Ulrich Türke

Efficient Methods for WCDMA Radio Network Planning and Optimization

TEUBNER RESEARCH

Advanced Studies Mobile Research Center Bremen

Das Mobile Research Center Bremen (MRC) erforscht, entwickelt und erprobt in enger Zusammenarbeit mit der Wirtschaft mobile Informatik-, Informations- und Kommunikationstechnologien. Als Forschungs- und Transferinstitut des Landes Bremen vernetzt und koordiniert das MRC hochschulübergreifend eine Vielzahl von Arbeitsgruppen, die sich mit der Entwicklung und Anwendung mobiler Lösungen beschäftigen. Die Reihe „Advanced Studies" präsentiert ausgewählte hervorragende Arbeitsergebnisse aus der Forschungstätigkeit der Mitglieder des MRC.

In close collaboration with the industry, the Mobile Research Center Bremen (MRC) investigates, develops and tests mobile computing, information and communication technologies. This research association from the state of Bremen links together and coordinates a multiplicity of research teams from different universities and institutions, which are concerned with the development and application of mobile solutions. The series "Advanced Studies" presents a selection of outstanding results of MRC's research projects.

Ulrich Türke

Efficient Methods for WCDMA Radio Network Planning and Optimization

TEUBNER RESEARCH

Bibliographic information published by Die Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.d-nb.de>.

Dissertation Universität Bremen, 2007

mrc
Mobile Research Center

Gedruckt mit freundlicher Unterstützung des
MRC Mobile Research Center der Universität Bremen

Printed with friendly support of
MRC Mobile Research Center, Universität Bremen

1st Edition September 2007

Readers: Ute Wrasmann / Anita Wilke

Deutscher Universitäts-Verlag and Teubner Verlag are companies of
Springer Science+Business Media.
www.duv.de
www.teubner.de

Cover design: Regine Zimmer, Dipl.-Designerin, Frankfurt/Main
Printed on acid-free paper
Printed in Germany

ISBN 978-3-8350-0903-5

Preface

When joining Siemens in 2001, I also extended my research interest towards radio network planning methodologies. This area of research brought together my personal interest in mobile communications and in the design of efficient algorithms and data structures. Between 2001 and 2003, I participated in the EU project Momentum, which was targeting the performance evaluation and optimization of UMTS radio networks. In this project my main focus was on Monte-Carlo snapshot techniques, shortly after I broadened my research scope towards network optimization techniques. Both constitute the basis of this thesis. Later I got the chance to bring in the ideas and methods of my thesis into a Siemens internal planning and optimization tool, the development of which I guided over the last two years.

First of all I would like to thank Professor Dr. Carmelita Görg for supervising this thesis, for her guidance, and her encouragement. I very much appreciate that she - although I, as an external Ph.D. student, was not part of her department at the University of Bremen - was always very open for discussions and provided me with a large amount of useful suggestions over the past years. I without a doubt greatly benefited from this, both in my scientific work and personally. I am also deeply grateful for the comments and suggestions by Professor Dr. Thomas Kürner, who is also the second examiner of this thesis. The discussions with him and his advice - in particular in the last phase of the work on this thesis - were of great help to me. I am deeply obliged to Helmut Mühlbauer for giving me the chance to conduct this thesis within Siemens, for his strong faith in me and my work, and for strengthening my view on economical aspects and the practical relevance of my research. Furthermore I am very grateful to Professor Dr. Ranjit Perera, who participated as a guest scientist in the EU project Momentum. I very much appreciated the long and fruitful discussions with him - not only on technical topics. Special thanks go to my colleagues Michael Koonert and Richard Schelb for many discussions, and not least for proof-reading parts of this thesis.

Ulrich Türke

Abstract

The planning of mobile radio networks requires fast and accurate methods for network performance evaluation. In this thesis novel methods for the performance analysis of WCDMA networks are developed and discussed.

The state-of-the-art approach for evaluating the performance of WCDMA networks is the Monte-Carlo snapshot simulation. However, current solutions based on snapshot simulations have major weaknesses with respect to accuracy, simulation time, and flexibility. In particular the size of networks that can be analyzed in an acceptable time frame with standard hardware resources is typically very limited.

The new snapshot analysis method developed in this thesis significantly outperforms the state-of-the-art methods. It allows for a detailed analysis of very large networks (N x 1000 base station sites) on a standard PC within running times acceptable for everyday planning. This is mainly achieved by a highly linear modeling of the system and by a reduction of the size of the sets of linear equations to be calculated in each snapshot. The dimension reduction is achieved by the application of a cell based transmission power and interference level calculation as opposed to the standard mobile based approach. Despite the significant speed-up achieved in the evaluation of a snapshot, the proposed approach also allows a very accurate and flexible modeling of system features, e.g. with respect to *Radio Resource Management* (RRM) and *Quality of Service* (QoS).

Increasing the performance of evaluating a single snapshot is only one approach to reduce the running time of the snapshot analysis. Besides this, also the number of snapshots required to yield a certain accuracy of results can be reduced by means of variance reduction techniques. In the area of snapshot simulations for network planning, variance reduction techniques have almost not been considered yet. In this thesis considerable effort has been put into evaluating the applicability of variance reduction techniques for cutting down the running time of the snapshot analysis. Based on the discussion of several variance reduction techniques known from theory, two variance reduction techniques are identified and tailored to an application in everyday planning. Both, *control variates* and *correlated sampling*, are applied to example planning scenarios. Considerable gains with respect to the number of required snapshots to yield a certain accuracy of results are obtained for the proposed variance reduction methods.

While providing very accurate results at a high level of detail, for some applications the snapshot simulation approach is too time consuming. In particular search based optimization techniques, which compare a large number of different network configurations, require methods with much shorter running times, while accepting a loss in level of detail and accuracy. Targeting such areas the analytical methods developed in this thesis can be effectively applied. Based on a static method two novel extended statistical methods, incorporating shadow fading statistics, are proposed. Results are presented showing significant improvements in approximating the results from snapshot simulations in terms of relevant system parameters. Finally, the application of the developed methods in network optimization is discussed. Two different optimization methods are developed. The presented results prove the value of the methods for an application in network optimization.

The majority of the developed methods is currently successfully applied in a radio network planning tool.

Contents

List of Abbreviations

3GPP	3rd Generation Partnership Project
AAP	Analysis of Adjacent Points
AC	Admission Control
AICH	Acquisition Indication Channel
AMC	Adaptive Modulation and Coding
AMR	Adaptive Multi–Rate
ARIB	Association of Radio Industries and Businesses
ARPU	Average Revenue Per Subscriber
ARQ	Automatic Repeat Request
ATIS	Alliance for Telecommunications Industry Solutions
BCCH	Broadcast Control Channel
BCH	Broadcast Channel
BHCA	Busy Hour Call Attemps
BLER	Block Error Rate
BMC	Broadcast/Multicast Control Protocol
BPSK	Binary Phase Shift Keying
BS	Base Station
C-Plane	Control Plane
CC	Congestion Control
CCCH	Common Control Channel
CCSA	China Communications Standards Association
CDMA	Code Division Multiple Access
CIR	Carrier to Interference plus Noise Ratio
CN	Core Network
CorS	Correlated Sampling
CPCH	Uplink Common Packet Channel
CPICH	Common Pilot Channel
CS	Circuit-Switched
CTCH	Common Traffic Channel
CV	Control Variates
D-AMPS	Digital Advanced Mobile Phone System
DCCH	Dedicated Control Channel
DCH	Dedicated Transport Channel
DL	Downlink

DPCCH	Dedicated Physical Control Channel
DPDCH	Dedicated Physical Data Channel
DS-CDMA	Direct Sequence Code Division Multiple Access
DSCH	Downlink Shared Channel
DTCH	Dedicated Traffic Channel
DTX	Discontinuous Transmission
E-AGCH	E-DCH Absolute Grant Channel
E-DCH	Enhanced Dedicated Channel
E-DPCCH	Enhanced Dedicated Physical Control Channel
E-DPDCH	Enhanced Dedicated Physical Data Channel
E-HICH	HARQ Indicator Channel
E-RGCH	E-DCH Relative Grant Channel
EGSM	Extended Global System for Mobile Communications
ETSI	European Telecommunications Standards Institute
EU	European Union
FACH	Forward Access Channel
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
GGSN	Gateway GPRS Support Node
GMSC	Gateway MSC
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HCS	Hybrid Coefficient Scaling
HLR	Home Location Register
HPI	Horizontal Projection Interpolation
HS-DSCH	Highspeed Downlink Shared Channel
HS-PDSCH	High-Speed Physical Downlink Shared Channel
HS-DPCCH	High-Speed Dedicated Physical Control Channel
HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
HSUPA	High Speed Uplink Packet Access
IIR	Infinite Impulse Response
IMT2000	International Mobile Telecommunications-2000
IP	Internet Protocol
IS	Importance Sampling
ISDN	Integrated Services Digital Network
ISI	Inter-Symbol Interference
ITU	International Telecommunications Union
LOS	Line of Sight
LTE	(3GPP) Long Term Evolution
MAC	Medium Access Control
ME	Mobile Equipment
MS	Mobile Station

MSC	Mobile Services Switching Center
NCS	Noise Coefficient Scaling
NLOS	Non Line of Sight
NOF	Maximum Neighbor Offset Factor
OCCS	Own Cell Coefficient Scaling
PBDG	Path loss Based Downgrading
PC	Personal Computer
PCCH	Paging Control Channel
PCCPCH	Primary Common Control Physical Channel
PCH	Paging Channel
PCPCH	Physical Common Packet Channel
PDC	Personal Digital Cellular
PDCP	Packet Data Convergence Protocol
PICH	Paging Indication Channel
PLMN	Public Land Mobile Network
PRACH	Physical Random Access Channel
PS	Packet-Switched
PSCH	Physical Downlink Shared Channel
PSTN	Public Switched Telephone Network
PHY	Physical Layer
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RACH	Random Access Channel
RDG	Random Downgrading
RLC	Radio Link Control
RNC	Radio Network Controller
RNS	Radio Network Sub-System
ROI	Return on Investment
RRC	Radio Resource Control
RRM	Radio Resource Management
RSCP	Received Signal Code Power
SCH	Synchronization Channel
SCCPCH	Secondary Common Control Physical Channel
SF	Spreading Factor
SGSN	Serving GPRS Support Node
SHO	Soft-Handover
SNR	Signal to Interference plus Noise Ratio
STD	Short-Term Dynamic Simulation
StrS	Stratified Sampling
TCP	Transmission Control Protocol
TD-SCDMA	Time Division-Synchronous Code Division Multiple Access
TDD	Time Division Duplex
TDMA	Time Division Multiple Access

TTA	Telecommunications Technology Association
TTC	Telecommunication Technology Committee
TTI	Transmission Time Interval
TX	Transmit
U-Plane	User Plane
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
URA	UTRAN Registration Area
USIM	UMTS Subscriber Identity Module
UTRA	Universal Terrestrial Radio Access
UTRAN	Universal Terrestrial Radio Access Network
VLR	Visitor Location Register
WBI	Weighted Bilinear Interpolation
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access

List of Symbols

Δ_{SHO}	SHO margin
Λ_{ij}	Fast fading contribution to link gain between i and j in decibel
$\underline{\Lambda}_{ij}$	Fast fading contribution to link gain between i and j in linear scale
$\Phi\left(x\right)$	Standard normal distribution
Ω	Antenna gain
Ω_G	Global antenna gain
$\Omega_H()$	Contribution of horizontal antenna pattern to total antenna gain
$\Omega_V()$	Contribution of vertical antenna pattern to total antenna gain
Σ_{ij}	Shadow fading contribution to link gain between i and j in decibel
Θ_{ij}	Total path loss between i and j
α_i	Orthogonality factor applicable for mobile/pixel i
β	Scaling constant (control variate constant)
$\gamma_i^{\uparrow}$	Required *Uplink* (UL) signal to interference plus noise ratio
$\gamma_i^{\downarrow}$	Required *Downlink* (DL) signal to interference plus noise ratio
$\hat{\gamma}_{i,\,min}$	Minimum required pilot signal to total interference plus noise ratio
$\delta_{i,p,s}$	Serving probability of cell i at pixel p for service s
$\delta_{i,p,s}^{\text{SHO}}$	SHO probability of best serving cell i at pixel p for service s
$\delta_{p,s}^{\text{SHO}}$	Total SHO probability at pixel p for service s
$\epsilon_i^{\uparrow}$	UL source activity factor for mobile/pixel i
$\epsilon_i^{\downarrow}$	DL source activity factor for mobile/pixel i
$\zeta_{ij}^{\downarrow}$	DL link coefficient
$\zeta_{i,j,p}^{*\downarrow}$	Mean DL link coefficient
$\eta_i^{\uparrow}$	Total noise perceived by cell i
$\eta_i^{\downarrow}$	Total noise perceived at mobile/pixel i
$\theta_{p,s}$	User density of service s in pixel p
$\kappa_{ij}^{\uparrow}$	Effective UL interference raise between mobile/pixel i and cell j

$\kappa_{ij}^{\downarrow}$	Effective DL interference raise between cell i and mobile/pixel j
ρ_L	Shadow fading correlation factor
λ	Wave length
λ	Arrival Rate
μ_X	Mean value of X, $\mu_X = \mathrm{E}\{X\}$
$\nu_i^{\downarrow}$	DL link inhomogeneity of mobile/pixel i
$\nu_{p,s}^{*\downarrow}$	Mean DL link inhomogeneity of pixel p for service s
ψ	Variance reduction coefficient
σ	Shadow fading deviation
σ_p'	Location specific shadow fading deviation
σ_{cp}''	Link specific shadow fading deviation
σ_X	Standard deviation of X
$\upsilon_{i,p,s}$	Coverage probability of best serving cell i at pixel p for service s
$\phi(x)$	Standard normal density function
τ_c	Total power for all common channels of cell c (static cell power)
$\tau_{c,s}$	Total power for all common channels on scrambling code s of cell c
$\vec{\tau}_c$	Vector of common channel power levels for each scrambling code of cell c
$\hat{\tau}_c$	Pilot channel power of cell c
$\check{\tau}_c$	Static channel power of cell c without pilot channel power: $\check{\tau}_c = \tau_c - \hat{\tau}_c$
$\xi_{ij}^{\uparrow}$	Total UL link gain between mobile/pixel i and cell j
$\xi_{ij}^{\downarrow}$	Total DL link gain between cell i and mobile/pixel j
$\xi_{i,p,s}^{*\downarrow}$	Median DL link gain between cell i and pixel p for service s
$\tilde{\xi}_{ij}$	Time varying link gain between transmitter i and receiver j (subject to fast fading)
ω_{ip}	Shadow fading component between cell c and pixel p
B	User bit rate
$\mathcal{C}$	Set of all cells in the system
C	Control variate
CIR	Carrier to interference ratio
$\widetilde{CIR}$	Time varying carrier to interference ratio (subject to fast fading)
$\mathrm{E}\{X\}$	Expected value of X
$\frac{E_b}{N_0}$	(Required) bit-energy to interference plus noise ratio

$\frac{E_c}{I_0}$	Received pilot power to total interference ratio plus noise ratio
I	Identity matrix
$\mathcal{M}$	Set of served mobiles/users
$\mathcal{M}_c$	Set of mobiles/users served by primary cell c
$\mathcal{M}_{c,s}$	Set of mobiles/users served by scrambling code s of primary cell c
$\mathrm{P}\{i\}$	Probability of a certain event i
PG	CDMA processing gain
$R_p\{\mathcal{S}_p\}$	Best serving cell at pixel p (random variable)
R_p^*	Static best serving cell at pixel p
$\mathcal{S}_p$	Serving set at pixel p
$T^{\uparrow}$	Characteristic UL coupling matrix
$T^{\downarrow}$	Characteristic DL coupling matrix
$T^{*\downarrow}$	Mean characteristic DL coupling matrix
W	Chip-rate (UMTS: $3.84 \cdot 10^6$)
c	Speed of Light
$\mathrm{cov}\{X, Y\}$	Covariance of X and Y
e	Mean absolute error
f	Frequency
h_{BS}	BS height
h_{MS}	MS height
$h_X(x)$	Probability density function of random variable X
$h_c^{\uparrow}$	Characteristic UL inhomogeneity of cell c
$h_c^{\downarrow}$	Characteristic DL inhomogeneity of cell c
$h_c^{*\downarrow}$	Mean characteristic DL inhomogeneity of cell c
$\vec{h}^{*\downarrow}$	Mean vector of characteristic DL inhomogeneities
$\vec{h}^{\uparrow}$	Vector of characteristic UL inhomogeneities per cell
$\vec{h}^{\downarrow}$	Vector of characteristic DL inhomogeneities per cell
$\vec{h}^{*\downarrow}$	Vector of mean characteristic DL inhomogeneities per cell
$i_c^{\uparrow}$	Total interference at cell c
$\breve{i}_c^{\uparrow}$	UL other system interference level at cell c
$\vec{i}$	Vector of all cell interference levels $i_c^{\uparrow}$
$\tilde{i}_i$	Time varying received interference at receiver i (subject to fast fading)

$i_i^{\downarrow}$	Total interference at mobile/pixel i
$\breve{i}_i^{\downarrow}$	DL other system interference at mobile/pixel i
$l_c^{\uparrow}$	UL cell load of cell c
$l_c^{\downarrow}$	DL cell load of cell c
$l_{c,\,max}^{\uparrow}$	Maximum allowed UL cell load of cell c
$l_{c,\,max}^{\downarrow}$	Maximum allowed DL cell load of cell c
$p_c^{\downarrow}$	Total output power of cell c
$p_c^{*\downarrow}$	Mean total output power of cell c
$\vec{p}$	Vector of total cell output power levels
$\vec{p}$	Mean vector of total cell output power levels
$p_{c,\,max}^{\downarrow}$	Maximum output power of cell c
$\vec{p}$	Vector of total cell power levels $p_c^{\downarrow}$
r	Correlation coefficient
$r_{ci}^{\uparrow}$	Received UL link power from mobile/pixel i at cell c
$r_{jc}^{\downarrow}$	Received DL link power from cell c at mobile/pixel i
$\hat{r}_i$	Received pilot power of primary serving cell at mobile/pixel i
$\hat{r}_{i,\,min}$	Minimum required pilot power at mobile/pixel i
$\tilde{r}_i$	Time varying received power of link i (subject to fast fading)
SF_i	Spreading factor of channel i
S_c	Code allocation matrix of cell c
S_{ij}^c	Element (i,j) of the code allocation matrix S_c
$t_i^{\uparrow}$	Output power of mobile/pixel i
$t_{i,\,min}^{\uparrow}$	Minimum output power of mobile/pixel i
$t_{i,\,max}^{\uparrow}$	Maximum output power of mobile/pixel i
$t_i^{\downarrow}$	DL transmit power to mobile/pixel i
$t_{i,p,s}^{*\downarrow}$	Mean DL transmit power required by serving cell c for pixel p and service s
$t_{i,\,max}^{\downarrow}$	Maximum DL transmit power to mobile/pixel i
$\tilde{t}_i$	Time varying transmit power of transmitter i (subject to fast fading)
var $\{X\}$	Variance of random variable X
$v_{c,s}$	Total output power of scrambling code s of cell c
$\vec{v}_c$	Vector of output powers $v_{c,s}$ for each scrambling code s of cell c

List of Figures

List of Tables

Chapter 1

Introduction

Since the introduction of the *Global System for Mobile Communications* (GSM), as the most widely deployed second generation mobile communication standard, about fifteen years ago, mobile communication has established itself as a remarkably large business. The number of subscribers worldwide has constantly been growing and has exceeded the two billion mark in the year 2005. While the penetration of citizens with a subscription to a mobile communication network has exceeded 80% in many countries in the western world and a further increase in total revenue will require an increase in *Average Revenue Per Subscriber* (ARPU), there are still large emerging markets, in particular in Asia, with a rather low penetration of mobile subscribers.

Despite the great consumer interest in mobile communication, network operators and vendors have struggled significantly in the last years. The reasons are manifold, but two main reasons can be seen in the general recession of the new economy following the end of the "dot-com" boom of the late 1990s and the huge investments in both licenses and infrastructure for the third generation mobile networks in the last years. As a consequence mobile operators are obliged to constantly increase their *Return on Investment* (ROI) figures. It is obvious that this requirement will last even though the economical situation may change.

An important lever for improving the cost efficiency of mobile networks is the optimization of the existing network infrastructure, in particular the air interface, and an advanced planning of new networks in order to fully exploit the infrastructure in terms of capacity and end-user quality. By an improved planning of the air interface, capacity gains of typically up to 40% compared to conventional planning are expected, resulting in significant reductions in capital expenditure required for a desired target network quality.

The air interface denotes the communication protocols between base stations and user terminals. A base station typically comprises several cells, each serving a set of users. The signal transmission is separated between cells by at least one of the following mechanisms: space division, frequency division, code division, and time division, e.g. [Rap02]. In most mobile networks, space division is applied by using separate directional antennas

associated with the individual cells, each of the antennas thereby pointing to a different horizontal direction. The orientation of the antenna is described by its *azimuth angle*, i.e. the horizontal direction of the main lobe of the antenna, and its *tilt angle*, i.e. the vertical angle of the antenna's main lobe. Azimuth and tilt angles are two of the most important parameters in network design. Further important parameters are the position of base stations, i.e. *site locations*, the number of cells per base station, the antenna types, power settings, and the parameters related to the frequency-division and time-division access schemes, e.g. frequency allocation plans. In addition there are parameters related to the *Radio Resource Management* (RRM), like handover (cell change due to user mobility) and call admission control parameters, which contribute to the overall network performance. The objective of the radio network planning process is to optimize all above mentioned parameters, such that, given a specific set of constraints like required perceived end user quality and maximum total network cost, the capacity of the network is maximized. Radio network planning generally constitutes a trade-off between coverage, e.g. determined by a sufficiently large signal level received from the serving cell, and capacity, which is highly related to the amount of interference experienced by the users. For example, if the received signal strength at the edge of a cell area is increased by tilting the cell's antenna up (improved coverage), also additional interference is introduced in neighboring cells, thus the capacity in those cells is reduced.

The current radio network planning process can be described as *computer-aided manual planning*. Planning tools are used to assess the network configuration which is successively and manually changed by the network planner. The performance assessment is based on predictions regarding path loss and traffic data. This process possesses three inherent weaknesses:

1. Current state-of-the-art planning tools have considerable deficiencies with respect to both accuracy and computational time required to obtain statistically significant results. In fact, this is the reason why many network planning engineers still tend to carry out a *coverage based planning*, focusing on the provisioning of sufficient signal strength in the cell area. While this was a valid approach for the planning of traditional GSM networks, which apply *Frequency Division Multiple Access* (FDMA) with comparably large frequency reuse factors between cells, modern GSM and *Universal Mobile Telecommunications System* (UMTS) networks are operated at low frequency reuse factors (UMTS: reuse factor 1) and interference reduction is of eminently greater importance.
   The choice of the method for performance assessment is a compromise between accuracy and performance in terms of computational complexity. Dynamic simulations, simulating the network including user mobility and traffic dynamics as well as the dynamics of the involved physical processes, e.g. fading effects, over some period of time provide the most accurate results. However, they are at least for the currently available computer hardware too time and memory consuming to be applied in everyday planning. The dynamic simulation of a typical small cluster com-

prising 19 cells can easily take several days. Network planning requires the evaluation of networks comprising up to multiple thousand base stations in a reasonable time frame. The state-of-the-art approach for the evaluation of radio networks is thus the Monte-Carlo snapshot analysis. Available commercial solutions have major weaknesses with respect to the maximum size of networks that can be considered, simulation time, accuracy, and flexibility in modeling the system. Furthermore, in particular for an application in automated network optimization, also snapshot simulations are too time-consuming to be applied. Thus, for such applications even more time-efficient methods are required.

