DEEP LEARNING FOR LOCATION BASED BEAMFORMING WITH NLOS CHANNELS

Luc Le Magoarou, Taha Yassine, Stéphane Paquelet, Matthieu Crussière

To cite this version:
Luc Le Magoarou, Taha Yassine, Stéphane Paquelet, Matthieu Crussière. DEEP LEARNING FOR LOCATION BASED BEAMFORMING WITH NLOS CHANNELS. 2021. hal-03503074

HAL Id: hal-03503074
https://hal.science/hal-03503074
Preprint submitted on 27 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Massive MIMO systems are highly efficient but critically rely on accurate channel state information (CSI) at the base station in order to determine appropriate precoders. CSI acquisition requires sending pilot symbols which induce an important overhead. In this paper, a method whose objective is to determine an appropriate precoder from the knowledge of the user’s location only is proposed. Such a way to determine precoders is known as location based beamforming. It allows to reduce or even eliminate the need for pilot symbols, depending on how the location is obtained. The proposed method learns a direct mapping from location to precoder in a supervised way. It involves a neural network with a specific structure based on random Fourier features allowing to learn functions containing high spatial frequencies. It is assessed empirically and yields promising results on realistic synthetic channels. As opposed to previously proposed methods, it allows to handle both line-of-sight (LOS) and non-line-of-sight (NLOS) channels.

1. INTRODUCTION

Machine learning (ML) techniques have been applied successfully to wireless communications in recent years (see [1, 2] for exhaustive surveys). In particular, channel estimation and beamforming in the context of massive MIMO systems have benefited from ML [3, 4, 5]. Although such approaches are very promising, their main drawback is that most need pilot symbols to be sent in order to estimate channel state information (CSI) or beam training procedures [6]. Both approaches induce a consequent overhead and limit overall system efficiency.

Is CSI or beam training absolutely needed for a massive MIMO system to operate? Not necessarily, since appropriate precoders could be chosen at the base station based on other sources of information about users. In particular, it has been proposed to determine precoders based on estimated user locations, giving rise to the so-called location based beamforming (LBB) [7, 8, 9]. For such approaches, the location of users is directly used to determine precoders. This allows to spare a lot of resources for communications and enhances the physical layer security [10]. However, the main drawback of LBB with respect to CSI based approaches is that it assumes existence of a line of sight (LOS) propagation path on which the precoder is based. This LOS path may not exist in some scenarios, thus limiting the applicability of LBB. Nevertheless, the central element of LBB is the mapping that associates each location to a suitable precoder. This mapping can be learned by a neural network, in which case it is not limited to LOS channels.

Contributions. In this paper, a supervised learning method allowing to map the user location to an appropriate precoder is proposed. It requires a labeled database containing user locations and the associated channels in order to train a neural network realizing the mapping. The obtained precoders handle with similar success LOS and NLOS channels. The neural network has a specific structure based on random Fourier features (RFF) [11] that allows it to learn functions containing high spatial frequencies. This is beneficial to the task at hand because of the very fast variations of channel coefficients with respect to the user location due to fading.

Note that the location determination step is not studied in this paper, the location being considered known at the base station. This makes the proposed method totally independent of the specific location determination method that is used. In practice, location determination can be done either with help of a global navigation satellite system (GNSS), using a radar, with video cameras looking at the scene, or by sending pilot symbols (much less than for CSI acquisition) in order to determine directions of arrival.

Related work. Location based beamforming has been proposed a decade ago [7]. In its original formulation, it relies only on the main angle of departure (estimated by sending pilots) in order to synthesize an estimated single path channel which is then used to determine a precoder (without any learning). Such an approach is inherently limited to LOS channels, or at least channels comprising a dominant path. LBB Methods that were subsequently proposed [10, 8, 9] introduce refinements to the original idea but do not get rid of this inherent limitation. The method proposed in this paper overcomes this limitation to LOS channels since it directly learns the location/precoder mapping, but as a counterpart, it requires a training phase to calibrate this mapping. The proposed approach pertains to integrated sensing and communication (ISAC) [12], and more specifically to sensing-assisted communication which aims at leveraging sensors at the base station to enhance communication [13].

