Thesis for the degree of Licentiate of Engineering

Channel Gain Prediction for Cooperative Multi-Agent Systems

Markus Fröhle

CHALMERS

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology

Gothenburg, Sweden 2016

Fröhle, Markus
Channel Gain Prediction for Cooperative Multi-Agent Systems

Department of Signals and Systems
Technical Report No. R005/2016
ISSN 1403-266X

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Telephone: + 46 (0)31-772 1865
Email: frohle@chalmers.se

This thesis has been prepared using LaTeX.

Printed by Chalmers Reproservice,
Göteborg, Sweden, October 2016.

Abstract

In a cooperative multi-agent system (MAS), agents communicate with each other using the wireless medium. As agents move in the environment in order to fulfill the MAS' higher level task, their location changes and so does the wireless communication channel they experience. To enable a successful coordination, it is paramount for the agents to retain connectivity among themselves. In order to achieve this, the availability of explicit channel knowledge for the MAS' future configuration is needed. Since the agents determine their location from sensors, any expected residual location uncertainty for the MAS' future configuration will have an implication on the channel knowledge. For this reason, a computationally attractive yet accurate method to predict the wireless ad-hoc communication channel for any configuration and location uncertainty of the agents is needed.

In this thesis, we employ Gaussian processes (GPs) for learning of channel model parameters and for channel prediction at arbitrary (unvisited) transmitter (TX) and receiver (RX) locations. In an indoor measurement campaign, we investigate the ad-hoc wireless communication channel and its properties with respect to path-loss and shadowing from obstacles. We derive a suitable GP model, where we incorporate spatial correlation of communication links caused by shadowing. The effectiveness of our approach in a cooperative MAS is demonstrated, where the bit error rate (BER) among the agents' communication links is minimized. Furthermore, we extend our GP framework allowing to make distributed predictions using a consensus scheme.

We found that the incorporation of location uncertainty into channel prediction allows to outperform approaches where this is neglected. The incorporation of location uncertainty at both, the TX and the RX location, leads not only to robust estimates of the underlying channel parameters, but also to realistic channel predictions with respect to the agents' true location uncertainty. Applied to a cooperative MAS, we see that the BER and BER uncertainty can be significantly reduced. Finally, with a distributed channel prediction, we observe a trade-off between computation complexity and accuracy of prediction.

Natural extensions of our GP channel prediction framework could include distributed parameter learning and efficient methods to handle a high number of measurements.

Keywords: Gaussian processes, spatial correlation, channel prediction, parameter learning, multi-agent systems, wireless ad-hoc networks, distributed algorithms.

List of Included Publications

The thesis is based on the following appended papers:

[A] M. Fröhle, T. Charalambous, I. Nevat, and H. Wymeersch, "Channel Prediction with Location Uncertainty for Ad-Hoc Networks," submitted to *IEEE Transactions on Signal and Information Processing over Networks*, 2016.

[B] M. Fröhle, L. S. Muppirisetty, and H. Wymeersch, "Channel Gain Prediction for Multi-Agent Networks in the Presence of Location Uncertainty," in *Proceedings of 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3911–3915, Shanghai, China, Mar. 2016.

[C] V. P. Chowdappa, M. Fröhle, H. Wymeersch, and C. Botella, "Distributed Channel Prediction for Multi-Agent Systems," in preparation for *IEEE International Conference on Communications (ICC)*, 2017.

Other contributions by the author (not included in this thesis):

[D] F. Meyer, H. Wymeersch, M. Fröhle, and F. Hlawatsch, "Distributed Estimation With Information-Seeking Control in Agent Networks," in *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2439–2456, Nov. 2015.

[E] S. Zhang, M. Fröhle, H. Wymeersch, A. Dammann, and R. Raulefs, "Location-Aware Formation Control in Swarm Navigation," in *Proceedings of 2015 IEEE Globecom Workshops (GC Wkshps)*, San Diego, CA, 2015.

[F] M. Fröhle, and H. Wymeersch, "On the Separation of Timescales in Radio-based Positioning," in *Proceedings of 2015 International Conference on Location and GNSS (ICL-GNSS)*, Gothenburg, Sweden, 2015.

[G] M. Fröhle, A. A. Zaidi, E. Ström, and H. Wymeersch, "Multi-Step Sensor Selection with Position Uncertainty Constraints," in *Proceedings of 2014 IEEE Globecom Workshops (GC Wkshps)*, pp. 1439–1444, Austin, TX, 2014.

Acknowledgements

If a door [to funding] closes, somewhere else another door will open[1]. This happened in 2013, when I started my journey with the move to Sweden to join the Communication Systems group at Chalmers. I would like to take this opportunity to express my deepest gratitude to my supervisor Prof. Henk Wymeersch for his guidance, discussions, patience, and pep talks. Please, keep on doing this! Also, I would like to thank my co-supervisor Prof. Erik Ström for making the CS group really what it is.

Thanks to all the former and current COOPNET team members, in alphabetical order: Ali, Christopher, Erik, Gabo, Hamed, Jie, Mohammad, Mouhamed, Nil, Rahul, Robert, Rocco, Srikar, and Themis for the great discussions, collaboration attempts and collaborations. A great thanks to the administrative staff, especially Agneta and Katarina; as well as Lars who cannot be with us anymore. A special thanks to all my coffee break victims, the holy grail materialized in a coffee preparation apparatus, and the bubble machine for the times it actually provides sparkling water.

Furthermore, I would like to thank all the fellows and seniors within the MULTI-POS project. We had many great meetings and I hope our paths cross in the future. Special thanks to my colleagues at Honeywell and DLR for the nice work experience there.

I would like to thank my family for their endless support and trust in me. Probably I am your example for lifelong learning at academia, but I will get a real job soon, I promise. Finally, the most important thank you belongs to you, Isabel, because you chose to ride along with me in this journey!

Markus Fröhle
Gothenburg, October 2016

This work is supported, in part, by the European Research Council under grant no. 258418 (COOPNET), by EU FP7 Marie Curie Initial Training Network MULTI-POS (Multi-technology Positioning Professionals) under grant no. 316528, and by the EU-H2020 project HIGHTS (High Precision Positioning for Cooperative ITS Applications) under grant no. MG-3.5a-2014-636537.

[1] The author's free interpretation of Alexander Graham Bells famous quote.

Acronyms

BER	bit error rate
CC	command center
cdf	cumulative distribution function
C-MAS	MAS with direct communication among agents
GPS	global positioning system
I-MAS	MAS only communicating with infrastructure
LOS	line-of-sight
NLOS	non-LOS
MAP	maximum a posteriori
MAS	multi-agent system
ML	maximum likelihood
MQAM	amplitude modulation with M bits per symbol
MSE	mean squared error
pdf	probability density function
RX	receiver
SNR	signal-to-noise ratio
SR	search-and-rescue robot
TX	transmitter
UAV	unmanned aerial vehicle

Contents

Part I

Overview

Chapter 1

Introduction

1.1 Motivation

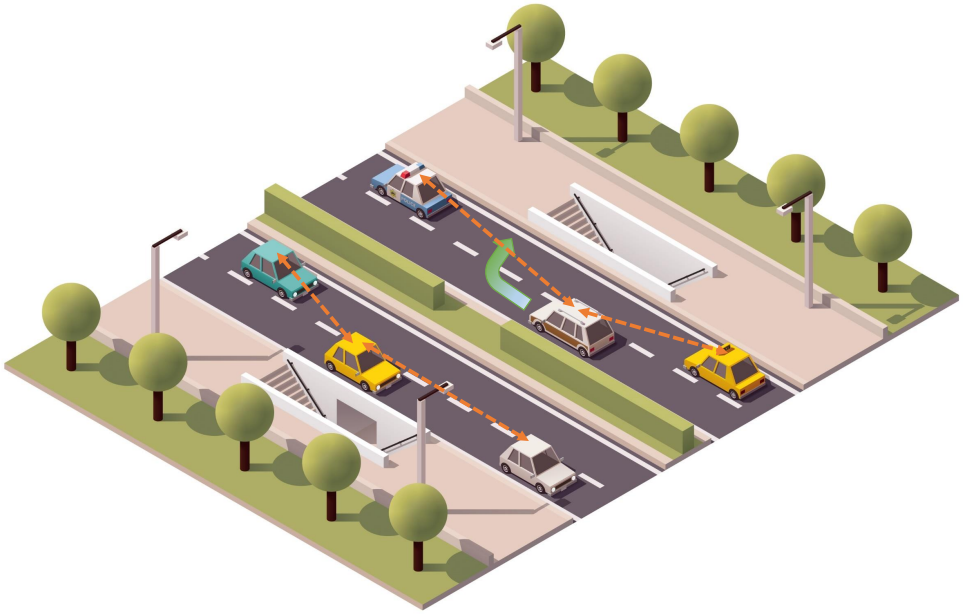
Multi-agent systems (MAS) employ the wireless medium for communication and often for positioning as well. Typical examples of MAS include mobile wheel-driven robots for surveillance and search-and-rescue operations, or unmanned aerial vehicles (UAVs) for remote sensing, e.g., to detect forest fires [1] or for chemical plume tracking [2]. We can distinguish two types of communication in such systems: (a) agents directly communicating with infrastructure (e.g., an anchor node or a command center), denoted by I-MAS and (b) cooperative MAS, denoted by C-MAS, where agents also communicate directly among each other. These two different systems are outlined in Fig. 1.1.