2. Even if fast and accurate methods for performance evaluation were commercially available, the time required by human network planners to manually optimize a network of decent size would be very high and hence expensive. Furthermore, the complexity of modern radio networks is so high, that the interdependencies can hardly be fully perceived by a human being. Hence, automatic optimization techniques are required to provide high quality results in a reasonable time frame.
3. An inherent problem of the tool-based planning approach is the (potential) inaccuracy of the input data. This holds in particular for the path loss predictions (generated from path loss models) and the spatial distribution of users. Both greatly influence the results of the performance evaluation and can thus adulterate the results of the planning process. Path loss predictions can be improved by means of measurement data, e.g. measurements from drive tests. A methodology for extrapolating traffic distributions from live network measurements is presented in [WTK04]. Further details on the relevance of accurate input data is presented in [EGT05].

For the last point a general methodology can hardly be defined. It depends on the specific infrastructure and detailed requirements as well as the availability of data how a sufficient reliability of the input data can be assured. This thesis targets at points one and two. Several novel methods for a performance evaluation of *Wideband Code Division Multiple Access* (WCDMA) networks well suited for the different requirements in network planning are developed and discussed. Furthermore, methods for an automated network optimization are presented. Although all methods proposed in this thesis have been specifically designed for an application in the planning of WCDMA networks they can be (and partially have been) adapted to other air interface standards e.g. GSM with frequency hopping, GPRS, EDGE, and WiMAX.

In Chapter 2 the WCDMA air interface is described. The chapter focuses on the protocols and functionalities required for an understanding of the subsequent chapters.

In Chapter 3 the modeling of the radio link between base station antennas and user terminal is presented. An appropriate modeling of the radio link is an essential ingredient of any method for radio network performance evaluation. Special attention is paid to the consideration of shadow fading and fast fading, and the proper modeling of base station antennas.

In Chapter 4 the Monte-Carlo method and its basic properties are introduced. In particular the application of variance reduction schemes is discussed in detail and illustrated for a simple example.

In Chapter 5 a novel and elaborate snapshot analysis method is developed. It allows a flexible modeling of the system, e.g. with respect to RRM algorithms, and at the same time provides a very fast and accurate network performance analysis. As an extension of the basic method, two variance reduction methods are proposed. The variance reduction achieved by these methods is quantified for example scenarios.

While the snapshot analysis provides very accurate results, for an application in automated optimization of networks it is too time-consuming. The novel statistical methods derived in Chapter 6 are well suited for such applications. The results obtained by the methods are compared with the results from the snapshot analysis.

In Chapter 7 two methods for a network optimization are presented. The first method is based on a path loss analysis of the network and allows a very fast optimization of the network. The second applies a local search optimization, where the methods presented in the preceding chapter are used for evaluating the performance of a network configuration. The applicability of the methods is shown for an example scenario.

Finally, in Chapter 8 conclusions are drawn and an outlook on future work is given.

Chapter 2

The WCDMA Air Interface

WCDMA is the most widely adopted air interface standard for third generation (3G) mobile networks. 3G networks have been designed to supersede second generation (2G) networks, i.e. GSM (Europe and large parts of the world), cdmaOne (Americas), D-AMPS (Americas), and PDC (Japan). The main advances of 3G systems compared with 2G systems are [HT04]:

- Higher user peak data rates (several Mbps).
- Dynamic adaptation of user data rates ("bandwidth on demand").
- Asymmetric uplink and downlink data rates to fit usual *Packet-Switched* (PS) traffic characteristics, e.g. web browsing, ftp downloads.
- Support of *Quality of Service* (QoS) differentiation, e.g. in terms of error rates and delay.
- Multiplexing of services with different quality and data rate requirements on a single connection.
- Improved spectrum efficiency (higher system throughput per unit of allocated frequency spectrum).

Work on the development of 3G systems started in 1992 by the *International Telecommunications Union* (ITU). Original target of the ITU activities was to define one global *International Mobile Telecommunications-2000* (IMT-2000) standard that operates in a common frequency band around 2 GHz in the largest world markets for mobile communication, namely USA, Europe, and Asia. In the USA, however, the frequency band considered by the ITU for IMT-2000 was already occupied by 2G networks (GSM 1900) and for satellite communication (in the 2.1 GHz band). The 3G networks currently being deployed in the USA hence partly reuse the frequency spectrum originally auctioned for GSM 1900 but in particular utilize different frequencies bands. This drives the need for several types of base station equipment and costly multi-band user equipment to provide global handsets. This is similar to the present situation in 2G networks that ITU intended to overcome with 3G.

Several competing standards have been defined by the standardization bodies in the IMT-2000 framework. In 1999 the ITU approved five air interfaces for IMT-2000. The most relevant of these are WCDMA, CDMA2000, TD-SCDMA, and EDGE as a direct successor of GSM. CDMA2000 is the successor of cdmaOne in the USA. CDMA2000 utilizes a smaller frequency band than WCDMA to simplify the usage of the fragmented spectrum. TD-SCDMA is a standard to be deployed in China.

WCDMA is the air interface used by the *Universal Mobile Telecommunications System* (UMTS), which is specified and brought into ITU by the 3rd Generation Partnership Project (3GPP). The most important standardization bodies contribute to 3GPP, i.e. ARIB (Japan), ATIS (USA), CCSA (China), ETSI (Europe), TTA (Korea), and TTC (Japan). WCDMA is called Universal Terrestrial Radio Access (UTRA) within 3GPP. It covers two modes of operation, that is UTRA-FDD (Frequency Division Duplex) and UTRA-TDD (Time Division Duplex). UTRA-FDD uses paired frequency bands for *Uplink* (UL) and *Downlink* (DL) data transmission, whereas UTRA-TDD utilizes a common frequency band for both directions and adjusts the time domain portion assigned for UL and DL transmission dynamically. This thesis focuses on the UTRA-FDD mode as the UTRA-TDD mode is currently of no commercial relevance. The term WCDMA is thus in accordance with the respective literature used as a synonym for UTRA-FDD throughout the thesis.

In the following the UTRA-FDD air interface and its operations are described. This thesis bases on the major Release 99 of the 3GPP specification of UMTS and consequently also this chapter focuses on Release 99. Recent developments with respect to *High Speed Packet Access* (HSPA) specified in 3GPP Releases 5 and 6 are addressed in Section 2.2.4.

2.1 UMTS Network Architecture

The UMTS network comprises a number of logical network elements. In the 3GPP standards their functionalities as well as the interfaces between them are specified. The principle architecture of a UMTS network, generally referred to as *Public Land Mobile Network* (PLMN), is presented in Figure 2.1.

A PLMN is operated by a mobile network operator. It can be logically divided into *User Equipment* (UE), *Universal Terrestrial Radio Access Network* (UTRAN), and *Core Network* (CN). In Figure 2.1 the external networks as potential peering partner for a PLMN illustrate the transition to other PS or CS networks. Although the elements of the PLMN are in principle defined as logical units, the specification of open interfaces gives rise to the fact that the logical elements typically also correspond to physical network elements, which can be purchased from competing manufacturers.

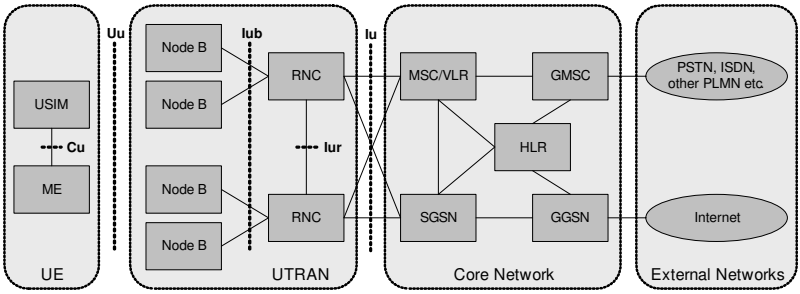


Figure 2.1: Basic architecture of a UMTS network (PLMN) [HT04]

User Equipment (UE)

The UE is the only mobile element in the UTRAN. Velocity and location of the UE considerably influence the performance of the network.

From standardization perspective the UE consists of two elements: The *Mobile Equipment* (ME) that represents the actual user terminal, and the *UMTS Subscriber Identity Module* (USIM), which is an application that runs on the smart card. The USIM holds subscriber information and stores authentication and encryption keys. USIM and ME communicate via the Cu interface. The UE accesses the fixed part of the network through the Uu interface.

UMTS Terrestrial Radio Access Network (UTRAN)

The UTRAN is responsible for managing and operating the radio access to the UE. It comprises one or more *Radio Network Sub-System* (RNS) each of which consists of one *Radio Network Controller* (RNC) and one or (typically) more Node Bs. The Node B denotes the *Base Station* (BS) in a UMTS network. It basically performs the physical layer operations such as channel coding, interleaving, modulation, etc., but also conducts some basic RRM operations (cf. Section 2.3). The introduction of HSPA (cf. Section 2.2.4) is accompanied by a significant increase of functionality in the Node B.

The RNC hosts most of the RRM functionality, e.g. call admission, and (intra-system) handover control (cf. Section 2.3). The RNC interfaces with the Node B through the Iub interface and with the CN via the Iu interface. The UTRAN allows *Soft-Handover* (SHO), i.e. a user can be connected to multiple Node Bs at a time. If these Node Bs belong to separate RNS, the Iur is used to exchange information relevant for the SHO operation.

Core Network (CN)

In contrast to UE and UTRAN, the CN has been adopted from GSM in order to allow a smooth interoperability between GSM and UMTS. The main elements of the CN are presented in Figure 2.1.

The *Home Location Register* (HLR) is a central database in the PLMN that stores all relevant subscriber information. This information includes associated telephone numbers, supplementary services, access priorities, and security keys. Besides these quasi-permanent data, the HLR also stores the current location of the subscriber, which is dynamically updated.

The CN comprises at least one *Mobile Services Switching Center* (MSC). The MSC is an intelligent digital switching center which is responsible for the call handling of *Circuit-Switched* (CS) calls: call setup, routing, inter-system mobility management, etc. The MSC interfaces through the Iu (Iu-CS) with the RNC and via the *Gateway MSC* (GMSC) with mobile users in other mobile or fixed networks. Typically integrated into the MSC is a *Visitor Location Register* (VLR). The VLR is a database which holds copies of the service profiles and position information of subscribers currently served by the MSC. These information are basically a subset of the information stored in the HLR. Whereas MSC and GMSC are dedicated to the CS domain, *Serving GPRS Support Node* (SGSN) and *Gateway GPRS Support Node* (GGSN) conduct equivalent operations for the PS traffic (PS domain).

2.2 UTRA-FDD Protocols

The high level purpose of the UTRAN is to provide UMTS bearer services to the CN. A bearer service is characterized by the supplied resources, in particular in terms of data rates, and its QoS requirements.

The UTRAN protocol stack is shown in Figure 2.2. It comprises the *Physical Layer* (PHY) (layer 1), the data link layer (layer 2), and the network layer (layer 3) [3GP99e]. Figure 2.2 only contains those parts of the protocol stack that are part of the UTRAN. In layer 3 there are additional higher level protocols and functionalities, e.g. call control and mobility management, which are subject of the CN, transparent to the UTRAN, and thus not described here. With Release 99 layer 1 is located in the Node B. Layer 2 and the UTRAN part of layer 3 are located in the RNC. Within the UTRAN, the *User Plane* (U-Plane) that denotes the part that processes user data and the *Control Plane* (C-Plane) that handles control information are distinguished.

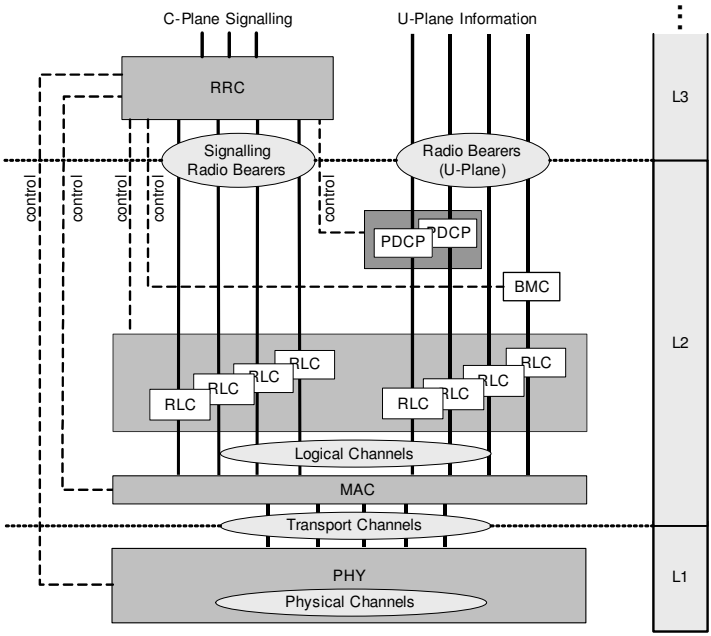


Figure 2.2: UTRAN protocol stack [3GP99e]

2.2.1 Physical Layer

The PHY owes its name to the fact that it is responsible for the actual physical transmission of data, i.e. baseband processing, channel coding, modulation, etc. The PHY provides transport channels for the transmission of layer 2 data and control information and maps them onto corresponding physical channels. The Release 99 transport channels and physical channels and the mapping between them are shown in Figure 2.3.

Transport Channels

The transport channels can be divided in dedicated and common transport channels. While a dedicated channel carries information for one user only, common channels carry information for multiple users. Depending on the type of common channel, the transmitted data might either contain time-multiplexed data dedicated to individual users or broadcast information.

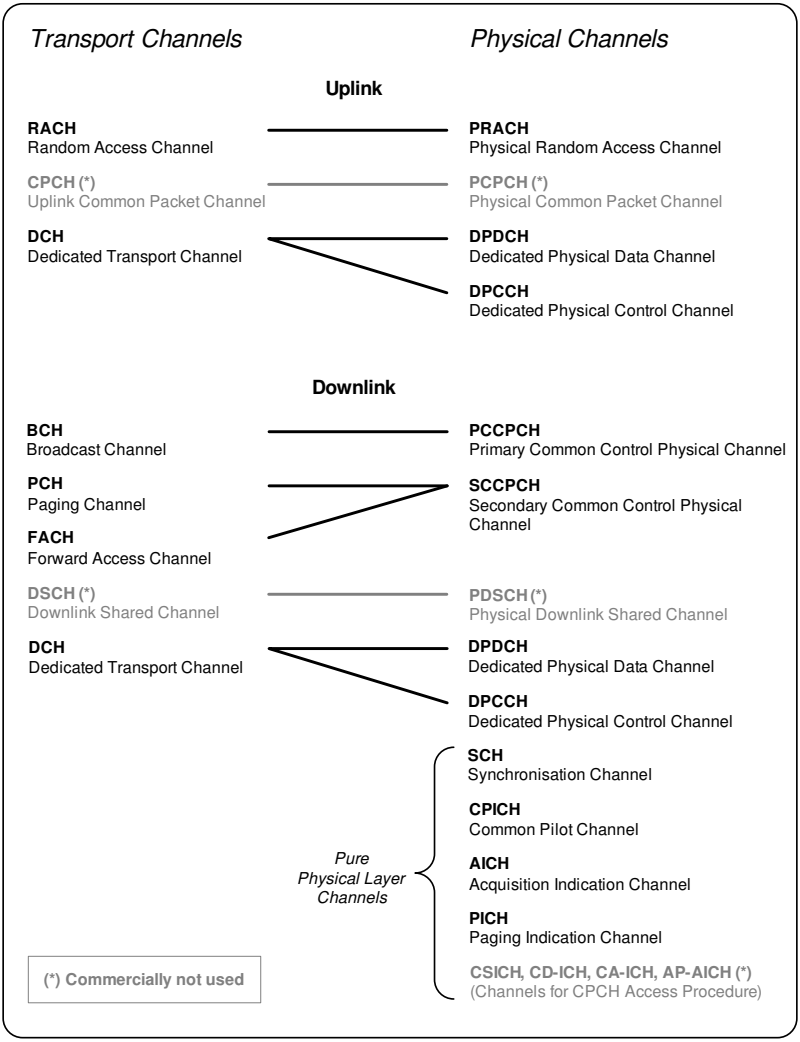


Figure 2.3: UTRAN transport channel to physical channel mapping (3GPP Release 99)

The individual Release 99 transport channels are:

DCH The *Dedicated Transport Channel* (DCH) is the only dedicated channel in Release 99. The DCH is typically a duplex channel. It carries user data and higher layer control information for a particular user.

BCH The *Broadcast Channel* (BCH) transmits information that are specific for the particular cell or *Universal Terrestrial Radio Access* (UTRA) network. Among these are the codes and slots to be used in the cell for the *Random Access Channel* (RACH) procedure (see below). Without these essential information a terminal would not be able to register with the network. The BCH thus needs to be received in the whole cell even by low-end terminals.

FACH The *Forward Access Channel* (FACH) is a DL low data rate channel basically used to transmit control information to particular users, e.g. during call setup. In addition it can be used to transmit small amounts of user data. It is possible to configure multiple FACH in a cell. At least one of them needs to be receivable by all terminals in the cell.

PCH The *Paging Channel* (PCH) is a DL channel used in the paging procedure, i.e. to locate and wake up terminals the network wants to initiate a communication with (e.g. for call setup).

RACH The RACH is a UL channel which is used to transmit control information from the terminal to the network, e.g. to register with the network or to set up a mobile user originated call.

CPCH/DSCH The *Uplink Common Packet Channel* (CPCH) and the *Downlink Shared Channel* (DSCH) are shared channels, which were included in the Release 99 specification. They are, however, not mandatory for the operation of the network and have never been (widely) used commercially. With the recent introduction of HSPA in Release 5 and Release 6 (cf. Section 2.2.4) they are obsolete and will in the following not be further considered.

Spreading and Scrambling

With UTRA-FDD, UL and DL channels are transmitted in separate frequency bands. *Direct Sequence Code Division Multiple Access* (DS-CDMA) [Pro95, Vit95, LM98, Rap02] is applied to separate different transmissions from the same source, i.e. the individual cells in DL and user equipments in UL. Each bit from the input data streams is multiplied with a *channelization code*, resulting in a stream of higher data rate. Consequently, the original narrowband input signal is transformed into a signal of larger frequency spectrum. The procedure is thus widely called *spreading*, and the channelization code is also referred to as *spreading code*. Channelization codes are taken from a tree of orthogonal codes (cf. Figure 2.4) and chosen in such a way that the spreading results in the same characteristic physical channel bit rate for all channels, called chip rate (UMTS: 3.84 Mbps), and thus the same characteristic width of the resulting frequency spectrum. For the transmission of different channels from the same source, channelization codes are taken from parallel branches of the code tree. These codes are orthogonal and the spread signals, if synchronized, theoretically do not interfere with each other. In practice however, this orthogonality is not perfect due to multi-path propagation effects and imperfect receiver hardware, resulting in a certain amount of mutual interference.

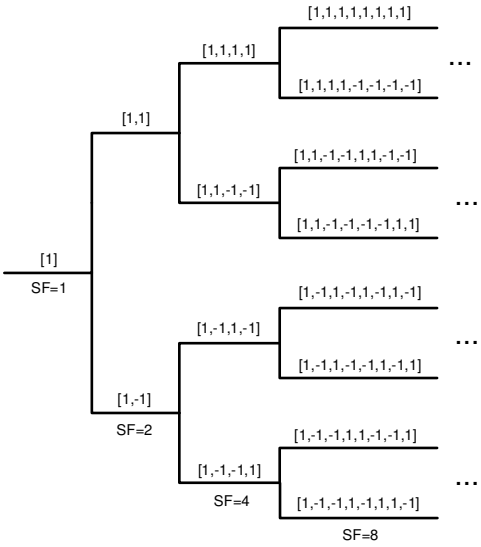


Figure 2.4: Tree of orthogonal spreading codes

The signals from different transmitters are on top of the spreading multiplied with *scrambling codes*. Scrambling codes are pseudo-random bit sequences which have the same bit rate as the signals they are multiplied with. They are known to both transmitter and receiver. In this way the signals from different transmitters can be separated, although they potentially use identical channelization codes and transmit unsynchronized. A fundamental characteristic of CDMA is, that despreading and descrambling raise the signal from the interference by the *Spreading Factor* (SF), i.e. the length of the channelization code in bits (cf. Figure 2.4).

In case the code tree of a transmitter is fully utilized, an additional code tree in combination with a different scrambling code is used. In this case the signals from one transmitter (but different code trees) interfere in the same way as the signals from different transmitters.

Physical Channels

The DCH channel is mapped on a combination of *Dedicated Physical Control Channel* (DPCCH) and *Dedicated Physical Data Channel* (DPDCH). These are power controlled with a fixed power offset between DPCCH and DPDCH. The DPCCH has a fixed data rate and contains control information, in particular the power control commands. The DPDCH is used for the actual (user) data transmission and has a variable data rate, which

might even be adapted dynamically. A DPDCH requires a DPCCH in both directions, so that for a unidirectional DCH three physical channels need to be set up.

The transmission of DPCCH and DPDCH slightly differs in UL and DL direction. In UL DPCCH and DPDCH are transmitted in parallel. For the DPCCH a channelization code with a fixed SF of 256 is used. The DPDCH can have a spreading factor of 4 to 256, which is in UL dynamically adapted to the instantaneous data rate. DPCCH and DPDCH are processed in parallel and then mapped on the I and Q branches of the *Quadrature Phase Shift Keying* (QPSK) constellation, which is used as modulation scheme [3GP99h]. The latter procedure is referred to as I-Q/code multiplexing. The I-Q/code multiplexing in UL is applied in order to avoid audible interference due to *Discontinuous Transmission* (DTX) [HT04], as for example known from GSM. In DL audible interference is no problem, because there is typically a considerably larger distance between transmitter and the equipment that might be sensitive to audible interference, e.g. television sets, radios, ATM. Hence, DPCCH and DPDCH are in DL time multiplexed on a channel with fixed spreading factor with a possible range between 4 and 512. Variable data rates are in DL achieved by applying DTX with varying activity factors for the data transmission. For the DL transmission of DPDCH and DPCCH a QPSK modulation is applied.

The RACH is mapped on the *Physical Random Access Channel* (PRACH). A user requests for data transmission on the RACH by sending a preamble with an initial power setting calculated from a pilot channel based path loss estimation. The transmission is answered by the UTRAN by transmitting an acknowledgment on the *Acquisition Indication Channel* (AICH). If no acknowledgment is received, the mobile retransmits the preamble with higher power. This iterative procedure is referred to as *open-loop power control* (cf. Section 2.3). Once the acknowledgment is received, the mobile starts to transmit the actual data part with a spreading factor provided by the UTRAN. The complete RACH procedure is handled by the physical layer. The AICH is thus not visible above layer 1.

The BCH is mapped on the *Primary Common Control Physical Channel* (PCCPCH). For the PCCPCH a fixed channelization code with spreading factor 256 is used. The BCH needs to be received by all mobiles in the cell. For the operation of the UTRAN it is thus essential that the PCCPCH power is configured to be large enough.

FACH and PCH are both mapped on the *Secondary Common Control Physical Channel* (SCCPCH). They can be either time-multiplexed on one SCCPCH or mapped to distinct SCCPCH. In particular if more than one FACH is set up, also multiple SCCPCH need to be set up. The spreading factor of the SCCPCH can be chosen depending on the required data rate.

Common Pilot Channel (CPICH), *Synchronization Channel* (SCH), and *Paging Indication Channel* (PICH) are signaling channels for PHY signaling only. The SCH is used by the terminals to synchronize with the cell and in addition transmits information to simplify the identification of the scrambling code used by the cell, i.e. during *cell search procedure*. The SCH is the only channel that does not use the cell specific primary scram-

bling code. The CPICH is used by the mobiles to estimate the DL channel conditions, which are important for initial power settings, handover operations, etc. The PICH is used together with the PCH (transmitted on the PCCPCH) to save battery lifetime. Instead of constantly listening to the PCCPCH for potential paging messages, a mobile only needs to periodically listen to a flag transmitted on the PICH for the *paging group* it has been assigned to. With this flag the UTRAN indicates that there will be a paging message for the particular paging group. Hence only in this case mobiles need to carry out the power consuming decoding of the message on the PCCPCH.

