



Location-To-Channel Mapping Using Model-Based Learning

¹ G. Sudheer Kumar, ² Y. Chandana,

¹ Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.

² MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

ABSTRACT

Accurate channel estimate is crucial for modern communication systems to ensure dependable and efficient delivery of information. A neural network may be used to convert the user's geographic coordinates into the coefficients of the communication channel, as this information is highly dependent on the user's position. These latter, however, are changing location-dependently at a rate on the order of wavelengths. Because classical neural networks have a spectrum bias that makes them better at learning low-frequency functions, learning such a map is quite challenging. To get over this issue, this research introduces a cheap model-based network that divides the target mapping function into low-frequency and high-frequency components. The result is a hypernetwork design in which the neural network learns sparse coefficients with low frequencies from a lexicon of components with high frequencies. On realistic synthetic data, the suggested neural network outperforms conventional methods, according to the simulation findings. Index Terms—Information Modeling, Channel Estimation, Spectral Bias, and Implicit Neural Representations

1. INTRODUCTION

For many years, people have turned to classical signal processing techniques when they needed help with data processing issues. These approaches rely on models, and as all models have their flaws, they might be quite biased. Nevertheless, the minimal complexity of such approaches is a plus. A new paradigm has emerged in machine learning: instead of utilizing models, one might use data-learning generic neural networks. While these approaches' inherent flexibility makes them bias-free, training them may be computationally and sample-complexly intensive. By adapting signal processing models for use in machine learning's initialization, structure, and training, model-based machine learning [1] aims to combine the advantages of both low complexity and low bias. Since several models have been created to characterize communication systems, the discipline of communication engineering is well suited to employ model-based (MB) machine learning. Gaining precise channel estimate is possible, in particular, by using propagation channel models [2, 3, 4, 5]. There are a lot of potential uses for learning the location-to-channel mapping, as propagation channels are strongly tied to the user's location. The trained neural network's weights and biases would store the radio environment surrounding a base station (BS), making location-based channel prediction practical and allowing for radio-environments compression. Beamformer prediction, jamming detection, resource allocation, and secure communication techniques are among other possible uses. Unfortunately, learning this mapping is very complicated since the spatial dependency of the underlying model changes with the order of the wavelength. The fact that traditional neural networks have a bias toward learning low frequency functions (spectral bias) was shown in [6]. Helpful hints. In this research, a model-based neural architecture is derived from a channel model based on physics with the goal of learning the location-to-channel mapping under supervision. The suggested neural network has a spectrum separation step that divides the target mapping function's low-frequency and high-frequency material, much as designs from the implicit neural representation (INR) literature [7, 8, 9]. The spectral bias problem may be circumvented in this way. In contrast to

INR designs, the suggested architecture makes use of a second hypernetwork to sparsely learn the activation coefficients of the high-frequency components. In comparison to traditional and INR neural networks, the suggested design is tested using synthetic data that is both realistic and artificial. Not only does it significantly reduce the amount of parameters to learn, but it also produces a significant increase in reconstruction results. Tasks that are similar. For picture reconstruction[8,10,11] and 3D scene reconstruction from 2D images[9,12,13], the INR community has delved further into learning mappings with neural networks. The development of architectures capable of learning details at high frequencies has received a lot of attention [7, 8, 9]. Additionally, channel estimation using machine learning algorithms has garnered a lot of attention recently [14, 15, 16]. Learned location or pseudolocation for beamformer mapping has been the subject of prior research [17, 18]. But no prior study that the authors are aware of has concentrated on learning the location-to-channel mapping.

2. PROBLEM FORMULATION

This study examines a SISO monocarrier situation with a single base station (BS) situated at $\mathbf{x}_1 = (x_1, y_1)$ and a single user equipment (UE) situated at an arbitrarily defined location $\mathbf{x} = (x, y)$. For the sake of clarity and simplicity of explanation, only 2D locations are taken into account here. It is possible, however, to easily apply the suggested approach to the 3D scenario. With L_p virtual propagation pathways in mind, the BS-UE link's channel coefficient at a certain frequency may be represented as: in where λ is the wavelength and γ_l and d_l are the complex attenuation and propagation distance of the l th route, respectively. By modeling the propagation interactions using the image source theory and increasing the η_l term, the following results are obtained [19, Chapter 1, pp.47-49]:

$$h(\mathbf{x}) = \sum_{l=1}^{L_p} \frac{\alpha_l e^{j\beta_l}}{\|\mathbf{x} - \mathbf{x}_l\|_2} e^{-j\frac{2\pi}{\lambda} \|\mathbf{x} - \mathbf{x}_l\|_2}, \quad (2)$$

