Marie Skłodowska Curie Action

WINDMILL

Machine Learning for Wireless Communications

H2020-MSCA-ITN-ETN
Grant Agreement Number: 813999

WP4–Prediction Schemes and Anticipatory Optimization for Fast-Varying Processes

D4.2–Intermediate report on machine learning techniques for wireless channels

Contractual Delivery Date:	August 31, 2021
Actual Delivery Date:	August 30, 2021
Responsible Beneficiary:	AALTO
Contributing Beneficiaries:	AALTO, EURECOM, Ericsson AB, AAU, KTH
Dissemination Level:	Public
Version:	1.0

PROPRIETARY RIGHTS STATEMENT
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813999.

PROPRIETARY RIGHTS STATEMENT
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813999.

Document Information

Document ID:	WP4/D4.2
Version Date:	August 30, 2021
Total Number of Pages:	49
Abstract:	We have considered machine learning methods for channel learning and prediction. Beam selection in mmWave 5G systems with two approaches is investigated. First, contextual information from from sensors muonted on the vehicles is leveraged to introduce an efficient mmWave beam selection method. Second, from received signals of users at the base station a radio map is constructed in an unsupervised manner called Channel Chart (CC). By constructing beam-wise CCs and training SNR predictors for different beams given CC locations, we predict the SNR difference of different beams, thereby predicting the next best beam for a moving user. In a non stand alone mmWave system, given CC of the sub-6 GHz base station, we devise a best mmWave beam prediction model based on the CC locations, as well. Similarly, in beyond 5G systems, concepts such as Large Intelligent Surfaces (LIS) arises, which act as an accurate tool for wireless sensing. Its vast amount of antenna elements allows the creation of radio images for describing the environment which leads to the intersection of the computer vision and wireless sensing fields to understand channel information. Lastly, we study correction of RF hardware impairments in context of beyond 5G (6G) systems. Specifically the problem related to learning detailed and interpretable models of grey- and black-box systems and its solution via Bayesian optimisation of combinatorial structures.
Keywords:	Beam-management, beam prediction, channel charting, non-standalone systems, radio resource management, large intelligent surface, computer vision, sensing

Authors

Full name	**Beneficiary/ Organisation**	**e-mail**	**Role**
Cristian Jesús Vaca Rubio	AAU	cjvr@es.aau.dk	Contributor
Zecchin Matteo	EURECOM	zecchin@eurecom.fr	Contributor
Sergey Tambovskiy	Ericsson AB / KTH	sergeyta@kth.se	Contributor
Parham Kazemi	AALTO	parham.kazemi@aalto.fi	Contributor

Reviewers

Full name	Beneficiary/ Organisation	e-mail	Date
Hanan Al-Tous	Aalto	hanan.al-tous@aalto.fi	August 17, 2021
Hugo Tullberg	Ericsson AB	hugo.tullberg@ericsson.com	August 27, 2021

Version history

Version	Date	Comments
0.1.	August 16, 2021	Draft version for review
0.2.	August 20, 2021	Updated after first round of review comments
1.0.	August 30, 2021	Final version, updated after second round of review comments

Table of Contents

List of Figures

List of Acronyms

5G 5^{th} Generation

AAC Antenna Array Calibration

AAS Advanced Antenna System

ABF Analog Beamforming

ADC Analog-to-Digital Converter

AoA Angle of Arrival

BNP Bayesian nonparametric

BS Base Station

CC Channel Chart

CMD Collinearity Matrix Distance

CSI Channel State Information

DAC Digital-to-Analog Converter

DL Downlink

DNN Deep Neural Network

DPD Digital Predistortion

DR Dimensionality Reduction

ECDF Empirical Cumulative Distribution Function

FDD Frequency Division Duplexing

GP Gaussian process

GPR Gaussian Process Regression

HO Handover

HPA High Power Amplifier

IA Initial Access

ICBH Intra Cell Beam Handover

ICH Inter Cell Handover

IMD Intermodulation Distortion

IRS Intelligent Reflective Surface

KNN K-Nearest Neighbour

LIS Large Intelligent Surface

LM Levenberg-Marquardt

LTE Long Term Evolution

LTI Linear Time Invariant

LUT Lookup Table

MIMO Multiple Input, Multiple Output

ML Machine Learning

mmWave Milimeter Wave

MSE Mean Squared Error

NAS Neural Architecture Search

NN Neural Network

NR New Radio

NSA Non Stand Alone

OTA Over-the-Air

PCA Principal Component Analysis

PIMC Passive Intermodulation Correction

QuaDRiGa Quasi Deterministic Radio Channel Generator

RF Radio Frequency

RIS Reconfigurable Intelligent Surface

RMSE Root Mean Squared Error

RX Receiver

SGD Stochastic Gradient Descent

SNR Signal-to-Noise Ratio

SVM Support Vector Machine

t-SNE t-Distributed Stochastic Neighbor Embedding

TDD Time Division Duplexing

TX Transmitter

UE User Equipment

UL Uplink

1. Introduction and Motivation

5th Generation (5G) New Radio (NR) supports beam-oriented cellular systems in which beams transmitted towards and / or received by Base Station (BS) and User Equipment (UE) can be flexibly configured. Beam selection is mainly performed by exhaustive and iterative search methods where each has own pros and cons [1].Through an Initial Access (IA) phase a UE establishes a physical link connection to a BS. After the IA in order to ensure connectivity of a moving user, tracking the best beam pair between the UE and BS is necessary. Therefore, in addition to inter cell Handover (HO), intra cell beam HO is needed.

As a direct consequence of the narrow Milimeter Wave (mmWave) beam-width, 5G systems need to rely on precise alignment and tracking procedures to establish a reliable and high throughput communication links. In the absence of estimates of channel matrices, which are extremely costly to obtain in case of numerous antennas, beam selection is typically carried out using iterative search procedures. These procedures introduce a large communication overhead, especially in the the vehicular-to-infrastructure (V2I) communications systems, for which mmWave communication is envisioned to be a key technology [2]. For this reason, in chapter 2 we introduce a simple V2I system model to frame the mmWave beam selection problem and we explain the limitations of conventional iterative search procedure. We then list successful examples in which contextual information from sensors mounted on the vehicles and the infrastructure is leveraged to reduce the beam selection overhead of iterative search procedures, rendering mmWave communication efficient also in the V2I scenarios.

Recently, data driven methods which are capable of extracting meaningful information from large volume of data has gained considerable attention in wireless communication [3]. With presence of massive amount of Channel State Information (CSI) information, a novel framework called channel charting which uses unsupervised learning tools to produce radio maps of high dimensional CSI, can be exploited for radio resource management problems [4]. Channel Chart (CC) can be used in such applications as user grouping and HO that do not require absolute location information. Beam selection is primarily carried out based on the CSI feedback. Thanks to Machine Learning (ML) approaches and the CC concept the problem could be solved without any feedback overhead. In this regard, first a stand alone mmWave system is adopted and CC's of beams are constructed. Then, using different ML techniques, a Signal-to-Noise Ratio (SNR) predictor is trained based on CC of one beam to predict the SNR of a UE at other neighboring beams. Depending on the SNR difference of current beam and neighboring beams a HO decision will be made. Second, in a Non Stand Alone (NSA) system where a sub-6 GHz assist mmWave system to provide a robust communication is investigated and best mmWave beam is predicted based on CSI measured at the sub-6 GHz BS. The sub-6 GHz system's CC is constructed and annotated with mmWave system best beam information. A best mmWave beam predictor is trained on the given information, connecting microwave band CSI with a predicted best mmWave beam.

On a related note, in massive Multiple Input, Multiple Output (MIMO), the base station is equipped with a very large number of antennas. With the aim of pushing their benefits to the limit and looking towards post-5G, researchers are defining a new generation of base stations that are equipped with an even larger number of antennas. This gives rise to the concept of Large Intelligent Surface (LIS) which designates a large continuous electromagnetic surface able to transmit and receive radio waves. These large surfaces can be placed on walls/ceilings and are easily integrable into the surroundings. In practice, an LIS is composed of a collection of closely spaced tiny antenna elements. While the potential for communications of LIS is being investigated [5, 6], these devices offer possibilities which are not being under study accurately, i.e., environmental sensing based on radio images [7]. The large aperture and high number of antenna elements of LIS can be used for performing an accurate environment sensing while providing a huge amount of data. In this way, ML and deep learning are useful for understanding radio environmental maps. These radio maps can be mappings of the signals into an image structure, describing the propagation environment. This allow merging the computer vision field into wireless communications taking advantage of state-of-the-art solutions for understanding image information. The input to this computer vision algorithms can be addressed from different perspectives: CSI or reconstruction of radio environmental maps to exploit convolutional neural networks which will be discussed in details in chapter 4. Finally, we note that LIS is one of the technologies being considered for future 6G systems, which may change the relevant cost/benefit analysis in that any sensing functionality is then expected to be added onto the system rather than requiring explicit investment on extra dedicated hardware.

Correction of various Radio Frequency (RF) hardware impairments (both linear and non-linear in nature) plays an important role in the operation of any modern wireless system, especially in cases of beyond 5G (6G) systems. As without units such as Digital Predistortion (DPD), Passive Intermodulation Correction (PIMC) and Antenna Array Calibration (AAC) radio link performance will suffer significant degradation. Mathematical models for aforementioned algorithms are learnt though process of system identification. Detailed information about the problem of hardware impairment correction is described in D4.1.A – Joint Probabilistic Modelling of Wireless Channels and Hardware Impairments.

The remainder of this report is organized as follows. In chapter 2 mmWave beam selection based on contextual information from sensors mounted on the vehicles in a V2I system model is studied. In chapter 3 beam SNR prediction in a mmWave stand alone system and best mmWave beam prediction in a NSA system are discussed. In chapter 4 the potential of LIS as a beyond massive MIMO system is studied by leveraging its major sensing capabilities under a specific industrial use case. In chapter 5 we study the problem of learning detailed and interpretable models of grey- and black-box systems. Once learnt they can be used for controller design or as parts of simulation environments. Specifically, this chapter targets the problem of structure and parameter estimation of grey-box systems given a limited number of observed samples. To solve this problem we employ a method of Bayesian optimisation of combinatorial structures (BOCS). We demonstrate its performance in application to grey-box systems and discuss how the BOCS technique can be further extended to support structure learning in dynamical black-box systems.

2. Contextual information for mmWave beam selection in vehicle-to-infrastructure communication

In vehicular-to-infrastructure (V2I) communications, for which mmWave communication is envisioned to be a key technology [2], beam selection and tracking are particularly challenging due to the high mobility of the receivers, which leads to reduced beam coherence time [8]. In this scenario, conventional beam selection techniques, such as beam sweeping or multi-level beam selection [9, 10] impose a significant overhead. Therefore, more efficient beam selection techniques that can reduce the cost of iterative search procedure by exploiting contextual information are of great interest. It has been shown that contextual information from sensors mounted on the vehicles and the infrastructure can be leveraged to reduce the beam selection overhead. In the following, we introduce a simple system model to frame the V2I beam alignment problem, we illustrate the conventional iterative search procedure and then show provide examples of different types of contextual information that can be leveraged to reduce the alignment overhead.

2.1. System Model

To illustrate the beam selection problem we may consider an orthogonal frequency-division multiplexing (OFDM) mmWave system with analog beamforming capabilities, where the BS located on the street curb serves a vehicle in its coverage area utilizing N_c subcarriers. Both the transmitter and the receiver ends are equipped with antenna arrays with a single radio frequency (RF) chain and fixed complex beam codebooks, which we denote by $\mathcal{C}_t = \{\mathbf{f}_i\}_{i=1}^{C_t}$ and $\mathcal{C}_r = \{\mathbf{w}_j\}_{j=1}^{C_r}$, respectively. The downlink channel matrix from the BS to the vehicle over the n'th subcarrier is denoted by $\mathbf{H}_n$.
For each precoder and combiner vector pair $(i, j) \in \mathcal{C}_t \times \mathcal{C}_r$, the resulting channel gain at subcarrier n is determined by $\mathbf{w}_j^H \mathbf{H}_n \mathbf{f}_i$, where $(\cdot)^H$ denotes the conjugate transpose. We then define the effective power gain matrix $\mathbf{G} \in \mathbb{R}_+^{|\mathcal{C}_t| \times |\mathcal{C}_r|}$, whose (i, j)-th entry contains the aggregate power gain over the N_c subcarries for the transmitter-receiver codebook pair (i, j), as

$$\mathbf{G}_{i,j} = \sum_{n=1}^{N_c} |\mathbf{w}_j^H \mathbf{H}_n \mathbf{f}_i|^2. \tag{2.1}$$

The optimal pair of precoding and combining vectors is the one that maximizes the channel gain,

$$(i^*, j^*) = \underset{(i,j)}{\operatorname{argmax}}\, \mathbf{G}_{i,j}. \tag{2.2}$$

Without side information, the transmitter and receiver need to perform a search through the $C_t \times C_r$ beam pairs in order to identify (i^*, j^*). In particular, the current 5G release exploits an iterative search that sweeps in the angular domains at both receiver and transmitting ends and measures the relative channel quality. In this setting, exhaustive search over all beam pairs can be extremely costly given the size of the beam codebooks. In order to reduce the overhead, hierarchical search procedures have been considered. These resort on a multi-tier search, during which coarser and larger beams are firstly probed to roughly localize the