Typical examples in an I-MAS setting may include that agents maintain connectivity with infrastructure [3], or that data is proactively cached for a mobile user to balance channel load [4, 5]. For C-MAS, typical examples may include a swarm of UAVs navigating through an environment [1, 6], autonomous cars cooperating on sensor data or for throughput maximization at an intersection [7, 8].

Since mobile agents employ the wireless medium for communication, it is important to have accurate, yet computationally efficient, channel models allowing them to pursue their higher level tasks successfully. This includes as well channel prediction, i.e., when knowledge of the channel is needed for configurations of the agents where no measurements are available. The I-MAS wireless channel depends on the environment, the fixed anchor position, as well as the mobile agents' position. Its properties have been well studied and accurate channel and propagation models have been developed, many based on cellular communication systems [9]. C-MAS wireless channel models, on the other hand, are less developed. One reason for this is, that demand for C-MAS has been low in the past. A C-MAS channel model is more complex, since it needs to consider mobility of both link endpoints (mobile agent to mobile agent communication) in contrast to an I-MAS channel model.

In general, the wireless channel can be considered as being composed of three parts: a deterministic part due to path-loss, a part due to shadowing because of obstacles, and a final part due to multi-path [9]. Shadowing is spatially correlated and measurements

(a) Example of an I-MAS.



(b) Example of a C-MAS.

Figure 1.1: Examples of different types of MAS. In the I-MAS case, the agents (cars) communicate only with infrastructure (street light). In the C-MAS case, the agents communicate directly with each other. The wireless communication link is indicated by an orange dashed line.

have shown that it decorrelates in the order of 50–100 m outdoors [9] and 1–5 m indoors [10, 11]. Multi-path, on the other hand, decorrelates on a much smaller scale within a few wavelengths [9]. Based on this, several different channel models have been developed. The most simple model is the disc model, where communication quality is assumed *good* if the receiver (RX) is located within a radius r of the transmitter (TX) and *bad* otherwise. The next step towards a more accurate model considers deterministic path-loss. The incorporation of such simple channel models into the higher level control problem for, e.g., connectivity maintenance has been extensively studied by the control community [12–18]. More sophisticated channel models incorporate shadowing and multi-path as well. These components are modeled either in a probabilistic or a deterministic way (e.g., through ray-tracing [9]). Although more accurate, such models come with some drawbacks. For instance, a ray-tracing based channel model is not applicable in real-time applications due to the high computational demand. Furthermore, it is not always possible to estimate all model parameters from measurements, e.g., for ray-tracing this means that reflection properties of all surfaces in the environment need to be estimated.

An attractive tool for channel prediction is through the usage of Gaussian processes (GPs) [19], where deterministic path-loss and spatially correlated shadowing is modeled. Through the incorporation of measurements in the form of a measurement database, an accurate and still computationally efficient channel prediction is possible. Additionally to channel prediction, the GP framework allows to estimate the underlying model parameters from measurements. For accurate channel prediction and parameter estimation the location of the mobile agents needs to be known exactly. In reality, this is seldom the case, since mobile agents need to determine their location from sensors. For instance, for a GPS sensor the positioning quality varies with the number of visible satellites, their geometry, and the operation environment (land, urban, sub-urban, indoors). In [20], a GP based channel predictor was presented, which allows the incorporation of the agents location uncertainty in an I-MAS setting. Although useful, it cannot be applied in a C-MAS setting due to the mobility of both the TX and the RX. Furthermore, the whole measurement database needs to be available at each agent. This is not practical in a C-MAS context, where a distributed approach using only local information is preferred.

As part of this thesis, we develop a GP based channel predictor for a C-MAS setting, where we model location uncertainty of both link endpoints. This allows us not only to estimate the underlying channel parameters of ad-hoc wireless channels in a robust manner, but also to predict the channel at arbitrary (unvisited) locations. Furthermore, a distributed version of the channel predictor is developed. Distributed processing over the wireless channel is thereby achieved by a consensus scheme [21]. Our channel predictor is incorporated into a higher level optimization problem, where we aim to find the optimal configuration of mobile robots such that the bit error rate (BER) of the wireless communication network is minimized.

1.2 Thesis Outline

This thesis is divided into two parts. In the first part, we give a brief introduction on channel gain prediction in a C-MAS setting. We present in Chapter 2 ad-hoc channel

measurements and a suitable statistical channel model. This is followed by Chapter 3, where we give a general introduction on GPs. We address learning of model parameters from measurements, centralized and distributed prediction, as well as the incorporation of the presented statistical channel model allowing to explicitly model spatial correlation of shadowing. In Chapter 4, we present an application example of the proposed GP based channel prediction framework, aiming at improving the communication quality among C-MAS. Finally, in Chapter 5, we summarize our contributions. In part two of this thesis, the included papers are appended.

Chapter 2

The Wireless Ad-Hoc Propagation Channel

In this chapter, we present channel measurements and the C-MAS channel model we rely on. Towards this end, we first present the measurement campaign we have performed. This is followed by the presentation of a known empirical statistical channel model, which we relate to the collected measurements and discuss briefly its validity. Following that, we highlight two spatial correlation models, one where only the RX is mobile corresponding to an I-MAS setting, and one where TX and RX are mobile corresponding to a C-MAS setting.

2.1 Ad-Hoc Channel Measurements

In order to justify any model assumptions in the development of our channel prediction framework, we have performed a measurement campaign, where we recorded indoor ad-hoc channel measurements in a hallway at the Department of Signals and Systems at Chalmers University of Technology. For this campaign, we placed the RX at several different positions along a line following the hallway. On a hallway perpendicular to it, we placed the TX on several different positions also along a line, where the first location thereby corresponds to position 6.5 m on the RX line. The full measurement scenario is outlined in Fig. 2.1. For every position of the TX and RX, the received power in dBm was recorded using commodity hardware radios of type *Netgear N150 Wireless adapter*. The measurement parameters are stated in Table 1.

In Fig. 2.2 the received signal power in dBm is plotted over the distance between the TX and RX locations. Note that the hardware used provides the received power readings only in integer dBm values. From the figure, we can observe a power loss with increasing distance, but due to the complexity of the indoor propagation environment it is difficult to come up with an accurate and computationally efficient channel model. We are therefore interested in a simple channel model, which captures the signal propagation sufficiently well with respect to its application. It should be noted that, any such model

Table 2.1: Measurement parameters

Parameter	Value
Total number of different RX positions	10,900
Total number of different TX positions	11
Spatial resolution RX positions	0.02m
Spatial resolution TX positions	1m
Shortest TX-RX distance	1m
Transmit power P_{TX}	20dBm
Antenna height	0.85m
TX communication frequency	2.422GHz

is only an approximation of the real channel.

2.2 Statistical Channel Model

Consider a workspace $\mathcal{W} \subset \mathbb{R}^2$. The TX located at $\boldsymbol{q}_{\mathrm{TX}} \in \mathcal{W}$ emits a signal with power $P_{\mathrm{TX}}^{\mathrm{lin}}$ through the wireless channel. Until the signal is received at the RX located at $\boldsymbol{q}_{\mathrm{RX}}$, it experiences distance-dependent path-loss, large-scale fading caused by obstacles in the propagation path, and small-scale fading due to multi-path. The received signal power can be expressed as [22]

$$P_{\mathrm{RX}}^{\mathrm{lin}}(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}) = P_{\mathrm{TX}}^{\mathrm{lin}} g_0 \, ||\boldsymbol{q}_{\mathrm{TX}} - \boldsymbol{q}_{\mathrm{RX}}||_2^{-\eta} \psi(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}) \, |h(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}})|^2, \tag{2.1}$$

where g_0 is a constant capturing antenna and other propagation gains, η is the path-loss exponent, ψ is the position dependent shadowing and h captures small-scale fading. Let us assume that measurements average out small-scale fading, either in time (measurements taken over a time window), frequency (measurements represent average power over a large frequency band), or space (by using multiple antennas) the received signal power in dBm is then expressed as

$$P_{\mathrm{RX}}(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}) = L_0 - 10\eta \log_{10} ||\boldsymbol{q}_{\mathrm{TX}} - \boldsymbol{q}_{\mathrm{RX}}||_2 + \Psi(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}), \tag{2.2}$$

where constant $L_0 = P_{\mathrm{TX}} + 10 \log_{10}(g_0) + 10\eta \log_{10}(d_0)$, d_0 is a reference distance, and $\Psi(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}) = 10 \log_{10} \psi(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}})$. We assume that large-scale fading (shadowing) follows a log-normal distribution, i.e., $\Psi(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}) \sim \mathcal{N}(0, \sigma_\Psi^2)$, where σ_Ψ^2 is the shadowing variance [9].