For all channels either turbo or convolutional channel coding is applied, where convolutional coding is used for the lower data rate channels [HT04].

Physical Layer Procedures

Apart from the basic operations on the channels, e.g. spreading, modulation, and coding, a set of essential procedures with higher complexity is also carried out by the PHY. Among them is a part of the power control functionality. Power control is part of the RRM and is described in more detail in Section 2.3. Further physical layer procedures are

- the *cell search procedure* carried out by the mobile to synchronize with the network.
- the *paging* procedure applied by the network to locate a particular UE.
- the *RACH procedure* for transmitting UL data on the PRACH.
- the *transmit diversity* procedure.
- several channel *synchronization and reconfiguration* procedures.

The complete set of physical layer operations is specified in [3GP99d].

2.2.2 Data Link Layer

The data link layer (layer 2) provides *radio bearer* services for transmitting user data and *signaling radio bearer* services for transmitting control information to the network layer. The data link layer comprises two sub-layers: the *Medium Access Control* (MAC) for mapping the radio bearers to transport channels and the *Radio Link Control* (RLC), which is responsible for an appropriate handling of individual data streams, e.g. retransmission of erroneously received packets, and segmentation and reassembly of data packets. Additionally there are two optional protocols which might be used between RLC and network layer, i.e. the *Packet Data Convergence Protocol* (PDCP) and the *Broadcast/Multicast Control Protocol* (BMC). The PDCP is applied for PS radio bearers in the U-Plane. The PDCP performs *Internet Protocol* (IP) header compression and it handles the forwarding of data in case the UE is leaving the serving area of one RNC (“SRNS relocation”) [3GP99c]. The BMC is responsible for transmitting cell broadcast messages, which can be received by all users in a cell [3GP99a]. In Release 6 the *Multimedia Broadcast Multicast Service* (MBMS) is introduced for transmitting messages to a group of users [3GP04a].

The main advantage over the BMC is that it provides the possibility to restrict the user groups and thus allows charging.

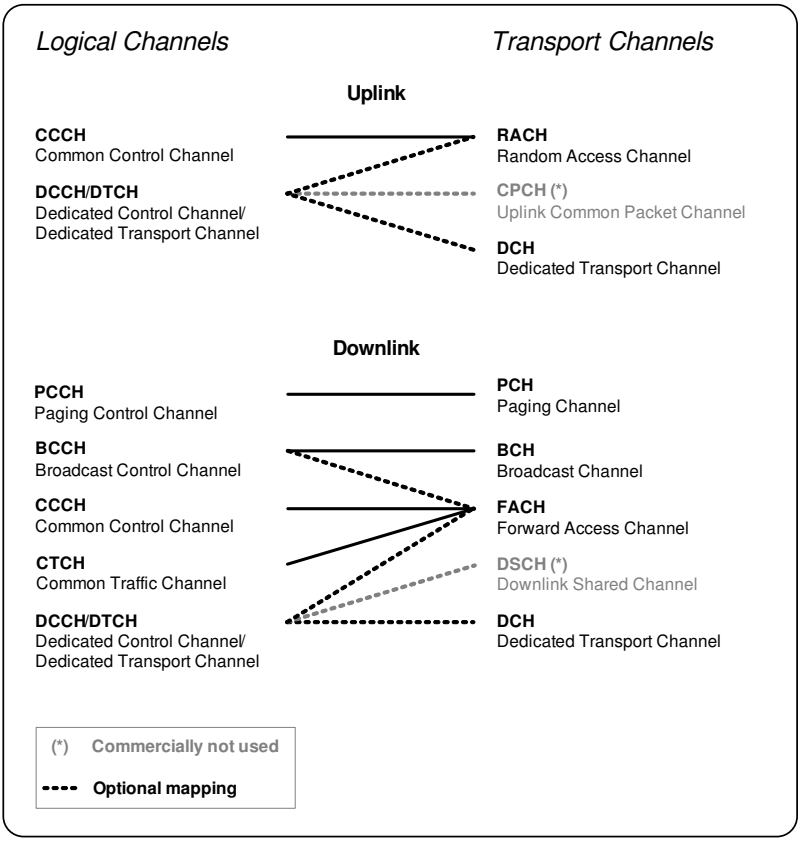


Figure 2.5: UTRAN logical channels to transport channel mapping (3GPP Release 99)

Medium Access Control (MAC)

The overall purpose of the MAC is to map logical channels onto transport channels (cf. Figure 2.5):

DCCH/DTCH The *Dedicated Control Channel* (DCCH) and the *Dedicated Traffic Channel* (DTCH) are dedicated logical channels used for transmitting control and user data respectively. If only small data volumes are to be transmitted, they can be

mapped on FACH and RACH in order to save air interface resources. Larger data volumes are transmitted via the DCH.

CTCH The *Common Traffic Channel* (CTCH) is a DL point-to-multi-point traffic channel mapped on the FACH. It is in particular used by the BMC protocol.

CCCH The *Common Control Channel* (CCCH) is a bidirectional control channel used for signaling between UE and UTRAN. It is mapped on RACH and FACH respectively.

BCCH On the *Broadcast Control Channel* (BCCH) system control information are broadcast in DL. It is mapped on the BCH with the option to transmit additional information also on the FACH.

PCCH The *Paging Control Channel* (PCCH) is the logical DL channel for transmitting paging information. It is always mapped on the PCH.

Besides the channel mapping, there is a large number of operations – most closely connected to the channel mapping – which are carried out by the MAC [3GP99b]. These in particular include the *priority handling* between data streams to/from the same UE, *multiplexing/demultiplexing* of higher layer data, and *channel type* and *transport format switching* [HT04].

Radio Link Control (RLC)

The RLC layer provides segmentation/assembly services to layer 3 or BMC/PDCP respectively [3GP99f]. It can be operated in *acknowledged mode*, *unacknowledged mode*, and *transparent mode*. In transparent mode the RLC does not add any control information to the data. If erroneous layer 1 data units are received, they are either marked or discarded. Unacknowledged mode adds sequence numbers to the data packets. Hence a reordering of data units and a detection of multiple received data units can be carried out. The acknowledged mode, which in contrast to the other modes is bidirectional, introduces *Automatic Repeat Request* (ARQ) to trigger the retransmission of lost or corrupted data units. Hence, the correctness and consistency of the data provided to the higher layers can be guaranteed. The selection of the appropriate mode depends on the type of service, e.g. for streaming services typically the transparent mode is used, whereas for packet data the acknowledged mode is applied.

2.2.3 Network Layer

The *Radio Resource Control* (RRC) is a layer 3 protocol responsible for most of the signaling carried out in the UTRAN [3GP99g]. The RRC messages contain all parameters required to establish, modify and release layer 1 and layer 2 entities, and the RRC also handles all mobility signaling (handover, cell updates, etc.). Most of the RRM functionality described in Section 2.3 is part of the RRC.

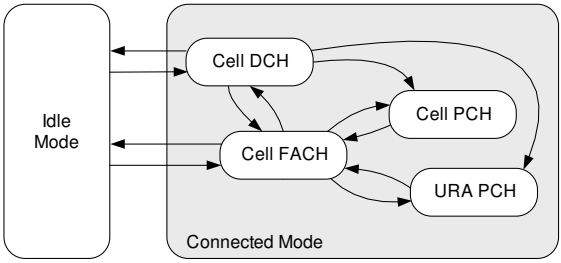


Figure 2.6: UE Modes and RRC states in connected mode [HT04]

In the context of RRC, the UE can be in two basic modes of operation: *idle mode* and *connected mode*. Being in idle mode, also referred to as "camping on a cell", the UE is synchronized to the cell, and listens to its control channels. However, there are no resources allocated in the UTRAN for the particular mobile, and the location of the mobile is only known in the CN in terms of the *location area* which typically consists of a large amount of cells [HT04]. For a RRC connection setup, a paging is required in the complete location area.

The connected mode comprises four states: *Cell DCH*, *Cell FACH*, *UTRAN Registration Area* (URA) *PCH*, and *Cell PCH* (cf. Figure 2.6). In Cell DCH state, a dedicated channel is allocated to the UE. Serving cell(s) are known by the RNC. In FACH State, FACH and RACH are used for the transmission of U-Plane and C-Plane data. In case of PS services the UE may be switched between Cell FACH and Cell DCH states dynamically depending on the amount of data to be transmitted.

In Cell PCH state the mobile has no active connection with the network, but RRC entities are allocated and the cell ID of the UE is stored in the RNC. The UE listens on the PCH for potential paging information. The URA PCH state is very similar to the Cell PCH state, with the difference that in URA PCH state the exact location of the UE is not known. The UE is only known to be located somewhere in the URA. The URA comprises several cells, and can but does not necessarily need to be equal to the location area. The main difference between URA PCH and Cell PCH state is, that in the latter case, the UE has to indicate its new serving cell every time it changes from one cell to the other, whereas in URA PCH state this only needs to be done when leaving the URA area. The selection between Cell PCH and URA PCH state is thus a trade-off between signaling traffic due to mobility and signaling traffic for the paging operation.

2.2.4 High Speed Packet Access Extensions

The WCDMA standard has recently been extended towards HSPA. With 3GPP Release 5 and Release 6, *High Speed Downlink Packet Access* (HSDPA) and *High Speed Uplink*

Packet Access (HSUPA) have been introduced. They provide high data rate UL and DL channels mainly applicable for PS services. HSDPA and HSUPA have the following characteristics in common [HT06, 3GP01, 3GP05]:

Fast Node B based scheduling: The scheduling of resources to users is adapted dynamically by the Node B. For this, a part of the MAC protocol is moved from the RNC to the Node B. The scheduling principles, however, differ between HSUPA and HSDPA.

Adaptive Modulation and Coding (AMC): HSDPA and HSUPA support various different coding schemes, which can be adapted to the current link conditions. HSDPA also allows to switch the modulation from QPSK to 16QAM.

Shorter TTI: The *Transmission Time Interval* (TTI) denotes the length of blocks the input data is segmented into for coding and transmission. With Release 99 the TTI can be between 10 ms and 80 ms. A TTI of 2 ms is used with HSDPA, with HSUPA either 2 ms or 10 ms are possible. A shorter TTI allows to react faster to changes in the link conditions and to quickly change the transmission between users depending on instantaneous link conditions. Also the delay for a retransmission can be reduced (when applying ARQ).

Hybrid ARQ: Both, HSDPA and HSUPA, apply *Hybrid Automatic Repeat Request* (HARQ). HARQ is equivalent to ARQ with the difference that the erroneously received data blocks are not discarded but (soft) combined with the blocks received in successive retransmissions. Hence, retransmissions can be sent with lower power and a higher system capacity can be achieved.

With HSDPA theoretical peak data rates of up to 14.4 Mbps can be achieved. HSUPA supports theoretical data rates of up to 5.76 Mbps.

High Speed Downlink Packet Access (HSDPA)

HSDPA introduces a new transport channel type, i.e. the *Highspeed Downlink Shared Channel* (HS-DSCH). The HS-DSCH is a shared channel, allowing the time- and code-multiplexing of several users. It is not power controlled, instead adaptive modulation (QPSK and 16QAM) and coding is used to adapt to the current link condition. Link condition estimates are reported by the UE via a feedback channel. The scheduling is conducted on TTI basis. Up to 15 parallel codes (SF 16) can be allocated in parallel to the HS-DSCH channel. The number of codes per user depends on the terminal capabilities. In order to limit the signaling overhead, the HS-DSCH does not support SHO.

The HS-DSCH is mapped on three physical channels. The actual data transmission is carried out on the *High-Speed Physical Downlink Shared Channel* (HS-PDSCH). In parallel the *High-Speed Shared Control Channel* (HS-SCCH) is transmitted, which provides information on the content of the HS-PDSCH, like the currently served UE, the modulation and coding scheme, etc. The UE sends feedback on the current channel conditions

in terms of a *Channel Quality Indicator* (CQI) on the *High-Speed Dedicated Physical Control Channel* (HS-DPCCH).

High Speed Uplink Packet Access (HSUPA)

HSUPA introduces an advanced dedicated UL transport channel, denoted as *Enhanced Dedicated Channel* (E-DCH). The E-DCH is mapped on pairs of physical channels, namely the *Enhanced Dedicated Physical Data Channel* (E-DPDCH) and the *Enhanced Dedicated Physical Control Channel* (E-DPCCH). The E-DPDCH is used for the actual user data transmission. Coding scheme and spreading factor of the E-DPDCH can be dynamically adapted on TTI basis to the instantaneous link quality. The E-DPCCH signals the coding scheme and spreading factor associated with the momentary data transmission to the Node B. The E-DPDCH requires Release 99 DPCCH channels for transmission of power control commands and for synchronization purposes [HT06]. For the *Hybrid Automatic Repeat Request* (HARQ) operation, the Node B uses the new *E-DCH HARQ Indicator Channel* (E-HICH) to send acknowledgments to the UE.

The scheduling of resources is with HSUPA significantly different than in case of HSDPA. Instead of time-multiplexing users, the data rate assigned to HSUPA users is controlled in terms of granted power levels that the UE is allowed to transmit the E-DPDCH channels with. This power is thereby defined by a power offset with respect to the DPCCH. The UE dynamically adapts the data rates of the E-DPDCH channels to best utilize the granted power. HSUPA uses two physical channels to set and modify the individual power grants. The *E-DCH Absolute Grant Channel* (E-AGCH) is a broadcast channel that allows to set an absolute power offset for typically a group of UE. On the *E-DCH Relative Grant Channel* (E-RGCH) power up and down commands are transmitted to the individual UE. Whereas the UE only listens to the E-AGCH of one distinct primary serving cell, relative grants can be transmitted also by the other SHO cells. The objective of the scheduling is to exploit the cell resources in terms of maximum UL interference levels as much as possible, but to also assure a certain fairness among users and cells. The actual scheduling algorithm is vendor specific. Some more details can be found in [HT06].

2.3 Radio Resource Management

The *Radio Resource Management* (RRM) is responsible for an efficient utilization of the air interface resources and for provisioning and guaranteeing QoS. The WCDMA RRM algorithms comprise [HT04] *power control*, *handover control*, *admission control*, *congestion control (load control)*, and *packet scheduling*. The RRM functionalities are spread over all three UTRAN protocol layers and make use of the protocols described in the previous section, in particular the RRC protocol. An appropriate consideration of the RRM mechanisms is of essential importance for the accuracy of the performance analysis methods developed in this thesis.

Apart from Section 2.2.4 and Section 5.7.2 which briefly describe the principles and the modeling of the HSPA features introduced in 3GPP Release 5 and Release 6, this thesis is targeting at the analysis of Release 99 WCDMA networks. Hence, packet scheduling, which is introduced with HSPA, will not be further considered. The other RRM algorithms are in the following addressed in more detail.

2.3.1 Soft Handover

Handover operations, i.e. changes between cells, are required in cellular networks to support the mobility of users. In WCDMA, three handover types are distinguished: *intra-frequency handover*, *inter-frequency handover*, and *inter-system handover*. The latter two handover types are for changing between different carriers of the same WCDMA network and for changing between a WCDMA network and a network of different technology standard (e.g. between UTRAN and GSM) respectively. Both are *hard handovers*, that is, at any time of the handover operation only one connection is maintained. In contrast in case of intra-frequency handovers typically *Soft-Handover* (SHO) is applied. With SHO, in the overlapping regions between cells, multiple links are maintained in parallel. The main advantages of SHO are a seamless handover and extended cell ranges because the UE can in SHO transmit with lower power. As this thesis mainly focuses on the performance analysis of one carrier of a WCDMA network, inter-frequency and inter-system handover will not be further considered in this chapter. The modeling of the hard handovers is discussed together with the extension towards a multi-carrier analysis in Section 5.7.1.

SHO can be classified into *softer handover* and *soft handover*. With softer handover a UE finds itself in multiple overlapping cells of one Node B. In DL direction the UE communicates simultaneously through separate channels (DL codes) for each cell. In UL the signal transmitted by the UE is decoded by each of the associated cells. The separate communication paths are generally considered as the *branches* of a connection in SHO. In softer handover the received signals from the different branches are combined within one receiver (in the UE for DL and in the Node B for UL respectively).

The UE is referred to as being in soft handover, but not softer handover, if the UE communicates with multiple cells of different Node Bs. In DL direction communication is similar to the procedure in softer handover, i.e. the signals from the different cells are combined in the receiver of the UE. In UL direction the SHO branches cannot be processed in one receiver, as the Node Bs are physically separated. The different branches are consequently despread and decoded separately and routed to the RNC for combining. The RNC picks the error-free blocks based on the outcome of the cyclic redundancy check. This procedure is called *selection combining*.

In general soft and softer handover may take place simultaneously for one connection. The decision on which cells to communicate with is triggered by the RNC based on the received pilot channel power to interference ratio, denoted Ec/I0. The UE constantly mea-

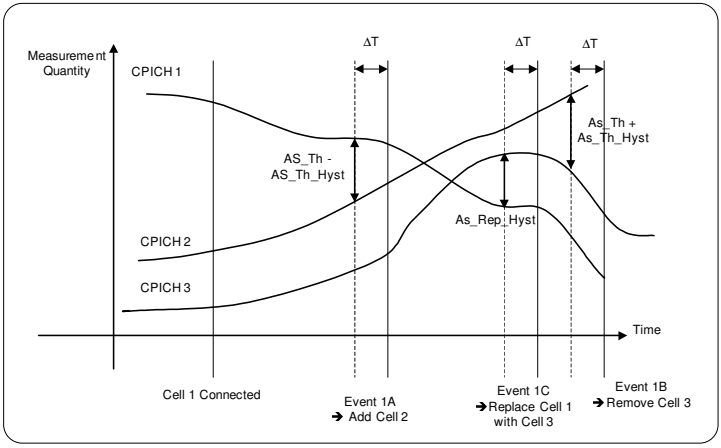


Figure 2.7: Illustration of the WCDMA soft handover algorithm [3GP04b]

sures the pilot channels of the neighbor cells and signals them to the RNC [3GP04b]. The following terminology is used:

Active set: The cells in the active set form a soft handover connection to/from the mobile station.

Candidate set: List of cells that are currently not in the active set but which are strong enough to be added to the active set.

Monitored set: List of cells that are continuously measured by the mobile station.

The basic SHO algorithm is illustrated in Figure 2.7. The algorithm takes those cells from the monitored set to the candidate set, whose Ec/I0 is in a relative margin from the Ec/I0 of the best cell. The candidate set is limited in size (typical maximum size: 3) so that among the cells fulfilling the criterion only the best ones are chosen. In order to avoid "ping-pong" effects, a hysteresis is introduced. As long as the active set is not full, cells from the candidate list are successively added to the active set beginning with the best cell (*Event 1A*), provided that sufficient resources are available in the cell. Cells are removed from the active set if its quality with reference to the best cell is worse by more than a certain margin (*Event 1B*). If the active set is full, the cell in the active set received with worst quality is replaced by the best cell from the candidate set if the measured Ec/I0 of the candidate cell exceeds the Ec/I0 of the active cell by more than a certain threshold (*Event 1C*).

2.3.2 Power Control

Managing power is of particular relevance for the operation of WCDMA networks. Users share the same bandwidth at the same time, thus interfere with each other. Power control is applied in order to set the *Signal to Interference plus Noise Ratio* (SIR) for each connection to its target level. The target level is the minimum level for which the desired link quality is retained and hence the interference is maximally reduced. Besides its role in interference reduction, power control also helps to overcome effects like the near-far effect well known from literature [Rap02]. Within WCDMA, open- and closed-loop power control are distinguished.

Open-loop Power Control

Open loop power control is used during the RACH procedure, where no direct feedback loop between UE and Node B is available. The transmit power has to be adjusted autonomously by the UE. The principle procedure is, that the UE starts to transmit with some initial power, which is derived from a pilot based channel estimate. The UE subsequently retransmits the data with increasing power until an acknowledgment from the corresponding cell is received.

Closed-Loop Power Control

Closed-loop power control is applied for adapting the power levels of DCH, E-DCH (and DSCH). It is used to adapt the transmit power to the time varying radio link conditions. Feedback from the receiving station is directly used to set the power levels at the transmitting station for both, the data channel and the corresponding control channel, which have a fixed relative power offset. The closed-loop power control can be characterized by a two stage process, comprising two feedback loops: the *inner-loop* and the *outer-loop*.

Inner-loop power control, also referred to as the *fast power control* loop, is between UE and Node B. It aims at adapting the transmit power to the fast varying radio link conditions mainly caused by fast fading. The SIR value is estimated at the receiving station (the Node B in UL, and the UE in DL) and a power up or down command is sent to attain the SIR target. Power control commands are sent at a rate of 1500 per second. If the UE is in SHO, the mobile receives for the UL power control commands from multiple cells. All power control commands are decoded, the power is increased only if all cells are requesting a power up, otherwise a power down step is performed. This situation does not apply in DL power control, because the same power control command is sent by the mobile to all serving cells. However, in DL a problem arises if the power control command is received incorrectly by a part of the serving cells and cells consequently do not maintain the same power levels. In order to overcome this effect, referred to as power drifting, a power balancing function is proposed in [3GP99d].

The outer loop is subject to the RNC. It aims at controlling the SIR target of the inner-loop power control in such a way that the *Block Error Rate* (BLER) of the transmission matches its target value. For this purpose the RNC measures the BLER, adapts the SIR target, and communicates it to the Node B. SIR target variations are mainly due to different flavors of user equipment, mobility effects, and different propagation environments.

2.3.3 Admission Control

The capacity of WCDMA networks is power and interference limited. An appropriate decision on the addition of new bearers is of great importance for a well working system. The *Admission Control* (AC) has to take care that QoS requirements of all existing bearers in combination with the new candidate bearer can be maintained. AC is invoked by several functions including bearer setup, bearer modification, channel switching and branch addition during soft handover.

Most proposed AC algorithms, e.g. [LE94, HY96, HCLS01], for WCDMA networks are load based. Focus is to keep the individual cell loads and hence the network load below some value. The cell load is defined for UL and DL separately. In DL direction it is defined straight forward as the total output power of the Node B divided by its maximum transmission power. In UL direction the load is defined as one minus the total receiver noise over the total received interference plus noise. Hence, if there are no interferers in the system, the UL load is zero, in contrast a load of one would imply an infinite interference level.

AC algorithms are proprietary and not part of the 3GPP specification. In this thesis a simple but yet realistic AC algorithm is considered. It is in alignment with the “Wideband Power-based Admission Control Scheme” described in [HT04]. In case of an admission decision the cell load after virtually admitting the call is estimated based on current load measurements. A call is admitted, if the maximum load level, which is a parameter of the algorithm, is not exceeded. In order to allow prioritization between services, different maximum load levels for different traffic types can be defined. A special situation occurs if a new high priority call arrives that cannot directly be served, because admitting this call would lead to an excessive load level. In this case bearers with lower priority could be assigned a lower data rate or they might even be dropped.

In addition to the load based admission procedure, code restrictions and hardware limitations, e.g. channel cards in the base station, have to be taken into account by the AC.

2.3.4 Congestion Control

Congestion is defined as a situation where the cell load becomes too high. A congestion may occur due to traffic fluctuations, user mobility or channel variations. The task of the *Congestion Control* (CC) is to resolve and prevent such situations, which might in

particular in UL even lead to an instability of the system. The strategy behind CC is to lower the power levels in the system, hence forcing the system back to a stable state.

Suitable indicators for congestion are that the measured interference levels (UL) or the total output power levels are above some thresholds. Actions to lower these levels and thus to resolve congestion are based on traffic priorities and similar to the actions taken by the AC, i.e. lowering of data rates and (eventually) dropping of users. After the congestion has been resolved, downgraded users can be upgraded to provide them the requested data rate.