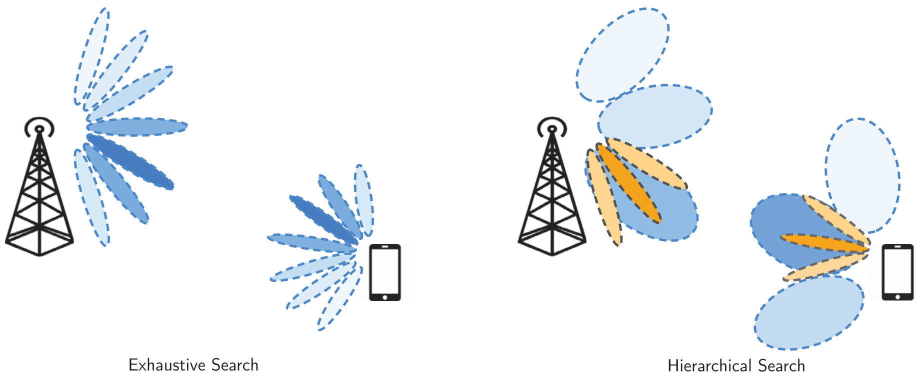

Figure 2.1: Exhaustive search procedures (left) sweep all the available beams at the transmitting and receiver ends in order to find the best beam combination. Hierarchical search procedures (right) work by refining the search space: beams with larger width (blue) are used to find the direction with the largest gain and then narrower beams (orange) are used to search within that angular sector.

receiving end and then they are refined using more directional beams in the most prominent direction given by the previous search step. The exhaustive and hierarchical approaches are depicted in Fig. 2.1.
Alternatively when side information is available, the goal is to refine the search space by inferring a small subset of k beam pairs $\mathcal{S}_k \subset \mathcal{C}_t \times \mathcal{C}_r$ such that $(i^*, j^*) \in \mathcal{S}_k$. This results in a reduction of $\frac{k}{C_t \times C_r}$ in the search space of the beam selection procedure. Two metrics to gauge the quality of $\mathcal{S}_k$ as a function of its size k, are the top-k accuracy and top-k throughput ratio. The top-k accuracy is formally defined as

$$A(k) = \mathbb{E}\left[\mathbb{1}\{(i^*, j^*) \in \mathcal{S}_k\}\right], \tag{2.3}$$

and it measures the fraction of instances for which the best beam index is in the top-k selector output. On the other hand, the top-k throughput ratio is defined as

$$T(k) = \frac{\mathbb{E}\left[\max_{(i,j)\in\mathcal{S}_k} \log_2(1 + \mathbf{G}_{i,j})\right]}{\mathbb{E}\left[\log_2(1 + \mathbf{G}_{i^*,j^*})\right]} \tag{2.4}$$

where $\mathbf{G}$ is the effective power matrix defined in (2.1) and all expectations are with regard to the inherent randomness introduced by vehicles' positions and channel realization. Note that the top-k throughput ratio is a very informative metric for the problem at hand. In fact, the numerator represents the throughput that can be achieved (at a zero dB transmit SNR) by searching only among the top-k beams suggested by the contextual information; while the denominator is a normalizing factor representing the maximum throughput achievable by an exhaustive beam sweeping approach.

2.2. Contextual Information for mmWave beam alignment

We now give notable examples of contextual information that can be used to enhance the beam search procedure.

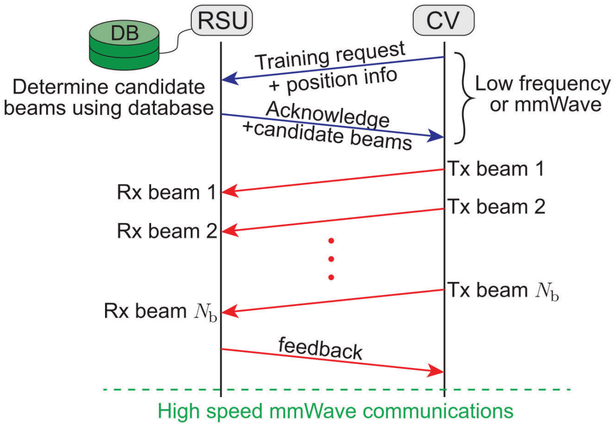

Figure 2.2: Beam training based from inverse fingerprinting. [11].

2.2.1. Position Information

The most commonly available side information at vehicle side is the position information acquired through the global positioning system (GPS). In [11], this information is used to apply an inverse fingerprint scheme that is able to discard less promising direction and sweep only those that are more likely to yield a robust communication link. The proposed approach leverages a fingerprint database that contains channel characteristics such as signal strength or multipath signature at different access locations. The database can be obtained at the road side unit (RSU) by requesting communicating vehicles (CV) to perform beam training at different location in the coverage area. Once the dataset is generated, the proposed beam alignment scheme works as depicted in Fig. 2.2. The CV probes the RSU and sends its position information, upon reception, the RSU queries the pre-fetched database and sends a sequence of candidates beams back to the CV for training. It is experimentally shown that even modest size databases (in the order of 200-300 beam training phases) are sufficient to attain satisfactory performance. Moreover, measurements obtained during heavy traffic conditions positively transfer to mild and low traffic scenarios. Exploiting the position based inverse fingerprinting, the training phase can be reduced to 30 beams in a system with originally 256 available pairs.

2.2.2. Inertial Information

Inertial sensors measure the acceleration and angular speed of objects over different axes of motion. Vehicles are often equipped with these sensors for safety and control purposes [12]. In the context of mmWave communication, inertial data can also be used to reconfigure antenna arrays [13]. In the case of sharp beams, vibrations due to road asperities and the titling of vehicles can misalign the receiver and transmitting ends, trigger costly search procedures. Exploiting inertial sensor, it is possible to track the orientation of the vehicle and its pitch angle to compensate for this unwanted impairments. Using real world measurements it is shown that thanks to inertial sensors it is possible to attain throughput almost equals to the perfect alignment scenario. Furthermore, positional and motion information can be jointly processed to further reduce the alignment overhead [14].

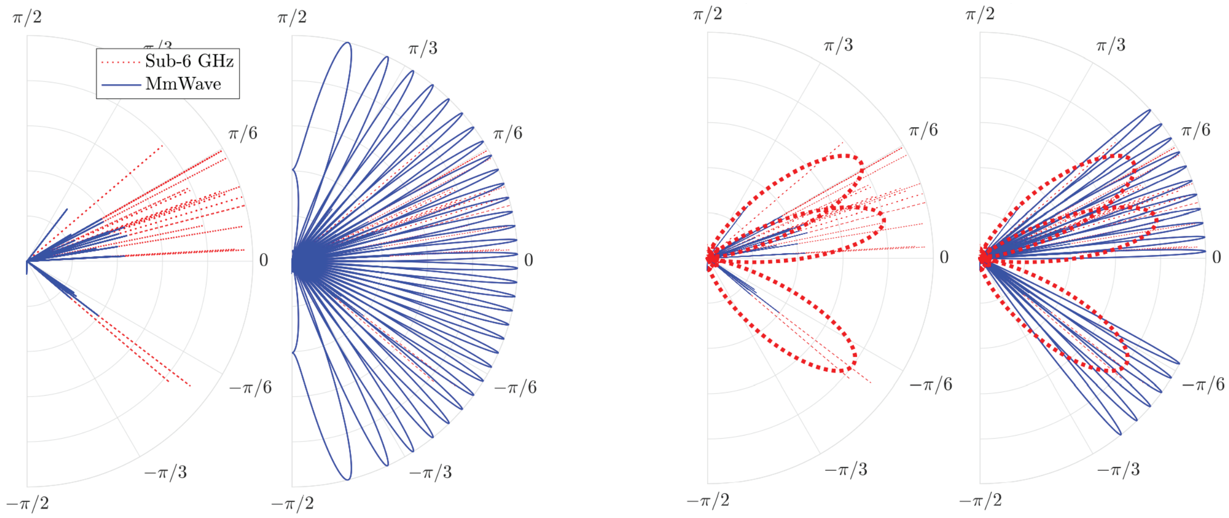

Figure 2.3: On the left, the sub-6GHz and mmWave dominant paths. On the right, the sub-6GHz beams associated to the strongest paths and the candidate mmWave beams using the spatial information. Image from [16].

2.2.3. sub-6GHz Information

In future networks, the weaknesses of mmWave communication technologies will be mitigated by the existing sub-6GHz apparatus. Other than rendering the system more robust, equipment operating in other frequency bands can be leveraged to provide side information. For instance, a radar located at the BS and operating outside the millimeter spectrum can be used to estimate the direction of arrival of vehicles and aid the beam search [15]. In [16], the authors assume the presence of digital MIMO system operating in the sub-6GHz spectrum to back up the mmWave communication. The spatial information is then extracted by the out-of-band measurements as shown in Fig. 2.3 and the most prominent directions for the mmWave beams are predicted. The relation between sub-6GHz and the mWave channel measurements can be contrived and difficult to characterize analytically. For this reason, data-driven methods have been proposed to learn it directly from data. Artifical neural network can be trained using sub-6GHz channel measurements in order to predict the best mmWave beams and blockage probabilities [17]. Tested and trained on 3D ray-tracing datasets, data-driven methods are able to predict the best beam without training procedure 50% of the times and they can predict blockage with an error probability close to the irreducible error.

2.2.4. LIDAR Information

Thanks to the recent surge of autonomous driving technologies, high dimensional sensory information is nowadays commonly available also at the vehicle side. For instance, light detection and ranging (LIDAR) is commonly used for autonomous navigation. LIDAR uses a laser to produce a depth map of the environment and surrounding obstacles using delay measurements of the back-scattered signal. Because of the data dimensionality and the lack of analytical models that would relate LIDAR depth map to mmWave beams quality,

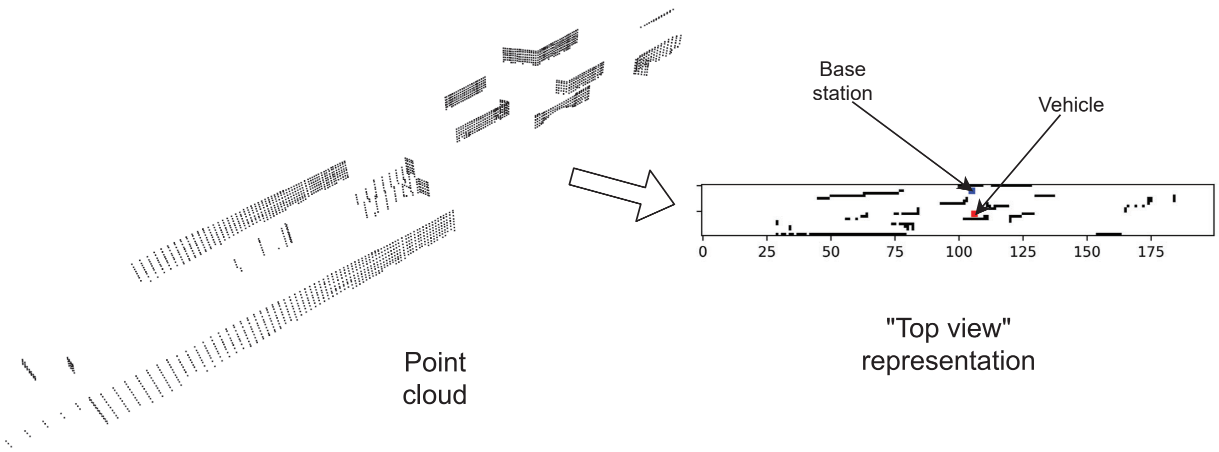

Figure 2.4: Preprocessing of the LIDAR point cloud.

Table 2.1: Performance comparison between the state-of-the-art LIDAR based beam alignment solutions.

Model	$A(1)$	$T(1)$	$A(5)$	$T(5)$	# params.
[18, 19]	$31.5 \pm 2.6\%$	$46.1 \pm 2.6\%$	$71.9 \pm 2.2\%$	$76.1 \pm 1.9\%$	403677
[20]	$52.3 \pm 1.9\%$	$70.3 \pm 2.6\%$	$85.3 \pm 0.9\%$	$90.8 \pm 1.5\%$	7462
[21]	$59.5 \pm 0.5\%$	$79.9 \pm 0.8\%$	$87.0 \pm 0.3\%$	$94.6 \pm 0.8\%$	30872

data-driven methods have been considered to effectively process LIDAR signals as side information for beam search. In [18, 19], a NN architecture was trained over simultaneous LIDAR and ray-tracing channel datasets with a top-k classification metric to identify k beam directions that most probably include the beam resulting in the largest channel gain. In order to reduce the computational cost and NN model size, a simplified classifier architecture that can be trained in a distributed fashion using federated learning was proposed in [20].
Later, in [21] authors proposed a simplified LIDAR preprocessing technique reported in Fig. 2.4 and a convolutional neural network (CNN) architecture that was trained to exploit LIDAR and positional data in order to identify the best beam directions and reduce the beam search overhead in V2I communication. This solution introduced various novelties in order to increase the beam classification accuracy while at the same time reducing the number of trainable parameters and computational burden, such as:

- A novel loss function, inspired by the knowledge distillation (KD) techniques [22], which not only maximizes the prediction accuracy of the best beam index, but also its corresponding power gain.

- A non-local attention scheme [23], which improves the beam classification accuracy, specifically for the non-of-sight (NLOS) case.

- A curriculum training strategy [24] which improves both the convergence speed and the final beam prediction accuracy.

Simulation results on benchmark datasets show that, utilizing solely LIDAR data and the receiver position, the NN-based beam selection scheme can achieve 79.9% throughput of an exhaustive beam sweeping approach without any beam search overhead and 95% by searching among as few as 6 beams. The key performance indicator of the above discussed solution are summarized in Table 2.1.