2.3 Measurement-Model Relationship

We assume the RX provides us with a noisy observation of the received signal power in dBm such that

$$y = P_{\mathrm{RX}}(\boldsymbol{q}_{\mathrm{TX}}, \boldsymbol{q}_{\mathrm{RX}}) + n, \tag{2.3}$$

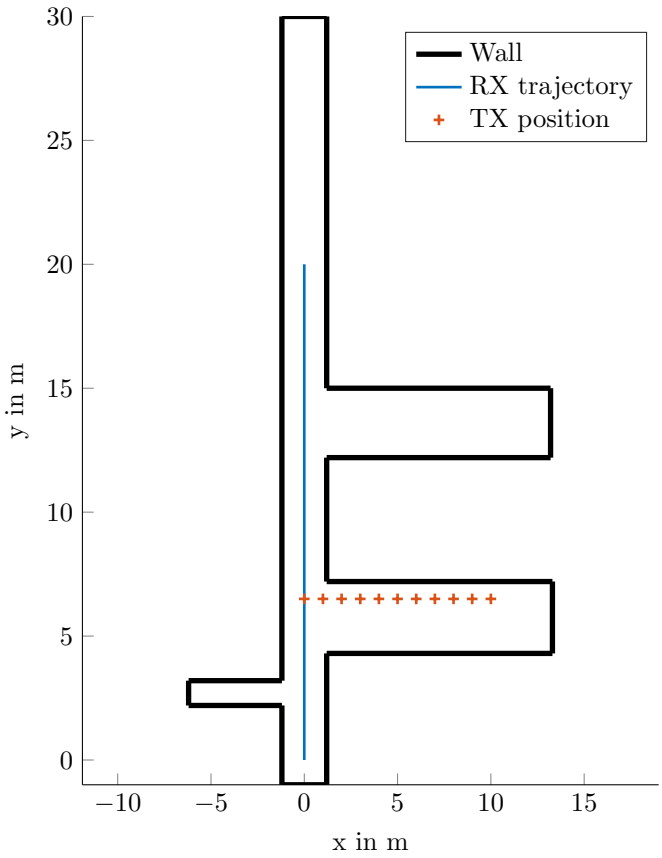


Figure 2.1: Floorplan together with TX and RX locations where channel measurements have been recorded. All walls consist of reinforced concrete. For every TX position, the RX moved along the indicated RX trajectory recording every 0.02 m the received signal power in dBm.

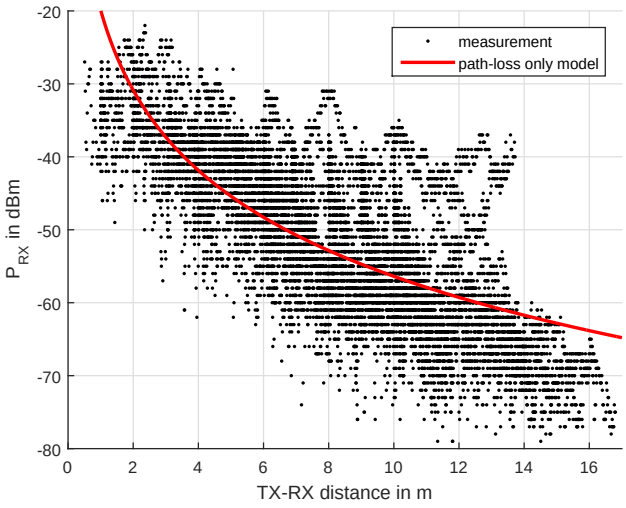


Figure 2.2: Measured received power in dBm with respect to TX-RX Euclidean distance. The estimated deterministic path-loss parameters are $\hat{L}_0 = -19.88$ dB and $\hat{\eta} = 3.65$. The path-loss only model ignoring any present spatial correlation is plot in red.

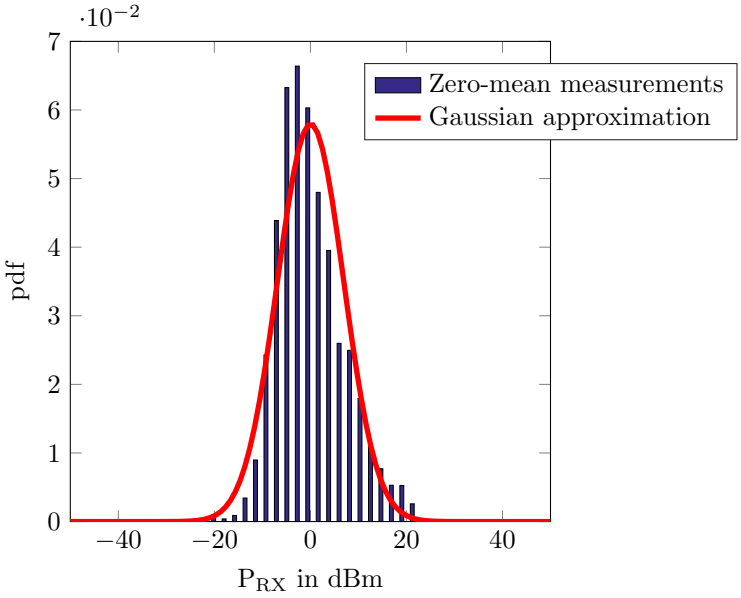


Figure 2.3: Empirical pdf of zero-mean measurements $\boldsymbol{y}_c$ plot in blue bars. A Gaussian approximation is plot in red.

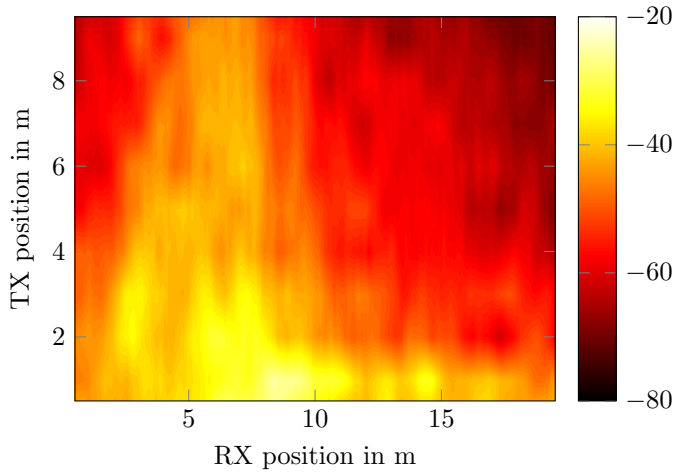


Figure 2.4: Recorded received power in dBm for different TX and RX positions. The RX moved perpendicular to the TX starting at TX position 6.5 m. Measurements have been spatially averaged to remove small-scale fading.

where $n \sim \mathcal{N}(0, \sigma_n^2)$ and σ_n^2 is the measurement noise variance. In order to make (2.2) applicable, we need to ensure multi-path is not present in y. For this, we have performed spatial averaging over a distance of 0.4 m along the RX trajectory. Assuming now (2.2) holds, we can determine the channel constants collected in a vector $\boldsymbol{\alpha} = [L_0, \eta]^{\mathrm{T}}$ via least-squares (LS) estimation with solution

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F})^{-1}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{y}, \tag{2.4}$$

where $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]^{\mathrm{T}}$, $\boldsymbol{F} = [\mathbf{1}_N, -\boldsymbol{d}]$, $\boldsymbol{d} = 10[\log_{10}\|\boldsymbol{q}_{\mathrm{TX},1} - \boldsymbol{q}_{\mathrm{RX},1}\|, \log_{10}\|\boldsymbol{q}_{\mathrm{TX},2} - \boldsymbol{q}_{\mathrm{RX},2}\|, \ldots, \log_{10}\|\boldsymbol{q}_{\mathrm{TX},N} - \boldsymbol{q}_{\mathrm{RX},N}\|]^{\mathrm{T}}$ and $\mathbf{1}_N$ is a column vector with N ones. From our measurements we have obtained $\hat{L}_0 = -19.88$dBm and $\hat{\eta} = 3.65$. The model (2.2) is plot in red in Fig. 2.2 using the estimated parameters. In (2.2), shadowing Ψ is assumed Gaussian with variance σ_Ψ^2. To verify this is true, we first obtain the zero mean measurements by

$$\boldsymbol{y}_c = \boldsymbol{y} - \boldsymbol{F}\hat{\boldsymbol{\alpha}}, \tag{2.5}$$

and then investigate its probability density function (pdf). In Fig. 2.3, we have plotted the empirical pdf of $\boldsymbol{y}_c$ together with a Gaussian fit, where we found $\Psi \sim \mathcal{N}(0.04, 6.88^2)$.[1] From the figure, we observe that shadowing is not exactly Gaussian distributed. It has a positive skewness but, with respect to accuracy and complexity of the channel model, a Gaussian approximation may still be reasonable.

For a static propagation environment and homogeneous TX and RX units, the transmission channel can be considered reciprocal [23]. This means if the role of TX and RX

[1] Note, since σ_n^2 in (2.3) is in general very low compared to σ_Ψ^2, we have omitted it here and therefore making it a part of σ_Ψ^2.

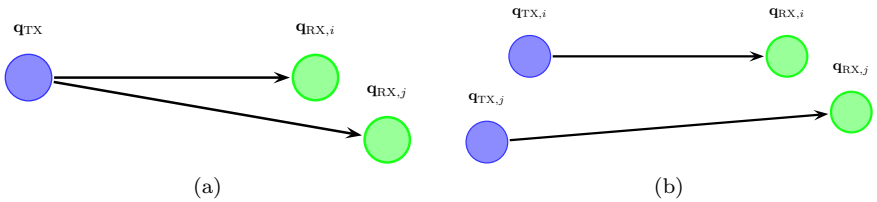


Figure 2.5: Left: The Gudmundson correlation model considers wireless communication links between different RX locations for a common TX location. Right: An ad-hoc correlation model needs to consider wireless communication links where the TX locations differ as well.

units is interchanged, the instantaneous channel characteristics remain the same. In (2.2), the part due to path-loss only depends on the distance between TX and RX and is therefore independent of the TX/RX roles. The second part of (2.2) models shadowing and to ensure channel reciprocity holds we need to ensure that $\Psi(\mathbf{q}_{\mathrm{TX}}, \mathbf{q}_{\mathrm{RX}}) = \Psi(\mathbf{q}_{\mathrm{RX}}, \mathbf{q}_{\mathrm{TX}})$ (c.f. [9, 22]).