$\mathbf{x}_l \in \mathbb{R}^2$ is the location of the picture source linked to the l th route, where $\forall l > 1$. When thinking about a Line of Sight (LoS) route, $\alpha_1 = 1$ and $\beta_1 = 0$, which stand for the small-scale attenuation and phase shift of the l th path, respectively. The large scale fading of the l th route is represented by the $1/\|\mathbf{x} - \mathbf{x}_l\|_2$ attenuation. The study's objective is to calibrate

$$f_{\theta} : \mathbf{x} \rightarrow h(\mathbf{x}), \quad (3)$$

assigns the value of $h(\mathbf{x})$ to the position \mathbf{x} . The high frequency spatial dependency of the propagation model under consideration is shown by the exponential argument in Eq. (2). The wavelength decreases as the carrier frequency increases; at the typical frequencies used in communication systems (below 6 GHz), λ is less than a few millimeters. Learning the location-to-channel mapping is therefore particularly challenging since even a little change to the location \mathbf{x} under consideration causes a large change to its channel coefficient $h(\mathbf{x})$.

3. PROPOSED METHOD

Deep neural networks' status as universal function approximators is well-established [20, 21]. Classical neural networks, on the other hand, exhibit spectral bias, which makes them unsuitable for learning functions with fast variations in frequency [6, 22]. The authors of this work suggest using the propagation model as a starting point for developing a model-based machine learning paradigm-based neural architecture for location-to-channel mapping learning [1]. Approximation of a local planar curve. Equation (2)'s channel model depicts waves reflecting off of different barriers as they travel spherically from the BS antenna. Just to review, the l th picture source is located at $\mathbf{x}_l \in \mathbb{R}^2$. Planar wavefronts may be used as local approximations to those spherical ones around any given reference position \mathbf{x}_r in \mathbb{R}^2 . To demonstrate this, one may use a Taylor expansion [23]. At $\mathbf{x} = \mathbf{x}_r$, the function $\xi(\mathbf{x})$ is differentiable, for every $\mathbf{x} - \mathbf{x}_l$. The result of plugging \mathbf{x}_r into the first-order Taylor expansion of $\xi(\mathbf{x})$ is:

$$\begin{aligned} \xi(\mathbf{x}) &\simeq \xi(\mathbf{x}_r) + \nabla \xi(\mathbf{x})|_{\mathbf{x}_r} \cdot (\mathbf{x} - \mathbf{x}_r) \\ &= \|\mathbf{x}_r - \mathbf{x}_l\|_2 + \mathbf{u}_{(\mathbf{x}_r - \mathbf{x}_l)} \cdot (\mathbf{x} - \mathbf{x}_r), \end{aligned} \quad (4)$$

where the unit norm vector in the direction of $(\mathbf{x}_r - \mathbf{x}_l)$ is denoted as $\mathbf{u}_{(\mathbf{x}_r - \mathbf{x}_l)}$. The result, for \mathbf{x} near \mathbf{x}_r , of injecting Equation (4) into Equation (2) is:

$$h(\mathbf{x}) \simeq \sum_{l=1}^{L_p} \alpha_l e^{j\beta_l} h_l(\mathbf{x}_r) \frac{e^{-j\frac{2\pi}{\lambda} \mathbf{u}(\mathbf{x}_r - \mathbf{x}_l) \cdot (\mathbf{x} - \mathbf{x}_r)}}{1 + \frac{\mathbf{u}(\mathbf{x}_r - \mathbf{x}_l) \cdot (\mathbf{x} - \mathbf{x}_r)}{\|\mathbf{x}_r - \mathbf{x}_l\|_2}}, \quad (5)$$

