RX locations, namely the links $\ell = 1$ (red) and $\ell = 2$ (green), corresponding to a target now moving along the LOS $0.25\text{ m} \leq x \leq 3.75\text{ m}$ ($y = 0$). The generated samples are again compared with RSS measurements where the target is set to move along the LOS path with a constant speed approximated as 0.25 m/s .

The average error $\varepsilon_\ell(\mathbf{p}_k) = \text{E}_t[S_{\ell,t}(\mathbf{p}_k)] - \text{E}_t[\hat{S}_{\ell,t}(\mathbf{p}_k)]$ between the RSS values $\hat{S}_{\ell,t}(\mathbf{p}_k)$ predicted by the C-VAE model and the corresponding measurements $S_{\ell,t}(\mathbf{p}_k)$ are summarized in Tab. 4 for varying target locations \mathbf{p}_k along the LOS path. The corresponding error $\delta_\ell(\mathbf{p}_k) = \sigma_\ell(\mathbf{p}_k) - \hat{\sigma}_\ell(\mathbf{p}_k)$ between the true $\sigma_\ell = \sqrt{\text{E}_t[S_{\ell,t}(\mathbf{p}_k) - \text{E}_t[S_{\ell,t}(\mathbf{p}_k)]]^2}$

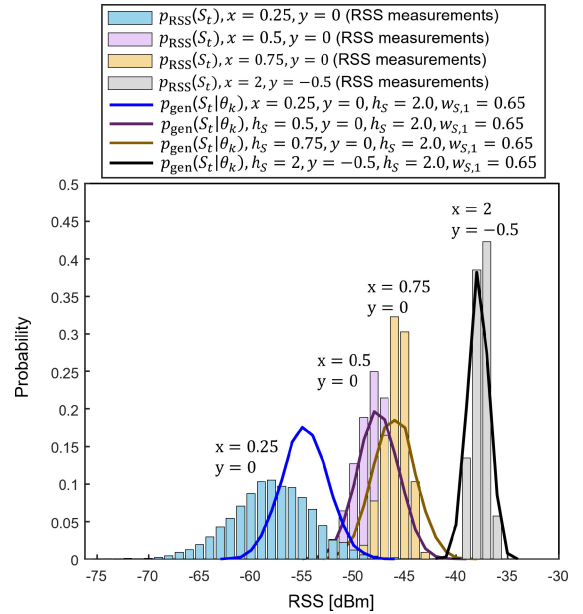


FIGURE 9. Sample probabilities $p_{\text{gen}}(S_{\ell}|\theta_k)$ of generated RSS samples using C-VAE (18) compared with histograms $p_{\text{RSS}}(S_{\ell})$ obtained from true RSS measurements at 2.4GHz. Target size is $h_s = 2\text{ m}$, $w_{s,1} = 0.55\text{ m}$, and orientation φ is uniformly distributed in $0 \leq \varphi \leq \pi/6$, to model small/involuntary movements in the surrounding of the nominal positions. The subject is standing while performing small movements around 4 nominal positions $\mathbf{p}_k = (x, y)$, namely $x = 0.25\text{ m}$, $y = 0$ (blue color), $x = 0.5\text{ m}$, $y = 0$ (violet color), $x = 0.75\text{ m}$, $y = 0$ (yellow color), and outside Fresnel area $x = 2\text{ m}$, $y = -0.5$ (black color). A single link is considered ($\ell = 2$).

and predicted $\hat{\sigma}_\ell = \sqrt{\text{E}_t[\hat{S}_{\ell,t}(\mathbf{p}_k) - \text{E}_t[\hat{S}_{\ell,t}(\mathbf{p}_k)]]^2}$ standard deviation values are shown as well. The generative model can be effectively used to predict the true RSS values for both the considered links with average error of 2.7dB and max standard deviation error of 2.9dB. On the other hand, as also observed in Section V-A, the C-VAE generation tool seems to over-estimate ($\varepsilon_\ell < 0$) the observed RSS values for target positions close to the transmitter or receiver, i.e., $x = 3\text{ m}$, since it is trained using diffraction models [15].

