Dissertation

The Random-Cluster Model –

A Stochastic MIMO Channel Model for Broadband Wireless Communication Systems of the 3rd Generation and Beyond

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der
Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von
Dipl.-Ing. Nicolai Czink

Wien, Dezember 2007

TECHNISCHE UNIVERSITÄT WIEN

VIENNA UNIVERSITY OF TECHNOLOGY

Supervisor

Prof. Ernst Bonek

Institut für Nachrichtentechnik und Hochfrequenztechnik

Technische Universität Wien

Examiner

Dr. Juha Ylitalo

Elektrobit Corp.

Finland

ftw. Dissertation Series

Nicolai Czink

The Random-Cluster Model –

A Stochastic MIMO Channel Model for Broadband Wireless Communication Systems of the 3rd Generation and Beyond

telecommunications research center vienna

This work was carried out with funding from **K***plus* in the ftw. project I0.

Supervisor at ftw.: Thomas Zemen

This thesis has been prepared using LaTeX.

Dezember 2007
2. Auflage

Herausgeber: Forschungszentrum Telekommunikation Wien
Printed in Austria
ISBN 3-902477-16-4

Acknowledgements

First of all, I want to thank the Elektrobit Corp., Finland. Their co-funding over a period of three years made this thesis possible. Especially, I want to acknowledge the continuous and strong support of their research team, Dr. Juha Ylitalo, Jukka-Pekka Nuutinen, Lassi Hentilä, and Pekka Kyösti. Our cooperation was a true pleasure for me. I also want to express my gratitude to the Centre for Wireless Communications at the University of Oulu, Finland, who were so kind to offer their sounder for the MIMO channel measurements I conducted. Special thanks to Mikko Alatossava and Veli-Matti Holappa for their help with the measurements (and for teaching me my first words in Finnish: Kolme, kaksi, yksi)!

I am also very grateful to the Forschungszentrum Telekommunikation Wien (ftw.) for co-funding this thesis. Both ftw. and the Institut für Nachrichtentechnik und Hochfrequenztechnik at Technische Universität Wien provided me with a wonderful scientific working environment. I was free to follow all my ideas, and to let them blossom and grow. In the hectic nowadays world, this freedom was really invaluable to me. I owe many thanks to all my colleges who could help me with the many questions I had. I especially want to mention Prof. Christoph Mecklenbräuker, for hiring me at ftw., and for his continuous support of my research work, Dr. Thomas Zemen for many fruitful (and sometimes also quite heated) debates, and Prof. Franz Hlawatsch for teaching me the DOs and DON'Ts in signal processing.

Very special thanks go to my supervisor, or rather: mentor, Prof. Ernst Bonek. Even though he has already been retired, he took the burden of guiding me through the jungle of science, engineering, cooperation, funding, and doubt. The maybe most important thing he taught me was that science must be built on doubt, hence concepts need to be challenged before we can accept them as valid. His unique way of asking me controversial questions, rather than stating his opinion, was extremely helpful to recognise the right view of bewildering findings.

I was very lucky to meet and cooperate with so many world-class researchers in the international community. I am very grateful for successful cooperation, and also for interesting discussions with the following people: Dr. Xuefeng Yin, Prof. Bernard Fleury, Prof. Reiner Thomä, Dr. Giovanni Del Galdo, Markus Landmann, Prof. Pertti Vainikainen, Veli-Matti Kolmonen, Dr. Andreas Richter, Prof. Andreas Molisch, Dr. Fredrik Tufvesson, Shurjeel Wyne, Johan Karedal, Ruiyuan Tian, Prof. Claude Oestges, Prof. Alister Burr, Hui (Julie) Xiao, Prof. Mark Beach, Prof. Luis Correia, Prof. Vittorio Degli Esposti, Dr. Jari Salo, Pierluigi Cera, and Prof. Mary Ann Ingram. Some of them that had a very particular impact on my research are mentioned again in the text of my thesis.

Last but not least I want to thank my family and my friends. Special thanks go to my father who supported me in every way possible, for which I am immensely grateful. I also want to express my gratitude for the loving support of my grandmother. My best friends, Stefan Hummel, Karin Chmel, Ursula Simo, and Werner Chmel were my foothold in good and also in distressing times. Finally, I want to say "Thank you for everything!" to my love, Karoline Sindelar.

Abstract

This thesis presents the *Random-Cluster Model* (RCM), a novel geometry-based stochastic model for frequency-selective and smoothly time-variant MIMO radio channels.

The RCM uses the concept of clusters, i.e. groups of multipath components (MPCs), to model the propagation environment. In the RCM, the environment is solely specified by the *multivariate distribution* of the *cluster parameters*, such as the cluster positions, the cluster movement, and the cluster spreads. In this way, even correlations between the cluster parameters are easily reflected.

The most significant feature of the RCM is that it is parametrised directly from *channel measurements* by an automatic procedure. In this way, the RCM is *specific to the environment*. It closes the gap between channel measurements and channel modelling.

By using clusters, the model parametrisation *complexity* becomes *quite low*. Compared to a single propagation path, needing 6 parameters per modelled time instant, a cluster is described by as few as 21 parameters for the whole cluster lifetime (over many time instants). Since a cluster usually consists of 6 to 20 propagation paths, the reduction in the number of parameters is significant.

Of course, the cluster distribution needs to be parametrised accurately. Identifying clusters from time-variant MIMO channel measurements constitutes the basis to parametrise the RCM consistently. Since visual clustering algorithms are highly subjective and cumbersome to use, I concentrated on the long cherished goal of *automatic clustering*. This thesis compares different clustering approaches regarding their suitability for multipath clustering, i.e. visual clustering, hierarchical clustering, K-means clustering, and Gaussian Mixture Model clustering. Eventually, I present a new and complete framework that provides a solution for automatic clustering and tracking of MPCs, which bases on

(i) an initial-guess estimator choosing clusters to be as separate as possible,

(ii) the *KPowerMeans* algorithm, an extension of the K-means algorithm, that takes the power of MPCs into account and handles the ambiguity of the angular domain, and

(iii) a Kalman filter for cluster tracking.

Using this framework, the RCM is automatically parametrised from measurements.

To account for non-discrete contributions in the MIMO channel, the RCM is the first channel model to include the concept of *diffuse multipath* (DMP) modelling.

I will argue that *validation* is an *important task* when completing a channel model. I validate the RCM by comparing its fit to measurements using the following validation metrics: (i) mutual information, which will turn out to be no distinctive validation metric, (ii) channel diversity, (iii) the Demmel condition number of the MIMO channel matrices, and (iv) my Environment Characterisation Metric (ECM) comparing directly the discrete propagation paths in the channel. It turns out that the RCM shows a very close fit to measurements, making it well suitable to simulate channels in the kind of measured ones.

To satisfy the never-ending need for measurements, I conducted a *MIMO channel sounding campaign* at the University of Oulu, Finland. I measured a total number of 28 scenarios in three different indoor environments: offices, larger rooms, and a big hall. To compare the frequency dependence of the model parameters, each measurement route was sounded at 2.55 GHz and at 5.25 GHz. The thesis contains a comprehensive documentation of the measurement campaign.

Kurzfassung

Diese Dissertation stellt das neue *Zufalls-Cluster-Modell* (“Random-Cluster Model”, RCM) vor, ein geometriebasiertes, stochastisches Modell für frequenzselektive, zeitvariante MIMO Funkkanäle.

Das RCM verwendet das Konzept von sogenannten Clustern, d.h. Gruppen von Ausbreitungspfaden, um die Ausbreitungsumgebung zu modellieren. Im RCM ist diese Umgebung ausschließlich durh die *multivariate Verteilung* der *Clusterparameter*, wie die Clusterposition, Clusterbewegung und Clusterdispersion, beschrieben. Dadurch können auch Korrelationen zwischen Clusterparametern einfach wiedergegeben werden.

Die bedeutendste Eigenschaft des RCM ist, dass es durch eine automatische Methode direkt aus Kanalmessungen parametrisiert wird. Dadurch ist das RCM *spezifisch bezüglich der Ausbreitungsumgebung*. Es schließt damit die Lücke zwischen Kanalmessungen und Kanalmodellierung.

Die Verwendung von Clustern *reduziert die Komplexität* der Modellparametrisierung. Verglichen mit einem einzelnen Ausbreitungspfad, der 6 Parameter pro modelliertem Zeitpunkt benötigt, wird ein Cluster nur durch 21 Parameter für seine gesamte Existenzdauer (mehrere Zeitpunkte) beschrieben. Da ein Cluster für gewöhlich aus 6 bis 20 Ausbreitungspfaden besteht, wird die Anzahl der Parameter erheblich verkleinert.

Natürlich muss die Clusterverteilung sorgfältig beschrieben werden. Die Identifikation von Clustern aus zeitvarianten MIMO Kanalmessungen bildet die Basis für eine konsistente Parametrisierung des RCM. Da die visuelle Clusteridentifizierung in hohem Maße subjektiv und mühsam in der Anwendung ist, habe ich mich auf das lang gehegte Ziel der *automatischen Clusteridentifikation* konzentriert. Diese Dissertation vergleicht die Anwendbarkeit unterschiedlicher Clusterindentifikationsmethoden auf Ausbreitungspfade, insbesondere visuelle Clusteridentifikation, hierarchische Identifikation, K-means Identifikation, und Gauss’sche Mischmodell-Identifikation. Letztendlich stelle ich ein neues, vollständiges System zur automatischen Identifikation und zum Verfolgen von Clustern vor, das auf folgenden Teilen basiert:

(i) einem Anfangswertschätzer, der Cluster so weit wie möglich separiert,

(ii) dem *KPowerMeans*-Algorithmus, eine Erweiterung des K-means Algorithmus, der die Leistung von Ausbreitungspfaden berücksichtigt und die Mehrdeutigkeit von Winkeln ausnützt, und

(iii) einem Kalman-Filter zum Verfolgen von Clustern.

Durch dieses System wird das RCM automatisch aus Messungen parametrisiert.

Um nichtdiskrete Anteile im MIMO Kanal zu berücksichtigen, bietet das RCM als erstes Kanalmodell eine Möglichkeit *diffuse Mehrwegekomponenten* (“diffuse multipath”, DMP) zu modellieren.

Ich rege an, dass die *Modellvalidierung* eine *wichtige Aufgabe* darstellt, um ein Kanalmodell zu vollenden. Um das RCM zu validieren, vergleiche ich modellierte mit gemessenen Kanälen anhand von folgenden Validierungsmetriken: (i) Transinformation, die, wie sich herausstellen wird, keine markante Validierungsmetrik ist, (ii) Diversität, (iii) die Demmel-Konditionsnummer der MIMO Kanalmatrizen, sowie (iv) die von mir erdachte Umgebungscharakterisierungsmetrik (“Environment Characterisation Metric”, ECM), die direkt diskrete Ausbreitungspfade im Kanal vergleicht. Es stellt sich heraus, dass das RCM absolut geeignet ist, Kanäle zu simulieren, die gemessenen sehr ähnlich sind.

Um das endlose Bedürfnis nach Messungen zu stillen, habe ich eine *MIMO Kanalmesskampagne* an der Universität Oulu in Finnland durchgeführt. Ich habe insgesamt 28 Szenarien in drei verschiedenen Innenumgebungen gemessen: Büroräume, größere Räume, und einen große Halle. Um die Frequenzabhängigkeit der Modellparameter zu vergleichen, wurde jede Messroute jeweils bei 2.55 GHz und bei 5.25 GHz gemessen. Diese Dissertation enthält eine umfassende Dokumentation der Messkampagne.

Contents

List of Figures

List of Tables

Part I

The Random-Cluster Model

"You can always find an environment that fits your model"

Jørgen Bach Andersen

1 Introduction

The abbreviation MIMO stands for "multiple-input multiple-output". This rather inexpressive construct makes only sense in connection with the subject that is MIMO-ized. However, in common understanding, MIMO denotes a communication system that is able to exploit the features coming with multiple antennas at both link ends.

Publications in the area are talking about MIMO systems, MIMO communications, MIMO transceivers, MIMO algorithms, MIMO features, MIMO *everything*. The whole research field is exploding — a Google search returned approximately 3.2 million hits for "MIMO system", and Google scholar found 185.000 hits for "MIMO"[1].

In this introduction, Section 1.1 will first discuss what MIMO is, and why MIMO experienced such a hype in the field of radio communications. I will also motivate the need for channel models that accurately reflect the propagation environment. Subsequently, Section 1.2 will give a brief overview of the different approaches of modelling the MIMO channel and will discuss the "perfect" channel model. A common way to model the MIMO radio channel is by using "clusters". This controversial topic is fought out in Section 1.3. A novel approach of including the non-discrete part of the radio channel, the so-called diffuse multipath, is introduced in Section 1.4.

I will provide a first overview of cluster-based MIMO channel models currently available in Section 1.5. This section is to be seen in connection with Chapter 5, where I will compare the different cluster-based models with the new *Random-Cluster Model* presented in this work.

Finally, Section 1.6 will provide an overview of the three parts of this thesis.

1.1 Why MIMO?

So, why is there such a hype about MIMO? A very good introductory paper that addresses this point visions MIMO as "A Key to Gigabit Wireless" [1]. The reasoning for this was the finding of Telatar [2], which stated that when having a radio system employing N_{Tx} antennas at the transmitter (Tx) and N_{Rx} antennas at the receiver (Rx), one can achieve the $\min(N_{\mathrm{Tx}}, N_{\mathrm{Rx}})$-fold channel capacity of a single-link channel, *provided that all the channels between the Rx and Tx antennas are uncorrelated and Gaussian distributed*. This assumption is also called the *"iid. assumption"*.

This precarious promise of boosting the channel capacity was sufficient to attract enough attention for instantaneously becoming a new field of research. Hundreds of papers have been published on the unverified assumption that the propagation channel behaves in the way assumed by Telatar, mostly because the maths were tractable. The well-known Random Matrix Theory could be used straightforward.

[1] See `http://www.google.com`, and `http://scholar.google.com`. The number of hits were evaluated on 20/09/2007.

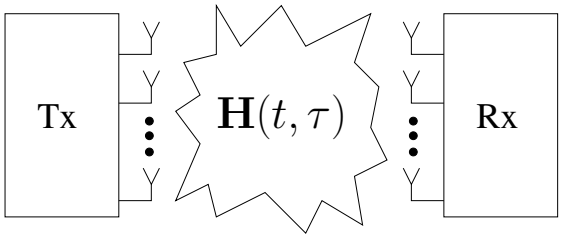


Figure 1.1: Schematic representation of the MIMO channel

Most surprisingly Telatar's main assumption, the iid. assumption was already disproved in [3, 4, 5][2], which was *12 years before Telatar's findings*. One should not forget that the iid. assumption was necessary to Telatar to come to closed-form solutions of his problem. He did not claim, that this assumptions would hold in real-world channels. Nonetheless a great number of researchers and engineers blindly took the iid. assumption for granted.

Fortunately, some groups also started to challenge the iid. assumption. A number of different research groups investigated the MIMO channel capacity from channel measurements almost at the same time [6, 7, 8, 9]. They all came to the same result: the MIMO capacity decreases significantly with higher correlation between subchannels. This correlation is due to both closely spaced antenna array elements, and correlated propagation due to line-of-sight transmission.

1.1.1 MIMO radio channel

Figure 1.1 presents a schematic view of the MIMO radio channel. Considering a MIMO channel with N_{Tx} antennas at the Tx, and N_{Rx} antennas at the Rx, the time-variant channel impulse response between the jth Tx antenna and the ith Rx antenna is denoted by $h_{ij}(t,\tau)$, with t denoting absolute time and τ denoting the time delay, the MIMO channel response is given by the $N_{\mathrm{Rx}} \times N_{\mathrm{Tx}}$ *MIMO channel matrix* $\mathbf{H}(t,\tau)$ with

$$\mathbf{H}(t,\tau) = \begin{bmatrix} h_{11}(t,\tau) & h_{12}(t,\tau) & \cdots & h_{1N_{\mathrm{Tx}}}(t,\tau) \\ h_{21}(t,\tau) & h_{22}(t,\tau) & \cdots & h_{2N_{\mathrm{Tx}}}(t,\tau) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_{\mathrm{Rx}}1}(t,\tau) & h_{N_{\mathrm{Rx}}2}(t,\tau) & \cdots & h_{N_{\mathrm{Rx}}N_{\mathrm{Tx}}}(t,\tau) \end{bmatrix}.$$

Given that the signal $s_j(t)$ is sent from the jth transmitter, the signal observed at the ith Rx antenna is given by

$$r_i(t) = \sum_{j=1}^{N_{\mathrm{Tx}}} h_{ij}(t,\tau) * s_j(t) + n_i(t),$$

where $*$ denotes the convolution operation, and n_i is additive noise at the Rx.

[2] Lee showed in [3] that the antenna correlation depends on the surrounding propagation environment, while the authors of [4, 5] showed that signals arriving at an antenna array are correlated, depending on the antenna spacing.

For a flat-fading channel[3] (i.e. there is no dependence on the delay), this system equation reduces to

$$\mathbf{r}(t) = \mathbf{H}(t)\mathbf{s}(t) + \mathbf{n}(t), \tag{1.1}$$

where $\mathbf{r}(t) = [r_1(t) \dots r_i(t)]$, $\mathbf{s} = [s_1(t) \dots s_j(t)]$, and $\mathbf{n}(t) = [n_1(t) \dots n_i(t)]$.

Equation (1.1) already shows that it should be possible to decode the transmitted signals at the receiver (provided $\mathbf{H}$ has full rank).

The theoretic assumption of Telatar for his derivations was that, for a flat fading channel, all channel coefficients h_{ij} are circular-symmetric complex-Gaussian independent identically distributed (iid.). In this case, the MIMO channel matrix has full rank, allowing for estimating the transmitted signal in a straightforward way, and the channel capacity rockets high. However, when the channel coefficients are *correlated*, this is no longer the case.

These channel coefficients are *determined by the underlying wave propagation* in the radio channel. Simply speaking, the richer the scattering in the environment, the less correlated the channel coefficients are. Because of this fact, the Gaussian iid. assumption is sometimes also termed as "rich-scattering assumption".

In non-line-of-sight (NLOS) environments, there is no direct path between the Tx and Rx. The transmitted waves undergo reflections, scattering, diffraction, etc., until they impinge at the Rx. By this effect, the channel coefficients become uncorrelated. Under these conditions, MIMO algorithms, especially spatial multiplexing[4] and space-time coding[5] schemes, perform extremely well. However, there are also many NLOS environments, where there is much scattering, but measurements showed that the channel coefficients are still significantly correlated [10, 11].

In line-of-sight (LOS) environments, there is a strong direct path between the transmitter and receiver antennas. Especially in these scenarios, the channel coefficients are correlated. MIMO algorithms suffer from this effect [12].

What is often forgotten is the influence of the signal-to-noise ratio (SNR) on the link quality. Under LOS conditions, the SNR is much higher than in NLOS conditions. From this perspective, LOS is most often advantageous over NLOS, also for MIMO communications [13]. The major advantage of MIMO systems is the efficient use of the rich scattering channel in NLOS environments to combat poor SNR by making use of the rich scattering in the channel. Beamforming and space-time coding schemes are of great help, here. In some indoor environments, there is even significant scattering although the stations have LOS. Of course, these scenarios offer the best gains.

MIMO-enabled devices[6] will have to work properly in all these kinds of scenarios. MIMO algorithms need to be flexible and need to deal will all different kinds of channels the environment will come up with. So, the algorithms have to adapt to the underlying channel. One way to achieve this flexibility is to simply switch between schemes given certain performance criteria or feedback information.

Concluding, MIMO systems truly offer significant advantages, but the key point is: *It is the radio channel that determines the ultimate performance bounds of a MIMO system.*

[3] When considering an OFDM system, the subchannels can be considered as frequency flat.
[4] Transmitting more than one data stream over multiple antennas
[5] Using diversity to improve the bit error ratio
[6] Yikes! Another MIMO-ization!

1.2 MIMO channel modelling

Since the MIMO channel is the determining part of the whole system, accurate models are strongly required. One needs to distinguish between models for MIMO deployment, and models for MIMO link level and system simulation.

Models for MIMO deployment target MIMO network planning, e.g. on optimizing MIMO base-station positions.

Models for system simulation try to reflect the general properties of the MIMO channel for many different scenarios. These models are designed for verifying signal processing algorithms and testing prototypes. This thesis focusses on this kind of models.

According to [14, 15], one can distinguish between two different kinds of MIMO channel models for system simulation, *physical models* and *analytical models*.

1.2.1 Physical models

These models have their roots in electromagnetic wave propagation. The channel is most commonly modelled by so-called "multipath components" (MPCs). One MPC is seen as a propagation path where a wave is originating at the Tx, undergoing several physical propagation effects, and impinging at the Rx. It is important to note that one MPC describes a discrete link between the stations characterised by physical parameters, like path loss, phase, delay, direction of departure (DoD), and direction of arrival (DoA). These unique links between DoAs and DoDs are a special property of every radio link, also called the "double-directional radio channel" [16]. Most of these models also assume that the transmitted and arriving waves have a plane wave front ("plane-wave assumption"). As a consequence, handling the waves as "rays", described by their direction, are useful.

The problem is to model the parameters of the MPCs, which can be seen as a task that is more art than science. The following paragraphs will provide a short overview of the most prominent approaches.

After having modelled the parameters of the MPCs, they are passed through the system model, which generates the MIMO channel matrix by including bandwidth, antenna arrays and other parameters discussed in Section 2.4.

Deterministic physical models

A deterministic way to model the channel is by ray tracing or ray launching. In this concept, the Tx and Rx are placed in a fixed geometry, and move with certain trajectories through this environment. Rays are undergoing reflection, diffraction and scattering on objects between the stations.

Another way to generate a scattering geometry is measurement-based deterministic MIMO channel modelling [17, 18]. The idea is to use MIMO channel measurements to characterise a propagation scenario but generate new channel realisations showing the same spatial properties as the measurements have. In this approach the location of scatterers are identified from measurements. Hence, for this method, MIMO measurements must be used, where the Tx and/or the Rx are *moving* along *known trajectories*. The measurements are then post processed using a high-resolution

algorithm (e.g. [19]), in order to estimate the delay, direction of arrival, and direction of departure of the underlying propagation paths. Using signal processing algorithms known from radar research, the (x/y/z)-location of the involved scatterers can be estimated [20], under the assumption of double scattering[7]. With the resulting positions of the scatterers, a "measured" environment is generated. Rx and Tx can then also be moved along trajectories in the environment that are *different* from the measured ones, generating new channel realisations.

An even more deterministic way to model the channel is to replay previously stored measurements, however this approach is limited to the system configuration that was used while conducting the measurements [21].

Geometry-based stochastic models

In this group of models, the parameters of the propagation paths are partly, or completely, defined in a stochastic way, while still taking a geometry of the channel into account. All models of this kind use the concept of *multipath clusters*.

The novel *Random-Cluster Model* presented in this thesis belongs to this type of MIMO channel models. For this reason I will comprehensively discuss the concepts of multipath clusters in Section 1.3, while the prominent models using the cluster concept are provided in Section 1.5.

Non-geometric stochastic models

Non-geometric stochastic models describe and determine physical parameters (DoD, DoA, delay, etc.) in a completely stochastic way by prescribing underlying probability distribution functions without assuming an underlying geometry (an example is the Saleh-Valenzuela model [22]).

1.2.2 Analytical models

Analytical models characterise the channel matrix directly in a statistical way, without dealing with physical wave propagation (yet sometimes faintly touching it).

The most simple model is to assume the channel coefficients as uncorrelated, iid. Gaussian, which does not at all reflect most of the scenarios[8].

The model accuracy improves when including correlations *at* the transmitter and the receiver, or even better, *between* the transmitter and the receiver. Channel matrices are then generated by a noise-colouring process assuming a Rayleigh-fading channel. Depending on the accuracy and complexity of the model, correlation of the individual links between Tx and Rx antennas are introduced.

A very popular, yet not very accurate analytical model is the Kronecker model [23]. It assumes the channel correlation matrix (holding the correlations between all links of the MIMO channel), to be separable into a transmitter correlation matrix and a receiver correlation matrix. This assumption also implies that the double-directional angular power spectrum, describing the linking of the

[7] The double-scattering assumption must be used for location estimation from full spherical data, because in three-dimensional space it is difficult for two lines to intersect. In contrast, single scattering can be easily identified in two-dimensional space, where lines always intersect (except when being parallel).

[8] However, that fact does not make this simple model less attractive for algorithm designers.

directions between Tx and Rx, is determined by an outer multiplication of the marginal spectra. In other words, the Rx receives always from the same directions, no matter in which direction the Tx transmits. This assumption may lead to a significant model mismatch as described in Section 1.5.3.

A very comprehensive discussion on the different kinds of analytical models and their applicability in different scenarios can be found in [14], while [24] provides the analytic connections between the family of such models.

However, these models are only accurate, when assuming the channel to fade according to a Rayleigh distribution. When dealing with LOS channels (or channels having a constant dominant path from a specular reflection), one observes Ricean fading. In this case, the correlation-based models need to be extended accordingly. Without this extension, some of the analytical models, like the Weichselberger model [25], still represent the spatial structure of the channel accurately, but not the fading characteristics. Unfortunately, this fact is often woefully neglected.

The analytical models usually focus on flat-fading channels. An interesting extension to frequency-selective channels was introduced by Costa and Haykin in [26]. A further extension and generalisation was suggested in [27].

1.2.3 What is the perfect model?

Every channel model needs a *focus*. All-in-one models claiming to be suitable for every purpose are usually unfeasible all-rounders.

In my opinion, the "perfect" model should

- focus on the need of the model user,
- be able to accurately model the kinds of environments the model focuses on,
- fit measurement data, and
- perform with tractable complexity.

Of course this list is extensible in all respects, but it provides a good basis for developing a MIMO channel model.

1.3 Multipath clusters — a controversial topic

The question "What is a cluster?" ignited extensive, controversial discussions. While Part II will provide a comprehensive overview of clustering algorithms, the following introduction will provide an overview of opinions, both of the supporters and the sceptics. Let us in the first stage use a general definition of clustering [28]:

> *"The process of grouping a set of physical or abstract objects into classes of similar objects is called clustering. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. A cluster of data objects can be treated collectively as one group in many applications."*

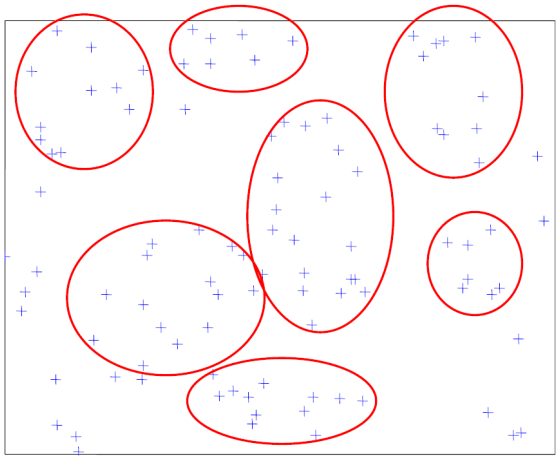


Figure 1.2: Are there clusters? (Figure taken with friendly permission from [29].)

Why should clusters exist?

Discussions start already about the existence of clusters. Using the previous definition, why should there be waves coming in clusters?

Let's go back to the basics of wave propagation. When the wavelength is in the order of the size of the surrounding objects, scattering will be the dominant mechanism, rather than reflections. At the receiver this scattering can be observed as a dispersed angular power spectrum (APS), or as a number of paths having similar directions.

Having a number of such scatterers in the considered environment, the APS will show a number of dispersed peaks. The paths belonging to these peaks could already be called clusters according to the definition above.

Also, a group of co-located reflections, e.g. from a big building far away, can be well seen as a cluster.

Doubts

The concept of clusters is not solved so straightforward as in the previous paragraph. Some people[9] played fast and loose with identifying clusters from figures showing a group of paths.

The discussion about "clusters" was blurred for a while, when there was not made the clear distinction between physical scattering objects, from which clusters originated and the MPC clusters themselves.

An impressive example was presented by Alister Burr[10], who used to be the strongest opponent to clusters. He showed following example (see Figure 1.2). A number of points are shown in two-dimensional space. On the first glimpse, one might think that there are clusters. The points are somehow grouped, there are also some outliers. One observes similar structures when looking at measurement data.

[9] Admittedly, I cannot exclude myself from this group during the time when I was a clustering greenhorn.
[10] Who gave his friendly permission to use his plot in this thesis

But, when he disclosed the fact that the location of these points were drawn from a uniform distribution, the identified clusters, of course, do not make any sense any more.

This example demonstrates that the visual identification of clusters is highly subjective, especially when it is not clear whether clusters do exist or not. Note that Burr did not consider the path powers in his example.

Very recently, his group developed a statistical test to identify whether there are clusters in measurements, or not [30]. This test includes the power of the paths, which is a necessary precondition. In their paper, they used their technique on a subset of the measurements to be presented in Part III, where they "found that the technique has good performance in automatically identifying clusters from multi-dimensional MIMO measurement data."

I interpret this result that (i) Prof. Burr has become a convert, (ii) clusters do exist in most realistic environments, and (iii) they can be identified automatically and even statistically.

Still, clusters will remain a controversial topic. In [31], the authors show that, when estimating propagation paths from measurements, clusters might be an artefact of an incomplete data model. A further discussion on this finding is provided in Section 6.3.2.

How to identify clusters?

Multipath cluster identification methods can be characterised between two extremes, visual clustering methods, and fully automatic clustering algorithms. In between, there are approaches that use a combination of automatic clustering and visual validation or alteration of the results.

Traditionally, *visual clustering* was used quite extensively. Even some prominent channel models, like the IEEE 802.11 TGn model [32] is based on results from visual clustering algorithms. However three significant problems were encountered:

(i) Even with a good definition of how to identify a cluster, any visual approach is highly subjective. Different people will always obtain different results.

(ii) Clustering a large number of scenarios and snapshots is very cumbersome.

(iii) The human eye is incapable of joint clustering in more than 2 parameter dimensions.

For these reasons, *automatic clustering* algorithms attracted the researchers' interest. The advantage of automatic approaches is that they offer a clear definition of what is identified as a cluster. However, the algorithmic definition of a cluster may not necessarily match the one that was intended at the start. Other interesting problems are to find an (optimum) number of clusters and to track the clusters over time. Part II is devoted to a comprehensive discussion of the available automatic clustering algorithms providing solutions to the problems described here.

Summary

Clusters simplify MIMO channel models. Instead of describing a large number of individual paths, only their superset, the clusters, need to be parametrised. For this reason, clusters became very popular. Even more, [33, 34, 35] showed that MIMO channel models disregarding clustering effects might significantly overestimate capacity.

Despite the popularity of clusters, it was long neglected to identify and parametrise them automatically and accurately. For this reason I invested significant effort to find a coherent clustering-and-tracking algorithm, which will be presented in Chapter 7.

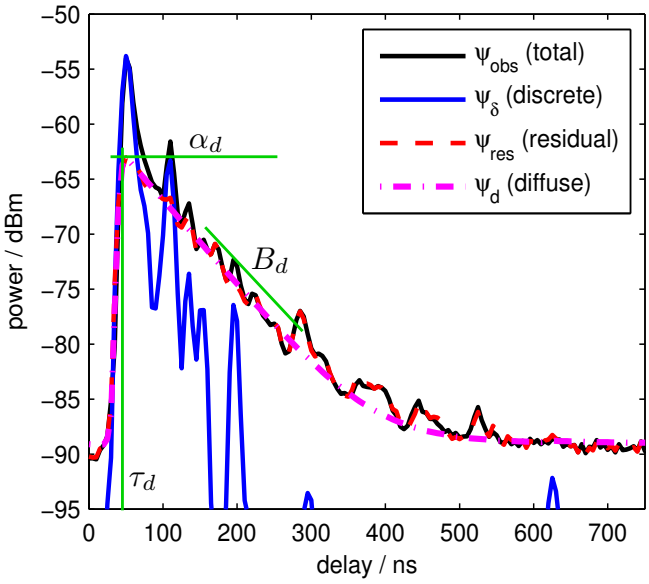


Figure 1.3: Diffuse power

An interesting question is, whether clusters correspond to real-world scattering objects. In a joint work with Giovanni Del Galdo, we tried to solve this issue, where it turned out that *most* clusters *do indeed* correspond to real-world scattering objects [20]. However, in complex environments, this is not always the case.

But let me ventilate the following counterquestion: Assuming I can use these clusters to model the MIMO channel accurately, is it necessary that one cluster uniquely corresponds to one physical object?

1.4 Diffuse multipath

The current advanced MIMO channel models use the concept of describing the radio channel by multiple paths that are clustered but discrete. This approach raises two questions: (i) How many paths need to be included in the model? (ii) How to model signal components that are not discrete (in delay, angles, Doppler)?

Recently, Richter [36] observed from MIMO channel measurements that, after estimating and subtracting discrete paths, a residual with a specific structure, the *diffuse multipath* (DMP), remains (due to e.g. reverberation of the room [37] or distributed scattering). The DMP is sufficiently described by only three parameters, which reduces the complexity of the modelling significantly.

Using the DMP in modelling provides answers to both questions raised above. The complexity can be significantly reduced by taking only dominant paths into account, while the DMP describes the non-discrete signal components in the radio channel. There is a certain lack of definition what paths are considered to be dominant. In my thesis, I consider paths that can be identified from

measurements with a certain minimum peak-to-noise ratio to be dominant (see Section 6.3.2 for estimation details).

After subtracting a number of discrete paths from measured channel impulse responses, an exponentially decaying residual power delay profile (PDP) remains. Figure 1.3 shows an example, where the *total observed Rx* PDP (averaged over all N_t transmit and N_r receive antennas), $\psi_{\mathrm{obs}}(\tau)$, (black line) can be approximated by estimated *discrete paths* with PDP $\psi_\delta(\tau)$ (blue), and a *residual* PDP $\psi_{\mathrm{res}}(\tau)$ (red). This residual PDP can be well approximated by an exponential PDP $\psi_d(\tau)$ describing the DMP (magenta).

Denoting the measured channel matrices with $\mathbf{H}_{\mathrm{obs}}(\tau)$ and channel matrices generated from discrete paths $\mathbf{H}_\delta(\tau)$, these PDPs can be calculated as

$$\psi_{\mathrm{obs}}(\tau) = \frac{1}{N_r N_t} \|\mathbf{H}_{\mathrm{obs}}(\tau)\|_{\mathrm{F}}^2, \tag{1.2}$$

$$\psi_\delta(\tau) = \frac{1}{N_r N_t} \|\mathbf{H}_\delta(\tau)\|_{\mathrm{F}}^2, \tag{1.3}$$

$$\psi_{\mathrm{res}}(\tau) = \frac{1}{N_r N_t} \|\mathbf{H}_{\mathrm{obs}}(\tau) - \mathbf{H}_\delta(\tau)\|_{\mathrm{F}}^2, \tag{1.4}$$

where $\|\cdot\|_{\mathrm{F}}^2$ denotes the squared Frobenius matrix norm.

Following [38], the exponential DMP PDP approximating the residual PDP can be described by

$$\psi_d(\tau) = \begin{cases} 0, & \tau < \tau_d \\ \alpha_d/2, & \tau = \tau_d \\ \alpha_d e^{-B_d(\tau-\tau_d)}, & \tau > \tau_d, \end{cases} \tag{1.5}$$

where τ is the (sampled) delay, B_d is the decay factor, α_d denotes the maximum diffuse power, and τ_d is the (sampled) base delay. This equation was initially suggested by [39] to model the total PDP of a channel with small bandwidth.

DMP has not been coherently implemented in any MIMO channel model up to now. I will show in Section 4.3.2.3 that models disregarding DMP will result in a mismatch in channel diversity. For this reason, I decided to include DMP in the Random-Cluster Model.

1.5 Cluster-based MIMO channel models

This section provides an overview of 4 prominent MIMO channel models, the 3GPP Spatial Channel Model (SCM), the WINNER II channel model, the IEEE 802.11 TGn channel model, and the COST 273 channel model. These four models all base on the concept of multipath clusters, but use the concept in a different fashion.

While the next paragraphs will introduce the basic concepts and important features of the models, I will compare them to the Random-Cluster Model in Chapter 5.

1.5.1 3GPP SCM model

The 3GPP spatial channel model (SCM) defined in [40] and comprehensively described in [41] consists of two parts, a *calibration model* and a *simulation model*.

The very simple calibration model is just used for verifying whether signal processing algorithms are implemented correctly. The more interesting simulation model, is thought for link-level simulations of one (or more) mobile stations (MS) connected to a base station (BS) in a single cell. So, I will concentrate on the simulation model. To describe the basic concepts I consider the single-link case.

The simulation is carried out for "drops", where a drop is defined as a simulation run over a specified (small) number of frames. During a drop the channel undergoes fast fading according to the motion of the MS, but the global parameters are assumed to be constant during a drop.

The SCM is a generic model[11] that differentiates between three different kinds of environments, suburban macro, urban macro, and urban micro, and is parametrised for each of these environments individually.

The geometry of the environment needs to be specified by the user, but is basically limited to the distance between MS and BS, and the BS antenna orientation. The path loss is calculated by the well-known COST 231 models [42], depending on the environment chosen.

Surprisingly, in all environments there is a fixed number of 6 clusters[12], and a number of 20 paths within each cluster. All the clusters show *the same azimuth spreads*, and *no delay dispersion*. So, all paths within a cluster are placed at deterministic offsets to the cluster centre position to exactly match the desired azimuth spread.

The SCM defines the global delay spread and the global azimuth spreads for the different environments. The *cluster* mean delays are chosen such that they represent the global delay spread of the channel. Also the *cluster* centre positions are chosen such that they represent the global azimuth spread of the channel at the Tx and at the Rx, correctly.

The SCM has several optional features: (i) a polarization model, (ii) far scattering clusters, (iii) a LoS component for the micro-cellular case, and (iv) a modified distribution of the angular distribution at the MS, which emulates propagation in an urban street canyon, (v) including correlations of the global parameters. All these options lead to a more accurate parametrisation of the cluster position, but do not change any of the facts stated before.

Summarising, the 3GPP SCM model focuses on BS to MS links. Unfortunately, it is parametrised for only three very general scenarios that were foreseen for MIMO communications. Also, the concept of zero-delay clusters is quite questionable.

1.5.2 WINNER II channel model

The WINNER channel models are an extension of the 3GPP SCM model and comprise of all the features that the SCM model has, including a significant number of extensions. I will discuss the latest WINNER Phase II interim channel model [43], since it is the most complete and mature one. This channel model is also intended for system simulations with many BSs and MSs. It consists of two different kinds of approaches: a generic simulation model, and a "clustered delay line" model having lower computational complexity but also lower accuracy. Again, I will use the generic simulation model for a single link to discuss its concepts.

[11] A generic model has the same (generic) model structure for all environments

[12] Confusingly, a cluster is called "path" in the SCM, while the paths within a cluster are called "subpaths". For consistency, I will keep with my nomenclature in this thesis to allow for straightforward comparison.

The WINNER II model distinguishes between 19 environments[13].

The drop concept of the 3GPP SCM model is inherited, but, to enable *correlation between subsequent drops*, the large-scale parameters are correlated.

Based on the geometry, the path loss is calculated using individual formulas for the different kinds of environments.

The number of clusters varies between 4 and 20, while the number of paths within each cluster is again considered constant as 20. Also the cluster azimuth spreads are constant, and the azimuth position of each path within a cluster is tabularised.

A central assumption of the WINNER model is that the composite azimuth power spectrum is "wrapped Gaussian". For this reason, the cluster delays and azimuth spreads are chosen such that they match the composite distribution. Optionally, also elevation is considered in indoor environments.

One novelty of the WINNER II model is that the two strongest clusters show dispersion in delay. This is achieved by splitting each of these two clusters in three sub-clusters, where there is a constant delay offset of 5 ns between each sub-cluster. LOS links are more thoroughly considered than in the SCM model. For bad-urban channels, WINNER II also employs an updated far-cluster approach.

Summarising, the WINNER II interim model is an all-purpose model that is parametrised meticulously for a large number of scenarios. A shortcoming is the model's strong focus on the global parameters of the environment (delay spread and angular spreads), while clusters all show the same size. This model structure gives rise to the apprehension that the model will perform quite inaccurately in indoor environments. On the other hand, this model is the currently most advanced (quasi-standardised) one and should be able to reflect the properties of the environment best.

1.5.3 IEEE 802.11 TGn Channel model

The IEEE 802.11 TGn channel model focuses on indoor wireless LAN channels. The generic model is parametrised in detail for 6 different kinds of channels (A – F), distinguished by their increasing delay spreads.

The TGn model uses a number of clusters that are parametrised in the double-directional/delay domain. The shape of the clusters is Laplacian in angles and exponential in delay, where clusters may overlap in these domains. The model is implemented using a tap-delay line with up to 18 taps. Clusters extend over several taps, where their power is decaying in each tap.

The central point of the TGn model is that the MIMO channel matrix *for every tap* is generated by the Kronecker model, assuming that the double-directional power azimuth spectra (PAS) can be separated into the Tx PAS and Rx PAS. The Rx and Tx correlations can be calculated directly from the PAS given a certain antenna geometry. By this way, the system model is already included in the clustering approach.

This model performs well for taps containing only one cluster. However the TGn model creates huge errors, when there are more than one cluster in one tap. Figure 1.4 demonstrates this concept. In Figure 1.4a the double-directional APS clearly shows two clusters[14]. These two clusters

[13] These environments can be further differentiated into 4 groups, but each of the 19 scenarios is individually parametrised. No doubt, a vast number of measurements were necessary to achieve this.

[14] Model B was chosen from the TGn parameter sets. According to the model, the clusters show Laplacian shape.

are also reflected in the marginal distributions. When trying to model this scenario using the Kronecker model, the resulting APS is given by an outer multiplication of the marginal spectra. This effect is demonstrated in Figure 1.4b. While the original APS showed two independent directions, these directions are after modelling fully coupled. Resulting, the available channel diversity is significantly overmodelled.

Note that this problem could easily be overcome by modelling each individual cluster by a Kronecker-generated channel matrix. Finally, the channel matrices for multiple clusters in one tap just need to be summed up.

The interesting property of the TGn model is that its clusters show significant delay spread. This effect was not considered in the 3GPP SCM model, while the WINNER model considers a delay spread only for the two strongest clusters. For this reason, the TGn model is better suited for indoor environments.

The TGn model considers indoor Doppler spectra as well as Doppler spectra from passing cars. Including the Doppler modelling, the model is time variant for the duration of a transmission block.

Summarising, the TGn model is a model tailored to indoor WLAN environments. The concept of placing clusters is quite interesting, and the model is also well parametrised. However, the way of generating clusters by using the tap-wise Kronecker model is faulty by design. By a very simple extension, this problem could easily be overcome. By the implicit assumption of an antenna array structure, the TGn model is well-suited for algorithmic simulations, but not for simulating under real-world conditions.

1.5.4 COST 273 model

The COST Actions look back on a long and successful history of creating channel models for wireless systems. The latest model, the COST 273 channel model is again on the leading edge of research and provides a plenitude of different options and implementation aspects. Since the Random-Cluster Model is rooted in the COST 273 model, I will discuss its features in greater detail.

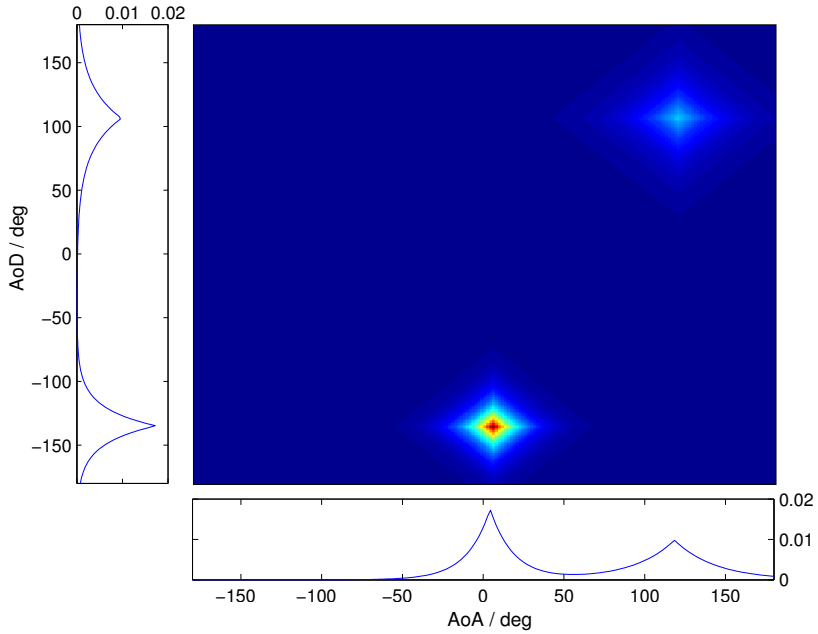
To introduce the COST 273 model, I would like to cite Prof. Pertti Vainikainen[15] (Helsinki University of Technology),

> *"The COST 273 model is rather a collection of good ideas than a model itself".*

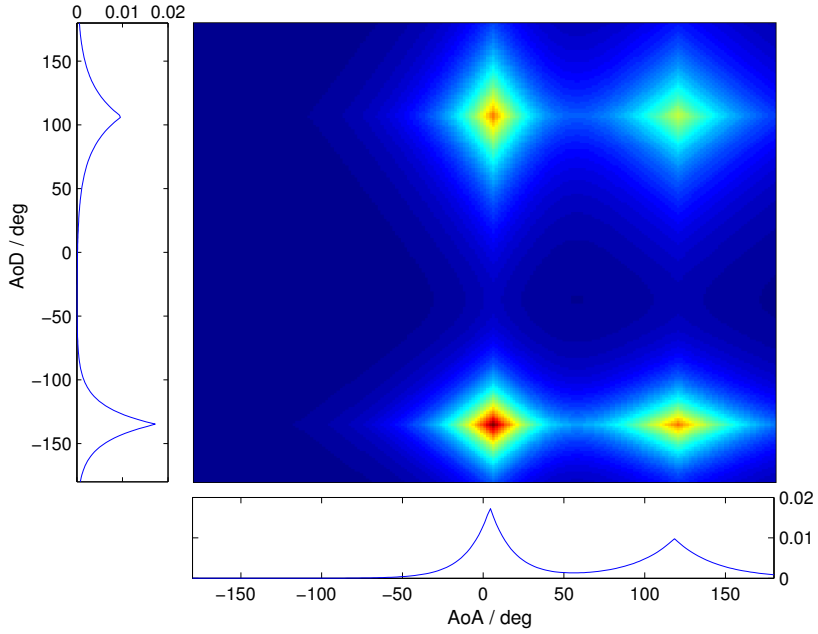
The COST 273 model emerged from its predecessor COST 259, hence one can observe many similarities. Indeed, one design issue of the COST 273 model was to be partially backwards-compatible. Despite the possible advantages of this intention, this can also be seen as a shortcoming, as the model clutches to old structures. Especially the process of cluster creation is a mixture of old concepts and new ideas, which makes this part of the model difficult to understand and to implement. Also, there are some parts of the model to be interpreted creatively.

However, the novel idea of multiple interaction clusters used in the COST 273 model [44] is a considerable improvement in MIMO channel modelling. The concept enables clearly linked

[15] With friendly permission by him; stated on the 3rd COST 2100 management committee meeting, Duisburg, Germany, September 2007.



(a) True double-directional azimuth power spectrum, including marginal spectra



(b) Artifacts introduced in the APS by using the Kronecker channel model

Figure 1.4: Effect of Kronecker-like modelling of more than one cluster. The Kronecker approach models the double-directional angular power spectrum by an outer multiplication of the marginal spectra. As result, diversity is significantly overmodelled.

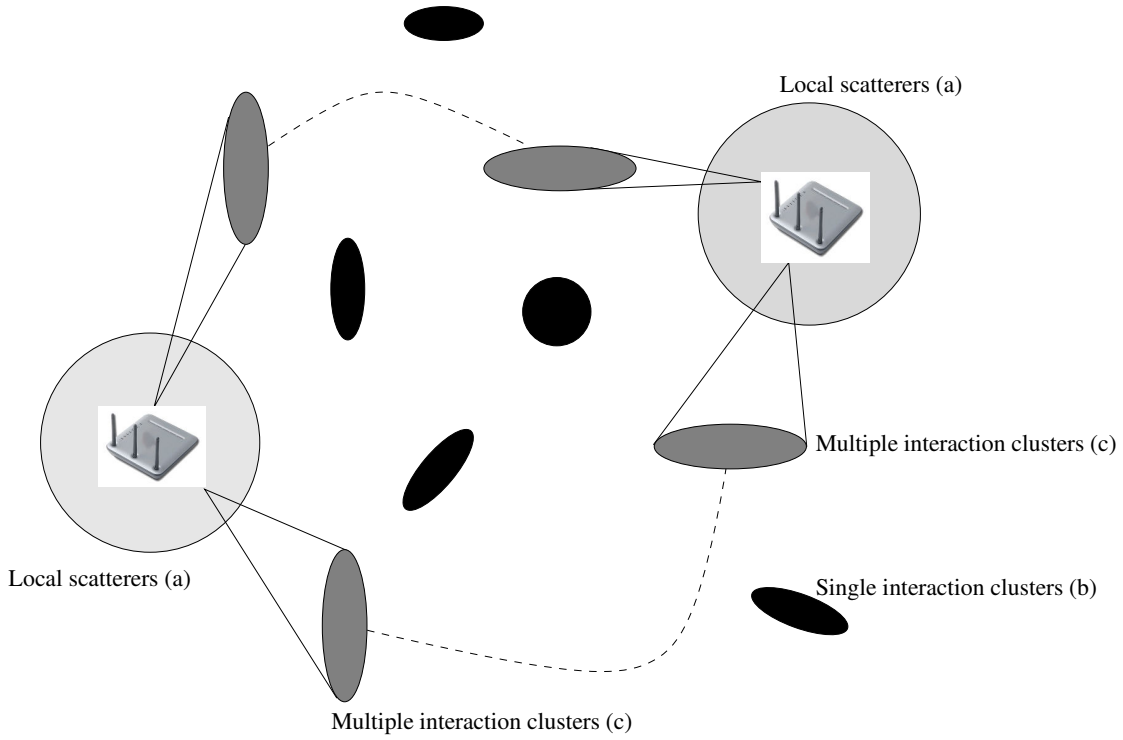


Figure 1.5: Different kinds of clusters in the COST 273 model: (a) Local scattering clusters, (b) Single interaction clusters, (c) Multiple interaction clusters

propagation paths as well as Kronecker-like[16] channels. For the multiple interaction clusters, the geometry of the channel is defined by *statistical means only*. Furthermore, the internal cluster structure also allows for an implementation with low complexity.

Also the COST 273 model is a generic model that is foreseen to be parametrised for a large number of scenarios. Unfortunately, the parameter tables are not completely filled, yet.

The next paragraphs provide an overview of a number of features that influenced me while developing the Random-Cluster Model.

Cluster $\neq$ Cluster

The COST 273 concept knows three different kinds of clusters, which are created by different means (Figure 1.5).

To my understanding the different cluster types are:

(a) Local scattering clusters

In the COST model, there is always a cluster around the MS, but only in some scenarios a cluster around the BS.

[16] Find the discussion about the effects of the Kronecker model in Figure 1.4.

(b) Single interaction clusters

These clusters were already introduced in the COST 259 model and were taken over into this model for compatibility reasons. Single interaction clusters are placed at a fixed position in space, where they do not move. “Visibility regions” define whether the clusters are currently active or inactive.

(c) Multiple interaction clusters (MIC) / Twin clusters

This kind of clusters was introduced recently in [44]. One can interpret this as an extension of the double-directional propagation model [16] to clusters. It does not matter for the Tx or Rx with which propagation mechanism the signal from one MIC has to cope with. MICs are only described by the mean parameters (delay, angles) and their spreads. This makes this approach computationally advantageous. These two “faces” of a twin-cluster are connected by a link delay. When trying to identify this link delay from measurements, one might obtain also negative values. This effect is due to the inaccurate identification of clusters. So, approaches, that use twin-clusters but do not rely on a geometrical interpretation might be advantageous.

The cluster generation in the COST model works as follows:

(i) Create the local cluster(s) at MS and/or BS, depending on the chosen environment

The (external) “selection parameter” K_{sel} describes the power ratio of single-interaction clusters to additional multiple-interaction clusters.

(ii) Create single-interaction clusters with power according to K_{sel}.

(iii) Create additional multiple-interaction clusters to fit K_{sel}.

Note that the COST model tries to combine three ideas to one concept, which might be too much of a good thing.

Cluster positions

The twin-clusters are positioned statistically, but there is no exact description how to place the other types of clusters.

Visibility region

A spatial region for each cluster defines whether the cluster is regarded as active (sometimes also called “existent”) or inactive (“non-existent”). If the Tx/Rx is moving into (or out of) the visibility region, a smooth transition from inactive to active (or vice-versa) is considered. The number of active clusters is determined by the number of visibility regions. This implies that the total number of clusters is large, but only a certain subset of these clusters are visible.

Time variance

Time-variance in the COST model is usually achieved by moving Tx and Rx along a movement route, because clusters are fixed in space. Alternatively, time variance can also be achieved by assigning twin-clusters a velocity vector for the Tx and Rx side separately.

Diffuse scattering

"Diffuse scattering[17]" is foreseen to be modelled by a mean diffuse power level and its standard deviation. The PDP of the diffuse component is uniform in azimuth and exponential in delay. Note that there has not been any implementation or attempt to parametrise the diffuse multipath in the COST 273 model yet.

Summary

Also the COST model claims to be an all-rounder. It is more general than the WINNER II model, but was not yet parametrised accurately. Since the description of the model is very general, there is no clear line of how to implement the COST 273 model. There is currently only a single implementation presented in [45], which does not contain all model features yet. Moreover the model parametrisation relies on "educated guesses"[18]. However, the COST 2100 Action, which is the successor of COST 273, has decided to finally parametrise and simplify this model.

1.6 Overview of this thesis

This thesis is organised in three parts.

Part I presents my main contribution, the *Random Cluster Model* (RCM), a novel wideband, propagation-based stochastic, cluster-based, smoothly time-variant MIMO channel model for system simulation that is parametrised from measurements. Admittedly, these are quite a lot of buzz words for a single model, but all of them together render the RCM. Chapter 2 describes the model structure of the RCM, and details how to use the model for system simulation. I outline the validation process and the validation metrics in Chapter 3. Model validation and parametrisation are closely connected, so I decided to present the RCM's validation jointly with the different parametrisation approaches in Chapter 4. Finally, I conclude this part by comparing the features of the RCM to other MIMO channel models in Chapter 5.

Part II discusses *multipath clustering*. Profound questions about the existence of clusters have already been discussed in this introduction. So, assuming that there are clusters, Chapter 6 concentrates on the different ways to identify clusters from measurements. For parametrising smoothly time-variant environments, clusters need to be tracked over time. Chapter 7 comprehensively describes a novel automatic framework for jointly identifying and tracking clusters from MIMO channel measurements, and provides resulting cluster parameters.

Part III describes the respective *measurement campaign*. Measurements are vital to coherently parametrise the RCM. This part describes the measurements that were used for the parametrisation and validation of the RCM. First, Chapter 8 outlines the objectives of the measurement campaign. Subsequently, Chapter 9 comprehensively documents the measurement campaign that I conducted at the University of Oulu with support of the specialists from EB and from the Centre of Wireless Communications (CWC). An overview of the measurement routes is provided in Appendix A, while each individual measurement route is described in detail in Appendix B.

[17] Furtheron I will use the more accurate term "diffuse multipath"

[18] Playing devil's advocate, one could also substitute with "wild guesses". Sometimes a good guess is better to having no parameter at all. When the model finally performs well with these parameters, the intuition was right. Otherwise one needs to guess again, or — what I would suggest — finally do measurements.

Concluding, **Chapter 10** contains my final remarks, thoughts, and ideas for possible future research in the fascinating field of MIMO radio channel modelling.

2 Random-Cluster Model description

2.1 Model Structure

The Random-Cluster Model (RCM) is based on the concept of multipath clusters. It is rooted in the COST 273 MIMO channel model [46]. I introduced two major improvements: (i) the environment is described only by *statistical* parameters, (ii) the effect of diffuse multipath is incorporated. Also I harmonised the clustering concepts and simplified the temporal implementation. The RCM including these improvements will be detailed in the following sections.

The RCM consists of two major parts, the *parametric channel model*, and the *system model* (see Figure 2.1):

(i) The *parametric channel model* (see Section 2.3) is used to generate clusters and paths according to the environmental parameters. In order to make the parametric model work most accurately, it has to be parametrised directly from measurements.

(ii) The *system model* describes the radio system, and is used to compile the frequency-dependent *channel matrix* from the parametric channel data (see Section 2.4). The radio system comprises the antenna arrays, the carrier frequency, the transmission bandwidth, and the RF system impulse response. Additionally, diffuse multipath (components that cannot be resolved by the communication system) are included according to [36]. More details are provided in Section 2.4.2.