Shadowing is, in general, spatially correlated, i.e., RX locations which are spatially close to each other experience similar shadow fading. For instance, if the TX-RX propagation path is blocked at one RX location, it may be also blocked at close-by RX locations with respect to the obstacle dimensions. This can be observed in Fig. 2. In the figure, the RX position indicates the position along the RX trajectory shown in Fig. 2.1 and similarly for the TX position. A line-of-sight (LOS) situation occurs when the RX is around 6.5 m resulting in high received signal power. For RX positions more distant than 6.5 m a non-LOS (NLOS) situation occurs due to the blockage from the wall.

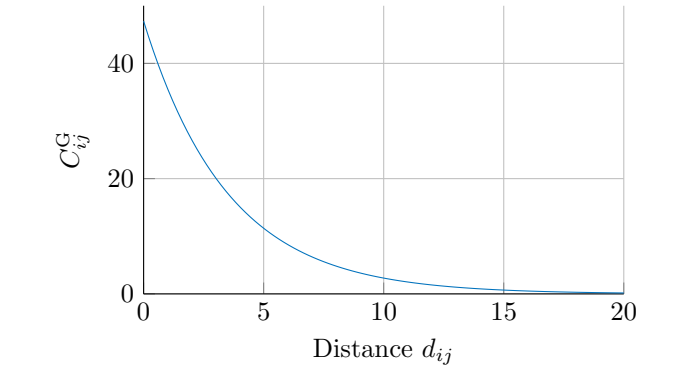
For a common TX endpoint $\mathbf{q}_{\mathrm{TX}} \in \mathcal{W}$, a well-known correlation model is the Gudmundson model for different RX locations $\mathbf{q}_{\mathrm{RX},i}, \mathbf{q}_{\mathrm{RX},j} \in \mathcal{W}$ [24]. This case is illustrated in Fig. 2.5a. The covariance between RX i and j with zero-mean measurements $\boldsymbol{y}_{c,i}$ and $\boldsymbol{y}_{c,j}$ follows an exponential decay, i.e.,

$$C_{ij}^{G}(d_{ij}) = \mathbb{E}[\boldsymbol{y}_{c,i}\boldsymbol{y}_{c,j}] = \sigma_{\Psi}^{2} \exp\left(-\frac{d_{ij}}{d_c}\right), \tag{2.6}$$

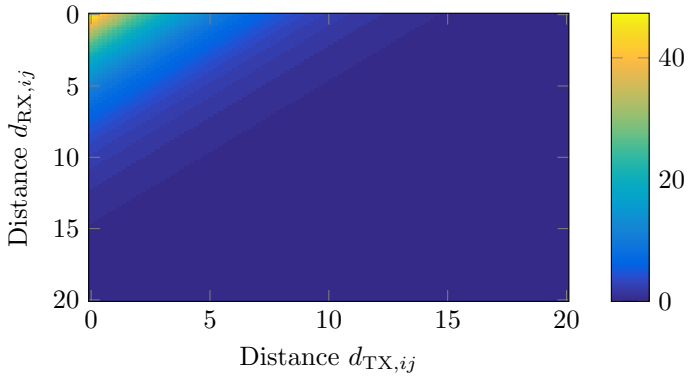
where $d_{ij} = \|\boldsymbol{q}_{\mathrm{RX},i} - \boldsymbol{q}_{\mathrm{RX},j}\|$ and d_c is the decorrelation distance. In Fig. 2.6a, the covariance under this model is plotted as a function of the distance between the RXs. The parameters $\sigma_{\Psi} = 6.88^2$ and $d_c = 3.51$ have been used.

This model can be extended to account for ad-hoc communication links with non-common TX endpoints $\mathbf{q}_{\mathrm{TX},i}, \mathbf{q}_{\mathrm{TX},j} \in \mathcal{W}$ (see, for example [25, 26]), under the assumption that the TX–RX distance is large compared to the displacement between the TX endpoints and the RX endpoints. This case is illustrated in Fig. 2.5b. The covariance under this model becomes [25]

$$C_{ij}^{W}(d_{\mathrm{TX},ij}, d_{\mathrm{RX},ij}) = \mathbb{E}[\boldsymbol{y}_{c,i}\boldsymbol{y}_{c,j}] = \sigma_{\Psi}^{2} \exp\left(-\frac{d_{\mathrm{TX},ij}}{d_c}\right) \exp\left(-\frac{d_{\mathrm{RX},ij}}{d_c}\right), \tag{2.7}$$



(a) C_{ij}^{G}



(b) C_{ij}^{W}

Figure 2.6: Top: covariance (2.6) as a function of distance d_{ij} between RX i and RX j. Bottom: covariance (2.7) as a function of the distances $d_{\mathrm{TX},ij}$ and $d_{\mathrm{RX},ij}$. The following parameters have been used: $\sigma_\Psi = 6.88$ and $d_c = 3.51$ m.

where $d_{\mathrm{TX},ij} = \|\boldsymbol{q}_{\mathrm{TX},i} - \boldsymbol{q}_{\mathrm{TX},j}\|$ and similarly for the RX. In Fig. 2.6b, the covariance under this model is plotted as a function of the distances between the TXs and the distance between the RXs. For brevity, we omit a discussion on the validity of these correlation models and instead refer the reader to [24–26]. Note, both presented models only depend on the distances between the link endpoints thus maintaining channel reciprocity.

Next, we will exploit the spatial correlation for channel gain prediction, where we make use of the channel model (2.2) together with the shadowing correlation model (2.7).

Chapter 3

Gaussian Process Regression

In this chapter, we introduce the GP based channel prediction framework. First, we define the GP model and address then the two important phases of GPs: (i) learning of the underlying model parameters, and (ii) prediction of a (non-linear) function at an unvisited location with the help of a measurement database. In general, the prediction phase makes use of the full measurement database, which is not practical in a C-MAS setting, where each agent has only access to its own local measurement database. For this reason, we adapt the centralized GP prediction allowing a distributed prediction in a wireless network using a consensus scheme. Finally, we make use of the GP framework in the context of channel prediction addressing both, the learning phase to determine the underlying channel parameters from measurements, and the prediction phase allowing to predict the wireless ad-hoc channel between arbitrary transmitter and receiver locations.

3.1 Standard Gaussian Process

3.1.1 Model

According to [27], a Gaussian process is defined as a probability distribution over functions $f(\boldsymbol{x})$ such that the set of values $f(\boldsymbol{x})$ evaluated at an arbitrary set of points $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ jointly have a Gaussian distribution.

We write the GP of a real stochastic process $f(\boldsymbol{x})$ with input $\boldsymbol{x} \in \mathbb{R}^D$ as [19]

$$f(\boldsymbol{x}) \sim \mathcal{GP}(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \tag{3.1}$$

with mean function

$$\mu(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})] \tag{3.2}$$

and covariance function (also called kernel function)

$$k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - \mu(\boldsymbol{x})(f(\boldsymbol{x}') - \mu(\boldsymbol{x}')]. \tag{3.3}$$

Many choices of kernel functions are available of which the γ-exponential family is commonly selected. It is given by [19]

$$k(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\left(\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|}{l}\right)^{\gamma}\right) \tag{3.4}$$

for $0 < \gamma \leq 2$, where the exponential kernel is obtained for $\gamma = 1$ and the squared exponential kernel for $\gamma = 2$. A necessary and sufficient condition for a kernel to be valid is that the resulting Gram matrix $\boldsymbol{K}$ with entries $\boldsymbol{K}_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$, $\forall \boldsymbol{x}_i, \boldsymbol{x}_j$ is positive semidefinite [27]. A way of constructing valid new kernels is to combine valid (simpler) kernels following some rules. As an example, consider two valid kernels $k_1(\boldsymbol{x}, \boldsymbol{x}')$ and $k_2(\boldsymbol{x}, \boldsymbol{x}')$. Then valid kernels are constructed by [27]

$$k_3(\boldsymbol{x}, \boldsymbol{x}') = c k_1(\boldsymbol{x}, \boldsymbol{x}'), \tag{3.5}$$

$$k_4(\boldsymbol{x}, \boldsymbol{x}') = k_1(\boldsymbol{x}, \boldsymbol{x}') + k_2(\boldsymbol{x}, \boldsymbol{x}'), \tag{3.6}$$

$$k_5(\boldsymbol{x}, \boldsymbol{x}') = k_1(\boldsymbol{x}_a, \boldsymbol{x}'_a) k_2(\boldsymbol{x}_b, \boldsymbol{x}'_b), \tag{3.7}$$

where $c > 0$ is a positive constant and $\boldsymbol{x} = [\boldsymbol{x}_a^{\mathrm{T}}, \boldsymbol{x}_b^{\mathrm{T}}]^{\mathrm{T}}$ (not necessarily disjoint). There exist more rules to construct valid kernels, but rules (3.5), (3.6), and (3.7) will turn out to be sufficient for our studies.

In order to apply GPs for regression, we need to take the noise on the observed function values into account. The measurement is given by

$$y = f(\boldsymbol{x}) + n, \tag{3.8}$$

where we assume $n \sim \mathcal{N}(0, \sigma_n^2)$. The GP model parameters are collected in a vector $\boldsymbol{\theta}$ and comprise the parameters of the mean function $\mu(\boldsymbol{x})$, the kernel function $k(\boldsymbol{x}, \boldsymbol{x}')$, as well as the measurement noise standard deviation σ_n.