B. PASSIVE LOCALIZATION USING RF GENERATED SAMPLES

In this section, we discuss an example of passive localization. The goal is to detect the distance d_R of the target from the multi-antenna RX device in real-time, as in typical radar applications. Given the RSS observations S_t over the same links considered in Fig. 8, we want to recover the estimated target state θ_k , in our case the position \mathbf{p}_k of the target relative to RX.⁴ The proposed use case is critical in industrial scenarios where human workers operate in areas featuring increasing level of safety or privacy. Enforcing safety/privacy constraints requires the real-time monitoring and tracking of the human subject.

The estimated position $\hat{\mathbf{p}}_k$ of the target relative to RX is obtained as in (2), replacing $\hat{\theta}$ with $\hat{\mathbf{p}}_k$. Next, the target distance is derived as $d_R = |\hat{\mathbf{p}}_k|$. The C-VAE generated attenuations are here used to reproduce the prior distribution.

⁴To simplify the reasoning, the target moves along the LOS path.

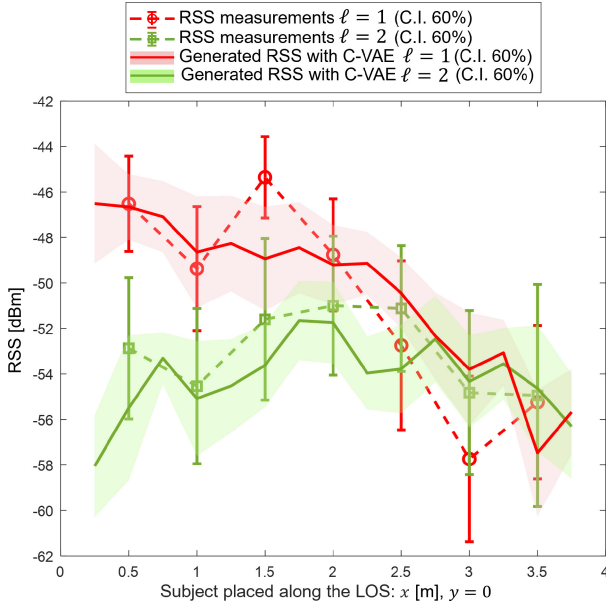


FIGURE 10. C-VAE generation of RSS samples observed for links $\ell = 1$ (red) and $\ell = 2$ (green) described in Fig. 8 and corresponding to a target moving along the LOS $0.25 \text{ m} \leq x \leq 3.75 \text{ m}$ ($y = 0$). Generated samples are compared with RSS measurements at 2.4GHz. Both the generated samples and the corresponding true measurements have Confidence Interval (CI) of 60%.

TABLE 4. Average ε_ℓ and standard deviation δ_ℓ error analysis between true RSS and predicted via C-VAE generative model (setup in Fig. 8).

Target $\mathbf{p}_k : (x, y = 0)$	Link $\ell = 1$		Link $\ell = 2$	
	$\varepsilon_1(\mathbf{p}_k)$	$\delta_1(\mathbf{p}_k)$	$\varepsilon_2(\mathbf{p}_k)$	$\delta_2(\mathbf{p}_k)$
$x = 0.5\text{m}$	0.17dB	-0.61dB	0.29dB	0.29dB
$x = 1\text{m}$	1.59dB	-1.70dB	1.10dB	-0.2dB
$x = 1.5\text{m}$	2.65dB	-1.81dB	0.31dB	2.89dB
$x = 2\text{m}$	-0.4dB	-0.74dB	1.36dB	0.61dB
$x = 2.5\text{m}$	-3.09dB	2.75dB	0.84dB	0.51dB
$x = 3\text{m}$	-4.53dB	0.77dB	-1.59dB	1.53dB
$x = 3.5\text{m}$	-0.36dB	-0.53dB	-1.15dB	-0.51dB