The novel approach of the RCM is to describe the geometry of the channel *only* by statistical cluster parameters. Clusters provide a compact way of describing the underlying propagation environment. To accurately parametrise the cluster parameters, they are extracted from measurements. An important feature of the MIMO channel also reflected by the model is the coupling between propagation paths in space and time, also known as the double-directional MIMO channel model [16]. To enable time-variance, Tx and Rx or clusters may move. By this, the RCM creates correlated snapshots in time of the propagation environment. The RCM also considers contributions that cannot be resolved by communication systems but still have significant impact on its performance. This so-called *diffuse multipath* is parametrised using a novel approach and generated according to [36].

Summarising, the model has following properties. It is

- cluster-based,
- propagation-based, but stochastic,
- double-directional,
- time-variant, and
- unresolvable components are described by the *diffuse multipath concept*.

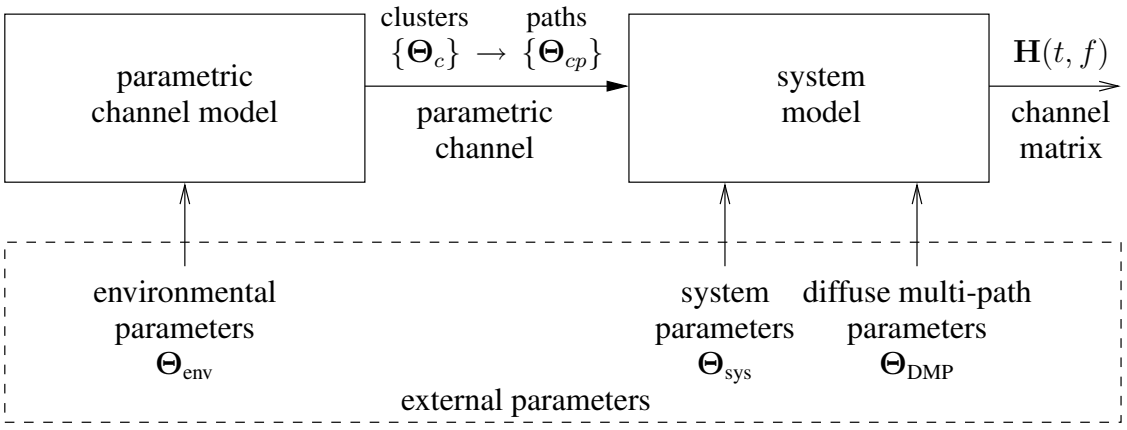


Figure 2.1: Basic structure of the Random Cluster Model (parameters will be detailed in the next sections)

Focus of the RCM

The main focus of the RCM is system simulation, for both algorithm testing and device testing. It is well suited to reflect time-variant scenarios that are similar, but not equal to ones measured before.

A major feature is that the parametrisation of the RCM originates from measurements. Having measurements available, the parametrisation is achieved automatically. In this way it perfectly fills the gap between channel *sounding* and channel *simulation*.

What this model does not provide

By the way it is parametrised, the RCM is very specific in reflecting a certain type of environment. Being rooted in the COST 273 model, one might think that the RCM is an all-purpose model. The model user shall be warned that it does not perform like this. Many aspects that make a model very general are intentionally neglected in the RCM in order to reduce complexity, e.g. a proper path loss calculation, or a description of general environments.

For some scenarios, the RCM will still perform better than other (even standardised) models available, but proper parametrisation is always necessary.

The RCM is definitely not intended for supporting MIMO deployment. Since the model does not include any geometry, it is not suited for channel prediction in specific environments, particularly in those that were not measured before.

In the following, the RCM together with the system model are described in more detail.

2.2 General model description

An important feature of the MIMO propagation channel is the occurrence of multipath components (MPCs) in clusters. The COST model[1] also uses the cluster concept to create the parametric

[1] In the following, the term "COST model" always refers to the most recent COST 273 MIMO channel model.

channel. Multiple clusters are placed in space to emulate the propagation environment. A *cluster* consists of *multiple propagation paths* with *the same parameter distribution*. The implementation of clusters and their parametrisation is described in the following sections.

The concept of clusters is introduced in the following way [47]: Clusters are examined from the Tx and the Rx side separately. They show a mean angular position and a mean angular spread[2] when seen from the Tx and Rx, respectively, as well as a mean delay and a delay spread. One can interpret this as an extension of the double-directional propagation model [16] to clusters. It does not matter for the description of a cluster, which underlying propagation effects it stems from. Clusters of multipath components are only described by their mean parameters (delay, angles, powers) and their spreads. This makes this approach computationally advantageous.

Additionally, the concept of diffuse multipath captures components that cannot be resolved by discrete paths.

Mathematically, the channel modelled according to the RCM can be expressed as

$$\begin{aligned}\mathbf{H}(t,f,\boldsymbol{\Theta}_{\text{env}}) &= \underbrace{\mathbf{H}_\delta(t,f,\boldsymbol{\Theta}_{\text{env}})}_{\text{discrete components}} + \underbrace{\mathbf{H}_d(t,f,\boldsymbol{\Theta}_{\text{DMP}})}_{\text{diffuse multipath}} \\ &= \sum_{c=1}^{N_c} \mathbf{H}_c(t,f,\boldsymbol{\Theta}_c) + \mathbf{H}_d(t,f,\boldsymbol{\Theta}_{\text{DMP}}) \\ &= \sum_{c=1}^{N_c} \mathbf{H}_c(t,f,\bar{\tau}_c,\bar{\varphi}_{\text{Tx},c},\bar{\varphi}_{\text{Rx},c},\bar{\theta}_{\text{Tx},c},\bar{\theta}_{\text{Rx},c},\sigma_{\tau_c},\sigma_{\varphi_{\text{Tx},c}},\sigma_{\varphi_{\text{Rx},c}},\sigma_{\theta_{\text{Tx},c}},\sigma_{\theta_{\text{Rx},c}},\dots) \\ &\quad + \mathbf{H}_d(t,f,\boldsymbol{\Theta}_{\text{DMP}}),\end{aligned} \tag{2.1}$$

where $\mathbf{H}_c(\cdot)$ denotes the cluster impulse response of the cth cluster, $\mathbf{H}_\delta(\cdot)$ denotes the impulse response from all clusters (which all contain discrete components only), $\mathbf{H}_d(\cdot)$ denotes the channel from diffuse (unresolvable) multipath. The environment is specify in the environment parameter $\boldsymbol{\Theta}_{\text{env}}$, cluster parameters are described by the cluster parameter set $\boldsymbol{\Theta}_c$, and $\boldsymbol{\Theta}_{cp}$ is the path parameter set. The parameters of the diffuse multipath are collected in $\boldsymbol{\Theta}_{\text{DMP}}$. All these parameter sets are described in the next sections in detail.

The *cluster impulse response* is given as the sum of its path responses as

$$\begin{aligned}\mathbf{H}_c(t,f,\boldsymbol{\Theta}_c) &= \sum_{p=1}^{N_p} \mathbf{H}_{cp}(t,f,\boldsymbol{\Theta}_{cp}) \\ &= \sum_{p=1}^{N_p} \mathbf{H}_{cp}(t,f,\tau_{cp},\gamma_{cp},\varphi_{\text{Tx},cp},\varphi_{\text{Rx},cp},\theta_{\text{Tx},cp},\theta_{\text{Rx},cp}),\end{aligned} \tag{2.2}$$

where the individual path parameters are explained in more detail in Section 2.3.1.1, while the cluster parameters are discussed in Section 2.3.1.2.

I apply the following three steps to parametrise the model:

[2] The term "spread" always refers to the rms spread values, and if not stated otherwise, for the *cluster* spread value [48, 11], in contrast to a global spread value [49].

(i) The environment is characterised exhaustively by $\mathbf{\Theta}_{\text{env}}$. The tremendous advantage of the RCM is that the *parametrisation of the environment is achieved by statistical means only.* The parameter $\mathbf{\Theta}_{\text{env}}$ describes the properties of the environment by a *probability density function* (pdf) of the cluster parameters.

(ii) The parametric model creates a number of clusters with corresponding cluster parameter sets $\{\mathbf{\Theta}_c\}$ in accordance with $\mathbf{\Theta}_{\text{env}}$.

(iii) According to the cluster parameter sets, the parametric model creates a number of MPCs with corresponding path parameter sets $\{\mathbf{\Theta}_{cp}\}$.

The parameters of the system model, and of the diffuse multipath, are collected in $\mathbf{\Theta}_{\text{sys}}$, and $\mathbf{\Theta}_{\text{DMP}}$, respectively.

The exact meaning of the parameters will be explained in the next sections. It should be noted that the strength of the channel model lies in the parametric model, which will provide these many parameters (contained in $\mathbf{\Theta}_{\text{env}}$, and $\mathbf{\Theta}_{\text{DMP}}$) accurately. The mapping from the parameter domain to the impulse responses is straight forward and accomplished by the system model (see Section 2.4).

Before going into detail with the parametrisation of the RCM, I will present two different interpretations of this model, in the angular and in the geometrical domains.

2.2.1 Cluster interpretation: angular domain

For the sake of simplicity I will describe the angular interpretation only in the AoA/AoD domain, omitting elevations, power, and delay. Figure 2.2 demonstrates the basic concept. The RCM places clusters in the azimuth domain, where the location and parameters are determined according to pdfs. Experiments suggest clusters usually exhibit an elliptical shape in the azimuth domain [50].

2.2.2 Cluster interpretation: geometric domain

The demonstration in azimuth domain already showed the simplicity and low complexity of the model. To understand time-variant modelling, I complement the angular interpretation by a geometrical interpretation as used in [47].

Of course, clusters show different (angular) positions and angular spreads. Therefore, the description of a cluster can be split up into its Rx and Tx properties when independently observed from Rx and from Tx. Figure 2.3 illustrates the idea of this *twin-cluster* concept. The angular position of the twin-cluster can be mapped easily to a position in a coordinate system[3] by means introduced in [47]. Figure 2.3 shows this interpretation in the $x/y/z$-domain. A cluster is virtually split up between Rx and Tx into a twin cluster. The twins are linked by a link delay, $\bar{\tau}_{\text{link}}$. The extent of a cluster in space is modelled by its delay spread. In this interpretation, clusters always have a circular shape, of some size, but the distances to Tx and Rx, and thus the angular spreads, are different. (Note that this is no contradiction to the angular interpretation.) The parameters of the MPCs *within* a cluster are *the same* when seen from Tx and Rx.

[3] This interpretation is only provided for better understanding, the representation in $x/y/z$-space is *not used* by the RCM.

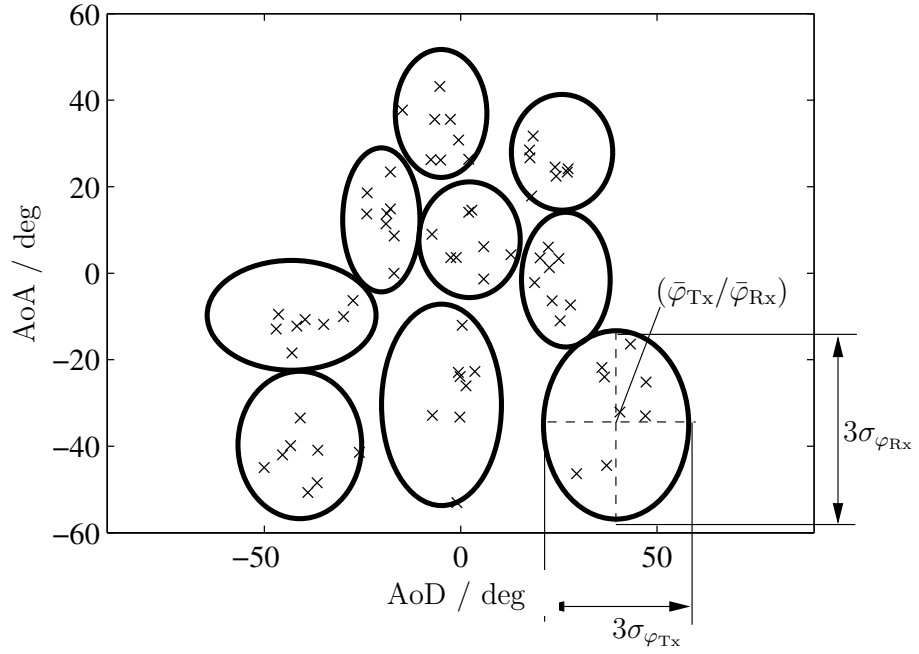


Figure 2.2: Cluster interpretation in the angular domain. Each cluster has a centre position $(\bar{\varphi}_{\mathrm{Tx}}, \bar{\varphi}_{\mathrm{Rx}})$ and angular spreads $\sigma_{\varphi_{\mathrm{Tx}}}, \sigma_{\varphi_{\mathrm{Rx}}}$. MPCs are placed in the cluster according to the parameters.

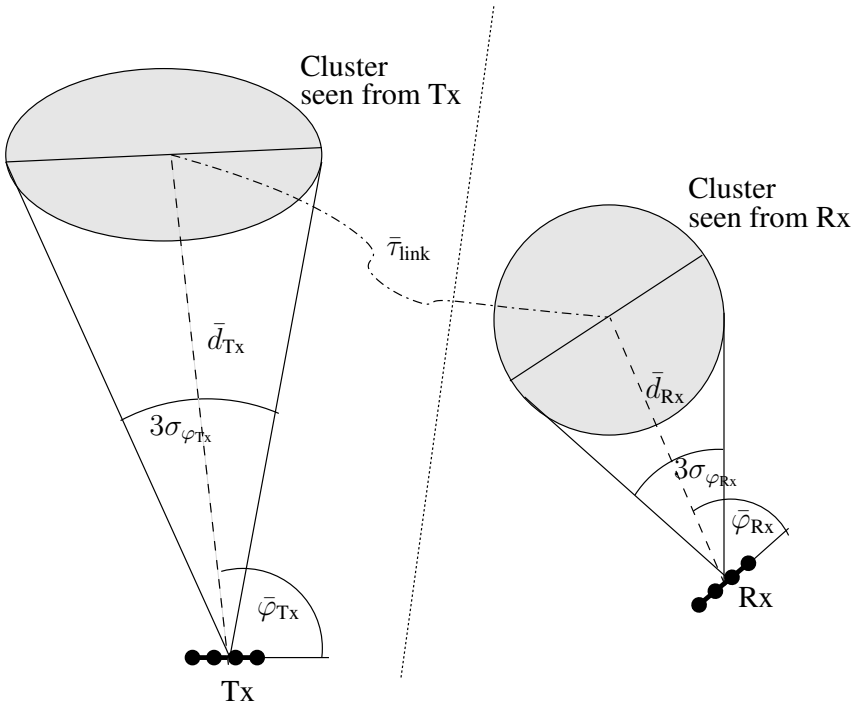


Figure 2.3: Geometrical interpretation of the RCM demonstrated for a single cluster.

By introducing the concept of moving clusters, consecutive *channel realisations are inherently correlated in time*.

A significant problem of using the twin-cluster representation for *modelling* [45] is to accurately reflect the link delay $\bar{\tau}_{\text{link}}$. Experiments showed that this delay might become negative for reasons of inaccurate cluster identification, which eventually lacks a physical interpretation..

For this reason, the RCM describes movements already in the AoA/AoD/delay-domain, in contrast to taking the detour via the $x/y/z$-domain. By this the modelling of the link delay is easily circumvented. Nevertheless, the twin-cluster concept remains valid also for the RCM as a *representation*.

2.3 Parametric model

The responsibility of the parametric model is twofold. First, it specifies external parameters that make it possible to characterise environments as accurate as desired, but as general as possible. Secondly, from these external parameters, it determines the internal parameters to describe the radio propagation down to the properties of individual propagation paths. The only external parameter for the parametric model is the environment parameter pdf $\mathbf{\Theta}_{\text{env}}$.

The environment parameter $\mathbf{\Theta}_{\text{env}}$ is a single, but multi-dimensional probability density function of the cluster parameters. First, cluster parameters are drawn from this pdf. Then, path parameters are drawn according to the cluster parameters.

To keep the RCM most general, the environment is specified by a function. But to be exact, this function also needs parameters that can of course be expressed numerically. In Section 4.2 I will present three ways to parametrise this environment pdf.

In the following I will detail the pdf dimensions, i.e. the cluster parameters, which are needed to accurately model the radio channel. I will start with describing the smallest entity, one MPC in Section 2.3.1.1. A number of MPCs are grouped in one cluster, its parameters are characterised in Section 2.3.1.2. One snapshot of the propagation environment consists of a number of clusters. The environment parameter, describing the propagation scenario and specifying the cluster parameters, is detailed in 2.3.1.3. Finally, I will provide a comprehensive description of how the model is invoked to actually obtain cluster and path parameters from the environment parameters.

2.3.1 Parameters

The following paragraphs describe the path parameters, cluster parameters and environment parameter pdf.

2.3.1.1 Path parameters

Each cluster consists of a number of $N_{p,c}$ MPCs. Each of these paths p in the cluster c is described by the path parameter set $\mathbf{\Theta}_{cp}$ which contains the parameters indicated in Table 2.1.

Parameter	Description	Unit
γ_{cp}	complex amplitude (path gain) of one MPC in a cluster	1
τ_{cp}	delay of one MPC in a cluster	s
$\varphi_{\mathrm{Tx},cp}$	azimuth of departure (AOD) of one MPC in a cluster	rad
$\varphi_{\mathrm{Rx},cp}$	azimuth of arrival (AOA) of one MPC in a cluster	rad
$\theta_{\mathrm{Tx},cp}$	elevation of departure (EOD) of one MPC in a cluster	rad
$\theta_{\mathrm{Rx},cp}$	elevation of arrival (EOA) of one MPC in a cluster	rad

Table 2.1: Path parameters

2.3.1.2 Cluster parameters

The concept of the cluster parameters is to describe the paths within the cluster with *fewer parameters*. These parameters are typically the mean cluster position and the cluster spread, but also some additional parameters are necessary to accurately parameterise the environment.

Each cluster c is described by a cluster parameter set $\mathbf{\Theta}_c$, which contains the following parameters.

Cluster mean position

These parameters specify the positions of the clusters in parameter space. A representation of these parameters for one exemplary cluster is shown in Figure 2.2. The position of each cluster c is specified by the parameters in Table 2.2.

Parameter	Description	Unit
$\bar{\tau}_c$	cluster mean delay	s
$\bar{\varphi}_{\mathrm{Tx},c}$	cluster mean azimuth position when seen from Tx	rad
$\bar{\varphi}_{\mathrm{Rx},c}$	cluster mean azimuth position when seen from Rx	rad
$\bar{\theta}_{\mathrm{Tx},c}$	cluster mean elevation position when seen from Tx	rad
$\bar{\theta}_{\mathrm{Rx},c}$	cluster mean elevation position when seen from Rx	rad

Table 2.2: Cluster mean position parameters

Cluster spread parameters

The cluster spread parameters allow for modelling the *extent of the clusters* in parameter space. A representation of these parameters for one exemplary cluster is shown in Figure 2.2. The following set of parameters describe the spreads of each cluster c in the individual dimensions as detailled in Table 2.3.

Parameter	Description	Unit	Parameter bounds
$\sigma_{\tau,c}$	cluster delay spread	s	≥ 0
$\sigma_{\varphi_{\mathrm{Tx},c}}$	cluster rms azimuth spread seen from Tx	rad	≥ 0
$\sigma_{\varphi_{\mathrm{Rx},c}}$	cluster rms azimuth spread seen from Rx	rad	≥ 0
$\sigma_{\theta_{\mathrm{Tx},c}}$	cluster rms elevation spread seen from Tx	rad	≥ 0
$\sigma_{\theta_{\mathrm{Rx},c}}$	cluster rms elevation spread seen from Rx	rad	≥ 0

Table 2.3: Cluster spread parameters

Power parameters

The power of a cluster $\sigma^2_{\gamma,c}$ describes the sum power of all paths within a cluster.

Another important parameter for modelling is the power of a snapshot[4] containing a number of clusters. In environments with different snapshot powers, cluster parameters can be quite different. For this reason, a cluster must also be parametrised by the snapshot power it exists in, ρ_c. This parameter subsequently allows for selecting the *kind of environment* the cluster occurs in. The cluster power parameters are specified in Table 2.4.

Parameter	Description	Unit
$\sigma^2_{\gamma,c}$	cluster power	dB
ρ_c	total snapshot power, where cluster c exists in	dB

Table 2.4: Cluster power parameters

Number parameters

Every cluster has two number parameters, (i) the number of MPCs within a cluster, and (ii) the number of clusters in the current snapshot coexisting with the described cluster (see Table 2.5.

The number of MPCs within a cluster $N_{p,c}$ determines how accurate the cluster dispersion parameters are reflected by the model. In theory, it strongly depends on the system assumptions. Using a large number of antennas, cluster dispersion parameters can be resolved very accurately. In contrast, when using a small number of antennas, the clusters' dispersion cannot be resolved very accurately and thus does not show so much impact on the modelled system. Also, the measurement bandwidth plays an important role. The larger the bandwidth, the better the delay resolution of the system. Hence, the number of MPCs within a cluster should be determined by the expected number of antennas and the bandwidth of the modelled system. Experiments showed that the number of paths strongly depends on the cluster spreads. Large clusters need more paths to be described accurately enough.

Another parameter characterising the environment is the number of clusters in the current snapshot. Again, every individual cluster c is characterised by the total number of clusters $N_{c,c}$ that coexist during the cluster's lifetime.

[4]The problem of selecting the powers of individual clusters within a snapshot is a very general one. A different approach for obtaining cluster powers is discussed in [45].

Parameter	Description	Unit	Parameter bounds
$N_{p,c}$	number of paths within the cluster	1	≥ 1
$N_{c,c}$	total number of clusters, existing with cluster c	1	≥ 0

Table 2.5: Number parameters

Cluster movement

The RCM is capable of modelling smoothly time-varying scenarios under the assumption that one station (typically the Rx) is fixed, while the other station (the Tx) moves[5]. This time variance is implemented by the movement of clusters.

The change rates specified in Table 2.6 are defined as the change of the respective parameter when the Tx travels one wavelength.

Parameter	Description	Unit
$\Delta\sigma^2_{\gamma,c}$	rate of change of the cluster power	dB per travelled wavelength
$\Delta\bar{\tau}_c$	rate of change of cluster mean delay	s per travelled wavelength
$\Delta\bar{\varphi}_{\mathrm{Rx},c}$	rate of change of cluster mean AOA	rad per travelled wavelength
$\Delta\bar{\varphi}_{\mathrm{Tx},c}$	rate of change of cluster mean AOD	rad per travelled wavelength
$\Delta\bar{\theta}_{\mathrm{Rx},c}$	rate of change of cluster mean EOA	rad per travelled wavelength
$\Delta\bar{\theta}_{\mathrm{Tx},c}$	rate of change of cluster mean EOD	rad per travelled wavelength
$\lvert\gamma_{\mathrm{att}}\rvert^2$	cluster attenuation factor for cluster fading in and out	dB

Table 2.6: Change rates of cluster parameters for time-variant modelling

The cluster attenuation factor is used for the cluster birth and death process to let a cluster smoothly enter and leave a scenario.

Note that the rate of change of the cluster mean delay intrinsically introduces cluster Doppler shift. Following assumptions are made on the cluster movement: (i) All paths within a cluster are assigned the same rate of change of their delay (i.e., they show the same Doppler shift), and (ii) the effects of rotating antenna arrays are neglected in the RCM.

With these assumptions, the cluster Doppler shift can then be calculated as

$$\nu_c = \frac{f_0}{c_0} v_c,$$

where v_c denotes the velocity component of the Tx towards cluster c. Since only the Tx is regarded as moving, the changes in the cluster mean delay $\Delta\bar{\tau}_c$ are due to changes of the distance between the Tx and cluster c (see Figure 2.3). Hence, the velocity of the Tx towards cluster c can be expressed as $v_c = -\Delta\bar{\tau}_c c_0 \cdot v_{\mathrm{Tx}}$. The speed of the Tx v_{Tx} will be defined in Section 2.3.1.4. Finally, the cluster Doppler shift is obtained as

$$\nu_c = -f_0 \cdot v_{\mathrm{Tx}} \cdot \Delta\bar{\tau}_c. \tag{2.3}$$

Note that this approach does not need to calculate a position in $x/y/z$-domain as suggested in [47], but still uses the twin-cluster concept.

[5] In the following, I will always use the Tx as the moving station.

Cluster lifetime

The cluster's lifetime (see Table 2.7) describes for how many wavelengths a cluster exists. This parameter is necessary for modelling the birth/death process of smoothly time-varying channels.

Parameter	Description	Unit	Parameter bounds
Λ_c	cluster lifetime	multiple of t_Λ	≥ 1

Table 2.7: Cluster lifetime

2.3.1.3 Environment parameters

The environment parameters specify where clusters occur, what their spreads and their velocities are, how many paths the clusters are composed of, and what the power of the individual clusters is. The RCM models these environment parameters by *the pdf of the cluster parameters*. Note that by this definition, the environment parameter $\boldsymbol{\Theta}_{\text{env}}$ is a *function*.

It is not surprising that the cluster parameters show significant correlations (see Section 4.2.1.2). For example, a cluster coming with short delay will contain more power than a cluster with very large delay. For this reason, the environment pdf is modelled using *one single multi-dimensional pdf*,

$$\boldsymbol{\Theta}_{\text{env}} = p_{\boldsymbol{\Theta}_c}(\boldsymbol{\theta}_c), \tag{2.4}$$

where $p_{\boldsymbol{\Theta}_c}$ denotes the multi-dimensional pdf of the cluster parameters, and $\boldsymbol{\theta}_c$ denotes the argument of this pdf.

At the first glance, this multi-dimensional pdf may seem clumsy to handle. However, it can be estimated easily from measurements using a kernel-density estimator, e.g. [51], as outlined in Section 4.2.1.3.

By drawing cluster parameters from this pdf, the individual scenarios are determined. This approach shows the high-grade impact that a scenario is *only* described by statistical parameters, in contrast to deterministically placing clusters in an $x/y/z$-plane. This is *the* tremendous advantage of this approach, as it elegantly solves the problem of the positioning of clusters in space. It is obvious that the accuracy of the model strongly relies on the accurate parametrisation of this pdf from comprehensive measurements (see Chapter 4).

In the following I will discuss the *marginal* distributions of the individual parameters separately. A discussion about the correlation between the different dimensions [52] is provided with the discussion about the model parametrisation in Section 4.2.1.2. The exact way of how to generate an environment from this pdf, I will detail in Section 2.3.2.

Cluster positions

Clusters positions are described by their probability in the angular/delay domain according to *cluster position pdfs* given in Table 2.8.

Parameter	Description
$p(\bar{\tau})$	pdf of the cluster mean delay
$p(\bar{\varphi}_{\mathrm{Tx}})$	position pdf of azimuth cluster position at Tx
$p(\bar{\varphi}_{\mathrm{Rx}})$	position pdf of azimuth cluster position at Rx
$p(\bar{\theta}_{\mathrm{Tx}})$	position pdf of elevation cluster position at Tx
$p(\bar{\theta}_{\mathrm{Rx}})$	position pdf of elevation cluster position at Rx

Table 2.8: Cluster position pdfs

The cluster position pdfs already specify the directivity of the channel. For instance, a uniform azimuth position pdf mean clusters are all around a station, while a peaky position pdf indicates predominant directions.

Cluster spreads

The cluster size in the angle/delay-space is described by the rms spreads of the paths within the cluster. Details about the distribution functions of the cluster spreads are described in Section 2.3.2.1. The associated external environment parameters are again the distribution functions of the cluster spreads described in Table 2.9.

Parameter	Description
$p(\sigma_{\tau})$	pdf of cluster delay spread
$p(\sigma_{\varphi_{\mathrm{Tx}}})$	pdf of cluster azimuth spreads seen from Tx
$p(\sigma_{\varphi_{\mathrm{Rx}}})$	pdf of cluster azimuth spreads seen from Rx
$p(\sigma_{\theta_{\mathrm{Tx}}})$	pdf of cluster elevation spreads seen from Tx
$p(\sigma_{\theta_{\mathrm{Rx}}})$	pdf of cluster elevation spreads seen from Rx

Table 2.9: Cluster spread pdfs

Cluster powers

The instantaneous power of the individual clusters (which is determined by the sum of the MPCs) has to be chosen carefully. The cluster power is influenced by its mean delay, shadow fading, and small-scale fading.

The multi-dimensional environment pdf manages to model all three effects in an elegant way, since the cluster power is strongly correlated with the other parameters. For example, clusters with small delay will show much larger power than clusters with long delay.

Also the total snapshot power ρ, in which the regarded cluster occurs is strongly correlated with all other cluster parameters (cf. the discussion about the cluster parameter ρ_c). Together with the number of clusters, this snapshot power is used to select the propagation environment in the model. The pdfs of the parameters are summarised in Table 2.10

Parameter	Description
$p(\sigma_\gamma^2)$	pdf of the cluster mean power
$p(\rho)$	pdf of the total snapshot power, in which a cluster occurs in

Table 2.10: Cluster power parameter pdfs

In its present form, the RCM does not take the property of polarisation into account. A possible extension is to model horizontal and vertical polarisation, including cross-polarisation effects of the channel. The received signals from different polarisations are then computed by the system model. The strict separation of parametric model and system model are of help, here.

Number of clusters and paths

The number of clusters N_c has an important impact in the model. For example, in scenarios with large-spread clusters, there will usually be only few clusters, while in scenarios with small-spread clusters there will usually be many clusters.

For this reason, the number of clusters it is the second parameter (besides snapshot power) that mainly determines which kind of scenario is modelled, as will become clear in Section 2.3.2.1. The number of clusters is strongly correlated with all other cluster parameters. In Section 2.3.2.1 I will detail how to determine the number of cluster to obtain a good match with a given scenario.

The number of paths also plays an important role. Large-spread clusters will usually consist of more paths than small-spread clusters. The number of paths can also determine the multipath richness in a channel. However, it turned out that the channel can also be well modelled with a low number of paths in a cluster, as long as the cluster statistics are still properly mapped. In the RCM, the number of paths within a cluster are identified from measurements, leading to consistent results.

The corresponding number parameter pdfs are collected in Table 2.11.

Parameter	Description
$p(N_c)$	pdf of the number of clusters
$p(N_p)$	pdf of the number of paths within a cluster

Table 2.11: Number parameter pdfs

Cluster movements

The cluster position change rates are also modelled by their pdfs summarised in Table 2.12. Note that these distributions will be strongly correlated with each other and with other parameters, such as the cluster delay, position, and the cluster spread.

Note that the clusters' Doppler shift is *not* modelled by a distribution. It is intrinsically provided by the cluster movement parameters as detailed in Section 2.3.1.2.

Parameter	Description
$p(\Delta\sigma^2_{\gamma,c})$	pdf of change rate of cluster power
$p(\Delta\bar{\tau}_c)$	pdf of change rate of cluster mean delay per travelled wavelength
$p(\Delta\bar{\varphi}_{\mathrm{Rx},c})$	pdf of change rate of cluster mean AOA per travelled wavelength
$p(\Delta\bar{\varphi}_{\mathrm{Tx},c})$	pdf of change rate of cluster mean AOD per travelled wavelength
$p(\Delta\bar{\theta}_{\mathrm{Rx},c})$	pdf of change rate of cluster mean EOA per travelled wavelength
$p(\Delta\bar{\theta}_{\mathrm{Tx},c})$	pdf of change rate of cluster mean EOD per travelled wavelength

Table 2.12: Cluster movement parameter pdfs

Cluster lifetime

The cluster lifetime is also parametrised by its pdf (see Table 2.13). Cluster birth and death are modelled by a birth-death process, which provides N_c clusters in the mean. I will further discuss the birth-death process of clusters and the cluster movement when detailing the smooth time-variance implementation of the RCM in Section 2.3.2.2.

Parameter	Description
$p(\Lambda)$	pdf of the cluster lifetime

Table 2.13: Cluster lifetime pdf

2.3.1.4 Time-variance parameters

Additional global parameters indicated in Table 2.14 are necessary to include time variance.

Parameter	Description	Unit
Δt_s	Sampling time interval	s
Δt_Λ	Birth/death interval	multiple of t_s
v_{Tx}	speed of Tx	$\lambda/\Delta t_s$
$p(\chi_{\mathrm{birth}})$	pdf of the cluster births per cluster lifetime Δt_Λ	

Table 2.14: Time-variance parameters

The RCM uses two time bases as demonstrated in Figure 2.4:

- Δt_s denotes the channel sampling interval. In these time steps, clusters move. The channel matrices are calculated on this basis.
- Δt_Λ denotes the cluster lifetime interval. Cluster lifetimes can only be a multiple of this value. The reason for this is that newly born clusters have to fade in and dying clusters need to fade out smoothly. Note that Δt_Λ is an integer multiple of Δt_s.

The cluster velocities depend on the speed of the Tx movement modelled by v_{Tx}. Note that the direction of movement is already characterised by the cluster movement parameters. Hence, the Tx velocity is only described by a scalar number rather than by a vectorial (directional) movement.

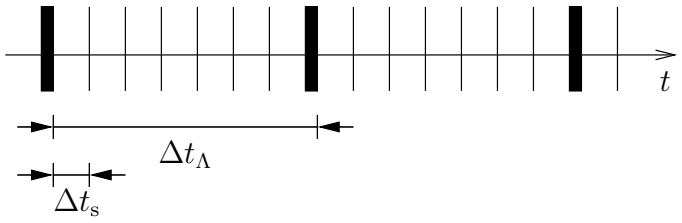


Figure 2.4: Time-variance periods: Changes of the number of clusters can occur in the lifetime periods Δt_Λ, while clusters move in the sampling time periods Δt_s. Note that Δt_Λ is taken as a multiple of Δt_s.

The lifetime of a cluster is already limited by the parameter Λ_c, but also new clusters have to be generated. This is implemented by the cluster birth rate χ_{birth} described by its pdf. Since a cluster birth does not occur too often, the normalisation of the birth rate to the cluster lifetime interval is numerically advantageous and results in a well-balanced birth-rate pdf.

2.3.2 Applying the parametric model

The parametric channel model creates cluster parameter sets $\boldsymbol{\Theta}_c$ and path parameter sets $\boldsymbol{\Theta}_{cp}$ from the environment pdf $\boldsymbol{\Theta}_{\text{env}}$. In other words, it determines the actually modelled wave-propagation environment from the environment parameter pdf.

First, I will describe how to model *random-access channels* (Section 2.3.2.1). Then, the problem of *smooth time variance* is treated in Section 2.3.2.2.

2.3.2.1 Random-access model

In a *random access scheme*, considered in this section, a device communicates using (relatively) short transmission blocks and larger pauses between these blocks. This leads to the assumption that the channel stays constant for the duration of the block transmission, but changes completely after the transmission. The following sections describe how to use the model to generate channels for *one data block*.

The general procedure is this: The model generates clusters for one scenario. Subsequently, it places paths within the clusters. The parameters of the clusters and the paths are assumed to be constant over the whole transmission block.

Obtaining the cluster parameter sets $\{\boldsymbol{\Theta}_c\}$

The multi-dimensional environment pdf $\boldsymbol{\Theta}_{\text{env}}$ in (2.4) provides a representation of *all possible kinds* of clusters in a scenario.

Figure 2.5 sketches how to obtain cluster parameters from the environment pdf.

The number of clusters and intended snapshot power determine the kind of environment of the regarded snapshot. These parameters need to be drawn only once per snapshot by following procedure:

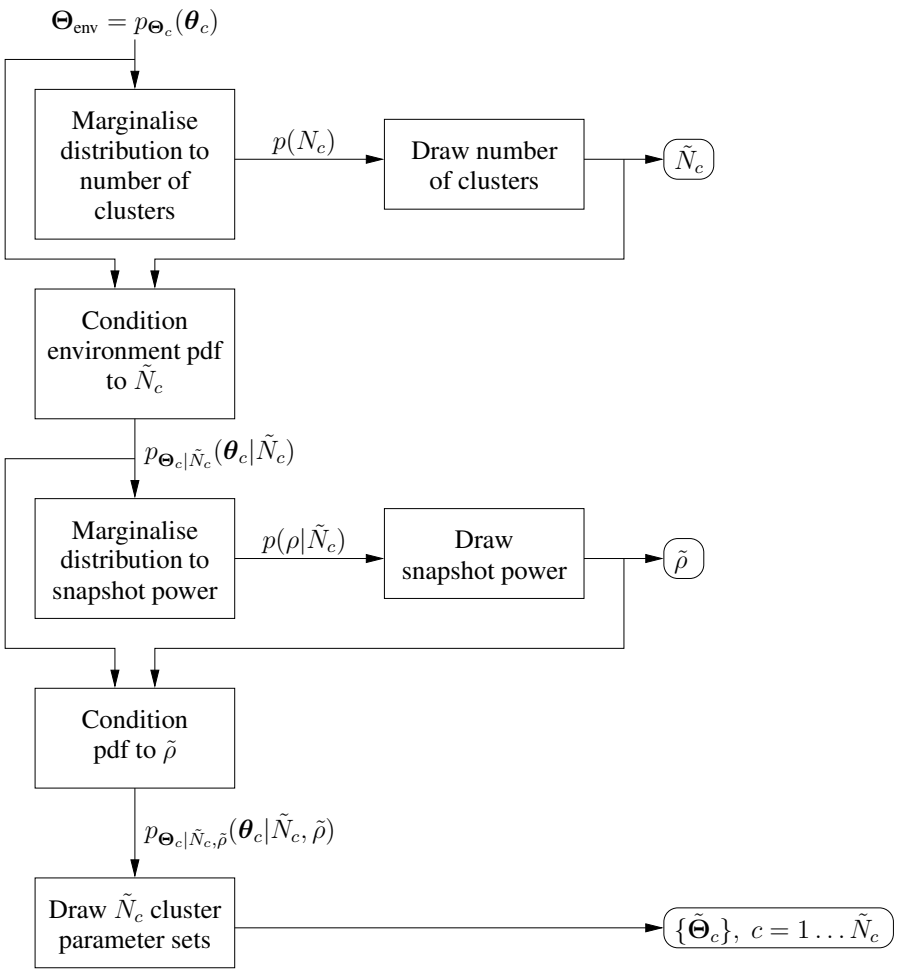


Figure 2.5: Flow diagram of obtaining cluster parameters from the environment pdf

First, a number-of-cluster pdf is obtained from the environment pdf by integrating over all other dimensions. Then a realisation[6] of the number of clusters $\tilde{N}_c$ is drawn from the marginal pdf $p(N_c)$. Since the sample is real-valued, the number of clusters is obtained by rounding to the closest integer. This number of clusters is then fixed for this snapshot.

Having determined $\tilde{N}_c$, the intended snapshot power has to be fixed next. For this, the environment pdf is marginalised to the intended snapshot power given the number of clusters $\tilde{N}_c$. From this resulting pdf $p(\rho|\tilde{N}_c)$, the current intended snapshot power $\tilde{\rho}$ is drawn. This intended power is then fixed for this snapshot.

After determining the intended snapshot power and the number of clusters, the cluster parameters are drawn in the following way:

The environment pdf is conditioned on $\tilde{\rho}$ and $\tilde{N}_c$, hence

$$\boldsymbol{\Theta}_{\text{env}}^{(\tilde{N}_c,\tilde{\rho})} = p_{\boldsymbol{\Theta}_c|N_c,\rho}(\boldsymbol{\theta}_c|N_c = \tilde{N}_c, \rho = \tilde{\rho}). \tag{2.5}$$

A number of $\tilde{N}_c$ *new clusters* is created by drawing cluster parameter sets $\{\tilde{\boldsymbol{\Theta}}_c\}$, $c = 1 \dots \tilde{N}_c$ from this reduced distribution.

Note that in this step the model determines *all parameters for all clusters* in the considered snapshot. The cluster powers are also drawn from the reduced pdf. For this reason, the total *intended* snapshot power $\tilde{\rho}$ drawn before is only used to select the kinds of clusters that occur with this power. The actual total power in the snapshot may, of course, differ from the one drawn before.

Since the parameters are modelled by a distribution, sometimes invalid cluster parameters may be drawn. In this case, this invalid cluster parameter is mapped to its boundary value that is provided with the cluster parameters in Section 2.3.1.2. Also, for all cluster parameters that are integer values, the drawn number is rounded to the closest integer.

Obtaining the path parameter sets $\{\boldsymbol{\Theta}_{cp}\}$

Using the parameter sets $\tilde{\boldsymbol{\Theta}}_c,\ c = 1 \dots \tilde{N}_c$, the model creates paths by drawing a number of path parameters for each cluster c. Note that the number of paths within each cluster $\tilde{N}_{p,c}$ was already determined in the previous step by drawing the cluster parameters.

The actual realisations are drawn from following distributions:

$$\begin{aligned}
\tau_{cp} &\sim \mathcal{N}(\bar{\tau}_c, \sigma^2_{\tau,c}),\\
\varphi_{\text{Rx},cp} &\sim \mathcal{N}^\circ(\bar{\varphi}_{\text{Rx},c}, \sigma^2_{\varphi_{\text{Rx},c}}),\\
\varphi_{\text{Tx},cp} &\sim \mathcal{N}^\circ(\bar{\varphi}_{\text{Tx},c}, \sigma^2_{\varphi_{\text{Tx},c}}),\\
\theta_{\text{Rx},cp} &\sim \mathcal{N}^\circ(\bar{\theta}_{\text{Rx},c}, \sigma^2_{\theta_{\text{Rx},c}}),\\
\theta_{\text{Tx},cp} &\sim \mathcal{N}^\circ(\bar{\theta}_{\text{Tx},c}, \sigma^2_{\theta_{\text{Tx},c}}),\\
\text{abs}(\gamma_{cp}) &= \sqrt{\sigma^2_{\gamma,c}/(|\gamma_{\text{att}}|^2 \cdot \tilde{N}_{p,c})},\\
\arg(\gamma_{cp}) &\sim \mathcal{U}(\ [-\pi \dots \pi)\ ),
\end{aligned}$$

[6] Realisations of a random variable are denoted by the tilde symbol.

where $\mathcal{N}$ and $\mathcal{U}$ denote the Gaussian and uniform distribution, respectively, and $\mathcal{N}^\circ$ denotes the Gaussian distribution, where the realisations are mapped to their principal value in $(-\pi, \pi]$ in the angular domain.

These is no trustworthy prior information about the distributions of the paths within a cluster. When using high-resolution estimation algorithms to identify paths from measurements, the underlying distribution of the path parameters within a cluster is not represented correctly [53]. So, the Gaussian distribution is most appropriate since it maximises the entropy [54, 55].

In the angular domain, also the Laplace distribution has been suggested [56, 46]. This distribution was identified in measurement data, however it is questionable whether this finding was an artefact of the identification algorithm [53]. Taking the approach to choose the entropy-maximising distribution, the Von Mises distribution [57] is the correct choice, which is well approximated by the principal-value mapped Gaussian distribution.

In the delay domain, paths within a cluster are often modelled by the exponential distribution. This assumption is usually reasoned with a large scattering object, where the first reflections with shortest delay must be the strongest ones. I did *not* observe this effect in path estimations from measurements. The paths within a cluster were rather Gaussian distributed in delay. For this reason, the RCM also uses the (entropy-maximising) Gaussian distribution for the path delays.

The distribution of the path weights is responsible for the cluster power and for the cluster fading behaviour. A simple, yet efficient approach is to assign the same power to all paths within a cluster with a random phase.

Having modelled the paths, the system model in Section 2.4 is used to calculate the time- and frequency-selective channel matrices for the regarded transmission block.

2.3.2.2 Extension to smoothly time-variant channels

When signal processing algorithms track the properties of the channel, the channel model must support smoothly time-variant channels.

The RCM supports two types of time-and-spatially smooth time-variance: The first type is an extension of the random-access model, introducing moving clusters, but the number of clusters stays fixed for the whole transmission block period. After the transmission block, the channel changes completely as in the random access model; a new realisation of the environment is drawn from the environment pdf.

More challenging is the second type, the complete spatially-smooth model, where an environment is drawn once at the beginning and then changes smoothly. Determined by the cluster lifetime, new clusters appear and older ones disappear. Since the first type is just a simplified version of the complete model, I will only describe the latter one.

Movement

The RCM uses moving clusters to model smooth time variance. The movement parameters of the clusters in the scenario are determined in the (previously drawn) cluster parameters $\tilde{\mathbf{\Theta}}_c$ while the cluster is born.

The RCM assumes all paths within a cluster to move in the same way, so only the path parameters need to be updated. In every sampling instant Δt_s, the path parameters are updated as:

$$\begin{aligned}
\tau_{cp}((k+1)\Delta t_s) &= \tau_{cp}(k\Delta t_s) + \Delta\bar{\tau}_c \cdot v_{\mathrm{Tx}}\Delta t_s,\\
\varphi_{\mathrm{Rx},cp}((k+1)\Delta t_s) &= \varphi_{\mathrm{Rx},cp}(k\Delta t_s) + \Delta\bar{\varphi}_{\mathrm{Rx},c} \cdot v_{\mathrm{Tx}}\Delta t_s,\\
\varphi_{\mathrm{Tx},cp}((k+1)\Delta t_s) &= \varphi_{\mathrm{Tx},cp}(k\Delta t_s) + \Delta\bar{\varphi}_{\mathrm{Tx},c} \cdot v_{\mathrm{Tx}}\Delta t_s,\\
\theta_{\mathrm{Rx},cp}((k+1)\Delta t_s) &= \theta_{\mathrm{Rx},cp}(k\Delta t_s) + \Delta\bar{\theta}_{\mathrm{Rx},c} \cdot v_{\mathrm{Tx}}\Delta t_s,\\
\theta_{\mathrm{Tx},cp}((k+1)\Delta t_s) &= \theta_{\mathrm{Tx},cp}(k\Delta t_s) + \Delta\bar{\theta}_{\mathrm{Tx},c} \cdot v_{\mathrm{Tx}}\Delta t_s,\\
\mathrm{abs}(\gamma_{cp}((k+1)\Delta t_s))^2\,[\mathrm{dB}] &= \mathrm{abs}(\gamma_{cp}(k\Delta t_s))\,[\mathrm{dB}] + \Delta\sigma^2_{\gamma,c} \cdot v_{\mathrm{Tx}}\Delta t_s,\\
\arg(\gamma_{cp}((k+1)\Delta t_s)) &= \arg(\gamma_{cp}(k\Delta t_s)).
\end{aligned}$$

Since all movement parameters are normalised to the travelled wavelength, the multiplication of $v_{\mathrm{Tx}}\Delta t_s$ is necessary.

Cluster birth and death

Cluster birth and death are governed by the parameters χ_{birth} and Λ_c. To allow for cluster fading in and fading out, the cluster number can only change in Δt_Λ time periods.

Cluster death Already when a cluster is generated, its cluster lifetime Λ_c is determined. The cluster fades away when its lifetime has ended. In the following, the number of died clusters within one cluster lifetime is denoted by χ_{death}.

Cluster birth Every Δt_Λ the number of new-born clusters $\tilde{\chi}_{\mathrm{birth}}$ is drawn from the cluster birth distribution $p(\chi_{\mathrm{beta}})$. For all new-born clusters, cluster parameters $\tilde{\boldsymbol{\Theta}}_c$ are drawn from the environment pdf $\boldsymbol{\Theta}_{\mathrm{env}}$ as described for the random-access channel model in Section 2.3.2.1. Note that the number of clusters on which the environment pdf has to be conditioned on is given as $N_c^{(\mathrm{new})} = N_c^{(\mathrm{old})} + \tilde{\chi}_{\mathrm{birth}} - \chi_{\mathrm{death}}$.

Fading in and out Several ways have been proposed to let a cluster fade in and out of a scenario [58, 46]. A common method is to use the arc-tangent function since it provides a smooth behaviour. In contrast, the RCM takes a different approach that has even lower complexity. For fading in the cluster attenuation factor $|\gamma_{\mathrm{att}}|^2$ is logarithmically decreasing from $40\,\mathrm{dB}$ to $0\,\mathrm{dB}$ (i.e. linearly increasing in dB-domain) every Δt_s over a period of one Δt_Λ. Similarly, for clusters fading out, their attenuation factor is increased from $0\,\mathrm{dB}$ to $40\,\mathrm{dB}$.

2.4 System model

The system model is split in two parts. First, the impulse response of the discrete paths $\mathbf{H}_\delta$ is calculated from the modelled clusters. Subsequently the diffuse multipath response $\mathbf{H}_d$ is generated. Finally, the complete channel $\mathbf{H}$ is obtained by adding discrete and diffuse components as in (2.1).

2.4.1 Filtering discrete paths

The radio system is described by the following three parameters collected in $\boldsymbol{\Theta}_{\mathrm{sys}}$:

(i) Antennas
The number of antennas, the array geometry, the polarisation and the radiation patterns are defined in the antenna array steering vectors $\mathbf{a}_{\mathrm{Rx}}(\varphi_{\mathrm{Rx}}, \theta_{\mathrm{Rx}})$ and $\mathbf{a}_{\mathrm{Tx}}(\varphi_{\mathrm{Tx}}, \theta_{\mathrm{Tx}})$.

(ii) Bandwidth, number of frequencies
The system bandwidth B defines the delay resolution of the channel impulse response, while the number of frequencies M determines the maximum resolvable delay by $\tau_{\max} = \frac{M}{B}$. Note that this equation can also be used in the other direction to specify the minimum number of frequencies necessary to capture the maximum delay by $M \geq B \cdot \tau_{\max}$.

(iii) System impulse response
The radio system itself may also have a specific transfer function at both Tx and Rx. The system model incorporates these into the RCM.

Using these system parameters, the time and frequency-dependent channel matrices of the discrete paths can be computed as indicated in [36], by

$$\mathbf{H}_{\delta}(t, \Delta f) = \sum_{c=1}^{N_c} \sum_{p=1}^{N_{p,c}} \Big(\gamma_{cp} \cdot \cdot \mathbf{G}_{\mathrm{Rx}}(\Delta f) \mathbf{a}_{\mathrm{Rx}}(\varphi_{\mathrm{Rx},cp}, \theta_{\mathrm{Rx},cp}) \cdot \mathbf{a}_{\mathrm{Tx}}^{\mathrm{T}}(\varphi_{\mathrm{Tx},cp}, \theta_{\mathrm{Tx},cp}) \mathbf{G}_{\mathrm{Tx}}^{\mathrm{T}}(\Delta f) \cdot \cdot \mathrm{e}^{-j2\pi\, \Delta f\, \tau_{cp}} \Big), \tag{2.6}$$

where $\mathbf{a}_{(\cdot)}$ denotes the (complex-valued) antenna array response, and $\mathbf{H}_{\delta}(t, \Delta f)$ denotes the time- and frequency-selective MIMO channel matrix calculated from the discrete paths with dimensions $N_{\mathrm{Rx}} \times N_{\mathrm{Tx}}$. The system transfer functions of the Tx (Rx) are described by the diagonal matrices $\mathbf{G}_{\mathrm{Rx/Tx}}(\Delta f)$, where the diagonal elements describe the transfer functions of the specific transmitter (receiver) chain. Should the system response be the same for all antenna elements, these matrices reduce to scalar values. The parameters of the system model are summarised in Table 2.15.

Note that Doppler shifts do not need to be included in this system model. The circular rotation of the phases of the paths, introduced by the Doppler shift, is intrinsically included by the rate of change of the path delays.

Parameter	**Description**
$\mathbf{a}_{\mathrm{Tx}}$	Antenna array description of the Tx
$\mathbf{a}_{\mathrm{Rx}}$	Antenna array description of the Rx
$\mathbf{G}_{\mathrm{Rx}}$, $\mathbf{G}_{\mathrm{Tx}}$	System impulse responses at Rx and Tx
B	System bandwidth
M	Number of frequencies

Table 2.15: System model parameters

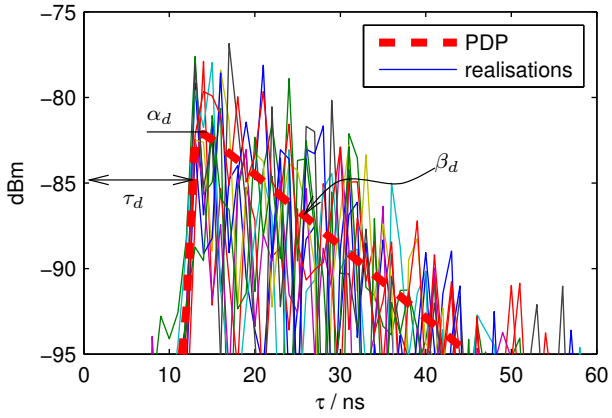


Figure 2.6: Example of DMP realisations in comparison to the PDP

2.4.2 Modelling diffuse multipath

As outlined in Section 1.4, the DMP is an inherent part of the channel. This effect has been distressingly neglected in all recent GSCM models. The RCM is the first model to include DMP[7], with the result that the channel diversity is matched much better than by other models [60].

The following subsection will show how to generate the DMP. Subsequently, the second subsection will detail how to model the necessary DMP parameters.

2.4.2.1 Generating DMP realisations

For better understanding I will first describe how to actually generate realisations for the DMP.

As already indicated in Section 1.4, the DMP is modelled by its PDP (cf. (1.5)), which is specified by the three parameters describe in Table 2.16, and illustrated in Figure 2.6.

Parameter	Description
α_d	Peak of DMP PDP
β_d	Exponential decay constant of DMP PDP
τ_d	Base delay of the DMP PDP

Table 2.16: Diffuse multipath parameters

The actual realisations are obtained using a noise-colouring filter in the frequency domain. For this, I consider the Fourier transform of (1.5) ($\tau \Leftrightarrow \Delta f$)

$$\Psi_d(\Delta f) = \frac{\alpha_d}{B_d + \mathrm{j}2\pi\Delta\mathrm{f}} \mathrm{e}^{-\mathrm{j}2\pi\Delta f\tau_d}. \tag{2.7}$$

[7] The IlmProp channel modelling tool by Del Galdo [59] outlines a promising preliminary method to model the DMP parameters. However, this method has not been tested against measurements yet, as Del Galdo writes, “to carry out a rigorous validation and to develop an algorithm to determine the parameters [...] goes beyond the scope of this work, and is therefore left to future work”.

To implement a generator for the DMP, I introduce the normalised parameters $\beta_d = B_d/B$, $\tau'_d = \tau_d B/M$, and a sampled version of (2.7)

$$\boldsymbol{\kappa} = \frac{\alpha_d}{M}\left[\frac{1}{\beta_d} \quad \frac{\mathrm{e}^{-j2\pi\tau'_d}}{\beta_d + j2\pi\frac{1}{M}} \quad \cdots \quad \frac{\mathrm{e}^{-j2\pi(M-1)\tau'_d}}{\beta_d + j2\pi\frac{M-1}{M}}\right], \tag{2.8}$$

having M samples, and bandwidth B. Then, a Toeplitz matrix is created from (2.8) and its Cholesky decomposition computed as

$$\mathbf{R}_d = \mathrm{toep}\left(\boldsymbol{\kappa}, \boldsymbol{\kappa}^{\mathrm{H}}\right) = \mathbf{L}_{\mathrm{d}}\mathbf{L}_{\mathrm{d}}^{\mathrm{H}}. \tag{2.9}$$

Finally, the *realisations* of the DMP are obtained according to

$$\mathbf{h}_d = \mathbf{L}_d\mathbf{w}, \tag{2.10}$$

where $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}) \in \mathbb{C}^{M\times 1}$ is a realisation of a circular, complex, i.i.d., zero-mean Gaussian distributed process. If the number of frequencies M is large, it is reasonable to exploit the Toeplitz structure of (2.9) as detailed in [36, pp. 152].

Note that $\mathbf{h}_d$ contains *one* realisation of the DMP, stacked in the frequency domain, i.e. $\mathbf{h}_d \in \mathbb{C}^{M\times 1}$. So, the generation step (2.10) has to be done for all $N_{\mathrm{Tx}} \times N_{\mathrm{Rx}}$ links separately and independently. Figure 2.6 presents an example of 10 realisations that were generated according to the underlying PDP.

This implies that the RCM currently considers the DMP to be spatially and temporally white. Since the RCM is currently focussing on indoor modelling, this stipulation can be expected to be met, since diffuse power will most likely be scattered from all directions around the Tx and Rx. For outdoor cases, this stipulation has yet to be proven, or the DMP model needs to be improved.

2.4.2.2 Model of diffuse-multipath parameters

The previous section showed how to actually obtain realisations of the DMP. The open problem discussed in this section is how to obtain the DMP parameters needed for the DMP generation.

In Section 4.3.2, I will substantiate that following quantities are strongly correlated:

(i) the base delay of the peak of the discrete PDP, $\check{\tau}_\delta = \arg\max_\tau\{\psi_\delta(\tau)\}$ is correlated with the base delay of the DMP

(ii) the discrete PDP peak power $\check{\psi}_\delta = \max\{\psi_\delta(\tau)\}$ is correlated with the DMP peak power,

(iii) the total rms delay spread σ_τ [49] of the impulse response is correlated with the ratio between the total diffuse power P_d and the total Rx power P_{Rx}.