Chapter 3

Modeling the Wireless Transmission Channel

The transmission of information via a wireless channel is typically modeled as shown in Figure 3.1 [Sha48, Pro95, Sau99]. The upper part of the figure applies to any communication system: The information generated by the *source* is transformed into an adequate signal by the *transmitter* and sent via the communication *channel*. The channel might modify or even corrupt the signal. In the receiver the original information is aimed to be recovered and sent to the destination. Typical operations of a transmitter in a wireless communication system are: *source coding*, *channel coding*, and *modulation*. In the receiver the inverse operations are carried out.

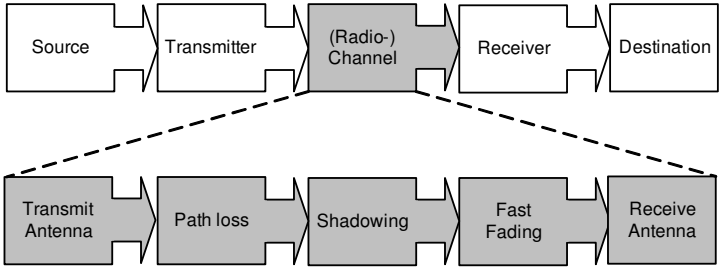


Figure 3.1: Radio propagation channel

In mobile networks, signals are sent from the *Base Station* (BS) to the *Mobile Station* (MS) and vice versa[1]. The communication path from the BS to the MS is denoted as DL, the path from the MS to the BS is called UL. The channel is modeled by five components: *transmit antenna*, *path loss, shadow fading*, *fast fading*, and *receive antenna*.

[1] The general terms BS and MS correspond to the WCDMA specific terms Node B and UE which have been introduced in Chapter 2

The median attenuation between transmit and receive antenna is called path loss. For simulations it is typically obtained based on models.

In addition there are two statistical components which multiplicatively contribute to the attenuation. The *shadow fading*, also referred to as *shadowing* or *slow fading*, statistically describes the attenuation by obstacles on the transmission path. Shadow fading models both, the uncertainty in the path loss component, if given by a path loss model, and the long term variations of the propagation environment, e.g. caused by moving obstacles (cars, etc.). The uncertainty in the path loss predictions is due to the attenuation by buildings, differences in terrain heights, vegetation on the propagation path, etc., which cannot be fully considered by the propagation models. The magnitude of the error strongly depends on the model which is used, in particular its complexity.

In contrast to shadowing, *fast fading* is – as suggested by its name – a very fast varying component. Fast fading is caused by destructive and constructive interference between signal components that reach the receiver via multiple transmission paths. Fast fading is often also called *multi-path fading*.

Whereas the path loss, shadowing, and fast fading are phenomena imposed by nature and can hardly be influenced by the system design (only their impact can be reduced by intelligent equipment and network design), antennas are an important means for improving the system quality. The main objective of network quality optimization is to reduce the attenuation of desired signals and to further attenuate interfering signals. As a consequence, directional high quality antennas are widely used in particular on the BS site. The modeling of antennas and antenna gain is a major ingredient of a radio performance analysis.

In this chapter the modeling of the above mentioned components of the wireless channel is explained. Focus is on the aspects required for an understanding of the subsequent chapters of this thesis.

3.1 Macro Path Loss Prediction

Accurate path loss predictions are of significant importance for the network design process (cf. Chapter 1). Various prediction models can be found in literature. In this section the basic concepts behind path loss prediction models are described and two of the most popular empirical models are presented.

3.1.1 Basic Propagation Models

Basic propagation models calculate the path loss under very simplified and ideal assumptions regarding the propagation environment. Due to these simplifications, the results are not of great significance for network planning. Nevertheless, basic propagation models are useful for an understanding of the more complex propagation models.

Free Space Loss

Free space loss describes the path loss experienced by a radio wave in free space. Assuming an ideal isotropic radiator transmitting power P_t, the power P_r received by an ideal receiver with isotropic antenna characteristic located at distance D from the transmitter is (e.g. [DS94]):

$$\frac{P_r}{P_t} = \left(\frac{\lambda}{4\pi D}\right)^2 = \left(\frac{c}{4\pi f D}\right)^2 \tag{3.1}$$

where c is the speed of light, f and λ are frequency and respective wavelength of the transmitted wave. Equation (3.1) is known as the *Friis transmission formula*.

The path loss in decibels consequently is

$$\Theta_d = 32.4 + 20\log(D) + 20\log\left(\frac{f}{\text{MHz}}\right) \tag{3.2}$$

The Friis formula particularly shows that the path loss increases with the square of the distance. Friis formula gives a lower bound on the propagation loss, but is too optimistic to be used for practical applications. For very rough approximations sometimes (3.2) is modified to yield a more realistic formula:

$$\Theta_d = 32.4 + 10 \cdot \beta \cdot \left(\log(D) + \log\left(\frac{f}{\text{MHz}}\right)\right) \tag{3.3}$$

where for the path loss exponent β depending on the environment typically a value between 3.5 and 5 is used.

Plane Earth Loss

In the plane earth loss model it is assumed that the signal received by the receiver consists of two rays which are superimposed at the receiver (Figure 3.2). The first ray is the direct ray the attenuation of which is given by free space loss, the second is the ray reflected from the surface below transmitter and receiver.

Assuming horizontally polarized antennas and a reflection coefficient of 1, the total path loss can be shown to be [Sau99]

$$\Theta_D = 40\log(D) - 20\log(h_{MS}) - 20\log(h_{BS}) \tag{3.4}$$

With plane earth loss the path loss increases with 40 dB per decade compared to 20 dB per decade in case of free space loss.

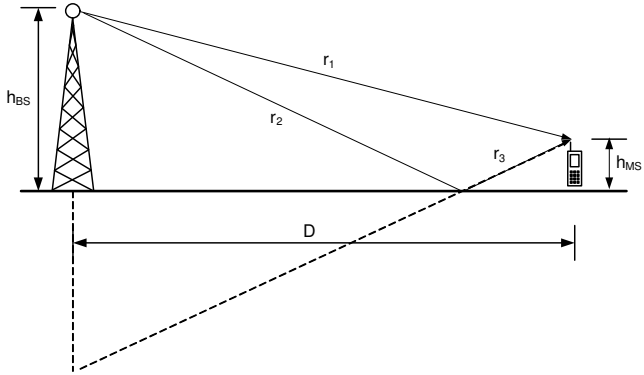


Figure 3.2: Considered constellation for plane earth loss

Dual-Slope Model

A common approach [CCK+02] is to combine free space loss and plane earth loss into a hybrid model. For distances below the *breakpoint* the free-space loss is used, for larger distances the plane earth loss model is applied. The breakpoint D^* is thereby defined as the distance at which the two models give an equal path loss value:

$$D^* = \frac{4\,\pi\,h_{MS}\,h_{BS}}{\lambda} \tag{3.5}$$

3.1.2 Empirical Prediction Models

The basic propagation models estimate the path loss based on some very simple assumptions. They in particular help to understand the principles of radio wave propagation and can be applied for very rough estimations. For planning purposes, however, they are inadequate. For network planning purposes, empirical or semi-deterministic models are widely used. These models are typically based on large measurement campaigns in different environments. From the measurements mathematical models are extrapolated.

Cost 231 - Hata

The most widely accepted propagation model is the Okumura Hata model. The model is based on measurements carried out by Okumura in and around Tokyo published in 1968 [OOKF68]. These measurements were later fitted by Hata into mathematical formulas [Hat80]. The resulting model is referred to as the Okumura Hata model.

Hata's formulas are only valid for frequencies up to 1000 MHz. They have been extended within the *Cost 231* activities [Kür99] to frequencies between 1500 MHz and 2000 MHz,

which are particularly of relevance for GSM1800 and UMTS. The resulting model is called the Cost 231-Hata model. Its basic path loss formula is

$$\begin{aligned}\Theta(D) &= 46.3 + 33.9 \log\left(\tfrac{f}{\text{MHz}}\right) - 13.82 \log\left(h_{BS}\right) - a\left(h_{MS}\right) \\ &\quad + \left(44.9 - 6.55 \log\left(h_{BS}\right)\right) \cdot \log\left(\tfrac{D}{\text{km}}\right) + C_m\end{aligned} \tag{3.6}$$

where the mobile height specific term $a\left(h_{MS}\right)$ is defined as follows:

$$a\left(h_{MS}\right) = \left(1.1 \cdot \log\left(\frac{f}{\text{MHz}}\right) - 0.7\right)\ h_{MS} - \left(1.56 \cdot \log\left(\frac{f}{\text{MHz}}\right) - 0.8\right) \tag{3.7}$$

The parameter C_m is a correction factor which is chosen based on the particular propagation environment:

$$C_m = \begin{cases} 0 \text{ dB} & \text{for medium sized cities} \\ & \text{and suburban centers with} \\ & \text{medium tree density} \\ 3 \text{ dB} & \text{for metropolitan centers} \end{cases} \tag{3.8}$$

The application of the Cost 231-Hata model is restricted to macro cells above roof-top and cell radii between 1 km and 20 km (cf. Table 3.1).

Cost 231 - Walfish Ikegami

For micro cells and urban scenarios with small distances between BS and MS antennas, the Cost 231-Walfish-Ikegami model has been developed within Cost 231. The Cost 231-Walfish-Ikegami model is a combination of the Walfish-Bertoni and the Ikegami model [Sau99]. Compared with the Cost 231-Hata model it takes considerably more information on the propagation environment into account. In particular it distinguishes between *Line of Sight* (LOS) and *Non Line of Sight* (NLOS) conditions. In the LOS case, e.g. if the BS and the MS are located in one street canyon, the path loss is calculated as follows [Kür99]:

$$\Theta(D) = 42.6 + 26 \log\left(\frac{D}{\text{km}}\right) + 20 \log\left(\frac{f}{\text{MHz}}\right) \qquad \text{for } D \geq 20m \tag{3.9}$$

In the NLOS case the path loss is calculated as

$$\Theta(D) = \begin{cases} \Theta_0 + \Theta_{rts} + \Theta_{msd} & \text{for } \Theta_{rts} + \Theta_{msd} > 0 \\ \Theta_0 & \text{otherwise} \end{cases} \tag{3.10}$$

where Θ_0 corresponds to the free space loss at distance D. The term Θ_{msd} reflects the multiple knife-edge diffraction between BS antenna and final building, Θ_{rts} accounts for the single knife-edge diffraction and scattering from the final building down to the user's location. Both terms require multiple input parameters, such as width of street canyons, street orientation, number of floors of buildings. A complete description is given for example in [Kür99] and [Sau99].

Parameter	Cost 231 - Hata	Cost 231 - Walfish Ikegami
f	1500 MHz ... 2000 MHz	800 MHz ... 2000 MHz
h_{BS}	30 m ... 200 m	4 m ... 40 m
h_{MS}	1 m ... 10 m	1 m ... 3 m
D	1 km ... 20 km	0.02 km ... 5 km

Table 3.1: Applicability of empirical path loss models

3.1.3 Improved Models

Empirical models as the ones described above serve as a basis for most propagation models in planning tools. However, planning tools often make use of additionally available topographic and land use data to further enhance the accuracy of predictions. A commonly used approach is to add a position specific diffraction loss. This is typically calculated by a multi-knife diffraction model [Sau99], which considers the height profile between BS antenna and the mobile. The height profile is generated based on the topology height data to which clutter specific height values are added.

The path loss obviously depends on the propagation environment. In order to take this into account in a more detailed way than considered e.g. by the Cost 231-Hata model, in planning tools often clutter correction factors are applied on top of the basic path loss, e.g. [Air]. As they typically consider only the clutter type the mobile is located in, clutter correction factors only take account of the propagation environment in the vicinity of the mobile. The correction factors are for example obtained from model tuning algorithms which make use of measurements to adapt the correction factors used for a certain BS or region.

In most situations, propagation models and according parameters can be selected such that they yield good results for certain environments, e.g. urban or sub-urban environments. However, a model which is well tuned for arbitrary environments can hardly be found. Thus, in several recent approaches the model to be applied for a certain path loss prediction is decided on depending on the type of environment. In the model proposed in [KFW96] the path loss model and the corresponding parameters are derived from the clutter profile between BS and user location.

If detailed information about the scenario, in particular vector building data, are available, also ray-tracing tools can be considered to be used for estimating the path loss. Ray-tracing tools simulate the propagation of radio waves including reflections, diffraction, etc. Several paths between BS antenna and mobile are considered and combined. Well calibrated ray tracing tools have proven to provide very accurate results.

3.2 Shadow Fading

In general, models that describe the macro path loss empirical or deterministic models that give the path loss between base station antenna and user equipment essentially as a function of distance and height of antennas. In such models the path loss is the same for users located at different positions but at the same height and with equal distance to one antenna. Although in practice the clutter along the propagation paths make each path different, propagation models only very roughly take this into account, e.g. in terms of diffraction losses. The variation with respect to the nominal value suggested by the prediction model is called *shadowing* or *slow fading* and is usually modeled statistically as a log-normal distributed random variable. The respective normal distributed random variable in the logarithmic domain has a mean of zero and a standard deviation of σ [Sau99]. The justification for the shadowing to be log-normal distributed is that the path loss along the propagation path(s) in decibel is the sum of a large number of independent contributions. However, it has to be noted that the independence of the individual components is strictly speaking only approximately true.

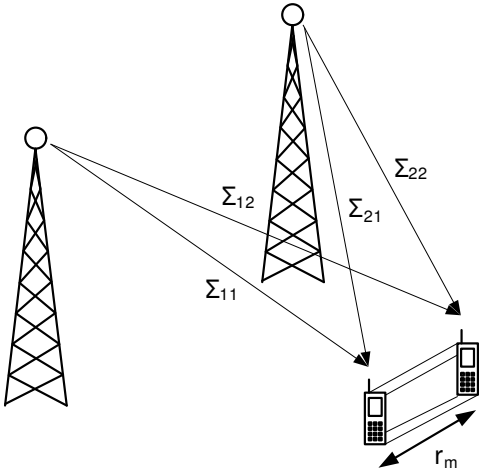


Figure 3.3: Illustration of the notation used for describing shadowing

Depending on the environment the correlation of shadowing between nearby propagation paths is to be considered. Two types of correlation can be distinguished, namely *spatial correlation* and *link correlation*. Spatial correlation is the correlation of propagation paths between nearby user locations and one base station antenna, whereas link correlation is considered as the correlation between links from one user position to different base station antennas. The term “link” here not necessarily corresponds to an actual communication link, but generally speaking describes the transmission of energy from one point to the

other. In that sense a link might for example also represent the transmission of interference from one antenna to the other.

3.2.1 Modeling of Spatial Correlation

Spatial correlation describes the rate at which the shadowing changes over time as the user moves. The spatial correlation between two positions is described as a function of the distance r_m between the two positions (cf. Figure 3.3):

$$\rho_S(r_m) = \frac{\mathrm{E}\left\{\Sigma_{11}\Sigma_{12}\right\}}{\sigma^2} \tag{3.11}$$

where Σ_{ij} denotes the random variable that describes the shadow fading between base station i and user j in logarithmic scale. In (3.11) it is assumed that the standard deviation of the shadowing σ is the same at both locations. This is in particular reasonable in the case where r_m is small.

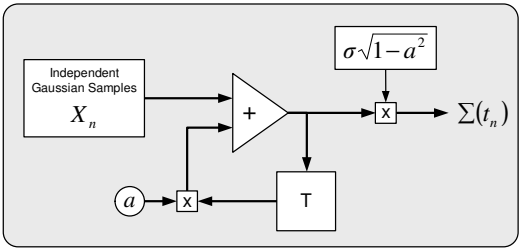


Figure 3.4: Generation of spatially correlated shadowing values

In [Sau99] it is suggested to model the shadowing correlation in simulations by an *Infinite Impulse Response* (IIR) filter. The shadowing $\Sigma(t_n)$ on the path from an arbitrary base station to a user evaluated at discrete simulation times t_n is calculated as (cf. Figure 3.4)

$$\Sigma(t_n) = a\,\Sigma(t_{n-1}) + \sigma\sqrt{1-a^2}\,X_n \tag{3.12}$$

where X_n is an independent Gaussian sample with unit standard deviation and zero mean.

The coefficient a is defined as

$$a = \frac{1}{\sigma^2} e^{-\frac{r_m}{r_c}} \tag{3.13}$$

where r_m is the distance the user has moved within $\Delta t = t_n - t_{n-1}$ and r_c denotes the correlation distance, which is obtained empirically and depends on the (propagation) environment. Typical values of r_c vary between a few meters and several tens of meters (cf. Table 3.2).

3.2.2 Modeling of Link Correlation

In addition to the spatial correlation between transmission paths, the link cross-correlation is to be considered:

$$\rho_L = \frac{\mathrm{E}\left\{\Sigma_{11}\Sigma_{21}\right\}}{\sigma^2} \tag{3.14}$$

where Σ_{11} and Σ_{21} denote the respective random variables that describe the shadow fading between two base station antennas and one user location in logarithmic scale.

In simulations the correlated shadowing on the links from each base station that influence a particular user are to be calculated. In the following three different possibilities for modeling the correlated fading are presented.

Equal Cross-Correlation Model

In this model the cross-correlation ρ_L between any pair of links for one user is assumed to be the same. For each path, e.g. between base station antenna i and user j, the shadowing Σ_{ij} is calculated as

$$\Sigma_{ij} = \sigma\left(\rho_L X_j + \sqrt{1-\rho_L^2}\, Y_{ij}\right) \tag{3.15}$$

where X_j is a user specific and Y_{ij} is a link specific independent Gaussian random variable with zero mean and unit standard deviation. This model is widely used in literature.

Angular Model

Obviously the cross-correlation between disjunctive links is a function of the angle between the links. In [Sau99] the following scheme for estimating the cross-correlation between links is suggested:

$$\rho_L = \begin{cases} \sqrt{\frac{r_1}{r_2}} & \text{for } 0 \le \Phi < \Phi_T \\ \left(\frac{\Phi_T}{\Phi}\right)^{\gamma} \sqrt{\frac{r_1}{r_2}} & \text{for } \Phi_T \le \Phi < \pi \end{cases} \tag{3.16}$$

where r_1 is the greater of the distances r_1 and r_2 from the two respective base station antennas to the user's position. If the distance between the two base station antennas is denoted as r_{12}, the angle Φ between the two links can be calculated as:

$$\Phi = \arccos\left(\frac{r_1^2 + r_2^2 - r_{12}^2}{2 r_1 r_2}\right) \tag{3.17}$$

The threshold angle Φ_T is related to the maximum angle at which both paths fully overlap meaning that the two paths cross the same obstacles. It is suggested to calculate Φ_T as follows:

$$\Phi_T = 2 \arcsin\left(\frac{r_c}{2r_1}\right) \tag{3.18}$$

where r_c is the correlation distance already introduced in Section 3.2.1. Finally, the correction exponent γ is chosen based on empirical observations (cf. Table 3.2).

From (3.16) the variance-covariance matrix that expresses the correlation of all links to be considered for user i can be obtained. This matrix can then be used to generate random variables of the desired correlation. A possible algorithm for the generation is described in Appendix A.

Hybrid Model

Whereas the angular model presented above offers a great amount of detail, the simple modeling according to (3.15) requires a much lower computational effort. Especially in case of a large amount of base stations to be considered, the computational effort for calculating the shadowing becomes very high for the angular model. Hence, typically the equal cross-correlation model is used. For a more detailed modeling of shadowing in this section a hybrid model that combines the advantages of both methods is proposed.

For each user i the links of concern are sorted by their importance, e.g. based on path loss. The links are then divided into a set $\mathcal{M}_i^H$ of links with high importance and a set $\mathcal{M}_i^L$ of links with low importance. The partition can for example be carried out in such a way that the number of links in $\mathcal{M}_i^H$ is constant for all mobiles.

For the links in $\mathcal{M}_i^H$ the shadowing is generated based on the angular model. Subsequently the shadowing for the other links is calculated with a modified version of the equal cross-correlation model:

$$\Sigma_{ij} = \sigma\left(\rho_L \frac{\Sigma_{cj}}{\sigma} + \sqrt{1-\rho_L^2}\, Y_{ij}\right) \tag{3.19}$$

where Σ_{cj} is one of the shadowing values obtained in the first step, e.g. the shadow fading generated for the most important link.

3.2.3 Implementation in Simulations

The simulation approaches within this thesis consider snapshots of the system. In particular there is no explicit user mobility within the simulations. It is proposed to calculate the shadowing for each user in a snapshot individually based on the hybrid model described in Section 3.2.2. A spatial correlation between several users is not considered. The reason

for considering users independently is mainly motivated by the assumption that the statistical effect of cross-correlation between the shadowing for several users is low compared to the effect of cross-correlation between path loss values of one user to/from several base station antennas. Also the fact that propagation data and traffic data are generally given in pixel granularity, but pixel sizes in turn are often greater than the shadowing distance, gives rise to the assumption of considering users independently with regard to shadowing.

From an algorithmic point of view the generation of shadowing can be described as follows:

1. Prior to the actual simulation, the parameters required to generate the shadowing are calculated for each pixel i:
   (a) Firstly the set $\mathcal{M}_i^H$ of the most important links is generated. In order to limit complexity in both, computation and memory, it is proposed to use a fixed size for the set, which is a parameter of the simulation.
   (b) For each pixel i the variance-covariance matrix for the links in $\mathcal{M}_i^H$ is calculated.
   (c) If the algorithm described in Appendix A is used to generate the correlated random variables, the Cholesky-Factorization of the matrix can already be carried out for each pixel.
   (d) The parameters calculated for each pixel are stored appropriately.
2. During simulation for each user in each snapshot the shadowing is independently and randomly generated using the parameters already stored in (1) for the pixel the user is located in.

The set of parameters for the modeling of shadowing is an input parameter to the simulation. The values presented in Table 3.2 were obtained and used by the EU project MOMENTUM [WTL+03] and shall give an indication on typical values used. Although not further considered in this thesis, for the sake of completeness also default parameters for the spatial distribution are included in the table.

Parameter	Default Value	Comment
σ	6-10dB	Standard deviation of shadow fading
ρ_L	0.5	Link cross-correlation (if not explicitly calculated)
r_c	5-20m	Spatial correlation: Correlation distance
γ	0.3	Spatial correlation: Correction exponent

Table 3.2: Typical shadow fading parameters

3.3 Fast Fading

Fast fading is caused by multi-path fading effects. In general it has to be distinguished between *narrowband fast fading* and *wideband fast fading*. Narrowband fast fading is

applicable if the delay between the different paths is very low, i.e. the delay is small compared with the symbol length. All rays arrive at the receiver at essentially the same time but with a certain phase shift, which leads to constructive and destructive interference among waves. Narrowband fast fading is typically either modeled by a Rayleigh or a Rice distribution depending on whether there is a NLOS or LOS condition between transmitter and receiver [Rap02, Sau99, BA02].

In contrast, if the delay becomes larger, a distortion of consecutive symbols occurs. This effect is called *Inter-Symbol Interference* (ISI). As the effect is only of significance if the symbol length is small and hence symbols have a wider representation in the frequency domain, the corresponding fading is referred to as wideband fast fading. Narrowband and wideband fast fading cannot be fully separated. In a wideband mobile communication channel typically several “beams”, each comprising several rays with very low delay spread, reach the receiver on different paths with considerable delay. The “beams”, often also denoted as *taps*, represent the wideband fast fading. The aggregation of rays within one tap (caused by scattering effects on one path) can be described by narrowband fast fading.