2.3. Conclusion

Beam alignment in the context of mmWave communication introduces a large communication overhead due to its costly search procedures and it represents one of the major challenges for 5G. This problem is exacerbated in the V2I communication domain, as mobility reduces beam coherence time. For this reason, in this chapter we have formalized the beam selection problem in a V2I system, and we explained why agnostic iterative search solution are doomed to fail in case of large antenna regime. We then listed alternative solutions that embody in their search procedures contextual information such as position, inertial data, sub-6GHz information and LIDAR scans. These solution are shown to be practical, to greatly reduce the length of the search procedures and, in some cases, even rendering search not necessary at all.

3. Channel Charting-Based Beam SNR Prediction

Wireless communication using mmWave, with gigabit-per-second data rate promise, have received considerable attention in the context of 5G and beyond. Beamforming [25] and massive MIMO antenna arrays are expected to combat high path loss of mmWave bands as well as other losses due to oxygen absorption and higher noise floor resulting from larger bandwidth.
Beam management issue is discussed in [26], where a Kalman filter approach is proposed to reduce alignment error and ensure more reliable connectivity as long as UE mobility is moderate. Recently, out-of-band information has shown to be beneficial in mmWave channel estimation [27]. Sub-6-GHz systems are more robust than mmWave bands in hard propagation conditions, and suffer less from blockages. Thus, in order to have a seamless and reliable connection integrating 4G Long Term Evolution (LTE) and 5G mmWave radio is advantageous.
In cellular networks user mobility is supported by HO. In 5G mmWave systems Intra Cell Beam Handover (ICBH) which switches beams among one BS beams is considered in addition to Inter Cell Handover (ICH) where the serving BS is changed. With assumption of autonomous beamforming at UE side, the UE beam is not under control of the BS. Impact of BS beam change is not clear without determining all BS-UE beam pairs. Therefore, the BS is not capable of determining the SNR of the best beam by measuring the UE transmission in all the BS beams.
An efficient beam selection method is required to reduce the exhaustive beam swiping method overhead since, in 5G NR a large number of beams are used for communication. As introduced in Section 2 Location information has been shown to be beneficial for speeding up the beam training. Location assisted beam management combined with Machine Learning techniques is discussed in [28]. By leveraging CC which is a dimensionality reduction technique applied to collection of massive MIMO CSI to construct radio map of the cell, we can devise a new method for prediction of the beam SNR. In [29] ICH based on the predicted SNR is investigated; the SNR of a UE from a neighboring BS is predicted based on the relative location information provided by CC. We explore the possibility of relative location information for predicting best beam of a mmWave system.
This study is twofold; First, a stand alone mmWave system is investigated. The promising result of ICH based on the predicted SNR motivated us to extend it to beam SNR prediction based on CC. Second, in a NSA system best mmWave BS beam prediction is studied. We consider best mmWave beam prediction for a UE, based on long term CSI measured by a possibly different BS at a sub-6 GHz carrier.

3.1. System Model

For the sub-6 GHz communication, the BS has an array of P antenna elements and the UEs have a single omni-directional antenna. In the sub-6 GHz frequency, the UEs perform pilot transmissions and the BS measures the channel $\mathbf{h_u} \in \mathbb{C}^{P\times 1}$. The CSI covariance of UE u at the sub-6 GHz BS is computed as:

$$\mathbf{R}_u = \mathbb{E}\left\{\mathbf{h}_u\mathbf{h}_u^H\right\}, \tag{3.1}$$

where the expectation $E\{.\}$ is over spatial and frequency samples and $(.)^H$ denotes the Hermitian conjugate.
For both stand alone and NSA systems we consider the mmWave communication system with a single BS and U UEs. The BS has an array of M antennas and UE has N element array antenna.
The MIMO channel between the BS and UE u at a subcarrier is $\mathbf{H}_u \in \mathbb{C}^{M\times N}$. The BS and UEs are capable of beamforming. We assume that the BS codebook has M beams and UE codebook has N beams. Beamforming vectors for the BS are $\mathbf{w}_m \in \mathbb{C}^{M\times 1}$ with $m = 1 \dots M$ and for a UE are $\mathbf{v}_n \in \mathbb{C}^{N\times 1}$ with $n = 1 \dots N$. Wideband beams are assumed at the BS, i.e., the same beam is used for all frequency samples. The UE chooses its best beam for every subcarrier. The Discrete Fourier Transform (DFT) based codebook of size M and size N is used at the BS and UE respectively. The codebook, $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_Q]$ is shown as:

$$\mathbf{c}_q = \frac{1}{\sqrt{Q}}[1, e^{j2\pi\frac{q}{Q}}, \dots, e^{j2\pi\frac{(Q-1)q}{Q}}]^T,\ q = 1, \dots, Q, \tag{3.2}$$

where $Q = M$ for the BS and $Q = N$ for the UE.
The received signal from UE u when the BS uses beam m and UE uses beam n on the subcarrier is then

$$y^u_{m,n} = \mathbf{w}^H_m \mathbf{H}_u \mathbf{v}_n x + z_u = h^u_{m,n} x + z_u, \tag{3.3}$$

where $h^u_{m,n}$ represents the effective channel coefficient for BS beam m and UE beam n, x is the transmitted symbol with $E\{|x|^2\} = 1$ and z_u is additive white Gaussian noise. The effective channel vector for receiving from UE u using beam $\mathbf{v}_n$, measured from all BS beams is:

$$\mathbf{h}^u_n = \mathbf{W}^H \mathbf{H}_u \mathbf{v}_n, \tag{3.4}$$

where $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_M]$ is the matrix with BS beamforming vectors as columns.
The UE selects it's best beam to transmit towards / receive from a BS beam $\mathbf{w}_m$ using the function:

$$\hat{n} = \hat{n}(m) = \underset{n}{\operatorname{argmax}}\, |h^u_{m,n}|^2. \tag{3.5}$$

Hence, the effective channel for a UE transmission depends on which BS beam it it assumes to be transmitting to.
The average received SNR at BS beam m from a transmissions of UE u towards this beam is then

$$\gamma_{m,u} = \frac{1}{\sigma^2}\mathbb{E}\left\{\left|h^u_{m,\hat{n}(m)}\right|^2\right\}, \tag{3.6}$$

where σ^2 is the noise power. The expectation is taken over frequency / subcarrier samples and over temporal samples taken from the fast fading process within a short time interval. For a transmission towards BS beam m, the effective channel vector from UE u, measured from all BS antennas then becomes

$$\mathbf{h}^u_{\hat{n}(m)} = \mathbf{W}^H \mathbf{H}_u \mathbf{v}_{\hat{n}(m)}. \tag{3.7}$$

If a UE has a single antenna, the best BS beam can be directly measured at the BS. However, when the UE has multiple antennas, and autonomously uses a beamformer, the BS cannot unilaterally measure and find the best beam towards the user. The BS can measure the elements in (3.7) from UE transmissions, but the BS cannot measure from transmissions towards beam $\mathbf{w}_m$ what the channel coefficients would be if the UE were to transmit towards $\mathbf{w}_{m'}$ with $m' \neq m$. There is a *beam-mismatch problem* arising from UE autonomous precoding; the precoder $\mathbf{v}_{n(m')}$ may or may not be the same as $\mathbf{v}_{n(m)}$.

3.1.1. Basics of Channel Charting

A CC [4, 30] preserves the relative neighbor relations of the UEs in the physical domain by creating radio maps. We use channel features to create CCs, which capture large-scale channel effects. For instance, a covariance matrix at the BS can be used to find the channel features of the the UEs.
In order to construct a CC, first a dissimilarity matrix $\mathbf{D}$ at the BS side is formed which contains the pairwise CSI dissimilarities between UEs in the cell. Using a dimensionality reduction technique e.g. t-Distributed Stochastic Neighbor Embedding (t-SNE) and dissimilarity matrix $\mathbf{D}$, a representation of UE in the cell is obtained.
To perform CC in a beam-based manner, effective channels must be measured from multiple beams. This can be done, if the BS is equipped with multiple RF chains. For simplicity, we assume that the BS is able to measure the effective channel from all of its M beams, obtaining an $M \times 1$-dimensional effective channel $\mathbf{h}^u_{\hat{n}(m)}$ when it measures the output from all M beams with the UE transmitting towards BS beam m using $\hat{n}(m)$. According to which BS beam $\mathbf{w}_m$ the UE is transmitting towards, the effective channel will vary. As a result of autonomous UE beamforming, the effective channel at the BS is determined by the BS beam that the UE transmits towards. Hence, a separate CC is considered for each BS beam, using the received signals from (3.7) with UEs transmitting towards the particular BS beam $\mathbf{w}_m$ using UE beam $\mathbf{v}_{\hat{n}(m)}$ that is best for it.
Beam-specific CCs are obtained by calculating the beam-m specific covariance matrices from effective channels of UEs that are transmitting towards beam $\mathbf{w}_m$:

$$\mathbf{R}_{m,u} = \mathbb{E}\{\mathbf{h}^u_{\hat{n}(m)}(\mathbf{h}^u_{\hat{n}(m)})^H\}, \tag{3.8}$$

where u is the UE index. Then beam-specific dissimilarity matrix $\mathbf{D}_m$ is created using the Collinearity Matrix Distance (CMD) metric [31]; the dissimilarity for two UEs u and u' with beam based covariances $\mathbf{R}_{u,m}$ and $\mathbf{R}_{u',m}$ is the Frobenius norm of normalized covariance matrices:

$$d_m(\mathbf{R}_{m,u}, \mathbf{R}_{m,u'}) = 1 - \frac{\mathrm{Tr}(\mathbf{R}_{m,u}\mathbf{R}_{m,u'})}{||\mathbf{R}_{m,u}||_\mathrm{F}||\,\mathbf{R}_{m,u'}||_\mathrm{F}}, \tag{3.9}$$

where Tr indicates trace operator. A specific CC for beam m is a 2D representation of the CSI of the set of UEs by feeding the dissimilarities to the Dimensionality Reduction (DR) algorithm.
In the NSA system, as we have single antenna UEs in the sub-6 GHz band, there would be only one CC for the cell. Dissimilarities between the sub-6 GHz covariance matrices $\mathbf{R}_u$ of the UEs as features, and the Log-Euclidean distance to measure feature dissimilarity is used. The dissimilarity between the covariance features $\mathbf{R}_u$ and $\mathbf{R}_{u'}$ of two UEs u and u' using Log-Euclidean distance measure is given by [32]:

$$d_{u,u'} = ||\log \mathbf{R}_u - \log \mathbf{R}_{u'}||_\mathrm{F}, \tag{3.10}$$

where the log indicates matrix logarithm. Then, the CC can be constructed from the dissimilarity.

3.2. Framework for CC-Based Best Beam Prediction

In 5GNR, a moving user served by a BS may experience switching between several beams of the BS. Beam HO is assumed to performed in a network centric manner, where the BS selects the target beam for a UE. In the stand alone system where only mmWave BS are present, the basic principle is that each BS has a CC constructed offline and the CC locations are annotated with other neighboring beam's SNRs. Then, a ML model is trained to predict the SNR of a target beam. Finaly, the annotated CC and the SNR prediction model are used to make an ICBH decision.
Figure 3.1 illustrates the network centric model. During training phase, CCs are created, annotated with neighbor beam SNRs , and SNR predictors are trained. In this regard, for each beam we have the CC locations and the beam SNRs as the input of SNR predictor and the output is the neighboring beams SNR. Then, in the online phase, using the current beam CC location and SNR, the target beam SNR is predicted. In this phase, a UE first establishes its connection to a beam. Then an SNR mapping function is used to predict the SNR of a target beam. After predicting SNR of all possible beam targets for the UE, an ICBH decision based on SNR difference is made.

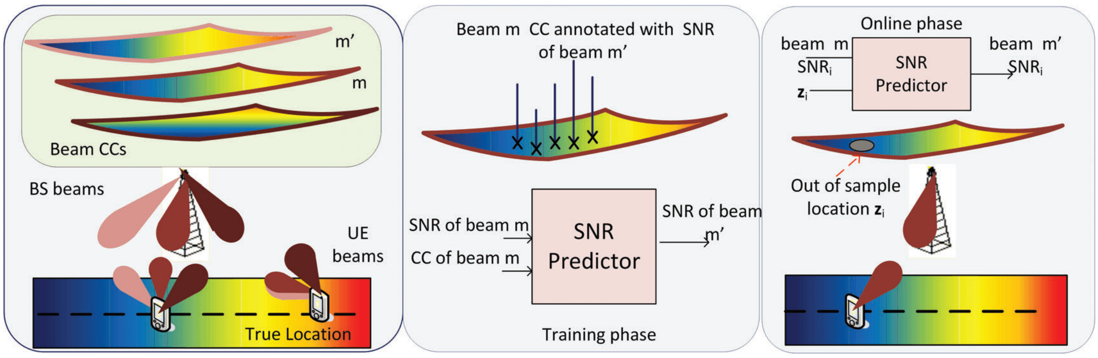

Figure 3.1: Beam SNR prediction based on beam CC: (Left); Constructing of beam based CC. (Middle); CC annotation and SNR prediction (the offline phase). (Right); Beam SNR prediction based on beam CC (online phase) [33].

In the NSA system, for the sub-6 GHz BS a CC is constructed and the mmWave system's best beam information is added to CC locations. During an offline phase all information is gathered in a central control unit. Therefore, in the offline phase CC locations are annotated with mmWave BS best beam IDs and MLs are trained. Then, in the online phase for a new UE a prediction algorithm predicts the best beam ID of mmWave system based on the CC location of the UE. Figure 3.2 shows the different phases of best beam ID prediction in a NSA system. Three parts of training phase are: CC construction, annotation withe mmWave best beam ID, and training of beam ID predictor. In the online phase, first a UE is mapped to a CC location and then the best beam ID is predicted based on the CC location.

