RIS-Based Radio Localization in Rich Scattering Environments: Harnessing Multi-Path with ANN Decoders

Philipp del Hougne

2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) | 978-1-6654-2851-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/SPAWC51858.2021.9593167

RIS-Based Radio Localization in Rich Scattering Environments: Harnessing Multi-Path with ANN Decoders

Philipp del Hougne
Univ Rennes, CNRS, IETR - UMR 6164, F-35000, Rennes, France
philipp.del-hougne@univ-rennes1.fr

***Abstract*—Radio localization is a key enabling technology for situational awareness but conventional techniques based on elaborate ray-tracing approaches naturally struggle in rich scattering environments (inside rooms, metro stations, planes, vessels, ...). Here, we discuss a completely different approach to radio localization: instead of attempting to understand rich scattering wave propagation in terms of rays, we harness the overwhelming complexity because it assigns unique wave fingerprints to each object position. We interpret wave propagation as a physical encoder of the sought-after localization information in multiplexed measurements and detail artificial neural network (ANN) architectures suitable to decode these measurements for a single or multiple, discrete or continuous, sought-after location variable(s). Capitalizing on recent physics-driven experiments, we clarify that the proposed technique is very robust to measurement noise and capable of achieving deeply sub-wavelength localization precision. The discussed technique can be implemented with multiplexing across spatial, spectral or configurational degrees of freedom, corresponding to sensor networks, broadband measurements and RIS-programmable environments, respectively. Specifically, multiplexing across a fixed *random* sequence of RIS configurations enables single-frequency localization with a single node. Finally, we propose an end-to-end vision of the technique in which programmable RIS elements take the role of physical weights in a hybrid analog-digital ANN. Thereby, relevant information for the localization task can be discriminated from irrelevant information already in the measurement process, enabling substantial latency improvements.**

***Index Terms*—Radio Localization, Reconfigurable Intelligent Surface, Rich Scattering Environment, Multi-Path, Artificial Neural Network**

I. Introduction

Localization information is a fundamental ingredient of situational awareness and underlies several emerging technologies expected to impact society at large, including the areas of mobility (autonomous vehicles and robots), health-care (ambient-assisted living [1]), human-machine interaction (gesture-based commands, virtual reality), retail (asset tracking, customer analytics), and wireless communication (beamforming) [2]. Radio technology is attractive for localization tasks because on the one hand it is independent of scene lighting, optical colors and optically opaque layers (clothes, fog, ...) while on the other hand it can often leverage existing wirless communication infrastructure, endowing it with a dual communication and sensing functionality. Conventional radio localization relies on ray-tracing, making use of information such as a radio wave's angle or time of arrival. These techniques were conceived for operation in quasi-free space but inherently struggle when confronted with the rich scattering that is typical for environments in which the above-listed applications are desired. In rich scattering environments, wave energy is not incident from one or a few well-defined directions at well-defined moments in time; instead, it is statistically spread across all possible directions and polarizations throughout the propagation environment [3]. The amount of scattering depends on the specific environment's materials and geometry, as well as on the radio wave's frequency; for typical frequencies of a few GHz, as used in WiFi, many common scenarios taking place inside rooms, metro stations, planes or vessels are confronted with strong reverberation and multipath effects.

Dealing with rich scattering environments from a ray tracing perspective, in particular distinguishing objects of interest from other strong reflectors, is highly challenging. Current proposals tackle mildly scattering environments in which it may be feasible to account for known scatterers with identified position in ray tracing [4], [5], to conduct a statistical analysis of the channels [6], [7], to leverage large-scale antenna arrays [8], [9], or to use an array of programmable reflectors (a "Reconfigurable Intelligent Surface", RIS) as known and controllable reference or anchor points [10]. "Single-bounce" assumptions [11] can be suitable simplifications of the multipath problem at higher frequencies. Some of the above-described approaches have been tested experimentally in an indoor environment using 2-GHz-bandwidth signals [4], [5], [12] or using a base-station equipped with a large amount of antennas [8].