The ratio between the total diffuse power P_d and the total Rx power P_{Rx} acts as an auxiliary quantity $r_d = P_d/P_{\mathrm{Rx}}$, where P_d is given by

$$P_d = \int_0^\infty \psi_d(\tau)\mathrm{d}\tau = \frac{\alpha_d}{\beta_d}, \tag{2.11}$$

and similar definitions hold for the total Rx power and the total power of the discrete PDP, P_δ. The power of the diffuse components is then given as

$$P_d = \frac{r_d}{1 - r_d} P_\delta. \tag{2.12}$$

Using the correlations indicated above, the DMP parameters are modelled by their probability distribution conditioned on the parameters they depend on, resulting in the probability distributions given in Table 2.17.

Parameter	**Description**
$p(\alpha_d[\mathrm{dBm}] \vert \check{\psi}_p)$	PDF of the peak of the DMP PDP
$p(\tau_d \vert \check{\tau}_p)$	PDF of the DMP base delay
$p(r_d \vert \sigma_\tau[\mu s])$	PDF of the DMP power ratio
$p(\sigma_\tau)$	pdf of the total delay spread of the channel

Table 2.17: Diffuse multipath parameter pdfs

Realisations of the diffuse-multipath parameters can be obtained in the following way:

- By drawing realisations from the distributions $p(\tau_d \vert \check{\tau}_p)$ and $p(\alpha_d[\mathrm{dBm}] \vert \check{\psi}_p)$, the base delay τ_d, and the peak power α_d are determined directly.
- The decay factor β_d is then determined by drawing a realisation of the diffuse power ratio from $p(r_d \vert \sigma_\tau[\mu s])$, obtaining the total diffuse power from (2.12), and calculating the decay factor by (2.11).

Section 4.3.2 will present how these DMP pdfs can be determined.

Note that the delay spread of the channel is a key parameter here. The delay spread cannot be modelled well using clusters, but is modelled easily by the DMP.

2.5 Model summary

2.5.1 Model flow diagram

Figure 2.7 provides a detailed summary of the model's flow.

2.5.2 Complete list of RCM parameters

Table 2.18 summarises the external parameters of the RCM.

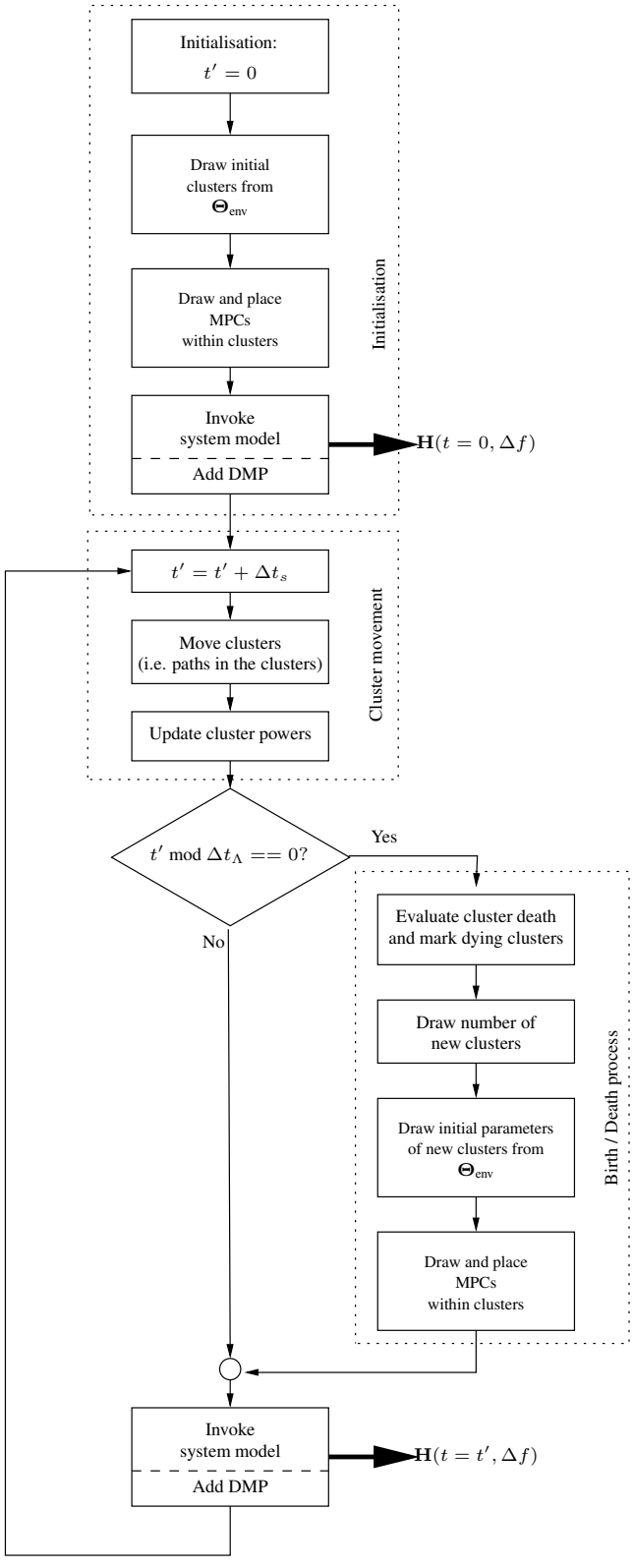


Figure 2.7: Flow chart of the smoothly time-variant Random-Cluster Model

Parameter	Description
Marginal distributions of the environment pdf Θ_{env}	
$p(\bar{\tau})$	pdf of the cluster mean delay
$p(\bar{\varphi}_{\text{Tx}})$	position pdf of azimuth cluster position at Tx
$p(\bar{\varphi}_{\text{Rx}})$	position pdf of azimuth cluster position at Rx
$p(\bar{\theta}_{\text{Tx}})$	position pdf of elevation cluster position at Tx
$p(\bar{\theta}_{\text{Rx}})$	position pdf of elevation cluster position at Rx
$p(\sigma_\tau)$	pdf of cluster delay spread
$p(\sigma_{\varphi_{\text{Tx}}})$	pdf of cluster azimuth spreads seen from Tx
$p(\sigma_{\varphi_{\text{Rx}}})$	pdf of cluster azimuth spreads seen from Rx
$p(\sigma_{\theta_{\text{Tx}}})$	pdf of cluster elevation spreads seen from Tx
$p(\sigma_{\theta_{\text{Rx}}})$	pdf of cluster elevation spreads seen from Rx
$p(\sigma_\gamma^2)$	pdf of the cluster mean power
$p(\rho)$	pdf of the total snapshot power, in which a cluster occurs in
$p(N_c)$	pdf of the number of clusters
$p(N_p)$	pdf of the number of paths within a cluster
$p(\Delta\sigma_{\gamma,c}^2)$	pdf of change rate of cluster power
$p(\Delta\bar{\tau}_c)$	pdf of change rate of cluster mean delay per travelled wavelength
$p(\Delta\bar{\varphi}_{\text{Rx},c})$	pdf of change rate of cluster mean AOA per travelled wavelength
$p(\Delta\bar{\varphi}_{\text{Tx},c})$	pdf of change rate of cluster mean AOD per travelled wavelength
$p(\Delta\bar{\theta}_{\text{Rx},c})$	pdf of change rate of cluster mean EOA per travelled wavelength
$p(\Delta\bar{\theta}_{\text{Tx},c})$	pdf of change rate of cluster mean EOD per travelled wavelength
$p(\Lambda)$	pdf of the cluster lifetime
Other environment parameters	
$p(\chi_{\text{birth}})$	pdf of the cluster births per cluster lifetime Δt_Λ
Δt_s	Sampling time interval
Δt_Λ	Birth/death interval
v_{Tx}	speed of Tx in $\lambda/\Delta t_s$
System parameters	
$\mathbf{a}_{\text{Tx}}$	Antenna array description of the Tx
$\mathbf{a}_{\text{Rx}}$	Antenna array description of the Rx
$\mathbf{G}_{\text{Rx}}$, $\mathbf{G}_{\text{Tx}}$	System impulse responses at Rx and Tx
B	System bandwidth
M	Number of frequencies
DMP parameters	
$p(\alpha_d[\text{dBm}]\vert\check{\psi}_p)$	PDF of the peak of the DMP PDP
$p(\tau_d\vert\check{\tau}_p)$	PDF of the DMP base delay
$p(r_d\vert\sigma_\tau[\mu s])$	PDF of the DMP power ratio
$p(\sigma_\tau)$	pdf of the total delay spread of the channel

Table 2.18: List of external parameters of the RCM

3 Principles of model validation

Any channel model has to be tested whether it is able to model scenarios correctly. Especially for MIMO systems, the model has to represent the physical structure of the channel, since it significantly determines the performance limits. The ultimate test for any model is the comparison to measurements.

According to [46, Chap. 6.7], the result of careful model validation should give answer to the following two questions: (i) Where are the limitations of the model used? (ii) Are some models better suited to predict a certain aspect of MIMO system performance than others?

In this thesis I will validate the RCM using a number of performance metrics. In this Section, I will discuss the suitability of some well-known validation metrics such as the mutual information (note that the term "channel capacity" is heavily misused in this respect). Finally, I will introduce a new metric, the environment characterization metric (ECM).

The following sections shall provide the link between the description of the RCM and its different ways of parametrisation. The parametrisation and the actual validation will be described in Chapter 4, where I will evaluate the model fit to measurements using the various validation metrics to show the impacts of the different parametrisation methods.

3.1 Validation procedure

The RCM is very specific to measurements. For this reason, *modelled channels* from the RCM can be compared with *measured channels*, which makes the validation straight-forward[1].

I use the following procedure to validate the RCM (see Figure 3.1):

1. Do channel measurements in representative scenarios
2. Estimate propagation paths [19] and DMP parameters (Section 4.3.1) from the measurements for every snapshot of the channel.
3. Identify and track MPC clusters (Chapter 7)
4. Parametrise the RCM (Chapter 4)
5. Generate *reference channels* using estimated paths and DMP parameters
6. Generate *modelled channels* by invoking the RCM
7. Compare the modelled channels with the reference channel according to the validation metrics.

[1] Note that validation is much more difficult for general channel models that target on reference environments. In this case, the models are not directly parametrised from measurements and can thus not directly be compared with them.

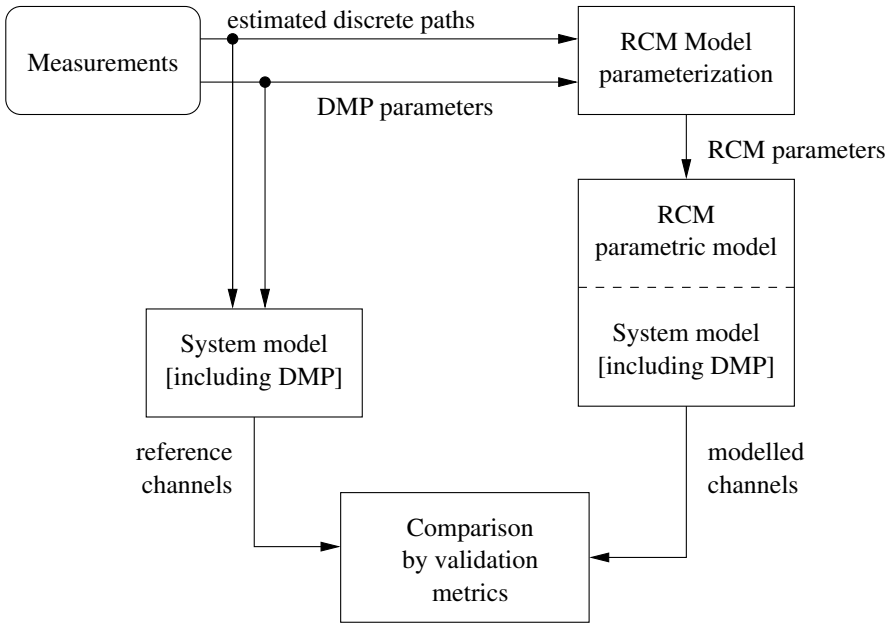


Figure 3.1: Validation framework. Whether DMP is included depends on the focus of the validation and is individually set in the respective sections.

Do channel measurements

Obviously, the comparison to measurements is only possible when having meaningful measurements [13]. For this reason, I conducted a measurement campaign comprehensively described in Part 3. The next section on the RCM parametrisation will detail which measurement routes were selected to be representative[2].

Estimate propagation paths

In the channel measurements, blocks of 4 complete channel snapshots closely spaced in time were recorded. To do the estimation, I combined always two of these bursts for three reasons: (i) to increase the SNR for estimation by averaging over two blocks, (ii) to allow for a basic estimation of the path-wise Doppler shift, (iii) to mitigate phase-noise effects of the sounding hardware [61, 62].

To obtain accurate estimation results, a number of prerequisites are vital:

- The sounding equipment has to be end-to-end calibrated, and these end-to-end calibration data must be used in the estimation algorithm.
- The antenna arrays, including the corresponding switches must be calibrated in an anechoic chamber in the full spherical domain, and in both polarisations. Note that the antenna response to both polarisations is even necessary for single-polarised antennas! This antenna calibration must be used in the estimation algorithm.

[2] In the last years, quite a lot of discussion was devoted to the term "typical" measurements. The COST 273 community agreed at some point that the typical environment is the environment of your home university. I do not want to claim that deviating to the term "representative" helps — admittedly, the measurements were conducted at the University of Oulu, the home university of the channel sounder available. However, note that the scenarios were chosen consciously to represent a variety of different types of environments.

- The measurements must have a certain minimum SNR that allows the estimation of discrete propagation paths, and must not be corrupted by interference[3].

Note that for the estimation of propagation paths, either very high computational power, or much time is necessary. To evaluate the measurement data used in this thesis, two state-of-the-art personal computers were running on full load for 12 months.

The estimation of propagation paths has to be performed very conscientiously. The parameters used for the estimation algorithm along with a discussion about the pitfalls of high-resolution parameter estimation is provided in Section 6.3.2.

Identify and track clusters

The whole Chapter 7 is devoted to the cluster identification and tracking algorithm. Speaking of validation, I have to point out already here that the clustering-and-tracking algorithm may also have a significant impact on the model accuracy, which will be discussed in Section 4.2.2.

Parametrise the RCM

The way of parametrising the RCM governs the model accuracy as will become clear in the next chapter.

Generate reference channels and modelled channels

When it comes to comparison of the RCM with measurements there are two options. The validation metrics to be introduced will process either the channel matrices or the path parameters, directly.

- For comparing the discrete paths only, the estimated paths from measurements and the modelled paths from the RCM are used.
- When comparing channel matrices, a fair comparison becomes difficult. I distinguish between following cases:
  1. The RCM is used for modelling discrete paths only, i.e. *without* diffuse multipath:
     In this case, a fair comparison is only possible when considering just the discrete (estimated) paths from the measurements. So, both the estimated paths from measurements and the modelled paths from the RCM are *processed by the same system model* to obtain comparable channel matrices.
  2. The system model is different from the measurement system:
     A fair comparison can be again achieved by using the *same system model* for both the estimated paths and the modelled paths. Additionally, the DMP (estimated and modelled) must be considered.

     Note that this case is the usual one, since there is no need to model the kind of antenna structure and system bandwidth the measurements were done with. However, this approach implies that the estimators for the paths and the DMP must provide accurate results.
  3. The complete system model of the sounding data is used:
     In this case, the modelled data can be compared directly to the measured channel matrices. A fairness problem arising here is the handling of polarisation. In our case the

[3] Care must be taken when planning measurements at 2.45 GHz in a WLAN-serviced area!

polarisation directions of the antennas we measured with were $\pm 45°$. This means that every element observed contributions from *both* polarisations in the channel. Since the RCM does not support the modelling of dual-polarised channels, such a comparison is quite unfair, there would be an intrinsic error that can not be quantified.

One must also consider that this approach is usually the most complex one, since the number of antennas and frequencies used in the measurements are much larger than the usually modelled ones. Also when choosing only a subset of the used antennas, the data handling is computationally demanding.

Note that for some channel sounders, the true channel matrix (impulse response) is not available, but a filtered version. In this case this filter must also be applied to the modelled channels. These can then be compared to the measured ones using the validation metrics.

In this thesis I will discuss the first and the second case. For reasons of a fair comparison and complexity I decided against the third approach.

Compare modelled channels with reference channels

Each of the validation metrics to be introduced in the next section describe the instantaneous MIMO channel by a single figure of merit. I compare these numbers using the empirical *cumulative distribution function* (cdf).

3.2 Validation metrics

I investigated the following validation metrics whether they are useful for comparing the RCM with measurements:

- Mutual information
- Diversity metric
- Demmel condition number
- Environment characterization metric

The next sections will discuss these metrics in detail. Note that there are also other methods that base on the *correlation matrix* of the channel coefficients, like the Correlation Matrix Distance [63]. However, to obtain a trustworthy estimate of a correlation matrix, the underlying channel must be (quasi-)stationary. In the random-access mode, the RCM does not provide stationary snapshots of the channel, so correlation-based metrics will not reflect the properties of the channel correctly. This will become obvious from the results of the diversity metric.

The question of the specificity of the other metrics will also be discussed together with the validation results in the next chapter.

3.2.1 Mutual information

A large number of publications use “Capacity” for model validation, while they usually mean mutual information (MI) [2]. But even MI has several disadvantages that disqualifies it as a good metric for model validation. To substantiate this opinion, I will compare the validation performance of MI to other validation metrics in this thesis.

3.2.1.1 Definition

The *narrowband MI* at frequency Δf and time t is defined as

$$I(t, \Delta f) = \log_2 \det \left[\mathbf{I} + \frac{\mathrm{SNR}}{N_t} \mathbf{H}_n(t, \Delta f) \mathbf{H}_n^{\mathrm{H}}(t, \Delta f) \right]. \tag{3.1}$$

When using the narrowband MI as metric in a cdf, all time realisations and frequencies are used as samples.

The *wideband MI* is obtained by averaging over the frequency domain as

$$I(t) = \frac{1}{B} \sum_{\Delta f} I(t, \Delta f) \tag{3.2}$$

where the sum is computed over the frequency bins. All time instants are taken as samples for the cdf.

The channel matrices $\mathbf{H}_n(t, \Delta f)$ denote *normalised* channel matrices, such that the SNR is reflected correctly, hence $\mathbf{H}_n = \mathrm{const} \cdot \mathbf{H}$. Note that two kinds of normalisations lead to relevant results:

1. Normalisation to *constant transmit power* over the whole simulation period. In this case,

$$\mathbf{H}_n(t, \Delta f) = \frac{1}{\frac{1}{MT} \sum_{t=1}^{T} \sum_{\Delta f} \|\mathbf{H}(t, \Delta f)\|_{\mathrm{F}}^2} \mathbf{H}(t, \Delta f). \tag{3.3}$$

   Using this normalisation, the differences in power in the channel are reflected by the validation metric.

2. Normalisation to *constant instantaneous SNR*. This corresponds to perfect power level control at the Tx. In this case, every time instant is normalized separately as

$$\mathbf{H}_n(t, \Delta f) = \frac{1}{\frac{1}{M} \sum_{\Delta f} \|\mathbf{H}(t, \Delta f)\|_{\mathrm{F}}^2} \mathbf{H}(t, \Delta f). \tag{3.4}$$

   Here, the validation metric reflects the spatial structure of the channel.

I chose an SNR of 10 dB for the following validation evaluations.

3.2.1.2 Deficiencies of MI as validation metric

I will demonstrate the deficiencies of using MI as validation metric by a meaningful example. This example will also show the difference between *average MI* and *ergodic capacity*.

Let us consider a deterministic, discrete scenario described by a number of propagation paths with their parameters power, AoA, AoD, and delay. We now create channel realisations of this scenario using the system model. Then, we create further channel realisations by just changing the phases of the paths randomly, but do not alter any of the other parameters. Note that the spatial structure of this channel does not change at all. Finally, we calculate the MI for all these realisations. For comparison, we repeat this whole experiment using a completely different scenario.

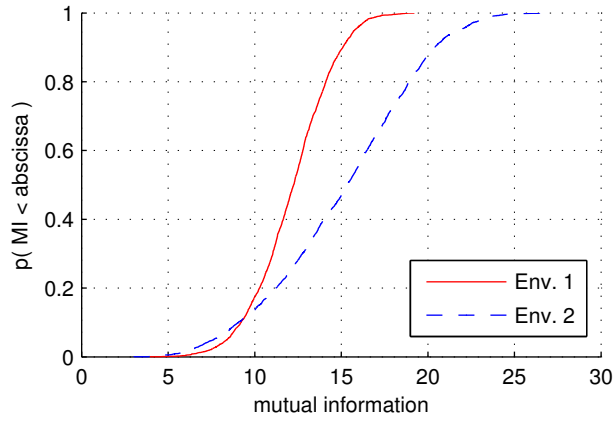


Figure 3.2: Why MI is no good validation metric: The MI cdf was computed from the same spatial environment, i.e. the same paths, with random phases. Two different environments are provided for comparison.

The measurements used for this experiment were TxR8 and TxR22, both at 2.55 GHz, where the 50th measured snapshot was considered. Details about the measurement routes are provided on pages 162 and 176, respectively.

Figure 3.2 shows the cdf of the MI (in short referred to as "MI cdf") evaluated from the two exemplary scenarios. For both scenarios, the *MI varies strongly*, even though the *spatial structure of the channel is the same*. This leads to following conclusions:

Mutual information fails to reflect the spatial structure of an environment.

The *slope* of the MI cdf represents the spatial diversity available in the channel. A steep slope corresponds to high diversity. The mean of the MI cdf can be interpreted as *ergodic capacity*. Note that this capacity was achieved by creating random samples of channel matrices by treating the phases of the MPCs as random[4]. In this case, the term *ergodic capacity* is justified. However, the *average mutual information* is by no means comparable to it, since the measurements were done in an non-stationary (and hence, non-ergodic) environment.

Is there still a way of using mutual information as performance metric? — From the previous findings, I would suggest that the *cdf of ergodic capacity* (instead of the cdf of instantaneous mutual information) would be one possible validation metric for judging on the multipath structure of the channel. This would imply that one needs to calculate a significant number of channel matrices for every single time instant simulated for this purpose. Unfortunately, this way is computationally much too demanding. Moreover, the environment characterisation metric to be introduced in Section 3.2.4 provides a more straight-forward way to judge on the multipath structure of the channel.

Nevertheless, MI is generally used for validating channel models, so, for reasons of comparison, I will use also instantaneous MI in this thesis.

[4]Note that this method was introduced by [64] in a very similar context.

3.2.2 Diversity metric

Spatial diversity describes the number of independent fading links between the Tx and Rx antenna arrays. In a full-diversity system, where all links between the Tx and Rx arrays are independent, one observes a spatial diversity of $N_{\mathrm{Tx}}N_{\mathrm{Rx}}$ [65]. This diversity is directly linked with the uncoded bit-error ratio (BER) performance of MIMO systems.

Correlation between the channels reduces the diversity significantly. Ivrlac and Nossek provided the *diversity measure*, a way to quantify the available diversity directly from the MIMO channels without taking the detour via BER simulations [66]. I will use this measure to quantify the diversity in both the measured and the modelled channels to subsequently compare the results.

The diversity measure $D(\mathbf{R})$ of a MIMO system described by the channel matrix $\mathbf{H}$ with correlation matrix $\mathbf{R} = \mathrm{E}\{\mathrm{vec}(\mathbf{H})\mathrm{vec}(\mathbf{H})^{\mathrm{H}}\}$ is given by

$$D(\mathbf{R}) = \left(\frac{\mathrm{tr}(\mathbf{R})}{\|\mathbf{R}\|_{\mathrm{F}}}\right)^2, \tag{3.5}$$

where $\mathrm{tr}(\cdot)$ denotes the matrix trace, and $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius matrix norm.

The usage of the channel *correlation* matrix implicitly assumes the channel to be stationary, at least over a certain period of time. This assumption is, of course, problematic for judging on the fit of the random-access mode of the RCM. However it turns out to be a reasonable metric for assessing the smoothly time-variant mode of the RCM. For the latter case, a sliding time window is used to approximate the channel correlation matrix. Also, all modelled frequencies are used as realisations of the channel matrix $\mathbf{H}$.

3.2.3 Demmel condition number

The Demmel condition number [67] of the instantaneous channel matrix was proposed as a parameter to decide whether to use this channel for spatial multiplexing rather than for diversity transmission [68]. Kyösti et. al. were the first to use this measure as validation metric in [69], when judging on the fit of the WINNER channel model.

The Demmel condition number is defined as

$$\kappa_D = \frac{\|\mathbf{H}\|_{\mathrm{F}}}{\lambda_{\min}(\mathbf{H})}, \tag{3.6}$$

where $\lambda_{\min}(\cdot)$ denotes the smallest singular value of the matrix in the argument. Using the Demmel condition number, measured and simulated channels are compared by their invertibility. Again, all modelled frequencies and time instants are used as realisations of $\mathbf{H}$.

This metric is *inapplicable for measured channel matrices*, since their smallest singular value is governed by the noise floor[5]. For the validation approach in this thesis, there is no noise floor, so the metric is applicable. This, however, does not tell whether the metric is specific or not.

[5]The measurement noise of course influences the accuracy of the path estimator, which is not validated in this thesis.

3.2.4 Environment Characterisation Metric

The novel *environment characterisation metric* (ECM) [70] allows to compare two scenarios in terms of their discrete propagation paths. Hence, the great advantage of the ECM is its *independence of the underlying system model*. It evolved from a metric that was quantifying the dispersion in the directions on one link end, only [71].

Starting point are both the propagation paths estimated from measurements $\{\hat{\boldsymbol{\Theta}}_{cp}\}(t)$ and the propagation paths from the modelled environments $\{\tilde{\boldsymbol{\Theta}}_{cp}\}(t)$, where t represents the corresponding time index of the measurement or the model. The following equations are applied for both measured (estimated), and modelled paths for all time indices t. For better readability, I will (i) enumerate all paths in each time instant from $l(t) = 1 \dots L(t)$, disregarding cluster structures for the time being, and (ii) skip the time index t in the following derivations.

As the metric has to cope with path parameters in different units (angular and delay), it is essential to transform the path parameters by proper scaling. For every path l, the angular data is transformed into coordinates on the unit sphere for both Rx and Tx. For angles of arrival the transformation is given as

$$\begin{bmatrix} x_{\mathrm{Rx},l} \\ y_{\mathrm{Rx},l} \\ z_{\mathrm{Rx},l} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sin(\varphi_{\mathrm{Rx},l}) \cdot \sin(\theta_{\mathrm{Rx},l}) \\ \sin(\varphi_{\mathrm{Rx},l}) \cdot \cos(\theta_{\mathrm{Rx},l}) \\ \cos(\theta_{\mathrm{Rx},l}) \end{bmatrix}, \tag{3.7}$$

for angles at the Tx it reads similarly. The scaling is done such that the maximum Euclidean distance between two paths is limited to 1.

Delay is scaled by the maximum delay that occurs in the considered snapshot [72], hence

$$\tilde{\tau}_l = \frac{\tau_l}{\max_l \tau_l} \tag{3.8}$$

Every path is now described by seven (dimensionless) parameters collected in

$$\boldsymbol{\pi}_l = [x_{\mathrm{Rx},l}\ y_{\mathrm{Rx},l}\ z_{\mathrm{Rx},l}\ x_{\mathrm{Tx},l}\ y_{\mathrm{Tx},l}\ z_{\mathrm{Tx},l}\ \tilde{\tau}_l]^{\mathrm{T}} \tag{3.9}$$

and its power $|\gamma_l|^2$. When considering propagation in the azimuthal plane only, the z-direction must be excluded.

The mean parameter vector is then given as

$$\bar{\boldsymbol{\pi}} = \frac{\sum_{l=1}^{L} |\gamma_l|^2 \boldsymbol{\pi}_l}{\sum_{l=1}^{L} |\gamma_l|^2}. \tag{3.10}$$

We define the novel *environment characterization metric* (ECM) as the empirical covariance matrix of the path parameter vector $\boldsymbol{\pi}$, so that

$$\mathbf{C}_{\boldsymbol{\pi}} = \frac{\sum_{l=1}^{L} |\gamma_l|^2 (\boldsymbol{\pi}_l - \bar{\boldsymbol{\pi}})(\boldsymbol{\pi}_l - \bar{\boldsymbol{\pi}})^T}{\sum_{l=1}^{L} |\gamma_l|^2}. \tag{3.11}$$

This metric shows the following properties:

- The metric is *system independent* as it is calculated from the propagation paths directly.
- The main diagonal contains the directional spreads [49] of the single components ($x/y/z$) at Rx and Tx and the (normalized) rms delay spread.
- The trace tr$\{\mathbf{C}_{\boldsymbol{\pi}}\}$ is the sum of the directional spreads at Rx and Tx plus the (normalized) delay spread. Note that the trace is dominated by the *large* SVs.
- The determinant $\det\{\mathbf{C}_{\boldsymbol{\pi}}\}$ has similar importance as detailed in [73, 71]. It describes the volume spanned in the parameter space. Since the value is dominated by the *small* SVs, it provides information about the most compact domain.

The singular values (SV) of $\mathbf{C}_{\boldsymbol{\pi}}$ can be interpreted as *fingerprint* of the scenario, by which one can judge the compactness of the paths in the channel. This genuine property makes the SVs of the ECM (SV-ECM) well suited for comparing channels.

In the following I use the cdfs of the SV-ECM in dB for comparing different scenarios. The statistics for the cdf are collected over all time indices t. Note that this approach results in 7 (or, when considering azimuth propagation only, 5) cdfs. For just three dimensions, these singular values could be seen as the lengths of the main axes of an ellipsoid (which becomes quite difficult in the 5-D or 7-D case).

The strongest SV-ECM describes the largest-spread dimension, which in some sense describe how dispersed the propagation environment is. In contrast, the weakest SV-ECM describes the smallest dispersion observed in the channel, which can be interpreted as a kind of directivity. In this sense, these two values are the most interesting ones for validation purposes.

4 Random-Cluster Model parametrisation and validation

Because of the statistical nature of the RCM, the model parametrisation is vital for the model to provide a good match to real-world propagation scenarios.

The two crucial parameters are (i) the multi-variate environment pdf $\boldsymbol{\Theta}_{\text{env}}$, (ii) the diffuse-multipath parameters $\boldsymbol{\Theta}_{\text{DMP}}$. The parameters of the system model $\boldsymbol{\Theta}_{\text{sys}}$ are required to fit the system to be simulated with the preferred accuracy, but these parameters are are (more or less) deterministic.

The next sections will make use of the validation procedure and metrics introduced in the previous chapter. Section 4.1 will detail how to set the system parameters, Section 4.2 will present three approaches to parametrise the environment pdf. In this context I will also discuss the specificity of the RCM to the environment pdf. I also discuss the impact of the clustering algorithm to the performance of the RCM. Finally, Section 4.3 will show a novel approach to parametrise diffuse multipath.

4.1 System parametrisation

The system parametrisation describes the radio system in terms of centre frequency, transmission bandwidth, the antenna geometry, and the RF responses of the transceivers. These parameters are up to the model user's own discretion and system demands. First, the following section provides a short overview of the properties of these parameters, which are important for the correct use of any propagation-based MIMO channel model. Then, I will describe the system used for validation in this thesis.

4.1.1 Discussion of system parameters

The following paragraphs provide a guideline to carefully choose the system parameters and to pave the way over some technical pitfalls.

Centre frequency

Usually, the system model, describing the radio system, is considered to be independent of the parametric model, which describes the propagation environment. However, one must not forget that the properties of the underlying wave propagation may change significantly at different frequencies. For example, I will show in Section 7.4.1 that I observed more, but smaller clusters at 5.25 GHz than at 2.55 GHz. In this respect, when changing the centre frequency in the system model, the environment parametrisation must also be adapted.

System bandwidth and number of frequencies

The modelling approach of the path delays assumes infinite bandwidth of the system. By introducing the system bandwidth B, the intrinsic delay resolution of the system is determined by $\Delta\tau = 1/B$. The frequency spectrum is sampled with M frequencies in this bandwidth. Using the discrete Fourier transform, the number of taps is then also equal to the number of frequencies. Also the maximum resolvable delay $\tau_{\max} = M/B$ is fixed in this way.

Of course, the maximum resolvable delay needs to be adjusted to the underlying propagation environment. A useful approximation is to set the maximum delay according to the expected delay spread $\tau_{\max} = 3\sigma_\tau$, hence $M \geq B \cdot 3\sigma_\tau$.

Antenna responses

Note that the antenna responses $\mathbf{a}_{\text{Rx}}(\varphi_{\text{Rx}}, \theta_{\text{Rx}})$ and $\mathbf{a}_{\text{Tx}}(\varphi_{\text{Tx}}, \theta_{\text{Tx}})$ are vector-valued complex functions of the directions of arrival and departure, respectively.

These functions can be determined in two ways:

1. When using *theoretical (ideal) antenna arrays*, the antenna responses can be calculated analytically. The analytical function can be directly implemented into the RCM.
2. When using *measured antenna patterns*, a good interpolation function is needed to represent the antenna responses at intermediate angles that were not measured. The well-known Effective Aperture Distribution Function (EADF) [74, 75] is tailor-made for this purpose.

RF responses

By using the RF response matrices $\mathbf{G}_{\text{Rx/Tx}}(\Delta f)$, the transmission system can be assigned an additional RF system response for every individual antenna train. Also the response of a matched filter can be included here.

4.1.2 System model used in this thesis

Table 4.1 describes the system models that I used in this thesis.

	System I	**System II**	**System III**
Antennas	4×4 ULA	4×4 ULA	8×8 ULA
Centre frequency	2.55 GHz	5.25 GHz	2.55 GHz
Bandwidth	B = 20 MHz		B = 100 MHz
Samples in frequency domain	M = 32		M = 64
System response	$\mathbf{G}_{\text{Rx/Tx}}(\Delta f) = \mathbf{I}$		

Table 4.1: System models used in this thesis

System 1 and System 2 are mainly used for validating the channel model with respect to the environment parametrisation (see Section 4.2). System 3 is used for validating the performance of the DMP model. For all three systems, I consider propagation, in the *horizontal plane only*[1] (i.e. no elevation).

[1] Note that this is the system model used in the RCM, i.e. for *modelling* the channels. Of course, for parameter estimation, another model is used, and elevation is not disregarded, there.

4.2 Environment parametrisation and validation

The performance of the RCM is depending on the accuracy of the environment parametrisation. I will present three possible approaches to parametrise the environment pdf in the order of increasing accuracy, and compare them in terms of the validation metrics in Section 4.2.1. Then, Section 4.2.2 details the impact of the clustering algorithm to the model accuracy. Finally, I will validate the smoothly time-variant behaviour of the RCM in Section 4.2.3.

4.2.1 Methods to parametrise the environment pdf

The environment pdf is a statistical representation of the cluster parameters. This distribution can be modelled in many ways. I will present three methods that were either used in literature, or turned out to be useful for modelling the cluster parameters.

The first method considers the different *cluster parameters uncorrelated*, and models only the second-order statistics (mean and variance). Using these second-order statistics, the Gaussian distribution is chosen as underlying distribution, as it is maximising the entropy.

The second method takes correlation between the cluster parameters into account. This correlation is modelled by the *second-order correlation matrix*. Using this method, one cluster is modelled by a single, but multi-dimensional Gaussian distribution, with correlated dimensions.

The third method models the environment pdf by a mixture of Gaussian distributions. In contrast to the previous methods, the underlying distribution can now fit any distribution with arbitrary accuracy. The most significant advantage of using a Gaussian mixture distribution is the ability to model multi-modal distributions (pdfs showing two or more peaks), which turns out to be necessary for specific indoor scenarios.

All three approaches have in common that they rely on the knowledge of the underlying wave propagation of the scenarios to be modelled. The first method can (to some extent) still be parametrised from literature. Parameters for the second methods are also already available, however they are still very scarce. The third method needs a powerful cluster identification-and-tracking algorithm to allow for a significant parametrisation at the current time.

First, I will describe the three parametrisation methods in detail. Then I will discuss the RCM's performance using these three methods. This comparison will also show the RCM's specificity to its accurate parametrisation.

The RCM comparison is the more sensitive to its parametrisation when I omit the DMP[2]. So, to scrutinise the parametric channel model and its parametrisation, I will validate the RCM *without DMP*.

4.2.1.1 Method 1: Mean and variance

This method models the different cluster parameters statistically by their mean and variance, but mutually *un*correlated.

[2]The overall fit would of course improve for all three methods when I include DMP. The question to answer here is how different the three models are.

Since the cluster parameters are only described by their second-order statistics (mean and variance), a Gaussian distribution (maximising the entropy) is imposed for each of the parameters. In terms of the RCM, the environment parameter pdf $\mathbf{\Theta}_{\text{env}}$ is described *only* by the marginal pdfs of the cluster parameters, which are described in Section 2.3.1.3.

This method does not take any correlation between the cluster parameters into account. To provide an example, there is no connection between the cluster rms spreads taken into account. One can expect that this approach will intrinsically show a certain model mismatch.

Still, this method is the current work horse found in literature throughout a great number of GSCM models, like the COST 259 or the COST 273 models.

The advantage of this method lies in the easy way to estimate the mean and the variance of the cluster parameters from measurements. Given a large number of identified clusters with their parameters $\hat{\mathbf{\Theta}}_c,\ c = 1 \dots C$, the mean and variance of the cluster parameters can be obtained by the common estimators.

The mean cluster parameters are then collected in $\boldsymbol{\mu}_{\mathbf{\Theta}_c}$. The cluster parameter spreads are collected in $\mathbf{C}_{\mathbf{\Theta}_c}$, which is a diagonal matrix because of the assumption that the cluster parameters are uncorrelated.

In the angular domain (mean AoA, mean AoD), care must be taken when estimating mean angles. Given a set of angles $\phi_l,\ l = 1 \dots L$, the mean angle can be calculated as

$$\bar{\phi} = \text{angle}\left(\sum_{l=1}^{L} \exp(j \cdot \phi_l)\right). \tag{4.1}$$

For estimating the variance of angles, the subtraction operation has to be mapped to the principal value $\text{pv}(\cdot)$ on $(-\pi, \pi]$, hence

$$\sigma_\phi^2 = \frac{1}{L-1} \sum_{l=1}^{L} \text{pv}(\phi_l - \bar{\phi})^2. \tag{4.2}$$

For applying (4.1) it is important to note that this calculation only holds true when the sum does not have 0 as result. This outcome is only possible in the pathological case of having only pairs of angles, each showing a difference of $180°$. In my measurements, I always observed more complex environments, hence (4.1) is a good and exact practical solution for estimating the mean angle. Another method overcoming these concerns is outlined in [76].

An advantage of this simple parametrisation is that the mean and the variance of many cluster parameters are already described in literature, e.g. [46, 77, 78, 11]. However all such-obtained cluster parameters must be considered deliberately before using them in the RCM.

4.2.1.2 Method 2: Correlated cluster parameters

A straight-forward improvement of the first method is to incorporate the correlation between the cluster parameters. Having identified cluster parameters from measurements, it is easy to estimate these correlations. I will first motivate this approach by presenting how cluster parameters are correlated, then I will describe how to model this effect.

Observation

I reported the correlation coefficients between cluster parameters the first time in [52], where I evaluated the static cluster parameters only. I extended this concept also to the cluster movement parameters.

The correlation coefficients are calculated as follows[3]: Let $\mathbf{C}_{\mathbf{\Theta}_c}$ be the covariance matrix of the cluster parameters, $\mathbf{C}_{\mathbf{\Theta}_c} = \mathrm{E}\{(\mathbf{\Theta} - \mathrm{E}\{\mathbf{\Theta}_c\})(\mathbf{\Theta}_c - \mathrm{E}\{\mathbf{\Theta}_c\})^{\mathrm{H}}\}$, then the (i,j)th element of the *correlation coefficients matrix* $\mathbf{\Phi}$ is defined as

$$\mathbf{\Phi}(i,j) = \frac{\mathbf{C}_{\mathbf{\Theta}_c}(i,j)}{\sqrt{\mathbf{C}_{\mathbf{\Theta}_c}(i,i) \cdot \mathbf{C}_{\mathbf{\Theta}_c}(j,j)}}. \tag{4.3}$$

This matrix is symmetric and all the diagonal elements are unity, since the power was normalized. The off-diagonal elements quantify the correlation between the respective dimensions in the range of $[-1, 1]$, where 0 denotes no correlation, and 1 (or -1) linear (inverse) dependence.

Figure 4.1 exemplifies the cluster correlations, which were evaluated for a measurement route in a cafeteria at 2.55 GHz (Route TxR22, see page 176).

1. *Power and number of clusters* The two key parameters of the RCM are the snapshot power and the number of clusters. These parameters are strongly correlated with many of the other cluster parameters. There is also a strong correlation between these two parameters, strong power is correlated with a low number of clusters. This fact is somewhat surprising, since one might think that many clusters will carry more power than few clusters.
2. *Cluster power ↔ number of paths* The cluster power is strongly correlated with the number of paths within a cluster. Intuitively, the more paths there are within a cluster, the stronger the cluster gets. There is one exception to this general behaviour, which is the LOS cluster. Since there is only one LOS cluster in each scenario (or up to three dominant clusters, coming from specular reflections) that has few paths but large power, this is statistically not relevant for the correlation coefficient.
3. *Cluster power ↔ cluster mean delay* From wave propagation it is clear that early arriving clusters carry more power than clusters with large delay. So this fact comes to no surprise.
4. *Cluster size: cluster delay spread ↔ AoA cluster spread ↔ AoD cluster spread* The "size" of a cluster in delay and angles are described by the cluster spreads. We observe strong correlation between the cluster azimuth spread at the transmitter, the cluster azimuth spread at the receiver, and the delay spread. If a cluster is "large" when seen from the Tx, it is also "large" at the Rx, so the delay spread must also be large. In this investigated scenario the correlation is quite strong since Tx and Rx were placed within the same room. For NLOS scenarios, this correlation will be weaker. In outdoor macro-cell environments, this correlation will be close to zero due to the small size of clusters seen from the base station.
5. *Number of clusters ↔ cluster size* Representatively for the other cluster size parameters, we comment on the correlation between the number of clusters and the AoA cluster spread. Here, we observe a strong inverse dependence, meaning that a large number of clusters results in smaller clusters. This observation is surprising in the sense that there is no straightforward physical explanation for it. The reason for this effect can be found in the

[3] $\mathrm{E}\{\cdot\}$ denotes the expectation operator

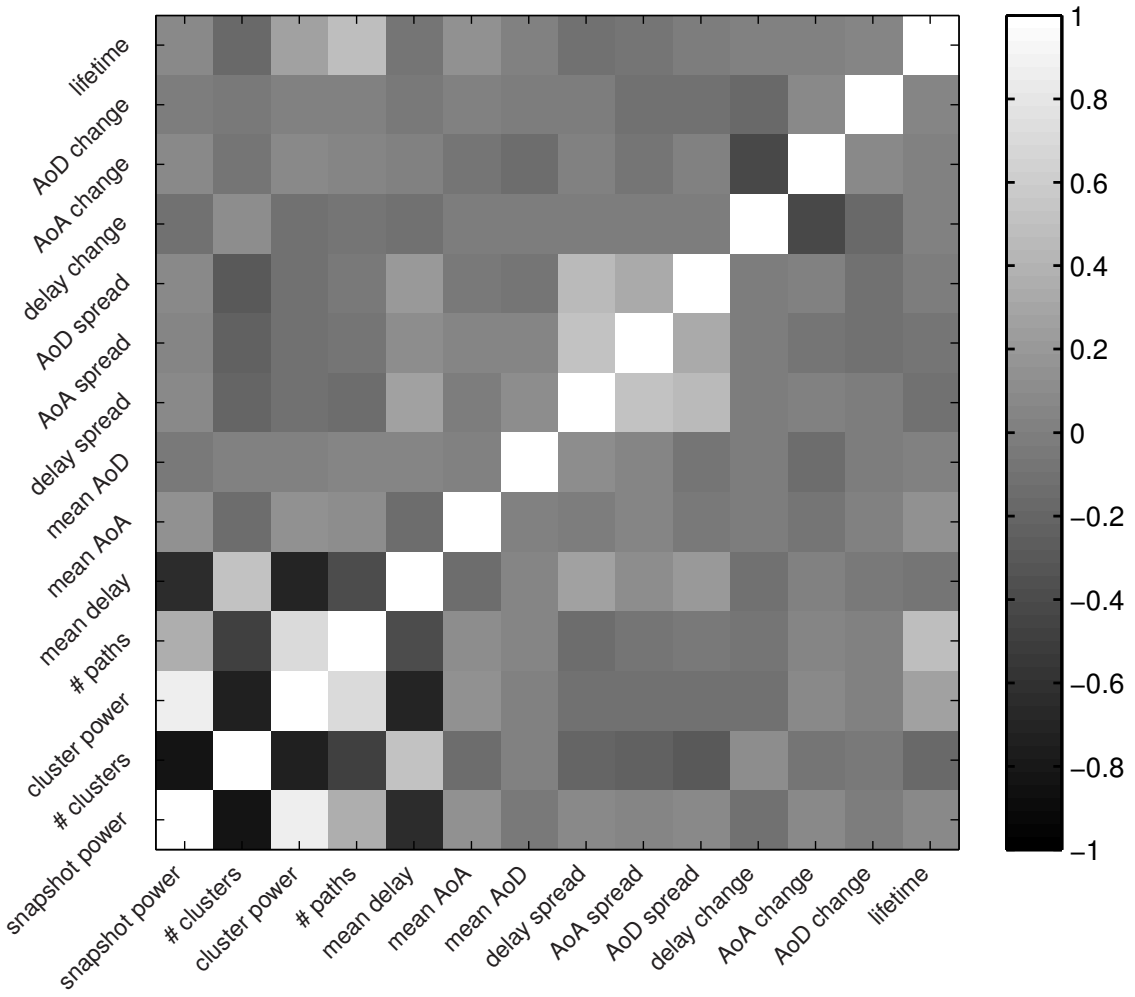


Figure 4.1: Correlation of cluster parameters exemplary evaluated for a measurement route in a cafeteria at 2.55 GHz (Route TxR22, see page 176).

clustering algorithm, which splits up the scenario into a number of clusters, until the number of clusters is large enough to represent the scenario accurately enough (cf. Section 7.2.2).

6. *Cluster movement parameters* When a mobile moves, the Tx and Rx angles of the clusters, as well as their delays are likely to change. A correlation between these cluster parameters is very likely, because of dominant movement directions. Of course, this correlation strongly depends on the environment. In the exemplary case the delay change is negatively correlated with the AoA position change.
7. *Cluster lifetime ↔ cluster power ↔ number of paths* The lifetime of a cluster is correlated with the cluster power and the number of paths within a cluster. This implies that strong clusters show a longer lifetime. Again, this is intuitive since strong clusters are mostly attributed to the LOS path or strong reflections. These paths dominant contributions can be tracked well.

Modelling correlated cluster parameters

Combining the knowledge of the correlation coefficients of the cluster parameters (4.3), the cluster mean parameters, $\boldsymbol{\mu}_{\boldsymbol{\Theta}_c}$, and the variance of the cluster parameters, $\mathbf{C}_{\boldsymbol{\Theta}_c}$ (which were detailed in the previous method), the underlying environment pdf is described using the multi-variant Gaussian distribution[4].

The resulting environment pdf is then given as

$$\boldsymbol{\Theta}_{\text{env}} = p_{\boldsymbol{\Theta}_c}(\boldsymbol{\theta}_c) = \frac{1}{(2\pi)^{D/2}|\mathbf{C}_{\boldsymbol{\Theta}_c}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{\theta}_c - \boldsymbol{\mu}_{\boldsymbol{\Theta}_c})^{\mathrm{T}}\mathbf{C}_{\boldsymbol{\Theta}_c}^{-1}(\boldsymbol{\theta}_c - \boldsymbol{\mu}_{\boldsymbol{\Theta}_c})\right), \tag{4.4}$$

where $D = 21$ denotes the number of dimensions (i.e. the number of cluster parameters). The elements on the main diagonal of $\mathbf{C}_{\boldsymbol{\Theta}_c}$ are given as described in Section 4.2.1.1, and the off-diagonal elements as

$$\mathbf{C}_{\boldsymbol{\Theta}_c}(i,j) = \boldsymbol{\Phi}(i,j) \cdot \sqrt{\mathbf{C}_{\boldsymbol{\Theta}_c}(i,i) \cdot \mathbf{C}_{\boldsymbol{\Theta}_c}(j,j)}. \tag{4.5}$$

4.2.1.3 Method 3: Gaussian mixture parametrisation

The third approach uses the concept of multivariate kernel density estimation [79] to provide the cluster parameter distribution function for the model.

The idea is to approximate the environment pdf using a sum of kernels

$$\boldsymbol{\Theta}_{\text{env}} = p_{\boldsymbol{\Theta}_c}(\boldsymbol{\theta}_c) = \frac{1}{n}\sum_{i=1}^{n} K(\boldsymbol{\theta}_c, \boldsymbol{\mu}_{\boldsymbol{\Theta}_i}, \mathbf{C}_{\boldsymbol{\Theta}_i}), \tag{4.6}$$

where $\boldsymbol{\mu}_{\boldsymbol{\Theta}_i}$ and $\mathbf{C}_{\boldsymbol{\Theta}_i}$ denote the mean and covariance of the ith kernel.

To parametrise the environment pdf for the RCM, I use Gaussian kernels, hence a Gaussian mixture pdf, such that

$$K(\boldsymbol{\theta}_c, \boldsymbol{\mu}_{\boldsymbol{\Theta}_i}, \mathbf{C}_{\boldsymbol{\Theta}_i}) = \frac{1}{(2\pi)^{D/2}|\mathbf{C}_{\boldsymbol{\Theta}_i}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{\theta}_c - \boldsymbol{\mu}_{\boldsymbol{\Theta}_i})^{\mathrm{T}}\mathbf{C}_{\boldsymbol{\Theta}_i}^{-1}(\boldsymbol{\theta}_c - \boldsymbol{\mu}_{\boldsymbol{\Theta}_i})\right). \tag{4.7}$$

[4] Following again the maximum-entropy argumentation, the Gaussian distribution is most appropriate.

The kernel parameters $\boldsymbol{\mu}_{\boldsymbol{\Theta}_i}$, and $\mathbf{C}_{\boldsymbol{\Theta}_i}$ need to be estimated. The input data for this estimation are the identified clusters from a measurement route.

A straight-forward way to find the kernel parameters is to choose the number of kernels equal to the identified number of clusters. Each individual cluster is used as (mean) parameter for an individual kernel. The variances of the kernel can then be estimated using the well-known minimum average mean integrated squared error (AMISE) criterion [79]. This parametrisation approach is the most accurate one, although the number of kernels may become quite large.

Another approach to parametrise the kernels is by using Gaussian mixture clustering methods. In this way, the number of kernels can be traded with the accuracy of the parametrisation.

When choosing the number of kernels as 1, this method reduces to the correlated cluster parameters method as described in the previous section.

The implementation of the RCM uses the code from Ihler [51], which allows to marginalise and condition the mixture pdf, and to draw new samples from the mixture pdf.

With this approach, distributions that do *not* follow a single (multi-variate) Gaussian distribution can be described, easily.

For demonstration I estimated the Gaussian mixture pdf for the measurement route in a cafeteria at 2.55 GHz (Route TxR22, see page 176), which was also used in the previous section to exemplify the strong correlation of the cluster parameters. Figure 4.2 presents different cross sections of this pdf. The marginalisation to two cluster parameters at a time is done in order to see the correlation of the parameters.

The figure reflects the correlations identified in Figure 4.1 nicely, but from some marginal distributions it becomes clear that a single multivariate Gaussian distribution is not sufficient to approximate the pdf. For example, the joint pdf of snapshot power and number of clusters clearly illustrates a more complicated correlation.

The great advantage of using a Gaussian mixture pdf is that also multimodal distributions[5] can be captured. Figure 4.3 shows the marginal pdfs from another measurement in an indoor NLOS office scenario (measurement route TxR11 at 2.55 GHz, see pg. 165). The figure illustrates that the underlying distribution cannot be approximated using a single multivariate Gaussian distribution, at all. The effect of this deficiency will become clearly visible during the validation in the following section.

4.2.1.4 RCM validation of the random-access mode

For the validation, I used the random-access mode of the parametric channel model. In this way, a number of different channels that all root in the same environment are generated. The validation of the smoothly time-variant mode of the RCM is considered separately in Section 4.2.3.

System models I and II, as described in Table 4.1, were used to calculate the channel matrices from the parametric channels, depending on the carrier frequency used in the underlying measurements. Since the main focus of this validation is on the parametric channel model, DMP was not considered in the channels.

[5] Multimodal distributions show more than one peak in their pdf

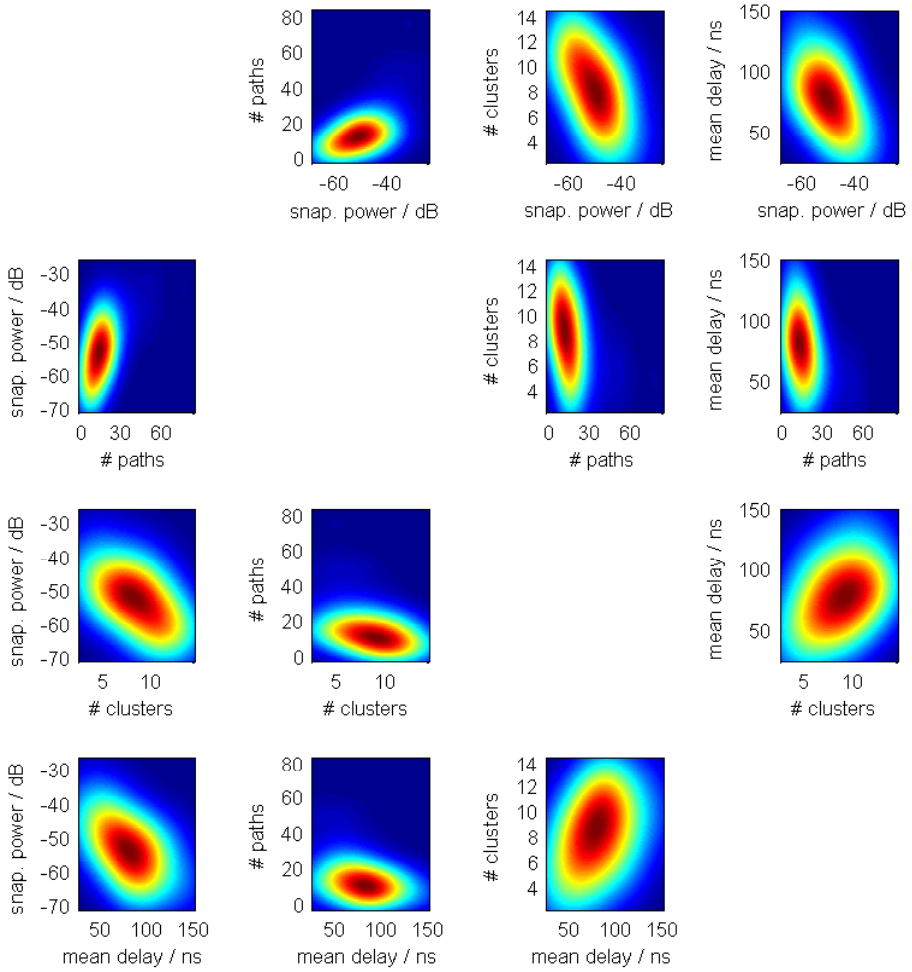


Figure 4.2: Cross sections of environment pdf, evaluated for measurement TxR22. Approximating this distribution by a single multivariate Gaussian distribution is difficult, but possible.