The joint distribution of measurements $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]^{\mathrm{T}}$ and inputs $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]$ is Gaussian with

$$p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{X}), \boldsymbol{C}), \tag{3.9}$$

where $\boldsymbol{\mu}(\boldsymbol{X}) = [\boldsymbol{\mu}(\boldsymbol{x}_1)^{\mathrm{T}}, \boldsymbol{\mu}(\boldsymbol{x}_2)^{\mathrm{T}}, \ldots, \boldsymbol{\mu}(\boldsymbol{x}_N)^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{C}_{ij} = \boldsymbol{K}_{ij} + 1_{\{i=j\}}\sigma_n^2$. Here, $1_{\{I=J\}} = 1$ for $I = J$ and 0 otherwise. The measurement database comprising all N measurements is denoted by $\mathcal{D} = \{\boldsymbol{X}, \boldsymbol{y}\}$. With $\mathcal{D}$, the GP model parameters $\boldsymbol{\theta}$ can be learned. Once the learning phase is complete, the function value $f(\boldsymbol{x}_*)$ for an arbitrary input $\boldsymbol{x}_*$ can be predicted. Learning and prediction phase are explained next.

3.1.2 Learning Phase

Here, we are interested in estimating $\boldsymbol{\theta}$ from $\mathcal{D}$. An estimator for $\boldsymbol{\theta}$ is obtained by minimizing the negative log-likelihood with respect to $\boldsymbol{\theta}$. It can be written as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \{-\log p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta})\} \tag{3.10}$$

$$\propto \arg\min_{\boldsymbol{\theta}} \left\{\log|\boldsymbol{C}| + (\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{X}))^{\mathrm{T}} \boldsymbol{C}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{X}))\right\}, \tag{3.11}$$

where $\boldsymbol{C}$ and $\boldsymbol{\mu}$ depend on $\boldsymbol{\theta}$. Note that the negative log-likelihood is in general non convex. For this reason, we decided to use a global search to find the optimal $\boldsymbol{\theta}$. Our approach is outlined in Sec. 3.4.1.

Although a global search is optimal, it comes with the drawback of a high computational cost. A computationally more attractive approach is to perform a local search via a gradient based method. In [28], it is pointed out that such an approach often leads to a local minimum which might not explain the data well. This can be impeded by performing an a posteriori (MAP) estimation over the hyperparameters instead of maximum likelihood (ML) estimation. In this way, the prior on the hyperparameters will act as a regularization term such that unlikely hyperparameters will not lead to a local minimum. Since we assume there is no prior knowledge on $\boldsymbol{\theta}$ available, we cannot follow this approach here.

3.1.3 Prediction Phase

With the GP model parameters $\boldsymbol{\theta}$ and the training database $\mathcal{D}$ at hand, we are ready to determine the predictive distribution of $f(\boldsymbol{x}_*)$ for the arbitrary input $\boldsymbol{x}_*$. It is given by [19]

$$p(f(\boldsymbol{x}_*)|\mathcal{D}, \boldsymbol{\theta}) = \mathcal{N}(m(\boldsymbol{x}_*), \sigma^2(\boldsymbol{x}_*)) \tag{3.12}$$

with mean

$$m(\boldsymbol{x}_*) = \mu(\boldsymbol{x}_*) + \boldsymbol{k}^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{X})) \tag{3.13}$$

and variance

$$\sigma^2(\boldsymbol{x}_*) = k_{**} - \boldsymbol{k}_*\boldsymbol{C}^{-1}\boldsymbol{k}_*^{\mathrm{T}}, \tag{3.14}$$

where $k_{**} = k(\boldsymbol{x}_*, \boldsymbol{x}_*)$ and $\boldsymbol{k}_* = [k(\boldsymbol{x}_1, \boldsymbol{x}_*), k(\boldsymbol{x}_2, \boldsymbol{x}_*), \ldots, k(\boldsymbol{x}_N, \boldsymbol{x}_*)]^{\mathrm{T}}$.

3.2 Gaussian Process with Uncertain Inputs

So far, we have considered the input $\boldsymbol{x}$ of a GP to be known and hence deterministic. In practice, this is seldom the case. Consider for instance that $\boldsymbol{x}$ represents the position of an agent. This quantity needs to be estimated from (noisy) measurements and therefore, we need to model it as a random variable. Then, the earlier presented GP framework, which considers $\boldsymbol{x}$ to be deterministic, cannot be applied anymore.

To see this, consider $\boldsymbol{x}_* \sim \mathcal{N}(\boldsymbol{u}_*, \Sigma_{\boldsymbol{x}_*})$. The predictive distribution $f(\boldsymbol{u}_*)$ is obtained by [29]

$$p(f(\boldsymbol{u}_*)|\mathcal{D}, \boldsymbol{u}_*, \Sigma_{\boldsymbol{x}_*}, \boldsymbol{\theta}) = \int p(f(\boldsymbol{x}_*)|\mathcal{D}, \boldsymbol{x}_*, \boldsymbol{\theta}) p(\boldsymbol{x}_*|\boldsymbol{u}_*, \Sigma_{\boldsymbol{x}_*}) d\boldsymbol{x}_*, \tag{3.15}$$

which is not Gaussian anymore since

$$p(f(\boldsymbol{x}_*)|\mathcal{D}, \boldsymbol{x}_*, \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2(\boldsymbol{x}_*)}} \exp\left(-\frac{(f(\boldsymbol{x}_*) - m(\boldsymbol{x}_*))^2}{2\sigma^2(\boldsymbol{x}_*)}\right), \tag{3.16}$$

which is nonlinear in $\boldsymbol{x}_*$. Following [28, 29], we can approximate the predictive distribution by a Gaussian

$$p(f(\boldsymbol{u}_*)|\mathcal{D}, \boldsymbol{u}_*, \Sigma_{\boldsymbol{x}_*}, \boldsymbol{\theta}) \approx \mathcal{N}(m(\boldsymbol{u}_*), \sigma^2(\boldsymbol{u}_*)) \tag{3.17}$$

and compute its expected moments instead. The expected mean function can be obtained by computing the expectation with respect to the location uncertainty

$$m(\boldsymbol{u}_*) = \int \mu(\boldsymbol{x}_*)p(\boldsymbol{x}_*)d\boldsymbol{x}_* \tag{3.18}$$

and the expected covariance function can be obtained by

$$\boldsymbol{K}_{ij} = \iint k(\boldsymbol{x}_i, \boldsymbol{x}_j)p(\boldsymbol{x}_i)p(\boldsymbol{x}_j)d\boldsymbol{x}_i d\boldsymbol{x}_j, \text{ for } i \neq j, \tag{3.19}$$

$$\boldsymbol{K}_{ii} = \int k(\boldsymbol{x}_i, \boldsymbol{x}_i)p(\boldsymbol{x}_i)d\boldsymbol{x}_i, \qquad \text{for } i = j. \tag{3.20}$$

Note, for the special case of the squared exponential kernel, where $\gamma = 2$ in (3.4) and $p(\boldsymbol{x}_*)$ is Gaussian, closed form expressions for the expected moments exist. Furthermore, uncertainty in the input $\boldsymbol{x}_*$ not only occurs in the prediction phase as we have highlighted here. Also uncertainty in the training set $\boldsymbol{X}$ might be present, which affects both learning of the model parameters and prediction. For ease of discussion, we omit the implications of uncertainty in the training set. However, we address this in detail in the appended Papers A and B.

3.3 Distributed Gaussian Process

Here, we focus on distributed prediction over multiple agents which are all part of a connected (wireless) communication network. Assume all I agents know $\boldsymbol{\theta}$ and $\boldsymbol{x}_*$. Additionally, each agent i has a local database $\mathcal{D}_i$ such that $\mathcal{D} = \bigcup_{i=1}^{I} \mathcal{D}_i$. Furthermore, assume there is no overlap between the databases such that $|\mathcal{D}| = \sum_{i=1}^{I} |\mathcal{D}_i|$. Then, in general

$$\begin{aligned} p(f|\mathcal{D}) &= p(\mathcal{D}_i|\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_{i-1}, \mathcal{D}_{i+1}, \ldots, \mathcal{D}_I, f) \\ &\quad \times p(\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_{i-1}, \mathcal{D}_{i+1}, \ldots, \mathcal{D}_I|f)p(f), \end{aligned} \tag{3.21}$$

where we have omitted the terms $\boldsymbol{x}_*$ and $\boldsymbol{\theta}$ for readability. For an efficient distributed prediction, we want that each agent only needs access to its own database. Therefore we perform the following approximation similar to [30]

$$p(\mathcal{D}_i|\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_{i-1}, \mathcal{D}_{i+1}, \ldots, \mathcal{D}_I, f) \approx p(\mathcal{D}_i|f). \tag{3.22}$$

According to [30], this is in general not true, but it might be a good approximation if the correlation among the databases is low. This can be achieved if mobile agents collect measurements for building their local database $\mathcal{D}_i$ for every agent i in geographically non-overlapping regions or, as pointed out by [30], by clustering the database $\mathcal{D}$ and assigning each database cluster $\mathcal{D}_i$ to a different agent i. We then have

$$p(f|\mathcal{D}) \propto p(f) \prod_{i=1}^{I} p(\mathcal{D}_i|f), \tag{3.23}$$

which yields after applying Bayes' rule

$$p(f|\mathcal{D}) \propto \frac{\prod_{i=1}^{I} p(f|\mathcal{D}_i)}{p(f)^{I-1}}, \tag{3.24}$$

where $p(f)$ is the a priori predictive pdf at the test input $\boldsymbol{x}_*$ with mean $\mu(\boldsymbol{x}_*)$ and variance k_{**}, i.e., without a training database. Since all distributions involved are Gaussian, the moments of the right hand side of (3.24) can be computed exactly in a distributed manner using neighbor communication only through a consensus scheme [21].