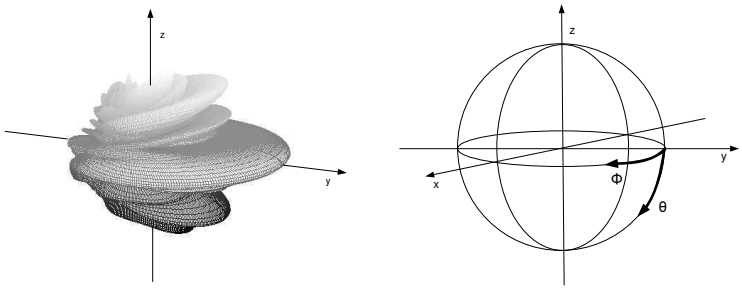
UMTS is a wideband mobile communication system. For its modeling wideband fast fading needs to be considered. The large scale delay spread is described by channel profiles defined for different user velocities and environment types, e.g. [IT97]. As suggested above, the fast fading caused by the superimposed rays within one tap is modeled as narrowband fast fading. The narrowband fast fading processes of the individual taps are considered as uncorrelated.

3.4 Antenna Modeling

As opposed to the antennas of users’ handsets, which are well modeled as isotropic radiators, base station antennas in mobile communication systems are generally directional, i.e. the antenna gain depends on the location relative to the antenna. A three dimensional representation of a typical antenna pattern is presented in Figure 3.5(a).

Important characteristics of antennas are their horizontal and vertical antenna patterns. The horizontal pattern represent the attenuation relative to the maximum antenna gain measured along the “equator” of the 3D antenna pattern, the vertical pattern corresponds to the “prime meridian” respectively.

The orientation of an installed antenna is described by its azimuth (horizontal orientation) and tilt (vertical orientation). In a cellular network tilting is commonly used to control the cell size and interference. However, mechanical tilting of the antenna basically changes the cell radius in direction of the main lobe of the antenna. The horizontal radiation pattern perpendicular to the front of the antenna keeps almost unchanged [Kat00]. This effect, sometimes referred to as “peanuting”, can be avoided by electrical tilting. That is, the



(a) 3D antenna representation (b) Angle definition

Figure 3.5: Three-dimensional antenna pattern

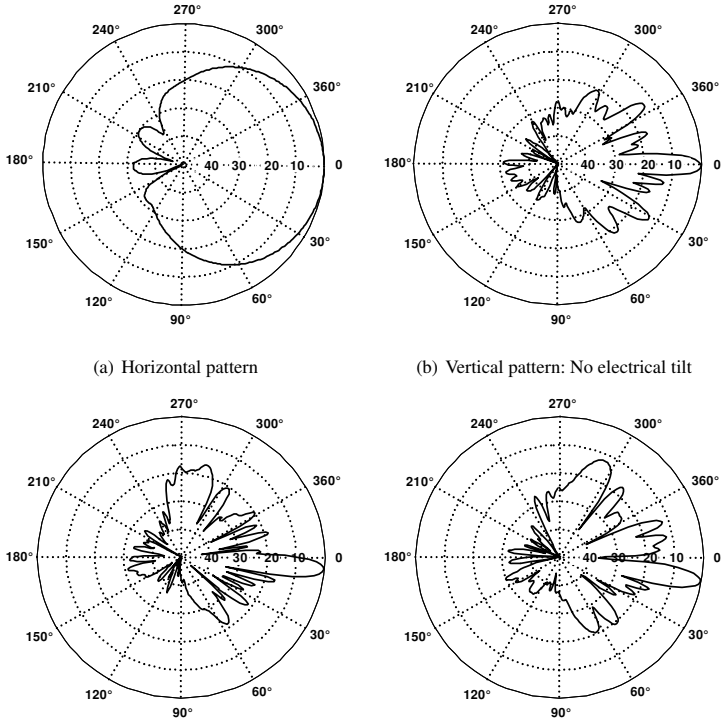
physical tilting of the antenna is emulated by modifying the (vertical) antenna pattern, such that the maximum gain is rotated down (or up) without changing the mounting angle of the antenna. Besides the avoidance of the "peanuting" effect, an important advantage of electrical tilting is that some antennas offer a variable electrical tilt, which can be either changed easily on site, or even remotely. In Figure 3.6 the antenna patterns for a typical antenna used in UMTS networks are shown. Besides the horizontal pattern, the vertical patterns for three different electrical tilt angles are presented. The respective values represent the attenuation in decibel with reference to the maximum antenna gain.

Both, azimuth and tilt angles, have a great influence on the performance of a cellular radio network. In network planning applications accurate three-dimensional antenna patterns are required to calculate proper power and interference levels. In the ideal case, the three-dimensional antenna pattern is provided by the antenna manufacturer. However, typically only the horizontal and the vertical pattern are available. In this case the three-dimensional pattern needs to be interpolated. Several methods for antenna interpolation have been developed, e.g. [Hop03, GCF+01, ZK05]. In [Böt06] and [Hop03] the four most popular interpolation methods are compared taking a measured three-dimensional pattern as reference.

The simplest interpolation scheme, in [Hop03] referred to as *Analysis of Adjacent Points* (AAP), calculates the total antenna gain $\Omega(\phi, \theta)$ as the sum of the values taken from the vertical pattern for the azimuthal angle ϕ and the value from the horizontal pattern for the polar angle θ:

$$\Omega(\phi, \theta) = \Omega_G - \Omega_H(\phi) - \Omega_V(\theta) \tag{3.20}$$

where $\Omega_H(\phi)$ and $\Omega_V(\theta)$ represent the values from the horizontal and vertical antenna patterns respectively. The antenna specific constant Ω_G denotes the global antenna gain, i.e. the gain in direction of the main lobe of the antenna. The definition of θ and ϕ in spherical coordinates is depicted in Figure 3.5(b).



(a) Horizontal pattern

(b) Vertical pattern: No electrical tilt

(c) Vertical pattern: Electrical tilt=5°

(d) Vertical pattern: Electrical tilt=10°

Figure 3.6: Horizontal and vertical antenna pattern of example antenna (Kathrein 742215: [Kat06])

AAP already provides very good results and according to the independent studies in [Böt06] and [Hop03] surprisingly outperforms much more complex algorithms like the *Weighted Bilinear Interpolation* (WBI) proposed in [GCF+01]. A further improved performance can be obtained from *Horizontal Projection Interpolation* (HPI) [Hop03]. The HPI algorithm takes the horizontal pattern as basis, but introduces a correction term that takes account of the vertical pattern:

$$\Omega(\phi,\theta) = \Omega_G - \Omega_H(\phi) + \left[\frac{\pi - |\phi|}{\pi} \cdot (\Omega_H(0) - \Omega_V(\theta)) + \frac{|\phi|}{\pi} \cdot (\Omega_H(\pi) - \Omega_V(\pi - \theta))\right] \quad (3.21)$$

The main advantage of the HPI algorithm is, that it takes values from the front-lobe ($0 \leq \theta < \pi$) and the back-lobe ($\pi \leq \theta < 2\pi$) of the antenna, whereas the AAP only takes values from the front-lobe (Note: Spherical coordinates: $0 \leq \theta < \pi$). This is also reflected

in the fact, that HPI is in particular more accurate than AAP for values of the back-lobe. As a consequence for the studies carried out in this thesis the HPI algorithm is used.

The antenna patterns considered so far are based on lab measurements carried out by the antenna manufacturers. Characteristic property of directional antennas is that they have deep dips in both, horizontal and vertical, patterns. The vertical pattern shows these dips even in direction of the front-lobe of the antenna. In a realistic outdoor propagation environment, however, this effect is reduced due to multi-path propagation, often referred to as "null-filling". Irrespective of ray-tracing models, multi-path effects are not explicitly considered by most prediction models. Most prediction models consider only the direct path to a certain position. It is thus proposed to modify the antenna pattern in such a way, that the antenna pattern itself incorporates multi-path effects in a statistical manner. As a result a smoothed antenna pattern is obtained. A simple approach for the smoothing of the vertical antenna pattern is proposed in [Böt06]. It has been derived empirically from comparisons between results obtained from ray-tracing and rather simple empirical models. In Figure 3.7 the modification of the vertical pattern is presented for the antenna already used earlier in this section.

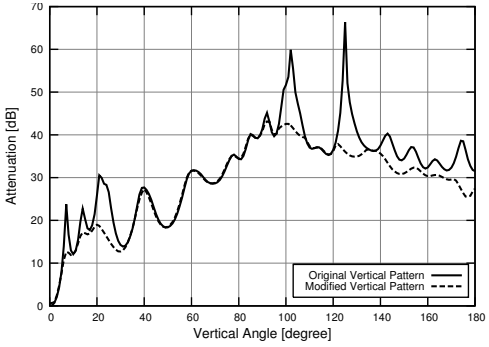


Figure 3.7: Modified vertical pattern due to multi-path effects

3.5 Signal and Interference Levels

The total radio transmission path comprises four components introduced earlier in this chapter: path loss, antenna gain, shadow fading, and fast fading. The total and in general time-varying link gain between a transmitter k and the receiver j in linear scale is described by the following equation:

$$\tilde{\xi}_{kj} = 10^{\frac{-\Theta_{kj} + \Sigma_{kj} + \Omega_{kj} + \Lambda_{kj}}{10}} \tag{3.22}$$

where Θ_{kj}, Σ_{kj}, and Ω_{kj} denote the path loss, the shadow fading, and the antenna gain respectively. The term Λ_{kj} denotes the fast fading contribution.

The radio transmission path underlies temporal variations on different levels of granularity. Firstly, there are long term variations caused by the mobility of users which change basically the contributions of all four components of the transmission path. On second level there are medium-term variations of the propagation environment described by shadow fading, and finally short-term variations in the order of kHz and above described by the fast fading.

The methods developed in this thesis (cf. Chapter 4, Chapter 5, and Chapter 6) are all essentially based on snapshots of the system, that is, the system state is for the analysis "frozen" at a certain point in time. Consequently, path loss values and antenna gains are given by the static position of the individual users. Shadow fading values are explicitly drawn as described in Section 3.2, and are considered by a fixed realization for each user within one snapshot. In contrast, to consider a random fast fading contribution for each link is not appropriate. The reason is, that a significant part of the network performance analysis is the modeling of the behavior of RRM algorithms. Except for the fast power control (cf. Section 2.3.2) RRM algorithms do not operate on such a short time frame, but rather consider mean values over a certain measurement period. It is consequently useful to average out fast fading effects. Details are given in the following.

At first a communication link including fast fading is considered. A signal is transmitted from a transmitter k to a receiver j. The time varying transmitted power $\tilde{t}_{kj}$ is received at the receiver with level $\tilde{r}_{kj}$:

$$\tilde{r}_{kj} = \tilde{\xi}_{kj} \cdot \tilde{t}_{kj} \tag{3.23}$$

where $\tilde{\xi}_{kj}$ is the link gain including the fast fading contribution.

The measure of greatest relevance for the quality of a communication link is the *Carrier to Interference plus Noise Ratio* (CIR) at the receiver, i.e. the CIR needs to be higher than a certain threshold. The CIR is defined as

$$CIR_{kj} = \frac{\tilde{r}_{kj}}{\tilde{i}_{kj}} \tag{3.24}$$

where $\tilde{i}_{kj}$ is the interference level experienced by the receiver of the communication link which is given by

$$\tilde{i}_{kj} = \sum_{k} \tilde{t}_k \tilde{\xi}_{kj} - \tilde{r}_{kj} + \eta_j \tag{3.25}$$

Here $\tilde{t}_k$ denotes the total power transmitted by transmitter k, η_j is the noise level at receiver j.

In a power controlled channel (cf. Section 2.3.2), the power is constantly adjusted, in order to achieve the required CIR target level CIR^T. Assuming perfect power control which adjusts the transmit power in an ideal way such that the CIR at the receiver side exactly matches the target level, it yields

$$CIR^T_{kj} = \text{const.} = CIR_{kj} = \frac{\tilde{r}_{kj}}{\tilde{i}_{kj}} \tag{3.26}$$

In this thesis static values, incorporating average fast fading contributions, are considered. For the static transmit power t_{kj}, the static receive level r_{kj}, and the static interference i_{kj} it yields

$$t_{kj} = \mathrm{E}\left\{\tilde{t}_{kj}\right\} \tag{3.27}$$
$$r_{kj} = \mathrm{E}\left\{\tilde{r}_{kj}\right\} \tag{3.28}$$
$$i_{kj} = \mathrm{E}\left\{\tilde{i}_{kj}\right\} \tag{3.29}$$

The static link gain is accordingly defined as

$$\xi_{kj} = \mathrm{E}\left\{\tilde{\xi}_{kj}\right\} \tag{3.30}$$
$$= \mathrm{E}\left\{10^{\frac{-\Theta_{kj}+\Sigma_{kj}+\Omega_{kj}+\Lambda_{kj}}{10}}\right\} \tag{3.31}$$
$$= 10^{\frac{-\Theta_{kj}+\Sigma_{kj}+\Omega_{kj}}{10}} \cdot \mathrm{E}\left\{\underline{\Lambda}_{kj}\right\} \tag{3.32}$$

where

$$\underline{\Lambda}_{kj} = 10^{\frac{\Lambda_{kj}}{10}} \tag{3.33}$$

If fast fading is defined by Rayleigh or Rice fading processes it yields

$$\mathrm{E}\left\{\underline{\Lambda}_{kj}\right\} = 1 \tag{3.34}$$

and hence

$$\xi_{kj} = 10^{\frac{-\Theta_{kj}+\Sigma_{kj}+\Omega_{kj}}{10}} \tag{3.35}$$

In radio network planning a static CIR target level $CIR_{kj}^{\text{static},T}$ is commonly used:

$$CIR_{kj}^{\text{static},T} = \frac{r_{kj}}{i_{kj}} = \frac{\mathrm{E}\left\{\tilde{r}_{kj}\right\}}{\mathrm{E}\left\{\tilde{i}_{kj}\right\}} \tag{3.36}$$

$CIR_{kj}^{\text{static},T}$ describes the average signal to average interference ratio which is required to assure a certain quality of the connection. Values are obtained from link level simulations and depend on multiple factors like the type of service, the propagation environment (channel model), and the imperfectness of power control. As the dynamics of interference effect the denominator of (3.36) there is in principle also a dependency on the constellation of interferers. This effect, however, is small and generally neglected. Nevertheless, please note that clearly $CIR_{kj}^{\text{static},T} \neq CIR_{kj}^{T}$. A fact which is often disregarded.

Signal to interference level requirements of WCDMA links are typically expressed in terms of $\frac{E_b}{N_0}$ targets. Hence, in the subsequent chapters the following definition is used:

$$\frac{E_b}{N_0} = PG \cdot CIR_{kj}^{\text{static},T} = PG \cdot \frac{\mathrm{E}\left\{\tilde{r}_{kj}\right\}}{\mathrm{E}\left\{\tilde{i}_{kj}\right\}} \tag{3.37}$$

The processing gain PG is defined as the ratio between chip-rate W and user bit rate B:

$$PG = \frac{W}{B} \tag{3.38}$$

3.5.1 Effective Interference Raise

The static modeling of fast fading not only effects the $\frac{E_b}{N_0}$ targets. It has also an impact on power and interference levels.

From (3.23) and (3.27):

$$t_{kj} = \mathrm{E}\left\{\tilde{t}_{kj}\right\} = \mathrm{E}\left\{\frac{\tilde{r}_{kj}}{\tilde{\xi}_{kj}}\right\} \tag{3.39}$$

Assuming perfect power control, received power $\tilde{r}_{kj}$ and link gain $\tilde{\xi}_{kj}$ are statistically independent, as the link gain variations are fully compensated by the transmit power. Hence

$$t_{kj} = \mathrm{E}\left\{\tilde{r}_{kj}\right\} \cdot \mathrm{E}\left\{\frac{1}{\tilde{\xi}_{kj}}\right\} = r_{kj} \cdot \mathrm{E}\left\{\frac{1}{\tilde{\xi}_{kj}}\right\} \tag{3.40}$$

Substituting (3.36) in (3.40) yields

$$t_{kj} = CIR_{kj}^{\text{static}} \cdot i_{kj} \cdot \mathrm{E}\left\{\frac{1}{\tilde{\xi}_{kj}}\right\} \tag{3.41}$$

Finally, substituting (3.22) and (3.35) results in

$$t_{kj} = CIR_{kj}^{\text{static}} \cdot i_{kj} \cdot \mathrm{E}\left\{10^{-\frac{-\Theta_{kj}+\Sigma_{kj}+\Omega_{kj}+\Lambda_{kj}}{10}}\right\} \tag{3.42}$$

$$= CIR_{kj}^{\text{static}} \cdot i_{kj} \cdot \frac{1}{\xi_{kj}} \cdot \mathrm{E}\left\{\frac{1}{\underline{\Lambda}_{kj}}\right\} \tag{3.43}$$

The term

$$\kappa_{kj} = \mathrm{E}\left\{\frac{1}{\underline{\Lambda}_{kj}}\right\} \tag{3.44}$$

is referred to as effective interference raise [SLSWJ99]. It describes the raise of interference caused by a power controlled signal that takes fast fading into account compared to the case where fast fading is not considered. The factor needs to be applied to the power evaluated for all power controlled channels. On the receiver side it needs to be added to the interference levels of all power controlled channels except the ones controlled by the respective receiver itself.

The effective interference raise depends on a large number of factors, in particular the user's velocity and the channel profile. The effective interference raise in this section was derived assuming perfect power control. In [SLSWJ99] values are given based on analytical investigations and simulations with imperfect power control.

3.5.2 Power Control Headroom

The maximum transmit power available for a connection is limited in both UL and DL direction. In UL the limitation is due to the maximum output power of the amplifier in the UE. In DL it is limited by the RRM. Fast fading causes the channel power to vary quite rapidly around the average transmit power. If the required output power is not available, the CIR target and consequently the quality target of the connection cannot be met. Thus, a certain margin between average and maximum transmit power is required in order to guarantee proper operation of the connection. This margin is commonly referred to as *power control headroom*. Similarly to the effective interference rise, also the fast fading margin depends on the mobile's velocity and the channel impulse response. Values obtained from simulations can be found in literature, e.g. [SLSWJ99, SJLSW99].

3.5.3 Soft-Handover Gain

The impact of SHO (Section 2.3.1) is manifold. Firstly, it allows for a reduction in transmitted output power, which is taken account of in simulations by a reduced $\frac{E_b}{N_0}$ target, widely denoted as *soft handover gain*. In addition SHO also has an impact on the effective interference raise and the power control headroom described above. Both quantities are reduced in case of SHO. The reduction is due to the statistical gain in combining a larger amount of independent fading paths. This is commonly referred to as *diversity gain*.

Chapter 4

Monte-Carlo Snapshot Analysis

The main objective of performance analysis for network planning is to estimate the grade of performance of a given network configuration based on assumptions regarding traffic distributions and propagation conditions. Network performance is assessed by various measures such as coverage probability, blocking rate, and packet delay.

For system performance evaluation, dynamic simulation techniques are widely used. Dynamic simulations are typically event driven and simulate the system over some period of time. Generally the level of detail as well as the correlation of system states in a simulation run is very high [LPTG03]. Due to the strong temporal correlations the gathered results are rather limited compared to the amount of calculations required for state handling, etc. A run giving significant results even for a small scenario can take several days or even weeks of computation. Dynamic simulations are very well suited especially for the analysis and development of system algorithms, where in general a small network can be studied in detail without being constrained by the simulation time. In network planning very large networks have to be analyzed in reasonable time. However, also the required level of detail is lower. As a consequence, the Monte-Carlo snapshot analysis has established itself as a common approach for network simulations in radio network planning [Air, For, Mar, WLSSJ99].

In this chapter the fundamentals of the Monte-Carlo method are introduced. The method itself is explained and its basic characteristics are illustrated with a simple planning example. Methods for variance reduction are discussed and applied to the simple example. The aim of this chapter is to provide a basis for the understanding of Chapter 5, in which a Monte-Carlo snapshot analysis method for the performance evaluation of WCDMA radio networks is developed.

4.1 The Monte-Carlo Method

Numerous publications on the Monte-Carlo method and its application in various areas of research can be found in literature. For an overview on the topic see e.g. [Rub81, Rub86, Dub99, Kal86].

4.1.1 General Concept

Many real world problems require the analysis of systems that are too complex to be analyzed analytically. This is mainly due to the large state space of the corresponding system models and the typically large number of random variables involved, e.g. for the modeling of input quantities. Such systems are often analyzed by means of stochastic simulations. According to [Rub81] *simulations experiment with the model over some time and as a rule observations are serially correlated*. From another perspective the dynamic simulation can be seen as a random walk through the state space. Due to the serial correlation, simulations are generally for some time bound to a certain region of the state space, which is explored in more level of detail than often required. In order to obtain reliable results, however, all regions of the state space have to be visited. As a result dynamic simulations while obtaining very detailed results usually require long simulation runs.

The Monte-Carlo method aims to overcome this drawback by taking uncorrelated probes, in the following referred to as *Monte-Carlo trials* or *snapshots*, from the state space. An exploration of the state space thus requires a much shorter time-frame. The term Monte-Carlo is not consistently used in literature. Sometimes Monte-Carlo is used for any simulation that involves randomness, i.e. stochastic simulations. In this thesis the Monte-Carlo method describes a technique that explicitly does not consider the simulation of a system over time, but in contrast analyzes uncorrelated random trials. For this approach also the terms *snapshot simulation*, *snapshot analysis*, and *static simulation* are widely used as synonym.

The applicability of the Monte-Carlo method strongly depends on the system to be analyzed and the applied system model. The following three prerequisites need to be fulfilled for an efficient application of the Monte-Carlo method:

1. The system to be analyzed is modeled stochastically.
2. The quantities of interest do not require a dynamic analysis of the system. If such quantities, e.g. transmission delays or cycle times, are to be analyzed, the *Short-Term Dynamic Simulation* (STD) approach [LPTG03] can be used. With this, starting from snapshots, a short period of dynamic simulation is carried out.
3. The probes, or snapshots, of the state space can be efficiently generated. In particular if a long transition phase is required to achieve a valid system state, the Monte-Carlo method becomes inefficient.

For an explanation of the basic principles of the Monte-Carlo method a single output quantity that is a function of a (multi-dimensional) random variable X is considered:

$$G = f(X) \tag{4.1}$$

The output quantity G is in general also a random variable. For the application of the Monte-Carlo method it is important to note, that the system is time-invariant and memoryless. That is, the current variate G_i only depends on the current realization X_i of the input variable:

$$G_i = f(X_i) \tag{4.2}$$

Expected value and variance of G are estimated by the sample mean m_G and the sample variance s_G^2 after N Monte-Carlo trials (e.g. [Kal86]):

$$m_G(N) = \frac{1}{N}\sum_{i=1}^{N} G_i \tag{4.3}$$

$$s_G^2(N) = \frac{1}{N-1}\left(\sum_{i=1}^{N} G_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} G_i\right)^2\right) \tag{4.4}$$

For $N \to \infty$ it yields

$$\lim_{N\to\infty} m_G(N) = \mathrm{E}\{G\} = \mu_G \tag{4.5}$$

$$\lim_{N\to\infty} s_G^2(N) = \mathrm{var}\{G\} = \sigma_G^2 \tag{4.6}$$

Similarly the higher moments of G can be estimated (e.g. [WM02]).

An estimation of μ_G according to (4.3) shall in the following be referred to as *Monte-Carlo sample-mean*.

4.1.2 Convergence Estimation

For practical applications of the Monte-Carlo method it is important to have a measure for the convergence of results, so that the number of Monte-Carlo trials required for a certain accuracy of results can be determined. Convergence can be appropriately kept track of by confidence intervals.

As the realizations G_i of G are statistically independent random variables, according to the central limit theorem [WM02] for $N \to \infty$, $m_G(N)$ is normally distributed with standard deviation $\sigma_G/\sqrt{N}$:

$$\lim_{N\to\infty} \mathrm{P}\left\{a \le \frac{\sqrt{N}\,(m_G(N) - \mu_G)}{\sigma_G} \le b\right\} = \int_a^b \frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \tag{4.7}$$

Assuming N to be sufficiently large, this formula can be used to estimate the probability of the estimate $m_G(N)$ to fall within a certain confidence interval.