Instead of struggling to deal with multi-path in ray-tracing frameworks, a completely different approach consists in harnessing the overwhelming scattering richness via wave fingerprinting (WFPing). The wave field in a rich scattering environment is extremely sensitive to geometrical perturbations such that the wave fields act as WFPs for different object positions. To unambiguously identify the WFPs, they must be "inked" with sufficient distinct measurements. These measurements of the wave field can multiplex the sought-after location information across spatial, spectral or configurational degrees of freedom (DoF) of the system. Spatial DoF are accessed via large-scale antenna arrays [13]–[15], spectral DoF via (ultra-) broadband measurements [16]–[20], and configurational DoF via a pre-defined series of *random* configurations

of a RIS-programmable propagation environment [21]. The latter has the advantage of using only a single node and a single frequency, resulting in a very light hardware layer and minimal spectrum allotment concerns. Moreover, unlike RIS-based localization proposals in quasi-free space [22], the use of random sequences for WFPing implies no need to model and optimize the RIS configuration. Not relying on a RIS model is of practical importance because to date no RIS models that faithfully capture all physically important details (inter-element coupling, dependence on angle of incidence, frequency and polarization, influence of nearby scatterers) is available.

Here, we discuss WFP-based radio localization in rich scattering environments from an information flow perspective in order to clarify important questions about its potential limitations (achievable localization precision, performance in dynamic environments). We interpret wave propagation in the rich scattering environment as physical encoder of the sought localization information, and discuss how this information can be decoded from the measurements. In particular, we argue that artificial neural networks (ANNs) are a good choice of digital decoder and discuss suitable architectures in various scenarios. Finally, the information flow perspective allows us to envision an end-to-end optimized localization scheme in which the programmable propagation environment plays the role of a first trainable ANN layer, holding the promise of significant latency improvements. With this unifying perspective on information flow, we bring together somewhat disjoint research threads on wave physics, metamaterials, and signal processing.

II. Information Flow

To start, we analyze how information flows through the physical and digital layers of the localization sensing pipeline. For concreteness, we consider the use of RIS-based configurational diversity in Fig. 1. Wave propagation encodes the sought after information in the multiplexed measurements, and the digital layer serves as decoder of the desired information.

Many choices of digital decoder are conceivable. When the measurements are complex-valued, a phase retrieval problem is avoided and the problem can be framed as a matrix inversion, which can be further refined via various regularization strategies. A digital decoder that works irrespective of the availability or not of phase information is a correlation-based dictionary search, sometimes also marketed as "virtual time reversal" if spectral DoF are used. A more modern strategy is to train an ANN to approximate the inverse function of the physical encoder. Refs. [23], [24] both observed that ANN decoders are significantly more robust to noise than their competitors. A further advantage, in comparison to inversion or correlation, is that during inference the forward pass through the ANN is likely faster.

III. Neural Network Architecture

In this section, we discuss the choice of layer architecture, activation functions and cost functions in various scenarios.

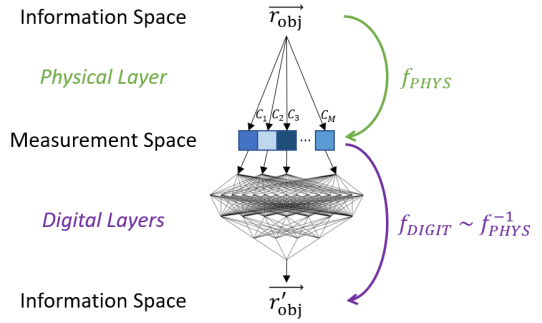

Fig. 1. Information Flow in WFP-Based Radio Localization using RIS-Based Configurational Diversity. Wave propagation in the RIS-parametrized rich scattering environment encodes the information about the object location, $\vec{r}_{obj}$, in a measurement vector $\mathcal{M} = f_{PHYS}(\vec{r}_{obj}, \{C_i\})$. The ith measurement is a single value (real or complex, depending on whether phase information is acquired) corresponding to the ith fixed random RIS configuration C_i. $\mathcal{M}$ concatenates the M measurements obtained with the fixed series of M random RIS configurations, $\{C_i\}$. The digital layer then attempts to decode $\mathcal{M}$ by approximating the inverse function of f_{PHYS} to obtain an estimate of the object position: $\vec{r}'_{obj} = f_{DIGIT}(\mathcal{M}) \approx \vec{r}_{obj}$.