where $h_l(\mathbf{x}_r)$ is the l th source's channel coefficient at the reference position \mathbf{x}_r . Keep in mind that when \mathbf{x}_r approaches \mathbf{x} , the interpolation term (the fraction) approaches 1. In addition, the equation of a planar wave, which is translating along the vector $(\mathbf{x}_r - \mathbf{x}_l)$, is the numerator of the interpolation term. So, on a small scale, Eq. (2) seems like a linear combination of flat wavefronts. Many array processing methods rely on this planar wave approximation of spherical waves, which is the basis for the famous steering vectors model or spatial signature [24, Chapter 7]. To approximate the channel coefficient $h(\mathbf{x})$ as a linear combination of planar wavefronts, one may use Eq. (5) for every point \mathbf{x} in R_2 . By rearranging the variables, we may rewrite Eq. (5) with the high-frequency components (coefficients) and the low-frequency components (planar wavefronts, which change on a wavelength scale) separated:

$$h(\mathbf{x}) \simeq \sum_{l=1}^{L_p} \underbrace{\frac{\alpha_l e^{j\beta_l} h_l(\mathbf{x}_r)}{1 + \frac{\mathbf{u}(\mathbf{x}_r - \mathbf{x}_l) \cdot (\mathbf{x} - \mathbf{x}_r)}{\|\mathbf{x}_r - \mathbf{x}_l\|_2}}}_{\text{Slowly varying}} \underbrace{e^{j\mathbf{k}_{r,l} \cdot \mathbf{x}_r}}_{\text{Fastly varying}} e^{-j\mathbf{k}_{r,l} \cdot \mathbf{x}}, \quad (6)$$

where the angular wave vector in the direction of $(\mathbf{x}_r - \mathbf{x}_l)$ is denoted by $\mathbf{k}_{r,l} = 2\pi/\lambda \mathbf{u}(\mathbf{x}_r - \mathbf{x}_l)$. It has already been shown that this linear approximation is limited to a small neighborhood around \mathbf{x}_r . We may tile the space with N hexagons that are encircled by circles with a radius and a center $\mathbf{x}_{r,i}$. Then, for every \mathbf{x} in H_i , the absolute value of the difference between $h(\mathbf{x})$ and $h'(\mathbf{x})$ must be less than or equal to ϵ . Here, H_i is the set of locations inside the hexagon with the given center $\mathbf{x}_{r,i}$, and $h'(\mathbf{x})$ is the Taylor-approximated channel, which is the right-hand term in Eq. (6). The spatial validity length of the Taylor approximation is strongly connected to this radius, as shown in Eq. (4). Once within each of these hexagons, Eq. (5) demonstrates that in order to calculate $h'(\mathbf{x})$, just L_p planar wavefronts are necessary. Consequently, for each $\mathbf{x} \in R_2$, the following may be done using a dictionary $\Psi(\mathbf{x}) = \{\psi_i(\mathbf{x})\}_D$ of size $D < L_p N$, which contains carefully selected planar wavefronts and an activation vector $\mathbf{w}(\mathbf{x}) \in \mathbb{C}^D$:

$$h(\mathbf{x}) \simeq \sum_{i=1}^D w_i(\mathbf{x}) \psi_i(\mathbf{x}), \quad (7)$$