RFFs [11] are a crucial element of the proposed neural network. Their usage in neural networks stems from the impossibility to learn mappings containing high frequencies with classical neural networks such as the multilayer perceptron (MLP) [14]. Coordinate based neural networks whose objective is to learn a
mapping starting from a very low-dimensional space particularly benefit from RFFs [14]. RFF based neural networks have been applied successfully to 3D scene reconstruction [15], where the network learns the mapping linking the 3D location and orientation to the radiance and density (which can be seen as the light channel).

2. PROBLEM FORMULATION

The method proposed in this paper applies to a wide variety of multi-user massive multiple input multiple output (massive MIMO) wideband systems [16, 17, 18], operating indifferently in time division duplex (TDD) or frequency division duplex (FDD), where the antennas at the base station are indifferently colocated or not (in which case it is a distributed MIMO system). Let us consider a base station antennas and a single subcarrier, and denote \( \mathbf{h} \in \mathbb{C}^A \) the downlink channel vector between the base station and any given user, and \( l \in \mathbb{R}^D \) its location, where \( D \) can be two or three, depending on whether or not the elevation of users is considered relevant. Moreover, the proposed method can be directly transposed to systems comprising several subcarriers, but is presented here for a single subcarrier for ease of exposition.

Based on a labeled database of \( N \) downlink channels associated with the corresponding user locations

\[
\{ (\mathbf{h}_i, l_i) \}_{i=1}^N,
\]

the objective in this paper is to build a location based precoding function (or simply precoding function), whose function is to focus energy on the intended user location. It is mathematically defined as

\[
\mathcal{P} : \mathbb{R}^D \rightarrow \mathbb{C}^A \\
1 \mapsto \mathbf{w} \triangleq \mathcal{P}(1),
\]

where \( \mathbf{w} \) is the predicted precoding vector (of unit norm).

**Performance measure.** In order to evaluate precoding functions, the normalized correlation between the precoder \( \mathbf{w} \) and the channel \( \mathbf{h} \) is used. It is expressed as

\[
\eta \triangleq \frac{|\mathbf{w}^H \mathbf{h}|^2}{\| \mathbf{h} \|^2_2}.
\]

It is between zero and one (for a perfect precoder), and is tightly linked to the downlink channel capacity (considering a single user), whose expression is

\[
\log(1 + \eta \cdot \text{SNR}_{\text{opt}}),
\]

for received signal of the form \( y = \sqrt{P} \mathbf{w}^H \mathbf{h} s + n \) where \( n \sim \mathcal{N}(0, \sigma^2) \) is additive noise, \( s \) is the sent symbol and \( P \) the transmit power. In that setting, \( \text{SNR}_{\text{opt}} \triangleq \frac{P \| \mathbf{h} \|_2^2}{\sigma^2} \) is the highest achievable signal to noise ratio (SNR) [19]. In summary, the correlation is a single number between zero and one allowing to determine the maximum achievable downlink spectral efficiency for any transmit power and noise variance considering a given precoder.

3. PROPOSED SOLUTION

**Neural precoding function.** Deep neural networks are known to be universal function approximators [20, 21] and have led to great practical successes [22]. It is proposed here to implement the precoding function \( \mathcal{P} \) as a deep neural network. However, it has been shown that classical neural networks known as multilayer perceptrons (MLP) [23] are in practice unable to learn functions of high frequency [14]. This phenomenon is known as spectral bias [24, 25] and has been characterized with help of the theory of neural tangent kernels (NTK) [26].