A. Fully Connected or Convolutional Layers?

Convolutional architectures are currently very popular in deep learning, including in WFP-based localization in rich scattering environments [15]. Their popularity originates from their success in computer vision, where convolutional layers excel at extracting local features from data [25]. Computer vision deals with optical images in which there is a one-to-one mapping between scene and image. In sharp contrast, there is *no* one-to-one mapping between scene and multiplexed measurements in WFP-based localization data. Information from all parts of the scene is scrambled via wave propagation in the rich scattering environment, and each scene component contributes to each measurement. Therefore, local scene features are very likely encoded in long-range correlations in the multiplexed data and hence fully connected layers capable of identifying global features in data from multiplexed measurements appear preferable. Indeed, recent work on the recovery of scrambled optical images after propagation through a multimode fiber confirmed this hypothesis [26], and successful WFP-based radio localization with a simple ANN consisting only of two fully connected layers was reported [3], [23].

B. Choice of Activation and Cost Functions

Depending on whether the sought-after location variable is discretized or continuous, the problem falls into the categories of classification or regression, respectively, with important consequences for the choice of activation and cost function.

Initial works [3], [21], [23] discretized the location variable via a grid of predefined positions on which the object was allowed to be located. For a single object, this is a standard classification problem that can be tackled with ReLU activation functions, except for the last layer on which a SoftMax activation should be applied, in combination with a cross-entropy cost function [23]. If more than one object is to

be localized, the classes are independent and not mutually exclusive, resulting in a multi-class multi-label classification problem. To that end, the activation function of the last layer should be a Sigmoid rather than a SoftMax function, again in combination with a cross-entropy cost function [3].[1]

More recently, in order to probe the achievable localization precision, Ref. [24] treated the location variable as continuous. For such a regression problem, it was heuristically observed that Sigmoid activation functions should be used after the fully connected layers; no activation function should be applied to the output layer. The cost function should be the mean error (or root-mean-square error), i.e. not a cross-entropy based cost function.

IV. Potential Limitations

In this section, we address two potential limitations of WPF-based radio localization.

A. Achievable Precision

It is often assumed that the localization precision can be at best half the operating wavelength because of the diffraction limit. However, the diffraction limit applies to energy focusing whereas the WFP-based localization operates with multiplexed measurements without any energy focusing. Moreover, the crucial roles of *a priori* knowledge and signal-to-noise (SNR) ratio are ignored by the diffraction-limit argument. In fact, reverberation in a rich scattering environment encodes deeply sub-wavelength information in the multiplexed measurements, and the longer the waves dwell in the environment, the more efficiently they encode such information [24]. Intuitively, this link between dwell time and localization precision is easily understood: the probability that a deeply sub-wavelength detail impacts the evolution of a wave is higher if the wave reverberates longer in the interaction domain. Specifically, this means that the localization precision in rich scattering media can vastly exceed that achievable in quasi-free space, underlining the potential of harnessing rich scattering. The challenge lies then only in extracting the sought-after information from the multiplexed measurements, and an ANN decoder is a suitable way of doing so [24]. There is hence *no* fundamental wavelength-induced bound on the achievable localization precision. Instead, in practice, the latter depends on physical parameters (dwell time and SNR), the available *a priori* knowledge, and the choice of digital decoder.

B. Operation in Dynamic Environments

A common objection to WFP-based localization are doubts about its ability to perform well in dynamic environments. Indeed, WFPing is based on the sensitivity of chaotic wave fields to the object to be localized, but the sensitivity of course equally applies to parasitic perturbations. In principle, the perturbation can be interpreted and treated as an effective source of noise. The worse the effective SNR, the less useful information is obtained per measurement. However, this can be counterbalanced by taking more measurements with distinct measurement modes. Therefore, even if the parasitic perturbations are stronger than those of the object(s) of interest, in principle that does not automatically preclude the successful implementation of WFP-based localization [23]. However, the stronger the parasitic perturber is, the more important it becomes to thoroughly characterize it such that the ANN can faithfully extract statistically relevant information from perturbed measurements. Specifically, this means that the training data set should capture a representative set of realizations of the perturbations that can be expected during on-line inference.