We know $p(f|\mathcal{D}_i) = \mathcal{N}(m_i(\boldsymbol{x}_*), \sigma_i^2(\boldsymbol{x}_*))$, where we have used (3.13) and (3.14), and similarly for the a priori pdf $p(f) = \mathcal{N}(m_{\mathrm{a}}(\boldsymbol{x}_*), \sigma_{\mathrm{a}}^2(\boldsymbol{x}_*))$, where in this case $\mathcal{D}_i = \varnothing$. Since the product of a Gaussian pdf is proportional to a Gaussian, and so is the ratio of two Gaussian pdfs [1] , we can rewrite (3.24) by

$$p(f|\mathcal{D}) \propto \frac{\mathcal{N}_f(\mu_{\mathrm{n}}, \sigma_{\mathrm{n}}^2)}{\mathcal{N}_f(\mu_{\mathrm{d}}, \sigma_{\mathrm{d}}^2)} \tag{3.25}$$

with

$$\mu_{\mathrm{n}} = \frac{\sum_{i=1}^{I} \sigma_i^{-2}(\boldsymbol{x}_*) m_i(\boldsymbol{x}_*)}{\sum_{i=1}^{I} \sigma_i^{-2}(\boldsymbol{x}_*)}, \tag{3.26}$$

$$\sigma_{\mathrm{n}}^2 = \left(\sum_{i=1}^{I} \sigma_i^{-2}(\boldsymbol{x}_*) \right)^{-1}, \tag{3.27}$$

$$\mu_{\mathrm{d}} = m_{\mathrm{a}}(\boldsymbol{x}_*) \tag{3.28}$$

$$\sigma_{\mathrm{d}}^2 = \frac{\sigma_{\mathrm{a}}^2(\boldsymbol{x}_*)}{I-1}, \tag{3.29}$$

such that

$$p(f|\mathcal{D}) \propto \mathcal{N}_f(\mu_{\mathrm{post}}, \sigma_{\mathrm{post}}^2) \tag{3.30}$$

with

$$\sigma_{\mathrm{post}}^2 = \frac{1}{\sigma_{\mathrm{n}}^{-2} + \sigma_{\mathrm{d}}^{-2}}, \tag{3.31}$$

$$\mu_{\mathrm{post}} = \sigma_{\mathrm{post}}^2 \left(\sigma_{\mathrm{n}}^{-2} \mu_{\mathrm{n}} - \sigma_{\mathrm{d}}^{-2} \mu_{\mathrm{d}} \right). \tag{3.32}$$

A distributed computation of (3.24) is achieved in the following way. Each agent computes locally first the denominator of (3.24), i.e., the product of the prior pdfs. Then, to compute the numerator of (3.24) we use average consensus [21]. In doing so, we initialize the variables

$$\alpha_i^{(0)} = I \sigma_i^{-2}(\boldsymbol{x}_*) m_i(\boldsymbol{x}_*), \tag{3.33}$$

$$\beta_i^{(0)} = I \sigma_i^{-2}(\boldsymbol{x}_*) \tag{3.34}$$

[1] Note that, $\mathcal{N}_f(\mu_1, \Sigma_1)\mathcal{N}_f(\mu_2, \Sigma_2) \propto \mathcal{N}_f\left(\left(\Sigma_1^{-1} + \Sigma_2^{-1} \right)^{-1} \left(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2 \right), \left(\Sigma_1^{-1} + \Sigma_2^{-1} \right)^{-1} \right)$ and $\frac{\mathcal{N}_f(\mu_1, \Sigma_1)}{\mathcal{N}_f(\mu_2, \Sigma_2)} \propto \mathcal{N}_f\left(\left(\Sigma_1^{-1} - \Sigma_2^{-1} \right)^{-1} \left(\Sigma_1^{-1}\mu_1 - \Sigma_2^{-1}\mu_2 \right), \left(\Sigma_1^{-1} - \Sigma_2^{-1} \right)^{-1} \right)$ [30].

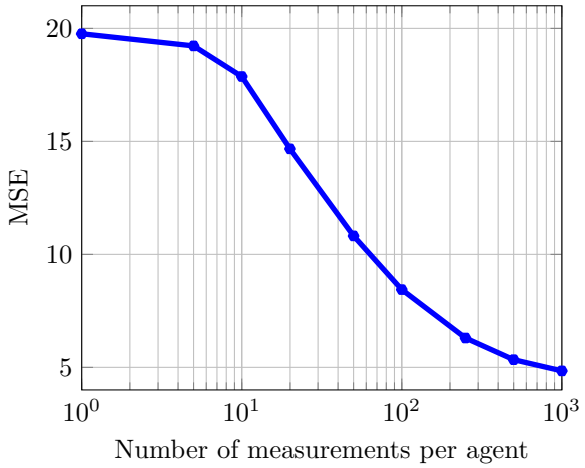


Figure 3.1: MSE versus size of the database per agent for a total database size of $1,000$. The figure is taken from appended Paper C.

for each agent i. The superscript of α and β denote the iteration step of the consensus algorithm, where

$$\boldsymbol{\gamma}_i^{(l)} = [\alpha_i^{(l)}, \beta_i^{(l)}]^{\mathrm{T}} \tag{3.35}$$

and the consensus update rule

$$\boldsymbol{\gamma}_i^{(l+1)} = \boldsymbol{\gamma}_i^{(l)} + \kappa \sum_{j \in \mathcal{N}_i} (\boldsymbol{\gamma}_j^{(l)} - \boldsymbol{\gamma}_i^{(l)}) \tag{3.36}$$

is applied. Here, $\kappa > 0$ is a small step-size depending on the communication graph and $\mathcal{N}_i$ denotes the neighborhood set of agent i. In this way, we have ensured that

$$\alpha_i^{(\infty)} = \sum_{i=1}^{I} \sigma_i^{-2}(\boldsymbol{x}_*) m_i(\boldsymbol{x}_*) \tag{3.37}$$

and

$$\beta_i^{(\infty)} = \frac{1}{\sigma_{\mathrm{n}}^2}, \tag{3.38}$$

allowing to retrieve (3.26) and (3.27).

Due to the assumption of independent databases $\mathcal{D}_i$, the prediction quality is deteriorated with respect to the fragmentation of $\mathcal{D}$. In Fig. 3.1 the mean squared error (MSE) between the centralized ($I = 1$) and the distributed case ($I > 1$) is plotted for different number of measurements per agent. Implementation and simulation details can be found in appended Paper C.

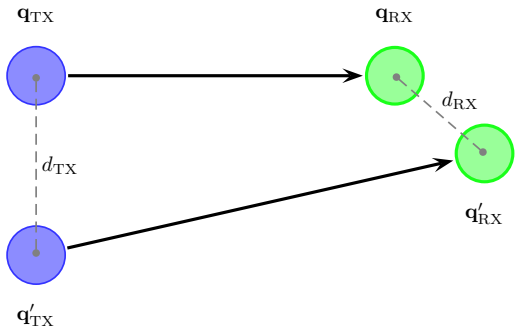


Figure 3.2: Illustration of TX distance d_{TX} and RX distance d_{RX} between two links $\boldsymbol{x}$ and $\boldsymbol{x}'$.

3.4 GP Regression for Channel Gain Prediction

Here, we are interested in modeling the received signal power (in dBm) as a GP with inputs $\boldsymbol{x} = [\boldsymbol{q}_{\mathrm{TX}}^{\mathrm{T}}, \boldsymbol{q}_{\mathrm{RX}}^{\mathrm{T}}]^{\mathrm{T}}$ consisting of the TX location $\boldsymbol{q}_{\mathrm{TX}} \in \mathcal{W}$ and the RX location $\boldsymbol{q}_{\mathrm{TX}} \in \mathcal{W}$, respectively. Then

$$P_{\mathrm{RX}}(\mathbf{x}) \sim \mathcal{GP}(\mu_{\mathrm{cGP}}(\mathbf{x}), c_{\mathrm{cGP}}(\mathbf{x}, \mathbf{x}')). \tag{3.39}$$

The mean $\mu_{\mathrm{cGP}}(\mathbf{x})$ is obtained by computing the expectation of (2.2) with respect to the large-scale fading $\Psi(\mathbf{x})$, yielding

$$\mu_{\mathrm{cGP}}(\mathbf{x}) = L_0 - 10\eta \log_{10} ||\mathbf{q}_{\mathrm{TX}} - \mathbf{q}_{\mathrm{RX}}||_2. \tag{3.40}$$