Confidence intervals are in this thesis used to assess the number of required Monte-Carlo trials for that the relative error of estimation is below a value ϵ with a certain probability $\delta_{\text{confidence}}$:

$$\mathrm{P}\left\{\left|\frac{m_G(N)-\mu_G}{\mu_G}\right| \leq \epsilon\right\} = \delta_{\text{confidence}} \tag{4.8}$$

Assuming a symmetric confidence interval with $a = -b$ ($b > 0$), (4.7) can be rewritten as

$$\lim_{N\to\infty} \mathrm{P}\left\{-b \leq \frac{\sqrt{N}\,(m_G(N)-\mu_G)}{\sigma_G} \leq b\right\} = \int_{-b}^{b} \frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \tag{4.9}$$

$$\Leftrightarrow \quad \lim_{N\to\infty} \mathrm{P}\left\{\left|\frac{\sqrt{N}\,(m_G(N)-\mu_G)}{\sigma_G}\right| \leq b\right\} = 1 - 2\int_{b}^{\infty} \frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \tag{4.10}$$

$$\Leftrightarrow \quad \lim_{N\to\infty} \mathrm{P}\left\{\left|\frac{m_G(N)-\mu_G}{\mu_G}\right| \leq \frac{b\,\sigma_G}{\sqrt{N}\,|\mu_G|}\right\} = 1 - 2\int_{b}^{\infty} \frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \tag{4.11}$$

Hence, for large N, the quantities ϵ and $\delta_{\text{confidence}}$ in (4.8) can be approximated as follows:

$$\epsilon = \frac{b\,\sigma_G}{\sqrt{N}\,|\mu_G|} \tag{4.12}$$

$$\delta_{\text{confidence}} = 1 - 2\int_{b}^{\infty} \frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \tag{4.13}$$

Two important observations can be made from (4.12). Firstly, the rate of convergence is proportional to $1/\sqrt{N}$. That is, if the error range is to be halved, four times as many Monte-Carlo trials need to be analyzed. Secondly, the number of required Monte-Carlo trials to assure a certain accuracy is proportional to the variance σ_G^2.

As an example for the application of confidence intervals the number of Monte-Carlo trials required to achieve a maximum relative error ϵ of 1% with a probability $\delta_{\text{confidence}}$ of 0.9974 is to be determined. First the value b corresponding to the desired confidence probability is to be looked up from a Normal distribution table ($b = 3$). This value is then used in (4.12) to obtain the number of required trials N:

$$0.01 = \frac{3\,\sigma_G}{\sqrt{N}\,|\mu_G|} \tag{4.14}$$

$$\Rightarrow \quad N = \left(\frac{300\,\sigma_G}{\mu_G}\right)^2 \tag{4.15}$$

In general μ_G and σ_G are not known a priori. Instead the respective estimates (4.3) and (4.4) for mean and variance are used in equation (4.12). This approximation again holds for large N only.

4.2 Simple Analysis Example

The Monte-Carlo method is in the following illustrated by means of a simplified planning scenario. The scenario comprises one cell of circular shape with radius R. Users are randomly distributed in the cell (cmp. Figure 4.1). The cell is served by a Node B with an omnidirectional antenna located in the center of the cell. The total *Transmit* (TX) power of the Node B is shared among all users. As typical quantities of interest in network planning, the mean total TX power of the cell and the probability that the power required to serve all users exceeds the maximum TX power ("blocking probability") shall be obtained. For the sake for simplicity the UL is not considered in this example.

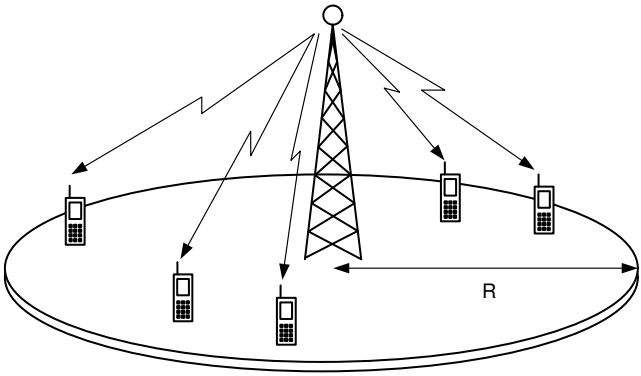


Figure 4.1: Simple one cell scenario

Call attempts in the cell are assumed to follow a Poisson process with arrival rate λ and service rate μ. The location of a call attempt is randomly chosen with uniform distribution over the cell area. Users are served by dedicated, power controlled channels, i.e. the power allocated to one user is adjusted so that the CIR target $\gamma^{\downarrow}$ is met. The total transmit power required to serve user i is calculated as follows:

$$t_i^{\downarrow} = \gamma^{\downarrow} \frac{i_i^{\downarrow}}{\xi_i^{\downarrow}} \tag{4.16}$$

where $\xi_i^{\downarrow}$ is the link gain between the base station and user i, and $i_i^{\downarrow}$ is the total interference experienced by user i. It has to be noted, that the total interference including the wanted signal is considered in the calculation although the wanted signal is in reality not visible in the receiver. The reason for this is a slight redefinition of the $\gamma^{\downarrow}$, which is introduced for an improved readability of formulas. The exact derivation of $\gamma^{\downarrow}$ is given in Section 5.5.

The median path loss between base station and user i located in distance D_i is described by the Cost 231-Hata model for urban areas (cf. Section 3.1.2). The total link gain consequently is

$$\begin{aligned}\xi^{\downarrow}_{i,[dB]} = &-46.3 - 33.9 \cdot \log\left(\frac{f}{\text{MHz}}\right) \\ &+ 13.82 \cdot \log\left(h_{BS}\right) \\ &+ \left((1.1 \cdot \log\left(\frac{f}{\text{MHz}}\right) - 0.7) \cdot h_{MS} - (1.56 \cdot \log\left(\frac{f}{\text{MHz}}\right) - 0.8)\right) \\ &- \left(44.9 - 6.55 \cdot \log\left(h_{BS}\right)\right) \cdot \log\left(D_i/1000\right) \\ &+ \Omega_i + \Sigma_i\end{aligned} \tag{4.17}$$

where Σ_i is the shadow fading contribution to the total link gain and Ω_i is the applicable antenna gain, both in decibels. Fast fading effects are not considered here. In accordance with Section 3.2 the shadow fading is modeled as a log-normal distributed random variable with zero mean and standard deviation σ in logarithmic scale. As the scenario consists only of one cell, a link correlation ρ_L does not apply. An isotropic transmit antenna is assumed, hence $\Omega_i = 0$.

A user i is assumed to be served if the received pilot power exceeds a minimum threshold $\hat{r}_{i,\,min}$:

$$\hat{\tau}\xi^{\downarrow}_i \geq \hat{r}_{i,\,min} \tag{4.18}$$

where $\hat{\tau}$ is the transmitted pilot power of the cell.

The total interference experienced by user i is the sum of the interference from the own cell (due to imperfect orthogonality) and receiver noise:

$$i^{\downarrow}_i = \alpha p^{\downarrow}\xi^{\downarrow}_i + \eta^{\downarrow} \tag{4.19}$$

All parameters and parameter settings are presented in Table 4.1.

The total cell TX power $p^{\downarrow}$ is the sum of the constantly transmitted power for the common channels τ and the individual transmit power levels $t^{\downarrow}_i$. If the total number of active users with a received pilot signal above the minimum threshold is denoted as M, the total TX power becomes:

$$p^{\downarrow} = \sum_{i=1}^{M}\left(\frac{\alpha\gamma^{\downarrow}p^{\downarrow} + \eta^{\downarrow}}{\xi^{\downarrow}_i}\right) \tag{4.20}$$

$$= M\alpha\gamma^{\downarrow}p^{\downarrow} + \sum_{i=1}^{M}\frac{\eta^{\downarrow}}{\xi^{\downarrow}_i} \tag{4.21}$$

$$= \frac{\eta^{\downarrow}}{1 - M\alpha\gamma^{\downarrow}}\sum_{i=1}^{M}\frac{1}{\xi^{\downarrow}_i} \tag{4.22}$$

Parameter	Symbol	Value
Frequency	f	2000 MHz
Height of BS antenna	h_{BS}	30 m
Height of MS antenna	h_{BS}	1.5 m
Cell radius	R	500 m
Distance between BS and MS	d	0 ... R
Transmitted pilot power	$\hat{\tau}$	2 W
Common channel power	τ	4 W
Maximum Node B TX power	$p^{\downarrow}_{max}$	20 W
Total Node B TX power	$p^{\downarrow}$	τ ... $p^{\downarrow}_{max}$
Orthogonality factor	α	0.4
Standard deviation of shadow fading	σ	6 dB
Link correlation of shadow fading	ρ_L	n.a.
Total MS receiver noise	$\eta^{\downarrow}$	-104.5 dBm
Minimum required pilot power receive level	$\hat{r}_{min}$	-115 dBm

Table 4.1: Parameters and parameter settings: Simple Example

The total cell TX power is bounded by the maximum output power $p^{\downarrow}_{max}$. That is, if the maximum output power is insufficient to serve all users with the amount of power required according to (4.16), the system is denoted to be in *blocking state*. The probability that the system is in blocking state is called *blocking probability*. If the system is in blocking state, appropriate measures have to be taken to reduce the required power, e.g. to drop users or to reduce data rates. In the following it is assumed that data rates are reduced. It is in particular important to note that in this example users are not explicitly blocked or dropped as this would have an impact on the traffic model (cf. Section 5.9.1).

From (4.22) two conditions that cause a blocking state can be deduced:

Condition 1: $p^{\downarrow} > p^{\downarrow}_{max}$. Appropriately serving the traffic would require a transmission power level that is not supported by the Node B.

Condition 2: $p^{\downarrow} < 0$, i.e. $M\alpha\gamma^{\downarrow} >= 1$. The amount of traffic is too large to be served by a finite output power. The system is above its *pole capacity*.

In this simplified scenario the total TX power and the blocking probability only depend on the amount of users, their distance from the base station antenna, and the shadow fading component of each user. If the arrival process is a Poisson process, the amount of active users in the system is Poisson distributed with mean $m = \lambda/\mu$. Hence, possible traffic realizations can be simply obtained by sampling appropriate random variables. Consequently, the Monte-Carlo method can be efficiently applied to analyze the system.

The system performance versus the amount of offered traffic has been analyzed for different settings of the CIR target $\gamma^{\downarrow}$, representing services with different bandwidth demands. As a trivial observation, total TX power (in Figure 4.2) and blocking probabil-

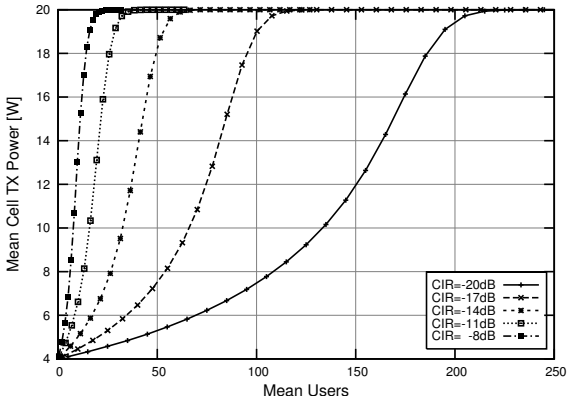


Figure 4.2: Mean cell transmission power vs. offered traffic

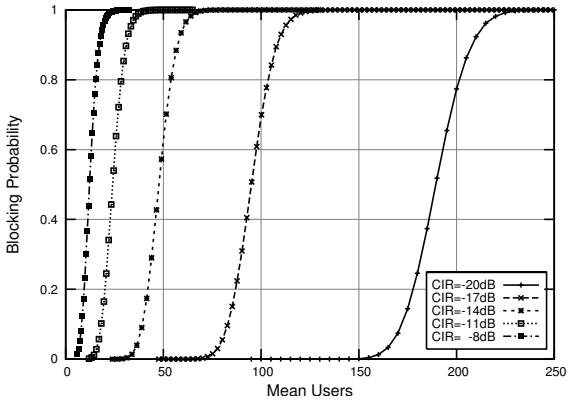


Figure 4.3: Blocking probability vs. offered traffic

ity (in Figure 4.3) increase with a higher bandwidth demand of users. More interestingly than the results themselves is the amount of required snapshots to obtain these results with sufficient statistical reliability. This is investigated in the following.

In Figure 4.4 the rate of convergence for one Monte-Carlo analysis run of 1,000,000 snapshots is depicted (CIR = -14 dB, mean users m = 40). In the figure, the relative error in the estimation of the mean TX power after each of the snapshots ("error of simulation trace") is compared with the maximum error given by (4.12) for the respective number of snapshots N ("99.74% error range"). In addition the error ranges evaluated based on a variance estimation are included ("99.74% error range (estimated)"). The latter was obtained substituting μ_G by $m_G(N)$ and σ_G by $s_G(N)$ in (4.12). The constant error and as a consequence the estimated error range of zero at the beginning of plot is due to the fact

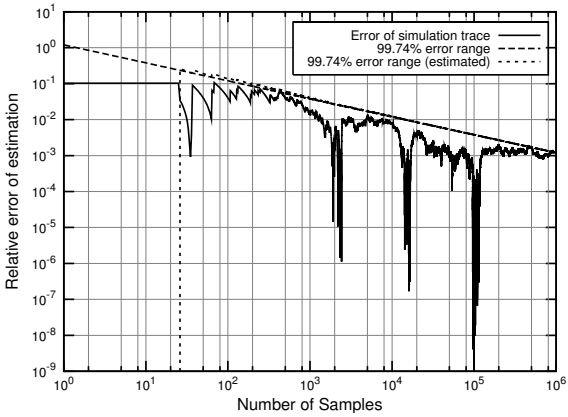


Figure 4.4: Mean power: Estimated and observed rate of convergence

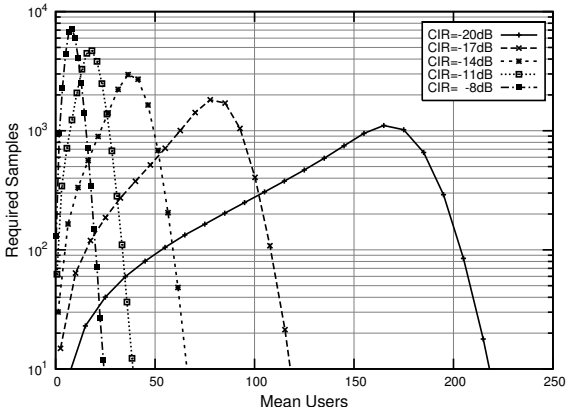


Figure 4.5: Required snapshots vs. offered traffic - estimation of mean total cell power

that for the first 25 snapshots the required TX power exceeded its maximum level. The power was thus clipped to the maximum value for all 25 snapshots, yielding a constant TX power estimate with a variance of zero. After an analysis of about 50 snapshots there is hardly any visible difference between estimated and exact error range. The plot verifies that confidence intervals can be effectively applied for an estimation of the accuracy of results. Moreover, the plot illustrates the fact that the rate of convergence is inversely proportional to $\sqrt{N}$.

In Figure 4.5 the required number of snapshots for an estimation of the mean cell TX power with a maximum relative error of 1% (99.74% confidence interval) is presented.

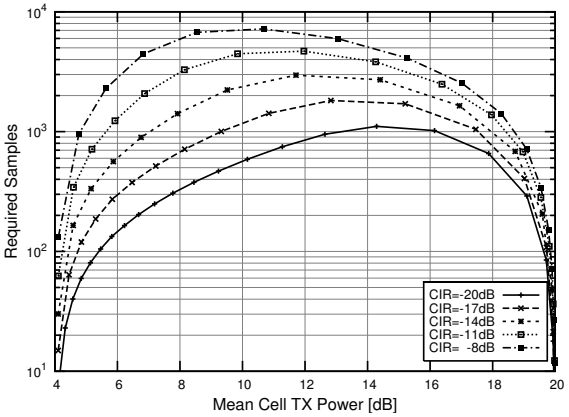


Figure 4.6: Required snapshots as a function of the estimated mean total cell power

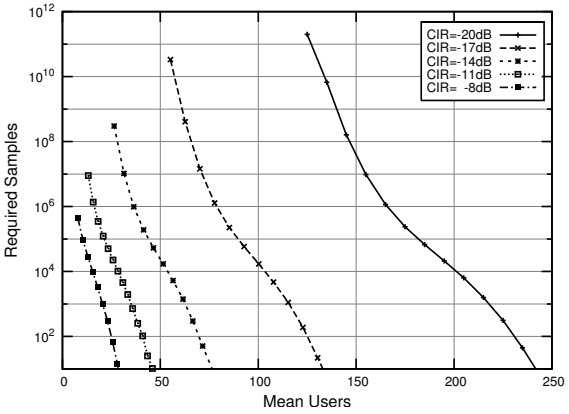


Figure 4.7: Required snapshots vs. offered traffic - estimation of blocking probability

The figures were determined following the procedure described in Section 4.1.2. For low offered traffic the amount of required snapshots is low because the output power is dominated by the common channel power. Consequently, the normalized standard deviation is small. In turn, for a large offered traffic the amount of required snapshots is also low because the total cell TX power is bounded by its maximum power and the variance of the cell power becomes smaller the higher the traffic. The number of required snapshots versus the mean output power is depicted in Figure 4.6. For realistic traffic scenarios (mean cell power roughly between 8 W and 16 W) a factor of 5 to 20 between required snapshots for the low CIR case (CIR =-20 dB) compared with the high CIR case (CIR =-8 dB) can be observed.

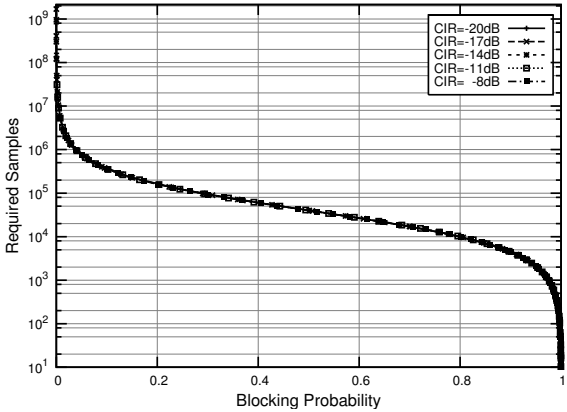


Figure 4.8: Required snapshots as a function of the estimated blocking probability

The number of required snapshots for the estimation of the blocking probability is for the different traffic cases presented in Figure 4.7. As well proven by the theory of *rare events*, for low traffic the number of required snapshots becomes extremely large. One might expect similar results also for blocking rates close to 1, because the non-blocking state becomes a rare event. However, it shall be recalled that in this analysis relative errors are considered. Hence, although the expected absolute error of estimation might for the high traffic case become similarly large as for the low traffic case, the expected relative error is significantly smaller and consequently less snapshots are required.

An additional interesting result can be obtained from Figure 4.8. The figure shows, that the amount of snapshots required for estimating the blocking probability only depends on the blocking probability itself, whereas it is (at least almost) independent of the traffic type. This is opposed by the power estimation case, where a strong dependency between the number of required snapshots and the traffic type can be observed (Figure 4.6).

4.3 Variance Reduction Techniques

The performance of the Monte-Carlo method strongly depends on the number of trials required to yield a certain level of convergence of results. In Section 4.1.2 it was shown that the convergence is proportional to the variance of the Monte-Carlo estimator. *Variance reduction* aims at reducing the variance of the estimated quantities, consequently reducing the number of required samples (or snapshots) and lowering the computational time required for an analysis. All methods for variance reduction have in common that they make use of known information about the model. This information is exploited in order to improve the estimation of the output quantities, e.g. by focusing on regions in the state space of the system which are known to be of more relevance. Vice versa if nothing

is known about the model, there will be no gain from variance reduction techniques, they might even cause the opposite effect.

In the following variance reduction is quantified by means of the *variance reduction coefficient*, which is defined as follows:

$$\psi = \frac{\text{var}\{m_G(N)\}}{\text{var}\{m_G^*(N)\}} \tag{4.23}$$

where $m_G(N)$ is the Monte-Carlo sample-mean estimator for the quantity G and $m_G^*(N)$ is the modified estimator imposed by the variance reduction method.

For a quantification of the reduction in number of samples, it is assumed that, similarly to $m_{G(N)}$, also $m_G^*(N)$ is unbiased and that its variance is inversely proportional to N (cf. Section 4.1.2). These assumptions hold true for all variance reduction methods presented in this thesis. The reduced number of required snapshots can be expressed as

$$N^* = \left\lceil \frac{1}{\psi} \cdot N \right\rceil \tag{4.24}$$

where N is the number of required Monte-Carlo trials to achieve a certain accuracy of results if a Monte-Carlo sample-mean estimator were applied. Accordingly N^* is the number of Monte-Carlo trials required for the same level of accuracy when the variance reduction scheme is used.

Several variance reduction techniques have been established in the past, in [Rub81] as many as eleven techniques are described. In the following four of them, which seem to be most promising in the scope of this thesis, are shortly introduced. Two of them will later be tailored and applied in order to improve the WCDMA radio network analysis (cf. Section 5.8).

4.3.1 Importance Sampling

The basic idea of *Importance Sampling* (IS) is to focus on values of the random input variables which are known to be of greater relevance for the quantities of interest. IS has been widely used for the estimation of error probabilities in communication channels, e.g. [Bal76, SB80, Jer84, WDT84]. Such error probabilities are typically very low, so that a Monte-Carlo sample-mean estimation with an acceptable relative error would in accordance with (4.12) result in a very large number of required Monte-Carlo trials. On the other hand, it is known that errors are more likely to occur for certain values of the input variables, e.g. for the tail of the receiver noise distribution. Hence, with IS these values of the distribution are sampled with increased probability. Obviously, as the sampling procedure is modified, an estimation of the mean of the sample results is not applicable anymore to obtain correct estimations of the output quantities. Instead a modified estimator is used. In the following, the mathematical background of IS is presented.

In contrast to the sample-mean Monte-Carlo estimator which directly takes the samples X_i in (4.2) from X, with IS a modified random input variable Y is considered:

$$\mathrm{E}\{G\} = \mathrm{E}\{f(X)\} = \int f(x)h_X(x)dx \tag{4.25}$$

$$= \int \frac{f(x)h_X(x)}{h_Y(x)}h_Y(x)dx \tag{4.26}$$

$$= \mathrm{E}\left\{\frac{f(Y)h_X(Y)}{h_Y(Y)}\right\} \tag{4.27}$$

where $h_X(x)$ and $h_Y(x)$ are the probability density functions of X and Y respectively.

Accordingly for the *importance sampling estimator* m_G^* follows:

$$m_G^*(N) = \frac{1}{N}\sum_{i=1}^{N} f(Y_i)\frac{h_X(Y_i)}{h_Y(Y_i)} \tag{4.28}$$

An important property of m_G^* is, that it is an unbiased estimator of μ_G and confidence intervals can be calculated according to Section 4.1.2.

It can be shown [Rub81] that the maximum variance reduction is achieved if $h_Y(x)$ is chosen as follows:

$$h_Y(x) = \frac{|f(x)|\,h_X(x)}{\int |f(x)|\,h_X(x)dx} \tag{4.29}$$

In particular, if $h_Y(x)$ is chosen according to (4.29) and if $f(x) > 0\,\forall\, x$, then var $\{m_G^*(N)\} = 0$.

An application of (4.29) for selecting $h_Y(x)$ is obviously useless as it practically requires a calculation of the quantity of interest, μ_G, which in turn would make the importance sampling worthless. Nevertheless (4.29) gives a suggestion on the desired properties of $h_Y(x)$. In order to minimize the variance, $h_Y(x)$ should be chosen such that its shape is similar to the shape of $|f(x)|\,h_X(x)$. Either some analytical investigation or a crude simulation can be used for this purpose.