V. End-to-End Vision

Reducing latency is a crucial goal for many localization applications. In terms of information flow, it is obvious that the above-discussed WFP-based schemes do not discriminate between relevant and irrelevant information during the physical encoding in multiplexed measurements. If only information relevant to the localization task was included in the measurements, fewer measurements would be needed to acquire the same amount of useful information, and less data would have to be processed, resulting in a double advantage in terms of latency.

If RIS-enabled configurational DoF are used for WFPing, it is possible to optimize the measurement modes. A *random* sequences of RIS configurations, as discussed thus far, yields pseudo-orthogonal measurement modes. This means that not only do the measurement modes not filter out irrelevant information but subsequent measurements in fact acquire to some extent redundant information due to the finite overlap of the measurement modes. This redundancy can be removed by optimizing the RIS sequence such that the measurement modes are truly orthogonal [27]. Such a sequence can be identified iteratively in the experiment, without needing to model the interaction of waves with the RIS. However, truly orthogonal measurement modes still fail to separate relevant from irrelevant information.

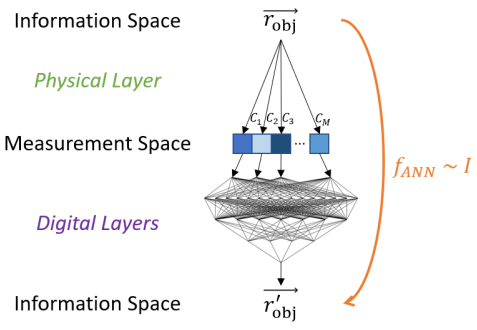

Fig. 2. "Learned Sensing" Architecture for RIS-Based Radio Localization. Learned Sensing is based on a hybrid analog-digital ANN, consisting of a physical layer on which the RIS elements' configurations are the trainable physical weights, followed by several traditional digital layers, with conventional digital weights. I denotes the identity function.

[1]In both cases, it is important in practice to make sure the last activation function is not accidentally applied twice; for instance, in Tensorflow, cross-entropy loss functions such as softmax_cross_entropy_with_logits compute internally the softmax of their inputs.