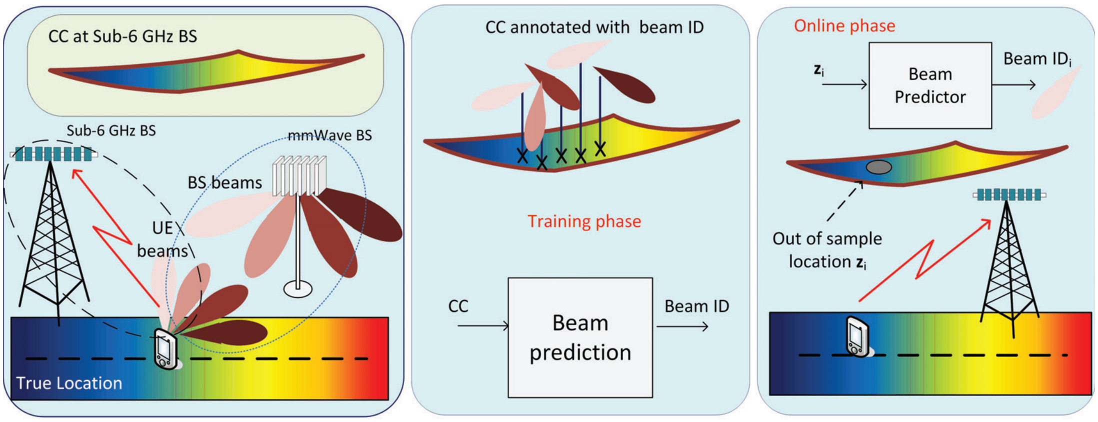

Figure 3.2: Beam prediction based on channel charting in Non-Standalone systems. (Left); a street segment served by a multi antenna sub-6 GHz and mmWave BSs. (Middle); Training phase with CC construction and annotation. (Right); Online phase showing the best beam prediction for a new UE [34].

3.2.1. Machine Learning Algorithms

In the stand alone system, SNR prediction is formulated as a regression problem. Given the beam m annotated CC, the BS finds a function that predicts the neighboring beam SNRs of a UE served by beam m. To do so, Neural Network (NN), Gaussian Process Regression (GPR), and K-Nearest Neighbour (KNN) are used to predict the target beam SNR. As CC is dimensionality deducted form of CSI, there is possibility to reduce the dimensionality of CSI to any number other than 2 (2D CCs are mostly used to show the neighboring information preserving ability of the CC). Thus, we evaluate the effect of higher degrees of dimensionality reduction on beam SNR prediction, as well. Also, for NNs different back propagation algorithms have various convergence speed and complexity. Stochastic Gradient Descent (SGD) and Levenberg-Marquardt (LM) [35] algorithms are used to minimize the output loss function. The LM has been shown to converge faster for moderate sized NNs but with higher computational cost.

In the NSA system, a classification problem is formulated and NN , Support Vector Machine (SVM) classifiers as well as KNN are considered. KNN as a low complexity predictor is considered for best beam ID prediction. The best beam ID classification is a multi class problem. Thus, in the SVM classifier which inherently is binary classifier, some changes are needed to be applied. One-vs-One method for the SVM is considered where every pair of two classes are classified by a binary classifier, and then based on a voting approach, the class that gains the majority of votes for an input is assigned to it.

3.3. Simulation and Results

The simulation is performed by generating a physical layout and CSI using the Quasi Deterministic Radio Channel Generator (QuaDRiGa) simulator. The simulation setting is summarized in Table 3.1. As for the stand alone system the BS is located at $[-114, -110]$ m, whereas for the NSA 2 GHz network is located at $[-100, -100]$ m, the mmWave BS of is

Table 3.1: Simulation Parameters

Parameter	sub-6 GHz	mmWave
Center Freq.	2 GHz	28 GHz
Subcarrier BW.	15 KHz	240 KHz
No. of Subcarriers	600	256
BS antenna	8 ULA	32 ULA
BS antenna element pattern	Omnidirectional	3GPP TR 36.873
UE antenna	Omni-directional	8 ULA
Scenario	3GPP 38.901 UMa-NLOS	

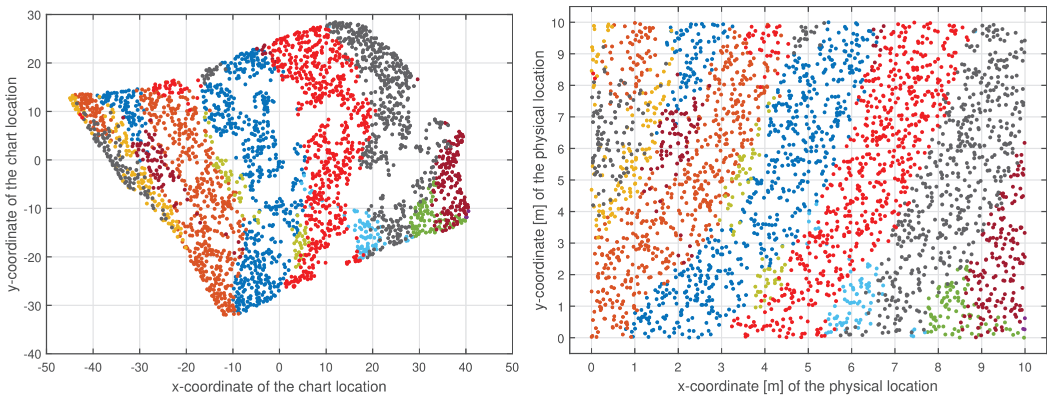

Figure 3.3: (Right); The true physical locations marked with their best BS beams with different colors. For example, beam 32 is the best beam for the locations marked with red. (Left); CC annotated with best beam information using t-SNE DR.

at $[5, 30]$ m and UEs are uniformly scattered in a street segment between [0 10] m on both the x-axis and the y-axis. For each UE , in order to compute the covariance matrix, 100 small-scale fading samples are collected.
In the stand alone system, for each beam one CC is constructed from the beam-specific dissimilarity matrix. CMD distance is used for computing the dissimilarities. t-SNE dimensionality reduction is used for CC construction. Then, the CC locations are annotated with neighboring beam SNRs during offline phase.
In the NSA system, 2 GHz channels are used to calculate dissimilarities, and a CC is constructed, annotated with mmWave best beam ID information in the offline phase. Similarly, t-SNE dimensionality reduction is used to construct CC. However, the dissimilarity matrix is obtained based on the log- Euclidean distance. Figure 3.3 shows the resulting annotated CCs obtained with t-SNE dimensionality reduction.

3.3.1. Beam SNR Prediction

NN, GPR, and KNN are used to create the SNR predictors. Mean Squared Error (MSE) is used as the loss function and an exponential kernel is used in the GPR predictor. NN

Table 3.2: Performance of Different Predictors Based on the 2D True Physical Location and 10D t-SNE CC.

Predictor	KNN 10	GPR	NN3 10 LM	NN3 100 SGD
Phys Loc.	0.26	0.25	0.24	0.47
t-SNE	0.35	0.29	0.32	0.52

activation function is hyperbolic tangent. Three layers are considered for NNs with various number of neurons and are called as ”NNHidden layers-Number of neurons in each hidden layer-Training algorithm”, e.g., NN3 10 LM is a NN with three hidden layers and 10 neurons in each hidden layer using LM algorithm. Root Mean Squared Error (RMSE) is used as the performance measure for predictors and the input to SNR predictors is assumed to be various dimension of CC from 2D to 10D. As in the area of interest there is only 8 dominant beam, the predictions are limited to these 8 beams and the reported average RMSE is averaged over 8×8 pairs of beam predictors. A more detailed discussion is provided in [33].

Figure 3.4 shows the average RMSE of target beam SNR prediction as a function CC dimension for the test data set. A slight improvement is achieved as we increase the dimension of the input CCs. Using SGD, increasing the network size is beneficial, however, it has basically twice the error in the LM algorithm. It is expected as LM is more effective in optimizing NN weights. Also, due to the smooth variation of SNR in the street segment, KNN is also showing a good performance.

We have considered relative location-based beam SNR prediction. As a benchmark, true location-based beam SNR prediction is also investigated. Comparison result is shown in Table 3.2. NN with LM is the best physical location based predictor with 0.24 dB RMSE. GPR is the best CC based SNR predictor, with only a 0.05 dB gap to the best true location-based prediction. Thus, the performance loss of CC based beam SNR prediction, as compared to prediction based on ground truth physical location, is negligible.

3.3.2. Best Beam ID Prediction

Beam prediction function using NN, SVM, and KNN predictors is created. Gaussian kernel for SVM and different NN sizes are used. As for KNN two cases are considered; either looking at the nearest neighbor for predicting the best mmWave beam ID or 10 nearest neighbors. Results of average prediction accuracy are shown in Table 3.3. The best NN structure for true location is NN3 20 and for t-SNE is NN3 50. The performance benefit from using NNs for beam prediction is that when predicting based on CC, the accuracy is equal to the case when using the true location.

By investigating top-2 and top-3 best beam accuracy, we reach 98% and 99% for both true location and CC based predictions. Accordingly, this information can be used to reduce the best beam search time [34].

3.4. Conclusion

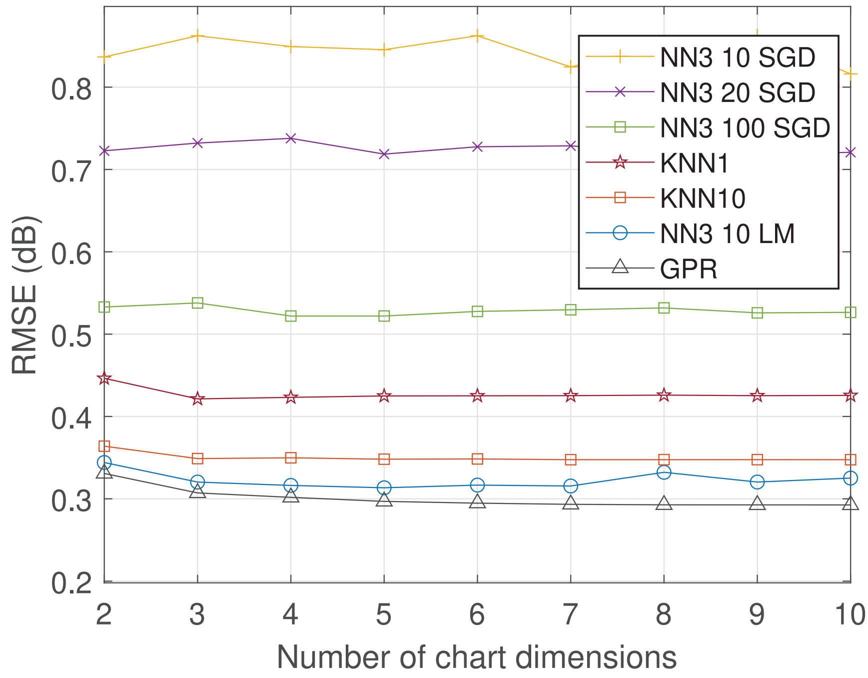

Figure 3.4: Average RMSE of different predictors as function of beam CC dimension for t-SNE DR.

Table 3.3: Prediction Accuracy of Different Predictors

Predictor	Best NN	SVM	KNN(10)	KNN(1)
t-SNE	90.2	88.2	86.6	86.8
True location	90.8	89.3	88.4	87.1

Neighboring beam SNR prediction and best beam prediction in mmWave systems based on CC were studied. In mmWave systems, narrow beams are needed to be tracked to ensure the connectivity of UEs. To do so, a handover process is needed to be initialized between different beams. In this regard first, a stand alone mmWave BS is considered where its beam's CCs are constructed and annotated with neighboring beams' SNR. Then in an offline phase a predictor is trained based on the CC locations to predict the SNR of neighboring beams. In an online phase, the trained SNR predictor is used to predict neighboring beam SNRs and based on the SNR difference an ICBH is decided. Second, in a NSA mmWave system best BS beam is predicted based on sub-6 GHz cell CC. A radio map of sub-6 GHz system i.e. CC is annotated with best beam information of mmWave BS during offline phase and a predictor is trained to predict the best mmWave beam based on sub-6 GHz BS CC. In the online phase, the best mmWave beam is predicted. In both cases it has shown that there is a small gap between performance of the CC-based predictions and true location based predictions.

4. Predictive Machine Learning for multiuser beamforming

MIMO technology has been widely studied in the literature since it can significantly improve data rates and reliability of wireless communication systems. The current tendency in wireless network designs is based on increasing the number of antennas at base stations i.e. acquiring a theoretical benefit of usage of MIMO systems. It comprises of using large arrays for a proportional increase in the SNR with respect to the number of elements N under the far field assumption [36].

By deploying base stations with a huge number of antennas, massive MIMO arises to obtain larger array gains and perform an accurate spatial multiplexing of users as well as exploit spatial diversity to take advantage of the channel hardening effect. Since the seminal paper by Marzetta [37], massive MIMO became a reality, being one of the key enablers for future wireless networks [38].

The promising benefits of these systems in principle have led to potential base station deployments with an increasing number of antennas beyond conventional massive MIMO. Hence, current research leads with new concepts such as *holographic MIMO*, *Large Intelligent Surface (LIS)* or *Intelligent Reflective Surface (IRS)*, also known as *Reconfigurable Intelligent Surface (RIS)*, emerging as a natural evolution of classical massive MIMO. In this deliverable, we focus on LIS which refers to a surface which integrates a vast number of antennas into a limited aperture (considered as a continuous electromagnetic surface) while IRS consists in a surface compounded by tunable circuits that are able of altering the characteristics of the incoming signals in a desired way. While the potential for communications of LIS is being investigated, these devices offer possibilities which are not being under study accurately, i.e., environmental sensing.