The problem simplifies to $\hat{\mathbf{p}}_k = \arg \max_k \{p_A(\theta_k)\}$: using simple Bayes rule considerations and eq. (2), the probabilities p_A are defined as:

$$\begin{aligned}
 p_A(\theta_k) &= \max_{\theta} \log [p(\mathbf{A}_{\theta} | \mathbf{S}_t, \theta_k)] \\
 &= \max_{\theta} \sum_{\ell} \log [p(S_{\ell,t} | A_{\ell,\theta}) \cdot p_{\text{gen}}^{\text{VAE}}(A_{\ell,\theta} | \theta_k)].
 \end{aligned} \tag{19}$$

Tab. 5 analyzes the precision and recall probabilities:

$$\begin{aligned}
 p_1(d_R) &= \Pr[\mathbf{p}_k \in \mathcal{L}(d_R) | \hat{\mathbf{p}}_k \in \mathcal{L}(d_R)] \\
 p_2(d_R) &= \Pr[\hat{\mathbf{p}}_k \in \mathcal{L}(d_R) | \mathbf{p}_k \in \mathcal{L}(d_R)]
 \end{aligned} \tag{20}$$

with $\mathcal{L}(d_R)$ being the region that contains the positions \mathbf{p}_k of the target at distance of d_R from the RX. The recall metric measures how often the algorithm correctly identifies the target distance from all the true positive counts, while the precision indicates how often the algorithm is correct when predicting the target distance. The table analyzes the

TABLE 5. Precision p_1 and recall p_2 probabilities for a target at varying distance d_R from the multi-antenna RX. Estimated prior from calibration data, C-VAE prior and uniform prior are compared.

Distance d_R	$p_1(d_R) / p_2(d_R)$		
	Estimated prior	C-VAE prior	Uniform prior
0.5 m	0.95/0.99	0.92/0.85	0.67/0.48
1 m	0.99/0.99	0.76/0.66	0.10/0.09
1.5 m	0.99/0.95	0.87/0.67	0.21/0.55
2 m	0.99/0.99	0.70/0.64	0.23/0.1
2.5 m	0.99/0.78	0.47/0.69	0.09/0.08
3 m	0.99/0.97	0.73/0.77	0.08/0.09
3.5 m	0.76/0.74	0.62/0.83	0.09/0.26

precision and the recall for varying distance d_R from the RX. Three cases are considered:

- i) *estimated prior* scenario: assumes the prior probability $p_{\text{gen}}(A_{\ell,\theta} | \theta_k)$ being estimated from calibration measurements;
- ii) *C-VAE prior*: adopts the probability $p_{\text{gen}}^{\text{VAE}}(A_{\ell,\theta} | \theta_k)$ as prior model with samples obtained using the C-VAE generator tool;
- iii) *uniform prior*: represents a case where no information on excess attenuation is available: the prior is replaced with a uniform probability function $\mathcal{U}_{-5\text{dB},15\text{dB}}$ with attenuations ranging from -5dB to 15dB .

Note that scenario i) gives the best performance, as expected; on the other hand, it requires time-consuming data collection and a calibration stage which might be not feasible in practice. Case iii) corresponds to the worst case scenario since no prior information on body-induced attenuations are available. Finally, case ii) does not need any calibration as it uses the C-VAE tool to real-time generate samples from the prior $p_{\text{gen}}^{\text{VAE}}(A_{\ell,\theta} | \theta_k)$. From the results in the Tab. 5, the performance of *C-VAE prior* scenario approaches the *estimated prior* case, with an average drop of about 10%.

VII. CONCLUSIONS AND FUTURE ACTIVITIES

The paper proposed the use of EM-informed Generative Neural Network (GNN) models to predict body-induced diffraction effects. Explored applications are in the field of passive radio sensing and localization. A Variational Auto-Encoder (VAE) tool, namely the Conditional VAE (C-VAE), is designed to generate samples of the targeted EM model through latent variable encoding/decoding neural network operations. The tool reproduces EM field samples corresponding to specific human body states that are user selectable and modifiable in real-time during model exploitation. Adaptations of Generative Adversarial Networks (GAN) are also considered for comparative analysis. Generated samples are set to reproduce both RF signal attenuations, i.e., Received Signal Strength (RSS), as well as base-band Channel State Information (CSI) for full EM analysis of human body blockage.