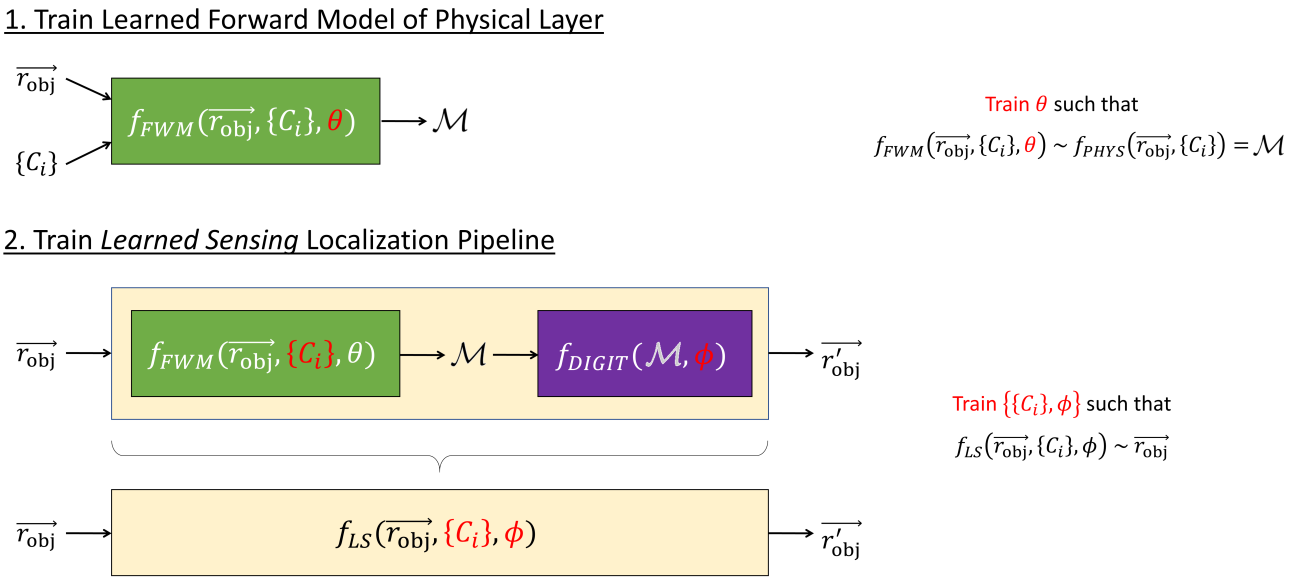

Fig. 3. Implementing a Realistic *Learned Sensing* Localization Pipeline. First, a learned forward model is established to approximate the physical layer. The physical layer performs the function $f_{PHYS}(\vec{r}_{obj}, \{C_i\}) = \mathcal{M}$, where $\vec{r}_{obj}$ is the object position, $\{C_i\}$ is a set of M RIS configurations, and $\mathcal{M}$ is the $1 \times M$ measurement vector (one measurement per RIS configuration). $\mathcal{M}$ can be real or complex valued, depending on whether phase information is measured or not. A neural network with trainable weights θ is now trained by optimizing θ such that $f_{FWM}(\vec{r}_{obj}, \{C_i\}, \theta) \approx f_{PHYS}(\vec{r}_{obj}, \{C_i\}) = \mathcal{M}$. Second, this learned forward model f_{FWM} is integrated into the learned sensing pipeline. Now, θ remains fixed to the optimized values from the first step. The learned sensing network has a set of trainable physical weights, namely the RIS configurations $\{C_i\}$, and a set of trainable digital weights, namely the weights of the digital decoding network ϕ. Overall, the learned sensing network is supposed to perform a function $f_{LS}(\vec{r}_{obj}, \{C_i\}, \phi) \approx \vec{r}_{obj}$. To that end, its trainable physical and digital weights are *jointly* optimized.

Following the recently introduced "learned sensing" paradigm in electromagnetic imaging and gesture recognition [28], to endow the physical layer with task-awareness, we propose to interpret the RIS-programmable rich scattering propagation environment as a trainable physical layer of a hybrid analog-digital neural network, followed by the digital layers discussed previously. The traditional WFP architecture from Fig. 1 also seeks to ultimately approximate an identity function by using $f_{DIGIT} \approx f_{PHYS}^{-1}$, however, therein f_{PHYS} is fixed and not trainable (a fixed sequence of random RIS configurations is used). In Fig. 1, only the digital layer can be optimized.

The architecture in Fig. 2 is reminiscent of an auto-encoder (see Ch. 14 in Ref. [29]) in that it seeks to map the information space to itself by approximating an identity function. It lacks, however, the common architectural symmetry of auto-encoders. Most importantly, the encoding step of the proposed "learned sensing" pipeline happens in the physical rather than digital world – a fundamental difference from conventional all-digital auto-encoders. Conventional auto-encoders are a signal processing tool aimed at dimensionality reduction of measured data, whereas here we aim at *optimal joint measuring and processing* of data given a variable of interest. Auto-encoders can also be trained to yield representations of corrupted input data, known as denoising auto-encoders [30]; this functionality is inherent in our proposal because the physical layer involves measurements that are inevitably affected by measurement noise.

In practice, an important challenge will be to obtain a forward model of the physical layer such that it can be integrated into the hybrid analog-digital end-to-end pipeline. Analytical forward models seem difficult to realize for localization in rich scattering environments for two reasons: (i) no models of RIS that faithfully capture the physics are available, and (ii) modelling a rich scattering environment via ray tracing is prohibitively expensive. Instead, it appears best to resort to a learned forward model [31]. To that end, as illustrated in Fig. 3, in a first step, an auxiliary neural network is trained such that the function f_{FWM} that it performs approximates the physical layer; specifically, f_{FWM} predicts the measurements $\mathcal{M}$ as a function of a set of RIS configurations $\{C_i\}$ and the object position $\vec{r}_{obj}$. To train this neural network, a representative sample of measurements obtained for different object positions and random RIS configurations is required. In a second step, this learned forward model is integrated into the "learned sensing" pipeline, where $\{C_i\}$ are the physical trainable weights and ϕ are the digital trainable weights. The learned sensing network is then trained by *jointly* optimizing $\{C_i\}$ and ϕ. Thereby, the RIS configurations learn to take into account the ultimate localization task, making it possible to discriminate between relevant and irrelevant information already in the measurement process. Since each measurement will then contain more information relevant to the localization task, fewer measurements will be needed to successfully perform the localization task. The shorter measurement process, combined with a lower data processing burden, will enable significant latency improvements. Moreover, the reduced number of necessary measurements will favorably impact other metrics