$$\text{with } \|\mathbf{w}(\mathbf{x})\|_0 = L_p, \forall \mathbf{x},$$

where the dictionary functions are constructed as

$$\psi_i(\mathbf{x}) = e^{-j\mathbf{k}_i \cdot \mathbf{x}}. \quad (8)$$

The spatial frequencies \mathbf{k}_i , which are similar to the angular wave vector but from a spectral viewpoint, determine the orientations of the planar wavefronts. For any i in J_1, DK , the sum of the absolute values of \mathbf{k}_i from 1 to 2 equals $2\pi/\lambda$, as can be shown from Eq. (5), which represents the angular wave vector. The spatial frequencies used to approximate the channel in Eq. (6) all include a one-dimensional manifold: the circle with a radius of $2\pi/\lambda$, as a result of that norm restriction. So, without the curse of dimensionality, it is feasible to construct a generic dictionary by sampling the circle while maintaining a close approximation. Also, regular neural network topologies should have no trouble learning the activation coefficients $w_i(\mathbf{x})$ because of how low frequency they are. Neuronal architecture based on models. In practice, one may use Eq. (7) to create the model-based neural network shown in Figure 1. This network generates low frequency coefficients from the position in two parallel branches and a dictionary with high frequency planar wavefronts. Hypernetwork [25, 26] of parameter set ϕ learns the low frequency coefficients in $\mathbf{w}_\phi(\mathbf{x}) \in \mathbb{C}^D$ in greater detail. Because the training loss becomes non-differentiable when the ℓ_0 requirement from Eq. (7) is integrated, one option is to transfer this sparsity constraint to the hypernetwork architecture. Figure 1 shows how this is accomplished using the softmaxC nonlinearity, which reduces the impact of non-significant coefficients and promotes sparsity in the activation vector $\mathbf{w}_\phi(\mathbf{x})$. The Fourier feature (FF) layer, shown in Figure 1, is used to build the dictionary $\psi_\phi(\mathbf{x})$ and is defined as follows:

$$\text{FF}_\varphi : \mathbf{x} \rightarrow [e^{-jk_1 \cdot \mathbf{x}}, \dots, e^{-jk_D \cdot \mathbf{x}}] \quad (9)$$

Keep in mind that this layer of embedding is only the nuanced version of the cos/sin embedding in RFF [7, 27]. According to the previous study, spatial frequencies in the dictionary may be learnt; however, in this work, they are maintained constant by randomly selecting a $2\pi \lambda$ circle.

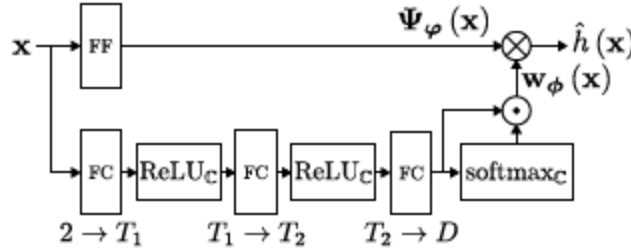


Fig. 1: Proposed model-based neural network architecture.

4. EXPERIMENTS

With $f = 3.5\text{GHz}$ ($\lambda \approx 8.5\text{cm}$), the suggested architecture needs to be tested on synthetic data in order to assess its performance. Create a dataset. To begin, a square space of 10 by 10 meters is created for the scenario. After that, we think about static image sources and a loss-of-service route to establish where the images came from. The image sources are given constant attenuation coefficients that are selected in a manner that allows them to reflect a significant portion of the incoming wavefronts. Typically, the range of these coefficients is $0.6 < a_i < 1$, with $i \neq 1$. The phase shift values are distributed evenly between 0 and 2π . The scene region is evenly sampled for the training dataset. For the test dataset, a uniform location grid is created with $\lambda/4$ spacing in both directions, yielding around 213k sites at the chosen frequency. Lastly, Eq. (2) is used to produce the training and test datasets. Analytics and models. We evaluate the suggested design in comparison to a traditional MLP and two RFF networks that drew inspiration from INR designs. You can see them in Figure 2.

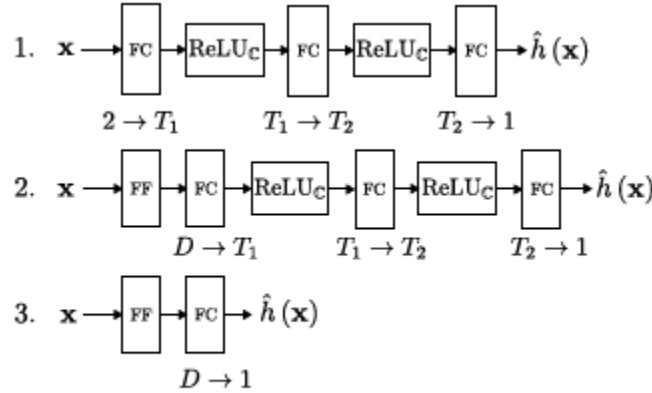


Fig. 2: 1. MLP; 2. RFF; 3. RFF lin.