GNNs produce observations sampled from the Bayesian prior probability which supports Bayesian estimation problems. Examples tailored for passive localization reveal the possibility of optimizing the generation process so to limit the use of time-consuming calibration stages and intensive EM computations. Generated samples might also serve as *synthetic* training data for supervised or semi-supervised machine learning tasks. They thus reduce the need of personal data collection, that could be used maliciously for person (re-)identification.

Beside the advantages, the proposed generators require hyper-parameter optimization to achieve a good compromise between average reproduction accuracy and generalization capability. The former measures how close the prediction is to the training observations, the latter quantifies the randomness of generated samples which is useful to predict effects not seen during training. When compared with real measurements, the generated tools appear to underestimate some human blockage effects, i.e., for small targets placed close to the receivers. This opens the room for possible improvements. The considered generative systems are currently trained to reproduce scalar diffraction effects. 2D EM absorbing sheets are also used to model 3D human bodies. Training the system with different EM blockage models, such as full-wave solutions or Method of Moments (MoM), and/or using more accurate body models, might increase the generalization capabilities.

Although still in their infancy, we expect physics-informed GNN models to become indispensable tools for designers in different scenarios. For example, future radio sensing tools will be paired with accurate EM modeling in high frequency bands as proposed in emerging wireless communication standards (6G and beyond). The possibility of generating large CSI tensor structures representing the full RF radiation field is also useful in emerging holographic methods and microwave imaging techniques based on Synthetic Aperture Radar (SAR). Finally, the proposed tools have been proved to be effective in reproducing body motions in user selectable locations. The property is instrumental to privacy selective sensing policies.

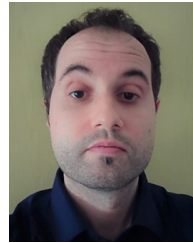
ACKNOWLEDGMENT

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Innovation Council and SMEs Executive Agency (EISMEA). Neither the European Union nor the granting authority can be held responsible for them. Grant Agreement No: 101099491.