such as power consumption and radiation exposure. We are confident that these latency trends which were evidenced in recent "learned sensing" studies of electromagnetic imaging in free space [28], [31] will also apply to localization in rich-scattering environments: on a very abstract level, localizing an object can be interpreted as imaging of an extremely sparse scene; the environment's rich-scattering nature can be interpreted as part of the "scene illumination hardware".

VI. Conclusion

To summarize, we discussed WFP-based radio localization from an information flow perspective, with a focus on the use of configurational DoF offered by a RIS-equipped propagation environment. We considered technical details of ANN-based digital decoders (layer architecture, choice of activation and cost functions) in various scenarios. Moreover, we clarified the absence of wavelength-induced bounds on the achievable localization precision, as well as that operation in dynamic environments is feasible. Then, we introduced an end-to-end vision for a hybrid analog-digital localization pipeline which treats the RIS-programmable environment as a first trainable physical layer. On the one hand, thereby wave propagation acts as analog processor. On the other hand, the measurement process pre-selects information relevant to the localization task which avoids the acquisition of irrelevant information. We expect the end-to-end pipeline to enable significant latency improvements. Looking forward, this vision can be further refined within the realm of reinforcement learning to update the sequence of RIS configurations in real time [32].

References

[1] K. Witrisal, P. Meissner, E. Leitinger, Y. Shen, C. Gustafson, F. Tufvesson, K. Haneda, D. Dardari, A. F. Molisch, A. Conti *et al.*, "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, 2016.

[2] A. Bourdoux, A. N. Barreto, B. van Liempd, C. de Lima, D. Dardari, D. Belot, E.-S. Lohan, G. Seco-Granados, H. Sarieddeen, H. Wymeersch *et al.*, "6G white paper on localization and sensing," *arXiv:2006.01779*, 2020.

[3] G. C. Alexandropoulos, N. Shlezinger, and P. del Hougne, "Reconfigurable intelligent surfaces for rich scattering wireless communications: Recent experiments, challenges, and opportunities," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 28–34, 2021.

[4] P. Meissner, E. Leitinger, and K. Witrisal, "Uwb for robust indoor tracking: Weighting of multipath components for efficient estimation," *IEEE Wirel. Commun. Lett.*, vol. 3, no. 5, pp. 501–504, 2014.

[5] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 12, pp. 5613–5629, Dec. 2019.

[6] 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.

[7] R. Mendrzik, F. Meyer, G. Bauch, and M. Z. Win, "Enabling situational awareness in millimeter wave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1196–1211, Sep. 2019.

[8] X. Li, E. Leitinger, M. Oskarsson, K. Åström, and F. Tufvesson, "Massive MIMO-based localization and mapping exploiting phase information of multipath components," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4254–4267, Sep. 2019.

[9] F. Wen, J. Kulmer, K. Witrisal, and H. Wymeersch, "5G positioning and mapping with diffuse multipath," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 2, pp. 1164–1174, 2020.

[10] H. Wymeersch, J. He, B. Denis, A. Clemente, and M. Juntti, "Radio localization and mapping with reconfigurable intelligent surfaces," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 52–61, Dec. 2020.

[11] F. Wen and H. Wymeersch, "5g synchronization, positioning, and mapping from diffuse multipath," *IEEE Wirel. Commun. Lett.*, vol. 10, no. 1, pp. 43–47, 2020.

[12] E. Leitinger, S. Grebien, and K. Witrisal, "Multipath-based slam exploiting aoa and amplitude information," in *IEEE ICC*. IEEE, 2019, pp. 1–7.