$T_1 = 4096$, $T_2 = 2048$, and $D = 2000$ for all baselines. The second and third layers of the FF network use the same uniformly sampled spatial frequencies across the $2\pi \lambda$ circle as the ones utilized in the suggested model-based network. The hyperparameters for the model-based network are set to 2000, with $T_1 = 256$ and $T_2 = 128$. The NMSE, where h represents the channel coefficients and \hat{h} represents the estimated channel coefficients, is the evaluation metric. It is expressed as $10 \log_{10}(\|h - \hat{h}\|_2^2 / \|h\|_2^2)$, and it is measured in decibels across the test dataset. The ℓ_2 loss is used for training all networks in the following manner:

$$\mathcal{L} = \mathbb{E} \left[\|f_\theta(\mathbf{x}) - h(\mathbf{x})\|_2^2 \right], \mathbf{x} \in \mathcal{D} \subset \mathbb{R}^2, \quad (10)$$

this scene's training location dataset is denoted by \mathcal{D} . The following is the whole training dataset: $\{x_i, h(x_i)\}$. With $N_d = 1$. Reconstruction across a given zone. One thousand training sites are represented by a density of $100 \text{ loci/m}^2 \approx 0.7 \text{ loci}/\lambda^2$. We take into account $L_p = 6$ pathways of propagation.

	MLP	RFF	RFF lin.	MB
Params.	16.8M	33.1M	4k	0.5M
NMSE _(dB)	0.16	-3.30	-3.04	-20.60

Table 1: NMSE over the test grid.

Table 1 shows that the suggested model-based architecture has many less parameters than the MLP, but it still manages to beat all of the baselines, including designs seen in the INR literature. Also, the other networks' high NMSE values indicate that they are unable to learn the mapping, therefore it's clear that the suggested model-based network is the only one that can. The hypernetwork is the key component of the suggested design, it should be noted. It is possible to see the RFF lin. network output by using model-based initialization of the spatial frequencies.

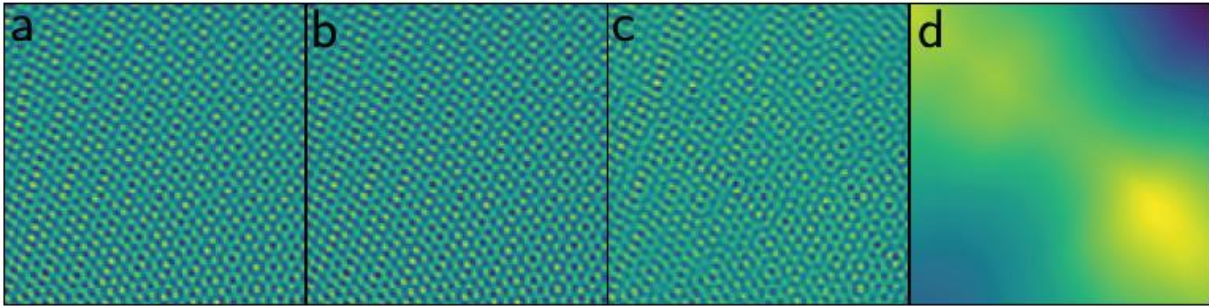


Fig. 3: Reconstruction performances (*real part*) over a small zone of the scene area (2.5m by 2.5m), a: Ground truth, b: MB, c: RFF, d: MLP