Under the assumption that spatial channel correlation is isotropic, it can be characterized by the Euclidean distance d_{TX} between the TX locations of two communication links $\boldsymbol{x}$ and $\boldsymbol{x}'$, and by their RX distance d_{RX}, respectively. This relationship is illustrated in Fig. 3.2. Therefore we define the covariance function

$$k(\mathbf{x}, \mathbf{x}') = \sigma_\Psi^2 \exp\left(-\frac{d_{\mathrm{TX}}^\gamma}{d_c^\gamma}\right) \exp\left(-\frac{d_{\mathrm{RX}}^\gamma}{d_c^\gamma}\right) + 1_{\{\mathbf{x}=\mathbf{x}'\}} \sigma_{\mathrm{proc}}^2, \tag{3.41}$$

where $d_{\mathrm{TX}} = \|\mathbf{q}_{\mathrm{TX}} - \mathbf{q}'_{\mathrm{TX}}\|$, $d_{\mathrm{RX}} = \|\mathbf{q}_{\mathrm{RX}} - \mathbf{q}'_{\mathrm{RX}}\|$, parameter σ_Ψ^2 denotes the variance of the shadowing process, and parameter σ_{proc}^2 is the variance of an additional white noise process able to model fluctuations of P_{RX}, which cannot be explained by spatial correlation and measurement noise only. This could, for instance, be a model mismatch with respect to real measurements. The parameter d_c is the decorrelation distance. For $\gamma = 1$ we have the model of [25], and if one of the link endpoints distance is zero, which is the case if $\boldsymbol{q}_{\mathrm{TX}} = \boldsymbol{q}'_{\mathrm{TX}}$ or $\boldsymbol{q}_{\mathrm{RX}} = \boldsymbol{q}'_{\mathrm{RX}}$, we obtain the Gudmundson model [24]. Note that kernel (6) is constructed from the γ-exponential family (3.4) by making use of the kernel construction rules (3.5), (3.6), and (3.7).

Assuming the noise power σ_n^2 is known, the hyper-parameter vector is

$$\boldsymbol{\theta} = [L_0, \eta, \sigma_\Psi, \sigma_{\mathrm{proc}}, d_c]^{\mathrm{T}}. \tag{3.42}$$

3.4.1 Learning Phase

As pointed out in Sec. 3.1.2, learning the hyperparameter vector $\boldsymbol{\theta}$ can be challenging. We therefore perform a simpler sub-optimal two-step approach [20, 31]. First, we perform a least-squares estimation on the deterministic path-loss parameters L_0 and η:

$$\hat{L}_0, \hat{\eta} = \underset{L_0,\eta}{\arg\min} \sum_{i=1}^{N} \left(y_i - \left(L_0 - 10\eta \log_{10} \left\|\boldsymbol{q}_{\mathrm{TX},i} - \boldsymbol{q}_{\mathrm{RX},i}\right\|\right)\right)^2 \tag{3.43}$$

leading to the closed-form expression (2.4). Following that, a ML estimation using the zero-mean measurements $\boldsymbol{y}_c = [y_{c,1}, y_{c,2}, \ldots, y_{c,N}]^{\mathrm{T}}$ is performed to find the remaining hyperparameters $\sigma_\Psi, \sigma_{\mathrm{proc}}$ and d_c. The zero-mean measurements $\boldsymbol{y}_c$ are obtained by (2.5). The ML estimation solves

$$\hat{\sigma}_\Psi, \hat{\sigma}_{\mathrm{proc}}, \hat{d}_c = \underset{\sigma_\Psi,\sigma_{\mathrm{proc}},d_c}{\arg\min} \left\{\log|\boldsymbol{C}| + \boldsymbol{y}_c^{\mathrm{T}} \boldsymbol{C}^{-1} \boldsymbol{y}_c\right\}, \tag{3.44}$$

where we have used a global search to find the shadowing hyperparameters similar to [20].

3.4.2 Prediction Phase

With the estimated hyperparameter vector $\boldsymbol{\theta}$, channel gain prediction can be performed as explained in Sec. 3.1.3. In doing so, the mean function (3.40) and the kernel function (3.41) are used.

3.4.3 Incorporation of Channel Reciprocity

The GP framework as introduced previously does not consider channel reciprocity per se. The reason for this is that, through the definition of the input vector $\boldsymbol{x} = [\boldsymbol{q}_{\mathrm{TX}}^{\mathrm{T}}, \boldsymbol{q}_{\mathrm{RX}}^{\mathrm{T}}]^{\mathrm{T}}$, an implicit ordering has been introduced. There are several ways how channel reciprocity can be ensured: by applying an operator on the input vector $\boldsymbol{x}$ to make it independent on the *link direction*, by modifying the kernel function, or by extending the measurement database by its reciprocal counterpart. In our work, we have chosen the latter approach, where for every measurement y_i recored at TX-RX location pair $\boldsymbol{x}_i$ an additional entry in the database $\mathcal{D}$ is made with the same measurement, but where the role of TX and RX are interchanged in $\boldsymbol{x}_i$.

Chapter 4

Application: Optimal Router Configuration Under Location Uncertainty

In this chapter, we apply the developed GP based channel prediction framework in a network context. In particular, we minimize the bit error rate (BER) along a given communication path. Consider the exemplary communication network outlined in Fig. 4.1, where there is a control center (CC), a mobile search-and-rescue robot (SR) and several mobile support robots acting as communication relays, which allow to increase the operating range of the search-and-rescue robot. Thanks to our channel prediction framework, optimal locations of each of the communication relays can be found such that the BER from CC to SR is minimized. Since every robot knowns its own location only with limited accuracy provided by its sensors, we need to consider this in the channel prediction framework as it has been introduced in Sec. 3.2.

4.1 Model

Consider L agents (robots) operating in the workspace $\mathcal{W}$. The agents build a communication network, which can be described by a graph $G = (V, E)$ with the agents as vertices V and the wireless communication between them modeled as edges $E \subset V \times V$. Assuming reciprocal channels, each edge (communication link) consists of a TX located at $\boldsymbol{q}_{i-1}$ and a RX located at $\boldsymbol{q}_i$ for $i = 2, 3, \ldots, L$. An edge is then described by the vector $\boldsymbol{x}_{i,i-1} = [\boldsymbol{q}_{i-1}^{\mathrm{T}}, \boldsymbol{q}_i^{\mathrm{T}}]^{\mathrm{T}}$. The set of all valid configurations is denoted $\mathcal{X} \subset \mathcal{W}$. For the sake of brevity, we omit here the case when robots locations are uncertain. Instead, we refer the reader to appended Paper A.

We are now ready to present the optimization objective. The BER can be defined as the probability that the received bit at agent i does not correspond to the transmitted bit at agent $i-1$. It is denoted $P_b(i|i-1)$. Assuming additive white Gaussian noise and

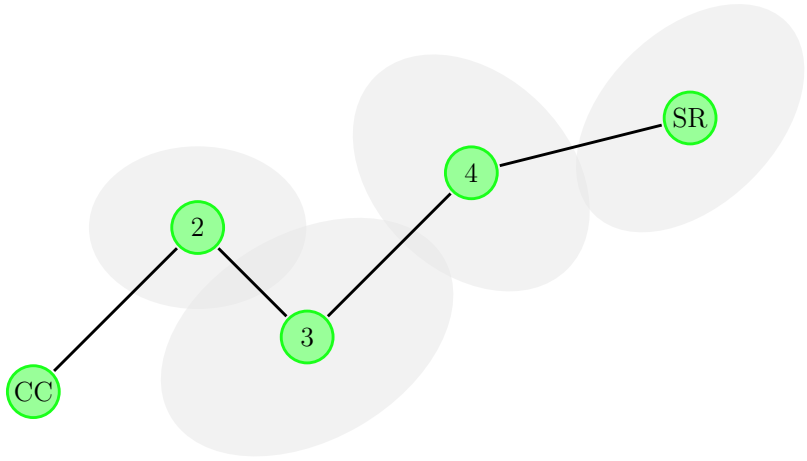


Figure 4.1: Communication network between a search-and-rescue robot SR, relay robots $2, 3$ and 4, and the command center CC. Each of the mobile robots has some location uncertainty illustrated by the gray error ellipses. This needs to be modeled in the channel prediction to achieve the higher level task such as, e.g., maximizing the communication quality.

quadrature amplitude modulation with M bits per symbol (MQAM) the BER between robot i and $i-1$ can be approximated by [9]

$$P_b(i|i-1) \approx 0.2 \exp(-c\gamma_{i,i-1}), \tag{4.1}$$

where $c = 1.5/(M-1)$, M is the constellation size, and $\gamma_{i,i-1}$ is the signal-to-noise ratio (SNR) between TX $i-1$ and RX i. Following [32], instead of minimizing the BER, we maximize the probability of correct reception $P_c(i|i-1) = 1 - P_b(i|i-1)$. For the whole network this yields

$$P_c \approx \prod_{i=2}^{L} P_c(i|i-1) \tag{4.2}$$

assuming the communication graph is a chain as illustrated in Fig. 4.1. Note that, bit flips leading to a correct reception are not considered and hence (4.2) acts as a lower bound on the probability of correct reception. Then, (4.1) combined with (4.2) yields the objective

$$J(\boldsymbol{X}) = \prod_{i=2}^{L} \left(1 - 0.2\mathbb{E}_{\gamma_{i,i-1}}\left[\exp(-c\gamma_{i,i-1})\right]\right), \tag{4.3}$$

which we seek to maximize over the locations $\boldsymbol{X} = [\boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_L]$. Here, we have assumed the communication links to be independent and the expectation is due to the fact that SNR $\gamma_{i,i-1}$ is a random variable. The expectation cannot be solved analytically,