4.1. Radio Image-based LIS sensing

Providing an integrated system which combines sensing and communication capabilities might be one of the key enablers for the future 6G networks. Its high spatial density of antennas and large array aperture can be exploited to perform sensing in an alternative way. In a wireless context, an LIS can be described as a structure which uses electromagnetic signals interacting with the scatterers (via reflections) in order to obtain a profile of the environment. Then, the resulting superimposed signal is used to obtain a high resolution image of the propagation environment. Note the LIS elements are observing the CSI information to conform these images. What is more, the LIS spatial resolution has been proven to provide e.g., high positioning accuracy [39]. Furthermore, one of the main advantages of using radio image-based LIS sensing is its ability to capture the environment in an accurate way. When using this approach, the complexity of the system is reduced to using information represented as an image. This provides a great opportunity to take advantage of computer vision for understanding radio image information.

Computer vision approaches are suitable in terms of understanding digital images and extraction of high-dimensional information from the real world [40]. The image data can take many forms, such as frames of a video or views from multiple cameras. In this case they can be seen as snapshots of the radio propagation environment.

On a related note, the goal of image processing is to strengthen or compress image/video information by using pixel-wise operations like transforming one image into another by filtering. Furthermore, there is no extraction of meaningful information from those pixel-wise operations. However, the goal of computer vision is to extract meaningful information from images/videos (i.e. understand the content of digital images) [40]. There is a vast range of applications in the literature to determine whether a specific object is present or not during a specific image or scene in a video, they are commonly known as object detectors [41]. These kind of algorithms can apply to take advantage of useful information of LIS images, for instance performing the tracking of the patterns obtained in the surface images. Computer vision is not limited to pixel-wise operations which are often far more complex than image processing.

Those complex operations are often summarized into feature detectors which can provide rich information about the contents of the image/video. The goal of ML (from which computer vision is a subset) is to optimize differentiable parameters so as that a specific loss/cost function is minimized. ML is strongly related to image processing and computer vision. As an example, convolutional neural networks (CNNs) are using all three techniques, convolutions are from image processing as they work on per small pixel neighborhood basis, the process of extracting image content is from computer vision while the kernel parameters are adjusted using ML techniques and common back-propagation optimization.

We conclude that the combination between computer vision algorithms and image processing techniques for the preprocessing of the input radio-based LIS images can be the key for exploiting the advantages of studying the channel characteristics as an image.

4.2. Computer vision-based LIS sensing

In the last sections, we highlighted one of the main enabling features of 6G will be radio sensing. With the aim of demonstrating the usefulness of LIS, radio based imaging and computer vision, we here present a baseline problem in order to analyze the sensing potential of LIS.

4.2.1. System model

Let us consider an industrial scenario where R robots are supposed to follow some predetermined fixed routes. Assume that, due to arbitrary reasons, we would like to monitor their positions in the route. The goal is to be able to track them based on the sensing signals transmitted by the R target devices.

In order to perform the detection, we assume that a M antenna elements LIS is placed along the ceiling, whose physical aperture comprise its whole area. In this way, we have a projected view of the environment in a 2D plane. The sensing problem reduces to determine, from the superposition of the received signals from each of the robots at every of the M LIS elements, the (x, y) coordinates (a.k.a the position) of the R robots involved in the scenario. The superposed complex baseband signal received at the LIS is given by

$$\mathbf{y} = \sum_{r=1}^{R} \mathbf{h_r} x_r + \mathbf{n}, \tag{4.1}$$

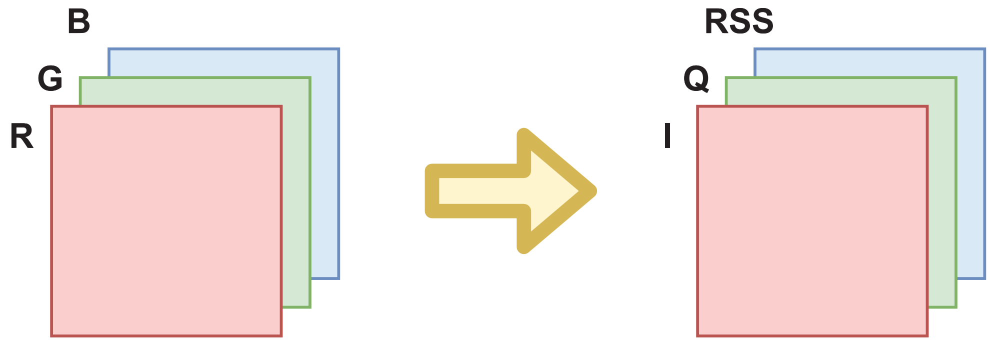

Figure 4.1: Illustration of performed image mapping

with x_r the transmitted (sensing) symbol from robot r (we consider $x_r = 1$ without loss of generality), $\mathbf{h_r} \in \mathbb{C}^{M\times 1}$ the channel vector from a specific position of robot r to each antenna-element, and $\mathbf{n} \sim \mathcal{CN}_M(\mathbf{0}, \sigma^2\mathbf{I}_M)$ the noise vector. Please note we are considering a narrow-band scenario.

4.2.2. LIS radio image generation

According to the radio-image sensing technique, we can interpret the received signal throughout the LIS as an image by mapping the complex resulting superposed signal in (4.1) to the RGB components of a color image. In this way, we can create an image by mapping the I/Q components of the signal and power values (I/Q/RSS) such that $I = R$, $Q = G$ and $RSS = B$. To that end, we apply min-max feature scaling for each layer of the color image (I/Q/RSS - R/G/B) to map the signals into the range of [0, 255], in which the value of each pixel $m_{i,k}$ for $i = 1, \ldots, M$ and $k = 1, \ldots, N_p$ positions of the robot R is obtained as

$$m_{i,k} = \left\lceil m_{\mathrm{MIN}} + \frac{(y^c_{i,k} - y^c_{\mathrm{MIN},k})(m_{\mathrm{MAX}} - m_{\mathrm{MIN}})}{y^c_{\mathrm{MAX},k} - y^c_{\mathrm{MIN},k}} \right\rceil, \tag{4.2}$$

where $y^c_{i,k}$ are the elements of $\mathbf{y}$, $c = \{I, Q, RSS\}$ component, in which $y^{RSS}_{i,k} = \|h_{i,k} + n_{i,k}\|^2$, $m_{\mathrm{MAX}} = 255$ and $m_{\mathrm{MIN}} = 0$, and

$$y^c_{\mathrm{MAX},k} = \max_{\{i=1,\ldots,M\}} \mathbf{y^c}_{i,k}, \quad y^c_{\mathrm{MIN},k} = \min_{\{i=1,\ldots,M\}} \mathbf{y^c}_{i,k} \tag{4.3}$$

are the maximum and minimum values from a point $\mathbf{p}_k$ along the surface. Figure 4.1 shows an illustration.
This strategy allows preserving the spatial information of the signal, as every antenna element is directly mapped to a pixel value, whose positions correspond to the one of the antenna in the physical LIS deployed in the ceiling. An exemplary image is shown in Figure 4.2 for $R = 1$.

4.2.3. Received signal and noise modeling

To simulate this scenario, we rely on ray tracing simulations to account for the multipath propagation phenomena and compute the rays in the most reliable way. For this matter, we use the commercially available software ALTAIR FEKO [42].

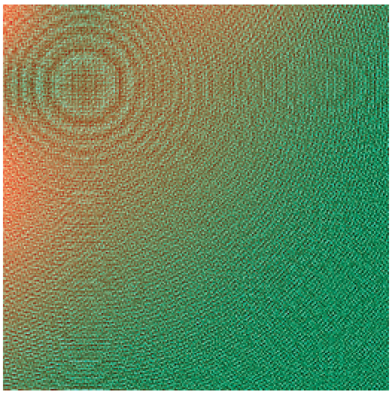

Figure 4.2: Exemplary RGB image obtained at the LIS in a noiseless scenario

In this way, the complex electric field arriving at the i-th antenna element, $\widetilde{E}_i$, can be regarded as the superposition of each path. Assuming isotropic antennas, the complex signal at the output of the i-th element is therefore given by

$$y_i = \sqrt{\frac{\lambda^2 Z_i}{4\pi Z_0}}\widetilde{E}_i + n_i, \tag{4.4}$$

with λ the wavelength, $Z_0 = 120\pi$ the free space impedance, Z_i the antenna impedance, and n_i is complex Gaussian noise with zero mean and variance σ^2. For simplicity, we consider $Z_i = 1\,\forall\, i$. Finally, in order to test the system performance under distinct noise conditions, the average SNR over the whole route, $\overline{\gamma}$, is defined as[1]

$$\overline{\gamma} \coloneqq \frac{\lambda^2}{4\pi Z_0 M T \sigma^2}\sum_{t=1}^{T}\sum_{i=1}^{M}|\widetilde{E}_i|^2, \tag{4.5}$$

where T is the number of time steps and M denotes the number of antenna elements in the LIS.

4.2.4. Model description

We introduce an ML model to perform the position estimation task based on the radio-based images obtained at the LIS. As Figure 4.2 shows, a concentric circular pattern is obtained result of the robot transmission in the scenario. This pattern has a characteristic shape that can be tracked by training an object detection algorithm in a simple way. Object detection is a technology related to computer vision and image processing that deals with detecting and locating instances of semantic objects of a certain class (such as humans, glasses, or buildings) in digital images and videos. In a similar way, There is a vast range of algorithms to perform the tracking of the concentric circle, but we have chosen to retrain YOLOv3 [41] due to its proven lower prediction time. The LIS deployed in the ceiling generates the radio images (according to Section 4.2.2) allowing using YOLOv3 to predict the bounding boxes around the robots transmission patterns. Then, by computing the center of the predicted bounding boxes, we obtain an inference of the (x, y) coordinates. Figure 4.3 shows the workflow in both the training and prediction stage.

[1] This is equivalent to averaging over all the points $\mathbf{p}_k$.

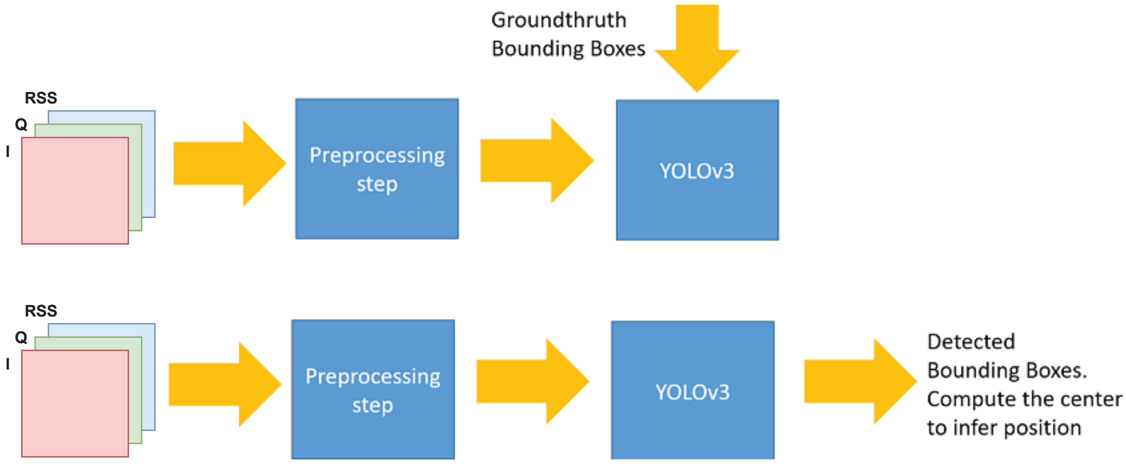

Figure 4.3: Computer vision algorithm workflow

In our considered problem, the dataset is obtained by sampling the I/Q components at each element of the LIS while the R robots move along the trajectories. Hence, the dataset consists of radio image snapshots. To train the algorithm, we provide the groundtruth bounding boxes generated from the positions of the robots in the scenario for every specific image in the dataset.

4.2.5. Noise averaging strategy

The presence of noise may be critical in the radio image sensing, since it impacts considerably in the image classification performance [43].
Trying to mitigate its impact, let us assume the system is able to obtain S extra samples at each channel coherence interval $\forall$ $\mathbf{p}_k$ to perform an S-averaging. Note that, if $S \to \infty$, then

$$\left|y'_{i,k}\right|_{S\to\infty} = \mathbb{E}[y_{i,k}|h_{i,k}] = h_{i,k}, \tag{4.6}$$

meaning that the noise variance at the resulting image has vanished, i.e., the received superimposed signal at each antenna (conditioned on the channel) is no longer a random variable. Observe that in this way the image preserves the pattern. This effect is only possible if the system would be able to obtain a very large number S of samples within each channel coherence interval.

4.3. Numerical results and Discussion

For the following results, we deploy an LIS in the ceiling whose physical aperture is $M = 259 \times 259$ and we use $S = 100$ extra samples to reduce the noise contribution. Here the underlying assumption of using the entire ceiling is necessary to perform a direct mapping between the real positions and the inferred (x, y) positions (pixels) in the image.