while combining planar wavefronts in a linear fashion. On the other hand, when it comes to memorizing maps, this network falls short. Even with the over-parametrized RFF, this is the case, as its output is a non-linear function of the planar wavefronts. Figure 3 shows that while this network does a decent job of reconstructing high frequency material, it is not yet flawless. The network's high spatial frequencies are a result of the FF layer initialization, thus this makes sense. However, the model analysis is not properly used. Figure 3 also shows that a model-agnostic MLP is unable to understand the mapping's high frequency material. Averaged throughout 100 trainings, reconstruction across a specified zone for various L_p and $loc.$ densities. This experiment allows for the simulation of numerous scenes by sampling fresh, randomly-placed picture source locations for each training. Additionally, every network is retrained from the beginning for every training. The model-based network surpasses the baselines for all route and location density configurations, as shown in Figure 4. For setups with a low location density, the model-based method also shows a failure scenario. This is because the training dataset in that regime does not include any places that are geographically near to each other, which causes the spatial content to change so quickly that it cannot be learned. A location density of $4 \text{ locs.}/\lambda^2$ is provided by the spatial Shannon-Nyquist criteria for flawless reconstruction... In sub-Shannon-Nyquist location density, the suggested MB network outperforms standard signal processing approaches and provides almost flawless reconstruction. Keep in mind that the wavelength decreases with increasing frequency, which means you'll need more training sites to maintain the same location density in $\text{locs.}/\lambda^2$. The number of possible propagation pathways decreases with increasing frequency, which might make learning location-to-channel mappings simpler. Using raytracing, reconstruct a single zone. Channels in a $10 \times 10 \text{ m}$ square scene area are created in this experiment in Paris, France, using the Sionna [28] ray-tracing module and the Etoile scenario. In contrast to the preceding studies, the scene area is used to randomly choose 15k training sites, with a density of $150 \text{ locs.}/\text{m}^2 \approx 1.1 \text{ locs.}/\lambda^2$. In both directions, the test spots are spaced $\lambda/4$ apart on a uniform grid. There can be no more than eleven propagation pathways within the picture, and no more than three successive reflections may be along any one path. This does not take diffraction and scattering into account. Table 2 shows that the MB network achieves better results than all baselines on more realistic channels and that its NMSE values are comparable to those in Table 1.

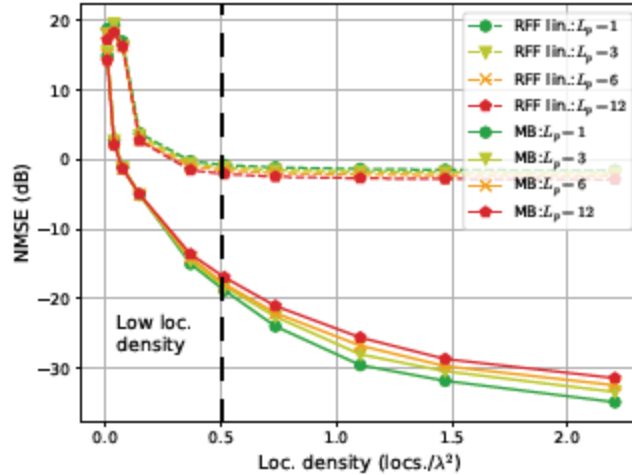


Fig. 4: Reconstruction performances

	MLP	RFF	RFF lin.	MB
Params.	16.8M	33.1M	4k	0.5M
NMSE _(dB)	0.14	-2.41	-2.21	-23.41

Table 2: NMSE over the test grid (ray-tracing channels).

5. CONCLUSION

To learn a so-called location-to-channel mapping, this article presented a model-based neural architecture. The design is based on a local planar approximation that was used to create a particular network architecture from a propagation model. Using realistic simulated data, the suggested network outperformed traditional and INR topologies with much less parameters, demonstrating a significant performance boost. To improve performance in the low location density regime, optimize the distribution of spatial frequencies, take into account multiple antennas, multiple subcarriers, and 3D locations, and further refine the hypernetwork architecture are all things that will be worked on in the future. These things should significantly improve the mapping learning capability.

6. REFERENCES

- [1] Nir Shlezinger, Jay Whang, Yonina C. Eldar, and Alexandros G. Dimakis, “Model-based deep learning,” *Proc. of the IEEE*, vol. 111, no. 5, pp. 465–499, 2023.
- [2] Hengtao He, Chao-Kai Wen, Shi Jin, and Geoffrey Ye Li, “Deep learning-based channel estimation for beamspace mmwave massive MIMO systems,” *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 852–855, 2018.
- [3] Xiuhong Wei, Chen Hu, and Linglong Dai, “Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems,” *IEEE Trans. on Commun.*, vol. 69, no. 1, pp. 182–193, 2021.
- [4] Taha Yassine and Luc Le Magoarou, “mpNet: Variable depth unfolded neural network for massive MIMO channel estimation,” *IEEE Trans. on Wireless Commun.*, vol. 21, no. 7, pp. 5703–5714, 2022.
- [5] Baptiste Chatelier, Luc Le Magoarou, and Getachew Redieteb, “Efficient deep unfolding for SISO-OFDM channel estimation,” in *IEEE Int. Conf. on Commun. (ICC)*, 2023.
- [6] Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville, “On the spectral bias of neural networks,” in *Int. Conf. on Mach. Learn.*, 2019, pp. 5301–5310.
- [7] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng, “Fourier features let networks learn high frequency functions in low dimensional domains,” *Advances in Neural Inf. Process. Syst.*, vol. 33, pp. 7537–7547, 2020.