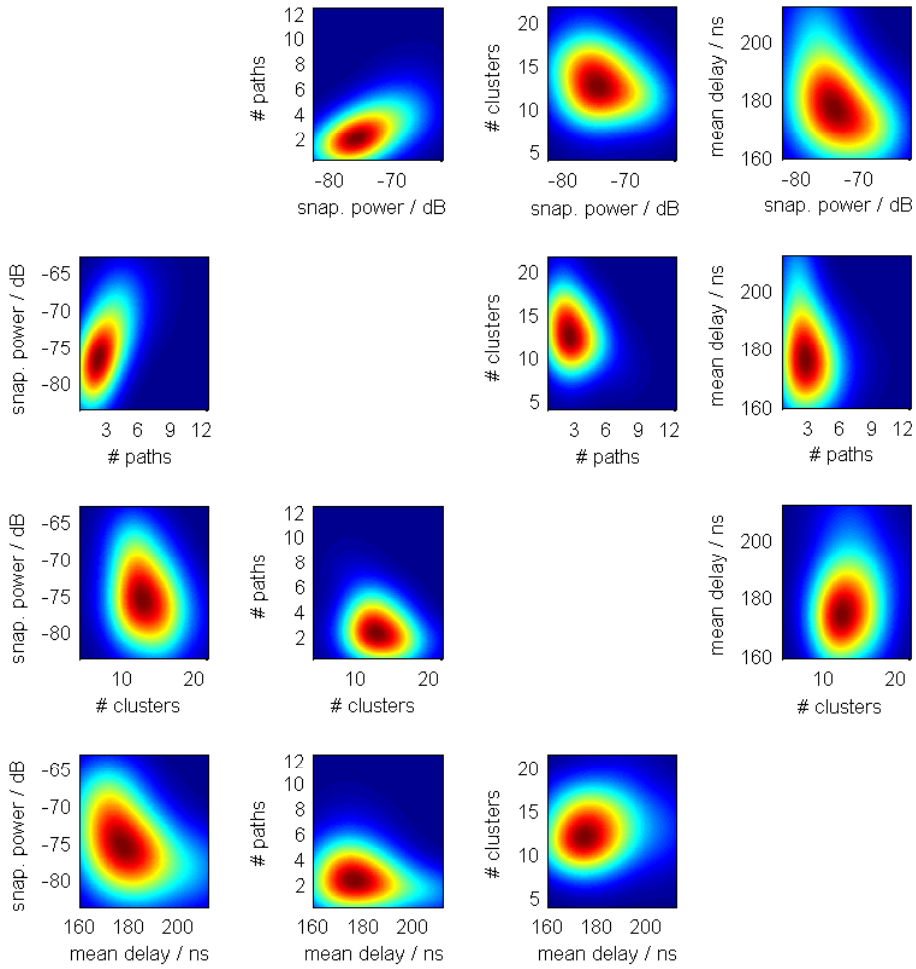


Figure 4.3: Cross sections of environment pdf, evaluated for measurement TxR11. Approximating this distribution by a single multivariate Gaussian distribution is not possible any more.

First, I compare the RCM's fit for the cafeteria environment (Route TxR22 at 2.55 GHz, see pg. 176) using the validation procedure as described in Section 3.1. The environment pdf was parametrised using the three methods, where I chose the maximum number of kernels for Method 3.

Figure 4.4 demonstrates the performance of the RCM. Figures 4.4a-e provide the SVs of the ECM (cf. Section 3.2.4), while Figure 4.4f shows the wideband mutual information metric for constant Tx power with 10 dB SNR (cf. Section 3.2.1). The black, solid line denotes the reference channels, the blue and red line denote method 1 (uncorrelated cluster parameters) and method 2 (correlated cluster parameters), respectively, and the magenta dashed line denotes method 3 (Gaussian mixture parametrisation, "KDE").

The ECM metric clearly demonstrates the superiority of the Gaussian mixture parametrisation. Both the largest and smallest SV-ECM exhibit a very good fit to the reference channels. This fit is also obvious in the MI metric.

Neither the correlated nor the uncorrelated parametrisation method are able to represent the scenario correctly. This is due to the strong directivity of the underlying channels. Note that the correlated parametrisation method fits better at small SVs-ECM, also the slope fits quite well. So, the correlated parametrisation method still shows a better fit than the uncorrelated parametrisation method. The underlying cluster distribution seems to be well approximated by a single multivariate Gaussian distribution, as already conjectured in Figure 4.2. This impact is also visible in the fit of the MI metric, where the mean is fit quite well, but the slope is not represented correctly.

Repeating this experiment with measurements in an NLOS office scenario (Route TxR9, see pg. 163 provided following results (see Figure 4.5).

In this scenario, the the distance between the Gaussian mixture parametrisation and the other two methods is even stronger expressed. The reason for this is the very directive channel around the Rx. Basically, there are two dominant directions, one looking into the corridor towards the Tx, the other is the backward wall reflection. This property cannot be captured by the two approaches that parametrise the scenario with a single Gaussian distribution.

Surprisingly, in this scenario the correlated cluster method is even inferior to the uncorrelated cluster method. The reason for this is that the underlying distribution is too complex to be described by a *single* multi-dimensional Gaussian pdf.

The Gaussian mixture method is again well able to parametrise the propagation environment. The SV-ECM curves show a much better fit, both in mean and slope.

Another interesting effect that can be observed in this scenario is the bad performance of MI as validation metric. Even though the Gaussian mixture parametrisation outperforms the other two methods in terms of the SV-ECM, MI does not reflect this.

The same effects become visible when further increasing the directivity of the channels. This can be observed in Figure 4.6, detailling another indoor NLOS office scenario, where the separation between Tx and Rx was already quite large (TxR11, see pg. 165).

The 5th SV-ECM substantiates the assumption that the underlying distribution consists of a sum of at least two distributions. This effect deteriorates the performance of the correlated parametrisation approach. The underlying distribution can obviously not be modelled correctly using a single multivariate Gaussian distribution, as already conjectured in Figure 4.3. In contrast, the Gaussian mixture parametrisation is able to reproduce this multi-modal distribution nicely.

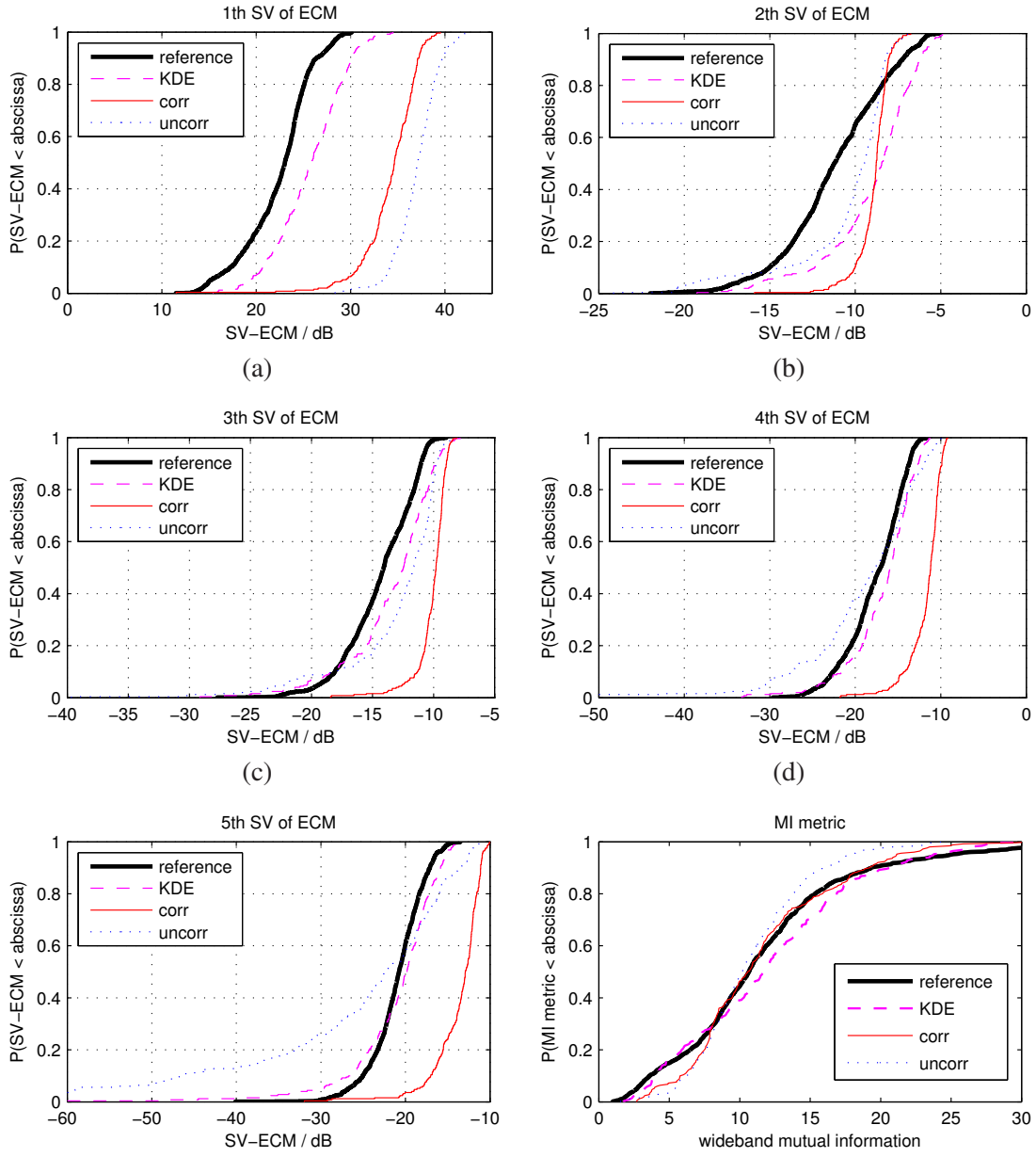


Figure 4.4: Comparison of the three parametrisation methods in Scenario TxR22. (a)-(e) Singular Values of the ECM, (f) MI metric with 10 dB SNR.

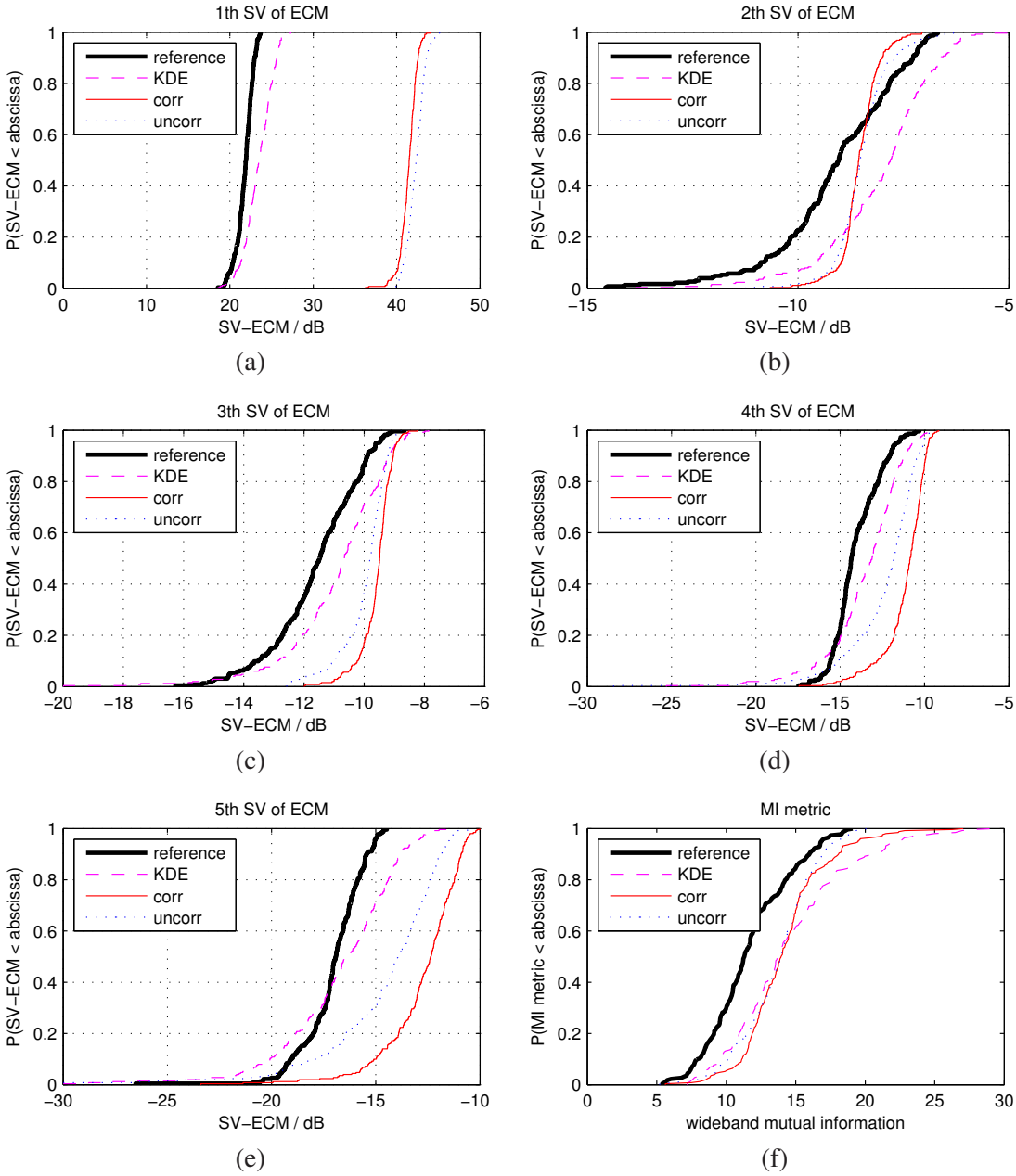


Figure 4.5: Comparison of the three parametrisation methods evaluated for the NLOS office scenario TxR9. (a)-(e) SVs-ECM, (f) MI metric.

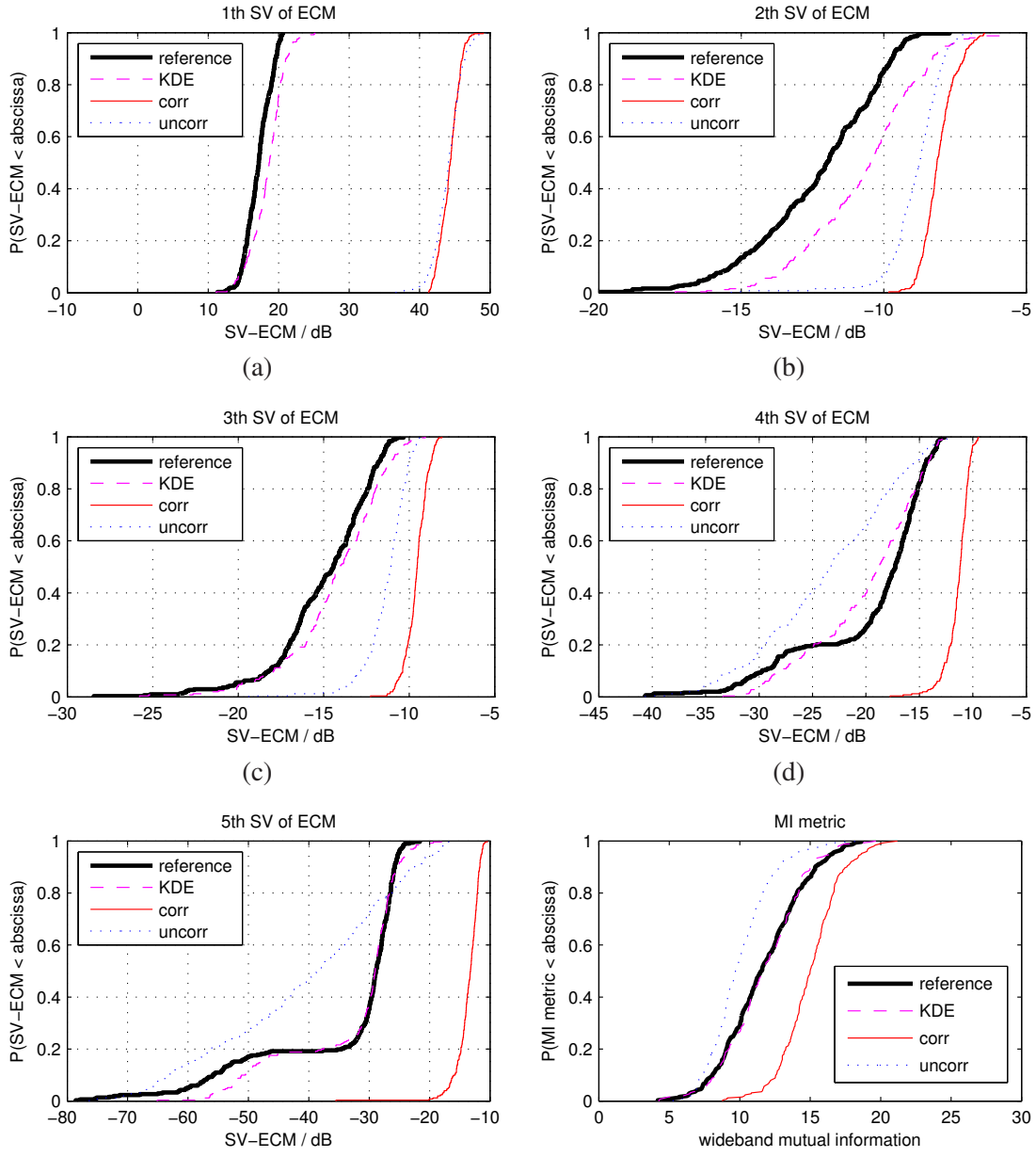


Figure 4.6: Comparison of the three parametrisation methods evaluated for the NLOS office scenario TxR11. (a)-(e) SVs-ECM, (f) MI metric.

Concluding, the *Gaussian mixture parametrisation provides overall the best fit* of the RCM with respect to the validation metrics. Parametrising the RCM using correlated cluster parameters, the performance depends whether the underlying scenario can be approximated using a single multi-dimensional Gaussian distribution. This method reproduces well the slope of both the SV-ECM cdf and the MI cdf. When using uncorrelated cluster parameters to describe the environment, the RCM still reflects the properties of the underlying channel, but with much lower accuracy. Here, the *specificity of the RCM* became clearly visible. The more detailed the parametrisation is, the more accurate the RCM will fit.

I also exemplified why *MI is not well suited as a validation metric*. It reflects the model fit to some extent but is not specific enough.

4.2.2 Impact of clustering algorithm parameters

The three introduced methods use cluster parameters from measurements to obtain the environment pdf. These clusters used for parametrisation need to be identified before.

Chapter 7 will present an automatic clustering-and-tracking algorithm that is well suited to this task. This algorithm needs two input parameters that have a strong impact on the clustering results and thus also the performance of the model:

1. The cluster threshold; it describes the minimum power that a cluster may have relative to the total snapshot power. The cluster threshold and the general problem of cluster thresholding is discussed in Section 6.4.3.
2. The window size; the snapshots of the measured channel are combined in order to achieve a better tracking performance. The window size describes the length of this combining window.

To see the impact of these clustering parameters on the performance of the RCM, I use the validation metrics. Again, the system models I and II were used to calculate the channel matrices, depending on the centre frequency of the underlying measurement. By modelling different kinds of scenarios, I will show which cluster threshold and window size have to be chosen for the RCM to accurately model the scenarios[6].

Impact of cluster threshold

Using a sliding window of 2 snapshots, I chose the cluster threshold from $[0.01, 0.03, \ldots, 0.09]$. The minimum cluster threshold of 0.01 was empirically set, for smaller numbers the resulting number of clusters would get prohibitively large. The maximum cluster threshold of 0.09 already results in a very low number of clusters and turned out to be a kind of upper limit.

Figure 4.7 shows the validation metrics evaluated for the indoor office measurement route TxR11 at 2.55 GHz (see pg. 165). The black bold line indicates the reference channels, while the other lines describe the fit of the RCM where clustering was performed using the different cluster thresholds.

For the first SV-ECM (Figure 4.7a), all kinds of clustering work quite well, however, only using cluster thresholds of 0.01 and 0.03 enables to approximate the distinct structure of the fifth

[6] I also want to acknowledge the work of my master student Paula Macarron Cuartero on this topic.

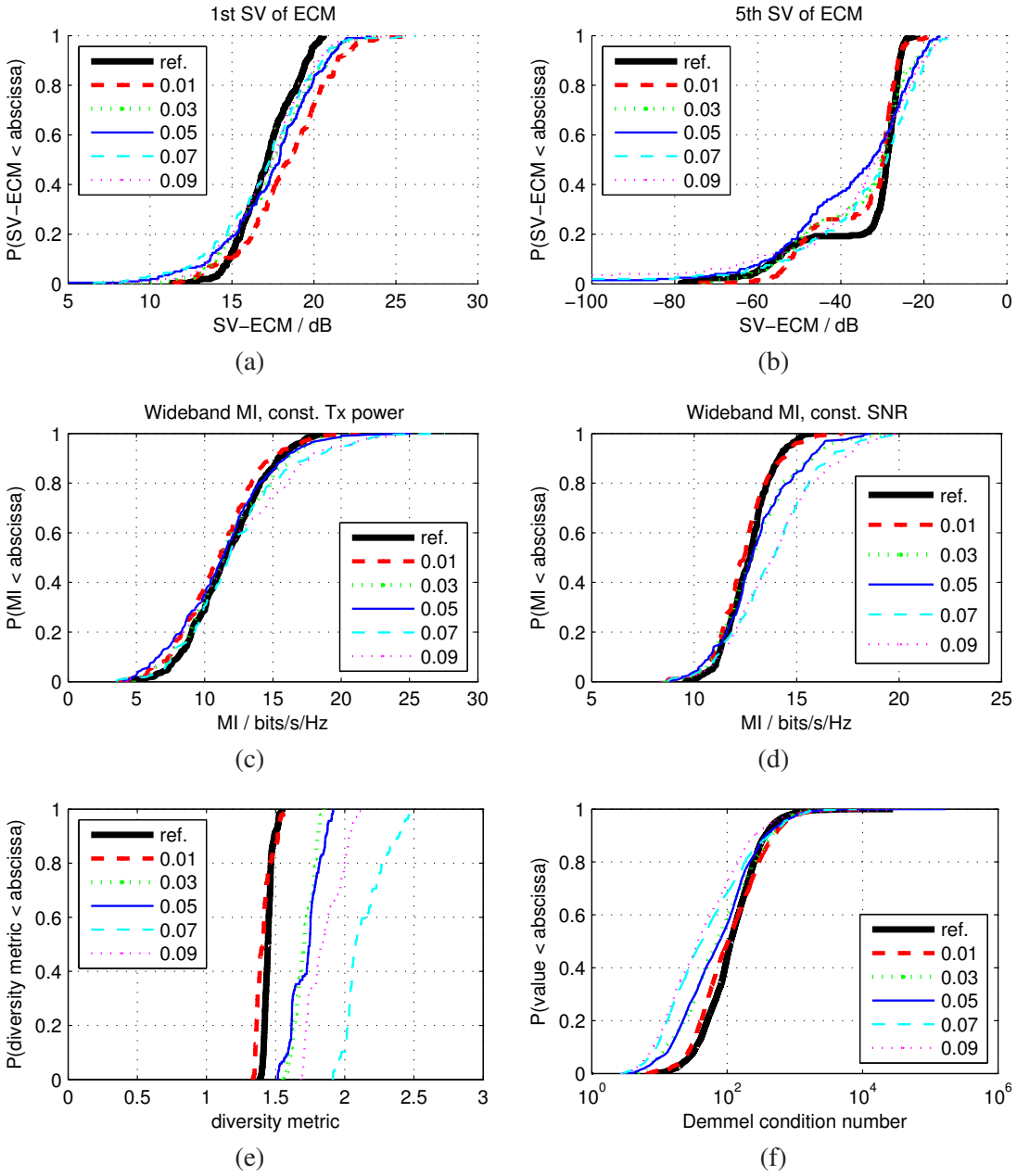


Figure 4.7: Impact of the cluster power threshold of the clustering algorithm on the RCM performance evaluated for measurement route TxR11 at 2.55 GHz (see pg. 165). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

SV-ECM (Figure 4.7b). In both MI metrics (normalisation to constant Tx power or to constant Rx SNR at 10 dB), a cluster threshold of 0.01 attains best results (Figures 4.7c-d), but cluster thresholds of 0.3–0.5 do fit not so bad, either. The same effect holds true for the diversity metric (Figure 4.7e), and for the Demmel condition number (Figure 4.7f).

Of course, these effects strongly depend on the scenario. Figure 4.8 shows a different result for the measurement route in the cafeteria at 2.55 GHz (TxR22, see pg. 176). In this scenario, there is no clear impact of the cluster threshold, all thresholds seem to fit quite well. In this environment there was (mostly) LOS, the channels are very directive, so the cluster threshold does not impact the results significantly.

Comparing this result to the same route measured at 5.25 GHz we see following effects (see Figure 4.9). The match between reference channel and modelled channels is still very good, though (for this very route) it matches a bit worse as it did at 2.55 GHz. Here, more clearly, a small cluster threshold improves the model fit.

Window size

To improve the tracking performance of the clustering-and-tracking algorithm, snapshots of the measured channel are combined in a sliding window. The window size describes the length of this sliding window.

I evaluated the impact of the window size on the model performance while using a cluster threshold of 0.01. Again, I did the validation for individual scenarios, altering the window size from 2 to 6 snapshots.

Figures 4.10–4.12 exemplify the results from these evaluations. The results all show that the window size does not have a significant impact on the RCM. On average, the results are incrementally better for smaller window sizes, but this effect is negligible.

Concluding, a small cluster threshold usually achieves best results, cluster thresholds between 0.01 and 0.05 seem useful for parametrising the environment pdf.

A large window size is beneficial for tracking clusters, but also increases the complexity of the clustering algorithm (cf. Section 7). A window size of 4 snapshots turned out to be a good tradeoff between clustering complexity and tracking accuracy.

4.2.3 Validation of the smoothly time-variant model

First, it is necessary to check, whether the birth/death process also reflects the propagation environment accurately. Then I will validate whether the process of smoothly changing the environment is reflected well. Finally I evaluate the Doppler spectra of the modelled channels.

4.2.3.1 Validating the birth/death process

I validated the reference channels against the modelled channels from the smoothly time-variant mode of the RCM using the Gaussian mixture parametrisation. I adjusted the following time-variance parameters of the RCM (see Section 2.3.1.4), to fit the parameters from the measurements: (i) the speed of the mobile v_{Tx}, (ii) the cluster birth rate, (iii) the cluster lifetime period (which I set equal to the sampling period, $t_\lambda = t_s$, for this evaluation).

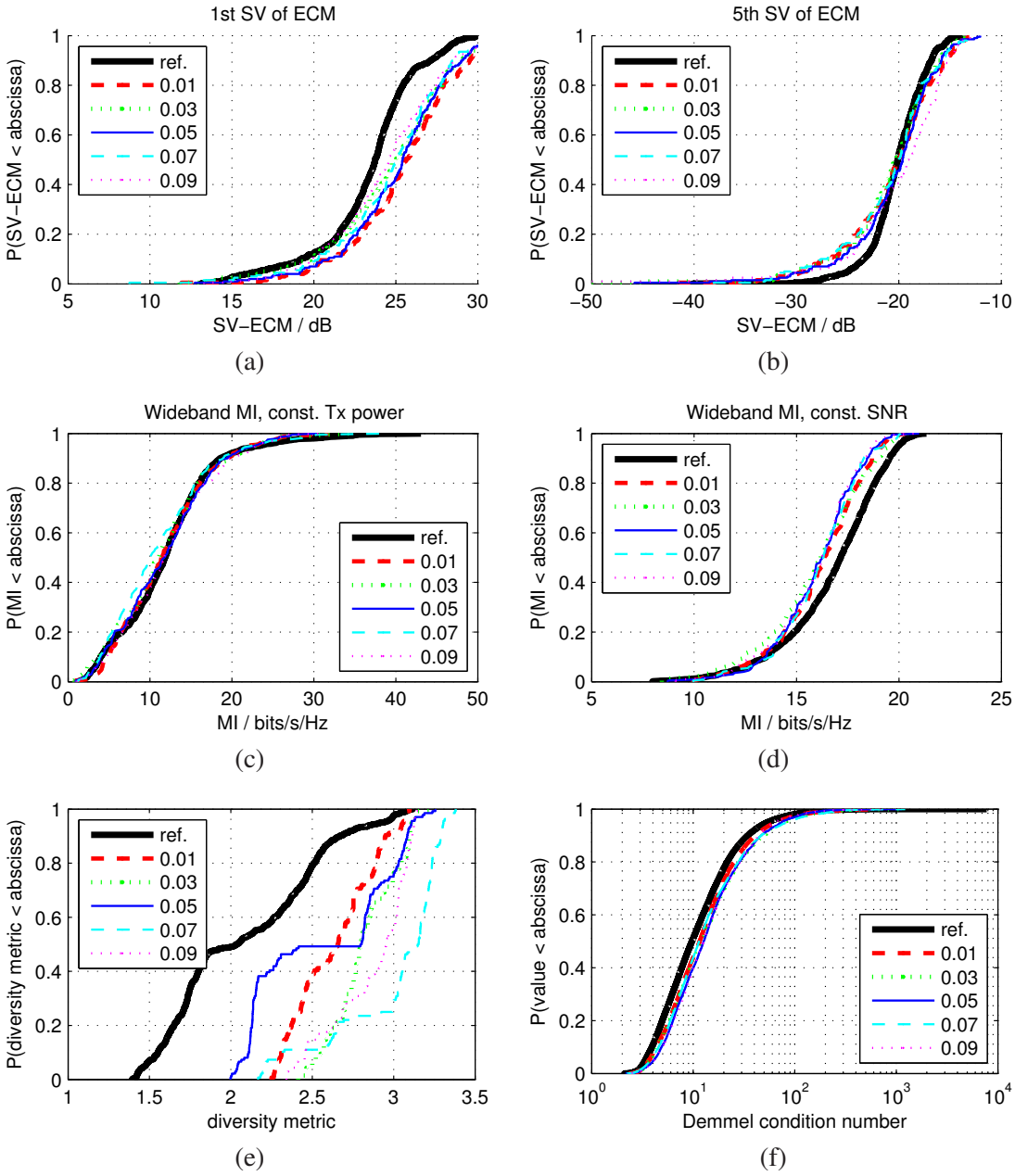


Figure 4.8: Impact of the cluster power threshold of the clustering algorithm on the RCM performance evaluated for measurement route TxR22 at 2.55 GHz (see pg. 176). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

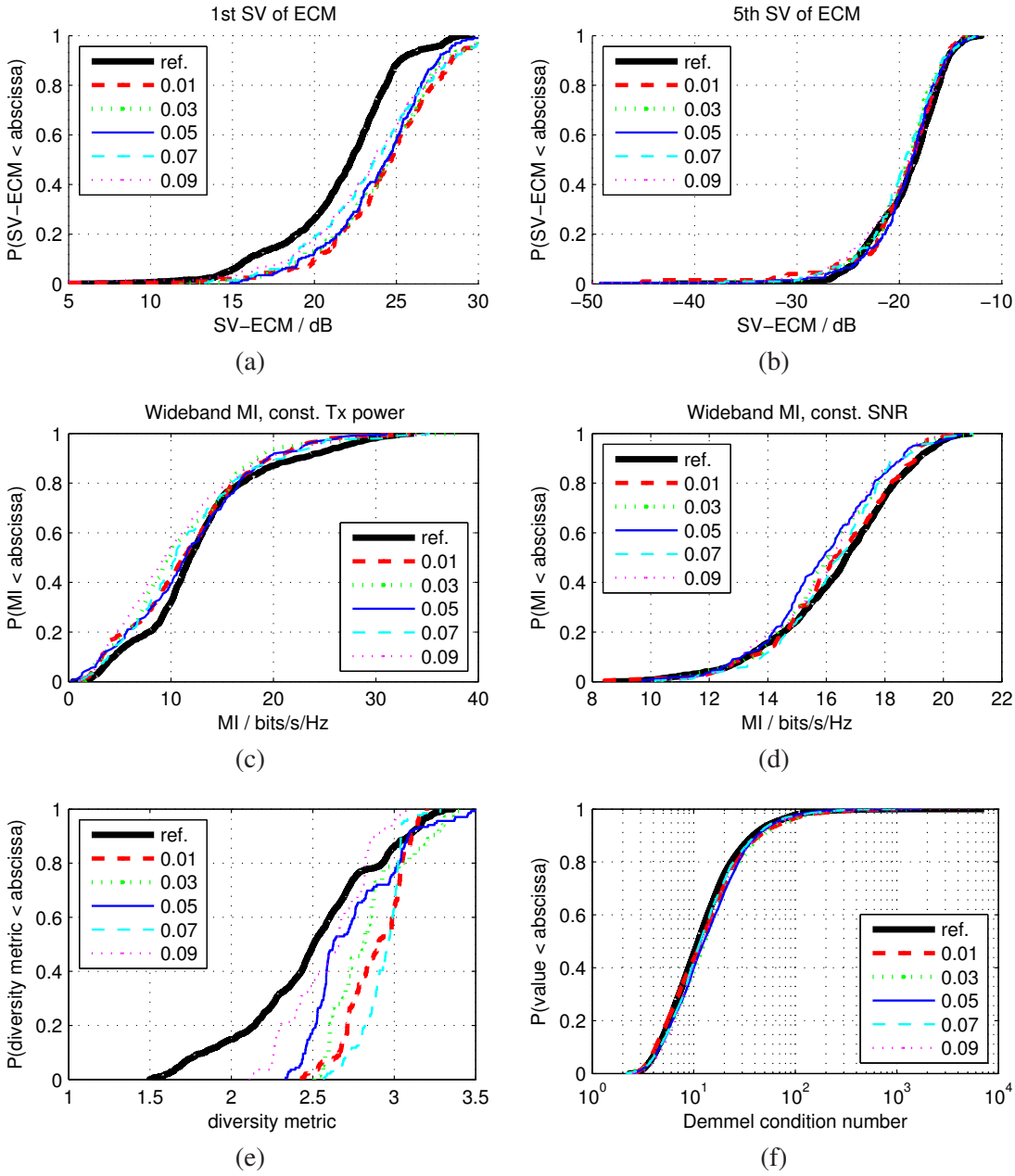


Figure 4.9: Impact of the cluster power threshold of the clustering algorithm on the RCM performance evaluated for measurement route TxR22 at 5.25 GHz (see pg. 176). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

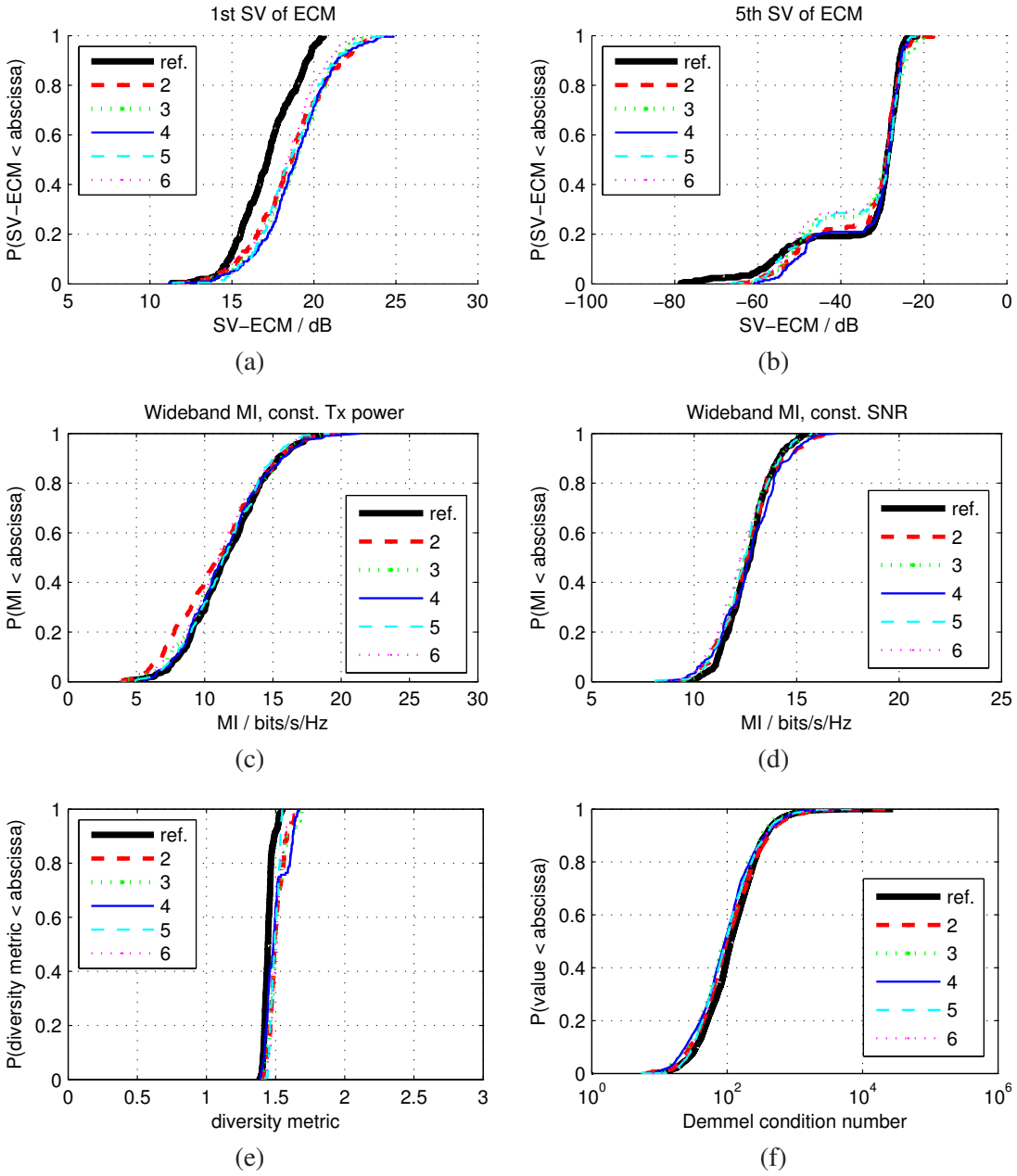


Figure 4.10: Impact of the window size parameter of the clustering algorithm on the RCM performance, evaluated for measurement route TxR11 at 2.55 GHz (see pg. 165). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

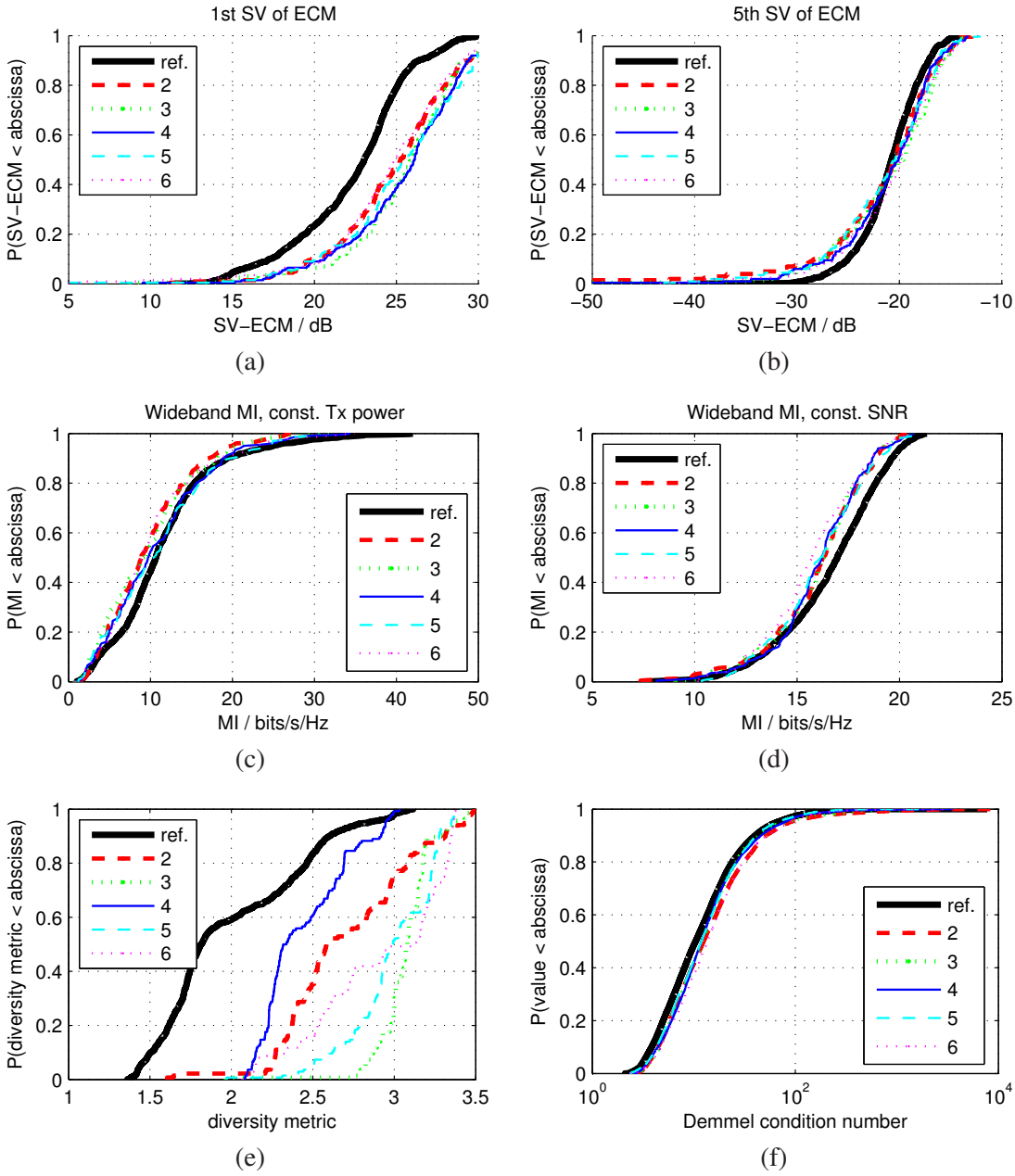


Figure 4.11: Impact of the window size parameter of the clustering algorithm on the RCM performance, evaluated for measurement route TxR22 at 2.55 GHz (see pg. 176). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

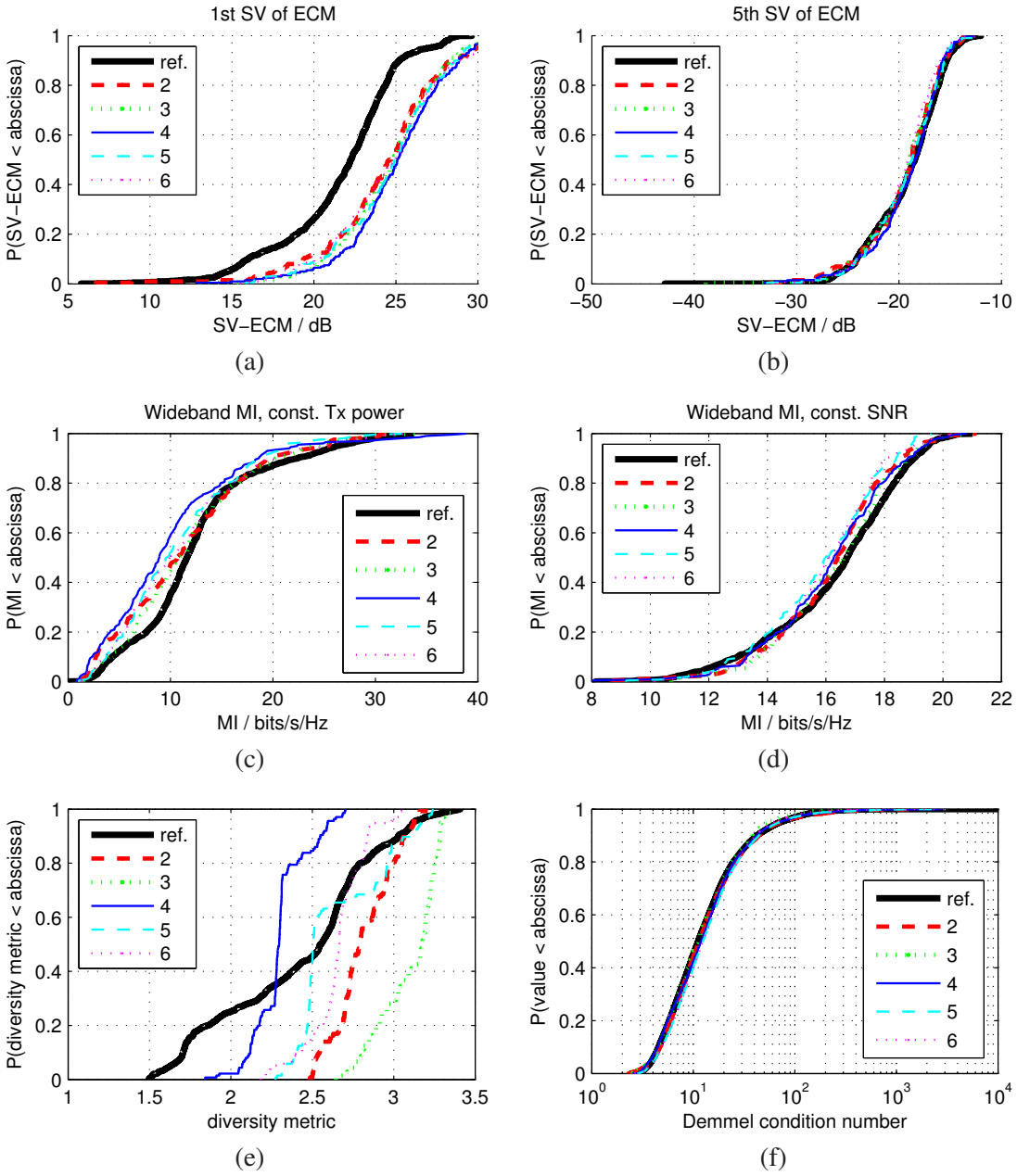


Figure 4.12: Impact of the window size parameter of the clustering algorithm on the RCM performance, evaluated for measurement route TxR22 at 5.25 GHz (see pg. 176). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

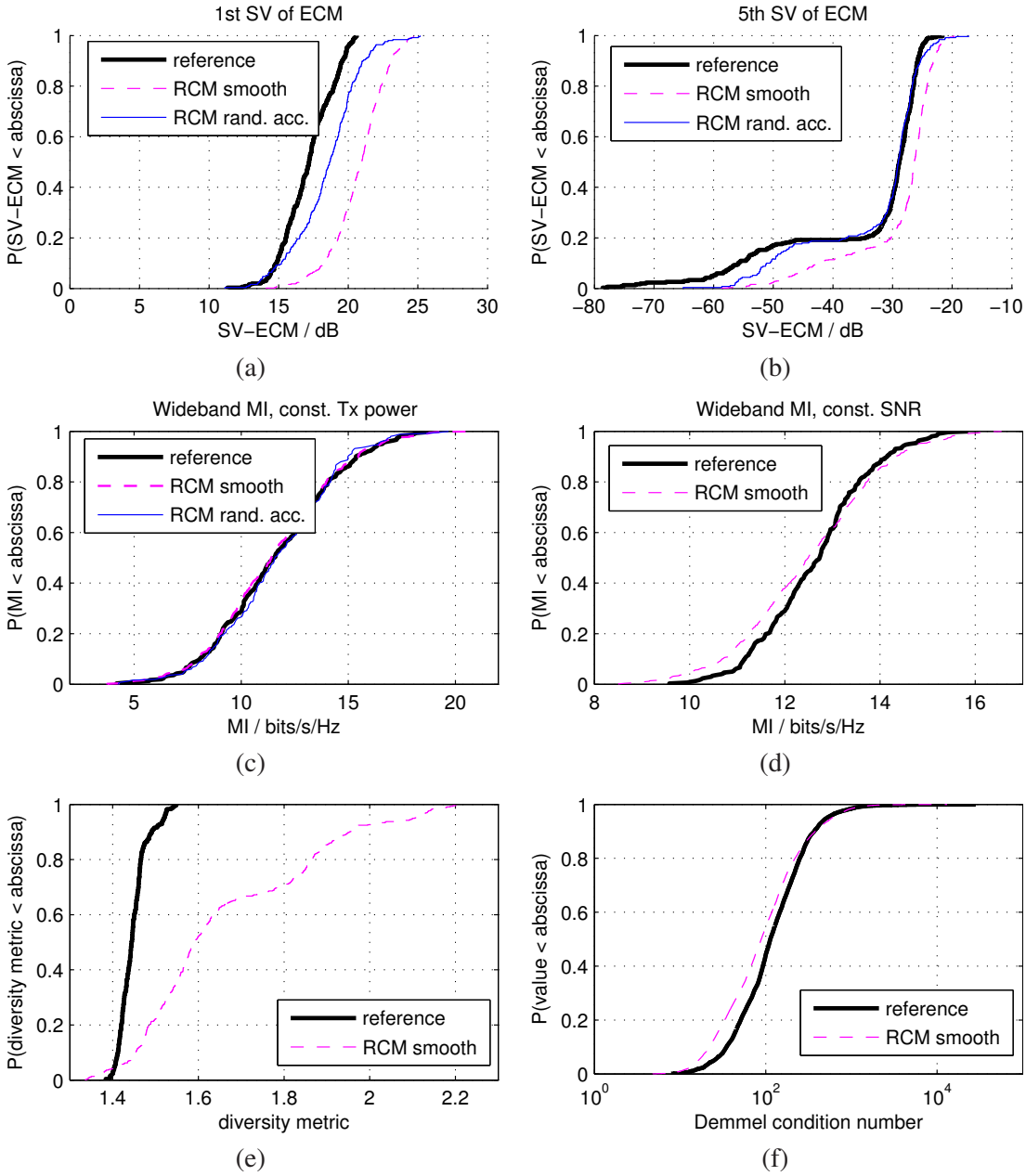


Figure 4.13: Validating the smoothly time-variant behaviour of the RCM, using measurement route TxR11 at 2.55 GHz (see pg. 165). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

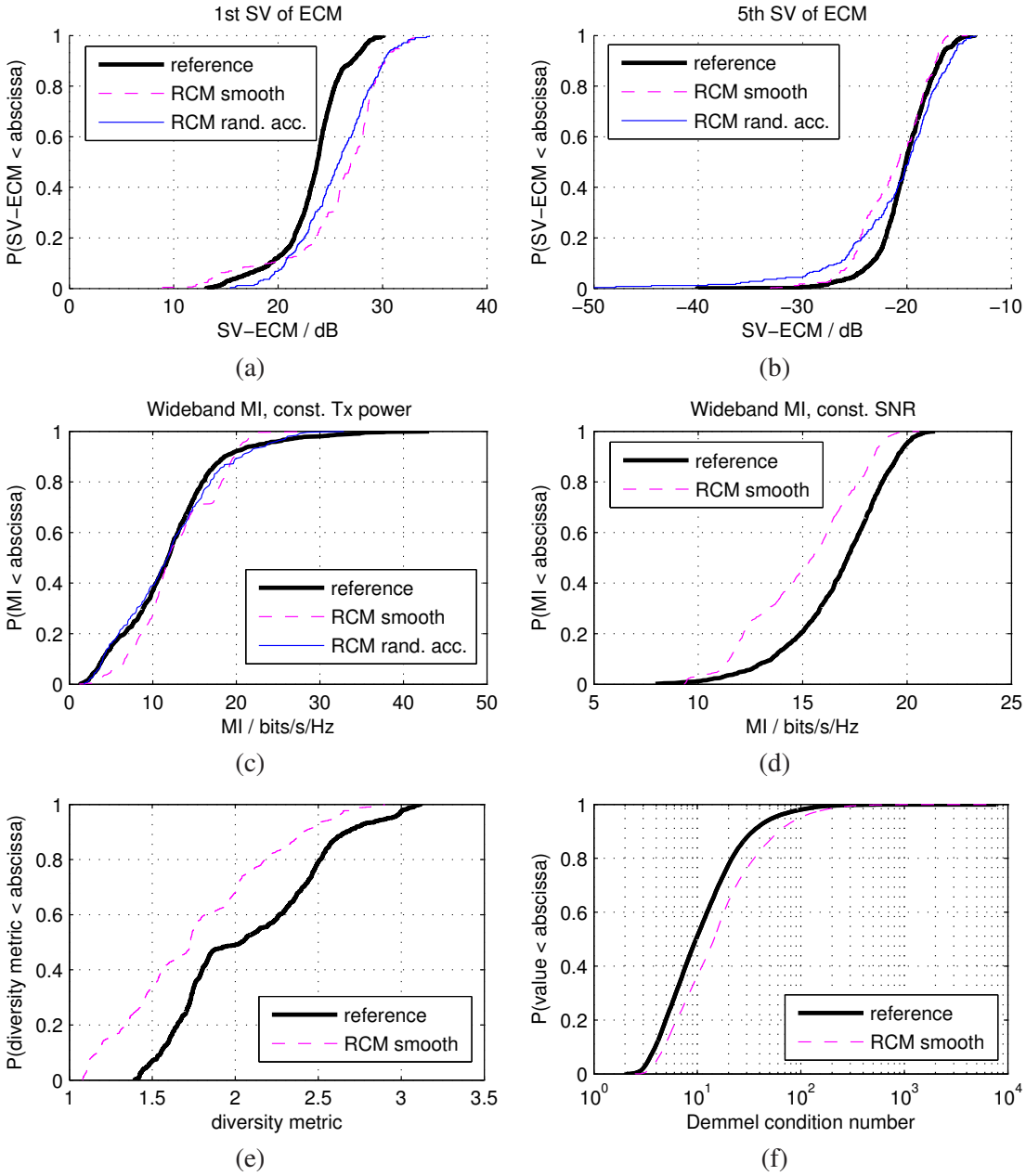


Figure 4.14: Validating the smoothly time-variant behaviour of the RCM, using measurement route TxR22 at 2.55 GHz (see pg. 176). (a)-(b) 1st and 5th SV-ECM, (c)-(d) MI metric, (e) diversity metric, (f) Demmel condition number

Again, the scenarios for the validation are (i) the same indoor NLOS office (TxR11, see pg. 165), and (ii) the cafeteria scenario (TxR22, see pg. 176).

The validation results are presented in Figures 4.13 and 4.14, respectively. Using the SV-ECM to compare these results with the results from the random-access mode in subfigures a–b, the smoothly time-variant mode matches the reference channels nearly as well as the random-access mode does. In contrast to the random-access mode, there is less randomness in the cluster generation. Clusters exist for a certain lifetime, during which they move. For this reason, the smoothly time-variant mode can never fit as well as the random-access mode does. I also compare the MI cdf in subfigues c, but there is no clear difference between the random-access mode and the smoothly time-variant mode. Again, diversity does not match very nicely, which seems to be a general problem of the diversity metric[7].

4.2.3.2 Validating the smooth changing of the environment

Currently, the literature does not provide meaningful measures to judge in which way a MIMO environment changes. Because of the lack of such measures, I validated the RCM using the previously introduced metrics and present the change of the metrics over time. In this way, one can observe the general trend how the reference scenario and the modelled scenario change.

It is important to note that, by the nature of the RCM, any movement is a random walk through a scenario. Clusters are drawn by their probability to exist, which results in scenarios *in the kind of* the observed one. This does not imply that the modelled scenarios change in exactly the same way as the observed ones do.

I used the same time-variance parameters as for the spatial validation presented in the previous subsection. This time I carried out the validation for scenarios TxR13 at 5.25 GHz and TxR22 at 2.55 GHz[8]. Figure 4.15 demonstrate the smoothly time-variant behaviour of the RCM. The left side of the figure describes TxR13, the right side TxR22.

Fist, I compare the SV-ECM (Figure 4.15a,d). It is noteworthy that the first SV-ECM exhibits some jumps. These are due to the setting of $t_\Lambda = t_s$, clusters appear and disappear quickly. The fifth SV-ECM in TxR13 is more steady in the reference channels than it is modelled. Here, the RCM models the impact of movement a bit too strongly. On the other hand, in the cafeteria scenario, changes in the fifth SV-ECM are reflected quite well.

The comparison of the MI metrics in Figure 4.15b-c,e-f illustrate that the *small-scale changes in mutual information* are reflected quite well. Particularly interesting is Figure 4.15f, where the RCM manages to reflect the rise of MI quite well. Of course, this result may be called a lucky coincidence, since the RCM chooses clusters purely randomly. Still, it is a useful result, which demonstrates that the RCM is well able to model this kind of behaviour.

At this point I want to come back to the theme of Part I, "You can always find an scenario that fits your measurements." — Of course, the performance of the RCM depends on the parametrisation of the underlying scenario. There are scenarios that are easily reproduced by the cluster distributions, few others are not. This is determined by the *choice of the ensemble* of cluster parameters that are

[7] In analytical channel models, it turned out that diversity is always overmodelled [80]. The RCM does not suffer from this effect.

[8] Since the office scenario used in the previous evaluations, TxR11, is a static one, where only people move, I decided to use another scenario, where the station moved instead.