Importance Sampling: Application Example

Importance sampling is in this section applied to the example in Section 4.2. In the one cell scenario, the output quantities can be considered to be functions of the three random variables X, V, W:

$$G = f\,(X, V, W) \tag{4.30}$$

where X is a one dimensional discrete random variable that represents the number of users in a snapshot, V is the random variable of user positions in the cell, and W corresponds

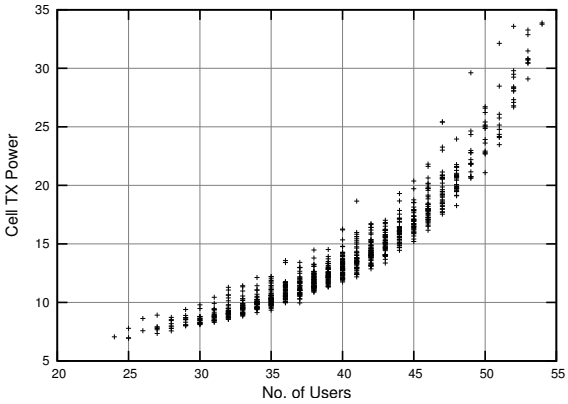


Figure 4.9: Correlation between total cell TX power and number of users for the simple analysis example

to the random shadowing values for the individual users. The dimension of V and W is the number X_i of actually drawn users in each traffic realization.

IS requires a random input variable that is strongly correlated with the output quantities. This random variable is then modified according to the description above. In Figure 4.9 the correlation between number of users X and the total cell TX power $p^{\downarrow}$ is illustrated (if cell TX power were unlimited). The blocking probability as second quantity of interest only depends on the total cell TX power, so that there is a strong correlation also between blocking probability and amount of offered traffic X. Hence the Poisson distributed random variable X is chosen as basis for the importance sampling of both quantities. In principle instead of X it would also be possible to select one of the multi-dimensional random variables V or W – or even all three random variables together. However, in this case finding appropriate modified random distributions would become extremely difficult. Also, to use a multi-dimensional random variable would require a complex and time-consuming generation of correlated random variables (cf. Appendix A).

In accordance with (4.29), given X to be used for the importance sampling, an optimal choice for the distribution function of the modified discrete random variable Y is

$$h_Y(n) = \frac{\mathrm{E}\left\{f\left(n,V,W\right)\right\} h_X(n)}{\sum\limits_{i=0}^{\infty} \mathrm{E}\left\{f\left(i,V,W\right)\right\} h_X(i)} \tag{4.31}$$

Of course, $\mathrm{E}\left\{f\left(n,V,W\right)\right\}$ needs to be estimated in practice. In fact, if an exact evaluation were possible without considerable effort, also $\mathrm{E}\left\{G\right\}$ could be easily calculated. For this example $\mathrm{E}\left\{f\left(n,V,W\right)\right\}$ was estimated based on a pilot run. It was observed that for the estimation of the mean total output power less than 100 Monte-Carlo trials in the pilot run were required in order to achieve almost maximum variance reduction. For the

estimation of blocking probabilities the amount of trials per pilot run strongly depends on the blocking probability to be estimated. However, the objective here is not to determine the best choice for the number of pilot trials, but to illustrate the principles and potentials of importance sampling. The number of trials per pilot run was thus chosen large enough to have no significant impact on the accuracy of results.

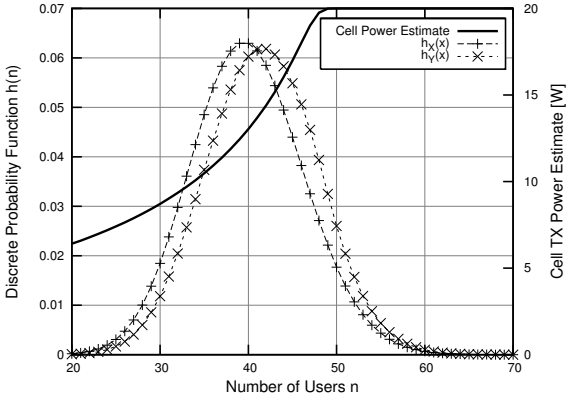


Figure 4.10: Application of importance sampling: Original distribution functions $h_X(n)$, cell TX power estimate $g(n)$, and modified distribution function $h_Y(n)$

In Figure 4.10, cell TX power estimate, original discrete probability function $h_X(n)$ and modified discrete probability function $h_Y(n)$ are exemplarily presented for the cell TX power estimation. The resulting discrete probability function $h_Y(n)$ has similar shape as the Poisson distribution $h_X(n)$, but has a larger mean.

Importance sampling was carried out for an estimation of both, mean cell TX power and blocking probability. The variance reduction coefficient ψ was evaluated for various amounts of users and several traffic types. The results are depicted in Figure 4.11 and Figure 4.12. Both plots show a considerable reduction in number of required snapshots. In particular for the estimation of low blocking probabilities very high gains can be achieved ("rare-event simulation"), whereas for larger blocking probabilities the gain becomes less significant. For the estimation of power levels, also quite significant reduction factors of typically larger than ten can be obtained. As an interesting result, the variance reduction in estimating the mean cell TX power levels only slightly depend on the traffic type.

General observation in this study was that even slight changes in the choice of $h_Y(n)$ can lead to significant performance degradations, in particular when estimating blocking probabilities. This sensitivity of IS towards its input parameters makes the method very difficult to handle.

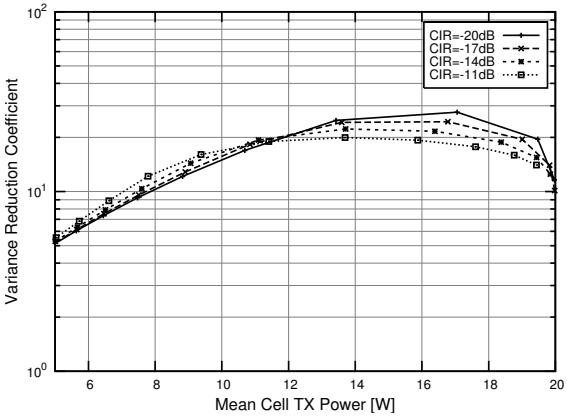


Figure 4.11: Variance reduction in estimating the mean TX power applying importance sampling

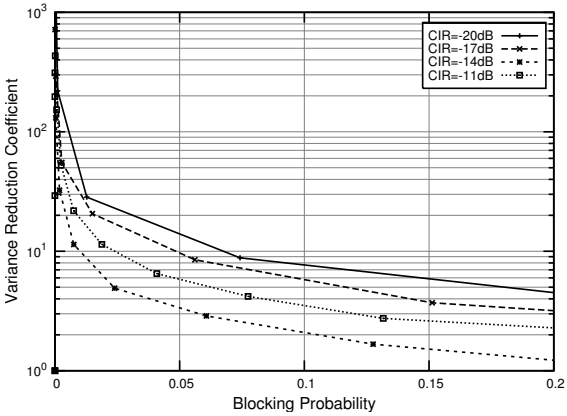


Figure 4.12: Variance reduction in estimating the blocking probability applying importance sampling

4.3.2 Stratified Sampling

Another widely used variance reduction technique is *Stratified Sampling* (StrS). The basic idea is to divide the input space into disjoint subregions which are analyzed separately by the Monte-Carlo method. Variance reduction is achieved by an appropriate choice of the number of Monte-Carlo trials for each subregion. More samples are taken in the more important regions, i.e. regions for which the variance of the output quantity is comparatively higher. As the idea of stratified sampling is very similar to that of importance sampling, it will not be further considered.

4.3.3 Control Variates

The application of *Control Variates* (CV) requires a random variable C that is correlated with the actual quantity of interest G. In other words, for each result G_i of the Monte-Carlo trial i there is a quantity C_i required that is correlated with G_i. Furthermore, the expected value μ_C of the *control variate* C needs to be known a priori. The control variate can make use of the same input variables that are also considered for the evaluation of G_i. For the generation of the control variate often a simplified system model is used, e.g. an analytic approximation of the analyzed system.

Variance reduction can be achieved by appropriately considering the control variate in a modified estimator for G. Instead of estimating G directly a modified random variable $G^*(\beta)$ is applied:

$$G^*(\beta) = G - \beta\,(C - \mu_C) \tag{4.32}$$

where β is an arbitrary scalar.

Accordingly an unbiased estimator for the expected value of G based on N Monte-Carlo trials is

$$m_{G^*(\beta)}(N) = \frac{1}{N}\sum_{i=1}^{N}\left(G_i - \beta\left(C_i - \mu_C\right)\right) \tag{4.33}$$

with

$$\lim_{N\to\infty} m_{G^*(\beta)}(N) = \mu_G \tag{4.34}$$

The variance of a Monte-Carlo sample-mean estimator is proportional to the variance of the estimate itself. Hence, variance reduction is achieved if the variance of $G^*(\beta)$ is lower than the variance of G:

$$\text{var}\left\{G^*(\beta)\right\} = \text{var}\left\{G\right\} - 2\beta\,\text{cov}\left\{G,\,C\right\} + \beta^2\text{var}\left\{C\right\} \tag{4.35}$$

Hence, the following condition needs to be satisfied:

$$2\beta\,\text{cov}\left\{G,\,C\right\} > \beta^2\text{var}\left\{C\right\} \tag{4.36}$$

The β^* that minimizes the variance can be easily calculated by differentiating (4.35):

$$\begin{aligned} &\frac{d\left\{\text{var}\left\{G^*(\beta^*)\right\}\right\}}{d\beta^*} = -2\,\text{cov}\left\{G,\,C\right\} + 2\,\beta^*\text{var}\left\{C\right\} = 0 \\ \Rightarrow\quad &\beta^* = \frac{\text{cov}\left\{G,\,C\right\}}{\text{var}\left\{C\right\}} \end{aligned} \tag{4.37}$$

For the minimum variance it accordingly yields

$$\text{var}\left\{G^*(\beta^*)\right\} = \left(1 - \frac{\text{cov}^2\left\{G,\,C\right\}}{\text{var}\left\{G\right\}\text{var}\left\{C\right\}}\right)\text{var}\left\{G\right\} \tag{4.38}$$

The more G and C are correlated, the greater is the variance reduction.

Besides finding a suitable control variate that has a high positive correlation with G, the choice of β is very crucial. Typically the covariance between G and C is unknown and needs to be estimated. Of course it can be estimated based on already carried out Monte-Carlo trials, however, the estimator (4.33) is then biased.

An important advantage of the application of control variates compared to importance sampling and stratified sampling is, that the input random variables do not need to be modified and thus multiple quantities can be estimated in one simulation run.

Control Variates: Application Example

Control variates can be applied to the example in Section 4.2 in a similar way as IS. The probability density function of the control variate C is chosen as

$$h_C(n) = \mathrm{E}\left\{f\left(n, Y, Z\right)\right\} \tag{4.39}$$

and accordingly

$$\mu_C = \sum_{n=0}^{\infty} \mathrm{E}\left\{f\left(n, Y, Z\right)\right\} h_X\left(n\right) \tag{4.40}$$

A positive correlation between C and Y is assured by using the same samples X_i for calculating G_i and C_i. Due to the high correlation (cf. Figure 4.9), β is chosen to one.

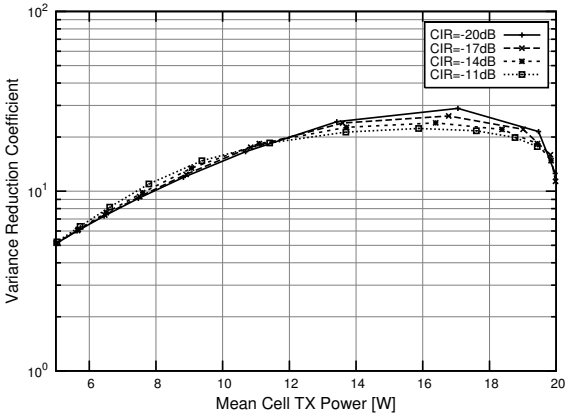


Figure 4.13: Variance reduction in estimating mean cell TX power applying control variates

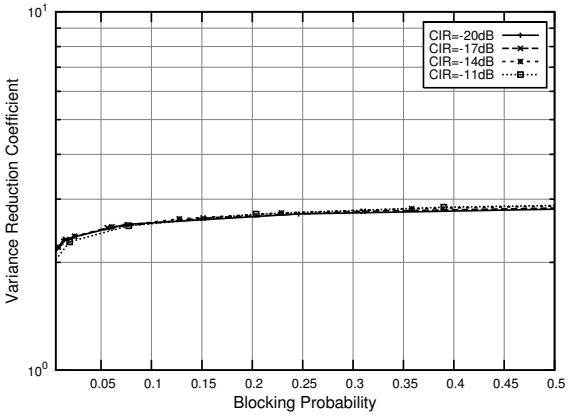


Figure 4.14: Variance reduction in estimating blocking probabilities applying control variates

The reduction factors in required snapshots compared with the sample-mean Monte-Carlo method are depicted in Figure 4.13 and Figure 4.14. As an interesting observation, the reduction in snapshots for estimating cell power levels is almost exactly the same as for the application of importance sampling. In contrast the reduction of snapshots required for estimating blocking probabilities is much lower than for importance sampling. Both are almost independent of the traffic type.

4.3.4 Correlated Sampling

The area of application of *Correlated Sampling* (CorS) is different compared with the three techniques described above. Correlated sampling can be applied if the differences between the output quantities of two similar systems are to be estimated [Rub81]. Instead of evaluating the two systems by means of two independent Monte-Carlo runs, the two series of Monte-Carlo trials are correlated with each other. This can in the easiest case be achieved by using the same samples of the input variables for both Monte-Carlo runs. In the scope of this thesis, CorS is applied to compare the performance of two different network configurations, e.g. for optimization purposes. For the evaluation of the performance of the two configurations the same random numbers, and consequently also the same traffic constellations (“snapshots”), are used.

Considered is the difference between the mean of two output quantities of two systems

$$D = \mathrm{E}\{S\} - \mathrm{E}\{U\} \tag{4.41}$$

where

$$S = f(X) \quad X \in R^n \tag{4.42}$$
$$U = g(Y) \quad Y \in R^n \tag{4.43}$$

For an estimation of the expected values of the random variables a Monte-Carlo sample-mean estimator is applied:

$$m_S(N) = \frac{1}{N}\sum_{i=1}^{N} f(X_i) \tag{4.44}$$
$$m_U(N) = \frac{1}{N}\sum_{i=1}^{N} g(Y_i) \tag{4.45}$$
$$m_D(N) = \frac{1}{N}\left(\sum_{i=1}^{N} f(X_i) - \sum_{i=1}^{N} g(Y_i)\right) \tag{4.46}$$

For the variance of $m_D(N)$ it yields

$$\text{var}\{m_D(N)\} = \text{var}\{m_S(N)\} + \text{var}\{m_U(N)\} - 2\,\text{cov}\{m_S(N), m_U(N)\} \tag{4.47}$$

A straightforward application of the sample-mean Monte-Carlo algorithm suggests an independent estimation of the two quantities S and U using two independent series of samples X_i and Y_i. In this case, $m_S(N)$ and $m_U(N)$ are statistically independent and $\text{cov}\{m_S(N), m_U(N)\} = 0$. On the other hand, if X and Y are positively correlated, e.g. by using the same generating random numbers, and $f(x)$ and $g(x)$ are similar, then $\text{cov}\{m_S(N), m_U(N)\} > 0$ and the variance of $m_D(N)$ can be significantly reduced.

Correlated Sampling: Application Example

Also correlated sampling is illustrated with the example scenario described in Section 4.2. Typical application, e.g. in network optimization, is to determine if one configuration is better or worse than another configuration. In this section, the original configuration is compared with a second configuration with increased base station antenna height. The differences in blocking and mean TX power are estimated.

Assuming that the difference of the two output quantities is greater than zero ($\mu_D > 0$), it is required that also for the estimate $m_D(N)$ it yields

$$m_D(N) > 0 \tag{4.48}$$

with sufficient probability (confidence).

From (4.7) with $a = -3$ and $b = \infty$ (99.87 % confidence) it yields

$$\mu_D - m_D(N) \leq \frac{3\sigma_D}{\sqrt{N}} \tag{4.49}$$

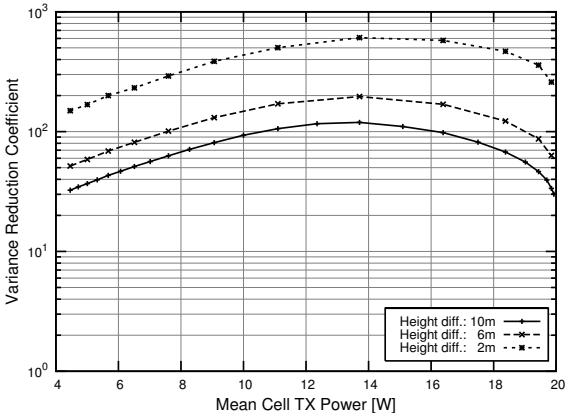


Figure 4.15: Variance reduction in estimating the difference in cell TX power for different changes in BS antenna height applying correlated sampling

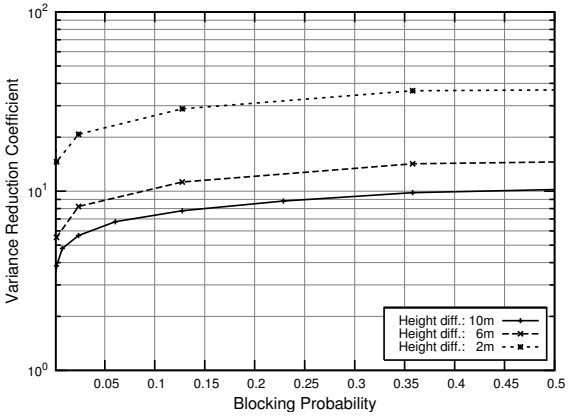


Figure 4.16: Variance reduction in estimating the difference in blocking probability for different changes in BS antenna height applying correlated sampling

For the edge of the error range ($m_D(N) = 0$), the following expression for the number of required samples is obtained from (4.49):

$$N > \left(\frac{3\sigma_D}{\mu_D}\right)^2 \tag{4.50}$$

Equivalently the same result can be deduced for $\mu_D < 0$.

The required number of snapshots using the same series of samples is in Figure 4.15 and Figure 4.16 compared with the number of required snapshots if independent series of

samples were used. Results are recorded for three different increases of antenna height: 2 m, 6 m, and 10 m. For the traffic demand a CIR of -14 dBm was chosen. Significant reduction factors of up to 600 for estimating the cell TX power level, and 35 for estimating the blocking probability are obtained. Similar to the results obtained for the application of control variates, the reduction for estimating mean cell TX power levels is higher than for estimating blocking probabilities. The speed-up is significantly higher, the smaller the difference in antenna height – hence the more similar the systems are.

4.4 Conclusions

In this chapter the Monte-Carlo method as a powerful approach for the analysis of systems like the mobile radio networks considered in the scope of this thesis has been introduced. The most crucial parameter is the number of analyzed Monte-Carlo trials, which is proportional to the runtime of the method. It is shown that the number of required trials to yield a certain statistical significance is proportional to the variance of the estimated output quantity. The latter variance depends on the analyzed problem and can hardly be obtained a priori. From the examples presented in this chapter, in particular a strong dependency on the analyzed traffic can be observed. Fortunately, the number of required trials can be effectively determined during the runtime of the method by means of confidence intervals.

Besides the pure estimation of the number of required trials, its reduction is an important means for improving the performance of the method. This can be achieved by variance reduction techniques, which exploit a priori knowledge about the problem, e.g. correlations between quantities. Four variance reduction techniques have been identified as being applicable to the problems considered in this thesis. Three of them have been applied to a simplified network planning scenario, yielding very high gains in different applications. In this study *Importance Sampling* (IS), while providing a high variance reduction, showed a very high sensitivity towards its input parameters. It also needs to be tuned towards a particular quantity of interest and is thus difficult to be applied to network planning scenarios, where typically a large number of output quantities need to be estimated at the same time. *Control Variates* (CV) and *Correlated Sampling* (CorS) do not impose this restriction. The example applications in this chapter make evident that with CV a high variance reduction in estimating the performance of a system can be achieved. CorS shows a very high variance reduction when comparing the performance of two systems, which is an important and recurring task in network planning. Moreover, CorS is very simple to implement and does not require an application specific tuning of its parameters. Both, CV and CorS are further considered for the more complex and realistic application scenarios in Chapter 5.

Chapter 5

Monte-Carlo Snapshot Analysis for WCDMA Network Performance Evaluation

The application of the Monte-Carlo method introduced in Chapter 4 constitutes the de facto standard for the analysis of WCDMA radio networks in network planning. The application of the Monte-Carlo method is in this case generally denoted as *snapshot simulation*. Sometimes also the term *static simulation* is used, as opposed to *static analysis*, which denotes analytical approaches for performance evaluation.

Although widely used and accepted, all of the snapshot simulator realizations known to the author still have severe weaknesses with respect to performance, i.e. the size of networks which can be handled in terms of memory consumption and simulation time is quite limited. Typical networks to be analyzed can easily consist of several thousand cells, and several tens of thousand active users per snapshot. In addition there are typically limitations with respect to the modeling of important system features, e.g. HSDPA, and with respect to the consideration of QoS mechanisms. In this chapter a very flexible and fast Monte-Carlo snapshot analysis approach for the performance evaluation of WCDMA radio networks is developed. The approach described in this chapter extends the approach published by the author in [TPL+03].

5.1 Basic Analysis Loop

The simulation approach is based on the Monte-Carlo method the properties of which are presented in the previous chapter. The basic analysis loop is depicted in Figure 5.1.

In each iteration of the loop independent *snapshots* of possible traffic constellations are generated by appropriate sampling of the input random variables. This process is in the following referred to as *snapshot generation*. During *snapshot evaluation* for each of the individual snapshots the most likely system state is estimated and the quantities of

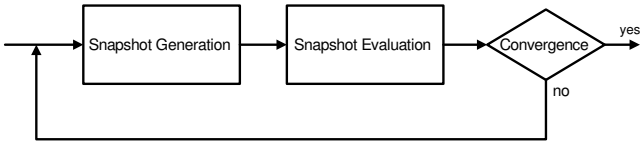


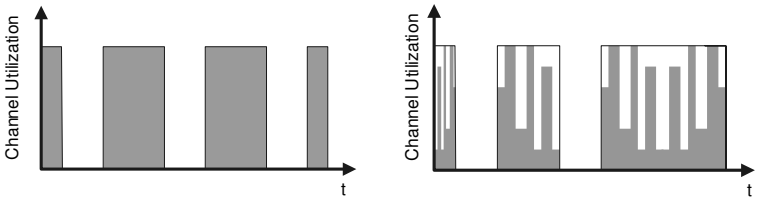
Figure 5.1: Basic snapshot simulation process

interest are calculated. A system state is thereby defined by the currently active users, the resource allocation to users (including blocking/dropping), and the corresponding power and interference levels, etc. In accordance with the Monte-Carlo method, by combining the results of a large number of snapshots, statistically significant results are obtained. The analysis loop is terminated after sufficient convergence of all quantities of interest has been achieved.

5.2 Snapshot Generation

A snapshot by definition represents the system in one particular instant in time. The task of the snapshot generation is to generate all required information that characterizes the state of the system. In radio network simulations the network itself is considered deterministic whereas traffic and communication links are modeled stochastically. The traffic is characterized by the active users in the system and the services demanded by them.

The offered traffic is composed of different traffic types. A traffic type is characterized by its service, equipment, mobility type, and user distribution.



(a) CS Source Model (b) PS Source Model

Figure 5.2: Modeling of CS and PS traffic sources

Both, PS and CS services are considered. For CS services an *on/off source model* is used (cf. Figure 5.2(a)). During the on-periods the service is assumed to fully utilize the allocated channel. The on-ratio is referred to as *connection activity factor*. A PS session is assumed to consist of a number of *packet calls* (cf. Figure 5.2(b)). Within a packet call the source not necessarily fully utilizes the allocated bandwidth, e.g. due to the intrinsic behavior of TCP/IP. The average fraction of transmitted data amount and maximum trans-

ferable data amount within packet calls is in the following referred to as *transmit activity factor*. The fraction of time for packet calls and the total session duration is in consistence with the CS case denoted *connection activity factor*.