- [8] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein, "Implicit neural representations with periodic activation functions," *Advances in Neural Inf. Process. Syst.*, vol. 33, pp. 7462–7473, 2020.
- [9] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Commun. ACM*, vol. 65, no. 1, pp. 99–106, 2021.
- [10] Mojtaba Bermana, Karol Myszkowski, Hans-Peter Seidel, and Tobias Ritschel, "X-fields: Implicit neural view-, light-and time-image interpolation," *ACM Trans. On Graphics (TOG)*, vol. 39, no. 6, pp. 1–15, 2020.
- [11] Yinbo Chen, Sifei Liu, and Xiaolong Wang, "Learning continuous image representation with local implicit image function," in *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 8628–8638.
- [12] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman, "Multiview neural surface reconstruction by disentangling geometry and appearance," *Advances in Neural Inf. Process. Syst.*, vol. 33, pp. 2492–2502, 2020.
- [13] Vincent Sitzmann, Michael Zollh"ofer, and Gordon Wetzstein, "Scene representation networks: Continuous 3d-structure-aware neural scene representations," in *Advances in Neural Inf. Process. Syst.*, 2019, vol. 32.
- [14] Xuanxuan Gao, Shi Jin, Chao-Kai Wen, and Geoffrey Ye Li, "ComNet: Combination of deep learning and expert knowledge in OFDM receivers," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2627–2630, 2018.
- [15] Mehran Soltani, Vahid Pourahmadi, Ali Mirzaei, and Hamid Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, 2019.
- [16] Eren Balevi, Akash Doshi, and Jeffrey G. Andrews, "Massive MIMO channel estimation with an untrained deep neural network," *IEEE Trans. on Wireless Commun.*, vol. 19, no. 3, pp. 2079–2090, 2020.
- [17] Luc Le Magoarou, Taha Yassine, St'ephane Paquelet, and Matthieu Crussi'ere, "Deep learning for location based beamforming with Nlos channels," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2022, pp. 8812–8816.
- [18] Luc Le Magoarou, Taha Yassine, St'ephane Paquelet, and Matthieu Crussi'ere, "Channel charting based beamforming," in *2022 56th Asilomar Conf. Signals, Syst., Comput.*, 2022, pp. 1185–1189.
- [19] David M Pozar, *Microwave engineering*, John Wiley & Sons, Second edition, 1998.
- [20] Kurt Hornik, Maxwell Stinchcombe, and Halbert White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [21] George Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [22] Yuan Cao, Zhiying Fang, YueWu, Ding-Xuan Zhou, and Quanquan Gu, "Towards understanding the spectral bias of deep learning," in *Proc. of the Thirtieth Int. Joint Conf. on Artif. Intell., IJCAI-21*, 2021, pp. 2205–2211.
- [23] Luc Le Magoarou, Antoine Le Calvez, and St'ephane Paquelet, "Massive MIMO channel estimation taking into account spherical waves," in *2019 IEEE 20th Int. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, 2019, pp. 1–5.
- [24] David Tse and Pramod Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [25] J"urgen Schmidhuber, "Learning to control fast-weight memories: An alternative to dynamic recurrent networks," *Neural Computation*, vol. 4, no. 1, pp. 131–139, 1992.
- [26] David Ha, Andrew M. Dai, and Quoc V. Le, "Hypernetworks," in *Int. Conf. on Learn. Representations*, 2017.
- [27] Ali Rahimi and Benjamin Recht, "Random features for large-scale kernel machines," in *Adv. Neural Inf. Process*, 2007, vol. 20.
- [28] Jakob Hoydis, Sebastian Cammerer, Fayc, al Ait Aoudia, Avinash Vem, Nikolaus Binder, Guillermo Marcus, and Alexander Keller, "Sionna: An open-source library for nextgeneration physical layer research," *arXiv preprint*, 2022.