REFERENCES

- [1] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proc. 13th Annu. ACM Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2007, pp. 222–229.
- [2] J. Wilson and N. Patwar, "Radio tomographic imaging with wireless networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 621–632, May 2010.
- [3] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, "Nuzzer: A large-scale device-free passive localization system for wireless environments," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1321–1334, Jul. 2013.
- [4] B. Rodrigues, C. Halter, M. Franco, E. J. Scheid, C. Killer, and B. Stiller, "BluePIL: A Bluetooth-based passive localization method," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manag. (IM)*, 2021, pp. 28–36.
- [5] S. Savazzi, S. Sigg, F. Vicentini, S. Kianoush, and R. Findling, "On the use of stray wireless signals for sensing: A look beyond 5G for the next generation of industry," *Computer*, vol. 52, no. 7, pp. 25–36, Jul. 2019.
- [6] O. Kaltiokallio, R. Hostettler, and N. Patwari, "A novel Bayesian filter for RSS-based device-free localization and tracking," *IEEE Trans. Mobile Comput.*, vol. 20, no. 3, pp. 780–795, Mar. 2021.
- [7] S. Savazzi, V. Rampa, F. Vicentini, and M. Giussani, "Device-free human sensing and localization in collaborative human-robot workspaces: A case study," *IEEE Sensors J.*, vol. 16, no. 5, pp. 1253–1264, Mar. 2016.
- [8] S. Palipana, B. Pietropaoli, and D. Pesch, "Recent advances in RF-based passive device-free localisation for indoor applications," *Ad Hoc Netw.*, vol. 64, pp. 80–98, Sep. 2017.
- [9] Z. Krupka, "The effect of the human body on radiation properties of small-sized communication systems," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 154–163, Mar. 1968.
- [10] M. Clemens et al., "Discrete electromagnetism with the finite integration technique," *Progress Electromagn. Res.*, vol. 32, no. 32, pp. 65–87, 2001.
- [11] A. Eleryan, M. Elsabagh, and M. Youssef, "Synthetic generation of radio maps for device-free passive localization," in *IEEE Global Telecommun. Conf. (GLOBECOM)*, 2011, pp. 1–5.
- [12] A. Abdelgawwad, A. C. Mallofré, and M. Pätzold, "A trajectory-driven 3D channel model for human activity recognition," *IEEE Access*, vol. 9, pp. 103393–103406, 2021.
- [13] G. Koutittas, "Multiple human effects in body area networks," *IEEE Antennas Wireless Propag. Lett.*, vol. 9, pp. 938–941, 2010.
- [14] Z. Wang, H. Liu, S. Xu, X. Bu, and J. An, "A diffraction measurement model and particle filter tracking method for RSS-based DFL," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2391–2403, Nov. 2015.
- [15] V. Rampa, G. G. Gentili, S. Savazzi, and M. D'Amico, "EM models for passive body occupancy inference," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, pp. 2517–2520, 2017.
- [16] V. Rampa, G. G. Gentili, S. Savazzi, and M. D'Amico, "Electromagnetic models for passive detection and localization of multiple bodies," *IEEE Trans. Antennas Propag.*, vol. 70, no. 2, pp. 1462–1745, Feb. 2022.
- [17] B. R. Hamilton, X. Ma, R. J. Baxley, and S. M. Matechik, "Propagation modeling for radio frequency tomography in wireless networks," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 55–65, Feb. 2014.
- [18] S. Kianoush, V. Rampa, S. Savazzi, and M. Nicoli, "Pre-deployment performance assessment of device-free radio localization systems," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, 2016, pp. 1–6.
- [19] Z. Zhao, J. C. Ye, and Y. Bresler, "Generative models for inverse imaging problems: From mathematical foundations to physics-driven applications," *IEEE Signal Process. Mag.*, vol. 40, no. 1, pp. 148–163, Jan. 2023.
- [20] T. Shan et al., "Physics-informed supervised residual learning for electromagnetic modeling," *IEEE Trans. Antennas Propag.*, vol. 71, no. 4, pp. 3393–3407, Apr. 2023.
- [21] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, "Physics-informed neural networks with hard constraints for inverse design," *SIAM J. Sci. Comput.*, vol. 43, no. 6, pp. B1105–B1132, 2021.
- [22] P. Di Barba, "Future trends in optimal design in electromagnetics," *IEEE Trans. Magn.*, vol. 58, no. 9, pp. 1–4, Sep. 2022.
- [23] M. Baldan, P. Di Barba, and D. A. Lowther, "Physics-informed neural networks for inverse electromagnetic problems," *IEEE Trans. Magn.*, vol. 59, no. 5, pp. 1–5, May 2023.
- [24] S. Savazzi, F. Fieramosca, S. Kianoush, V. Rampa, and M. D'Amico, "A physics-informed generative model for passive radio-frequency sensing," in *Proc. IEEE-APS Topical Conf. Antennas Propag. Wireless Commun. (APWC)*, 2023, pp. 56–61.