In order to remedy this fundamental weakness, it has been proposed [14] to use random Fourier features [11] to help the neural network building high frequencies. This line is followed in this paper, motivated by the fact that the optimal precoder may vary fast with respect to the user’s location, due to fading. This results in a neural network whose first layer is fixed and corresponds to random Fourier features expressed as

\[
\mathbf{B} \sim \mathcal{N}(0, s^2 \mathbf{I}),
\]

where the variance parameter \( s^2 \) controls the frequency range (the higher \( s^2 \) the more high frequencies are likely to appear). It is proposed to concatenate this RFF layer with a classical MLP using fully connected layers and rectified linear units (ReLU) activation functions [27] of width \( M \) (number of neurons per layer) and depth \( Q \) (number of layers). The last layer is of width \( 2A \), which corresponds to the real and imaginary parts of the predicted precoder \( \mathbf{w} \) stacked together. The structure of the proposed neural network is shown on Fig. 1. Simply put, the role of the RFF layer is to pre-build high frequencies in order to ease learning for the subsequent MLP.

In order to calibrate the weights, the cost function to minimize is expressed as

\[
\text{CF} \triangleq 1 - \frac{1}{N} \sum_{i=1}^{N} \frac{|\mathcal{P}(l_i)^H \mathbf{h}_i|^2}{\| \mathbf{h}_i \|^2_2}.
\]

It measures the misalignment of the predicted precoders with respect to the training channels and is between 0 (if the precoders are perfectly aligned with the channels) and 1 (if the precoders are orthogonal to the channels).

**Practical setup.** Two distinct phases are to be distinguished for the proposed method to operate in a real system:

- **Training phase:** At first, the database \( \{ (\mathbf{h}_i, l_i) \}_{i=1}^N \) has to be built. To do so, the base station can send pilots and estimate channels using any classical method. Moreover, it has to access users’ locations, using either a GNSS, video cameras or also pilots. The neural network can be trained with the built database in order to calibrate the precoding function \( \mathcal{P} \).
Adam optimization algorithm \cite{30} for the computation is chosen by cross validation and fixed so that the approach is inherently biased towards LOS channels. Using a deep learning approach with a simple neural network (without RFFs) yields a slight improvement (median of \(0.714\)). This very simple neural network is prone to too much variance. However, the proposed method using RFFs allows to attain a median of \(0.946\), which is a lot better than what is attained by the two aforementioned baselines, because of a better bias/variance trade-off.

It is also interesting to look at the spatial distribution of the attained correlation. This is depicted on Fig. 3 (for the classical LBB approach) and Fig. 4 (for the proposed approach). On these figures, the colors (blue to yellow) denote the obtained correlation for 50 000 test channels, and the base station location is denoted by a red cross. Note that on both figures, black areas on the left correspond to zero-norm channels. Comparing these figures, it is obvious that the classical LBB approaches (Fig. 3) yield good precoders (yellow) for areas where user are in line of sight with respect to the base station, but very bad precoders elsewhere (blue areas). On the opposite, the proposed approach (Fig. 3) yields good precoders everywhere on the considered area, even in locations that are in non-line of sight with respect to the base station. This clearly shows the benefit of the proposed method, which allows

\[ P(\eta < x) = \begin{cases} 0 & x = 0 \\ \frac{1}{Q} & x = 1 \end{cases} \]

Note that these two phases are not mutually exclusive, and can be intertwined in an online mode, in order to allow the base station to adapt to slow changes in the environment.

4. EXPERIMENTS

In this section, the proposed method is empirically assessed and compared to concurrent approaches on realistic multipath synthetic channels.

**Simulation settings.** The experiments are carried out using training multipath channels taken from the DeepMIMO dataset \cite{28}. The ‘O1’ urban outdoor ray-tracing scenario is chosen (with 5 paths per channel, see \cite[Figure 2]{28} for a schematic view of the scenario), with a single base station (BS 16) and a subset of all possible user locations (rows R1000 to R3852). The base station is equipped with a square uniform planar array (UPA) with \( A = 64 \) half-wavelength separated antennas at a frequency of 3.5GHz.