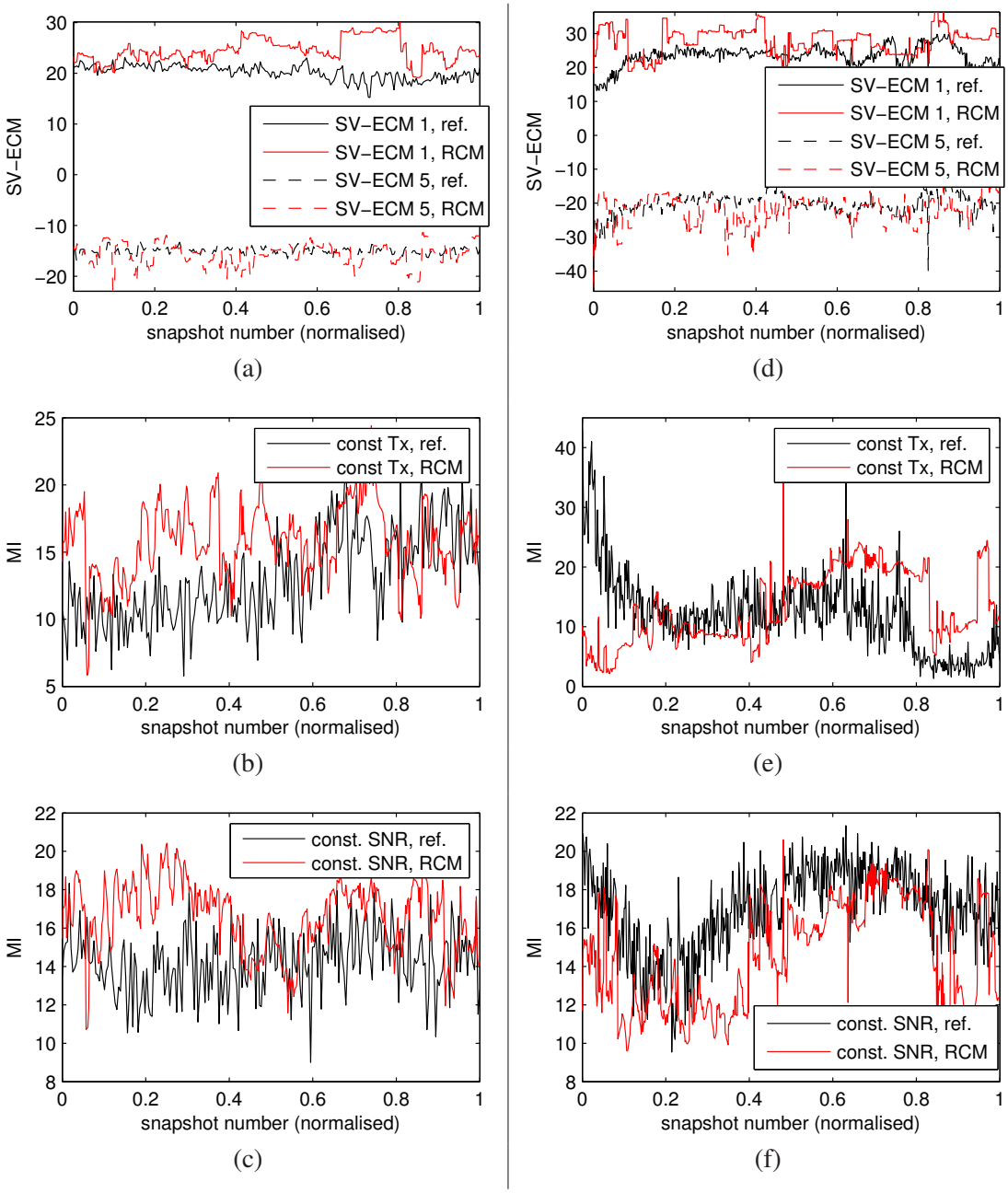


Figure 4.15: Comparing the evolution of the validation metrics over time evaluated for reference channels and modelled channels. (a)-(c): route TxR13 at 5.25 GHz (see pg. 167), (d)-(f): route TxR22 at 2.55 GHz (see pg. 176)

used to parametrise the environment pdf (in this case, a Gaussian mixture distribution). If the ensemble is able to describe the occurring clusters sufficiently, the RCM fits very well. If this is not the case, the RCM will of course show a certain mismatch.

4.2.3.3 Doppler spectrum evaluation

Unfortunately, I cannot present comparative results from the Doppler spectrum of reference and modelled channels. The estimation of the Doppler shift of the individual components was based on 2 snapshots in time (cf. Table 6.1 on pg. 101), leading to a very large estimation error. Instead, I will present the modelled Doppler spectra and compare them to well-known results in [81].

Again, I use the same time-variance parameters as before with $t_s = 10\,\text{ms}$ to be able to resolve Doppler frequencies in the range of $\pm 50\,\text{Hz}$. As the time window for evaluating the Doppler spectrum, I chose 64 samples in order to keep within the coherence time of the channel. Using these parameters, the Doppler resolution is 1.56 Hz. To see the evolution of the Doppler spectrum over time, I slid the Doppler window over all (smoothly time-variant) realisations.

To demonstrate the results I selected two measurement routes as input to the RCM. The first measurement route was recorded in an NLOS office environment (TxR8), at 2.55 GHz. Figure 4.16a presents the Doppler spectrum of the modelled channels over time. Clearly, a "horned" Doppler spectrum is visible. The spectrum matches surprisingly well to the ones presented in [81].

As second route I used the cafeteria measurements at 2.55 GHz. In this environment, we measured partly with LOS, partly with NLOS. This effect is clearly visible also in the modelled channels (see Figure 4.16b). The LOS part of the environment is correctly represented by a peaky Doppler spectrum [81]. In this case, the peaky spectrum changes smoothly to the horned spectrum, and back again.

This means, the RCM is able to model both, LOS and NLOS scenarios, as well as the transition between them, when parametrised accordingly.

4.3 Diffuse multipath parametrisation

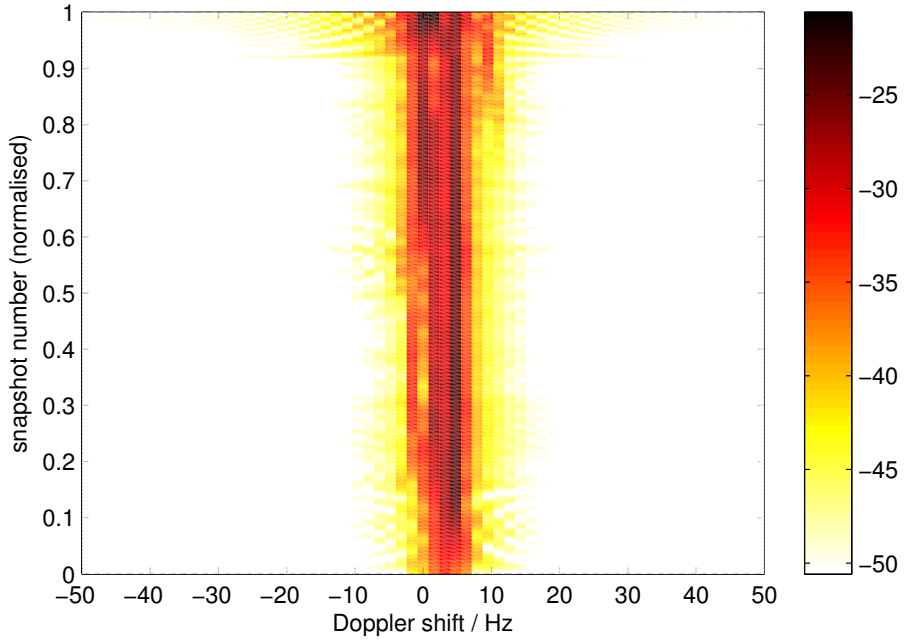
The effect of DMP in describing MIMO channel was considered only very recently by Richter [36]. He used the concept for improving the performance of a high-resolution path estimation algorithm. When considering MIMO channel models, DMP provides an efficient way to describe the diffuse part of the impulse response. Despite the need of including DMP in MIMO channel simulation, no models for the DMP parameters were available in literature.

Filling this gap, I presented the first DMP model that is parametrised from measurements[9] [60].

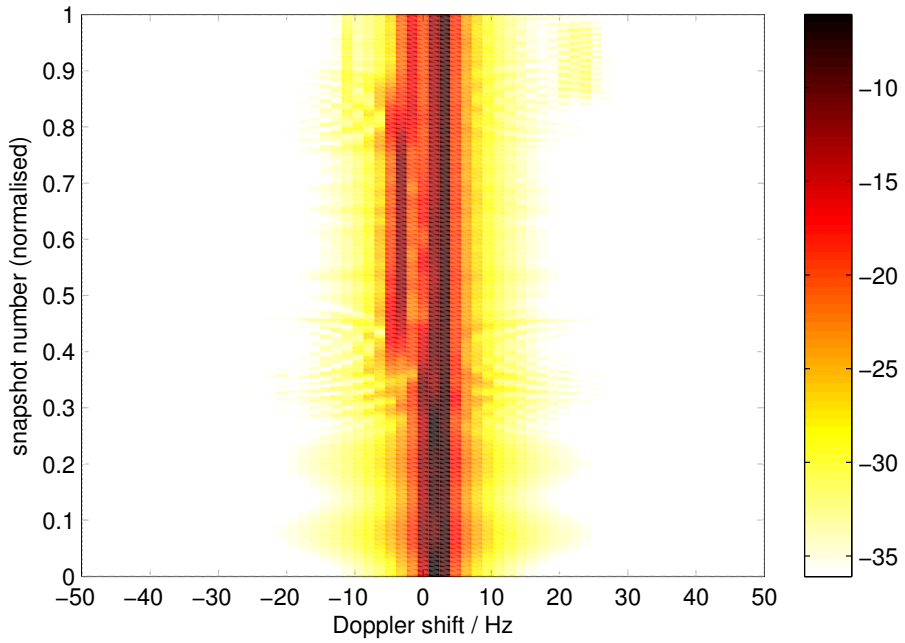
4.3.1 Estimation of DMP parameters from measurements

Salmi, Richter, and Koivunen [82] presented a high-resolution estimation algorithm that jointly estimates the discrete propagation paths and the DMP parameters. Their approach is very accurate, but also quite complex.

[9] Nearly at the same time, Giovanni Del Galdo presented a promising approach of modelling DMP in his PhD thesis [59]. However, he did not validate this model by any measurements. Also cf. the footnote on pg. 40.



(a)



(b)

Figure 4.16: Modelled doppler spectra over a sliding time window for two different parameterisations: (a) measurement TxR8 at 2.55 GHz (see pg. 162), (b) measurement TxR22 at 2.55 GHz (see pg. 176)

Since a very accurate high-resolution path estimator, the ISIS algorithm [19], was already available to me, I decided to extend this algorithm to estimate the DMP using a low-complexity approach [60].

The starting point is a measurement, where discrete propagation paths have been estimated from the impulse responses. The following procedure is performed for all snapshots, so for better readability I will omit the time index t.

From the *measured* channel matrices, $\hat{\mathbf{H}}_{\mathrm{Rx}}(\tau)$, and the channel matrix from the discrete (ISIS-)*estimated* paths, $\hat{\mathbf{H}}_{\delta}(\tau)$, the PDPs of these channels are calculated as

$$\hat{\psi}_{\mathrm{Rx}}(\tau) = \frac{1}{N_r N_t} \|\hat{\mathbf{H}}_{\mathrm{Rx}}(\tau)\|_{\mathrm{F}}^2, \tag{4.8}$$

$$\hat{\psi}_{\delta}(\tau) = \frac{1}{N_r N_t} \|\hat{\mathbf{H}}_{\delta}(\tau)\|_{\mathrm{F}}^2. \tag{4.9}$$

Then, the PDP of the *residual channel* is computed as

$$\hat{\psi}_{\mathrm{res}}(\tau) = \frac{1}{N_r N_t} \|\hat{\mathbf{H}}_{\mathrm{Rx}}(\tau) - \hat{\mathbf{H}}_{\delta}(\tau)\|_{\mathrm{F}}^2. \tag{4.10}$$

This residual PDP $\hat{\psi}_{\mathrm{res}}(\tau)$ shall now be approximated by the DMP PDP $\psi_d(\tau)$ as detailed in (2.7)–(2.9), using the DMP parameters α_d, β_d, and τ_d (see Figure 2.6).

To estimate these three parameters, I use a non-linear least-squares estimator [83] as

$$\{\hat{\alpha_d}, \hat{\beta}_d, \hat{\tau_d}\} = \underset{\{\alpha_d, \beta_d, \tau_d\}}{\arg\min} \sum_{\tau} |\hat{\psi}_{\mathrm{res}}(\tau) - \psi_d(\tau, \{\alpha_d, \beta_d, \tau_d\})|^2. \tag{4.11}$$

The DMP PDP can be easily calculated by first calculating $\mathbf{R}_d(\{\alpha_d, \beta_d, \tau_d\})$ as in (2.9), then using the Fourier transform as

$$\boldsymbol{\psi}_d(\{\alpha_d, \beta_d, \tau_d\}) = \mathrm{diag}[\mathcal{F}^{-1}\mathbf{R}_d(\{\alpha_d, \beta_d, \tau_d\})\mathcal{F}], \tag{4.12}$$

where $\boldsymbol{\psi}_d = [\psi_d(\tau = 0)...\psi_d(\tau = M/B)]^{\mathrm{T}}$, and $\mathcal{F}$ denoting the Fourier transform matrix.

In this way, I estimate of the DMP parameters for each snapshot of the channel.

4.3.2 Modelling DMP parameters

The model for the DMP parameters is observation driven. The following subsection will reveal a strong correlation between certain parameters of the DMP PDP and of the discrete PDP. The subsequent subsection will present a model based on these observations.

4.3.2.1 Observation

For the evaluations in this section I selected five scenarios with quite different propagation conditions. Three measurements were done in office rooms (TxR8, TxR9, and TxR11, see pages 162, 163, and 165, respectively), and two measurements were conducted in big rooms, with and without line of sight (TxR21 and TxR23, see pages 175 and 176, respectively). In all measurements

the Tx was moved while the Rx was fixed, except for TxR11 environment, through which people moved occasionally.

The measurements were all post-processed using the ISIS SAGE algorithm [84], with the parametrisation as in Table 6.1, in order to estimate the discrete propagation paths for every snapshot. Subsequently, the residual PDP was calculated as in (4.10) from which we estimated the DMP parameters by using (4.11).

To find a model for the DMP parameters, I related the following observable quantities from the PDP of the discrete components $\psi_\delta(\tau)$ to the diffuse-multipath parameters: (i) the base delay of the peak $\check{\tau}_p = \arg\max_\tau\{\psi_p(\tau)\}$, (ii) the peak power $\check{\psi}_p = \max\{\psi_p(\tau)\}$, and (iii) the rms delay spread σ_τ [49] of the total impulse response.

These parameters, related to the diffuse PDP, are shown in Figure 4.17. Scatter plots show the dependency of the DMP parameters to the quantities observed from the PDP of discrete components. Single snapshots are indicated by dots, where different colors indicate the different environments measured. In detail, the base delay of the diffuse components τ_d is correlated with the peak delay $\check{\tau}_p$ (Figure 4.17a). The diffuse-multipath peak parameter α_d is likewise correlated with the peak power $\check{\psi}_p$ (Figure 4.17b).

Finally, we use the ratio between the total diffuse power and the total Rx power, r_d, and relate it to the rms delay spread of the measured channels. Figure 4.17c reveals a positive correlation between these two quantities, not as good as for the two other parameters, but still good enough.

Immediately, a conclusion to be drawn from Figure 4.17 is that the correlations hold obviously for all environments investigated. Also, within a certain environment, represented by different dots of the same color, the correlations exist over large parameter ranges.

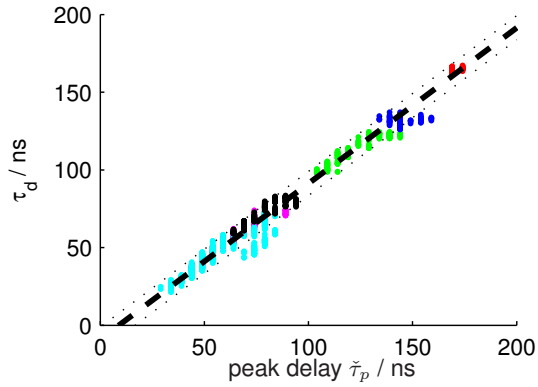
This conclusion might be surprising, but can be easily explained as follows:

The *base delay* of the DMP τ_d starts right with the first peak of the discrete paths, for two physical reasons: (i) local scattering around the terminal from scatterers that are within one delay bin (in our case corresponding to 3 m), (ii) scatterers that are close to the (quasi) line of sight. Both phenomena are very likely in indoor scenarios. The reason for some samples being below the regression line is that in some NLOS scenarios the strongest peak of the PDP comes later than the first contributions, i.e. the first *observable* discrete path has a larger delay than the first contribution to DMP.

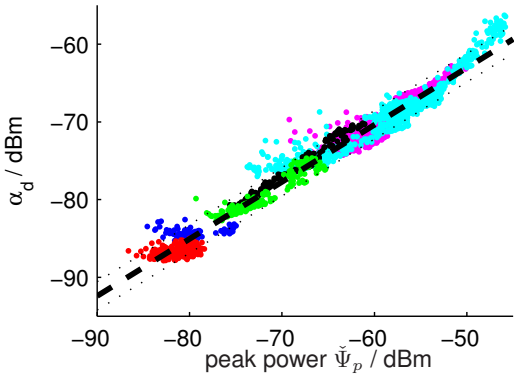
The *diffuse peak power* α_d must be smaller than the discrete peak power, otherwise the path estimator would not be able to estimate any paths. But, when there is more power, hence better SNR, at the receiver, more dominant paths can be resolved, calling for a larger gap between the diffuse and discrete peak power.

Finally, the *power ratio of the DMP and the received power* is likely to depend on the rms delay spread of the total response for the same reasons as stated above. The longer the impulse response is, the stronger the contribution of the diffuse power will be. Note that the rms delay spread of the discrete components cannot be used here, they usually do not represent the delay spread of the actual impulse response of the channel correctly. This is due to the properties of the high resolution parameter estimator that identifies stronger, but not the weaker ones, which significantly contribute to the delay spread.

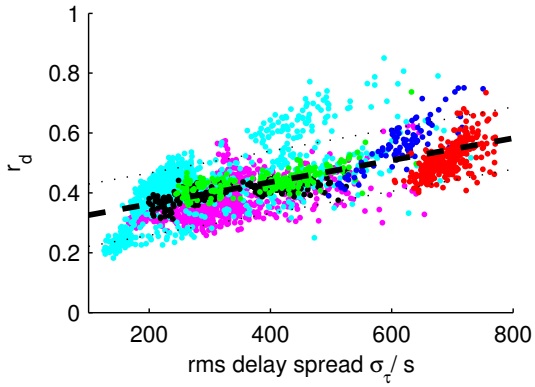
A more physical approach would be to correlate the diffuse peak power with the distance between Rx and Tx, where a correlation is very likely. But since we do not have the actual distances



(a)



(b)



(c)

Figure 4.17: Dependency of diffuse-multipath parameters. The black dashed line describes the linear regression curve, the dotted lines indicate the 3σ interval.

available, we resort to the correlation with the discrete peak power. The physical distance between Tx and Rx may also have significant correlation with the diffuse power ratio.

4.3.2.2 Model of diffuse-multipath parameters

We model the DMP parameters in the measured environments by Gaussian probability distributions, where we define the mean and variance using a first-order polynomial fit (corresponding to linear regression) of the observation data

$$p(\tau_d|\check{\tau}_p) = \mathcal{N}(\check{\tau}_p - 9\,\mathrm{ns}, 27.0\,\mathrm{ns}^2), \tag{4.13}$$

$$p(\alpha_d[\mathrm{dBm}]|\check{\psi}_p) = \mathcal{N}(0.73 \cdot \check{\psi}_p - 26.3\,\mathrm{dBm}, 1.49\,\mathrm{dBm}^2) \tag{4.14}$$

$$p(r_d|\sigma_\tau[\mu s]) = \mathcal{N}(0.037 \cdot \sigma_\tau + 0.29, 0.0088) \tag{4.15}$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a realisation of a Gaussian random variable with mean μ and variance σ^2.

Realisations of the diffuse-multipath parameters can be obtained in the following way: By drawing realisations from the distributions (4.13) and (4.14), the base delay τ_d, and the peak power α_d are determined directly. The decay factor β_d is then determined by drawing a realisation of the diffuse power ratio from (4.15), obtaining the total diffuse power from (2.12), and calculating the decay factor by (2.11).

4.3.2.3 DMP model validation

I validated the DMP model using system model III (see Table 4.1) and will present the validation for the stationary office environment (NLOS) and for the Cafeteria environment (LOS). Note that the RCM *parametric model* was parametrised by the measurements in the respective environments, but for the DMP model we used the parameters presented in Section 2.4.2, obtained from a much larger set of measured data.

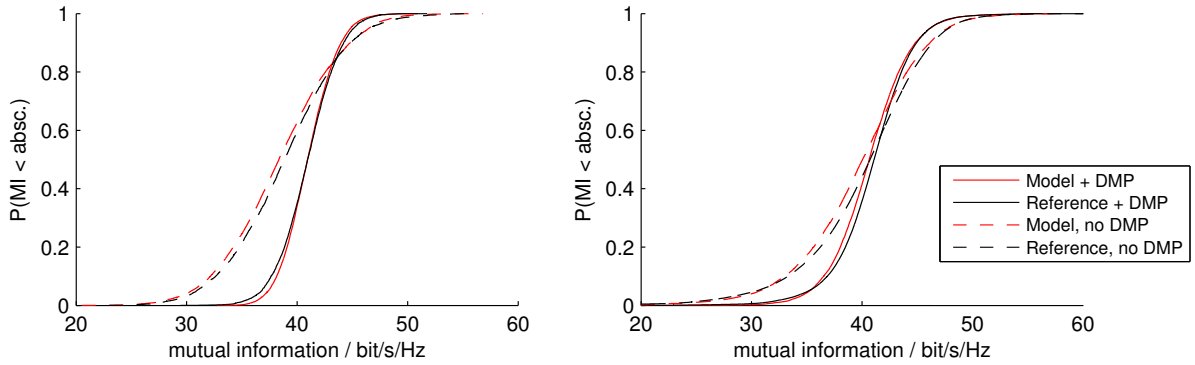
To assess the model fit to the reference channels I used the narrowband MI metric with 10 dB SNR (cf. Section 3.2.1), and the diversity metric (cf. Section 3.2.2). Since the ECM is directly applied to discrete paths, it cannot be used to validate the DMP model.

Figure 4.18 compares the MI metric for both scenarios and shows the impact of including the DMP in the model. Red colour indicate the modelled channels, blue colour indicate the reference channels, while solid lines denote the inclusion of the DMP, and dashed lines denote the disregard of DMP.

Figure 4.18a shows the MI evaluated for the Stationary Office environment. First, there is a significant difference between including the DMP in both the reference channels and in the model. The median MI changes considerably. We also find that the RCM fits the reference channels very well[10]. A similar behaviour is noticed in the cafeteria scenario in Figure 4.18b, with a significant change in outage capacity, while the ergodic capacity does not change too much, here.

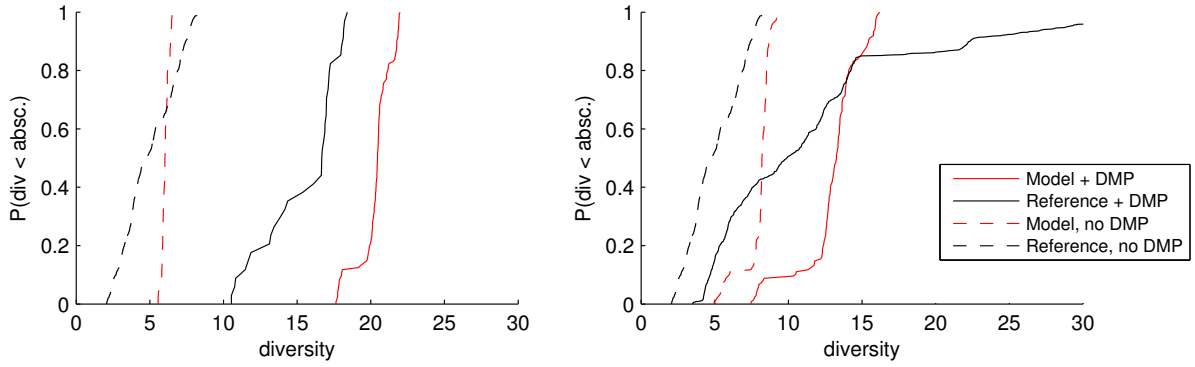
Since the MI curves indicated a strong difference in diversity, I also used the diversity metric (cf. Section 3.2.2) for comparing the channels. Again I assessed the impact of including the DMP in

[10]Note that for the DMP validation a more challenging system model (8×8 antennas, 100 MHz bandwidth) was used.



(a) Office scenario (TxR10) (b) Cafeteria scenario (TxR22)

Figure 4.18: Validating the DMP model using the mutual information metric



(a) Office scenario (TxR10) (b) Cafeteria scenario (TxR22)

Figure 4.19: Validating the DMP model using the diversity measure

the model. The slope of the MI cdf indicates a strongly increase in spatial diversity, the steeper the slope, the more diversity exists in the channel. Figure 4.19a shows that neglecting DMP has a significant impact on diversity. In this scenario, the model creates slightly more diversity than available in the channel. The same holds true for the Cafeteria scenario in Figure 4.19b.

Since this is the very first approach to model DMP, a validation by other groups with different measurements is necessary to adjust the empirical parameters. Should one be interested in best reflecting the DMP in a scenario, the modelling approach described above can be applied, using an individual measurement run to estimate the DMP parameters best fitting the considered environment.

5 Conclusions

5.1 Comparison to related MIMO channel models

I will compare the RCM to the following related models: the 3GPP SCM, the WINNER II interim channel model, the 802.11 TGn model, and the COST 273 MIMO channel model. An overview of selected features of these three models can be found in the Introduction in Section 1.5. The features I will focus on in this comparison are: (i) cluster generation, (ii) path loss calculation, (iii) smooth time-variance, (iv) model validation.

Cluster generation

The cluster generation is the most significant difference in all the models. While the 3GPP SCM and the WINNER II model consider clusters that all show the same spatial extents (cluster angular spreads) and no dispersion in delay, both the COST 273 model and the RCM use different cluster sizes within one scenario for modelling the radio channel.

The 3GPP SCM and the WINNER II model place the clusters in order to fit the specified global *marginal* azimuth power spectrum (APS) at the Tx and Rx. This would imply that the double-directional APS is separable[1] by definition, which would not be very physical. To actually obtain a separable APS using the WINNER II model, the modelled APS needs to be approximated very well by the placed clusters, i.e. many clusters are necessary to approximate a separable APS.

However, when only few clusters are placed according to the marginal APS, the resulting double-directional APS is not separable, intrinsically. The number of clusters in the WINNER II model depends on the chosen environment, ranging from 5 to 20 clusters, only. Additionally, clusters with larger delay show fewer powers. By using this concept, the WINNER II model ensures that the resulting APS is *not separable* and ensures a realistic modelling of the propagation phenomena.

A shortcoming of the WINNER II model is that the marginal APS are defined as "wrapped Gaussian distributions", which are not observed in indoor WLAN scenarios [85], thus showing a significant modelling error in these environments.

The COST 273 model knows three different kinds of clusters. In my opinion, the twin-cluster approach can be parametrised for all needs, thus the other cluster types are redundant. Unfortunately, the COST 273 model does not provide a clear algorithm on how or where to place the twin-clusters, but it seems that it suffers from the same problem of a possibly separable APS as the 3GPP SCM model does, when the number of clusters gets large.

The 802.11 TGn model parametrises the cluster spread and cluster position for a number of scenarios, but generates every tap using the Kronecker model, which again leads to a completely separable APS.

[1] This assumption is also known as the "Kronecker assumption" from analytical channel models.

To my knowledge, the RCM is currently the only model that manages to place many clusters in a way resulting in a non-separable APS[2]. This is due to the *correlated cluster parameters*, and the *Gaussian mixture parametrisation*. However, the employment of this features requires an accurate parametrisation of the environment by measurements. By this the model will match well the measurement, but not necessarily generally reflect the whole environment.

Path loss calculation

The other related models employ either the COST 259 path loss models, while the WINNER model parametrised its own models. In contrast, the RCM does not provide any path loss calculation.

This is because of the different focus of the models. The RCM strongly relies on a measurement-based model parametrisation. In this respect, the RCM will reproduce the path loss automatically well, when properly parametrised.

Since the other related models concentrate on a general description of the parametrised environments, a path loss model is obligatory for these.

Smooth time variance

With respect to modelling smoothly time-variant channels, the three related models perform quite differently.

The 3GPP SCM only models short time intervals ("drops") such that the global characteristics of the channel do not change, and considers time variation by Doppler shifts of paths to obtain fading realisations.

The WINNER II model improved this concept by correlating the global parameters between the drops to obtain coherent time variance by statistical modelling. Still, the channels do not adjust smoothly from one drop to the next one.

The COST 273 model considers smoothly time-variant channels by building an environment in a coordinate system. Clusters fade in and out depending on visibility regions, and move according to the stations movement.

The RCM combines the concept of statistically placing clusters in the azimuth/delay domain and generating smoothly time-variant channels by moving clusters. Since visibility regions are only useful in a Cartesian geometry, the RCM applies the concept of a birth/death process for generating new clusters (or letting old clusters die).

Model validation

> *"Model validation is so important because how can one otherwise be sure that one is modelling what one wants to model?"* [46, Chap. 6.7].

In my opinion it became bad practise to not validate models thoroughly. To my knowledge, neither the 3GPP SCM nor the WINNER models were validated sufficiently. There is only one publications on the WINNER I model [69] showing that there is a basic match to measurements judging

[2]Note that the RCM is anyway able to produce Kronecker-like channels, when this is observed in the measurements that are used for parametrisation.

from mutual information and Demmel condition number, but the channel diversity does not fit at all. The validation of the WINNER II model was foreseen to be done in the third part of the WINNER project, which unfortunately will not take place[3]. This decision showed that the focus seems to be on creating a new model, but there shall not be time and money to validate it accordingly.

There is no reference implementation for the COST 273 model, except for a version presented in [45], which is not publicly available. Hence, there can of course be no meaningful model validation by independent research groups, yet.

In this thesis I tried to motivate the imperative call for validating MIMO channel models with respect to both the underlying propagation environment as well as to the generated channel matrices. I do not want to claim that the validation metrics presented in this thesis provide “the ultimate truth”. Instead, they just seem to reflect the propagation environment well.

- *Mutual Information* is most often used as validation metric throughout many publications. I showed in Section 3.2.1 that neither instantaneous MI, nor the MI cdf *reflect the multipath structure of the underlying channel*. Fading influences MI too much.
- The *Diversity metric* is applied to the correlation matrix of the channels, so a stationary subset of channels needs to be selected to evaluate this metric. For this reason, it is only useful when comparing smooth time-variant channels. There it reflects diversity well.
- The *Demmel condition number* seems to fit always well, thus it is not very specific to the model match. Hence, it is no good metric for reflecting the multipath structure of the channel.
- The *Environment Characterisation Metric* is directly applied to both the modelled and to the estimated propagation paths. It does not directly consider channel matrices. For this reason, it is very specific to the multipath structure of the channel and can be well used as validation metric in this respect.

More validation metrics are necessary that focus on the significant *aspects of the underlying system* in focus. Such metrics need to be specific to the signal processing algorithms and should provide insight in how the algorithms are affected by different types of channels.

Also more validation metrics for time-variant channel models are badly needed in the scientific community. The new COST 2100 Action might be able to concentrate some research action in this field.

I validated the RCM using the metrics focussing on the underlying propagation environment (in contrast to focussing on an underlying communication system). Judging from these metrics, the RCM is quite able to reproduce many different measured environments well. The multipath structure is obviously reflected well. But the RCM is neither the ultimate truth, some environments could not be reflected as well as one could wish for. In these cases, the model reaches its intrinsic limits that are set by the structure of the parametric model, as discussed in the next section.

5.2 Comparison to “Playback Simulations”

The RCM is designed to closely resemble measured channels. The straight-forward approach of *Playback Simulations* [86] would also lead to similar results. For this reason I want to provide a

[3] Which is my state of knowledge in October 2007. Maybe there will be an extension later?

comparison between the RCM and the concept of Playback Simulations.

In Playback Simulations, a channel sounder is used to record the (typically time-variant) impulse response of the channel. Subsequently, the noise floor is removed from the impulse responses in a post-processing step. This is done by first estimating the noise floor from the tail of the impulse responses, and introducing the *"noise cut"* by adding typically 6 dB to the noise floor. Then, all samples of the impulse response smaller than the noise cut are set to zero.

These post-processed impulse responses are subsequently used for link-level simulations reflecting the measured channels perfectly. Note that measurement errors will have a significant impact, so planning and conducting the measurements need to be done with great care.

The concept of Playback Simulations has following disadvantages:

- The system model for the simulations is fixed by the measurement equipment.

  This also implies that the antenna configuration of the system to be simulated must already be used during the measurements.

- Only a limited number of snapshots can be recorded.

  Due to limitations of harddrive space and memory only a finite number of impulse responses can be stored. When doing bit error ratio simulations, a large number of channel realisations is necessary to obtain significant results for low bit error ratios. Note that replaying the recorded data does not enlarge the statistical ensemble.

The RCM provides a solution to both problems stated above. Since it uses clusters of propagation paths to model the channel, any antenna structure can be used for link-level simulations. Using the smoothly time-variant mode of the RCM, an infinite number of correlated channel impulse responses can be generated leading to a significant statistical ensemble.

5.3 Possible improvements of the RCM

Including path loss and global parameter correlations

When intending the RCM to become more general, a path loss model and correlation between the global parameters (e.g. mean delay, global rms delay spread, global rms azimuth spread) is imperative. The WINNER II model and the COST 273 model provide good guidelines on how to implement such features.

Polarisation

The use of multiple polarisation promises to increase the MIMO capacity while having very small antenna spacings [87]. Again, the underlying radio propagation determines how large the performance increase can get. Most significantly, the two polarisation directions suffer different path losses, while depolarisation (caused by physical scattering) leads to a performance reduction of MIMO systems [88, 89].

For these reasons accurate polarisation models are necessary. Luckily, the cluster concept enables one to model polarisation in a straight-forward way. The cluster parameters just need to be extended by the individual path loss of the vertical and horizontal polarisation, as well as by a depolarisation coefficient. Again, these values need to be parametrised accurately to reflect the underlying propagation environment.

Elevation

Especially in indoor scenarios, propagation paths may come with a significant elevation because of reflections from floors or ceilings. While the model description includes this property, the current *implementation* of the model does not account for elevation, yet. Again, the inclusion is straight forward: the mean cluster position parameters and cluster spread parameters need to be extended accordingly. Also, proper parametrisation is paramount.

Joint estimation of propagation paths and diffuse multipath

In the current state, the estimation of the DMP was based on the remainder of the impulse response. The quality of the estimation can be significantly improved when jointly estimating the discrete parameters and the DMP parameters as in [82], even with lower complexity and increased estimation performance.

Improving cluster correlations

Currently, the cluster parameters are correlated within one cluster with each other and correlated with the number of clusters and the total power of a snapshot.

The RCM would fit even better when including correlations between the parameters of *different* clusters. In this way, when creating a new cluster, its parameters depend on the already existing clusters. However one should note that this approach becomes even more focused on reproducing already measured channels rather than modelling channels of the kind of the measured one.

LOS handling

Another possible improvement is to model the LOS path *separately* from all other paths. This idea would lead to several modifications of the model.

For parametrisation, the LOS path needs to be separately identified from the measurements and must be excluded from the cluster parameter pdf. In the modelling procedure, first the NLOS part of the channel is generated, then the LOS path is considered separately, where one can individually choose the Ricean K-factor. This approach needs to include either a certain geometry of the channel (which runs against the concept of the RCM), or parametrise the LOS path by statistical properties.

Ultimate limits of the RCM model structure

The concept of the RCM is to create clusters according to an underlying cluster parameter pdf. One feature of this approach is that the smooth time variations of the channel correspond to a random walk in the scenario.

However, should one be interested in moving a straight line, a underlying geometry is necessary, which the RCM does not support. This may become interesting mostly in highly mobile scenarios such as in car-to-infrastructure and in car-to-car scenarios [90, 91].

Part II

Cluster identification from measurements

"Real-world clusters come in strange shapes, or have no simple shape at all. The nature appears to have no moral qualms on producing non-spherically or nonellipsoidally structured groups of multipaths. While such non-regular cluster shapes are relatively easily recognized by human eye, they are difficult to identify automatically by mathematical algorithms."

Jari Salo

6 Clustering methods

In this chapter I will describe different methods to identify multipath clusters from measurements. I will no longer discuss the existence of clusters, here. A discussion about this profound question is provided in the introduction (see Section 1.3).

Finding clusters is a bold venture. Referring to the theme of this part, clusters do come in strange shapes. They are usually not elliptically (let alone, spherically) shaped, and do not show any moral qualms of overlapping with each other.

The following sections will provide an overview of clustering approaches, both visual and automatic. Each of these approaches makes a certain assumption on the structure of the clusters, which leads to a number of *different cluster definitions*. This, of course, complicates the matter of finding some kind of optimal partitioning into clusters.

In the first stage, let us use a general definition of clustering [28]:

> *"The process of grouping a set of physical or abstract objects into classes of similar objects is called clustering. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. A cluster of data objects can be treated collectively as one group in many applications."*

From this very general definition, a number of definitions can be derived:

- Cluster definition for visual identification purposes
  Clusters are defined such that visual identification becomes less subjective. Section 6.1 will provide further discussion on this topic.
- Cluster definition by scattering objects
  A possible definition of this kind for clustering multipath components might be "A cluster is uniquely defined by a real-world scattering object". This definition is very valuable from a modelling perspective[1], but it complicated the discussion about clusters even more. For this reason, it is very important to distinguish this definition from the next one.
- Cluster definition by statistical means
  Another interesting model-based definition of a cluster is "A cluster consists of a group of multipath components, where the parameters of the components within the cluster have the *same* distribution". This definition provides a concept of how to place paths within a cluster. Note that this definition is used in the RCM (cf. Section 2.3.2.1).
- Cluster definition by clustering algorithms
  Automatic clustering algorithms intrinsically introduce their own definition of a cluster. I will present two of these in the following sections. When using automatic clustering algorithms to parametrise cluster-based model, it is important to validate whether the underlying

[1] An implementation of this concept is provided by the IlmProp channel modelling tool [59].

cluster definition of the clustering algorithm is compatible with the underlying cluster definition of the model. Section 6.4.2 will show that this is the case for the RCM.

- Cluster definition by fading characteristic
  Another kind of cluster definition, which is quite agreed, is that MPCs within a cluster should ideally show a correlated large-scale fading behaviour, while the large-scale fading of MPCs in different clusters should be uncorrelated. To evaluate this property, the high-resolution path estimator used for post-processing the measured impulse responses must be able to track MPCs over multiple snapshots with high estimation accuracy. Unfortunately, the high-resolution estimator used in this work did not have this capability, so I could not evaluate clusters in this respect.

This chapter is organised as follows. First, I will introduce an overview of visual clustering (Section 6.1) and semi-automatic clustering (Section 6.2). Then, I will present the problems, pitfalls, and solutions for automatic multipath clustering in Section 6.3. For the even more complex problem of joint clustering and tracking, I will spend the whole subsequent chapter.

6.1 Visual clustering

The first ones who introduced the concept of clusters to radio channel modelling were Saleh and Valezuela in [22]. Clusters were used to describe groups of paths in the delay domain. The latest standardised UWB channel models still use a similar concept with adapted parameters [92][2].

The COST 259 radio channel model [93, 94] was the first to introduce the clustering concept in the azimuth/delay-domain. Laurila et al. were the first to identify clusters in this 2-D domain in [95]. Spencer et al. used an improved method to model time and angle-of-arrival statistics of the radio channel [96]. Chong et al. extended this concept and presented a statistical wideband channel model in [97], that already took the spatial property of the channel into account, but only in a single-directional way. In the need for parameters for the upcoming 802.11 TGn channel model, Yu et al. identified *cluster angular spreads* of indoor WLAN channels in [77]. They used a heuristic visual-clustering approach. All these publications identified clusters in the delay/AoA domain.

I was the first to identify clusters in the AoA/AoD domain in [11]. Clusters can be much better resolved in this domain in indoor scenarios, since the impulse response is quite concentrated there and clusters can not be resolved well in delay. To identify clusters, we used the double-directional azimuth power spectrum jointly with SAGE estimates of the propagation paths with following rules:

- Each “cluster” is defined as a group of MPCs showing similar AoA and AoD.
- In a scatter plot of (AoA,AoD) estimates, clusters show dense estimated MPCs with similar powers, where the powers of the MPCs decrease from the cluster’s centre to the outskirts.
- In the double-directional azimuth power spectrum the cluster power distribution must also decrease from the centre to the outskirts.
- Clusters must not overlap.

[2] Since the 802.15.3a task group that developed this model has been dissolved, this citation seems most appropriate.

Using these rules we minimised the effects of human subjectivity.

The shortcomings of visual clustering get obvious when considering more than two dimensions, jointly. Three dimensions are still somewhat traceable, four dimensions are a mess. Also, when needing to cluster a large amount of data sets, visual clustering is utmost cumbersome.

For these reasons I concentrated my research on automatic clustering algorithms.

6.2 Semi-automatic clustering

The first semi-automatic clustering algorithm for tracking clusters in MIMO channel data was introduced by Salo [98]. A hierarchical clustering algorithm (cf. Section 6.3.4) was used, where clusters were identified in each domain sequentially and had to be joined visually, subsequently. The algorithm did not yet allow for *joint* clustering, i.e. clustering all all domains, jointly.

But, this algorithm provided a perfect basis for going forward to completely automatic clustering algorithms.

6.3 Automatic clustering algorithms

Automatic clustering is quite difficult, since (i) clusters come in strange shapes (or have no shape at all), (ii) there's no genie telling the number of clusters.

This section discusses the problem of clustering while assuming that the number of clusters is known[3]. Subsection 6.3.1 will define the problem, which is of course the same for all clustering approaches. All the clustering algorithms are applied to estimated propagation paths. The pitfalls and consequences of this path estimation is outlined in Section 6.3.2. Only when clustering jointly, automatic algorithms provide significant clustering results. To enable joint clustering, data needs to be scaled. For this, I use the Multipath Component Distance, which is presented in Section 6.3.3. Two different clustering algorithms, the hierarchical tree clustering, and KPowerMeans clustering, are presented in Sections 6.3.4 and 6.3.5, respectively. Section 6.3.6 compares these two approaches and discusses the quite significant differences[4].

6.3.1 Problem description

The starting point is a MIMO channel measurement. The measurements provide numerous snapshots of the impulse response of the – typically time-varying – radio channel. These measurements are fed to a high-resolution algorithm, e.g. the ISIS estimator [19], or ESPRIT [99] and its improved unitary variant [100], to estimate the channel parameters for each snapshot individually. It has been found in several MIMO studies that these parameters tend to appear in clusters, i.e.

[3] Sometimes clustering using this assumption is called "Genie-based clustering". The question is, when having a genie who could already tell how many clusters there are, the genie of course also knows the optimum clustering. So, why investigate automatic clustering? The answer is that I will dismiss the genie in Section 6.4.

[4] Indeed, it was Jari Salo who drew my attention to the field of automatic clustering. Together with him and my first Master student, Pierluigi Cera, I developed the automatic clustering algorithms presented in the next sections. I want to explicitly appreciate their hard work and significant contributions by writing the next subsections in the plural form.

in groups of multipath components (MPCs) with similar parameters, e.g. [77, 97, 96, 95]. The problem is to find an automatic procedure to identify and track these clusters.

We consider one data window with a number of L MPCs, where every single MPC is represented by its power, P_l, $l = 1 \dots L$, and a parameter vector, $\mathbf{x}_l$, containing the delay, τ, azimuth and elevation AoA, φ_{AoA}, θ_{AoA}, and azimuth and elevation AoD, φ_{AoD}, θ_{AoD}. The data for all paths are collected in the vector $\mathbf{P} = [P_1 \dots P_L]^T$ and the matrix $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_L]^T$.

6.3.2 Parameter estimation: garbage in, garbage out

Before going into detail with the clustering algorithms I want to issue a word of caution.

The input data of the clustering algorithms are *estimated* parameters of the propagation paths. High-resolution estimators such as the used ISIS algorithm [19] (but also others, like RIMAX [36], or an extended-Kalman-filter approach [82]), rely on an underlying signal model. The algorithms then try to estimate MPC parameters such that the reproduced signal fits best the recorded impulse response.

Already a small imprecision in the underlying signal model may lead to completely wrong estimation results.

Following parameters of the signal model have to be defined *most accurately*:

- Antenna array responses:
  The antenna array responses are used to estimate the directions of the individual paths. Should the array response be not available, direction estimations *can not* be trusted. Even more, following properties of the every individual element of the antenna arrays *have to be* calibrated accurately:
  - Fully-spherical complex antenna diagram (azimuth and elevation)
  - Fully-polarimetric antenna response (the response of the antenna when being individually excited by vertically and horizontally polarised waves.

  In [31], Landmann impressively demonstrates the effects of disregarding any of these calibrations in the signal model. His thesis [101] will provide further insight in this topic.
- System response:
  The impulse response of the measurement system already introduces a certain filter of the signal. In order to avoid artefacts from the measurement system, the response has to be included when estimating MPC parameters.

Let me repeat: The estimation results base on the calibration data, completely! Expressed in other words[5], "Garbage in, garbage out".

But also when ensuring perfect calibration, the results have to be considered with care. I observed following effects of performing the ISIS algorithm on simulated channel data without noise. Even though I only used the ISIS algorithm, I conjecture a same behaviour of all the other high-resolution estimators that work on a snapshot-by-snapshot basis.

[5] Actually, in the words of Wim Kotterman, a fellow researcher from TU Ilmenau to whom I'm very grateful for discussions on this very topic. I am also very grateful to Xuefeng Yin who managed to explain the concepts of SAGE estimation to me, very patiently.

- When estimating clustered paths, the exact parameters of the paths are *not correctly estimated*. Rather, the mean position of the cluster is represented very well, while the estimated spread varied around 50% of the true value.
- When a station moves towards clustered propagation paths, the estimates of the paths are not tracked, but a completely new group of paths is estimated in each snapshot. (Note that again the mean cluster position is accurate, but the spread may vary).
- When considering an inaccurate antenna response, the spread of one cluster may have a constant bias.

After all these concerns it may seem rather stupid to rely on estimated MPCs. However, the used algorithm is one of the most accurate parameter estimators available. Since my main focus is to identify *clusters*, and not individual propagation paths, it perfectly suits the needs. Also the variation on the cluster spread can be easily compensated by tracking clusters and using the median cluster spread value.

Nevertheless, an accurate calibration of the environment and the inclusion of this calibration in the signal model is vital for obtaining trustworthy estimation results.

Estimation parameters used for the evaluations

I used the following parameters to estimate discrete paths from the measurements (see Table 6.1):

Parameter	**Value**
Doppler-block size	2
Number of paths	100
Maximum dynamic range	40 dB
Noise threshold	6 dB
Estimate AoA	azimuth
Estimate AoD	azimuth and elevation
Estimate polarisations	full polarimetric coupling
Number of iterations	5

Table 6.1: High-resolution path estimation parameters

I combined 2 measured snapshots in a "Doppler block"[6]. From each of these groups of 2 snapshots, I estimated at most 100 paths, provided that these paths keep within the SNR thresholds. The power of the weakest path must not be more than 40 dB below the power of the strongest path[7] ("maximum dynamic range"). Moreover, it must be at least 6 dB stronger than the power of the noise floor ("noise threshold"). Note that this threshold determines the minimum power of

[6]The Doppler block determines the number of subsequently measured impulse responses from which the path-wise Doppler shift is estimated. When just using 2 snapshots, the Doppler estimation error will intrinsically be quite large, however by combining two snapshots, the estimation accuracy of the other parameters can be increased, moreover one can combat the effects of phase noise [61, 62]

[7]These 40 dB dynamic range are measured in the impulse response. Note that this value does not correspond to the actual SNR in the frequency domain. For typically exponentially decaying impulse responses, the power is concentrated at small delays, while at large delays only the noise floor remains. This brings a large peak-to-noise ratio, even when the SNR is quite low.

the propagation paths. The *"cluster power threshold"* to be introduced later in Section 6.4.3 will subsequently use a similar concept to determine the number of clusters.

I used the full spherical, full polarimetric antenna diagrams for the path estimation. Always, the full polarimetric coupling (VV, HH, HV, VH) was estimated for each path. At the Tx we could estimate the direction of each path in azimuth and elevation. At the Rx, we had to skip elevation estimation for reasons to become clear in Section 9.2.3. The number of iterations, determining the accuracy of the estimation result, was set to 5, which yields a reasonably good estimation performance.

6.3.3 Multipath Component Distance — enabling joint clustering

To enable joint clustering, we use the *multipath component distance* (MCD) [72] to scale the data.

The MCD allows to combine parameters that come in different units. This measure was first introduced in [72] as an intermediate measure on the way to quantify the complete multipath separation of the radio channel. Here we use the MCD to quantify the distance of two points in the parameter space.

For angular data it is given as

$$\mathrm{MCD}_{\mathrm{AoA/AoD},ij} = \frac{1}{2}\left| \begin{pmatrix} \sin(\theta_i)\cos(\varphi_i) \\ \sin(\theta_i)\sin(\varphi_i) \\ \cos(\theta_i) \end{pmatrix} - \begin{pmatrix} \sin(\theta_j)\cos(\varphi_j) \\ \sin(\theta_j)\sin(\varphi_j) \\ \cos(\theta_j) \end{pmatrix} \right|, \tag{6.1}$$

for AoA and AoD likewise, but separately. The scaling of $\frac{1}{2}$ normalises the maximum distance to 1.

The delay distance is obtained as

$$\mathrm{MCD}_{\tau,ij} = \zeta \cdot \frac{|\tau_i - \tau_j|}{\Delta\tau_{\max}} \cdot \frac{\tau_{\mathrm{std}}}{\Delta\tau_{\max}}, \tag{6.2}$$

with $\Delta\tau_{\max} = \max_{i,j}\{|\tau_i - \tau_j|\}$, τ_{std} being the standard deviation of the delays, and ζ being an suitable delay scaling factor to give the delay more "importance" when necessary. This has advantageous effects when clustering real-world data [102]. If not indicated otherwise, we chose[8] $\zeta = 5$. Also, in addition to the previous definition [72], we scale the delay distance with the normalized delay spread. Should the path delays be all quite similar, the delay spread will be quite small. In contrast, should the path delays be very different, e.g. occur in two different layers, the delay spread will be larger. Using the delay spread scaling we can account for these effects efficiently.

The resulting distance measure is given as

$$\mathrm{MCD}_{ij} = \sqrt{\|\mathrm{MCD}_{\mathrm{AoA},ij}\|^2 + \|\mathrm{MCD}_{\mathrm{AoD},ij}\|^2 + \mathrm{MCD}_{\tau,ij}^2}, \tag{6.3}$$

which can be interpreted as the radius of a (hyper-)sphere in the normalized multipath parameter distance space. We use this distance for joint clustering.

[8] This value was empirically chosen since it turned out to provide reasonable clustering results.

6.3.4 Hierarchical clustering algorithms

Hierarchical tree clustering [28], or agglomerative clustering, is an algorithm that constructs a binary cluster tree from the input data. The lengths of the branches of the tree quantify the distances between the nodes. Each node in the tree represents a group of data points. The leaf nodes are just the data points themselves, and the root node contains all data points.

Since for MPC clustering one data point is a vector, the metric calculating the distances between the data points must be able to take the properties of the completely different dimensions (delay and angular) into account.

To decide on the actual clusters (and their number), clusters can be visually assessed by plotting the tree in a *dendrogram*, or by seeking structural divisions in the data by comparing average distances between links below a node to the distance above it. It is also possible to use other validation criteria as described in Section 6.4.

However, automatic validation methods are somewhat artificial as tree clustering is a more natural technique for small data samples, where visual inspection of the cluster tree can reveal the inherent structure of the data.

In hierarchical tree clustering, distances between *clusters* (instead of points) need to be computed. So-called *linkage methods* have been proposed for finding the inter-cluster distances. We were using the group average linkage, where the inter-cluster distance is obtained as the average of all pair-wise distances between the points in the two clusters. The pair-wise distances, in turn, can be computed with any distance metric suitable to the problem. For this, the MCD fulfils the requirements to scale data correctly for joint clustering.

The complexity of hierarchical clustering increases quadratically in the number of data points, which renders it very inefficient for a large number of data points.

Note that the proposed algorithm inherently introduces its own description of a “cluster”:

> *For a given number of clusters, a cluster is defined as the grouping of MPCs that has largest distance to neighbouring clusters.*

The distance between clusters is measured by the average linkage method [28]. A drawback of this definition is that it does not take the MPCs powers into account. Furthermore the small- and large-scale fading characteristics of the identified clusters are disregarded. Hence, this algorithm cannot perform in an optimal fashion, but still in a heuristic way.

6.3.5 KPowerMeans clustering algorithm

To be able to identify clusters both more quickly and accurately, the concept of the K-means algorithm [28] is well suited for this challenge. A prerequisite is that one uses an appropriate distance function, for which the MCD is a tailor-made solution.

Another shortcoming of the agglomerative clustering algorithm is the disregard of path powers. Algorithm 1 describes the ***KPowerMeans*** algorithm, which introduces the novelty of regarding powers of the MPCs.

This algorithm iteratively minimizes the total sum of power-weighted distances of each path to its associated cluster centroid. In the following the single steps of the algorithm are described in more detail.

KPowerMeans clustering algorithm:

1. Randomly choose K initial centroid positions $\boldsymbol{\mu}_1^{(0)}, \ldots, \boldsymbol{\mu}_{N_c}^{(0)}$
2. For $i = 1$ To *MaxIterations*
   a. Assign MPCs to cluster centroids and store indices:
   $$\mathcal{I}_l^{(i)} = \arg\min_c \{P_l \cdot \mathrm{MCD}(\mathbf{x}_l, \boldsymbol{\mu}_c^{(i-1)})\}, \tag{6.4}$$
   $$\mathcal{I}^{(i)} = [\mathcal{I}_1^{(i)} \ldots \mathcal{I}_L^{(i)}], \mathcal{C}_c^{(i)} = \mathrm{Indices}_l(\mathcal{I}_l^{(i)}{=}c)$$
   b. Recalculate cluster centroids $\boldsymbol{\mu}_c^{(i)}$ from the allocated MPCs to coincide with the clusters' centres of gravity:
   $$\boldsymbol{\mu}_c^{(i)} = \begin{bmatrix} \frac{1}{\sum_{l \in \mathcal{C}_c^{(i)}} P_l} \cdot \sum_{l \in \mathcal{C}_c^{(i)}} P_l \tau_l \\ \mathrm{angle}(\sum_{l \in \mathcal{C}_c^{(i)}} P_l \exp(j \cdot \varphi_{\mathrm{Rx},l})) \\ \mathrm{angle}(\sum_{l \in \mathcal{C}_c^{(i)}} P_l \exp(j \cdot \varphi_{\mathrm{Tx},l})) \end{bmatrix}, \tag{6.5}$$
   c. If $\boldsymbol{\mu}_c^{(i)} = \boldsymbol{\mu}_c^{(i-1)}$ for all $c = 1 \ldots N_c$, then GoTo 3.
   Else Next i
3. Return $\mathcal{R}_{N_c} = [\mathcal{I}^{(i)}, \boldsymbol{\mu}_c^{(i)}]$

Alg. 1: KPowerMeans algorithm

Ad 1) The centroid starting positions are chosen randomly from the data $\mathbf{X}$.

Ad 2a) Every MPC is associated with a cluster centroid such that the function of the total sum of differences,

$$D = \sum_{l=1}^{L} P_l \cdot \mathrm{MCD}(\mathbf{x}_l, \boldsymbol{\mu}_{\mathcal{I}_l^{(i)}}), \tag{6.6}$$

is minimized[9]. We use the MCD as the basic distance function [72, 103] but also include the *power of the paths*, which has not been considered in previous works. It can be shown that the global distance (6.6) can be minimized by the introduced algorithm, when using (6.4). The index $\mathcal{I}_l^{(i)}$ is the cluster number for the lth multipath in the ith iteration step. Vice-versa, the set $\mathcal{C}_c^{(i)}$ contains the MPC indices belonging to the cth cluster in the ith iteration step.

By including power into the distance function, cluster centroids are pulled to points with strong powers. This is intuitive and yields massive performance improvements, which are demonstrated in Section 6.3.6. Considering receiver design one usually adresses the most dominant clusters, which are characterised by *power*. So, in development of MIMO transceiver algorithms, the weighting by power is quite natural. Furthermore, the global distance function (6.6) is an inherent definition for a cluster:

> *For a given number of clusters, clusters are chosen such that they minimize the total distance from their centroids.*

[9] The symbols are defined in Section 6.3.1.

This implies that, for a given N_c, clusters are selected such that *the cluster angular and cluster delay spreads are minimized*, which is again intuitive.

Ad 2b) In the second step of the iteration, the centroids move to the centres of gravity of the groups of MPCs allocated in the previous step. Note that moving centroids can result in a new group of MPCs that will be associated with the centroid in the next iteration step.

Ad 2c) If the centroids do not move any more the algorithm has converged to a stable solution. Should this procedure take too much time, it stops after a maximum number of iterations.

Ad 3) The output of the algorithm is the index set $\mathcal{I}^{(i)}$ and the associated cluster centroids $\boldsymbol{\mu}_c^{(i)}$, which were obtained by the last iteration.

In the common case that there is no additional algorithm providing an initial guess, the KPowerMeans algorithm is performed multiple times with random initial values. The best result is determined by the smallest value of (6.6).

6.3.6 Comparison

Figure 6.1 compares the Hierarchical Tree (HT) clustering algorithm with the KPowerMeans (KPM) clustering algorithm, where for both algorithms a number of 8 clusters was ingeniously inspired. In Figure 6.1a, the original dataset is presented. The dots describe propagation paths, where the dots' colours describe their power.