4.3.1. Scheduled vs simultaneous robot transmissions

We here present as a baseline a comparison of the sensing performance by tracking $R = 1$ vs $R = 3$ robots following their predefined routes to check the robustness of the obtained

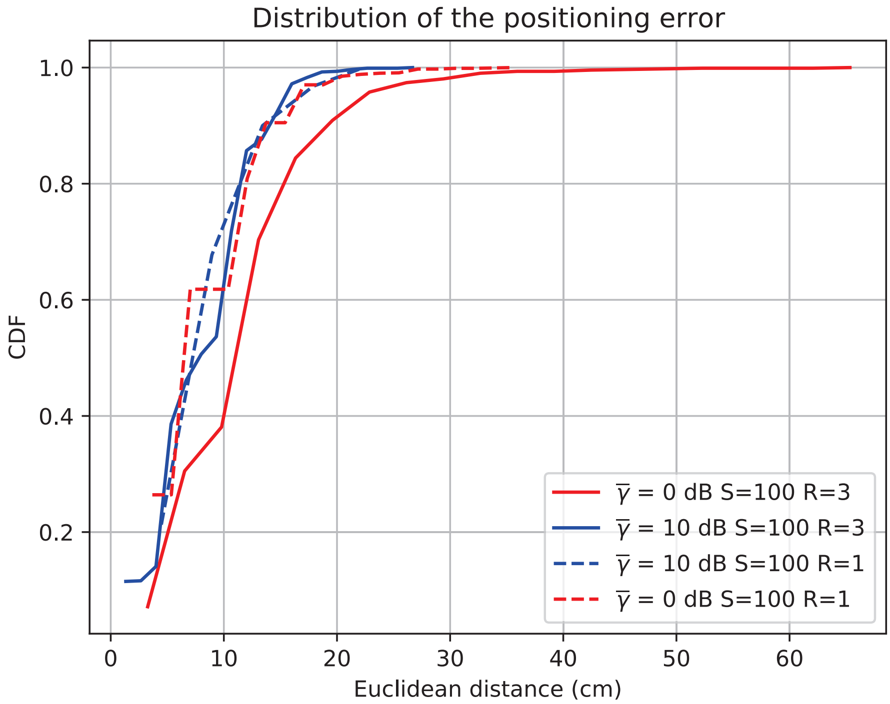

Figure 4.4: ECDF for radio image sensing for R = 1 vs R = 3 robots, with fixed LIS aperture of $M = 259 \times 259$ and S = 100 samples

pattern when every robot transmits by turns in a scheduled manner against a simultaneous transmission of the 3 under different $\overline{\gamma} = [0, 10]$ dB conditions.

Figure 4.4 shows the Empirical Cumulative Distribution Function (ECDF) of the Euclidean distance between the predicted and the groundtruth positions in the test set for a scheduled vs a simultaneous transmission. The results show the sensing system is able to perform a multi-robot tracking with a little compromise of the accuracy in positioning in comparison to a scheduled scheme. In the $\overline{\gamma} = 10$ case the accuracy in positioning is almost identical ($R = 1/3$), while in the worst $\overline{\gamma} = 0$ scenario the 80% of the results are under an error of 12 cm when R = 1 and around 16 cm when R = 3. This is thanks to the LIS physical aperture, which allows obtaining a high resolution image that creates a pattern discernible while 3 robots are transmitting in the scenario. The results seem quite promising as under 80% of the results are around 16 cm in the worst case while around 10 cm in the best case. Please note, the performance of the system may be improved by increasing the S extra samples, also increasing the antenna density in the physical aperture. The results show that even with a simultaneous robot transmission, the study of the channel characteristics for positioning under different $\overline{\gamma}$ scenarios seems like a promising solution to address the task. In terms of the detection performance, an mAP = 100% has been acquired for the 4 cases, meaning no robots were missed in the detection.

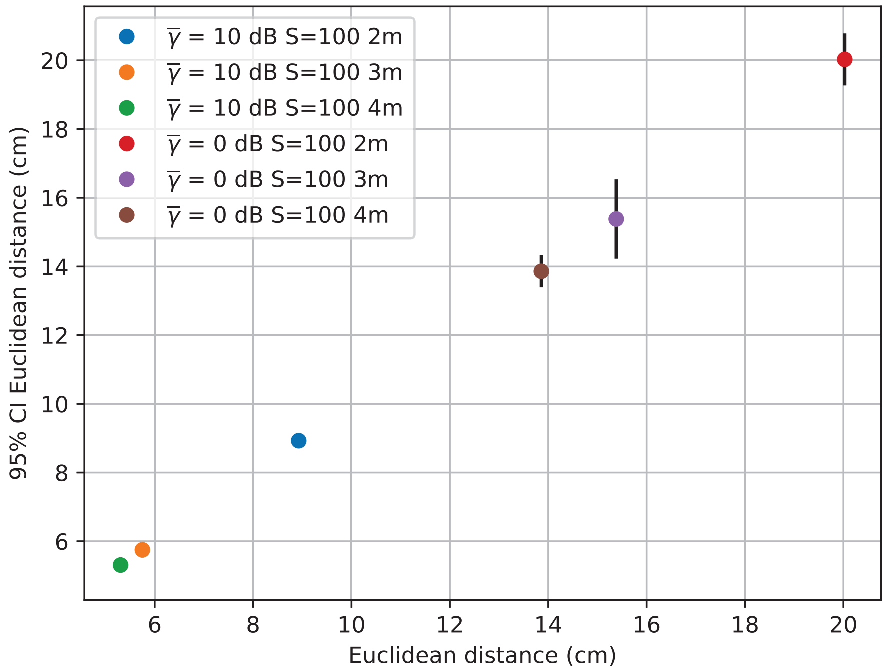

Figure 4.5: 95% Confidence Interval of distance error for radio image sensing for $R = 2$ robots, with fixed LIS aperture of $M = 259 \times 259$ and $S = 100$ samples

4.3.2. Distance among robots impact

Finally, we leverage the system performance when robots are separated at a certain distance among each other. We fix $R = 2$ robots transmitting their sensing signal in a $\overline{\gamma} = 10/0$ dB scenario.

Figure 4.5 shows the 95% confidence interval of the Euclidean distance between the predicted and the groundtruth positions in the test set for $R = 2$ different robot separations, showing the error in both axis. It shows that there is a common behavior regardless of the $\overline{\gamma}$ conditions. In this plot, the point represents the mean while the bars represent the variance w.r.t to it. Here, the 2 m case gets the worst performance and beyond 3 m the performance is almost the same. The highest mean error (2 m) is around 9 cm for the most suitable conditions case, which is an accurate performance. Similarly, in the worst conditions, the mean error is around 20 cm, which is also a good accuracy. Variance of the results augment in the poorest $\overline{\gamma}$ due to the impact of noise and interference due to the proximity of robot transmissions. Please note, a distance between robots lower than 2m may have no interest in the industrial setting. In terms of the detection, an mAP = 100% was obtained.[2]

4.4. Conclusion

[2]For some related works, please refer to [7]

Large Intelligent Surfaces are a key ingredient in current studies for improving communications in the forthcoming 6G paradigm. However, one of the main characteristics of 6G resides in the ability of sensing. LIS are a really useful tool for capturing the radio propagation environment as we have demonstrated this technology allows to represent the sense information as an image, something that decreases complexity and permits to use image processing tools for studying the characteristics of the radio environment. What is more, the presented use case shows that machine learning algorithms, concretely computer vision ones, are a powerful tool to take into account when using an image-based LIS sensing approach. Future lines can be of interest, for example, location assisted analog beamforming according to the user position radio map. Besides, a further analysis of the images can be of interest for determining the optimal pattern of antenna elements deployment along the whole surface. Please note the inter-antenna spacing might affect the performance once the number of elements is too few so the expected circular pattern is no longer recognizable. Autoencoders for image super resolution [7] can be of interest for reducing the physical aperture of the LIS deployed in the ceiling.

5. AAS Calibration

Successful deployment of novel Advanced Antenna Systems (AAS) can not be accomplished without correcting RF impairments within the hardware. These impairments include phase noise, active and passive intermodulation distortions, in-phase/quadrature imbalance and manufacturing imperfections of antenna arrays [44–47]. In the latter case, RF connections between Digital-to-Analog Converter (DAC) / Analog-to-Digital Converter (ADC) units and radiating elements can have varied lengths (group delays, phase delays) and other defects, resulting in different multiplicative complex gain and phase distortions across the bandwidth of the signal. Additionally, erroneous antenna element placement causes mutual coupling effects. As a result, in Time Division Duplexing (TDD) systems, Downlink (DL) and Uplink (UL) wireless channel responses (within the same coherence interval) become different and the reciprocity property no longer holds, which complicates the wireless channel estimation process [48].

These hardware defects also negatively affect Frequency Division Duplexing (FDD), Analog Beamforming (ABF), RIS and other systems. Notably, the general fidelity of beamforming patterns [47,49] and the quality of source detection with Angle of Arrival (AoA) estimation algorithms [48,50] degrade due to such hardware defects. The compensation of these specific distortions falls into domain of Antenna Array Calibration (AAC) models and algorithms.

5.1. Related Works and Historical Development

The early concepts of defining AAC, including mutual-coupling and blind methods, have first appeared in context of military radar arrays.

Efstathopoulos and Manikas, for example, improved the concept of Principal Component Analysis (PCA)-based blind array calibration by reducing the required amount of external UL signals to a single moving UE. This scheme only assumed the presence of a single angular component within the movement of the UE [51]. Such correction takes into account all RF impairments of the array without requiring any additional form of signal overhead. While this operating principle makes the proposed scheme applicable in both TDD and FDD systems, it does not support of Transmitter (TX) calibration, and its computational complexity significantly increases with the number of Receivers (RXs).

Mutual coupling (also known as self-calibration) methods represent a majority of recent works in the field. Luo *et al.* provided a comprehensive analysis of this methodology. They formulated the problem of finding an interconnecting network of mutual coupling measurements, and employed combinatorial optimization to prove that the “star” interconnection network is optimal for full calibration of AAS with an arbitrary geometry [48].

Wei *et al.* further exploited the Over-the-Air (OTA) concept by relying on UL training sequences to calibrate a network of analog phase shifters. Their solution allows to estimate phase deviations with affordable computational complexity, and propose algorithms that include Cramer-Rao lower bound estimations [49]. In a similar work, Tian *et al.* proposed analytical estimators for both amplitude scaling and phase drifts within RF chains [50]. Shan *et al.* applied an autoencoder Deep Neural Network (DNN) to the DL pilot matrix, which can reconstruct the measured RF impairments [52]. Similarly to other DNN solutions, there is no proof that solution will generalise to systems different from the one investigated by the authors.

Moon *et al.* addressed disadvantages of other OTA methods, by leveraging existing wireless channel estimation protocols [47]. In that paper, calibration is achieved by continuously changing the pre-coding values and collecting all the estimated phase responses with a shared reference. Specifically, the phase response of the array is measured while changing the phase shifter states, which allows to estimate the phase mismatches and calibrate the phased array. However, it is unclear how well this algorithm scales with the number of antenna elements and complexity of the precoder. Furthermore, there is a "cold-start" problem related to the convergence of the calibration algorithm in the context of dynamic environments, because the precoder has to be continuously updated to support moving UE.

In the works related to AAS calibration system model of impairments often assumes absence of Intermodulation Distortions (IMD). This is a valid assumption if we consider that algorithm units such as DPD and PIMC are present within the system. The first compensates active IMD caused by High Power Amplifier (HPA) [53, 54], in TX of both TDD and FDD systems. The second corrects passive IMD in RXs of FDD systems that are generated in defective components of RXs RF chains[1].

Predominant number of DPD and PIMC algorithms are model-based, where said models are defined by Linear Time Invariant (LTI) and nonlinear units (e.g. static memory polynomials, adaptive functions controlled by lookup table (LUT)). AAC can also be addressed in a similar way.

5.2. System Identification Perspective

But what defines system identification? By definition, it is a process, in which an interpretable model of a system with unknown dynamics and parameters is learned from labelled training examples [55, 56]. System identification also intertwines with statistical learning models and methods for time-series predictive analysis [57], such as Gaussian processes (GPs) [58] and other Bayesian nonparametric (BNP) methods. The latter becomes especially important if we want to predict outputs, or derive robust control algorithms for nonlinear dynamical systems, as BNP methods provide a unified view and tools for robust predictive models with uncertainty quantification properties. Examples of such nonlinear systems are diverse and range from robotic control to biological and wireless applications [56, 59–61].

In some cases it is possible to use prior knowledge to determine the model and then estimate its parameters from the data. For example, time series prediction models such as the nonlinear autoregressive exogenous model and the state-space model can be used to learn the dynamic and measurement models, while temporal versions of GPs can be directly used to form a regression on the data in the time domain. Alternatively, parametric models (e.g., DNNs) can be learnt automatically via an architecture search [62–64]. In other cases, explicit prior information is unavailable, limiting further application of known models or estimation methods. In the context of system identification, this scenario becomes even more complex if it is not possible to choose or control the input (excitation) signals.

[1] For more information visit project deliverable: *D4.1.A – Joint Probabilistic Modelling of Wireless Channels and Hardware Impairments*.

In this section we investigate a problem, where the structure of a nonlinear grey-box system together with it's parameters must be learnt in a data-driven fashion. We solve this task by approaching it from an optimisation perspective, and applying the framework of Bayesian optimisation of combinatorial structures (BOCS). Originally proposed by Baptista et al. [65] – it allows to obtain an approximate optimiser of the acquisition function, employing ideas from convex optimization [66, 67] and capturing the interaction of structural elements. We demonstrate its performance on several cases and discuss how the BOCS technique can be extended to support structure learning in black-box systems.

5.3. Problem Setting

We consider an open-loop identification problem of a nonlinear grey-box block-structured system. This system consists of two unit types: linear time-invariant and static nonlinear functions[2] [68] (figure 5.1 (a)). The system is generated randomly and unit connections are defined by an adjacency matrix $\mathbf{S}$ of directed acyclic graph. In it (5.1) each row represents feed-forward connection from one of above-mentioned units to others and zeros/ones indicate existence or absence of connections respectively. Each unit is parameterised randomly by a set of weights drawn from a Gaussian distribution – p_k, where k is the unit's number. This system is also assumed to be single input single output[3] and stable[4].
We are given vectors of N data points $\mathbf{X}_N, \mathbf{Y}_N = \{\mathbf{x}_n, y_n\}_{n=1}^N$, where x are delayed input samples of a fixed dimension and y are system's output scalar values [69]. Prior information includes only exact number and types of units within a system.
The task is twofold – estimate the adjacency matrix $\hat{\mathbf{S}}$ and system parameters $\{\hat{p}\}_{k=1}^K$ given a limited number of observations.