**Implementation details.** The method is implemented with help of the PyTorch library \cite{29}, so that gradients are computed automatically. The optimization of the precoding function is done according to the cost given in (4) on a database of \( N = 10000 \) gradient descent (minibatches of 100 training channels) using the Adam optimization algorithm \cite{30} for 50 epochs. The number of RFF is set to \( R = 1000 \), and the standard deviation used for their computation is chosen by cross validation and fixed so that \( \frac{1}{Q} = 0 \). The MLP following the RFF computation is of depth \( Q = 4 \) and width \( M = 512 \). Note that complex weights and inputs are handled classically by stacking the real and imaginary parts so that the neural network treats only real numbers.

**Baselines.** The proposed method is compared with two baselines:

- A classical LBB approach \cite{7, 8}, for which the precoder is chosen as the normalized LOS channel in the direction of the user, assuming the azimuth and elevation are perfectly known. This approach would be optimal in terms of correlation for LOS channels comprising a single path. This baseline is here to show the potential of the proposed method with respect to prior art.

- A deep learning approach using a simple MLP (without RFFs). The architecture of the MLP is exactly the same as the one of the proposed method, except the first layer (RFF computation) that is replaced by a simple fully connected layer. This baseline is here to show the interest of RFFs.

**Results.** The performance of the three compared methods on a test set of 10 000 channels (not used for training) is shown on Fig. 2. The cumulative distribution function (CDF) of the correlation is plotted (with an ideal precoder it would be equal to zero for all values of \( x < 1 \) and to one for \( x = 1 \)). The plot clearly shows the great improvement brought by the proposed method. Indeed, in such an urban environment where a LOS path is not available for every location, the classical LBB approach yields a median correlation of \(0.708\). This is explained by the fact that the approach is inherently biased towards LOS channels. Using a deep learning approach with a simple neural network (without RFF) yields a slight improvement (median of \(0.714\)). This very simple neural network is prone to too much variance. However, the proposed method using RFFs allows to attain a median of \(0.946\), which is a lot better than what is attained by the two aforementioned baselines, because of a better bias/variance trade-off.
to obtain good precoders in areas where classical LBB approach fails, again owing to the incapacity of handling NLOS channels.

In order to assess the applicability of the proposed method, it is interesting to vary the number of training channels \( N \) in order to determine which density of training channels is required. To that aim, the results in terms of average and median correlation are shown as a function of \( N \) on Fig. 5. From this figure, it can be seen that, obviously, the greater \( N \) the greater the test correlation. However, one can see that the proposed method starts to perform well (on test data) for \( N > 3000 \). The area in which users can be located spans a total surface of \( 20520 \, \text{m}^2 \), which means that the proposed method requires approximately one training channel every \( \frac{20520}{3000} \approx 7 \, \text{m}^2 \) on average to perform well. Such an empirical result is important in order to determine the time duration of the training phase in which both channels and locations have to be collected in order to build the database used to train the neural network.

5. CONCLUSION

In this paper, a location based beamforming method was proposed. It is based on a neural network comprising RFFs allowing to learn functions of high spatial frequency. The method was empirically assessed on realistic synthetic channels, yielding very promising results. Indeed, it is able to handle both LOS and NLOS channels whereas previously proposed methods perform well only for LOS channels. This shows the capacity of the proposed structure to interpolate well between training points thanks to RFFs which inject prior information on the learned mapping.

In the future, the method could be extended in several ways. First of all, the distribution of frequencies used to build the RFFs could be optimized to further enhance performance. It would also be interesting to aim at predicting not only an appropriate precoder but directly the channel vector. This would amount to summarize the whole propagation environment in the weights of a neural network, and could serve many applications, such as channel mapping [31, 32] (in space) or radio environment compression.

6. REFERENCES