The result of the HT clustering is shown in Figure 6.1c. Different colours of the propagation paths indicate a different cluster. Comparing to the result from KPM clustering in Figure 6.1d, there are quite noticeable differences.

In the KPM algorithm, the weak-powered group of paths at large delay is clustered with the larger group at smaller delay. Also, the larger group of paths between 26 and 27 ns was split up into 4 clusters. The reason for this is a significant difference in the power spectrum in this group. For better visibility, this group can better be observed in Figure 6.1b. It obviously exhibits 4 peaks in power, thus the KPM algorithm splits up this group.

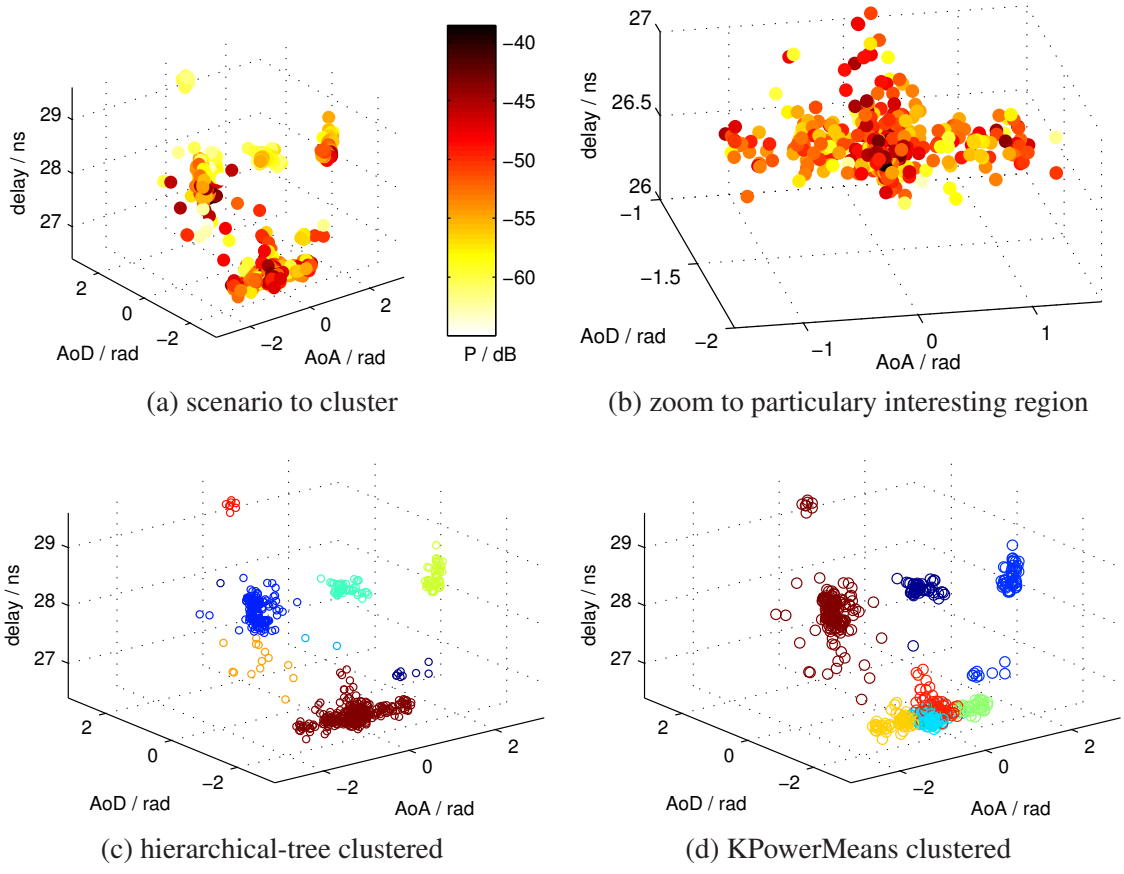
Note that this property will remain even when automatically determining the number of clusters as will be described in the next section. By including path powers, the KPM algorithm can identify the cluster structure of the scenario much more accurately.

6.3.7 Other promising clustering methods

In the meantime, quite a number of new approaches to automatically identify clusters have surfaced. This sections provides a brief overview of two particularly interesting methods.

Gaussian-mixture clustering

Gaussian mixture (GM) clustering superimposes a specific *structure* on the clusters. The cluster parameters are assumed to be Gaussian distributed, showing a higher path density in the centre than in its outskirts. For MPC clustering, this approach can be easily combined using path powers [104].



(a) scenario to cluster

(b) zoom to particulary interesting region

(c) hierarchical-tree clustered

(d) KPowerMeans clustered

Figure 6.1: Comparison of Hierarchical Tree clustering and KPowermeans clustering. (a) Unclustered scenario, (b) zoom to particularly interesting region, (c) clustered using hierarchical-tree algorithm, (d) clustered using KPowerMeans algorithm

Unfortunately, it turned out that this clustering method has its shortcomings when wanting to track clusters, since the clustering results are quite unstable.

Estimating clusters directly in the impulse response

Another interesting approach has been suggested by Yin et. al. in [55, 105]. The idea is to superimpose a structure on clusters in terms of a signal, and estimate these cluster signals directly from measured impulse responses.

Even though the approach is very promising, it is computationally not tractable at the moment.

6.4 Dismissing the genie — identifying the number of clusters

6.4.1 Heuristic approaches

There exist quite a number of heuristic approaches to determine the number of clusters. A good overview can be found in [106].

I selected two specific criteria well-known in literature, which we combined afterwards: the Caliñski-Harabasz index and the Davies-Bouldin criterion. Both of these indices and their proposed combination are described in the next paragraphs.

6.4.1.1 Caliñski-Harabasz index

When clustering L MPCs in N_c cluster, the Caliñski-Harabasz index (CH) is given as

$$\mathrm{CH}(K) = \frac{\mathrm{tr}(\mathbf{B})/(K-1)}{\mathrm{tr}(\mathbf{W})/(L-K)} \,,$$

which corresponds to the ratio between the traces of the *between-cluster scatter matrix* $\mathbf{B}$ and the *within-cluster scatter matrix* $\mathbf{W}$ [106]. Using the MCD as distance function, $\mathrm{tr}(\mathbf{B})$ and $\mathrm{tr}(\mathbf{W})$ are respectively given as

$$\mathrm{tr}(\mathbf{B}) = \sum_{k=1}^{K} L_k \cdot \mathrm{MCD}(\boldsymbol{\mu}_c, \bar{\boldsymbol{\mu}})^2 \,,$$

$$\mathrm{tr}(\mathbf{W}) = \sum_{k=1}^{K} \sum_{j \in \mathcal{C}_k} \mathrm{MCD}(\mathbf{x}_j, \boldsymbol{\mu}_c)^2 \,,$$

where L_k denotes the number of MPCs related to the kth cluster, and $\bar{\boldsymbol{\mu}}$ denotes the global centroid of the entire data set.

If we calculate the CH index for different values of N_c, e.g. in the range $[N_{c,\min},\ N_{c,\max}]$, the number of cluster N_c^{CH} corresponding to the best partition is achieved as

$$N_c^{\mathrm{CH}} = \arg\max_{N_c} \{\mathrm{CH}(N_c)\} \,, \tag{6.7}$$

corresponding to the partition with the most compact and separate cluster.

6.4.1.2 Davies-Bouldin index

The Davies-Bouldin index (DB) is a function of *intra-cluster compactness* and *inter-cluster separation* [106]. Using the MCD, the *compactness* S_k of the kth cluster is given as

$$S_k = \frac{1}{L_c} \sum_{l \in \mathcal{C}_c} \text{MCD}(\mathbf{x}_l, \boldsymbol{\mu}_c),$$

and the *separation*, i.e. the distance, between two centroids i and j, is defined as

$$d_{ij} = \text{MCD}(\boldsymbol{\mu}_i, \boldsymbol{\mu}_j)\,.$$

Finally the considered DB index is given as

$$\text{DB}(K) = \frac{1}{N_c} \sum_{i=1}^{N_c} R_i\,,$$

where

$$R_i = \max_{\substack{j=1,\ldots,N_c \\ j \neq i}} \left\{ \frac{S_i + S_j}{d_{ij}} \right\}\,.$$

When calculating the DB index for different values of K, the optimum number of cluster N_c^{DB}, corresponding to the best partition, is achieved as

$$N_c^{\text{DB}} = \arg\min_{N_c} \{\text{DB}(N_c)\}\,.$$

As for the CH index, also the DB index bases on seeking for the partition with the most compact but separated clusters.

6.4.1.3 Combined Validation

A combination of the two introduced validation criteria yield significant performance improvements [107]. The basic idea of the CombinedValidate (CV) index is to restrict valid choices of the optimum number of clusters by a threshold set in the DB index. Subsequently the CH index is used to decide on the optimum number out of the restricted set of possibilities.

We consider the set of feasible choices $\mathbf{F} = \{{N_c}^{(1)}, \ldots, {N_c}^{(N)}\} \subseteq [N_{c,\text{min}},\ N_{c,\text{max}}]$ containing only the values ${N_c}^{(i)}$ for which the following condition is satisfied,

$$\text{DB}(N_c^{(i)}) \leq t \cdot \min_{N_c} \{\text{DB}(N_c)\}\,,$$

where we chose $t = 2$. The optimum number of clusters N_c^{CV} is then obtained as

$$N_c^{\text{CV}} = \arg\max_{N_c \in \mathbf{F}} \{\text{CH}(N_c)\}\,.$$

In the unrestricted case $\mathbf{F} \equiv [N_{c,\text{min}},\ N_{c,\text{max}}]$ we obtain N_c^{CV} as in (6.7).

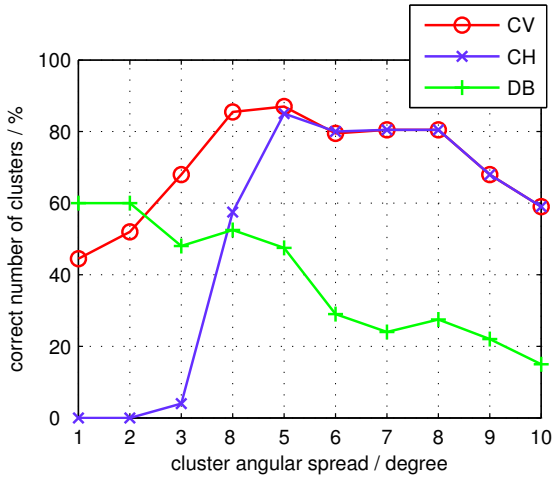


Figure 6.2: Comparing performance of validity indices

6.4.1.4 Comparison of the three methods

We tested the performance of the cluster validation schemes at different angular cluster spreads. For this we used synthetic MIMO channel data obtained from the 3GPP spatial channel model (SCM) [40], implemented by [108], but we extended the model to cope with varying angular spreads. For the following evaluation, we used 200 different samples of MIMO channels with 6 clusters, where each cluster consisted of 8 MPCs.

Fig. 6.2 demonstrates the performance of the different cluster validation indices, i.e. the novel CombinedValidate (CV), the Caliñski-Harabasz (CH), and the Davies-Bouldin (DB) index. The figure shows the fraction of the correctly estimated number of clusters versus the cluster angular spreads.

The CH index has troubles with finding the correct number of clusters with low cluster spreads. On the other hand the DB index decreases with larger cluster spreads. The CombinedValidate index always outperforms the CH index and outperforms the DB index for cluster angular spreads larger than 2.5°. We demonstrated in [102] that the clustering framework almost always finds the true (simulated) clusters as long as the correct number of clusters is detected.

6.4.2 Model-based approach

When modelling a scenario, it is computationally efficient to use as few clusters as possible. So, the optimum number of clusters is defined by the *lowest number of clusters* for which it is possible to *reflect the given scenario with a certain error threshold*.

This validation procedure is used as follows:

1. Loop for $N_c = [N_{c,\min} \dots N_{c,\max}]$
2. Cluster with N_c clusters
3. Extract the cluster parameters

4. Model the scenario using the extracted cluster parameters
5. Quantify how well the model fits the measurements
6. End loop

Note that the cluster identification can be done with any suitable clustering algorithm.

The following method models the scenario using the extracted cluster parameters: For the number of identified clusters, N_c, we extract their cluster parameters, $\hat{\mathbf{\Theta}}_c$, $c = 1 \dots N_c$, for this single snapshot. Subsequently, we generate a number of T modelled snapshots. The cluster parameters for the modelled snapshots are equal to the estimated cluster parameters, so, $\tilde{\mathbf{\Theta}}_c(t) = \hat{\mathbf{\Theta}}_c,\ t = 1 \dots T$. In contrast, the path parameters $\tilde{\mathbf{\Theta}}_{cp}(t)$ are drawn independently for every realisation t, as described in Section 2.3.2.1 on pg. 34. From this we obtain a number of paths for every realisation, $\{\tilde{\mathbf{\Theta}}_{cp}\}(t),\ c = 1 \dots N_c,\ p = 1 \dots \tilde{N}_{c,p},\ t = 1 \dots T$.

Finally, we calculate the mean mismatch between the modelled scenarios and the true one[10] directly from the modelled and estimated paths. As a measure how well paths match we use the singular values of the Environment Characterisation Metric (SV-ECM), σ_d^2, with d denoting the SV index, as described in Section 3.2.4.

Remember that the SV-ECM can be seen as "fingerprint" of the scenario. The mismatch between the true and the modelled scenarios is defined as

$$\mathcal{E} = \frac{1}{D} \sum_{d=1}^{D} |\hat{\sigma}^2_{d,[\mathrm{dB}]} - \tilde{\sigma}^2_{d,[\mathrm{dB}]}|,$$

with $\hat{\sigma}^2_{d,[\mathrm{dB}]}$ denoting the SV-ECM obtained from the true environment, and $\tilde{\sigma}^2_{d,[\mathrm{dB}]}$ denoting the SV-ECM obtained from the modelled environment. As we generate T snapshots with the model, we take the mean of the mismatch as the final measure how well the model fits.

Finally, the number of clusters is given by the lowest number of clusters where the mean mismatch is beyond a certain error threshold. We empirically set this threshold to 0.3 dB, since the mean mismatch begins to flatten around this value. Also, evaluations showed that the parameter estimator itself is only able to characterise simulated environments with a minimum error of around 1 dB [109].

This cluster validation algorithm works reasonably well, however it needs a large number of (tedious) clustering operations, which makes the algorithm very slow.

6.4.3 Power threshold criterion

Another approach that works (in connection with a trustworthy clustering algorithm) as well as the model-based validation is using a *cluster power threshold criterion*.

The concept is very simple: A cluster can only exist when it has more power than a certain ratio of the total snapshot power. Should a cluster drop below this threshold it is no longer regarded as a cluster, and clustering has to be repeated using a lower number of clusters.

As I will describe in Section 7.2.2, this power criterion can already be included in an initial guess, significantly reducing the clustering complexity. However, this approach is very sensitive to the correct choice of the clustering algorithm.

[10] Note that (i) the "true" scenario is also *estimated* from measurements, (ii) there is only *one* true scenario, but *many* modelled ones.

6.4.4 Complexity of validation approaches

This section compares the different validation approaches regarding their computational complexity. All these approaches have in common that they need clustering results for all possible numbers of clusters. Only subsequently they decide which number best matches the data.

- The *heuristic approaches* are all quite simple. Only few distance metrics need to be evaluated. Since these metrics are not tailored to the multipath clustering problem, they do not perform well in measured scenarios.
- The *model-based approach* is computationally very demanding. For every number of clusters, many snapshots need to be generated, and a singular value decomposition must be computed. Although this approach is very accurate, the high complexity is cumbersome when processing a large amount of measured data.
- The *power threshold criterion* is by far the least complex approach. Only the powers of the MPCs need to be summed up. Surprisingly, using a proper threshold level[11], the algorithm performs very well. It leads to similar results as the model-based approach does. Moreover, this criterion can be integrated into an advanced initial guess, rendering it ideal for the multipath clustering problem.

[11] See Section 4.2.2 for a discussion on valid threshold settings.

7 Time variance — joint clustering and tracking

In order to consistently parametrize the RCM (see Section 4), clusters must be identified from measurements. Since clusters are used to model time-variant scenarios, a consistent approach is required for joint cluster identification-and-tracking over time. A simple cluster tracking algorithm was presented in [110], but it did not take joint clustering and tracking into account. An alternative method is to track individual paths directly in the impulse response [82].

In this chapter we[1] developed a joint clustering-and-tracking framework that uses (i) a Kalman filter [83] to track and predict cluster positions together with (ii) a new initial-guess procedure allowing to include the prediction of the Kalman filter, and (iii) the KPowerMeans clustering algorithm using the MCD distance metric [107] to identify clusters. To test the framework we used three different sets of time-variant MIMO channel measurements, one indoor environment measured at 2.55 GHz and at 5.25 GHz showing rich scattering, an outdoor environment measured at 2.0 GHz showing few, very distinct propagation paths and many weak scattered paths, and another outdoor rural environment measured at 300 MHz. We found that this framework enabled us to extract the cluster characteristics from time-variant MIMO channel measurements consistently.

This chapter is organized as follows: Section 7.1 will describe the problem of clustering and tracking, and introduce the cluster data model. In Section 7.2 we provide a comprehensive description of the joint clustering-and-tracking framework. A brief overview of the three different MIMO channel measurements is provided in Section 7.3. Results from applying the framework to the measurement data are presented in Section 7.4. Finally, we conclude this chapter in Section 7.5.

7.1 Cluster data model

In the clustering algorithms introduced in the previous chapter, each snapshot is clustered independently [78, 107], and the clusters might be tracked afterwards [110]. The problem to solve is how to *combine clustering and tracking* in order to improve the clustering performance and to consistently track clusters.

For tracking it turned out to be advantageous, to combine the estimations from several snapshots using a sliding window. Doing so, the data becomes more stationary since only a subset of new MPCs is added and removed in every time instant. Note that this window size is an external parameter, which even has (a limited) impact on the final fit of the RCM (cf. Section 4.2.2).

[1] This clustering-and-tracking algorithm was developed in close cooperation with Ruiyuan Tian, my former master student, and has been significantly improved during a very fruitful cooperation in the framework of a COST 2100 short-term scientific mission with Lund University. For this reason, I decided to write this chapter using the plural form.

We consider N data windows, $n = 1 \dots N$, each with a number of $L^{(n)}$ MPCs, where every single MPC is represented by its power $P_l^{(n)}$, $l = 1 \dots L^{(n)}$, and a parameter vector $\mathbf{x}_l^{(n)} = [\tau_l^{(n)}\ \varphi_{\mathrm{Rx},l}^{(n)}\ \varphi_{\mathrm{Tx},l}^{(n)}]^{\mathrm{T}}$ containing the delay, azimuth AoA and azimuth AoD, respectively.

The data for all paths are collected in the power vector $\mathbf{P}^{(n)} = [P_1^{(n)} \dots P_L^{(n)}]^T$ and the matrix $\mathbf{X}^{(n)} = [\mathbf{x}_1^{(n)} \dots \mathbf{x}_L^{(n)}]^T$.

Each *cluster* is determined by following parameters:

1. A *unique cluster-ID* c.
2. The *cluster power* at time n. Denoting the set of path indices belonging to cluster c at time snapshot n by $\mathcal{I}_c^{(n)}$, the cluster power is calculated as $\gamma_c^{(n)} = \sum_{l \in \mathcal{I}_c^{(n)}} P_l^{(n)}$.
3. The *number of paths* within the clusters $L_c^{(n)} = |\mathcal{I}_c^{(n)}|$, where every path is assumed to belong to one cluster, uniquely.
4. The *cluster centroid* position in the angle-angle-delay domain $\boldsymbol{\mu}_c^{(n)}$. The cluster centroid position can be calculated as

$$\begin{aligned} \boldsymbol{\mu}_c^{(n)} &= [\tau_c^{(n)}\ \varphi_{\mathrm{Rx},c}^{(n)}\ \varphi_{\mathrm{Tx},c}^{(n)}]^{\mathrm{T}} = \\ &= \begin{bmatrix} \frac{1}{\gamma_c^{(n)}} \cdot \sum_{l \in \mathcal{I}_c^{(n)}} P_l^{(n)} \tau_l^{(n)} \\ \mathrm{angle}(\sum_{l \in \mathcal{I}_c^{(n)}} P_l^{(n)} \exp(j \cdot \varphi_{\mathrm{Rx},l}^{(n)})) \\ \mathrm{angle}(\sum_{l \in \mathcal{I}_c^{(n)}} P_l^{(n)} \exp(j \cdot \varphi_{\mathrm{Tx},l}^{(n)})) \end{bmatrix}, \end{aligned} \tag{7.1}$$

   where the mean angle is calculated by averaging angles over their respective complex representation.

   For tracking, also the centroid speed is of interest, so we combine the position and speed in the cluster tracking parameter vector $\boldsymbol{\theta}_c^{(n)} = [\tau_c^{(n)}\ \Delta\tau_c^{(n)}\ \varphi_{\mathrm{Rx},c}^{(n)}\ \Delta\varphi_{\mathrm{Rx},c}^{(n)}\ \varphi_{\mathrm{Tx},c}^{(n)}\ \Delta\varphi_{\mathrm{Tx},c}^{(n)}]^{\mathrm{T}}$.

5. The *cluster's joint spread* $\mathbf{C}_c^{(n)}$, which is the power-weighted covariance matrix of the parameters of the paths within one cluster at time n. The main diagonal contains the cluster spreads of the individual dimensions, i.e. the cluster delay spread, the cluster AoA spread and the cluster AoD spread. The off-diagonal elements describe the correlation between these spreads.

   The cluster spread matrix is calculated by

$$\mathbf{C}_c^{(n)} = \frac{\sum_{l \in \mathcal{I}_c^{(n)}} P_l^{(n)} (\mathbf{x}_l^{(n)} - \boldsymbol{\mu}_c^{(n)})(\mathbf{x}_l^{(n)} - \boldsymbol{\mu}_c^{(n)})^{\mathrm{T}}}{\gamma_c^{(n)}}. \tag{7.2}$$

   Note that in this equation, whenever adding or subtracting angles, the result must be mapped to the principal value in the interval of $(-\pi, \pi]$, which can be achieved easily by the operation

$$\mathrm{pv}(\varphi) = \mathrm{angle}(\exp(j\varphi)). \tag{7.3}$$

Based on this cluster data model, we will now introduce the *clustering-and-tracking framework*.

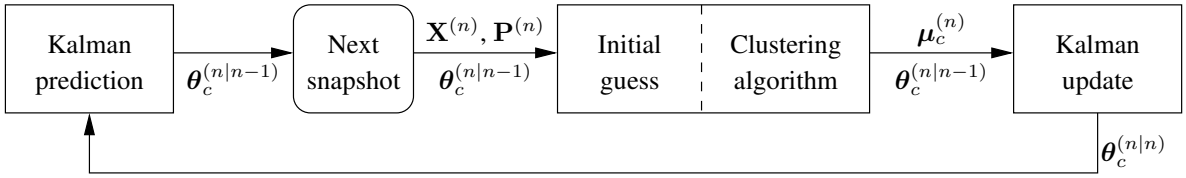


Figure 7.1: Clustering framework: Clusters with parameters $\boldsymbol{\theta}_c^{(n)}$ are identified and tracked in the input data $(\mathbf{X}^{(n)}, \mathbf{P}^{(n)})$, the Kalman filter updates the cluster parameters, the prediction provides an input to the initial guess.

7.2 Framework

For each time snapshot, the following steps are performed (see Figure 7.1):

1. A Kalman filter [83] both tracks the cluster position over time, and predicts the cluster position in the next snapshot.
2. The initial-guess routine provides a initial guess of the cluster centroids, taking the predicted cluster centroids into account.
3. The clustering algorithm identifies clusters in the measurement data based on the initial guess.

7.2.1 Kalman cluster tracking

7.2.1.1 State-space model

We consider linear cluster movement in delay and angles. The cluster joint spreads (size and rotation) are not used for tracking the position. In the Kalman tracking [83], only the cluster centroid position $\boldsymbol{\theta}_c$ is used. We use the following state equation

$$\boldsymbol{\theta}_c^{(n)} \quad = \quad \Phi\boldsymbol{\theta}_c^{(n-1)} + \mathbf{w}^{(n)}, \tag{7.4}$$

where $\mathbf{w}^{(n)}$ denotes the state-noise with covariance matrix $\mathbf{Q}$, and Φ is the state-transition matrix given by

$$\Phi = \mathbf{I}_3 \otimes \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

where identity matrices are denoted by $\mathbf{I}_d$ with d denoting the dimension, and $\otimes$ denotes the Kronecker matrix product.

Since we can observe only the cluster centroids and not their speed, we use the following observation model

$$\boldsymbol{\mu}_c^{(n)} \quad = \quad \mathbf{O}\boldsymbol{\theta}_c^{(n)} + \mathbf{v}^{(n)}, \tag{7.5}$$

where $\boldsymbol{\mu}_c^{(n)}$ describes the *observed* cluster centroid position, thus $\mathbf{H}$ is given by

$$\mathbf{O} = \mathbf{I}_3 \otimes \begin{bmatrix} 1 & 0 \end{bmatrix}, \tag{7.6}$$

and $\mathbf{v}^{(n)}$ denotes the observation noise with covariance matrix $\mathbf{R}$.

7.2.1.2 Tracking equations

The derivation of the Kalman filter is straight-forward and leads to following prediction and update equations[2]

$$
\begin{aligned}
&\textbf{Prediction:} \\
\boldsymbol{\theta}_c^{(n|n-1)} &= \Phi\boldsymbol{\theta}_c^{(n-1|n-1)}, && (7.7)\\
\mathbf{M}^{(n|n-1)} &= \Phi\mathbf{M}^{(n-1|n-1)}\Phi^{\mathrm{T}} + \mathbf{Q}, && (7.8)\\
&\textbf{Update:} \\
K^{(n|n)} &= \mathbf{M}^{(n|n-1)}\mathbf{O}^{\mathrm{T}}(\mathbf{O}\mathbf{M}^{(n|n-1)}\mathbf{O}^{\mathrm{T}} + \mathbf{R})^{-1}, && (7.9)\\
\boldsymbol{\theta}_c^{(n|n)} &= \boldsymbol{\theta}_c^{(n|n-1)} + K^{(n|n)}(\boldsymbol{\mu}_c - \mathbf{O}\boldsymbol{\theta}_c^{(n|n-1)}), && (7.10)\\
\mathbf{M}^{(n|n)} &= (\mathbf{I} - K^{(n|n)}\mathbf{O})\mathbf{M}^{(n|n-1)}. && (7.11)
\end{aligned}
$$

The identity matrix is used as initial value for $\mathbf{Q}$, $\mathbf{M}$ and $\mathbf{R}$.

7.2.1.3 Cluster association

A major problem in multi-target tracking is how to associate the predicted with the identified cluster centroids. Usually, such an association is based on the Euclidean distance in parameter space. Since we are tracking clusters that show a certain extent in parameter space, the Euclidean distance does not provide a good association. Instead, we use the following probability-based method:

- The distance between a cluster with parameters $(\boldsymbol{\mu}_c,\ \mathbf{C}_c)$ and a cluster centroid $\tilde{\boldsymbol{\mu}}$ is defined by

$$
\mathcal{G}_c(\tilde{\boldsymbol{\mu}}|\boldsymbol{\mu}_c, \mathbf{C}_c) = \frac{1}{(2\pi)^{3/2}\left|\mathbf{C}_c\right|^{1/2}} \cdot \exp\left(-\frac{1}{2}(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_c)^{\mathrm{T}}\mathbf{C}_c^{-1}(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_c)\right). \qquad (7.12)
$$

  Since a small distance between the two centroids now corresponds to a large value of this function, we refer to it as the *closeness function*.

- The closeness function is evaluated between all predicted and all new cluster centroids in both directions, i.e. between the old and the new centroids using the old covariance matrix, and between the new and old cluster centroids using the new covariance matrix.
- For each old cluster we determine the closest new cluster by finding the maximum value of the closeness function, and vice versa, for each new cluster we determine the closest old cluster in the same way.
- Whenever these two clusters are closest mutually, these two clusters are associated and being considered as the *tracked* cluster.
- Clusters that were not associated from the old snapshot stop to exist, clusters that were not associated from the new snapshot are considered as new clusters.

[2] Note that the principal-value calculation rules apply for the angular dimensions

7.2.2 Cluster initial guess

A crucial point in any iterative clustering algorithm is the initial guess of the cluster centroids. Our new method chooses the centroids by maximizing the distances between them. In the following we will present how to choose the initial-guess centroids $\hat{\boldsymbol{\mu}}_c$.

1. Initialization:
   - No cluster prediction available:
     The first centroid $\hat{\boldsymbol{\mu}}_1$ is chosen as the path having strongest power. Go to Step 2.
   - Cluster prediction available:
     Copy the initial-guess centroids from the predicted values. Go to Step 4.
2. Calculate a weighted distance between any path and all (initial-guess) centroids using the multipath component distance (MCD) [102] by

   $$D(\mathbf{x}_l^{(n)}, \hat{\boldsymbol{\mu}}_c) = \log_{10}(P_l^{(n)}) \cdot \mathrm{MCD}(\mathbf{x}_l^{(n)}, \hat{\boldsymbol{\mu}}_c).$$

   This leads to an $L \times c$ distance matrix $\mathbf{D}$ for every snapshot n. Here, the MCD is log-power weighted.
3. The new centroid is chosen from these paths, by selecting the one to which has the maximum minimum distance to any centroid, i.e. $l = \arg\max[\min_c \mathbf{D}]$, where $\arg\max[\cdot]$ returns the index of the maximum element.
4. Reallocate all paths to their closest centroid (as in the KPowerMeans algorithm) and calculate the cluster power. Note that, in this case, the power-weighted MCD is also used but the powers contribute linearly.
5. If the maximum number of clusters was not reached, and all centroid powers are larger then 1% of the total snapshot power, then repeat from Step 2.
   Else discard the last centroid and stop. This algorithm leads to a trustworthy identification of the number of clusters.

7.2.3 Clustering algorithm

We use the KPowerMeans clustering algorithm presented in Section 6.3.5 with following modifications: (i) we apply the initial guess as described above, (ii) since the initial guess is deterministic, the algorithm is performed only once.

Should the outcome result in clusters carrying less than 1% of the snapshot power, the result is discarded and the procedure is restarted with the initial guess, but reducing the maximum number of clusters by one. Note that in this algorithm the existence of singleton clusters is possible, as long as they show enough power. In this way we can also account for strong specular reflections.

7.3 Measurements

For this chapter, I decided to use also additional measurements than the ones presented in Part III. So we have data from three different channel sounding campaigns conducted by different institutions for comparison.

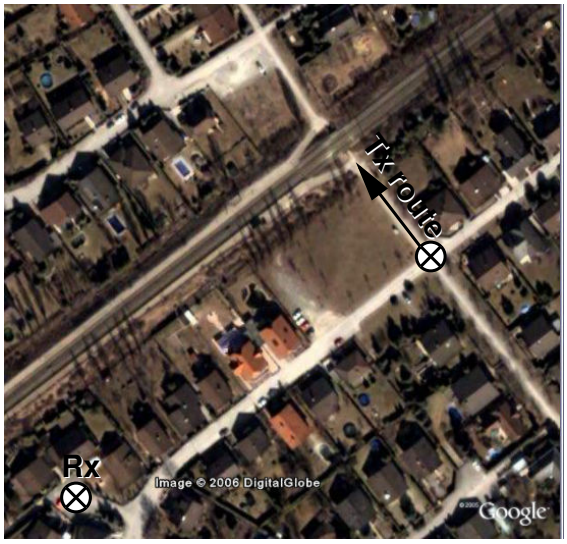


Figure 7.2: Sub-urban environment (Weikendorf, Austria)

7.3.1 Campaign I — indoor scenario

These measurements are comprehensively described in Part III. The following paragraph provides a short overview of the specific measurements used for the following evaluations, and for a direct comparison to the other environments.

The measurements were conducted at the University of Oulu, Finland using an Elektrobit Propsound CS^{TM} wideband radio channel sounder at a centre frequency of 2.55 GHz, and at 5.25 GHz [85]. The recorded impulse responses were post-processed using the ISIS (initialization-and-search improved SAGE) algorithm to estimate single propagation paths [111].

In this chapter we discuss results from a line-of-sight (LOS) measurement route in a cafeteria (route TxR22, see pg. 176). The Rx was placed on a table, while the Tx was moved along the indicated route in the room over a distance of $44\,\mathrm{m}$. Because of many metal chairs and tables, and the quite reflective walls, we expected rich scattering in the channel apart from the LOS component. However, it turned out that the observed channels were still pretty directive.

7.3.2 Campaign II — outdoor sub-urban environment

The data were collected in a small town called Weikendorf, northwest of Vienna, Austria. For the measurements we used a RUSK MEDAV channel sounder operating at a centre frequency of 2.0 GHz [110]. Snapshots of the radio channel were recorded approximately every 1.6λ. The recorded impulse responses were also post-processed using the ISIS algorithm to estimate the propagation paths.

The measurement route (see Figure 7.2) was along a road towards a railway tunnel, mostly with LOS, partly with obstructed LOS. The total distance traveled was $53\,\mathrm{m}$ corresponding to 353λ. The channel is quite directive at the Rx side, but shows rich scattering around the Tx.

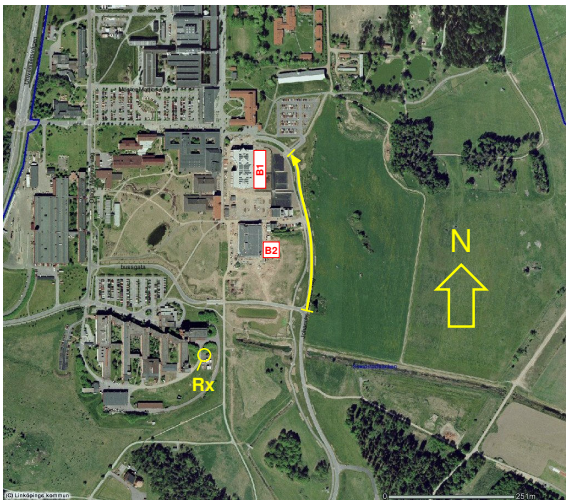


Figure 7.3: Rural environment (Linköping, Sweden)

7.3.3 Campaign III — outdoor rural environment

The third set of measurements were conducted by the Swedish Defense Agency (FOI) and Lund University in a rural environment in Linköping, Sweden. Using the RUSK Lund MIMO channel sounder, data was collected in the 300 MHz band. A description of the measurement campaign can be found in [112][3].

The measurement route used in this chapter (see Figure 7.3) is approximately 320 m long, corresponding to 320λ. The snapshot spacing used in this chapter is approximately 0.97λ. The measured impulse responses were post-processed by a SAGE algorithm [111] to obtain propagation paths for each snapshot of the channel.

7.4 Results

We applied our joint clustering-and-tracking framework to the three sets of measurements and found that the algorithm provides clusters that match well the time-varying physical propagation mechanisms observed in the measured scenarios. Exemplary plots from the indoor and the rural outdoor scenario are shown in Figures 7.4 and 7.5. The individual plots show the evolution over time. Propagation paths are marked by dots, their power is colour coded from red (strong power) to blue (weak power). Clusters are shown by ellipsoids (capturing 99.9% of the power of the included paths), where the colour describes the mean power of the included paths, and the numbers indicate the cluster IDs placed at the cluster centroid.

This section details the distribution of the cluster parameters, i.e. the cluster positions, the spreads, and the time variance parameters, such as how much do the cluster parameters change during the existence of the individual clusters.

[3] I would like to acknowledge Gunnar Eriksson from the Swedish Defence Agency (FOI) for providing these measurements.

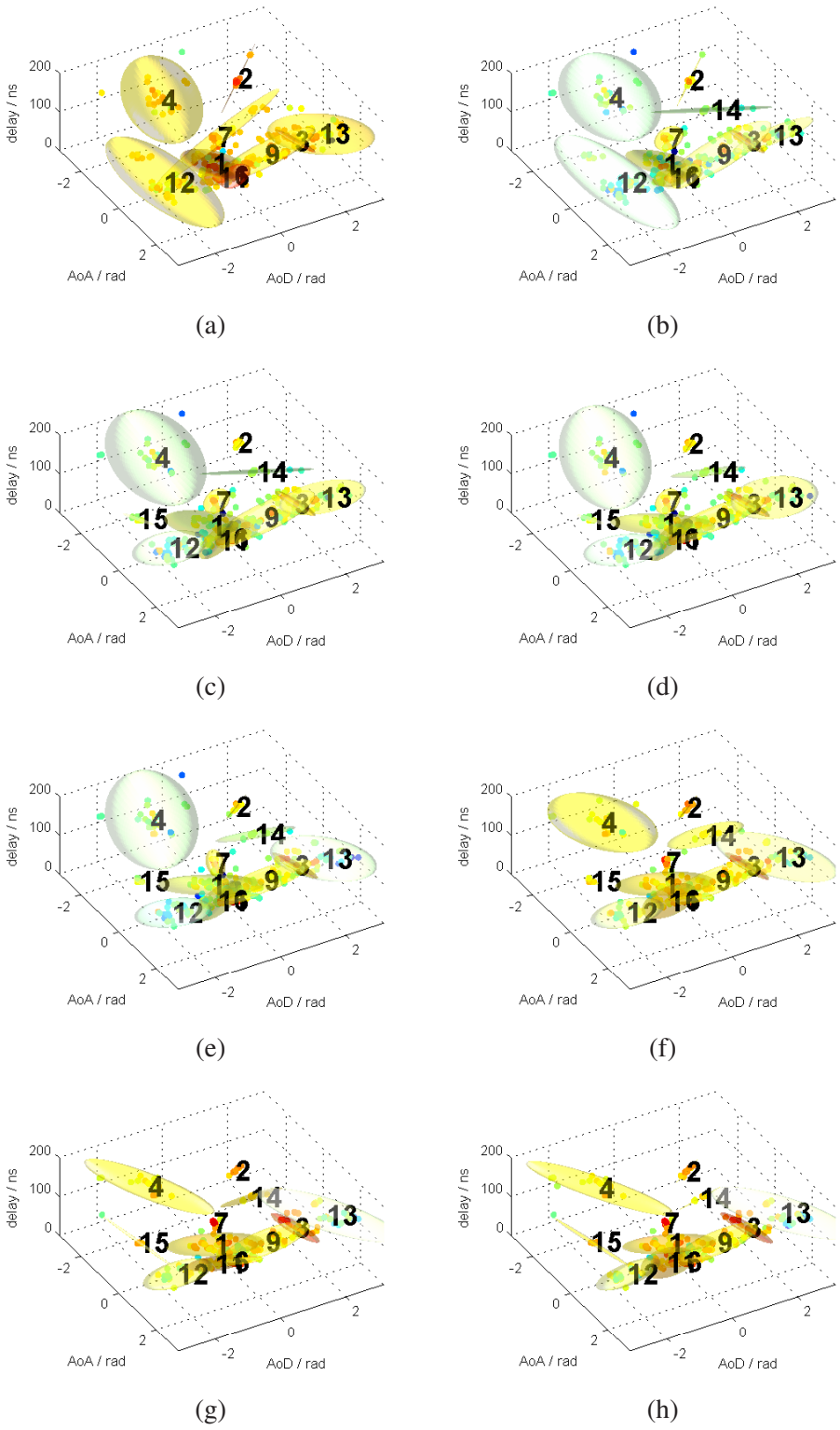


Figure 7.4: Tracked clusters from indoor scenario at 2.55 GHz; (a)-(h) show the clusters' evolution over time

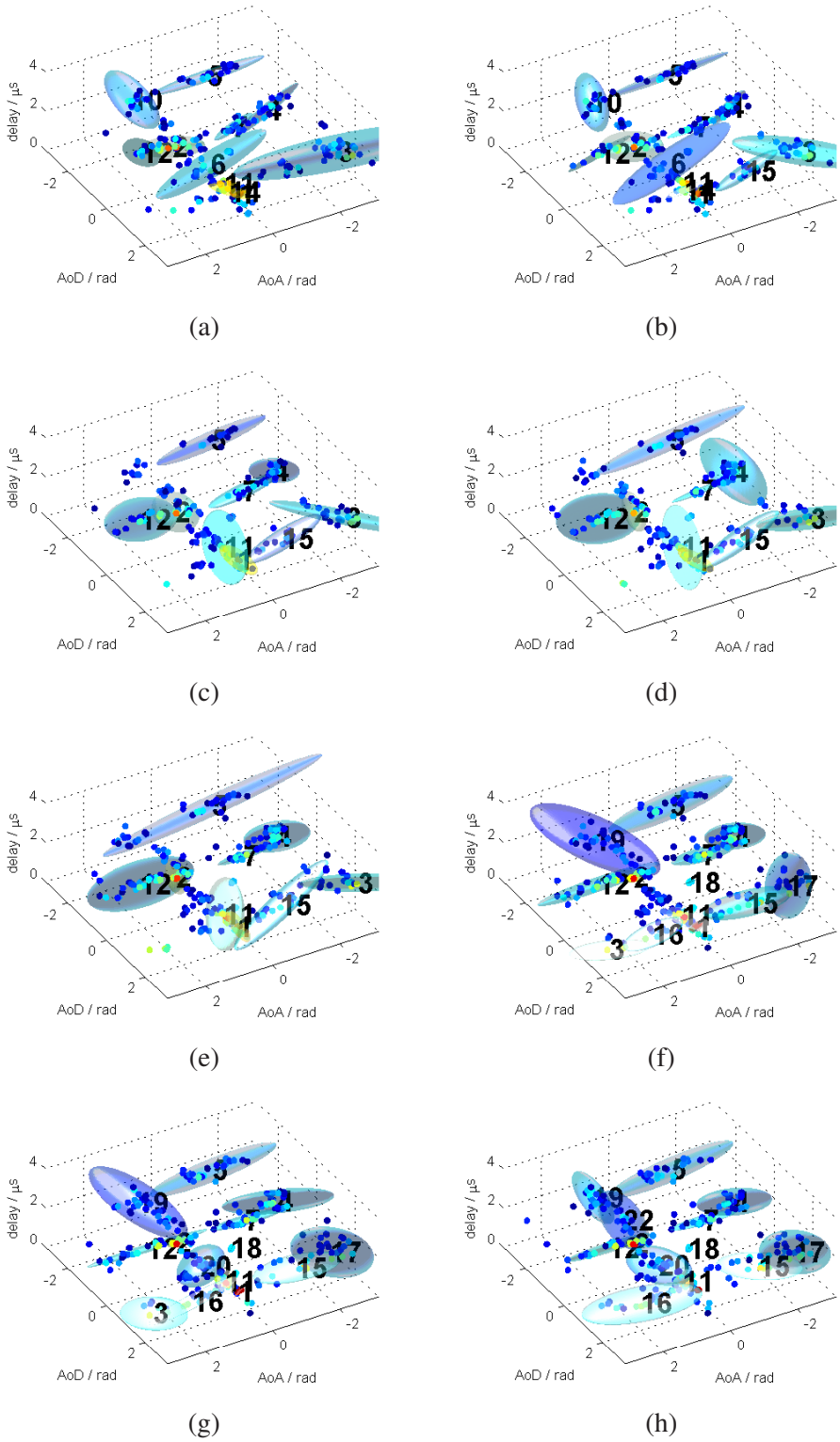


Figure 7.5: Tracked clusters from rural outdoor scenario; (a)-(c) show the clusters' evolution over time

We focus on (i) the cluster position and the cluster spreads, (ii) the change of the cluster position over the traveled wavelength ("cluster movement"), (iii) the change of the cluster spreads, (iv) the cluster lifetime and cluster birth and death rates.

7.4.1 Cluster position and cluster spreads

The most important parameters for the RCM are the cluster positions and the cluster spreads. They determine the multipath structure of the channel. These parameters are strongly correlated with each other, and with the number of paths within a cluster, as well as with the number of clusters in an environment.

As an example, I will compare these cluster parameters evaluated from the indoor measurement route, for the two centre frequencies. Figure 7.6 presents the pdfs of these cluster parameters: the cluster mean delay, and its delay spread, as well as the cluster mean azimuth and the azimuth at the Tx and Rx, respectively, the number of paths within a cluster, and the number of clusters in the environment.

The comparison of the cluster parameters provide an interesting insight into the propagation environment. We observe that the cluster spread distributions are more peaky in the 5.25 GHz measurements. The cluster spreads are also mostly lower at 5.25 GHz. To account for these "smaller" clusters, the number of clusters at 5.25 GHz is larger, while having fewer paths within a cluster. We also see that the cluster mean AoD is a bit more distinct at 5.25 GHz than at 2.55 GHz.

From this we can conclude that the propagation characteristics change when going from 2.55 GHz to 5.25 GHz. While small objects were scattering waves at 2.55 GHz, it becomes more probable at 5.25 GHz that they specularly reflect impinging waves. This leads to a more distinct propagation environment. We have speculated about this behaviour already in [113], when discussing the delay spreads of different environments for the two centre frequencies.

The mean parameters (mean cluster delay, mean cluster AoA, mean cluster AoD) determine the geometry of the room and the position of Tx and Rx within this room. The clusters are placed according to these distributions.

More interesting is the evaluation of the cluster spread parameters (cluster delay spread, cluster AoA spread and cluster AoD spread). These parameters can also be compared with the ones available in literature. The authors of [77] found cluster sizes (the whole extent of a cluster) ranging between 6° and 36° for comparable indoor scenarios. In [114] the authors identified only 2.3 clusters on average, but identified cluster spreads of around 27°. The spreading parameters of an assumed Laplacian distribution of the cluster parameters were investigated in [96]. The authors found spreading between 21.5° to 25.5°. Note that all these investigations were done based on *visual clustering*, where the definition of a cluster was quite unclear.

Note that these cluster parameters are all jointly correlated, as presented in Section 4.2.1.2 (cf. Figure 4.1 on page 60).

7.4.2 Cluster movement

For the following evaluations, we use the data sets from the three different kinds of environments. For the indoor environment, only the 2.55 GHz data was considered.

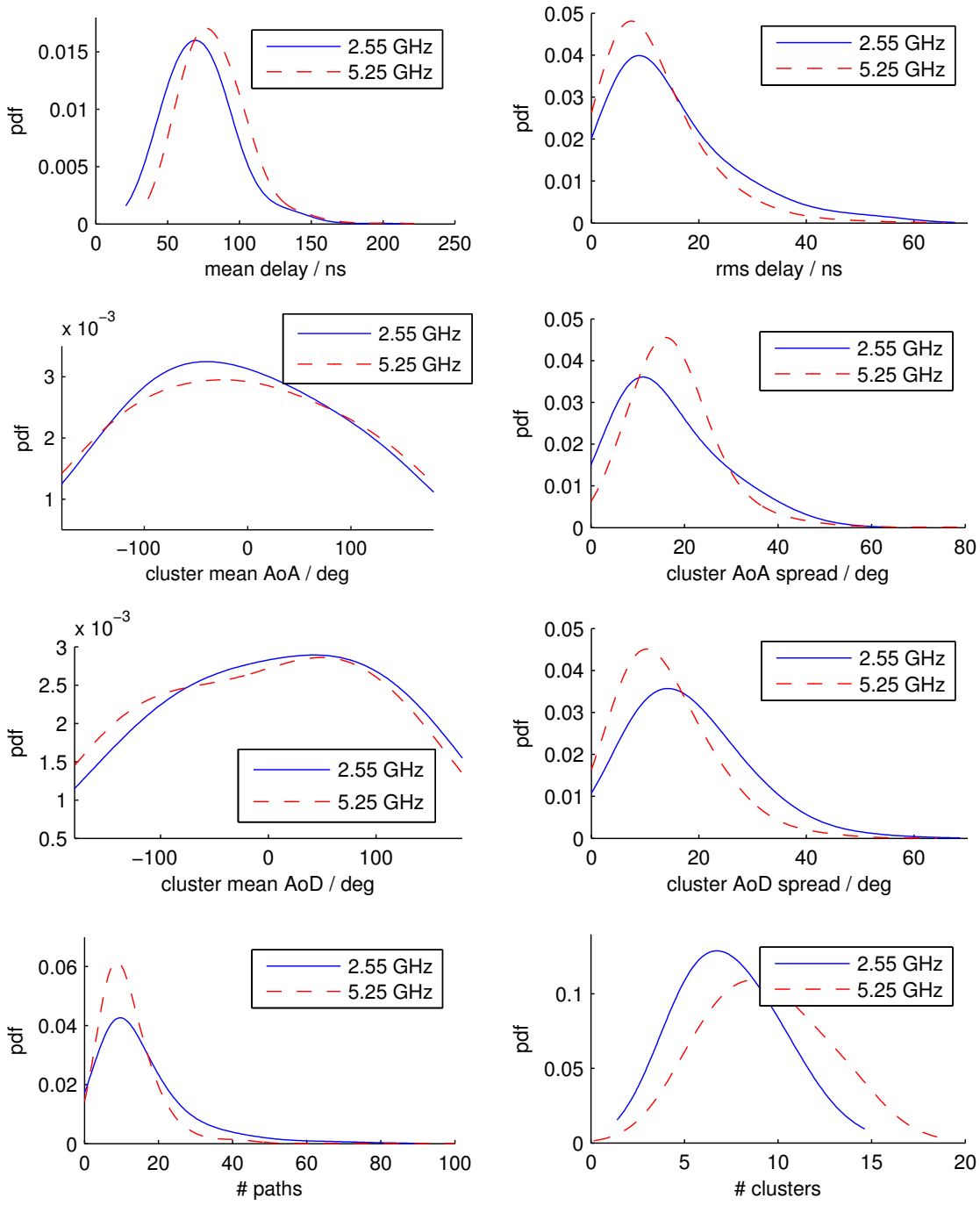


Figure 7.6: Cluster parameter pdfs evaluated for 2.55 GHz and 5.25 GHz from Route TxR22.

Figure 7.7 exemplifies the cluster mean delay for one individually selected cluster from each environment. The cluster mean delay varies significantly in the presented indoor environment (Figure 7.7a), it seems that the cluster jitters around a steady increase. This jitter can be attributed to a non-ideal clustering and tracking. Also in the sub-urban environment (see Figure 7.7b), the cluster is changing its position quite strongly over its lifetime. This cluster is likely to combine two scatterers. The steep decrease in delay is likely to come from a change in the propagation conditions, where the scatterer with shorter delay grew stronger, thus making the cluster delay smaller. In the rural environment (Figure 7.7c), the strong movement during the first few snapshots seems to be an artifact of the tracking algorithm. The Kalman tracking needs some training to keep track of the cluster. After these first snapshots, the cluster delay increases almost steadily, as it should.

We describe the cluster movement by the change of the cluster mean parameters related to travelling one wavelength with the transmitter. In the following we will present the average cluster movement of all clusters identified in the three environments. Since the sample mean of the parameters is strongly influenced by the artifacts, we decided to use the *median* of the sample instead[4]. Figures 7.8–7.10 show histograms of these movement parameters. The cluster mean delay changes within the range of $-5 \ldots 5$ ns per wavelength (see Figure 7.8). Strong changes can again be attributed to the combination of more than one propagation effect in one cluster.

The median cluster changes in AoA (see Figure 7.9) are particularly interesting, since we observe significant differences in the histograms for the different environments. This effect can be attributed to the very different propagation environments. In the indoor scenario, there was strong scattering around both the Rx and Tx, leading also to stronger cluster movement around the Rx. In the sub-urban scenario, the observed values are quite small. The Rx did not experience local scattering since it was placed on a crane overlooking the environment. In the rural environment, there was also strong scattering around the Rx, where the movement of the Tx led to cluster movement on the Rx side. The mean cluster changes in AoD shown in Figure 7.10 are quite similar for all environments.

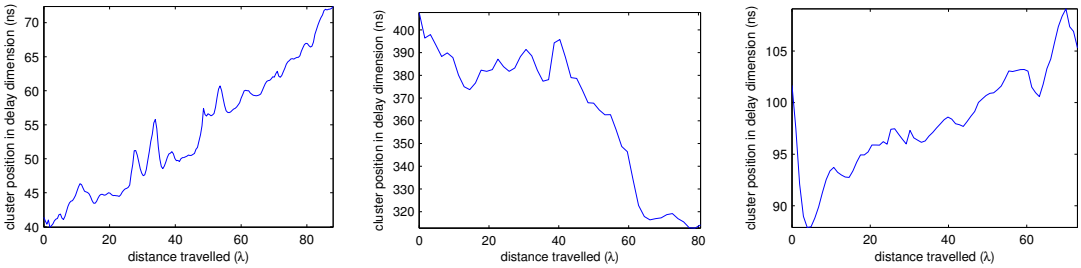
7.4.3 Change of cluster spreads

Another important figure is how much the cluster spreads change during the lifetime of a cluster. Figure 7.11 shows the cluster delay spread of one individually selected cluster for each of the environments. In all three environments we observe that the delay spread changes significantly over the travelled distance. These changes seem to be rather random. This effect is due to small-scale fading that cannot be resolved by the path estimator. A cluster can only be estimated with limited accuracy. When the phases of the paths within a cluster interfere destructively, the cluster will shrink in size, and vice versa, when the paths within a cluster interfere constructively, the cluster will grow. These changes can be up to 100 % of the median cluster size. For this reason, the median cluster size is a good measure to quantify the actual size of a cluster changing over time.

Moreover, we also sometimes observe some kind of outliers, which may again be an artefact of the clustering-and-tracking algorithm.

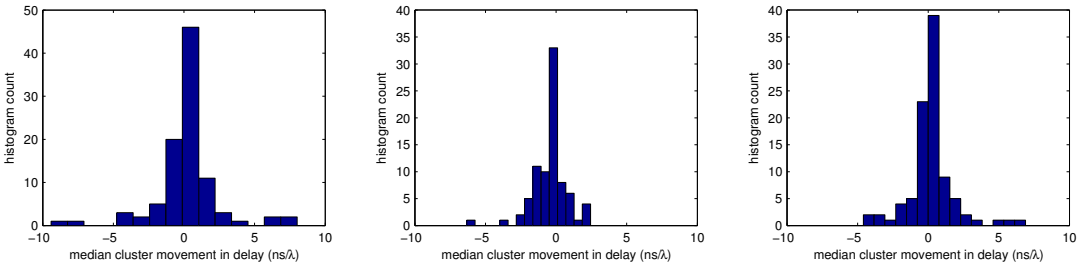
In order to quantify these changes, we use a deviation measure similar to the standard deviation of the cluster spread with following changes: (i) we use the median instead of the mean in order to

[4]The cluster mean delay (i.e. the cluster position in delay) changes over time. To describe this change, we use the *median* of the cluster mean delay.