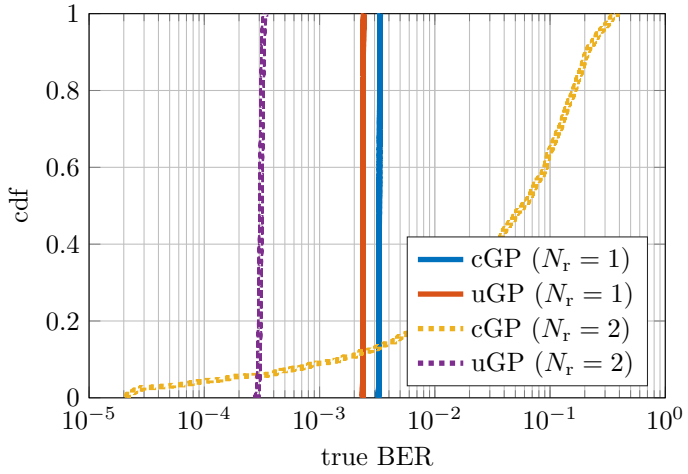


Figure 4.2: The cdf of the achieved end-to-end BER is plot for a multi-relay communication with cGP and uGP. The figure is taken from appended Paper A.

but a tight approximation exists (c.f. appended Paper A). The resulting optimization problem is

$$\underset{\boldsymbol{X}\in\mathcal{X}}{\text{maximize}}\ J(\boldsymbol{X}). \tag{4.4}$$

4.2 Results

In Fig. 4.2, the cumulative distribution function (cdf) of the true achieved BER for one ($N_r = 1$) and two ($N_r = 2$) relay robots is plotted for the case of using the GP based predictor, which is denoted cGP in the figure. Solving (4.4), when location uncertainty of the agents is explicitly modeled as highlighted in Sec. 3.2, leads to an improved performance with respect to the experienced BER. This case is denoted by uGP in the figure. Note that in the case of cGP, the true experienced BER is less with only one relay ($N_r = 1$) compared to using two relays ($N_r = 2$). This seems odd, especially since the inter-robot distance is less. The reason for this is the following. After solving (4.4), the optimal locations of the robots are found but once they steer to their optimal location, they will not exactly end up at this location due to the robots location uncertainty. Hence the true BER is different. This phenomenon is more pronounced with more relay robots having more sources of location uncertainty. For uGP this is not an issue, since any location uncertainty of the robots has been considered in solving (4.4).

Chapter 5

Contributions

Included Publications

1. **Paper A: "Channel Prediction with Location Uncertainty for Ad-Hoc Networks"**
   Multi-agent systems (MAS) rely on positioning technologies to determine their physical location, and on wireless communication technologies to exchange information. Both positioning and communication are affected by uncertainty, which should be accounted for. This paper considers an agent placement problem to optimize end-to-end communication quality in a MAS, in the presence of such uncertainties. Using Gaussian processes (GPs), operating on the input space of location distributions, we are able to model, learn, and predict the wireless channel. Predictions, in the form of distributions, are fed into the communication optimization problems. This approach inherently avoids regions of the workspace with high position uncertainty and leads to better average communication performance. We illustrate the benefits of our approach via extensive simulations, based on real wireless channel measurements. Finally, we demonstrate the improved channel learning and prediction performance, as well as the increased robustness in agent placement.

2. **Paper B: "Channel Gain Prediction for Multi-Agent Networks in the Presence of Location Uncertainty"**
   Coordination among mobile agents relies on communication over a wireless channel and can thus be improved by channel prediction. We present a Gaussian process framework to learn channel parameters and predict the channel between arbitrary transmitter and receiver locations. We explicitly incorporate location uncertainty in both learning and prediction phases. Simulation results show that if location uncertainty is not modeled appropriately, it has a degenerative effect on the prediction quality.

3. **Paper C: "Distributed Channel Prediction for Multi-Agent Systems"**
   Multi-agent systems (MAS) communicate over a wireless network to coordinate their actions or to report their mission status. Connectivity and system-level performance

can be improved by channel gain prediction. We present a distributed Gaussian process regression (GPR) framework, suitable for MAS. The framework combines a Bayesian Committee Machine with an average consensus scheme, thus distributing not only the memory, but also computational and communication loads. Through Monte Carlo simulations, we demonstrate the performance of the proposed GPR.

References

[1] B. J. Julian, M. Angermann, M. Schwager, and D. Rus, "Distributed robotic sensor networks: An information-theoretic approach," *The International Journal of Robotics Research*, vol. 31, no. 10, pp. 1134–1154, 2012.

[2] A. Nayak and I. Stojmenovic, "Wireless sensor and actuator networks," *John-Whiley & sons*, 2010.

[3] M. A. Hsieh, A. Cowley, V. Kumar, and C. J. Taylor, "Maintaining Network Connectivity and Performance in Robot Teams," *Journal of Field Robotics*, vol. 25, no. 1-2, pp. 111–131, 2008.

[4] L. S. Muppirisetty, J. Tadrous, A. Eryilmaz, and H. Wymeersch, "On Proactive Caching with Demand and Channel Uncertainties," in *53rd Annual Allerton Conference on Communication, Control, and Computing*, Sept 2015, pp. 1174–1181.

[5] H. Abou-Zeid and H. S. Hassanein, "Predictive Green Wireless Access: Exploiting Mobility and Application Information," *IEEE Wireless Communications*, vol. 20, no. 5, pp. 92–99, 2013.

[6] S. Waharte and N. Trigoni, "Supporting Search and Rescue Operations with UAVs," in *International Conference on Emerging Security Technologies*, Sept 2010, pp. 142–147.

[7] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil, "Vehicular Networking: A Survey and Tutorial on Requirements, Architectures, Challenges, Standards and Solutions," *IEEE Communications Surveys Tutorials*, vol. 13, no. 4, pp. 584–616, 2011.

[8] A. A. Zaidi, B. Kulcsár, and H. Wymeersch, "Traffic-adaptive Signal Control and Vehicle Routing Using a Decentralized Back-pressure Method," in *European Control Conference*, July 2015, pp. 3029–3034.

[9] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[10] A. Böttcher, P. Vary, C. Schneider, and R. S. Thomä, "De-Correlation Distance of the Large Scale Parameters in an Urban Macro Cell Scenario," in *6th European Conference on Antennas and Propagation (EUCAP)*, 2012, pp. 1417–1421.

[11] N. Jálden, "Analysis and Modelling of Joint Channel Properties from Multi-site, Multi-Antenna Radio Measurements," Ph.D. dissertation, KTH, Signal Processing, 2010.

[12] D. P. Spanos and R. M. Murray, "Motion Planning with Wireless Network Constraints," in *Proceedings of the American control conference*, vol. 1, 2005, pp. 87–92.

[13] B. M. M. Zavlanos, M. B. Egerstedt, and G. J. Pappas, "Graph-Theoretic Connectivity Control of Mobile Robot Networks," *Proceedings of the IEEE*, vol. 99, no. 9, pp. 1525–1540, 2011.

[14] M. M. Zavlanos and G. J. Pappas, "Distributed Connectivity Control of Mobile Networks," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1416–1428, 2008.

[15] A. Simonetto, T. Keviczky, and R. Babuška, "Constrained distributed algebraic connectivity maximization in robotic networks," *Automatica*, vol. 49, no. 5, pp. 1348–1357, 2013.

[16] J. Cortés, S. Martinez, and F. Bullo, "Spatially-distributed coverage optimization and control with limited-range interactions," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 11, no. 04, pp. 691–719, 2005.

[17] J. Cortés, "Coverage Optimization and Spatial Load Balancing by Robotic Sensor Networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 749–754, 2010.

[18] N. M. M. De Abreu, "Old and new results on algebraic connectivity of graphs," *Linear algebra and its applications*, vol. 423, no. 1, pp. 53–73, 2007.

[19] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. MIT Press, 2006.

[20] L. S. Muppirisetty, T. Svensson, and H. Wymeersch, "Spatial Wireless Channel Prediction under Location Uncertainty," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1031–1044, 2016.

[21] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan 2007.

[22] G. L. Stüber, *Principles of Mobile Communication*. Springer Science & Business Media, 2011.

[23] M. Neiman, "The principle of reciprocity in antenna theory," *Proceedings of the IRE*, vol. 31, no. 12, pp. 666–671, 1943.

[24] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, 1991.

[25] Z. Wang, E. K. Tameh, and A. R. Nix, "Joint Shadowing Process in Urban Peer-to-Peer Radio Channels," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 52–64, 2008.

[26] Z. Li, R. Wang, and A. F. Molisch, "Shadowing in Urban Environments with Microcellular or Peer-to-Peer Links," in *6th European Conference on Antennas and Propagation (EUCAP)*, 2012, pp. 44–48.

[27] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[28] P. Dallaire, C. Besse, and B. Chaib-draa, "An approximate inference with Gaussian process to latent functions from uncertain data," *Neurocomputing*, vol. 74, no. 11, pp. 1945–1955, 2011.

[29] A. Girard, "Approximate Methods for Propagation of Uncertainty with Gaussian Process Models," Ph.D. dissertation, University of Glasgow, 2004.

[30] V. Tresp, "A Bayesian committee machine," *Neural Computation*, vol. 12, no. 11, pp. 2719–2741, 2000.

[31] M. Malmirchegini and Y. Mostofi, "On the Spatial Predictability of Communication Channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 964–978, 2012.

[32] Y. Yan and Y. Mostofi, "Robotic Router Formation in Realistic Communication Environments," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 810–827, 2012.