- [25] S. Bond-Taylor, A. Leach, Y. Long, and C. G. Willcocks, "Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 7327–7347, Nov. 2021.
- [26] I. Goodfellow et al., "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [27] B. Bartlett, "A 'generative' model for computing electromagnetic field solutions," Lecture Notes CS229, Stanford Univ., Stanford, CA, USA, 2018. [Online]. Available: <https://cs229.stanford.edu/proj2018/report/233.pdf>
- [28] T. Shan et al., "Study on a fast solver for poisson's equation based on deep learning technique," *IEEE Trans. Antennas Propag.*, vol. 68, no. 9, pp. 6725–6733, Sep. 2020.
- [29] L. Lu, X. Meng, Z. Mao, and G. E. Karniadakis, "DeepXDE: A deep learning library for solving differential equations," *SIAM Rev.*, vol. 63, no. 1, pp. 208–228, 2021.
- [30] A. Khan and D. A. Lowther, "Physics informed neural networks for electromagnetic analysis," *IEEE Trans. Magn.* vol. 58, no. 9, pp. 1–4, Sep. 2022.
- [31] R. He, B. K. Lau, C. Oestges, K. Haneda, and B. Liu, "Guest editorial artificial intelligence in radio propagation for communications," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3934–3938, Jun. 2022.
- [32] S. Bakirtzis, J. Chen, K. Qiu, J. Zhang, and I. Wassell, "EM DeepRay: An expedient, generalizable and realistic data-driven indoor propagation model," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4140–4154, Jun. 2022.
- [33] P. Zhang, C. Yi, B. Yang, H. Wang, C. Oestges, and X. You, "Predictive modeling of millimeter-wave vegetation scattering effect using hybrid physics-based and data-driven approach," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4056–4068, Jun. 2022.
- [34] M. M. Rahman, S. Z. Gurbuz, and M. G. Amin "Physics-aware generative adversarial networks for radar-based human activity recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 3, pp. 2994–3008, Jun. 2023.
- [35] M. Zehni, and Z. Zhao, "An adversarial learning based approach for 2D unknown view tomography," *IEEE Trans. Comput. Imag.*, vol. 8, pp. 705–720, Aug. 2022.
- [36] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui, "WiFi CSI based passive human activity recognition using attention based BLSTM," *IEEE Trans. Mobile Comput.*, vol. 18, no. 11, pp. 2714–2724, Nov. 2019.
- [37] J. D. Kraus, *Antennas: From Theory to Practice*, 2nd ed. New York, NY, USA: McGraw-Hill, 1988.
- [38] V. Rampa, F. Fieramosca, S. Savazzi, and M. D'Amico, "An EM body model for Device-Free Localization with multiple antenna receivers: A first study," *IEEE-APS Topical Conf. Antennas Propag. Wireless Commun. (APWC)*, 2023, pp. 164–169.
- [39] D. J. Suroso, P. Cherntanomwong, and P. Sooraksa, "Deep generative model-based RSSI synthesis for indoor localization," in *Proc. 19th Int. Conf. Electr. Eng./Electron., Computer, Telecommun. Inf. Technol. (ECTI-CON)*, 2022, pp. 1–5.
- [40] W. Xia et al., "Millimeter wave channel modeling via generative neural networks," in *Proc. IEEE Globecom Workshops*, 2020, pp. 1–6.
- [41] "Code and Data: diffraction model generation pre-trained network and sample codes." Accessed: Apr. 28, 2023. [Online]. Available: https://github.com/labRadioVision/generative_diffraction_model
- [42] F. Fieramosca, V. Rampa, S. Savazzi, and M. D'Amico, "On the impact of the antenna radiation patterns in passive radio sensing," *IEEE Antennas Wireless Propag. Lett.*, vol. 23, no. 2, pp. 503–507, Feb. 2024.
- [43] F. Fieramosca, V. Rampa, S. Savazzi, and M. D'Amico, "Modelling of the floor effects in device-free radio localization applications," in *Proc. 17th Eur. Conf. Antennas Propag. (EuCap)*, 2023, pp. 1–5.
- [44] J. Benesty, I. Cohen, and J. Chen, "A brief overview of conventional beamforming," in *Array Beamforming With Linear Difference Equations*, vol. 20. Cham, Switzerland: Springer, 2021, pp. 13–21. [Online]. Available: https://doi.org/10.1007/978-3-030-68273-6_2
- [45] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2013, pp. 1–14. [Online] Available: <https://arxiv.org/abs/1312.6114>
- [46] H. Ham, T. J. Jun, and D. Kim, "Unbalanced GANs: Pre-training the generator of generative adversarial network using variational autoencoder," 2020, *arXiv:2002.02112*.
- [47] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Stat.*, vol. 22, no. 2, pp. 79–86, 1951.
- [48] Z. Weidong, W. Zhang, K. Liu, and J. Gu, "Learning to predict high-quality edge maps for room layout estimation," *IEEE Trans. Multimedia*, vol. 19, no. 5, pp. 935–943, May 2017.
- [49] "RSA500 series real time spectrum analyzers." Accessed: Sep. 27, 2022. [Online]. Available: <https://www.tek.com/en/products/spectrum-analyzers/rsa500>