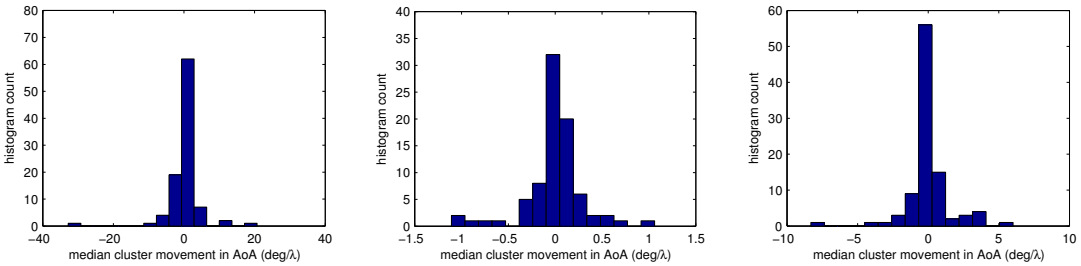
(a) indoor (b) sub-urban (c) rural

Figure 7.7: Cluster movement in delay dimension



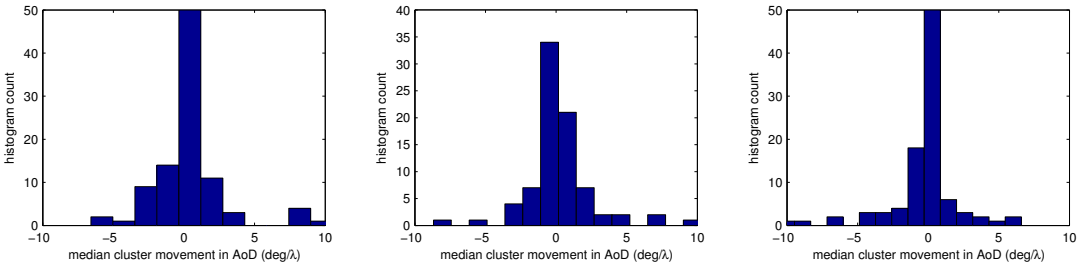
(a) indoor (b) sub-urban (c) rural

Figure 7.8: Median cluster movement in delay



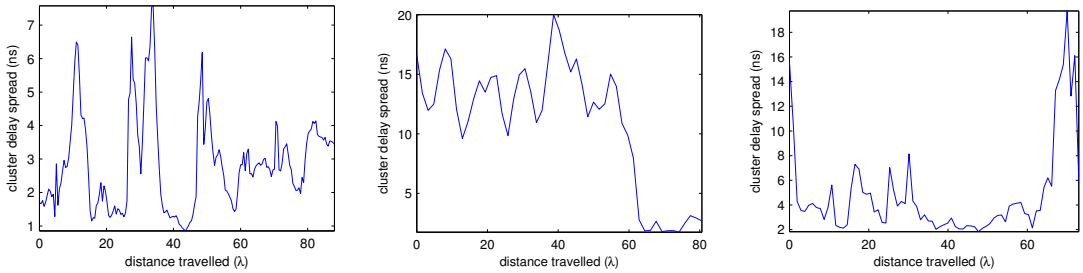
(a) indoor (b) sub-urban (c) rural

Figure 7.9: Median cluster movement in AoA



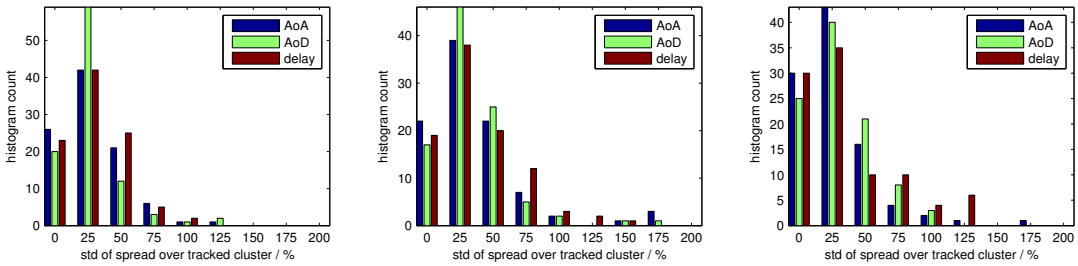
(a) indoor (b) sub-urban (c) rural

Figure 7.10: Median cluster movement in AoD



(a) indoor (b) sub-urban (c) rural

Figure 7.11: Change of the cluster delay spread



(a) indoor (b) sub-urban (c) rural

Figure 7.12: Standard deviations of the change of the cluster delay spreads in percent of the median delay spread

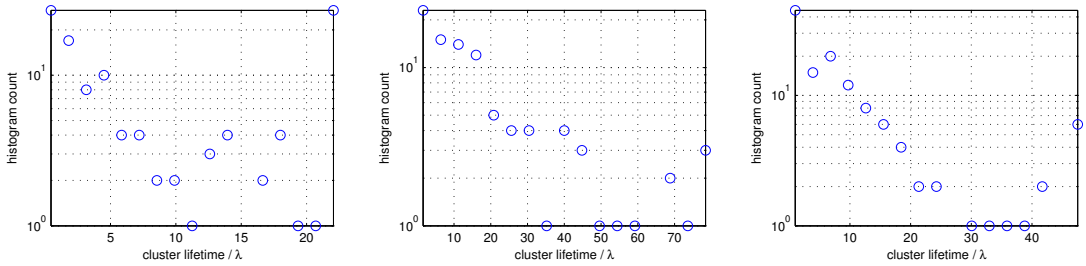
mitigate the effect of outliers, (ii) we relate the deviation to the median value of the cluster spread to obtain a spread in percent. We calculate this spread deviation as

$$\mathcal{D}_\tau = \frac{\sqrt{\frac{1}{N}\sum_{k=1}^{N}(\sigma_{\tau,k} - \bar{\sigma}_\tau)^2}}{\bar{\sigma}_\tau},$$

where $\sigma_{\tau,k}$ denotes the cluster delay spread at snapshot k, $\bar{\sigma}_\tau$ denotes the cluster *median* delay spread over all snapshots, and N is the lifetime of the regarded cluster. The deviations for the AoA and AoD cluster spreads are defined similarly.

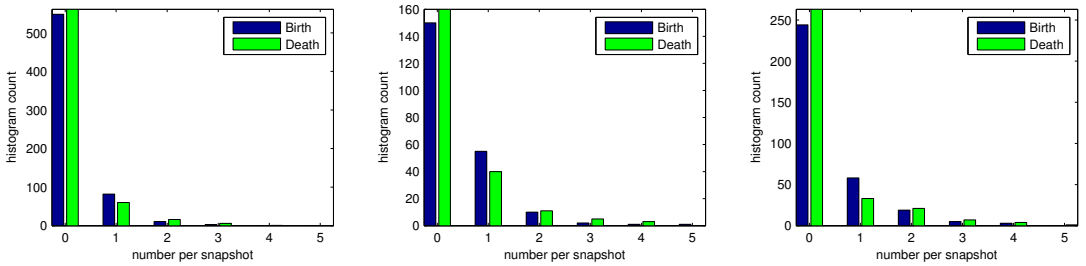
Figure 7.12 shows the the histograms of the deviation for all three cluster dimensions. We first observe a similar behaviour for all three dimensions AoA, AoD and delay. A standard deviation around 25 % of the cluster spread value is most probable. Surprisingly, the results are quite similar in all different environments. This results from the unresolvable small-scale fading.

The highly improbable values above 100 % indicate that some clusters tend to grow for short periods, where they show considerably larger spreads than the median spread. In these cases, the clustering algorithm (accidentially, or for a good reason) combines wider-spread paths into one cluster. This effect occurs when some weak outlying propagation paths exist for just a few snapshots and then vanish again. Such paths are allocated to the closest cluster.



(a) indoor (b) sub-urban (c) rural

Figure 7.13: Cluster lifetimes in terms of wavelengths



(a) indoor (b) sub-urban (c) rural

Figure 7.14: Cluster birth and death rate per snapshot

7.4.4 Cluster lifetimes, birth and death rates

Important parameters for cluster-based models are the lifetime of clusters and how strongly the number of clusters changes between different snapshots.

Figure 7.13 shows histograms of the cluster lifetimes for the different environments. Like in [110], the plots give rise to the assumption of an exponential distribution of the lifetime. However, we also observe a number of clusters with significantly larger lifetime. We particularly evaluated this effect for the rural environment and found that these long-living clusters come from the LOS path and from dominant reflectors in the channel. Thus, these clusters do have significant impact and must not be neglected when modelling the radio channel.

The birth rates and death rates of the clusters are evaluated in Figure 7.14. The histograms show the number of newly born or died clusters evaluated for all snapshots. It is evident that a change of one or two clusters in a snapshot is quite probable. Only in very few cases, three or more clusters are born or die at the same snapshot. Again, the results are fairly similar for all environments which leads to the conclusion that clusters can be tracked quite well.

7.5 Conclusions

This chapter presented (i) a novel joint clustering-and-tracking algorithm in order to identify time-variant cluster parameters for geometry-based stochastic MIMO channel models, and (ii) we evaluated the time-variant behavior of multipath clusters from MIMO channel measurements in three different environments.

Using a Kalman filter to track the clusters and to predict the cluster position for the next time instant significantly improves the ability to track clusters compared to empirical algorithms [110].

For tracking multiple clusters, we introduced a novel method for cluster association of predicted and identified clusters. Only by including the cluster spreads, we could enable the algorithm to associate clusters correctly.

Applying the framework on three highly different types of MIMO channel measurements led to consistent results. The combination of tracking and clustering allows to identify the time-variant properties of clusters coherently.

The resulting cluster parameters show that clusters move significantly in the parameter domain. Some of these movements are quite strong, which can be attributed to changing propagation conditions underlying these particular clusters.

The cluster spreads are also strongly varying over the lifetime of the individual clusters with a deviation of up to 50 % around their median value. This effect is twofold: (i) a variation of the cluster spread up to 50 % is due to small-scale fading that cannot be resolved by the path estimation algorithms, thatwhy it is important to track clusters over time for obtaining a median cluster spread, (ii) large variations are due to the allocation of short-living outlying paths to dominant clusters, which is an artefact of the clustering algorithm which is mitigated by using a median spread to describe the cluster extent in space.

Cluster lifetimes are approximately exponentially distributed. However, the line-of-sight cluster, and clusters from dominant reflections need to be accounted for individually, since they show a much longer lifetime.

Our results show that the environment plays a significant role for the cluster movement parameters and for the cluster lifetimes, while the mean deviation of the spreads is quite similar in all environments.

Part III

MIMO channel measurements

The difference between measurements and simulations is that nobody believes in simulations, except the one who did them, but everybody believes in measurements, except the one who did them.

general wisdom in science

8 Measurement objectives

My primary objective was to define and parametrize the RCM. These parameters have to be estimated from MIMO channel measurements.

The data from a previous measurement campaign on which I wanted to base these estimations on (INTHF at 2.45 GHz and Vienna Airport at 2.45 GHz) could not be used for three reasons: (i) the antennas that were used did not support an azimuth coverage of full 360°, (ii) the antenna calibration data was erroneous, which makes angular estimation impossible (cf. Section 6.3.2), (iii) interference from WLAN devices, especially at the Vienna Airport. Hence, new and reliable measurements were essential. The purpose of the measurements was to support the evolution of the RCM, and finally to parametrise the model accurately.

8.1 Planning of the campaign

The conducted measurement campaign was aimed to find the *cluster parameters* for the channel model. Finding clusters is a bold venture which has to be pursued very carefully, as shown in Part 2, since the mere assumption that clusters do exist might already bias the results. So, I planned to carry out measurements at standard environments, but chose the equipment to be able to identify clusters when there are some.

The measurements were be carried out with respect to following parameters:

- Environment parameters
  The environment parameters were chosen with focus on the channel model, w.r.t. indoor environments from offices to big halls.
  - Routes:
    Routes were chosen in an office environment (small office rooms, larger office rooms and office corridors), as well as medium-sized rooms (labs, meeting rooms, aulas) and in a big hall. Doors in corridors were usually open, other doors (to offices or labs) usually closed.
  - Furnishing:
    All environments were measured with their usual furnishing as the resulting parameters for the channel model should reflect the usual environments.
  - Moving people / Stationarity of the environment:
    The office environment was found to be static in general. Very rarely people were walking on the corridors. To account for this effect, some stationary and some non-stationary measurements were done. In medium-sized and large rooms, people were usually walking around, thus introducing time variance. So, it does not make sense to measure a stationary environment there as it would not reflect reality.

- System parameters
  - System bandwidth:
    To allow for best possible delay resolution we use the full bandwidth of 200 MHz which can be provided by the channel sounding equipment.
  - Center frequency:
    The model should be applicable at several center frequencies. So, we decided to measure the channel at 2.55 GHz and 5.25 GHz and compare the results.
  - Angular resolution:
    For accurate modelling we need to sound the environment with full azimuth field of view of 360° at both link ends. A good elevation field-of-view ($\pm 90°$) would be advantageous but is not essential. As we can then identify paths in the full angular domain, this makes the measurement results applicable to arbitrary antenna arrays, even other than 4x4.
  - Polarization:
    To describe polarization accurately, the measurements were be conducted using both, vertical and horizontal polarization directions, even though polarization finally was not considered in the model. Still, polarisation had to be used for estimating propagation paths (cf. Section 6.3.2).
  - Antennas:
    To be able to fulfil the requirements for angular resolution and polarisation, we needed to use omnidirectional dual-polarised antenna arrays which are calibrated accurately. Note that such antenna arrays usually have a large number of elements.
  - Time variance:
    We focused on the changes in the environment when moving one station or time-variance by an instationary channel (moving people). We did not focus on measuring Doppler shifts, because of the rather large array scan time, which is due to the large number of antenna elements at both link ends.

With this configuration I ensured to be able to identify and parametrise clusters, where they exist.

I recognised that there might be environments where I cannot find clusters, for several reasons: (i) especially in very small rooms there might be too many clusters, so I cannot distinguish between them, (ii) especially outdoors, cluster spreads might be too small, so only paths and not clusters would be identified, (iii) the high-resolution estimator is not able to identify a number of paths in a cluster because of insufficient dynamic range.

Anyway, this would *not* hurt the model, as the model uses *clusters as a tool to model the environment*, where clusters can have arbitrary sizes. Note that in the model clusters are again defined by propagation paths showing similar *statistics*.

Finally, it turned out that the automatic cluster identification was well able to find clusters in all these scenarios, which could all be modelled accurately using the RCM.

9 Measurement campaign

This chapter describes the parameters and environments of the performed indoor measurement campaign conducted within the cooperation of Elektrobit and Vienna University of Technology at the University of Oulu, Centre for Wireless Communications[1].

Measurements consist of following two measurement types performed at the two centre frequencies of 2.55 GHz and 5.2 GHz.

- Moving Tx/MS and fixed Rx/BS measurements to capture the time variance and movement of major propagation paths
- Spot measurements with fixed Tx/Rx to capture the time variance introduced by moving people

All measurements were conducted with omni-directional arrays allowing for full azimuth resolution within a field-of-view of 360°.

9.1 Measurement environments

Measurements were conducted at three different locations with many measurement routes. These locations differ in the room and corridor sizes to enable characterization of large- to small-room cluster behaviour. In the following I will denote the three different environments as "small" (S), "medium" (M) and "large" (L). Some measurements were with LOS between Rx and Tx, others with NLOS. These measurements should provide a sufficient set of measurement data to characterise these indoor environments. A summary of the measurement routes is given in Appendix A, details for every single route is provided in Appendix B.

Figure 9.1 shows the office-room environment (S) with small office rooms having plaster-board walls. Measurement routes were taken in several rooms and corridors. Plasterboard walls separated rooms and corridors, while external walls were brick built.

In Figure 9.2 the measurements will be conducted in the lecture halls which are medium-size rooms (M). Again plasterboard walls separated the larger rooms from the corridors, and brick walls served as external walls.

Figure 9.3 shows the large entrance hall of CWC (L), measurements shall be used here to capture large delays. In this scenario, only brick walls occured.

[1] At this point I want to acknowledge Lassi Hentilä (Elektrobit), Veli-Matti Holappa (CWC) and Mikko Alatossava (CWC) for their great help with the measurements. With respect to their significant assistance, this chapter is written using the plural pronoun "we" instead of "I".

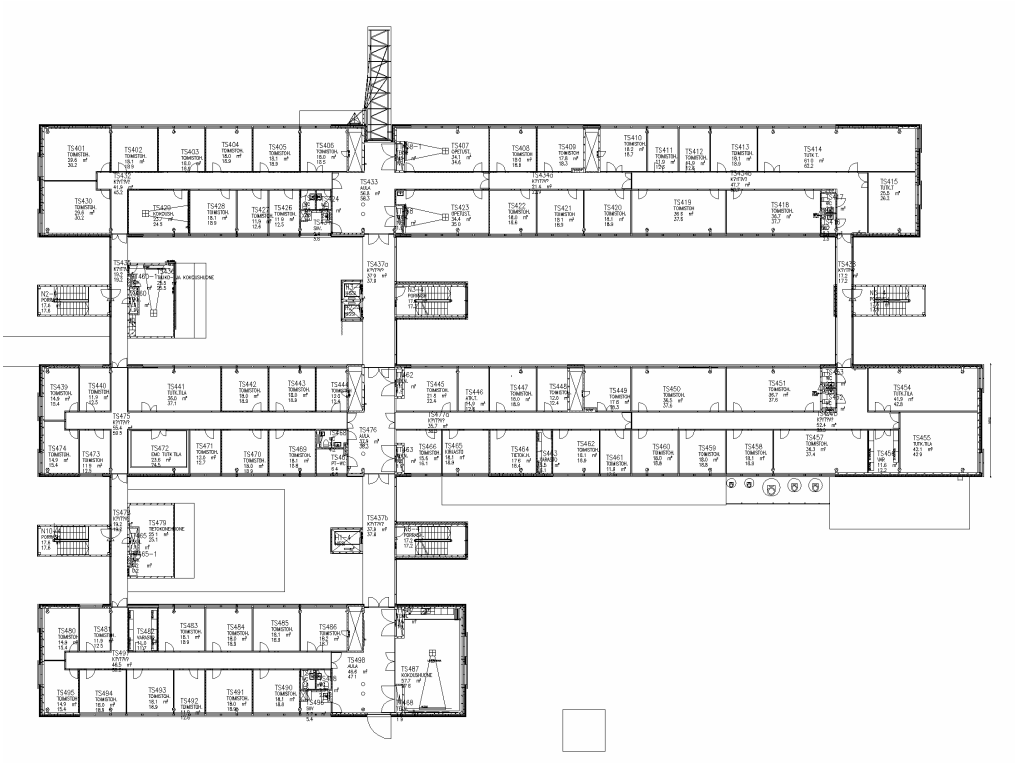


Figure 9.1: “Small” environment — Office rooms and corridors

Figure 9.2: “Medium” environment — Lecture rooms and aulas

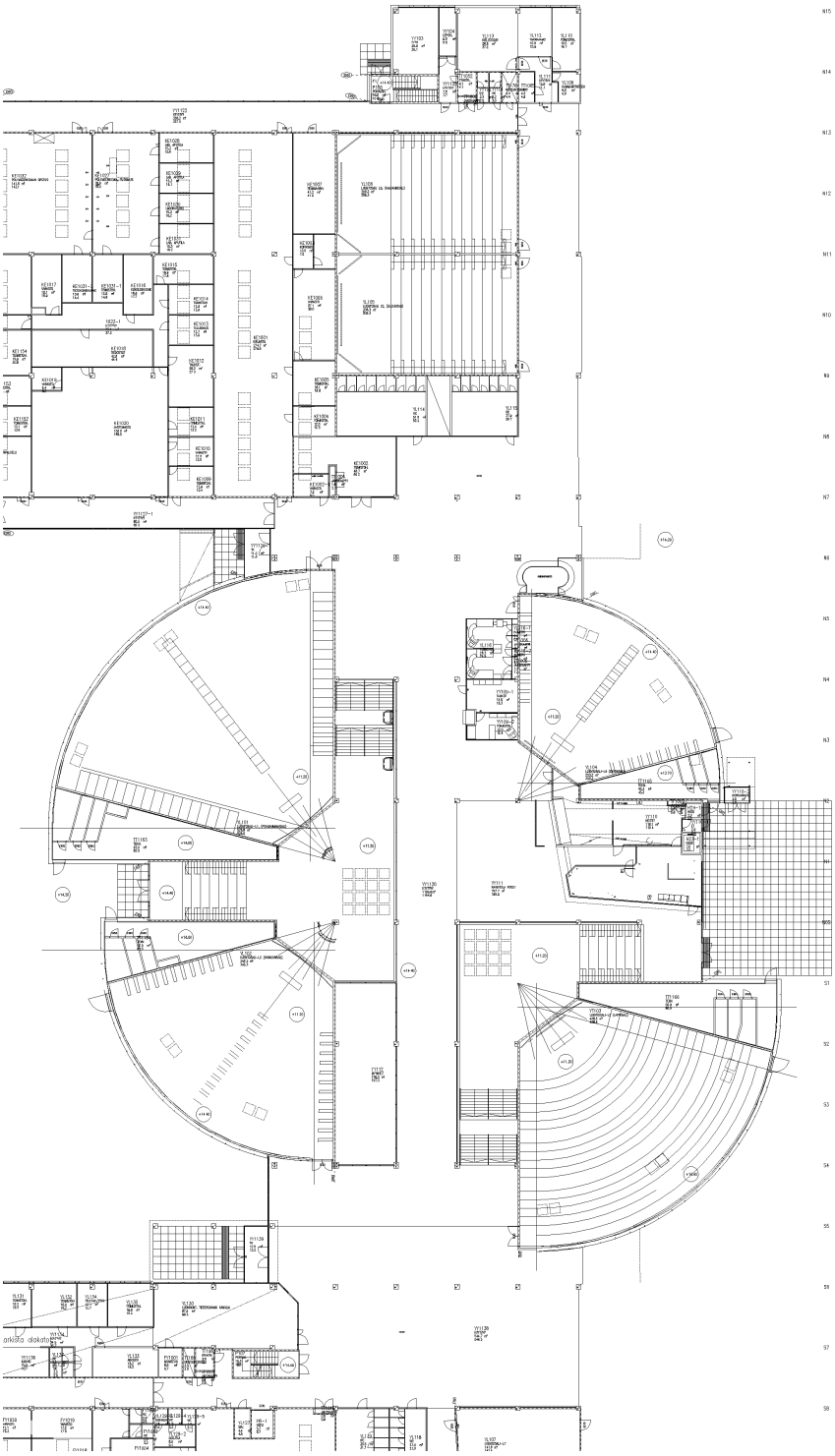


Figure 9.3: “Large” environment — Big hall

(a)

(b)

Figure 9.4: Measurement equipment, (a) author pulling Tx trolley with mounted antenna, (b) Rx trolley with detached antenna operated by Mikko Alatossava (CWC)

9.2 Measurement equipment

An Elektrobit PropSound CS^{TM} channel sounder was used to conduct the measurements. A description of this wide-band MIMO channel sounder can be found at [115].

The Tx and Rx modules of the sounder were assembled on two trolleys (see Figure 9.4). The Tx antenna was mounted on the trolley. For the Rx, a 10 m RF cable was used to place the antenna away from the rather large Rx equipment.

The following paragraphs provide details about the sounder modules and settings, and other equipment.

9.2.1 Radio modules

For using two carrier frequencies, two sets of transmitter and receiver modules were used in this campaign.

5.25 GHz:
TRU 5.1. . . 5.9 GHz
RRU 5.1. . . 5.9 GHz

2.55 GHz:
TRU 1.7. . . 2.7 GHz Dual frequency module
RRU 1.7. . . 2.7 GHz Dual frequency module

9.2.2 Sounder settings

Table 9.1 shows the setting used for sounding the environments at the two frequencies.

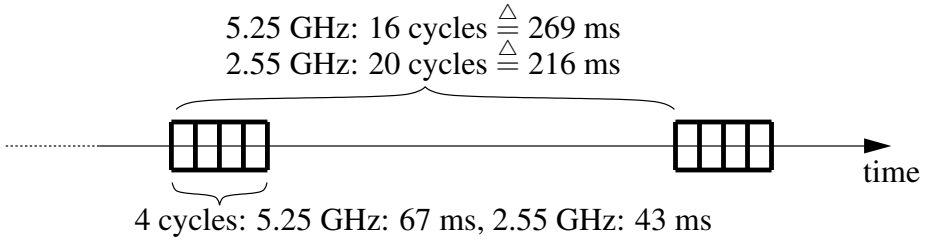


Figure 9.5: Burst mode: One bursts consist of 4 cycles, Burst period is 16 or 20 cycles, for 5.25 GHz and 2.55 GHz respectively.

Note that the null-to-null measurement bandwidth was 200 MHz for both carrier frequencies. At 2.55 GHz this still led to in-band interference from WLAN systems operating at 2.45 GHz. For this reason, we sometimes experienced a low dynamic range or even blocking of the receiver.

In order to cope with the large data rate we needed to use the burst mode (see Figure 9.5). We chose the burst period to be as small as possible to capture the time-variance of the channel as good as possible. The burst size was set to blocks of 4 cycles to allow for accurate Doppler estimation up to a maximum Doppler shift.

9.2.3 Antennas

To capture the full spatial structure of the radio channel, omni-directional antenna arrays are essential. Sounding was done sequentially between all antenna elements where the antenna switches were mounted on the antenna mast. The Tx switch supported 56 elements, the Rx switch 32 elements. Further details are given in Table 9.1.

The full azimuth field-of-view could be ensured at all link ends, unfortunately the elevation field-of-view was very limited.

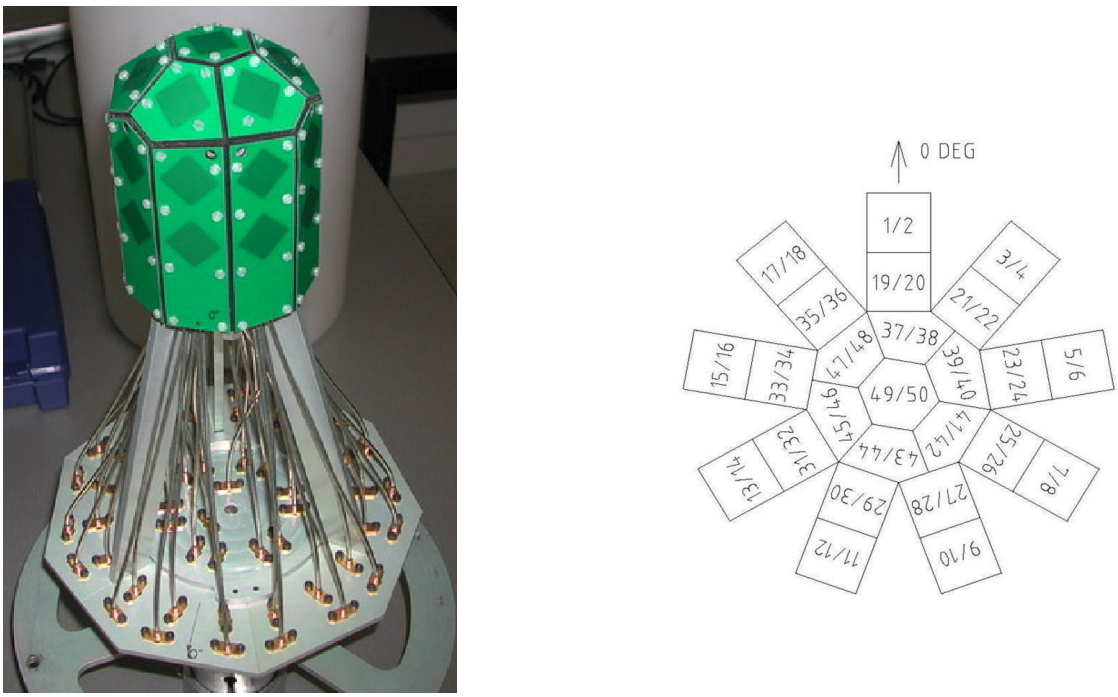
From all antenna arrays available we chose following sets for the different frequencies.

9.2.3.1 Antenna set 1: 5.25 GHz

Transmitter	
Antenna designation	**2x9ODA_5G25_T1**
Frequency / Bandwidth	5.25 GHz / 420 MHz
Radiation	$\pm 180°$ azimuth, -70°. . . +90° elevation
Antenna type	Dual polarized ($\pm 45°$) patch array, 50 elements (2x25)
Arrangement of elements	2 rings of 9 elements, slanted ring of 6 elements plus 1 element on top
Antenna elements in use	All elements
Receiver	
Antenna designation	**2x9ODA_5G25_T2**
Frequency / Bandwidth	5.25 GHz / 420 MHz
Radiation	$\pm 180°$ azimuth, -70°. . . +90° elevation
Antenna type	Dual polarized ($\pm 45°$) patch array, 50 elements (2x25)
Arrangement of elements	2 rings of 9 elements, slanted ring of 6 elements plus 1 element on top
Antenna elements in use	1. . . 18, 37. . . 50, i.e. lower ring + slanted ring + top

Basic Parameters			
Center frequency	**5.25 GHz**	**2.55 GHz**	
Transmit power	+26 dBm	+26 dBm	ALC / AGC enabled, FE att. 0 dB
Bandwidth	200 MHz	200 MHz	null-to-null
Chip rate	100 Mchip/s	100 Mchip/s	
Sampling frequency	200 MHz	200 MHz	
Measurement distance			
Code length	255 chips	255 chips	
Measurable excess delay	2.55 μs	2.55 μs	
Measurable excess distance	765 m	765 m	
Spatial resolution parameters			
Number of TX antenna elements	50	56	1. . . 50 for 5G25 and 1. . . 56 for 2G55
Number of RX antenna elements	32	8	1. . . 18, 37. . . 50 for 5G25 and 1. . . 8 for 2G55
Number of channels	1600	448	
Fast switching	Not used	51 chips	
Array scan time	8.42 ms	1.54 ms	
IR resolution parameters			
Channel sample rate (trigger rate)	59.4 Hz	92.62 Hz	
Maximum Doppler shift	29.7 Hz	46.3 Hz	
Data storage parameters			
Burst mode	Yes	Yes	
Burst length	4 cycles	4 cycles	
Burst time	67 ms	43 ms	
Burst way	3.3 cm	2.2 cm	
Burst period	16 cycles	20 cycles	one burst in 0.14 m for 5G25 and 0.11 m for 2G55
Storage data rate	19.1 MB/s	9.1 MB/s	
Other parameters			
Mobile speed	0.5 m/s	0.5 m/s	The Tx was moved at roughly
Scatterer speed	0.5 m/s	0.5 m/s	constant pace
Tx antenna height	1.53 m	1.53 m	
Rx antenna height	1.05 m	0.82 m	

Table 9.1: Sounder parameters for both carrier frequencies



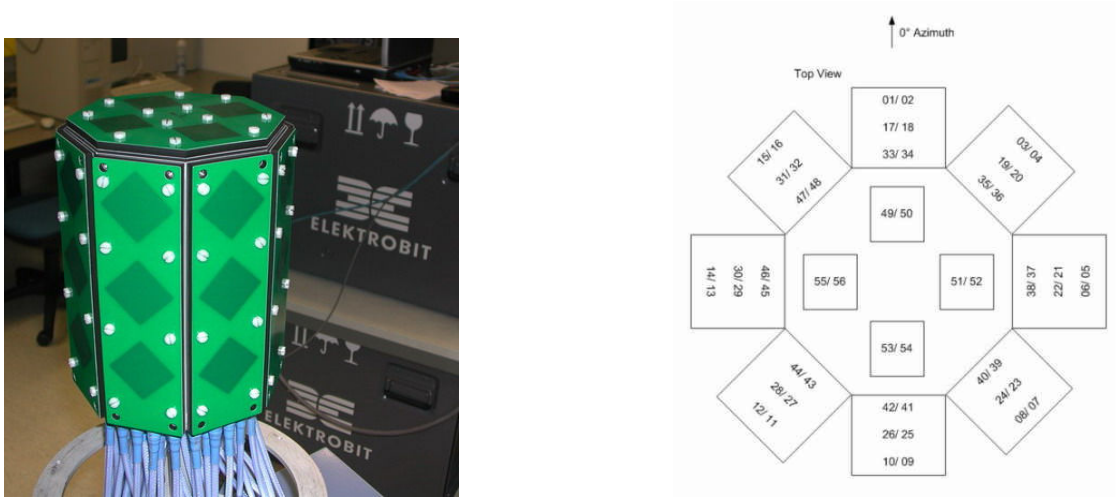
(a) (b)

Figure 9.6: Omni-directional patch arrays used at 5.2 GHz at both, Tx and Rx, (a) photograph, (b) schematic plot

Figure 9.6a shows a photograph of the antenna array. A schematic plot is shown in figure 9.6b.

As only 32 elements were available at the Rx switch, we had to choose a useful subset of the available antennas. There are two principal options for pruning antennas: polarization or number of patches. As the antennas had their polarisation axes in $\pm 45°$, we could not omit one polarisation, but we had to chose between patches.

The first idea was to exclude the lower ring, but the antenna array in use seemed to be not working correctly or was calibrated incorrectly on at least one of the elements of the upper ring. So we had to leave out the upper ring instead. However, there is one drawback: the antenna spacings between lower and slanted rings are larger than half a wavelength, leading to side-lobes in the vertical antenna patterns and in turn to unreliable estimation of the elevation. Because of this, we cannot estimate the elevation of arrival.



(a) (b)

Figure 9.7: Omni-directional patch array used as Tx at 2.55 GHz, (a) photograph, (b) schematic plot

9.2.3.2 Antenna set 2: 2.55 GHz

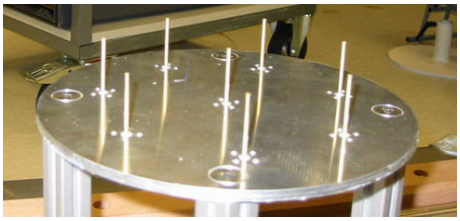
Transmitter	
Antenna designation	**3x8ODA_2G45_T1**
Frequency / Bandwidth	2.45 GHz / 200 MHz
Radiation	$\pm 180°$ azimuth, -70°. . . +90° elevation
Antenna type	Dual polarized ($\pm 45°$) patch array, 56 elements (2x28)
Arrangement of elements	3 rings of 8 elements, plus 4 elements on top
Receiver	
Antenna designation	**7+1_2G45_T3**
Frequency / Bandwidth	2.45 GHz / >200 MHz
Radiation	$\pm 180°$ azimuth, 0°. . . +90° elevation
Antenna type	Vertically polarized circular monopoles, 7+1 elements
Arrangement of elements	7 in a circle and one in the middle

Figures 9.7 and 9.8 show photographs and schematic plots of the antennas.

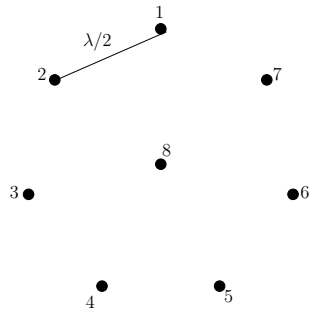
Only one omni-directional patch array was available for the time being, so we chose a uniform circular monopole array as the Rx, which also offers accurate azimuth resolution, but limited elevation resolution. Hence, we also skipped the elevation estimation for this antenna. As the UCA only offers vertical polarisation, we have to omit horizontal polarisation in the evaluations. Still, a comparison between 5.2 GHz and 2.55 GHz should be possible with respect to clusters.

9.2.4 Additional tools

Following additional tools were available for performing measurements faster and more conveniently:



(a)



(b)

Figure 9.8: Uniform circular monopole array used as Rx at 2.55 GHz, (a) photograph, (b) schematic plot

- Walkie-Talkies for quick communications
- Digital camera
- Multiple power outlet
- Duct tape to mark positions on the floor (and for a.o.b)
- Laser measurement device for distance measurements

9.3 Campaign practises

Following section describes the practises of the measurement campaign. By extensive planning and preparations the down time could be minimized.

9.3.1 Input files for the measurements

Following input files for the equipment were prepared before the measurements:

- Antenna calibration files
- PropSound settings
- Measurement description

PropSound settings for each measurement case are specified in Table 9.1. Input files were prepared beforehand based on the specified parameters. Required additional or changed information was filled in during the campaign. Note that this document only contains the final settings which were actually used.

9.3.2 Distance calibration

One LOS measurement with instantaneous evaluation determines the base delay between Tx and Rx at a fixed distance.

9.3.3 Field notes

For every measurement position field notes were taken including following points:

- Exact Rx and Tx positions/routes and waypoints were marked in the floor plan. The exact positions were measured with the laser-meter.
- Photographs of the measurement route as well as important objects. Photographs start with a screen shot showing the number of the current measurement to distinguish photographs afterwards easily.
- Nature of channel stationarity (moving people/objects?)
- Changes of the plan
- Exceptional happenings
- Errors
- Etc.

Field notes were filled in by pencil and paper in English language. After the measurements the field notes were scanned to the same directories as the measurement documentation.

9.3.4 Calibration

Calibration was done before and after every Rx position. Furthermore, the channel impulse response was checked at the sounder display at the start and end of each route, to decide if it is worth recording the specific channel (interference, noise floor).

9.3.5 Measurement

For each environment there were several locations for the BS (Rx) and several locations or routes for the MS (Tx). They are defined in more detail in Section A. Every route or measurement point was identified by a measurement index and produced an own data file. Calibration was performed before and after all measurements for a single Rx position.

While taking the measurements the trolley was moved at a slow pace with approximately constant speed. When moving around corners, we made sure to maintain the antenna orientation. As measurements were conducted at two carrier frequencies but same routes, these routes had to be measured taking care of very high location accuracy.

10 Final remarks

This last chapter is dedicated to my thoughts about future topics in this field of radio channel modelling.

Interference modelling

Communication technologies on the rise, like WiMAX, UMTS and (on the longer run) the long-term evolution (LTE) of 3GPP will all have to deal with a crowding of their networks. By this, the interference increases significantly. Of course, algorithms on the network layer will need to deal with this problem.

To design such algorithms, accurate channel models for the expected interference levels are required. Going to MIMO communications, this interference might even be correlated, which enables or simplifies interference mitigation.

There is currently a strong need for interference models validated by measurements, but, to my knowledge, only one measurement campaign has been conducted so far focusing on this problem [116]. More measurements and significant models are a key topic of future research.

Car-To-X communications

A long-lasting dream of road makers, car providers, and governments are the automatic road. Cars driving on their own, avoiding traffic jams, warning each other against dangers.

The technology is not that far, yet. An essential requirement to come one step closer to this goal is to communicate with the car from a station ("Car-to-Infrastructure"), and to enable communications between cars ("Car-to-Car"). Communication systems for this purpose face the problem of high mobility and high speeds (resulting in large Doppler shifts).

First of all, for developing robust communication systems, it is paramount to know the ultimate performance limits. Also, models for this high-mobility channel will become vital.

A number of measurements were already conducted for Car-to-Car and Car-to-Infrastructure channels [90, 91, 117, 118, 119]. Finding the ultimate performance bounds and creating significant channel models is an ongoing challenging process that has already started.

Sensor networks, cooperative distributed MIMO

An interesting idea from people in information theory is to distribute a large number of very small radio devices. When playing with the assumptions, e.g. that these devices can all act like in a big MIMO scheme, system capacity rockets high. However, these assumptions are currently irrational.

With a more realistic view, these self-organising systems are robust against the failure individual stations. Data packets can simply be diverted via other routes through the wireless network. Another aspect is that an infrastructure, like base stations, becomes redundant. This system scenario (longed for by users, but feared by mobile service providers), might bring another leap in radio technology.

An enabling requirement for these distributed systems is a proper scheduling and routing protocol. When assuming that each device wants to individually transmit a piece of information from one point to another in this wireless multi-hop network, the first question is where to transmit the packet. A good strategy is to transmit the packet into the direction of the receiver, but use a good channel with low bit error ratio.

So, to find the shortest and most beneficial path for a packet to reach its destination, information about the radio channel is paramount. In this case, it is also interesting to know whether the channels of different transmit paths are correlated. This is another new challenge of future wireless channel modelling.

Will we need radio channel modelling in the future?

Everything goes wireless. Since the market for voice communications is nearly saturated in Europe, the providers are looking hard for the next “killer application”. It seems that they identified wireless Internet as the next hype. Online, everytime, everywhere[1]. By the increasing need for bandwidth, the wireless systems need to operate close to their ultimate limits.

Thanks to Shannon, we know that these limits are *always* determined by the underlying channel. New transmission technologies, entering the last domains unexploited yet, will rely on accurate future channel models that accurately reflect the propagation scenarios.

There are still many open questions, still enough to do, and still a long way to go.

[1] When you’re in an airplane somewhere above the Atlantic Ocean, Skyping with your dad and trying to fix some problem on his computer using a remote administrator, and you’re even enjoying this, you know that you chose the right job in wireless communications.

Appendix A

Measurement maps & routes

A.1 Legend and notation

The measurement maps are shown on the following pages. While taking measurements, on-site maps were filled out with the exact Rx and Tx positions and routes.

The following floor plans (Figures A.1-A.6) give an overview over the measured routes. The exact location of the Rx and routes can be found in the documentation pages for each individual route (cf. Appendix B).

Moving measurements are named by their Tx route (TxR##), where ## denotes the route number. Every route number corresponds to a unique Rx position.

Waypoints are indicated by circles on the routes. When moving around corners, waypoints are set on every corner. On a straight route, waypoints were chosen to fit expected changes in the environment.

Spot measurements were conducted where significant changes in the channel could be expected by moving people. These spots are denoted as TxR##Pm, where m denotes the spot number.

Some routes were measured from different Rx positions, where special focus on location accuracy was taken.

A.2 Overview maps

The following pages provide an overview of the measurement routes. Exact locations and details for every single route can be found in Appendix B.

Figures A.1-A.4 describe the measurement positions for the small-size environment. Each map corresponds to an Rx position. People were working in the office rooms, but did not move significantly. The corridors were mostly empty. Hence, we chose to perform only moving measurements (except for one spot measurement) as the scenario was predominantly static.

Figure A.5 outlines the measurements at the medium-size environment. For each route the Rx position is indicated. We decided to measure different room types (student laboratory, computer room, small hall, cafeteria) with arbitrary Rx positions. As we encountered interference from WLAN systems, some routes could only be taken at 5.25 GHz.

Figure A.6 summarizes the measurements at the big-hall environment. The main hall of Oulu University was usually crowded with people moving along the main corridor. Hence, we chose to conduct both, moving and spot measurements.

Figure A.1: Routes for Rx1

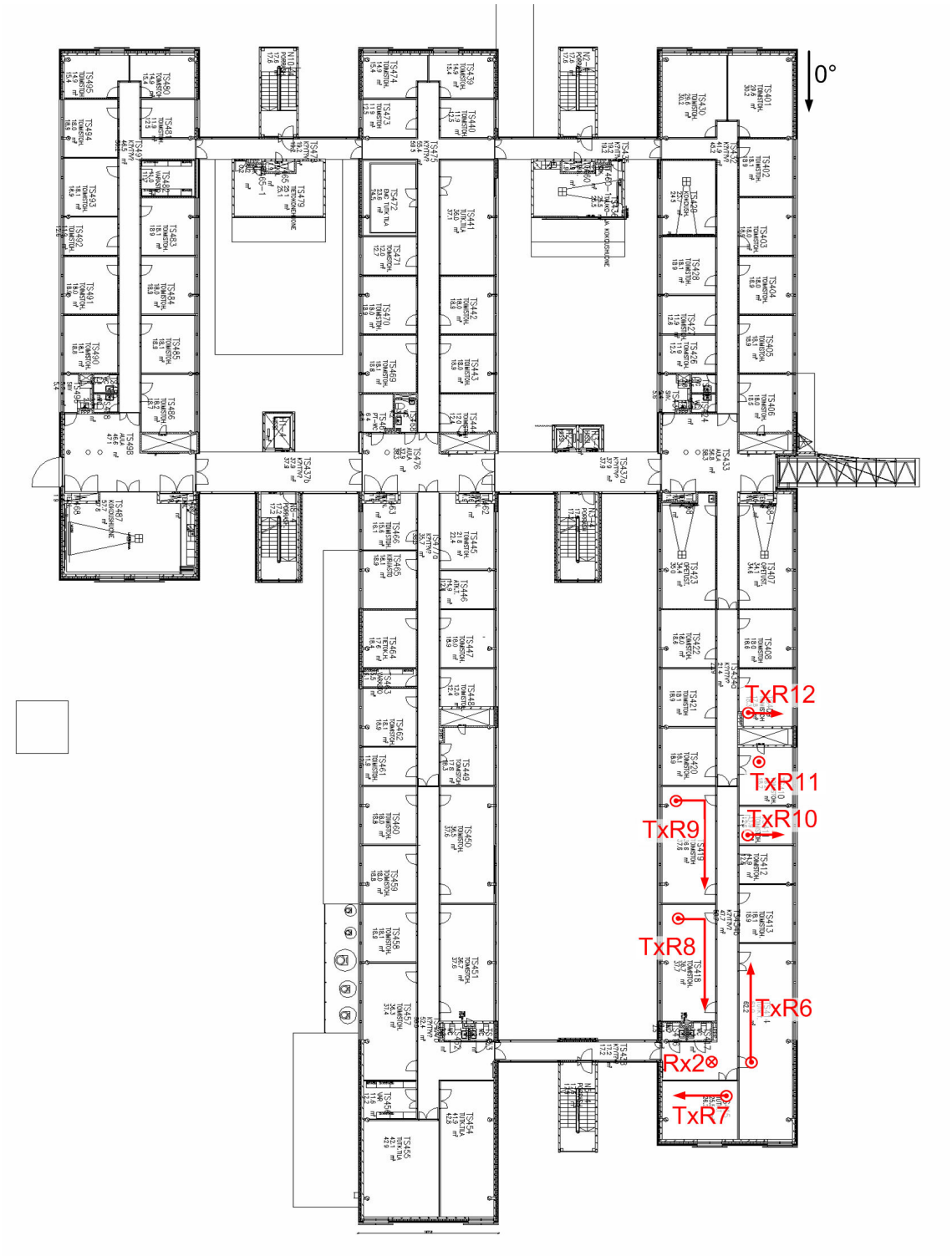


Figure A.2: Routes for Rx2

Figure A.3: Routes for Rx3

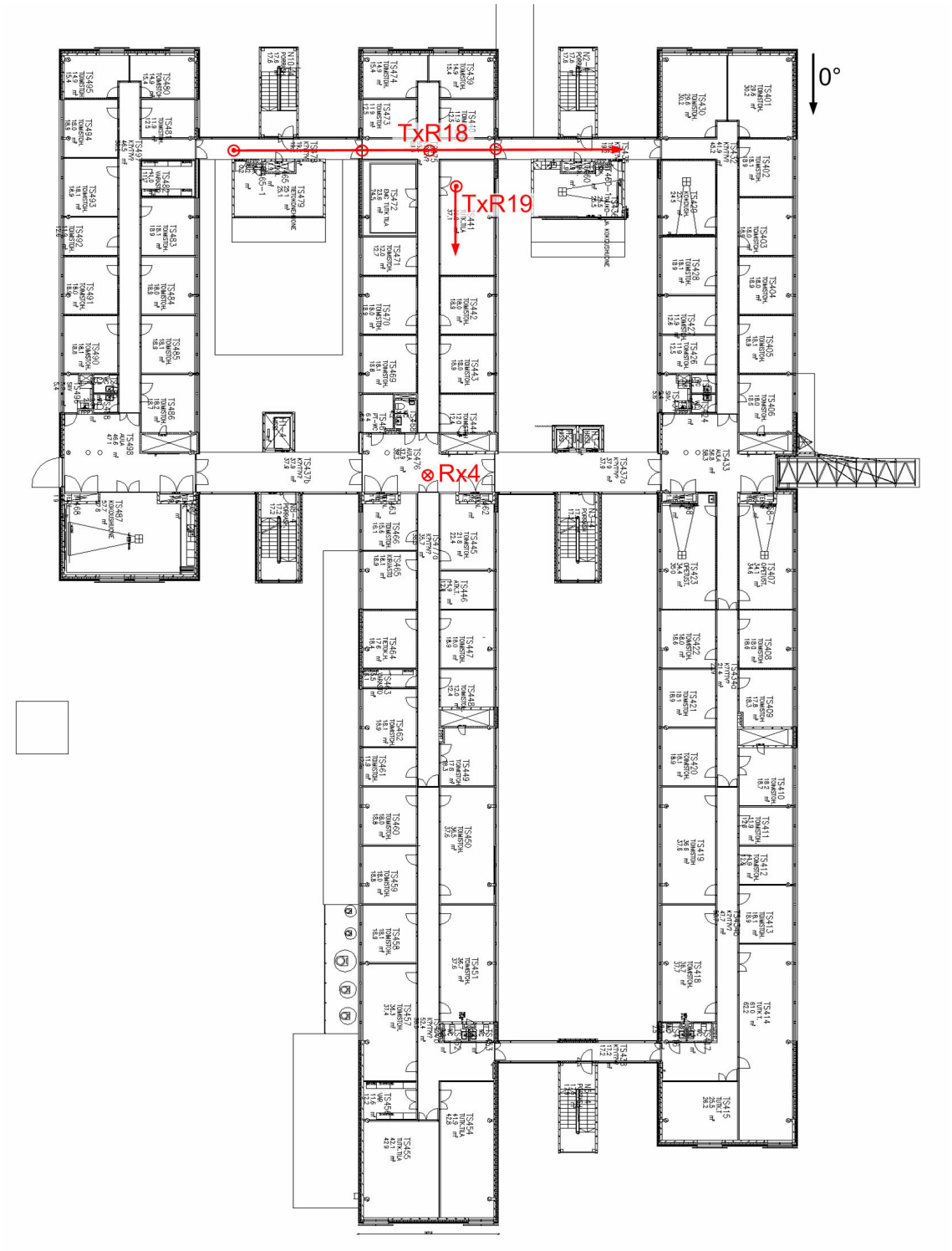


Figure A.4: Routes for Rx4

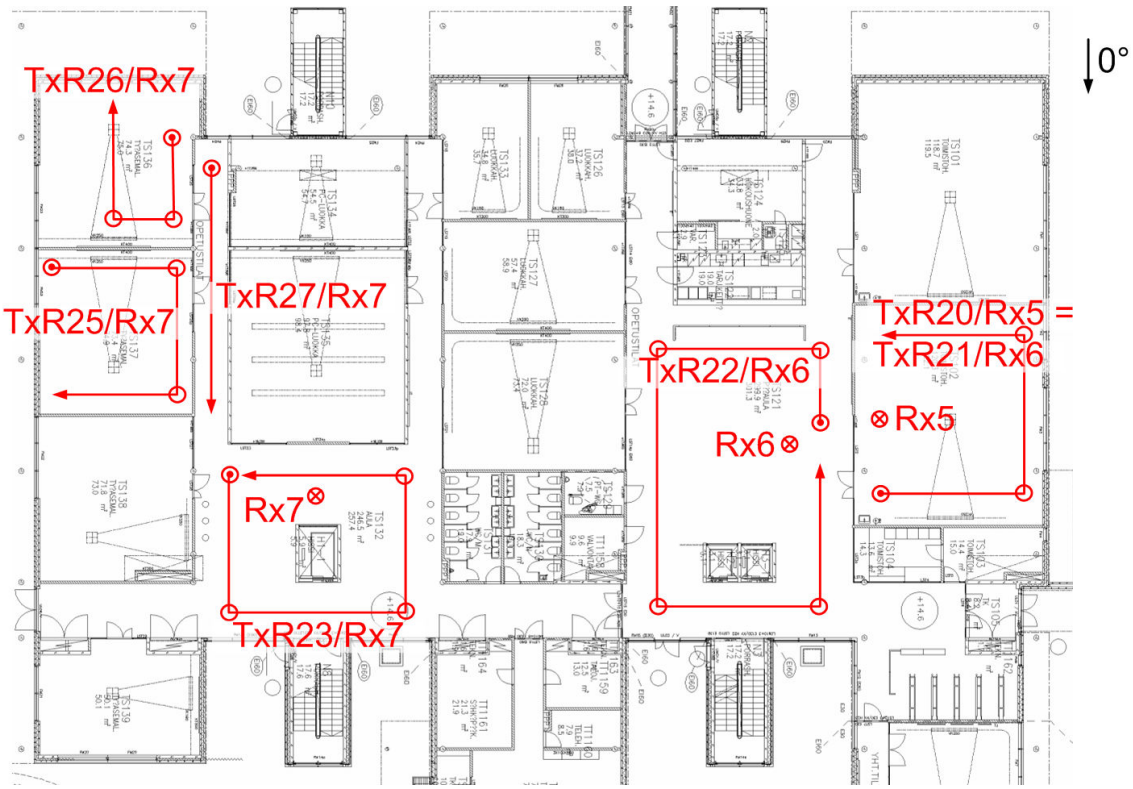


Figure A.5: Routes for Rx5 to Rx7

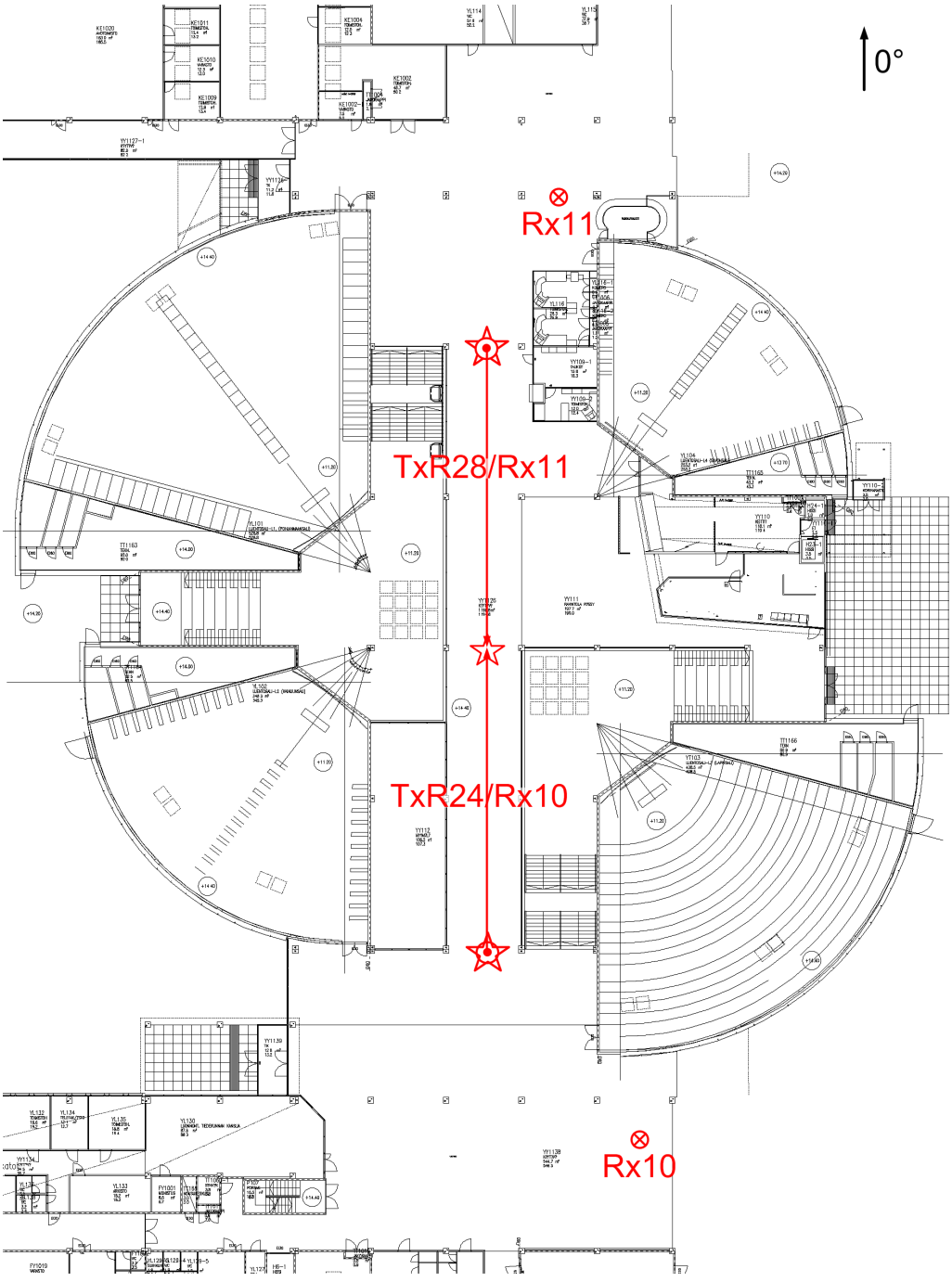


Figure A.6: Routes for Rx10 to Rx11

Appendix B

Route maps

The following pages detail the measurement routes done at Oulu University.

Distance_255

General Information

Tx Position/Route:	Distance_255
Rx Location:	Rx1
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

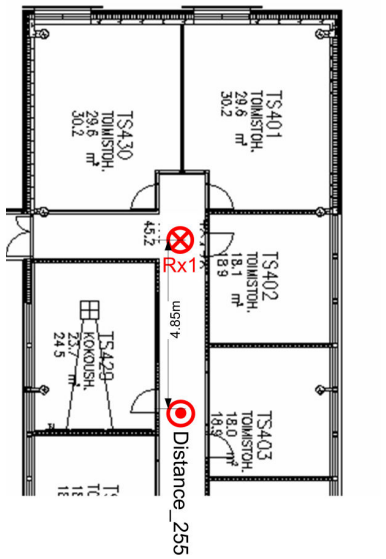
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	2	
Date	2005-11-22	
Front-End Att.	26 dB	
Calibration Att.	70 dB	
Cal. Index (before)	1	
Cal. Index (after)	8	

Tx Environment

General Notes:

Distance = 4.85 m



Route Map

TxR1

General Information

Tx Position/Route:	TxR1
Rx Location:	Rx1
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

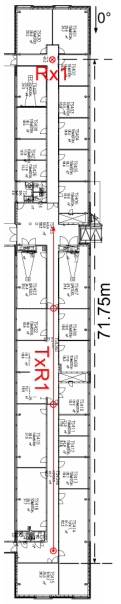
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	3	42
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	1	41
Cal. Index (after)	8	48

Tx Environment

General Notes:

Long corridor, dorrs open and close while passing through them. Small parts of the doors were kept open No people on corridor



Route Map

TxR2

General Information

Tx Position/Route:	TxR2
Rx Location:	Rx1
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	4	44
Date	2005-11-22	2005-11-24
Front-End Att.	20 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	1	41
Cal. Index (after)	8	48

Tx Environment

General Notes:

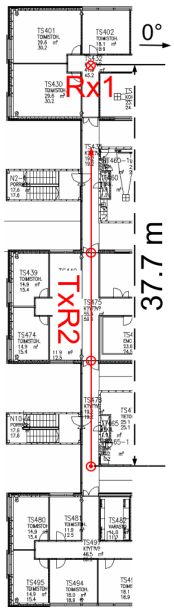
Same route like TxR18. Narrow corridor, open doors, no people.