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.1}$$

5.4. Bayesian Optimization of Combinatorial Structures

Consider an expensive-to-evaluate black-box function f over a discrete structured domain $\mathcal{D}$ of feasible points, where the goal to find the goal is to find a global optimizer $\mathrm{argmax}_{x\in\mathcal{D}} f(x)$.

[2] In polynomial form.
[3] BOCS methodology has a theoretical support for multiple inputs and outputs (MIMO) settings. On practice it would require to firstly, understand (through derivations) how to fit MIMO into the structural model of (5.1). Secondly derive additional constraints to limit the complexity of the problem.
[4] Stability is assured during the generation process within the simulation environment.

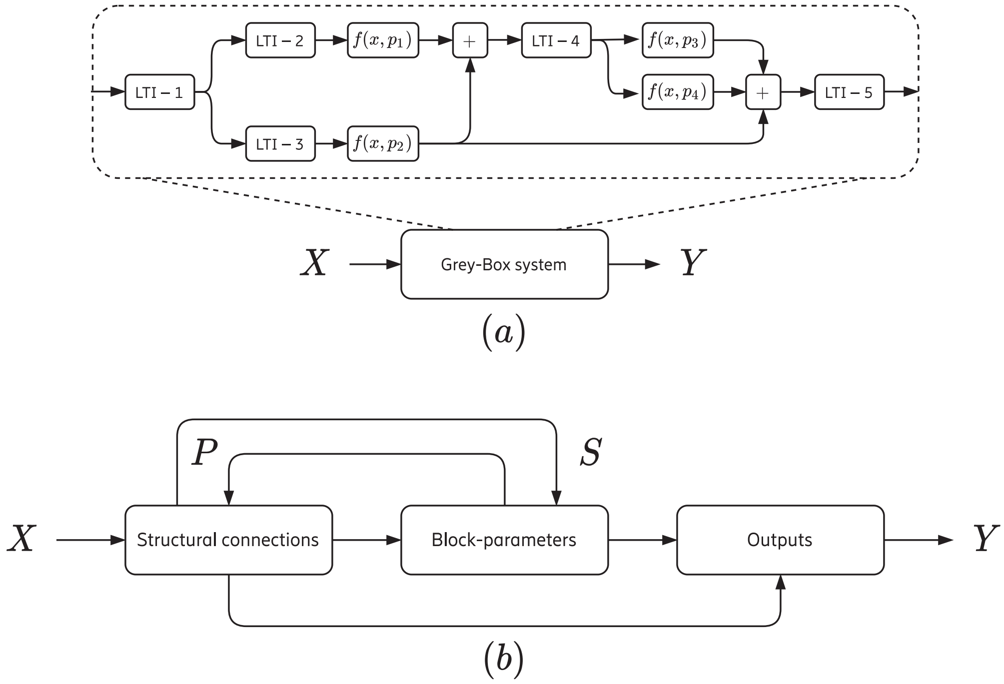

Figure 5.1: (*a*) – Example of grey-box system (9 units); (*b*) – Structural model within the BOCS framework; arrows indicate the flow of information between components.

Suppose that observing x provides independent, conditional on $f(x)$, and normally distributed observations with mean $f(x)$ and finite variance σ^2. For simplicity, assume that $\mathcal{D} = \{0, 1\}^d$, where x_i equals one if a certain element i is present in the design and zero otherwise. For example, we can associate a binary variable with possible coupling between two components in a multicomponent system, or with an edge in a graph-like structure.
The statistical model of BOCS is based on equation 5.2. The model consists of so-called "interaction" terms, quadratic in $x \in \mathcal{D}$, while regression model is linear in $\boldsymbol{\alpha} = (\alpha_i, \alpha_{ij}) \in R^p$ with $p = 1 + d + \binom{d}{2}$.

$$f_\alpha(x) = \alpha_0 + \sum_j \alpha_j x_j + \sum_{i,j>i} \alpha_{ij} x_i x_j \tag{5.2}$$

To quantify the uncertainty in the model, Bayesian treatment can be applied to $\boldsymbol{\alpha}$. For observations $\left(x^{(i)}, y^{(i)}\left(x^{(i)}\right)\right)$ with $i = 1, \ldots, N$, let $\mathbf{X} \in \{0, 1\}^{N \times p}$ be the matrix of predictors and $\mathbf{y} \in R^N$ the vector of corresponding observations of f. Using the data model, $y^{(i)}\left(x^{(i)}\right) = f\left(x^{(i)}\right) + \varepsilon^{(i)}$ where $\varepsilon^{(i)} \sim \mathcal{N}\left(0, \sigma^2\right)$, we have $\mathbf{y} \mid \mathbf{X}, \boldsymbol{\alpha}, \sigma^2 \sim \mathcal{N}\left(\mathbf{X}\boldsymbol{\alpha}, \sigma^2 I_N\right)$.
One disadvantage of using such model is that it has $\Theta\left(d^2\right)$ regression coefficients which may result in high-variance estimators for the coefficients in case of scarce data. To assert a good performance even for high-dimensional problems with expensive evaluations, it has been proposed to use a sparsity-inducing prior, specifically a heavy-tailed horseshoe prior [70],

$$\begin{aligned} \alpha_k \mid \beta_k^2, \tau^2, \sigma^2 &\sim \mathcal{N}\left(0, \beta_k^2 \tau^2 \sigma^2\right) \quad k = 1, \ldots, p \\ \tau, \beta_k &\sim \mathcal{C}^+(0, 1) \quad k = 1, \ldots, p \\ P\left(\sigma^2\right) &= \sigma^{-2} \end{aligned} \tag{5.3}$$

where $\mathcal{C}^{+}(0,1)$ is the standard half-Cauchy distribution. In this model, the global (τ) and the local (β_k) hyper-parameters individually shrink the magnitude of each regression coefficient. Additionally, auxiliary variables ν and ξ are introduced to re-parameterize the half-Cauchy densities using inverse-gamma distributions. Finally conditional posterior distributions for the parameters are defined as:

$$\begin{aligned}
\boldsymbol{\alpha} \mid \cdot &\sim \mathcal{N}\left(\mathbf{A}^{-1}\mathbf{X}^T\mathbf{y}, \sigma^2\mathbf{A}^{-1}\right)\\
\mathbf{A} &= \left(\mathbf{X}^T\mathbf{X} + \Sigma_*^{-1}\right), \Sigma_* = \tau^2 \operatorname{diag}\left(\beta_1^2, \dots, \beta_p^2\right)\\
\sigma^2 \mid \cdot &\sim IG\left(\frac{N+p}{2}, \frac{(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})}{2} + \frac{\boldsymbol{\alpha}^T\Sigma_*^{-1}\boldsymbol{\alpha}}{2}\right)\\
\beta_k^2 \mid \cdot &\sim IG\left(1, \frac{1}{\nu_k} + \frac{\alpha_k^2}{2\tau^2\sigma^2}\right) \quad k = 1, \dots, p\\
\tau^2 \mid \cdot &\sim IG\left(\frac{p+1}{2}, \frac{1}{\xi} + \frac{1}{2\sigma^2}\sum_{k=1}^{p}\frac{\alpha_k^2}{\beta_k^2}\right)\\
\nu_k \mid \cdot &\sim IG\left(1, 1 + \frac{1}{\beta_k^2}\right) \quad k = 1, \dots, p\\
\xi \mid \cdot &\sim IG\left(1, 1 + \frac{1}{\tau^2}\right).
\end{aligned} \tag{5.4}$$

The role of the acquisition function is to select the next sample point in every iteration. In case of BOCS, it is based on Thompson sampling (samples a point x with probability proportional to x being an optimizer of the unknown function). Since belief on the objective f at any iteration is given by the posterior on $\boldsymbol{\alpha}$, we sample $\boldsymbol{\alpha}_t \sim P(\boldsymbol{\alpha} \mid \mathbf{X}, \mathbf{y})$ and want to find an argmax $_{x\in\mathcal{D}} f_{\alpha_t}(x)$. Since applications often impose some form of regularization on x, problem is restated as argmax $_{x\in\mathcal{D}} f_{\alpha}(x) - \lambda\mathcal{P}(x)$, where $\mathcal{P}(x) = \|x\|_1$ or $\mathcal{P}(x) = \|x\|_2^2$ and thus cheap to evaluate. Then, for a given $\boldsymbol{\alpha}$ and $\mathcal{P}(x) = \|x\|_1$, the problem is to obtain an

$$\begin{aligned}
&\underset{x\in\mathcal{D}}{\operatorname{argmax}} \quad f_{\alpha}(x) - \lambda\mathcal{P}(x)\\
=&\underset{x\in\mathcal{D}}{\operatorname{argmax}} \quad \sum_j \left(\alpha_j - \lambda\right) x_j + \sum_{i,j>i} \alpha_{ij}x_ix_j,
\end{aligned} \tag{5.5}$$

where $x \in \{0,1\}^d$. Similarly, if $\mathcal{P}(x) = \|x\|_2^2$, the problem becomes argmax $_{x\in\mathcal{D}} \sum_j \alpha_j x_j + \sum_{i,j>i}\left(\alpha_{ij} - \lambda\delta_{ij}\right) x_i x_j$ Thus, in both cases we are to solve a binary quadratic program of the form

$$\underset{x\in\mathcal{D}}{\operatorname{argmax}} \quad x^TAx + b^Tx, \tag{5.6}$$

where $\mathcal{D} = \{0,1\}^d$.
Finally,the BOCS algorithm can be summarised as follows. Using an initial dataset of N_0 samples, BOCS first computes the posterior on f based on the sparsity-inducing prior. In the optimization phase, BOCS proceeds in iterations until the sample budget $N_{\max}$ is exhausted. During iterations $t = 1, 2, \dots$, it samples the vector $\boldsymbol{\alpha}_t$ from the posterior over the regression coefficients that is defined by the parameters in equation 5.4.
Afterwards, BOCS computes an approximate solution $x^{(t)}$ for $\max_{x\in\{0,1\}^d} f_{\alpha_t} - \lambda\mathcal{P}(x)$ as follows: first it transforms the quadratic model into an SDP, thereby relaxing the variables into

vector-valued variables on the $(d + 1)$ dimensional unit-sphere. This SDP is solved (with a predescribed precision) and the next point $x^{(t)}$ is obtained by rounding the vector-valued SDP solution. The iteration ends after the posterior is updated with the new observation $y^{(t)}$ at $x^{(t)}$.

5.5. Simulation Results

Similarly to other system identification problems, we consider a mean squared error (MSE) cost function defined in (5.7), where $\hat{y}(t \mid t-1, \theta)$ is the predictor of the output defined by the model structure and parameters.

$$F_N\left(\mathbf{X}_N, \mathbf{Y}_N, \hat{\mathbf{P}}, \hat{\mathbf{S}}\right) = \frac{1}{N}\sum_{t=1}^{N}(y(t) - \hat{y}(t \mid t-1, \hat{\mathbf{P}}, \hat{\mathbf{S}}))^2 \tag{5.7}$$

Here, the BOCS is an appropriate choice for the task. Since local search algorithms do not necessarily converge to a global optimum and mathematical programming (e.g., convex programming) typically cannot be applied to black-box functions nonlinear in parameters. Furthermore we can exploit combinatorial structure of the problem (figure 1-b). Figure 5.2 shows the BOCS convergence process for two randomly generated grey-box systems. In both cases, the optimum is reached within a limited number of function evaluations. Additionally, we must also note that extended set of simulations showed stability in MSE. Upon convergence, it remains within $10^{-5} - 10^{-7}$ range.

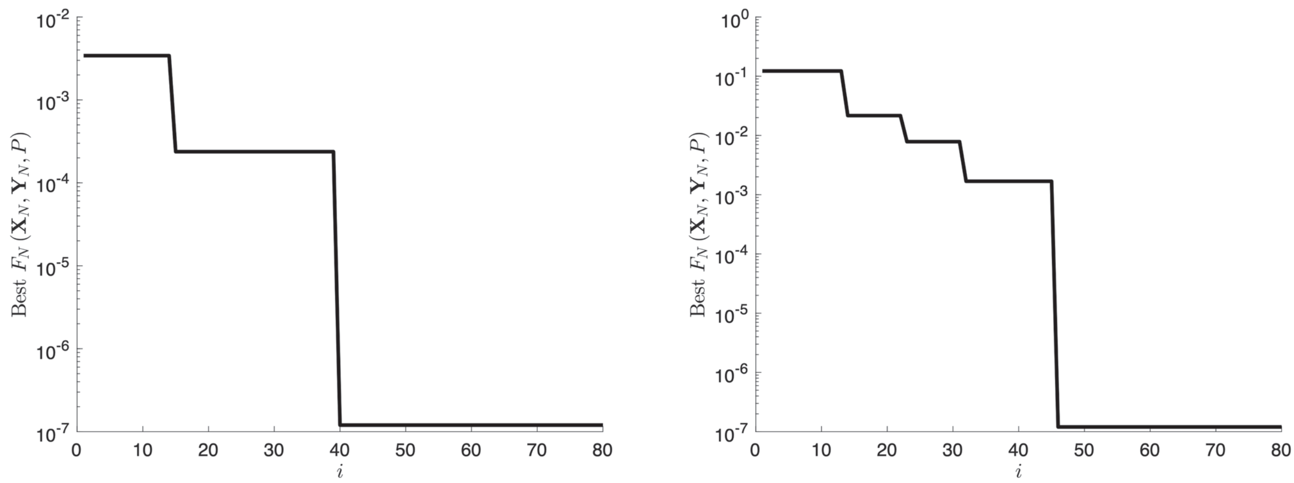

Figure 5.2: BOCS convergence results for grey-box system of different complexity: (*a*) 15 units, 100 parameters, (*b*) – 30 units, 200 parameters.[To be published in BAYSM 2021 proceedings]

5.6. Conclusion and Future Research

We have studied the performance of the BOCS framework for grey-box parameter estimation. Consequently this research serves as a foundation for study of methodology's applicability to black-box systems. Specifically, how well it will perform in the task of Neural Architecture Search (NAS) compared to alternatives, such as those described in works of Zhou et al. [71] and others [72–76] is an open question. Other areas of research include BOCS in online learning setting, where the structure of the system can change over time and the multi-objective case, in which the system system has multiple inputs and outputs. We think that the solution to these questions may be a combination of BOCS with other existing frameworks, for example constrained or "lifelong" Bayesian optimisation [77, 78].