[13] M. Vari and D. Cassioli, "mmWaves RSSI Indoor Network Localization," in *Proc. IEEE ICC*. IEEE, 2014, pp. 127–132.

[14] V. Savic and E. G. Larsson, "Fingerprinting-based positioning in distributed massive mimo systems," in *Proc. IEEE VTC2015-Fall*. IEEE, 2015, pp. 1–5.

[15] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive mimo fingerprint-based positioning," in *Proc. IEEE PIMRC*. IEEE, 2017, pp. 1–6.

[16] Y. Jin, N. O'Donoughue, and J. M. Moura, "Position location by time reversal in communication networks," in *Proc. IEEE ICASP*. IEEE, 2008, pp. 3001–3004.

[17] C. Steiner and A. Wittneben, "Efficient training phase for ultrawideband-based location fingerprinting systems," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6021–6032, 2011.

[18] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "You are facing the Mona Lisa: Spot localization using PHY layer information," in *ACM Proc. MobiSys 2012*, 2012, pp. 183–196.

[19] Z.-H. Wu, Y. Han, Y. Chen, and K. R. Liu, "A time-reversal paradigm for indoor positioning system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1331–1339, 2015.

[20] C. Chen, Y. Chen, Y. Han, H.-Q. Lai, and K. R. Liu, "Achieving centimeter-accuracy indoor localization on wifi platforms: A frequency hopping approach," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 111–121, 2016.

[21] P. del Hougne, M. F. Imani, M. Fink, D. R. Smith, and G. Lerosey, "Precise localization of multiple noncooperative objects in a disordered cavity by wave front shaping," *Phys. Rev. Lett.*, vol. 121, no. 6, p. 063901, 2018.

[22] Z. Abu-Shaban, K. Keykhosravi, M. F. Keskin, G. C. Alexandropoulos, G. Seco-Granados, and H. Wymeersch, "Near-field localization with a reconfigurable intelligent surface acting as lens," in *Proc. IEEE ICC*, Montreal, Canada, Jun. 2021.

[23] P. del Hougne, "Robust position sensing with wave fingerprints in dynamic complex propagation environments," *Phys. Rev. Research*, vol. 2, p. 043224, Nov 2020.

[24] M. del Hougne, S. Gigan, and P. del Hougne, "Deeply sub-wavelength localization with reverberation-coded-aperture," *Phys. Rev. Lett. (in press), arXiv:2102.05642*, 2021.

[25] Y. Zheng, J. Huang, T. Chen, Y. Ou, and W. Zhou, "Processing global and local features in convolutional neural network (cnn) and primate visual systems," in *Mobile Multimedia/Image Processing, Security, and Applications 2018*, vol. 10668. SPIE, 2018, p. 1066809.

[26] C. Zhu, E. A. Chan, Y. Wang, W. Peng, R. Guo, B. Zhang, C. Soci, and Y. Chong, "Image reconstruction through a multimode fiber with a simple neural network architecture," *Sci. Rep.*, vol. 11, p. 896, 2021.

[27] P. del Hougne, M. Davy, and U. Kuhl, "Optimal multiplexing of spatially encoded information across custom-tailored configurations of a metasurface-tunable chaotic cavity," *Phys. Rev. Applied*, vol. 13, no. 4, p. 041004, 2020.

[28] P. del Hougne, M. F. Imani, A. V. Diebold, R. Horstmeyer, and D. R. Smith, "Learned integrated sensing pipeline: Reconfigurable metasurface transceivers as trainable physical layer in an artificial neural network," *Adv. Sci.*, vol. 7, no. 3, p. 1901913, 2020.

[29] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[30] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. ICML*, 2008, pp. 1096–1103.

[31] H.-Y. Li, H.-T. Zhao, M.-L. Wei, H.-X. Ruan, Y. Shuang, T. J. Cui, P. del Hougne, and L. Li, "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns*, vol. 1, no. 1, p. 100006, 2020.

[32] V. Mnih, N. Heess, A. Graves, and K. Kavukcuoglu, "Recurrent models of visual attention," in *Proc. NIPS*, vol. 27. Curran Associates, Inc., 2014, pp. 2204–2212.