Special Notes 2.55 GHz:

Invalid waypoints

Special Notes 5.25 GHz:

waypoints correct. (Note: Index 43 had incorrect antenna direction)



Route Map

TxR3

General Information

Tx Position/Route:	TxR3
Rx Location:	Rx1
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	5	45
Date	2005-11-22	2005-11-24
Front-End Att.	20 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	1	41
Cal. Index (after)	8	48

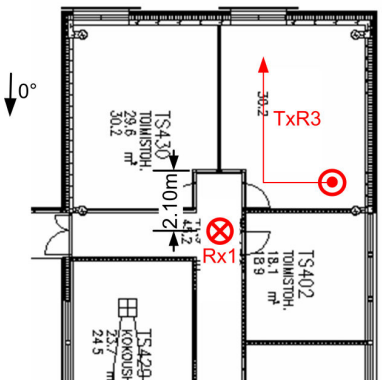
Tx Environment

General Notes:

Furnitured room, no people, door closed

Special Notes 2.55 GHz:

Waypoints in index 4!



Route Map

TxR4

General Information

Tx Position/Route:	TxR4
Rx Location:	Rx1
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

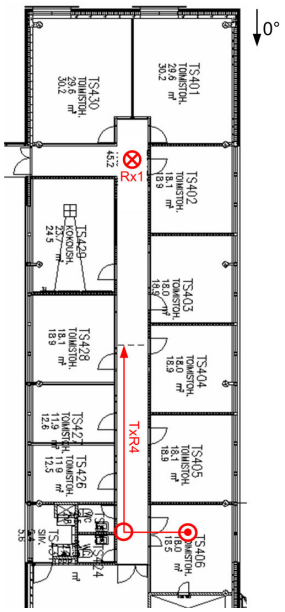
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	6	46
Date	2005-11-22	2005-11-24
Front-End Att.	20 dB	10 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	1	41
Cal. Index (after)	8	48

Tx Environment

General Notes:

NLOS → LOS, Furnitured room going out to corridor. Tx was close to one wall. No people, door open



Route Map

TxR5

General Information

Tx Position/Route:	TxR5
Rx Location:	Rx1
Tx Orientation:	180°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

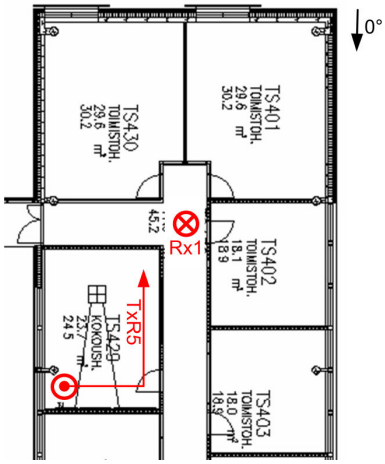
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	7	47
Date	2005-11-22	2005-11-24
Front-End Att.	20 dB	10 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	1	41
Cal. Index (after)	8	48

Tx Environment

General Notes:

Meeting room, tables and chairs in rectangular setting in the middle of the room. Glass wall with shades to the corridor. Shades closed, door closed. Rx equipment on corridor



Route Map

TxR6

General Information

Tx Position/Route:	TxR6
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

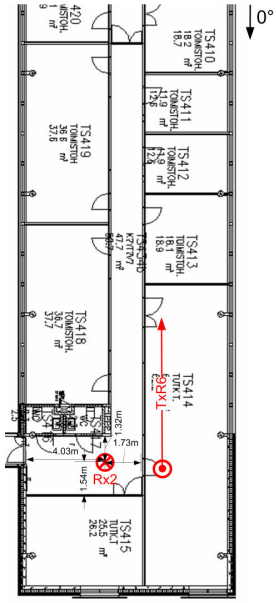
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	10	50
Date	2005-11-22	2005-11-24
Front-End Att.	20 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58

Tx Environment

General Notes:

Storage room, metal safes along one wall (∼90°), metal shelf on other wall (∼180°). Doors closed



Route Map

TxR7

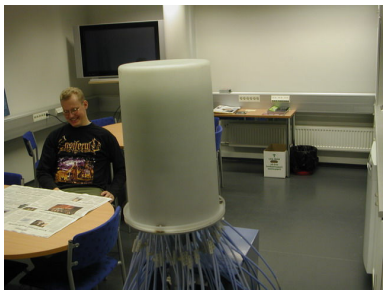
General Information

Tx Position/Route:	TxR7
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	11	52
Date	2005-11-22	2005-11-24
Front-End Att.	20 dB	10 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58



Tx Environment

General Notes:

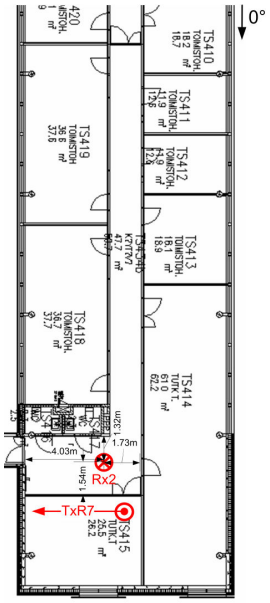
Coffee room, door closed.

Special Notes 2.55 GHz:

person was passing Rx and opening door to room shortly

Special Notes 5.25 GHz:

(Note: Index 51 had overload)



Route Map

TxR8

General Information

Tx Position/Route:	TxR8
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	12	53
Date	2005-11-22	2005-11-24
Front-End Att.	10 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58

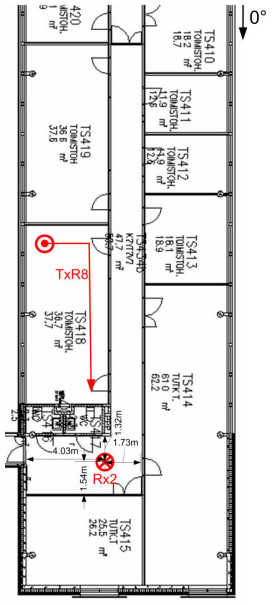
Tx Environment

General Notes:

Same route like TxR17. Office room, empty, rooms partitioned by book shelfs & desks, concrete wall at 0°, doors closed

Special Notes 5.25 GHz:

One person walking along corridor



Route Map

TxR9

General Information

Tx Position/Route:	TxR9
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

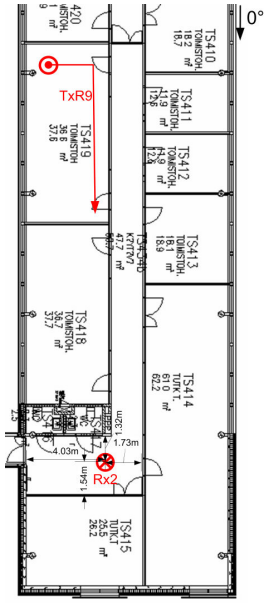
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	13	54
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58

Tx Environment

General Notes:

Same route like TxR16. Office room, empty, rooms partitioned by book shelfs & desks, doors closed



Route Map

TxR10

General Information

Tx Position/Route:	TxR10
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

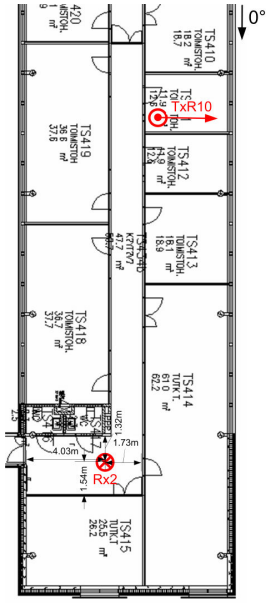
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	14	55
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58

Tx Environment

General Notes:

Small office room (Kimmo's office). Door closed, Kimmo working.



Route Map

TxR11

General Information

Tx Position/Route:	TxR11
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

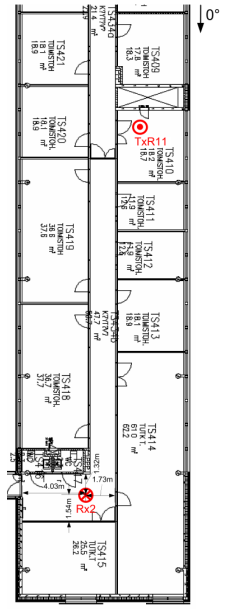
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	15	56
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58

Tx Environment

General Notes:

Stationary measurement. Small crowded room. People moving around Tx.



Route Map

TxR12

General Information

Tx Position/Route:	TxR12
Rx Location:	Rx2
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	16	57
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	9	49
Cal. Index (after)	17	58

Tx Environment

General Notes:

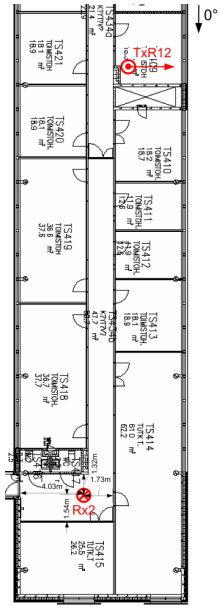
Tad Matsumoto's room. Office fully furnitured. (Note: pictures before spash-screen!)

Special Notes 2.55 GHz:

People working in the room. Door open, already low signal

Special Notes 5.25 GHz:

No people in the room. Door closed.



Route Map

TxR13

General Information

Tx Position/Route:	TxR13
Rx Location:	Rx3
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	19	60
Date	2005-11-22	2005-11-24
Front-End Att.	10 dB	10 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	18	59
Cal. Index (after)	25	66

Tx Environment

General Notes:

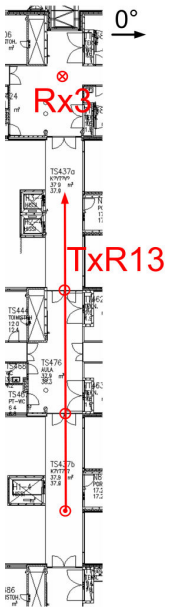
This route will be done twice (see TxR14). Doors open. In the middle, the corridor passes through a big hall. Rx is in office-corridor crossing.

Special Notes 2.55 GHz:

Person walking in the channel, opening door next to Rx.

Special Notes 5.25 GHz:

People in the channel



Route Map

TxR14

General Information

Tx Position/Route:	TxR14
Rx Location:	Rx3
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	20	61
Date	2005-11-22	2005-11-24
Front-End Att.	10 dB	10 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	18	59
Cal. Index (after)	25	66

Tx Environment

General Notes:

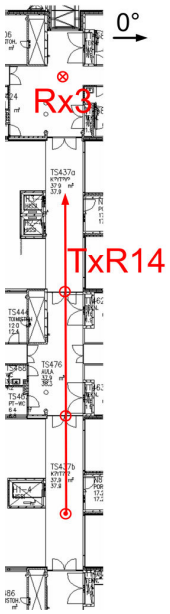
Second run (see TxR13)

Special Notes 2.55 GHz:

No persons in the channel

Special Notes 5.25 GHz:

People in the channel



Route Map

TxR15

General Information

Tx Position/Route:	TxR15
Rx Location:	Rx3
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

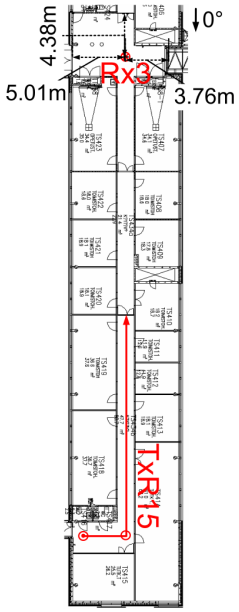
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	22	62
Date	2005-11-22	2005-11-24
Front-End Att.	10 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	18	59
Cal. Index (after)	25	66

Tx Environment

General Notes:

Corridor. First short time in NLOS, then LOS



Route Map

Special Notes 2.55 GHz:

(Note: Little IF overload at the end, Index 21 rejected because of much overload)

Special Notes 5.25 GHz:

Person in the channel

TxR16

General Information

Tx Position/Route:	TxR16
Rx Location:	Rx3
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

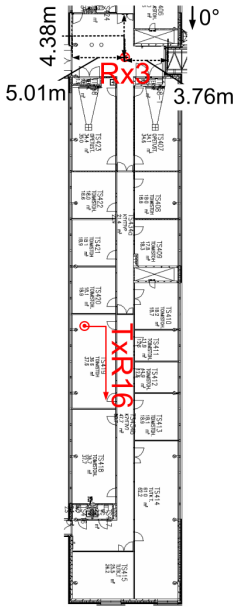
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	23	63
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	18	59
Cal. Index (after)	25	66

Tx Environment

General Notes:

Same route like TxR9. Office room, empty, rooms partitioned by book shelfs & desks, doors closed



Route Map

TxR17

General Information

Tx Position/Route:	TxR17
Rx Location:	Rx3
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

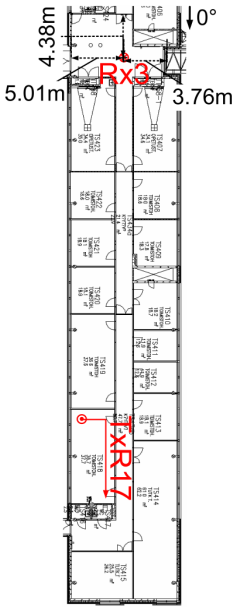
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	24	64
Date	2005-11-22	2005-11-24
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	18	59
Cal. Index (after)	25	66

Tx Environment

General Notes:

Same route like TxR8. Office room, empty, rooms partitioned by book shelfs & desks, concrete wall at 0°, doors closed



Route Map

TxR18

General Information

Tx Position/Route:	TxR18
Rx Location:	Rx4
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

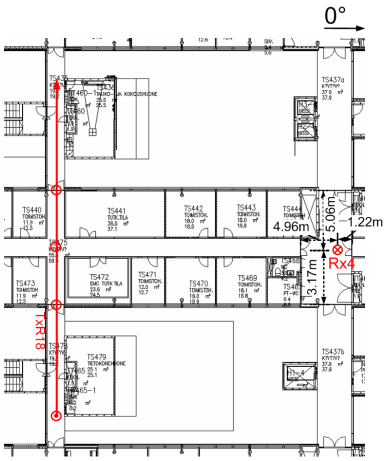
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	27	68
Date	2005-11-23	2005-11-24
Front-End Att.	16 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	26	67
Cal. Index (after)	29	70

Tx Environment

General Notes:

Same route like TxR2. Narrow corridor, open doors, no people. Note additional WP at LOS to the Rx



Route Map

TxR19

General Information

Tx Position/Route:	TxR19
Rx Location:	Rx4
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

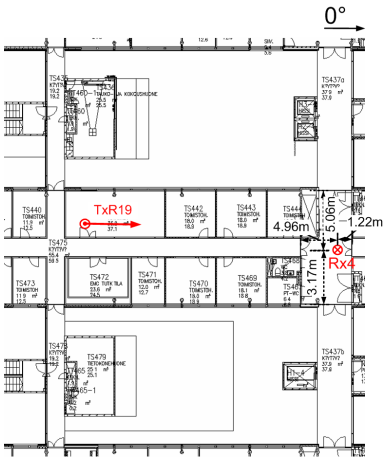
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	28	69
Date	2005-11-23	2005-11-24
Front-End Att.	16 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	26	67
Cal. Index (after)	29	70

Tx Environment

General Notes:

Furnitured room



Route Map

Special Notes 2.55 GHz:

One person working in the room, people on corridor.

Special Notes 5.25 GHz:

Three persons working in the room. People on corridor

TxR20

General Information

Tx Position/Route:	TxR20
Rx Location:	Rx5
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

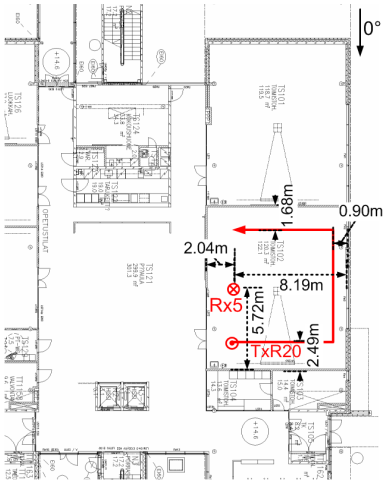
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	31	72
Date	2005-11-23	2005-11-25
Front-End Att.	24 dB	16 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	30	71
Cal. Index (after)	35	77

Tx Environment

General Notes:

Same route as TxR21. Student lab environment. Metal panels on eye-level. Rx in the same room. Precision of movement very accurate.



Route Map

Special Notes 2.55 GHz:

People working on some places

Special Notes 5.25 GHz:

One person working.

TxR21

General Information

Tx Position/Route:	TxR21
Rx Location:	Rx6
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

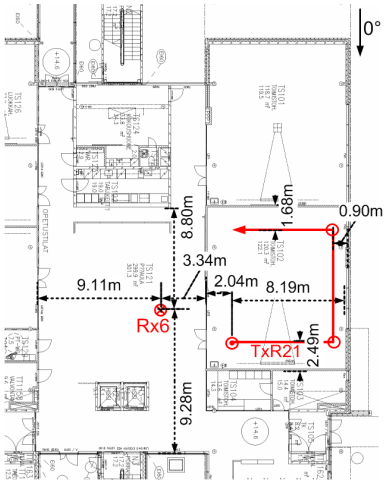
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	32	73
Date	2005-11-23	2005-11-25
Front-End Att.	16 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	30	71
Cal. Index (after)	35	77

Tx Environment

General Notes:

Same route as TxR20. Student lab environment. Metal panels on eye-level. Rx outside in a big hall. Door closed. Rx Antenna height = 1.67m. Precision of movement very accurate.



Route Map

TxR22

General Information

Tx Position/Route:	TxR22
Rx Location:	Rx6
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

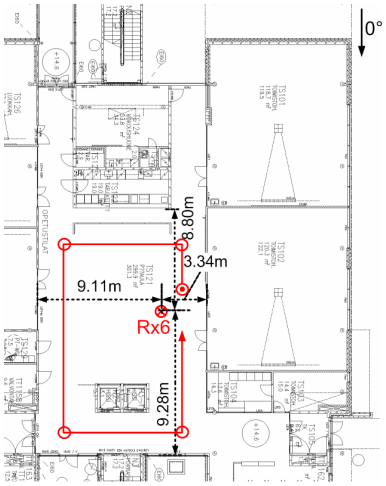
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	33	74
Date	2005-11-23	2005-11-25
Front-End Att.	24 dB	18 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	30	71
Cal. Index (after)	35	77

Tx Environment

General Notes:

Cafeteria, Rx going around Tx, with elevators shadowing sometimes. People sitting and walking around.



Route Map

TxR23

General Information

Tx Position/Route:	TxR23
Rx Location:	Rx6
Tx Orientation:	180°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

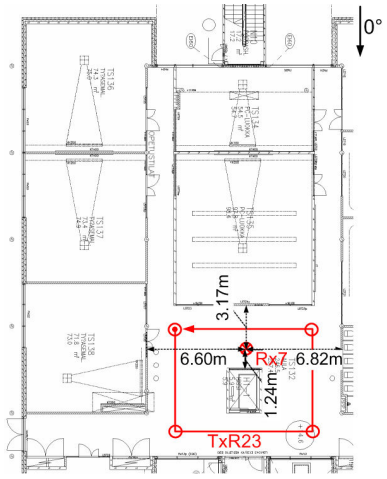
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	34	75
Date	2005-11-23	2005-11-25
Front-End Att.	30 dB	18 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	30	71
Cal. Index (after)	35	77

Tx Environment

General Notes:

Small aula, elevator shadows sometimes.



Route Map

Special Notes 2.55 GHz:

Some people passing. (Note: maybe there was interference at low frequencies!)

Special Notes 5.25 GHz:

No people.

TxR25

General Information

Tx Position/Route:	TxR25
Rx Location:	Rx7
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

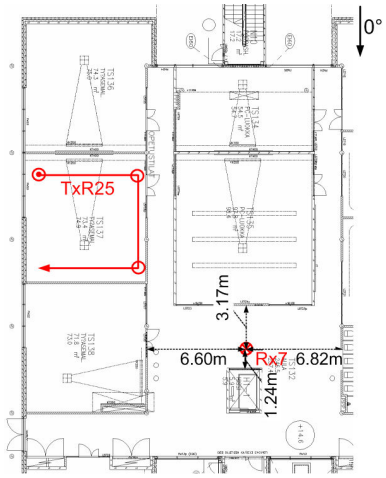
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index		76
Date		2005-11-25
Front-End Att.		0 dB
Calibration Att.		50 dB
Cal. Index (before)		71
Cal. Index (after)		77

Tx Environment

General Notes:

Computer room. no people, doors closed. Route was sometimes close to the wall



Route Map

TxR26

General Information

Tx Position/Route:	TxR26
Rx Location:	Rx7
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	VMH

Rx Environment

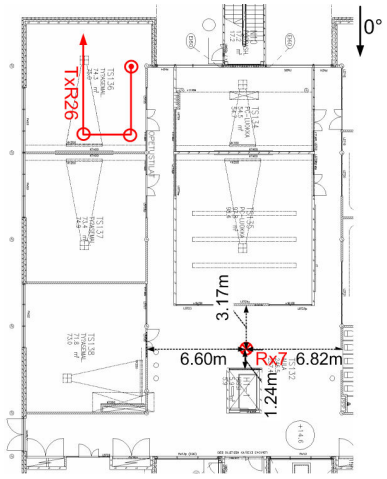
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index		78
Date		2005-11-25
Front-End Att.		0 dB
Calibration Att.		50 dB
Cal. Index (before)		77
Cal. Index (after)		80

Tx Environment

General Notes:

Computer room, Tables arranged in 3 clusters, door closed



Route Map

TxR27

General Information

Tx Position/Route:	TxR27
Rx Location:	Rx7
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	VMH

Rx Environment

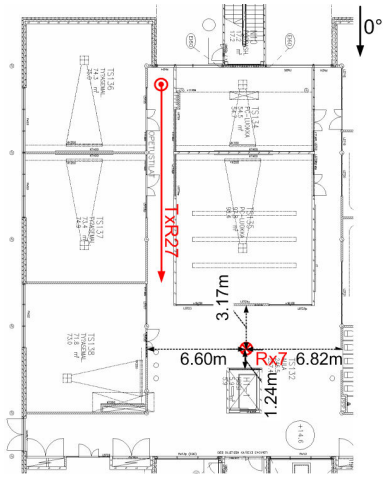
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index		79
Date		2005-11-25
Front-End Att.		0 dB
Calibration Att.		50 dB
Cal. Index (before)		77
Cal. Index (after)		80

Tx Environment

General Notes:

Corridor with NLOS to the Rx.



Route Map

TxR24

General Information

Tx Position/Route:	TxR24
Rx Location:	Rx10
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	37	82
Date	2005-11-23	2005-11-25
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	36	81
Cal. Index (after)	40	88

Tx Environment

General Notes:

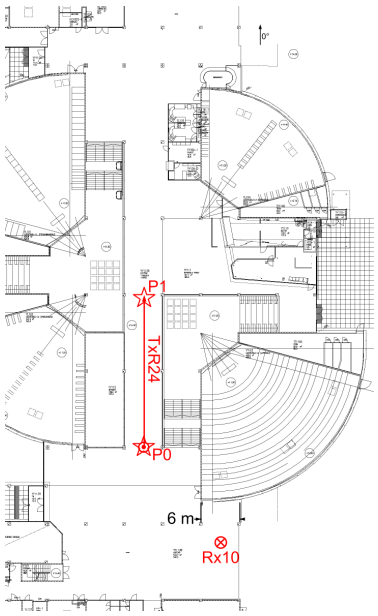
Oulu university main cafeteria (pizza slices). big hall

Special Notes 2.55 GHz:

crowded corridor

Special Notes 5.25 GHz:

corridor less crowded



Route Map

TxR24P1

General Information

Tx Position/Route:	TxR24P1
Rx Location:	Rx10
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

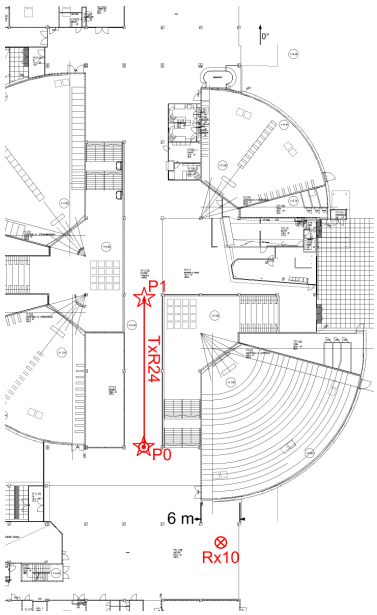
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	38	83
Date	2005-11-23	2005-11-25
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	36	81
Cal. Index (after)	40	88

Tx Environment

General Notes:

Stationary point, 1 min. See TxR24



Route Map

TxR24P0

General Information

Tx Position/Route:	TxR24P0
Rx Location:	Rx10
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

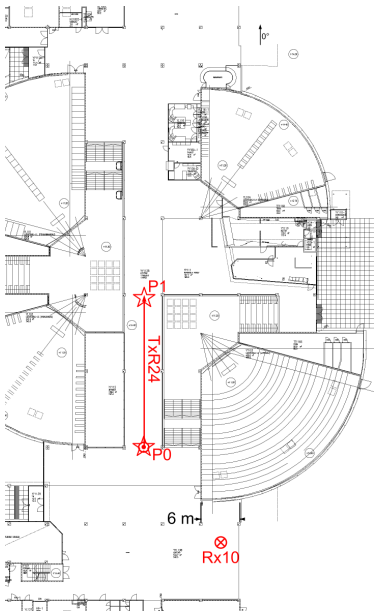
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index	29	84
Date	2005-11-23	2005-11-25
Front-End Att.	0 dB	0 dB
Calibration Att.	70 dB	50 dB
Cal. Index (before)	36	81
Cal. Index (after)	40	88

Tx Environment

General Notes:

Stationary point, 1 min. See TxR24



Route Map

TxR28

General Information

Tx Position/Route:	TxR28
Rx Location:	Rx11
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

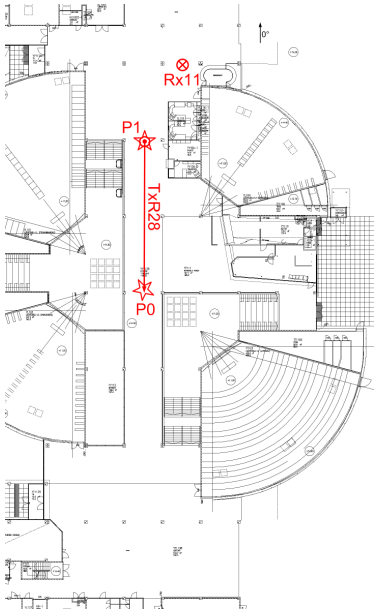
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index		85
Date		2005-11-25
Front-End Att.		0 dB
Calibration Att.		50 dB
Cal. Index (before)		81
Cal. Index (after)		88

Tx Environment

General Notes:

Oulu university main cafeteria (pizza slices). big hall. corridor sometimes croweded, sometimes not.



Route Map

TxR28P0

General Information

Tx Position/Route:	TxR28P0
Rx Location:	Rx11
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

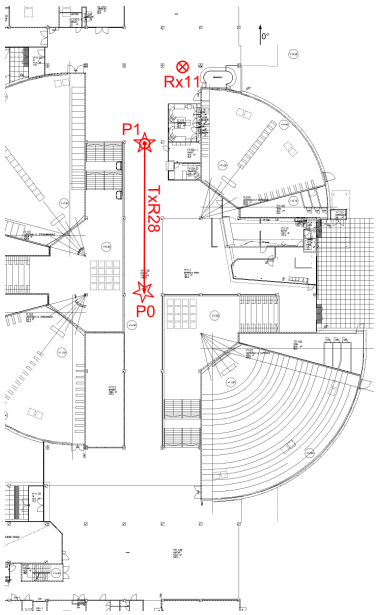
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index		86
Date		2005-11-25
Front-End Att.		0 dB
Calibration Att.		50 dB
Cal. Index (before)		81
Cal. Index (after)		88

Tx Environment

General Notes:

Stationary point, 1 min. See TxR28



Route Map

TxR28P1

General Information

Tx Position/Route:	TxR28P1
Rx Location:	Rx11
Tx Orientation:	0°
Rx Orientation:	0°
Initials:	Cz

Rx Environment

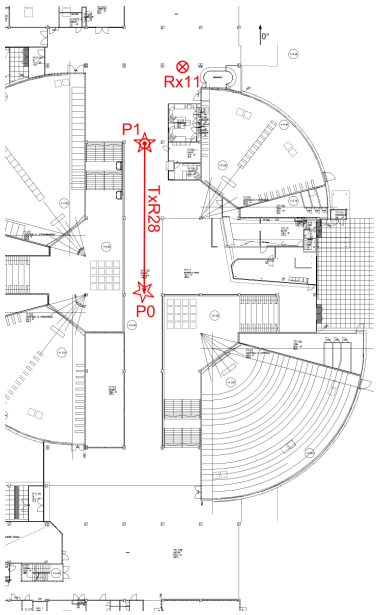
Frequency-Dependent Information:

	2.55 GHz	5.25 GHz
Index		87
Date		2005-11-25
Front-End Att.		0 dB
Calibration Att.		50 dB
Cal. Index (before)		81
Cal. Index (after)		88

Tx Environment

General Notes:

Stationary point, 1 min. See TxR28



Route Map

Symbols

Symbols used in Part I — The Random-Cluster Model

$\mathbf{\Theta}_{\text{env}}$	environment parameter pdf
$\mathbf{\Theta}_{\text{sys}}$	system parameter set
$\mathbf{\Theta}_{\text{DMP}}$	diffuse multipath parameter set
$\mathbf{\Theta}_c$	cluster parameter set for cluster c
$\mathbf{\Theta}_{cp}$	path parameter set for path p in cluster c
$\mathbf{H}(t, f)$	MIMO channel matrix (time variant, frequency selective)
$\mathbf{H}_\delta$	discrete part of MIMO channel matrix
$\mathbf{H}_d$	diffuse part of MIMO channel matrix
γ	complex amplitude
τ	delay
φ_{Tx}	azimuth of departure
φ_{Rx}	azimuth of arrival
θ_{Tx}	elevation of departure
θ_{Rx}	elevation of arrival
$\bar{\tau}$	cluster mean delay
$\bar{\varphi}_{\text{Tx}}$	cluster mean azimuth of departure (AOD)
$\bar{\varphi}_{\text{Rx}}$	cluster mean azimuth of arrival (AOA)
$\bar{\theta}_{\text{Tx}}$	cluster mean elevation of departure (EOD)
$\bar{\theta}_{\text{Rx}}$	cluster mean elevation of arrival (EOA)
$\sigma_{(\cdot)}$	cluster rms spread of respective argument
σ_γ^2	cluster power
ρ	snapshot power
N_p	number of paths within a cluster
N_c	number of clusters existing in a snapshot
$\Delta\sigma_\gamma^2$	rate of change of the cluster power
$\Delta\bar{\tau}_c$	rate of change of cluster mean delay
$\Delta\bar{\varphi}_{\text{Rx},c}$	rate of change of cluster mean AOA
$\Delta\bar{\varphi}_{\text{Tx},c}$	rate of change of cluster mean AOD
$\Delta\bar{\theta}_{\text{Rx},c}$	rate of change of cluster mean EOA
$\Delta\bar{\theta}_{\text{Tx},c}$	rate of change of cluster mean EOD
$\lvert\gamma_{\text{att}}\rvert^2$	cluster attenuation factor for cluster fading in and out
ν	Doppler frequency
v_{Tx}	transmitter speed of movement
Λ	cluster lifetime
$p(\cdot)$	probability density function (pdf) or respective argument
Δt_s	sampling time interval
Δt_Λ	cluster birth/death interval

continued on next page

continued from previous page	
$p(\chi_{\text{birth}})$	pdf of the cluster births
B	bandwidth
M	number of simulated frequencies
$\tilde{\cdot}$	realisation drawn from a pdf
Δf	frequency bin
$\mathbf{a}_{(\cdot)}(\cdot)$	antenna response
$\mathbf{G}_{(\cdot)}(\Delta f)$	system response
α_d	Peak of DMP PDP
β_d	Exponential decay constant of DMP PDP
τ_d	Base delay of the DMP PDP
ψ_{obs}	observed PDP at the receiver
ψ_δ	PDP of discrete contributions
ψ_{res}	PDP of residual power
ψ_d	PDP of DMP
$\mathbf{C}_{\boldsymbol{\pi}}$	Environment Characterisation Metric (ECM)

Symbols used in Part II — Cluster identification from measurements

$\mathbf{x}_l$	path parameter vector of lth MPC
τ_l	delay of lth MPC
$\varphi_{\text{Tx},l}$	AoD of the lth MPC
$\varphi_{\text{Rx},l}$	AoA of the lth MPC
$\mathbf{X}$	matrix of all path parameters in a snapshot
P_l	power of the lth path
γ_c	power of cluster c
$\mathbf{P}$	vector of all path powers in a snapshot
L	number of MPCs in a snapshot
L_c	number of paths within cluster c
$\boldsymbol{\mu}_c$	cluster centroid (mean position) of cluster c
$\mathcal{I}_l$	cluster number for the lth MPC
$\mathcal{C}_c$	indices of the paths contained in cluster c
$.^{(n)}$	parameter in the nth snapshot
$\boldsymbol{\theta}_c$	tracking parameter vector of cluster c
$\mathbf{C}_c$	joint spread (covariance matrix) of cluster c

Acronyms

AoA	azimuth of arrival
AoD	azimuth of departure
APS	azimuth power spectrum
COST	European cooperation in the field of scientific and technical research
DMP	diffuse multipath
DoA	direction of arrival
DoD	direction of departure
ECM	environment characterisation metric
GSCM	geometry-based stochastic channel model
LOS	line of sight
MCD	multipath component distance
MI	mutual information
MIMO	multiple-input multiple-output
MPC	multipath component
NLOS	non line of sight
OFDM	orthogonal frequency division multiplexing
pdf	probability density function
PDP	power-delay profile
RCM	Random-Cluster Model
Rx	Receiver, Receiver antenna(s)
SNR	signal-to-noise ratio
SV-ECM	singular values of the environment characterisation metric
Tx	Transmitter, Trasmitter antenna(s)

Bibliography

[1] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications — a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, February 2004.

[2] I. E. Telatar, "Capacity of multi-antenna gaussian channels," AT&T Bell Laboratories, Tech. Rep. BL0112170-950615-07TM, 1995.

[3] W. C. Lee, *Mobile communications engineering*. McGraw-Hill, 1982.

[4] F. Adachi, M. Feeny, A. Williamson, and J. Parsons, "Crosscorrelation between the envelopes of 900 MHz signals received at a mobile radio base station site," *Proc. Inst. Elect. Eng., pt. F*, vol. 133, pp. 506–512, 1986.

[5] M. Feeney and J. Parsons, "Cross-correlation between 900 MHz signals received on vertically separated antennas in small-cell mobile radio systems," *IEE Proceedings I, Communications, Speech and Vision*, vol. 138, no. 2, pp. 81–86, 1991.

[6] M. Steinbauer, A. Molisch, A. Burr, and R. Thomä, "MIMO channel capacity based on measurement results," in *European Conference on Wireless Technology*, Paris, France, October 2000.

[7] J. Wallace and M. Jensen, "Characteristics of measured 4x4 and 10x10 MIMO wireless channel data at 2.4 GHz," in *Proc. 2001 IEEE Antennas Propag. Soc. Intl. Symp.*, vol. 3, Boston, MA, USA, Jul 2001.

[8] R. Stridh, B. Ottersten, and P. Karlsson, "MIMO channel capacity of a measured indoor radio channel at 5.8 GHz," in *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, vol. 1, 2000, pp. 733–737 vol.1.

[9] P. Kyritsi and D. Cox, "Correlation properties of MIMO radio channels for indoor scenarios," in *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, vol. 2, 2001, pp. 994–998 vol.2.

[10] H. Özcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of the 'Kronecker´ MIMO radio channel model," *Electronics Letters*, vol. 39, no. 16, pp. 1209–1210, 2003.

[11] N. Czink, X. Yin, H. Özcelik, M. Herdin, E. Bonek, and B. Fleury, "Cluster characteristics in a MIMO indoor propagation environment," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1465–1475, April 2007.

[12] H. Bölcskei, D. Gesbert, and A. Paulraj, "On the capacity of ofdm-based spatial multiplexing systems," *EEE Transactions on Communications*, vol. 50, no. 2, pp. 225–234, 2002.

[13] E. Bonek, M. Herdin, W. Weichselberger, and H. Özcelik, "MIMO – study propagation first!" in *Proceedings of IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'03)*, Darmstadt, Deutschland, December 2003.

[14] H. Özcelik, "Indoor MIMO channel models," Ph.D. dissertation, Institut für Nachrichtentechnik und Hochfrequenztechnik, Technische Universität Wien, Vienna, Austria, December 2004, downloadable from www.nt.tuwien.ac.at/mobile/theses_finished.

[15] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D.Laurenson, G. Matz, A. Molisch, C. Oestges, and H. Özcelik, "Survey of channel and radio propagation models for wireless MIMO systems," *EURASIP Journal on Wireless Communications and Networking*, 2007.

[16] M. Steinbauer, A. Molisch, and E. Bonek, "The double-directional radio channel," *IEEE Antennas and Propagation Magazine*, vol. 43, no. 4, pp. 51 – 63, Aug. 2001.

[17] M. Landmann, M. Kaeske, R. Thomä, J. Takada, and I. Ida, "Measurement based parametric channel modeling considering diffuse scattering and specular components," in *International Symposium on Antennas and Propagation*, Niigata, Japan, August 2007 (Best Paper).

[18] R. Thomä, D. Hampicke, M. Landmann, A. Richter, and G. Sommerkorn, "Measurement-based channel modelling (mbpcm)," in *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA03)*, Torino, Italy, September 2003.

[19] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, no. 3, pp. 434–450, 17 1999.

[20] G. Del Galdo, V. Algeier, N. Czink, and M. Haardt, "Objects spatial localization from high-resolution parameter estimation on measurements," in *NEWCOM-ACORN Workshop*, Vienna, Austria, 2006.

[21] J. Kolu, J.-P. Nuutinen, T. Jämsä, J. Ylitalo, and P. Kyösti, "Playback simulations of measured MIMO radio channels," in *COST 273, TD(04)110*, Gothenburg, Sweden, June 2004.

[22] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128 – 137, Feb 1987.

[23] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, August 2002.

[24] W. Weichselberger, "Spatial structure of multiple antenna radio channels - a signal processing viewpoint," Ph.D. dissertation, Institut für Nachrichtentechnik und Hochfrequenztechnik, Technische Universität Wien, Dec. 2003, downloadable from www.nt.tuwien.ac.at/mobile/theses_finished.

[25] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 90–100, January 2006.

[26] N. Costa and S. Haykin, "A novel wideband mimo channel model and the wideband mimo software defined radio," in *Wireless Communications, Networking and Mobile Computing, 2006. WiCOM 2006.International Conference on*, 2006, pp. 1–4.

[27] M. Weis, G. D. Galdo, and M. Haardt, "A correlation tensor-based model for time variant frequency selective MIMO channels," in *Workshop on Smart Antennas*, Vienna, Austria, February 2007.

[28] J. Han and M. Kamber, *Data Mining, Concepts, and Techniques*. Morgan Kaufmann Publishers, 2001.

[29] A. Burr, “Some open questions in channel modeling,” in *NEWCOM Dept. 2 Workshop*, Vienna, Austria, 29-30 September 2005.

[30] H. Xiao, A. G. Burr, L. Hentilä, and P. Kyösti, “Statistical technique to identify clusters from multi-dimensional measurement data,” in *European Conference on Antennas and Propagation (EuCAP)*, Edinburgh, UK, November 2007.

[31] M. Landmann, W. Kottermann, and R. Thomä, “On the influence of incomplete data models on estimated angular distributions in channel characterisation,” in *European Conference on Antennas and Propagation (EuCAP)*, 2007.

[32] V. Erceg *et al.*, “TGn Channel Models,” IEEE P802.11 Wireless LANs, Tech. Rep., May 2004, http://www.802wirelessworld.com:8802/.

[33] B. Khatri, “Effect on capacity of clustering in indoor MIMO channels,” in *MIMO: Communications Systems from Concept to Implementations (Ref. No. 2001/175), IEE Seminar on*, 2001, pp. 15/1–15/7.

[34] K. Li, M. Ingram, and A. Van Nguyen, “Impact of clustering in statistical indoor propagation models on link capacity,” *IEEE Transactions on Communications*, vol. 50, no. 4, pp. 521 – 523, April 2002.

[35] A. F. Molisch, “Effect of far scatterer clusters in MIMO outdoor channel models,” in *Proc. 57th IEEE Vehicular Techn. Conf.*, 2003, pp. 534–538.

[36] A. Richter, “Estimation of radio channel parameters: Models and algorithms,” Ph.D. Dissertation, Technische Universität Ilmenau, Ilmenau, Germany, 2005.

[37] J. Andersen, J. Nielsen, G. Pedersen, G. Bauch, and M. Herdin, “Room electromagnetics,” *Antennas and Propagation Magazine, IEEE*, vol. 49, no. 2, pp. 27–33, 2007.

[38] A. Richter, J. Salmi, and V. Koivunen, “An algorithm for estimation and tracking of distributed diffuse scattering in mobile radio channels,” in *IEEE SPAWC*, Cannes, France, 2006.

[39] V. Erceg, D. G. Michelson, S. S. Ghassemzadeh, L. J. Greenstein, A. Rustako, P. B. Guerlain, M. K. Dennison, R. S. Roman, D. J. Barnickel, S. C. Wang, and R. R.Miller, “A model for the multipath delay profile of fixed wireless channels,” *IEEE JSAC*, vol. 17, no. 17, Mar. 1999.

[40] “Spatial channel model for Multiple Input Multiple Output (MIMO) simulations (3GPP TR 25.996), v6.1.0,” Sep. 2003. [Online]. Available: www.3gpp.org

[41] G. Calcev, D. Chizhik, B. Goransson, S. Howard, H. Huang, A. Kogiantis, A. Molisch, A. Moustakas, D. Reed, and H. Xu, “A wideband spatial channel model for system-wide simulations,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 389–403, March 2007.

[42] “COST 231 final report,” March 2001. [Online]. Available: http://www.lx.it.pt/cost231/final_report.htm

[43] “WINNER II interim channel models (D 1.1.1, V1.1),” Wireless World Initiative New Radio (WINNER II), November 2006. [Online]. Available: https://www.ist-winner.org

[44] N. Czink, H. Hofstetter, G. Steinböck, and A. Molisch, “Proposal for the COST 273 channel model: How to model multi-cluster environments,” *COST 273, TD(05)070*, 19–21 January 2005, Bologna, Italy.

[45] H. Hofstetter, "Characterization of the wireless MIMO channel," Ph.D. dissertation, Technische Universität Wien, Vienna, Austria, September 2006.

[46] L. Correia, Ed., *Mobile Broadband Multimedia Networks*. Academic Press, 2006.

[47] H. Hofstetter, A. F. Molisch, and N. Czink, "A twin-cluster MIMO channel model," in *European Conference on Antennas and Propagation (EuCAP) 2006*, Nice, France, 2006.

[48] A. Kuchar, M. Tangemann, and E. Bonek, "A real-time DOA-based smart antenna processor," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1279–1293, November 2002.

[49] B. H. Fleury, "First- and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2027–2044, September 2000.

[50] N. Czink, X. Yin, E. Bonek, and B. Fleury, "Cluster angular spreads in a MIMO indoor propagation environment," in *PIMRC'05*, Berlin, Germany, 2005.

[51] A. Ihler, "Kernel density estimation toolbox for Matlab," July 2007. [Online]. Available: http://ttic.uchicago.edu/~{}ihler/code/

[52] N. Czink, E. Bonek, L. Hentilä, P. Kyösti, J.-P. Nuutinen, and J. Ylitalo, "The interdependence of cluster parameters in mimo channel modeling," in *EUCAP 2006*, Nice, France, Nov. 2006.

[53] M. Bengtsson and B. Völcker, "On the estimation of azimuth distributions and azimuth spectra," *Proceedings of IEEE Vehicular Technology Conference*, vol. 3, no. 54, pp. 1612 – 1615, October 07–11, 2001, Atlantic City, NJ, USA.

[54] M. Debbah and R. Müller, "MIMO channel modelling and the principle of maximum entropy: An information theoretic point of view," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.

[55] X. Yin, T. Pedersen, N. Czink, and B. H. Fleury, "Parametric characterization and estimation of bi-azimuth dispersion of path components," in *Proceedings of the IEEE SPAWC 2006*, Cannes, France, 2006.

[56] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 2, pp. 437–447, March 2000.

[57] A. Abdi, J. A. Barger, and M. Kaveh, "A parametric model for the distribution of the angle of arrival and the associated correlation function and power spectrum at the mobile station," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 3, pp. 425–434, May 2002.

[58] L. Correia, Ed., *Wireless Flexible Personalised Communications (COST 259 Final Report)*. Chichester (UK): Wiley, 2001.

[59] G. Del Galdo, "Geometry-based channel modelling for multi-user MIMO systems and applications," Ph.D. dissertation, Technische Universität Ilmenau, Ilmenau, Germany, 2007.

[60] N. Czink, A. Richter, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "Including diffuse multipath parameters in mimo channel models," in *Vehicular Technology Conference 2007 Fall*, Baltimore, USA, 2007.

[61] A. Taparugssanagorn, X. Yin, J. Ylitalo, and B. H. Fleury, "Phase noise mitigation in channel parameter estimation for TDM switched MIMO channel sounding," in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2007.

[62] A. Taparugssanagorn, J. Ylitalo, and B. H. Fleury, "Phase-noise in TDM-switched MIMO channel sounding and its impact on channel capacity estimation," in *IEEE Global Communications Conference (GLOBECOM'07)*, Washington DC, USA, 2007.

[63] M. Herdin and E. Bonek, "A MIMO correlation matrix based metric for characterizing non-stationarity," in *Proceedings IST Mobile & Wireless Communications Summit*, Lyon, France, June 2004.

[64] A. F. Molisch, M. Steinbauer, M. Toeltsch, E. Bonek, and R. S. Thomä, "Capacity of MIMO systems based on measured wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 561–569, April 2002.

[65] C. Oestges and B. Clerckx, *MIMO Wireless Communications*. Elsevier Academic Press, 2007.

[66] M.T. Ivrlac and J.A. Nossek, "Quantifying Diversity and Correlation of Rayleigh Fading MIMO Channels," in *IEEE International Symposium on Signal Processing and Information Technology, ISSPIT'03*, Darmstadt, Germany, December 2003.

[67] J. W. Demmel, "The probability that a numerical analysis problem is difficult," *Mathematics of Computation*, vol. 50, pp. 449–480, 1988.

[68] R. W. Heath and A. Paulraj, "Switching between diversity and multiplexing in MIMO systems," *IEEE Transactions on Communications*, vol. 53, pp. 962–968, 2005.

[69] P. Kyösti, D. Laselva, L. Hentilä, and T. Jämsä, "Validating IST-WINNER indoor MIMO radio channel model," in *IST Mobile and Wireless Summit 2006*, June 2006.

[70] N. Czink, G. D. Galdo, X. Yin, E. Bonek, and J. Ylitalo, "A novel environment characterization metric for clustered MIMO channels used to validate a SAGE parameter estimator," *Wireless Personal Communications*, 2007, in press.

[71] A. Pal, M. Beach, and A. Nix, "A quantification of 3-d directional spread from small-scale fading analysis," *IEEE International Conference on Communications*, June 2006, Istanbul, Turkey.

[72] M. Steinbauer, H. Özcelik, H. Hofstetter, C. Mecklenbräuker, and E. Bonek, "How to quantify multipath separation," *IEICE Trans. Electron.*, vol. E85, no. 3, pp. 552–557, March 2002.

[73] A. Pal, M. Beach, and A. Nix, "A quantification of 3-d directional spread from small-scale fading analysis," in *COST 273 Post-Project Meeting*, Lisbon, Portugal, Nov. 2005.

[74] M. Landmann and G. Del Galdo, "Efficient antenna description for MIMO radio channel modelling and estimation," in *European Microwave Week*, Amsterdam, The Netherlands, October 2004.

[75] M. Landmann, A. Richter, and R. Thomä, "Performance evaluation of real antenna arrays for high-resolution DOA estimation in channel sounding - part 1: Channel parameter resolution limits," in *COST 273, TD(03)199*, Prague, Czech Republic, September 2003.

[76] W. Dong, J. Zhang, X. Gao, P. Zhang, and Y. Wu, "Cluster identification and properties of outdoor wideband mimo channels," in *IEEE Vehicular Technology Conference Fall 2007*, Baltimore, MD, USA, October 2007.

[77] K. Yu, Q. Li, D. Cheung, and C. Prettie, "On the tap and cluster angular spreads of indoor WLAN channels," in *Proceedings of IEEE Vehicular Technology Conference Spring 2004*, Milano, Italy, May 17–19, 2004.

[78] S. Wyne, N. Czink, J. Karedal, P. Almers, F. Tufvesson, and A. F. Molisch, "A cluster-based analysis of outdoor-to-indoor office MIMO measurements at 5.2 GHz," in *IEEE VTC Fall*, Montreal, Canada, 2006.

[79] D. W. Scott, *Multivariate Density Estimation*. Wiley, 1992.

[80] H. Özcelik, N. Czink, and E. Bonek, "What makes a good MIMO channel model?" in *VTC 2005 Spring*, Stockholm, Sweden, May 2005.

[81] J. Kivinen, X. Zhao, and P. Vainikainen, "Empirical characterization of wideband indoor radio channel at 5.3 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 49, no. 8, pp. 1192–1203, August 2001.

[82] J. Salmi, A. Richter, and V. Koivunen, "Enhanced tracking of radio propagation path parameters using state-space modeling," in *14th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, September 2006.

[83] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*. Prentice Hall, 1993.

[84] B. H. Fleury, X. Yin, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for communication systems equipped with antenna arrays," *Proc. 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands*, no. ISC-379, 2003.

[85] N. Czink, E. Bonek, L. Hentilä, J.-P. Nuutinen, and J. Ylitalo, "Cluster-based MIMO channel model parameters extracted from indoor time-variant measurements," in *IEEE GlobeCom 2006*, San Francisco, USA, Nov. 2006.

[86] J. Kolu, T. Jamsa, and A. Hulkkonen, "Real time simulation of measured radio channels," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, vol. 1, 2003, pp. 183–187 Vol.1.

[87] P. Kyritsi, D. Cox, R. Valenzuela, and P. Wolniansky, "Effect of antenna polarization on the capacity of a multiple element system in an indoor environment," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1227–1239, 2002.

[88] C. Oestges, M. Guillaud, and M. Debbah, "Multi-polarized mimo communications: Channel model, mutual information and array optimization," in *Wireless Communications and Networking Conference, 2007.WCNC 2007*, 11-15 March 2007, pp. 1057–1061.

[89] C. Oestges, D. Vanhoenacker-Janvier, and B. Clerckx, "Macrocellular directional channel modeling at 1.9 GHz: Cluster parametrization and validation," in *VTC 2005 Spring*, Stockholm, Sweden, May 2005.

[90] A. Paier, J. Karedal, N. Czink, H. Hofstetter, C. Dumard, T. Zemen, F. Tufvesson, C. Mecklenbräuker, and A. Molisch, "First results from car-to-car and car-to-infrastructure radio channel measurements at 5.2 GHz," in *IEEE PIMRC 2007 (invited)*, Athens, Greece, September 2007.

[91] A. Paier, J. Karedal, N. Czink, H. Hofstetter, C. Dumard, T. Zemen, F. Tufvesson, A. Molisch, and C. Mecklenbräuker, "Car-to-car radio channel measurements at 5 GHz:

Pathloss, power delay profile, and doppler delay spectra," in *ISWCS 2007*, Trondheim, Norway, October 2007.

[92] J. R. Foerster, M. Pendergrass, and A. F. Molisch, "A UWB channel model for ultra wideband indoor communications," in *Proc. of the International Symposium on Wireless Personal Multimedia Communications 2003 (WPMC 2003)*, October 2003.

[93] A. F. Molisch, H. Asplund, R. Heddergott, M. Steinbauer, and T. Zwick, "The COST 259 directional channel model – I. overview and methodology," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 3421–3433, 2006.

[94] H. Asplund, A. A. Glazunov, A. F. Molisch, K. I. Pedersen, and M. Steinbauer, "The COST 259 directional channel model – II. macrocells," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 3434–3450, 2006.

[95] J. Laurila, K. Hugl, M. Toeltsch, E. Bonek, K. Kalliola, and P. Vainikainen, "Directional wideband 3-D measurements of mobile radio channel in urban environment," in *COST 259, TD(99)092*, Leidschendam, The Netherlands, September 1999.

[96] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 347 – 359, March 2000.

[97] C.-C. Chong, C.-M. Tan, D. Laurenson, S. McLaughlin, M. Beach, and A. Nix, "A new statistical wideband spatio-temporal channel model for 5-GHz band WLAN systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 2, pp. 139 – 150, Feb. 2003.

[98] J. Salo, J. Salmi, N. Czink, and P. Vainikainen, "Automatic clustering of nonstationary MIMO channel parameter estimates," in *ICT'05*, Cape Town, South Africa, May 2005, Cape Town, South Africa.

[99] R. Roy, A. Paulraj, and T. Kailath, "Esprit — a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1340–1342, October 1986.

[100] M. Haardt and J. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1232–1242, May 1995.

[101] M. Landmann, "Limitations of experimental channel characterisation," Ph.D. dissertation, Technische Universität Ilmenau, 2008, yet unpublished.

[102] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "Improving clustering performance by using the multi-path component distance," *IEE Electronics Letters*, vol. 42, no. 1, pp. 44–45, Jan. 2006.

[103] ——, "Automatic clustering of MIMO channel parameters using the multi-path component distance measure," in *WPMC'05*, Aalborg, Denmark, Sept. 2005.

[104] R. Tian, "Joint clustering and tracking for mimo radio channel modeling," Master's thesis, Chalmers University of Technology, Gothenburg, Sweden, 2007.

[105] X. Yin, T. Pedersen, N. Czink, and B. H. Fleury, "Parametric characterization of biazimuth and delay dispersion of individual path components," in *EUCAP 2006*, Nice, France, 2006.

[106] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650–1654, Dec. 2002.

[107] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric MIMO channel data including path powers," in *IEEE Vehicular Technology Conference 2006 Fall*, Montreal, Canada, 2006.

[108] J. Salo, G. Del Galdo, J. Salmi, P. Kyösti, M. Milojevic, D. Laselva, and C. Schneider, "MATLAB implementation of the 3GPP spatial channel model (3GPP TR 25.996)," Online, jan 2005, available: http://www.tkk.fi/Units/Radio/scm/.

[109] N. Czink, G. Del Galdo, X. Yin, and C. Mecklenbräuker, "A novel environment characterisation metric for clustered MIMO channels used to validate a SAGE parameter estimator," in *IST Mobile Summit 2006*, Mykonos, Greece, 2006.

[110] N. Czink, G. D. Galdo, and C. F. Mecklenbräuker, "A novel cluster tracking algorithm," in *IEEE Personal Indoor and Mobile Radio Communications (PIMRC) 2006*, September 2006.

[111] B. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *2002 International Zurich Seminar on Broadband Communications*, Zurich, Feb. 2002, pp. 30–1 – 30–9.

[112] G. Eriksson, F. Tufvesson, and A. F. Molisch, "Propagation channel characteristics for peer-to-peer multiple antenna systems at 300 MHz," in *IEEE GlobeCom 2006*, San Francisco, USA, Nov. 2006.

[113] E. Bonek, N. Czink, V. Holappa, M. Alatossava, L. Hentilä, J. Nuutinen, and A. Pal, "Indoor MIMO measurements at 2.55 and 5.25 GHz - a comparison of temporal and angular characteristics," in *IST Mobile Summit 2006*, Mykonos, Greece, June 2006.

[114] A. S. Y. Poon and M. Ho, "Indoor multiple-antenna channel characterization from 2 to 8 GHz," *Proceedings of IEEE ICC*, vol. 5, pp. 3519–3523, May 2003.

[115] Elektrobit PropSim Homepage, http://www.propsim.com/.

[116] J. Koivunen, P. Almers, V.-M. Kolmonen, J. Salmi, A. Richter, F. Tufvesson, P. Suvikunnas, A. Molisch, and P. Vainikanen, "Dynamic multi-link indoor mimo measurements at 5.3 GHz," in *COST2100 Management Comittee Meeting TD(07)373*, Duisburg, Germany, September 10-12 2007.

[117] G. Acosta-Marum and M. A. Ingram, "Doubly selective vehicle-to-vehicle channel measurements and modeling at 5.9 GHz," in *Wireless Personal Multimedia Communications (WPMC) 2006*, 17–20 September 2006.

[118] J. Maurer, T. Fügen, and W. Wiesbeck, "Narrow-band measurement and analysis of the inter-vehicle transmission channel at 5.2 GHz," in *Vehicular Technology Conference Sping (VTC) 2002*, 2002.

[119] X. Zhao, J. Kivinen, P. Vainikainen, and K. Skog, "Characterization of doppler spectra for mobile communications at 5.3 GHz," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 1, pp. 14–23, January 2003.