6. References

[1] M. Giordani, M. Mezzavilla, C. N. Barati, S. Rangan, and M. Zorzi, "Comparative analysis of initial access techniques in 5G mmwave cellular networks," in *2016 Annual Conference on Information Science and Systems (CISS)*, pp. 268–273, 2016.

[2] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-wave vehicular communication to support massive automotive sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, 2016.

[3] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[4] C. Studer, S. Medjkouh, E. Gönültaş, T. Goldstein, and O. Tirkkonen, "Channel charting: Locating users within the radio environment using channel state information," *IEEE Access*, vol. 6, pp. 47682–47698, 2018.

[5] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Large intelligent surface (lis)-based communications: New features and system layouts," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2020.

[6] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Towards large intelligent surface (LIS)-based communications," *IEEE Transactions on Communications*, vol. 68, no. 10, pp. 6568–6582, 2020.

[7] C. J. Vaca-Rubio, P. Ramirez-Espinosa, K. Kansanen, Z.-H. Tan, E. De Carvalho, and P. Popovski, "Assessing wireless sensing potential with large intelligent surfaces," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 934–947, 2021.

[8] V. Va, J. Choi, and R. W. Heath, "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5014–5029, 2016.

[9] L. Wei, Q. Li, and G. Wu, "Initial access techniques for 5G NR: Omni/beam SYNC and RACH designs," in *2018 International Conference on Computing, Networking and Communications (ICNC)*, pp. 249–253, IEEE, 2018.

[10] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Multilevel millimeter wave beamforming for wireless backhaul," in *2011 Ieee Globecom Workshops*, pp. 253–257, IEEE, 2011.

[11] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath, "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4042–4058, 2018.

[12] H. Kuisma, "Inertial sensors for automotive applications," in *Transducers' 01 Eurosensors XV*, pp. 430–433, Springer, 2001.

[13] M. Brambilla, M. Nicoli, S. Savaresi, and U. Spagnolini, "Inertial sensor aided mmWave beam tracking to support cooperative autonomous driving," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, IEEE, 2019.

[14] I. Mavromatis, A. Tassi, R. J. Piechocki, and A. Nix, "Beam alignment for millimeter wave links with motion prediction of autonomous vehicles," in *Antennas, Propagation RF Technology for Transport and Autonomous Platforms 2017*, pp. 1–8, 2017.

[15] N. González-Prelcic, R. Méndez-Rial, and R. W. Heath, "Radar aided beam alignment in mmWave V2I communications supporting antenna diversity," in *2016 Information Theory and Applications Workshop (ITA)*, pp. 1–7, 2016.

[16] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1038–1052, 2018.

[17] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pp. 1–5, IEEE, 2020.

[18] A. Klautau, N. González-Prelcic, and R. W. Heath, "LIDAR data for deep learning-based mmWave beam-selection," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 909–912, 2019.

[19] M. Dias, A. Klautau, N. González-Prelcic, and R. W. Heath, "Position and LIDAR-aided mmWave beam selection using deep learning," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, 2019.

[20] M. B. Mashhadi, M. Jankowski, T.-Y. Tung, S. Kobus, and D. Gunduz, "Federated mmWave beam selection utilizing LIDAR data," *arXiv:2102.02802v1*, 2021. arXiv:2102.02802v1.

[21] M. Zecchin, M. B. Mashhadi, M. Jankowski, D. Gunduz, M. Kountouris, and D. Gesbert, "A novel look at lidar-aided data-driven mmwave beam selection," *arXiv preprint arXiv:2104.14579*, 2021.

[22] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *NIPS 2014 Deep Learning Workshop.*, 2014.

[23] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7794–7803, 2018.

[24] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proceedings of the 26th annual international conference on machine learning*, pp. 41–48, 2009.

[25] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, 2019.

[26] S. Jayaprakasam, X. Ma, J. W. Choi, and S. Kim, “Robust beam-tracking for mmwave mobile communications,” *IEEE Communications Letters*, vol. 21, no. 12, pp. 2654–2657, 2017.

[27] M. Polese, M. Mezzavilla, and M. Zorzi, “Performance comparison of dual connectivity and hard handover for LTE-5G tight integration,” *arXiv preprint arXiv:1607.05425*, 2016.

[28] M. Arvinte, M. Tavares, and D. Samardzija, “Beam management in 5G NR using geolocation side information,” in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, 2019.

[29] P. Kazemi, H. Al-Tous, C. Studer, and O. Tirkkonen, “SNR prediction in cellular systems based on channel charting,” in *2020 IEEE Eighth International Conference on Communications and Networking (ComNet)*, pp. 1–8, 2020.

[30] J. Deng, S. Medjkouh, N. Malm, O. Tirkkonen, and C. Studer, “Multipoint channel charting for wireless networks,” in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pp. 286–290, Oct. 2018.

[31] M. Kurras, S. Dai, S. Jaeckel, and L. Thiele, “Evaluation of the spatial consistency feature in the 3GPP geometry-based stochastic channel model,” in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, Apr. 2019.

[32] S. Akodad, L. Bombrun, J. Xia, Y. Berthoumieu, and C. Germain, “Ensemble learning approaches based on covariance pooling of CNN features for high resolution remote sensing scene classification,” *Remote Sensing*, vol. 12, no. 20, 2020.

[33] P. Kazemi, T. Ponnada, H. Al-Tous, Y.-C. Liang, and O. Tirkkonen, “Channel charting based beam SNR prediction,” in *2021 Joint European Conference on Networks and Communications 6G Summit (EuCNC/6G Summit)*, pp. 72–77, 2021.

[34] T. Ponnada, P. Kazemi, H. Al-Tous, Y.-C. Liang, and O. Tirkkonen, “Best beam prediction in non-standalone mm wave systems,” in *2021 Joint European Conference on Networks and Communications 6G Summit (EuCNC/6G Summit)*, pp. 532–537, 2021.

[35] M. T. Hagan and M. B. Menhaj, “Training feedforward networks with the marquardt algorithm,” *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.

[36] E. Björnson and L. Sanguinetti, “Power scaling laws and near-field behaviors of massive mimo and intelligent reflecting surfaces,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1306–1324, 2020.

[37] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE transactions on wireless communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[38] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, “Massive mimo is a reality—what is next?: Five promising research directions for antenna arrays,” *Digital Signal Processing*, vol. 94, pp. 3–20, 2019.

[39] S. Hu, F. Rusek, and O. Edfors, “Cramér-rao lower bounds for positioning with large intelligent surfaces,” in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–6, IEEE, 2017.

[40] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[41] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[42] “Feko, altair engineering, inc.” `https://www.altairhyperworks.com/feko`.

[43] P. Roy, S. Ghosh, S. Bhattacharya, and U. Pal, “Effects of degradations on deep neural network architectures,” *arXiv preprint arXiv:1807.10108*, 2018.

[44] Z. Liu, C.-h. Lee, W. Xu, and S. Li, “Energy-Efficient Design for Massive MIMO With Hardware Impairments,”

[45] X. Jiang, A. Decurninge, K. Gopala, F. Kaltenberger, M. Guillaud, D. Slock, and L. Deneire, “A Framework for Over-the-Air Reciprocity Calibration for TDD Massive MIMO Systems,” vol. 17, no. 9, pp. 5975–5990.

[46] K. Gopala, C. K. Thomas, and D. Slock, “Sparse Bayesian Learning for a Bilinear Calibration Model and Mismatched CRB,” in *2019 27th European Signal Processing Conference (EUSIPCO)*, pp. 1–5.

[47] T. Moon, J. Gaun, and H. Hassanieh, “Know Your Channel First, then Calibrate Your mmWave Phased Array,” vol. 37, no. 4, pp. 42–51.

[48] X. Luo, F. Yang, and H. Zhu, “Massive MIMO Self-Calibration: Optimal Interconnection for Full Calibration,” vol. 68, no. 11, pp. 10357–10371.

[49] X. Wei, Y. Jiang, Q. Liu, and X. Wang, “Calibration of Phase Shifter Network for Hybrid Beamforming in mmWave Massive MIMO Systems,” vol. 68, pp. 2302–2315.

[50] G. Tian, H. Tataria, and F. Tufvesson, “Amplitude and Phase Estimation for Absolute Calibration of Massive MIMO Front-Ends,” in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1–7.

[51] G. Efstathopoulos and A. Manikas, “A blind array calibration algorithm using a moving source,” in *2008 5th IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 455–458.

[52] C. Shan, X. Chen, H. Yin, W. Wang, G. Wei, and Y. Zhang, “Diagnosis of Calibration State for Massive Antenna Array via Deep Learning,” vol. 8, no. 5, pp. 1431–1434.

[53] W. Sneijers, “Doherty Architectures in UHF White Paper,” p. 14.

[54] R. Pengelly, C. Fager, and M. Ozen, “Doherty’s Legacy: A History of the Doherty Power Amplifier from 1936 to the Present Day,” vol. 17, no. 2, pp. 41–58.

[55] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning Series, The MIT Press, second edition ed.

[56] J. Schoukens and L. Ljung, “Nonlinear System Identification: A User-Oriented Road Map,” vol. 39, no. 6, pp. 28–99.

[57] S. Roberts, M. Osborne, M. Ebden, S. Reece, N. Gibson, and S. Aigrain, “Gaussian processes for time-series modelling,” vol. 371, no. 1984, p. 20110550.

[58] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning, MIT Press.

[59] T. B. Burghi, M. Schoukens, and R. Sepulchre, “System identification of biophysical neuronal models *,” in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 6180–6185.

[60] F. M. Ghannouchi and O. Hammi, “Behavioral modeling and predistortion,” vol. 10, no. 7, pp. 52–64.

[61] R. Calandra, S. Ivaldi, M. P. Deisenroth, and Peters, “Learning torque control in presence of contacts using tactile sensing from robot skin,” in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pp. 690–695.

[62] C. Cortes, X. Gonzalvo, V. Kuznetsov, M. Mohri, and S. Yang, “AdaNet: Adaptive Structural Learning of Artificial Neural Networks,” in *International Conference on Machine Learning*, pp. 874–883, PMLR.

[63] R. Y. Rohekar, Y. Gurwicz, S. Nisimov, and G. Novik, “Modeling Uncertainty by Learning a Hierarchy of Deep Neural Connections.”

[64] T. Elsken, J. H. Metzen, and F. Hutter, “Neural Architecture Search: A Survey.”

[65] R. Baptista and M. Poloczek, “Bayesian Optimization of Combinatorial Structures,” in *International Conference on Machine Learning*, pp. 462–471, PMLR.

[66] F. Hutter, H. H. Hoos, and K. Leyton-Brown, “Automated Configuration of Mixed Integer Programming Solvers,” in *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems* (A. Lodi, M. Milano, and P. Toth, eds.), Lecture Notes in Computer Science, pp. 186–202, Springer.

[67] S. Arora and S. Kale, “A combinatorial, primal-dual approach to semidefinite programs,” in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC ’07, pp. 227–236, Association for Computing Machinery.

[68] E. Zhang, M. Schoukens, and J. Schoukens, “Structure Detection of Wiener–Hammerstein Systems With Process Noise,” vol. 66, no. 3, pp. 569–576.

[69] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, illustrated edition ed.

[70] C. M. CARVALHO, N. G. POLSON, and J. G. SCOTT, “The horseshoe estimator for sparse signals,” *Biometrika*, vol. 97, no. 2, pp. 465–480, 2010.

[71] H. Zhou, M. Yang, J. Wang, and W. Pan, "BayesNAS: A Bayesian Approach for Neural Architecture Search," in *International Conference on Machine Learning*, pp. 7603–7613, PMLR.

[72] C. White, W. Neiswanger, and Y. Savani, "BANANAS: Bayesian Optimization with Neural Architectures for Neural Architecture Search."

[73] B. Ru, X. Wan, X. Dong, and M. Osborne, "Interpretable Neural Architecture Search via Bayesian Optimisation with Weisfeiler-Lehman Kernels."

[74] A. Camero, H. Wang, E. Alba, and T. Bäck, "Bayesian Neural Architecture Search using A Training-Free Performance Metric," vol. 106, p. 107356.

[75] V. Dalibard, M. Schaarschmidt, and E. Yoneki, "BOAT: Building Auto-Tuners with Structured Bayesian Optimization," in *Proceedings of the 26th International Conference on World Wide Web*, pp. 479–488, International World Wide Web Conferences Steering Committee.

[76] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," p. 9.

[77] B. Letham, B. Karrer, G. Ottoni, and E. Bakshy, "Constrained Bayesian Optimization with Noisy Experiments."

[78] V. Nguyen, S. Schulze, and M. A. Osborne, "Bayesian Optimization for Iterative Learning," p. 11.

www.windmill-itn.eu