STEFANO SVAZZI (Member, IEEE) received the M.Sc. and Ph.D. degrees (Hons.) in ICT from the Politecnico di Milano, Milano, Italy, in 2004 and 2008, respectively

He joined with Consiglio Nazionale delle Ricerche (CNR), in 2012, where he is a Senior Researcher with the Institute of Electronics, Computer and Telecommunication Engineering. He was a Visiting Researcher with Uppsala University, Uppsala, Sweden, in 2005, and the University of California at San Diego, La Jolla, CA, USA, in 2007. He has coauthored over 110 scientific publications (Scopus). His current research interests include distributed signal processing, distributed machine learning and networking aspects for the Internet of Things, radio localization, and vision technologies.

Dr. Savazzi was the recipient of the 2008 Dimitris N. Chorafas Foundation Award. He is Principal Investigator for CNR in Horizon EU projects Holden and TRUSTroke. He is also serving as an Associate Editor for *Frontiers in Communications and Networks*, *Wireless Communications*, and *Mobile Computing and Sensors*.



FEDERICA FIERAMOSCA received the M.Sc. degree in telecommunication engineering (cum laude) from Politecnico di Milano, Milano, Italy, in 2022, where she is currently pursuing the Ph.D. degree in information technology from the Department of Electronics, Information and Bioengineering.

She is currently working in the field of applied electromagnetics, focusing on electromagnetic propagation modeling for integrated sensing and communication.



SANAZ KIANOUSH received the Ph.D. degree in electronic engineering from the University of Pavia, Pavia, Italy, in 2014.

She is a Researcher with the Consiglio Nazionale delle Ricerche (CNR), the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT), Italy. In 2018, she was a Visiting Researcher with Aalto University, Espoo, Finland. Her research interests include statistical signal processing, sensor fusion in communication systems using deep learning and federated learning algorithms, design algorithms for device free, and reconfigurable intelligent surface-based localization/activity recognition in healthcare and industrial applications.

Dr. Kianoush won the IEIIT-CNR Young Research Award (PreGio) in 2018.



MICHELE D'AMICO (Senior Member, IEEE) received the M.Sc. degree from Politecnico di Milano, Milano, Italy, in 1990, and the Ph.D. degree in mathematics from the University of Essex, Colchester, U.K., in 1997.

Since 2002 he has been an Associate Professor of Applied Electromagnetics with the Department of Electronics, Information and Bioengineering, Politecnico di Milano. Since May 2013 he has been responsible for the experimental activities of the Spino d'Adda laboratory, Politecnico di

Milano, dedicated to the experimental investigation of EM waves in the troposphere. He has several patents on antennas and has authored or co-authored more than 140 papers published in international journals or conference proceedings. His research activities include antennas, electromagnetic wave propagation in the troposphere at frequencies above 10 GHz, and radar meteorology.



VITTORIO RAMPA (Senior Member, IEEE) received the M.Sc degree (Hons.) in electrical engineering from Politecnico di Milano (POLIMI), Milano, Italy, in 1984.

In 1986, he joined with the Consiglio Nazionale delle Ricerche (CNR) as a Researcher. From 1999 to 2014, he has been an Adjunct Professor with POLIMI, where he has taught courses on software defined radio algorithms and architectures, and navigation systems. From 2001 to 2022, he has been a Senior Researcher with the Institute of

Electronics, Computer, and Telecommunication Engineering (CNR-IEIIT). In 2006, he co-founded the CNR-POLIMI spin-off Wisytech Srl, where till 2014, he served as the CTO. Since 2018, he has been the member of the Steering Committee GTTS-4 of the National Cluster Fabbrica Intelligente. Since 2023, he has been a Senior Research Associate with CNR-IEIIT. His research interests include signal processing algorithms and architectures for wireless communications, wireless sensor networks for IoT applications, radio vision methods for detection, localization, and activity recognition of people, body models and machine learning tools for radio passive sensing, and federated learning algorithms for cooperative and green applications.

Open Access funding provided by “Consiglio Nazionale delle Ricerche-CARI-CARE-ITALY”
within the CRUI CARE Agreement