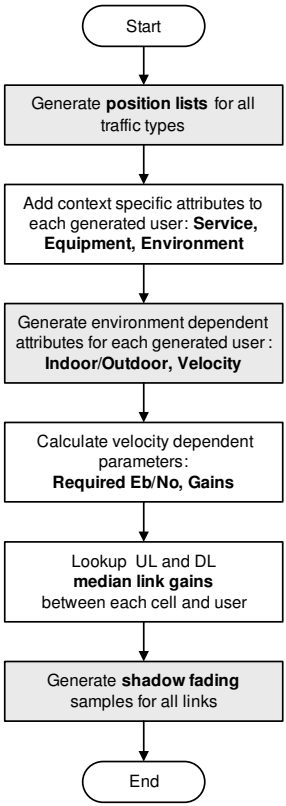


Figure 5.3: Snapshot generation process (gray shaded blocks comprise randomness)

The complete snapshot generation process is presented in Figure 5.3. At first, users and corresponding user positions are generated for all traffic types. Spatial distribution and average amount of users are given by pixel-oriented traffic maps. User arrival rates and call holding times are assumed to be negative exponentially distributed, the user arrival process hence is a Poisson process. If the blocking is low, the total offered traffic at a certain point in time can be assumed to be Poisson distributed with mean value m accumulated over all pixels:

$$p(x) = \frac{\mathrm{e}^{-m} m^x}{x!} \tag{5.1}$$

An efficient algorithm for the generation of user positions was developed in [LPTG03].

Depending on the users' positions and traffic types, users are in the next step assigned with a service type, an equipment type, and an environment type. Taking as input the indoor probability, which depends on the users' environment, it is randomly decided whether the user is indoor or outdoor. Indoor users differ from outdoor users by an additional contribution to the total path loss, denoted as *penetration loss*, and the velocity distribution. From the velocity distribution for each user a velocity is randomly drawn. Once the velocity is known, the velocity dependent parameters can be calculated. These in particular are the required UL and DL link quality targets (i.e. E_b/N_0) and fast fading contributions (cf. Section 3.5.1). For PS users it is in addition decided if they are currently active or inactive, where the probability of a user being active is given by the connection activity factor.

In the final two steps the link gains between all cells and users are generated. First, the pre-calculated median link gains are looked up. These consist of antenna gain and path loss as described in Chapter 3. In decibels these values describe the mean (and median) value of the normal distributed random shadowing variables. In linear representation they constitute the median values of the corresponding log-normal distributed random variables. In the next step, the shadow fading values are randomly generated according to the description in Section 3.2.3, yielding the total link gains.

5.3 Modeling of Dynamics

The WCDMA system underlies a great amount of dynamic effects. As the presented analysis method is based on snapshots of the system, each of which represents an instantaneous point in time, the modeling of the dynamics is not straightforward. The effects of dynamics have to be included indirectly.

In particular the following dynamics have a considerable impact:

Source Activity causes fluctuation in interference on different levels. Firstly, on/off sources change their state. Secondly, due to a non-constant bandwidth demand in on-periods the data rate also changes (e.g. DTX).

Fast fading, also often referred to as multi-path fading, leads to short-term variations of the link gains between cells and mobiles (cf. Section 3.3).

Mobility causes changes in the propagation environment and cell assignment.

Most of the quantities of interest are additively influenced by a number of independent and dependent random contributions. Additionally, the RRM algorithms do not operate on instantaneous measurements, but filter measurements over time. It is thus appropriate to model the interference of CS sources as an average over their on- and off-periods, whereas for PS services this does not hold. PS sources in general have a much lower activity, this is the reason for considering PS users to be active and to contribute to interference only during packet calls. During a packet call users also have a certain activity, which can be averaged for the same reason as for CS users.

To some extent the arguments given for CS sources support the decision to neglect fast fading and the use of average link gains. However, in the description of the *effective interference raise* in Section 3.5.1 it was shown that by neglecting fast fading the output power of a power controlled channel is systematically underestimated. Hence, fast fading is considered in terms of the effective interference raise in the calculations of power and interference levels.

The evaluation of connection power levels requires special attention. Connection powers at least in UL have hard limits on their minimum and maximum level. If connection powers were evaluated for the on-periods based on the mean interference, a comparison against hard limits is not reasonable, as the power considerably varies due to fast fading. For this reason the maximum TX power of a mobile for instance has to be reduced for comparison with the mean power value. The required margin is called *required power increase* or *power control headroom*, and has already been introduced in Section 3.5.2.

In addition to the effective interference raise and power control headroom, also the SHO gains as briefly introduced in Section 3.5.3 need to be considered in the analysis, resulting in a complex gain concept that comprises these three parameters as functions of various input quantities like user's velocity and number and relation of links in SHO. Further details can be found in [WTL+03] and [EFF+03].

5.4 System Constraints and Limits

A set of constraints and limits influence the overall quality of the WCDMA network and thus have to be considered. These can be grouped into those applicable on cell basis and those applicable on connection basis.

First of all there are limits applicable to individual connections. They are mainly due to natural limitations of the equipment and can hardly be controlled by the network.

Pilot Power

The pilot channel has to be received with an adequate strength in order to guarantee proper operation for each user i:

$$\hat{r}_i \geq \hat{r}_{i,\,min} \tag{5.2}$$

where $\hat{r}_i$ is the received pilot power, and $\hat{r}_{i,\,min}$ is the minimum required signal strength, which is subject to the equipment assigned to user i.

In addition, also the received pilot quality is required to be above a certain threshold $\hat{\gamma}_{i,\,min}$:

$$\frac{\hat{r}_i}{i_i^{\perp}} \geq \hat{\gamma}_{i,\,min} \tag{5.3}$$

where $i_i^{\downarrow}$ is the total interference perceived by user i. The minimum pilot quality level is in practice rather low and can generally be neglected compared with the other constraints. However, the pilot quality is considered as a measure for a well designed network and thus often a higher level than actually required is self-imposed by the network design process.

Minimum and Maximum Connection Power

The minimum output power of mobiles is restricted. This can lead to the situation where the mobile needs to transmit with higher power than actually required from the perspective of the power control procedure. In particular in indoor environments this can cause additional, undesired interference and needs to be taken into account. Obviously the maximum output power of a mobile is also limited. If the power control requires a larger power level than can be provided by the equipment, the transmission is disturbed and eventually the connection will be dropped. With these restrictions for the mean UL transmit power $t_i^{\uparrow}$ of user i it yields

$$t_{i,\,min}^{\uparrow} \leq t_i^{\uparrow} \leq t_{i,\,max}^{\uparrow} \tag{5.4}$$

where $t_{i,\,min}^{\uparrow}$ and $t_{i,\,max}^{\uparrow}$ are the respective minimum and maximum constraints imposed on the mean UL transmit power.

Also in DL the amount of power dedicated to one particular mobile is limited. This is mainly due to the fact that it is reasonable to limit the share of the total power allocated to one user:

$$t_i^{\downarrow} \leq t_{i,\,max}^{\downarrow} \tag{5.5}$$

In addition to the limits applicable on connection level, there are limitations on the resources available on cell basis.

Cell Load

The cell load in UL and DL is restricted by the network. The cell load of cell c in DL is defined as the ratio of the current transmission power $p_c^{\downarrow}$ to the maximum transmit power $p_{c,\,max}^{\downarrow}$ of the cell:

$$l_c^{\downarrow} = \frac{p_c^{\downarrow}}{p_{c,\,max}^{\downarrow}} \tag{5.6}$$

The cell load is restricted by the RRM algorithms to not exceed a certain maximum level $l_{c,\,max}^{\downarrow}$:

$$l_c^{\downarrow} \leq l_{c,\,max}^{\downarrow} \tag{5.7}$$

The UL cell load is defined as

$$l_c^{\uparrow} = 1 - \frac{i_c^{\uparrow}}{\eta_c^{\uparrow}} \tag{5.8}$$

where $\eta_c^{\uparrow}$ is the total receiver noise and $i_c^{\uparrow}$ is the total interference plus receiver noise perceived by cell c. In order to assure system stability and to limit cell breathing, also $l_c^{\uparrow}$ is limited by RRM algorithms:

$$l_c^{\uparrow} \leq l_{c,max}^{\uparrow} \tag{5.9}$$

UL and DL cell load levels might underlie further restrictions imposed by QoS mechanisms (cf. Section 5.6).

Code Limits

A cell has a number of scrambling codes available, each of which employs a tree of orthogonal spreading codes. The number of codes and consequently the number of users per tree is limited. If a tree is fully occupied either another scrambling code has to be used or the user has to be blocked. The use of an additional scrambling code however degrades the interference situation, as in general a considerable amount of orthogonality is lost.

Assuming dynamic reallocation of spreading codes, for the sum of all spreading codes i with corresponding spreading factor SF_i allocated to one scrambling code yields

$$\sum_i SF_i \leq 1 \tag{5.10}$$

Hardware Limits

The hardware resources in the base station (Node B) put a constraint on the number of users that can be served and the data rate that can be provided. These constraints are in the following not further considered as it is assumed that a proper dimensioning of the according hardware is carried out after the design of the radio network. The blocking due to insufficient hardware resources can thus be neglected.

5.5 Cell Based Radio Performance Evaluation

In this section the core of the snapshot analysis is derived: A fast evaluation method for power and interference levels given a traffic constellation and resource scheduling. The analysis algorithm described in detail in Section 5.6 will make use of this to evaluate the system, to check the cell and connection limits presented in the previous section, and to appropriately schedule resources.

Cell resources are controlled by the RRM. Dependent on the type of user being served, different possibilities for scheduling resources to them are available.

In Chapter 2 two principle ways to transmit data were introduced. Data can either be transmitted via dedicated channels, i.e. DCH, or shared channels. Examples for shared channels are the RACH in UL direction or the FACH and the HSDPA channels in DL direction. In this section it is assumed that active users in UL are always served by dedicated channels. In DL users can be either served by dedicated channels or the FACH. The FACH is a fixed bandwidth channel shared among the users served. The FACH has a fixed power allocation. An extension towards HSDPA is presented in Section 5.7.2. The RACH is not considered for user data transmission as it is a rather low-rate channel and most appropriate for carrying signaling traffic.

The DCH is power controlled, its data rate can be adapted (*bit rate adaptation*). The handling of CS and PS users is slightly different:

- CS users are considered as on/off sources with an activity rate ϵ. They are always mapped on a DCH. During an on-period they require a certain bit energy to interference plus noise ratio target ${}^{E_b}/_{N_0}$. During their off-period they are assumed to be inactive, thereby control channels are neglected. CS users can (typically) either be served at their demanded bandwidth and quality demand or have to be blocked/dropped.
- PS users demand a number of packet calls during their sessions. A packet call is considered as an amount of data to be transmitted. In contrast to CS users, there exist a number of possible bearer mappings for PS users. The channel type and bandwidth allocated to them is decided on by the system based on the current load situation. Users have an ${}^{E_b}/_{N_0}$ target and an activity factor ϵ that depend on the data rate allocated to them.

Evaluation of radio performance mainly means to evaluate the power levels allocated to the individual dedicated channels. The latter are power controlled, and as a consequence power levels mutually depend on each other. For each connection a linear equation describing this dependency can be obtained (see e.g. [HT04]). Though the original set of equations is linear, some non-linearities appear as the limitations and constraints presented in Section 5.4 are to be considered. Thus the state-of-the-art approach is to use an iterative method that also permits effective exploitation of the sparsity of the corresponding matrix [WLSSJ99].

In the following a novel approach is presented that significantly speeds up the calculation. In particular the size of the set of linear equations is drastically reduced to the number of cells rather than the number of connections. The fundamentals of this approach have been presented in [MH01]. In this chapter it is extended to include multiple scrambling codes in DL and adapted to the particular needs of the analysis method developed in this work.

A single carrier WCDMA system consisting of a set of cells $\mathcal{C}$ is assumed. All cells $c \in \mathcal{C}$ share the same frequency. Furthermore it is assumed that the resources have been

scheduled by the network. Each user i from the set of served users $\mathcal{M}$ has a certain $\frac{E_b}{N_0}$ target and activity factor ϵ_i in both UL and DL direction. At this stage it is assumed that each user is served by exactly one cell.The impact of macro diversity is explained further below. A user i is served by a certain primary serving cell c if the received pilot power of this cell exceeds the received pilot power of all other cells:

$$\hat{\tau}_c \xi_{ci}^{\downarrow} \geq \hat{\tau}_j \xi_{ji}^{\downarrow} \qquad \forall j \in \mathcal{C},\ j \neq c \tag{5.11}$$

where $\hat{\tau}_j$ is the pilot power assigned to cell j and $\xi_{ji}^{\downarrow}$ is the DL link gain between cell j and user i. In case of multiple equally received pilot signals, the primary serving cell is randomly chosen. According to their primary serving cells j, the total set of users $\mathcal{M}$ is divided into sets $\mathcal{M}_j$.

5.5.1 Uplink Equations

Assuming perfect power control, transmit power levels for dedicated channels are adjusted such that the received signal to interference plus noise ratio fulfills the quality requirement imposed by the required (static) $\frac{E_b}{N_0}$:

$$\frac{r^{\uparrow}}{i^{\uparrow} - r^{\uparrow}} = \frac{1}{PG} \cdot \frac{E_b}{N_0} \tag{5.12}$$

where $r^{\uparrow}$ is the received signal power of a certain user, $i^{\uparrow}$ is the total received cell interference and PG is the processing gain introduced in Section 3.5.

For the following derivations it is convenient to define the required signal to interference ratio $\gamma^{\uparrow}$ as received signal levels over total interference rather than over the visible interference:

$$\gamma^{\uparrow} = \frac{r^{\uparrow}}{i^{\uparrow}} \tag{5.13}$$

With this definition and (5.12) follows:

$$\frac{r^{\uparrow}}{i^{\uparrow} - r^{\uparrow}} = \frac{\frac{r^{\uparrow}}{i^{\uparrow}}}{1 - \frac{r^{\uparrow}}{i^{\uparrow}}} = \frac{\gamma^{\uparrow}}{1 - \gamma^{\uparrow}} = \frac{1}{PG} \cdot \frac{E_b}{N_0} \tag{5.14}$$

Hence, $\gamma^{\uparrow}$ can be expressed as

$$\gamma^{\uparrow} = \frac{\frac{E_b}{N_0}}{PG + \frac{E_b}{N_0}} \tag{5.15}$$

The total received UL power from user i at cell j is

$$r_{ji}^{\uparrow} = \kappa_{ji}^{\uparrow} t_i^{\uparrow} \xi_{ij}^{\uparrow} \tag{5.16}$$

where $\xi_{ij}^{\uparrow}$ is the UL link gain between user i and cell j, and $\kappa_{ji}^{\uparrow}$ is the *effective interference raise* explained in Section 3.5.1:

$$\kappa_{ji}^{\uparrow} = \begin{cases} 1, & \text{for } j \text{ serving cell} \\ > 1, & \text{otherwise} \end{cases} \tag{5.17}$$

The transmitted power of user i that is visible to its serving cell c is power controlled such that the signal to interference target is maintained:

$$t_i^{\uparrow} = \frac{r_{ci}^{\uparrow}}{\xi_{ic}^{\uparrow}} = \gamma_i^{\uparrow} \frac{i_c^{\uparrow}}{\xi_{ic}^{\uparrow}} \tag{5.18}$$

With (5.18) in (5.16) it yields

$$r_{ji}^{\uparrow} = \kappa_{ji}^{\uparrow} \xi_{ij}^{\uparrow} t_i^{\uparrow} \tag{5.19}$$

$$= \kappa_{ji}^{\uparrow} \gamma_i^{\uparrow} \frac{\xi_{ij}^{\uparrow}}{\xi_{ic}^{\uparrow}} i_c^{\uparrow} \tag{5.20}$$

The total received interference of cell c is the sum of the received power from all mobiles, the receiver noise $\eta_c^{\uparrow}$, and the total perceived interference from other systems $\check{i}_c^{\uparrow}$ (adjacent channel interference):

$$i_c^{\uparrow} = \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{M}_i} \epsilon_j^{\uparrow} r_{cj}^{\uparrow} + \eta_c^{\uparrow} + \check{i}_c^{\uparrow} \tag{5.21}$$

Substituting (5.19) and (5.20) in (5.21) yields

$$i_c^{\uparrow} = \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{M}_i} \epsilon_j^{\uparrow} \kappa_{cj}^{\uparrow} \xi_{jc}^{\uparrow} t_j^{\uparrow} + \eta_c^{\uparrow} + \check{i}_c^{\uparrow} \tag{5.22}$$

$$= \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{M}_i} \epsilon_j^{\uparrow} \kappa_{cj}^{\uparrow} \gamma_j^{\uparrow} \frac{\xi_{jc}^{\uparrow}}{\xi_{ji}^{\uparrow}} i_i^{\uparrow} + \eta_c^{\uparrow} + \check{i}_c^{\uparrow} \tag{5.23}$$

$$= \underbrace{\sum_{j \in \mathcal{M}_c} \epsilon_j^{\uparrow} \gamma_j^{\uparrow}}_{T_{cc}^{\uparrow}} \cdot i_c^{\uparrow} + \sum_{\substack{i \in \mathcal{C} \\ i \neq c}} \underbrace{\sum_{j \in \mathcal{M}_i} \epsilon_j^{\uparrow} \kappa_{cj}^{\uparrow} \gamma_j^{\uparrow} \frac{\xi_{jc}^{\uparrow}}{\xi_{ji}^{\uparrow}}}_{T_{ci}^{\uparrow}} \cdot i_i^{\uparrow} + \underbrace{\eta_c^{\uparrow} + \check{i}_c^{\uparrow}}_{h_c^{\uparrow}} \tag{5.24}$$

Equation (5.24) gives a set of $|\mathcal{C}|$ linear equations that can be expressed in matrix form:

$$\vec{i} = T^{\uparrow} \cdot \vec{i} + \vec{h}^{\uparrow} \tag{5.25}$$

$$= (I - T^{\uparrow})^{-1} \cdot \vec{h}^{\uparrow} \tag{5.26}$$

where $T^{\uparrow}$ is denoted as the characteristic *UL cell coupling matrix*, and $\vec{i}$ is the vector of cell interference levels $i_i^{\uparrow}$. The inhomogeneity vector $\vec{h}^{\uparrow}$ describes the amount of static interference, which is not subject to power control. Equation (5.26) makes evident that $\vec{h}^{\uparrow}$ multiplicatively contributes to the total cell interference. Hence, a reduction of $\vec{h}^{\uparrow}$, e.g. by reducing the receiver noise figures (enhanced equipment), is a powerful means for improving the overall network performance.

5.5.2 Downlink Equations

The calculation of the DL power levels is more complicated compared to the UL calculations. In particular multiple scrambling codes per cell need to be considered. For this, $\mathcal{M}_c$ is divided in sets $\mathcal{M}_{c,k}$ based on the respective serving scrambling codes k.

All signals to users served by the same scrambling code are in principle orthogonal and thus do not interfere with each other. In practice the orthogonality is not perfect - mainly due to multi-path propagation effects. The interference is consequently multiplied by an orthogonality factor denoted α. In accordance with (5.12) the DL relation between received signal strength, interference, and required DL $\frac{E_b}{N_0}$ is

$$\frac{r^{\downarrow}}{i^{\downarrow} - \alpha \cdot r^{\downarrow}} = \frac{1}{PG} \cdot \frac{E_b}{N_0} \tag{5.27}$$

Hence, for the DL an expression equivalent to (5.15) can be derived:

$$\gamma^{\downarrow} = \frac{r^{\downarrow}}{i^{\downarrow}} \tag{5.28}$$

$$\Rightarrow \quad \frac{r^{\downarrow}}{i^{\downarrow} - \alpha \cdot r^{\downarrow}} = \frac{\gamma^{\downarrow}}{1 - \alpha \cdot \gamma^{\downarrow}} = \frac{1}{PG} \cdot \frac{E_b}{N_0} \tag{5.29}$$

$$\Rightarrow \quad \gamma^{\downarrow} = \frac{\frac{E_b}{N_0}}{PG + \alpha \cdot \frac{E_b}{N_0}} \tag{5.30}$$

With this prerequisite the characteristic DL equations can be derived.

The total received interference experienced by user i served by primary cell c and scrambling code k is the sum of all received interference levels plus the receiver noise $\eta_i^{\downarrow}$ and the other system interference $\breve{\imath}^{\downarrow}$:

$$i_i^{\downarrow} = \alpha_i \kappa_{ci}^{\downarrow} \xi_{ci}^{\downarrow} v_{c,k} + \sum_{\substack{j=1 \\ j \neq k}}^{K_c} \kappa_{ci}^{\downarrow} \xi_{ci}^{\downarrow} v_{c,j} + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \kappa_{ji}^{\downarrow} \xi_{ji}^{\downarrow} p_j^{\downarrow} + \eta_i^{\downarrow} + \breve{\imath}_i^{\downarrow} \tag{5.31}$$

where $p_j^{\downarrow}$ is the total transmit power of cell j, $v_{c,k}$ is the total transmit power required for scrambling code k of cell c, and K_c is the total number of utilized scrambling codes of cell c. The DL effective interference raise $\kappa_{ci}^{\downarrow}$ is in contrast to the UL definition in (5.17) always greater than zero, because the DL channels of serving cell c – apart from the channel between c and i – are not power controlled with respect to the link between cell c and user i. Hence the effective interference raise is also applicable for the interference from the own cell.

For the average required transmit power for user i follows:

$$t_i^{\downarrow} = \frac{r_{ic}^{\downarrow}}{\xi_{ci}^{\downarrow}} = \epsilon_i^{\downarrow}\gamma_i^{\downarrow}\frac{i_i^{\downarrow}}{\xi_{ci}^{\downarrow}} \tag{5.32}$$

$$= \alpha_i\epsilon_i^{\downarrow}\gamma_i^{\downarrow}\kappa_{ci}^{\downarrow}v_{c,k} + \sum_{\substack{j=1\\ j\neq k}}^{K_c}\epsilon_i^{\downarrow}\gamma_i^{\downarrow}\kappa_{ci}^{\downarrow}v_{c,j} + \sum_{\substack{j\in\mathcal{C}\\ j\neq c}}\epsilon_i^{\downarrow}\gamma_i^{\downarrow}\kappa_{ji}^{\downarrow}\frac{\xi_{ji}^{\downarrow}}{\xi_{ci}^{\downarrow}}p_j^{\downarrow} + \frac{\epsilon_i^{\downarrow}\gamma_i^{\downarrow}}{\xi_{ci}^{\downarrow}}\left(\eta_i^{\downarrow} + \breve{\imath}_i^{\downarrow}\right) \tag{5.33}$$

The total transmitted power on scrambling code k of cell c is the sum of all common channel powers allocated to code k and all power levels required for the users allocated to k:

$$v_{c,k} = \tau_{c,k} + \sum_{i\in\mathcal{M}_{c,k}} t_i^{\downarrow} \tag{5.34}$$

$$= \tau_{c,k} + \sum_{j=1}^{K_c} S_{kj}^c v_{c,j} + \sum_{\substack{j\in\mathcal{C}\\ j\neq c}} A_{kj}^c p_j^{\downarrow} + b_k^c \tag{5.35}$$

where

$$A_{kj}^c = \begin{cases} 0 & \text{for } j = c \\ \sum\limits_{i\in\mathcal{M}_{c,k}} \epsilon_i^{\downarrow}\gamma_i^{\downarrow}\kappa_{ji}^{\downarrow}\frac{\xi_{ji}^{\downarrow}}{\xi_{ci}^{\downarrow}} & \text{for } j \neq c \end{cases} \tag{5.36}$$

$$b_k^c = \sum_{i\in\mathcal{M}_{c,k}} \frac{\epsilon_i^{\downarrow}\gamma_i^{\downarrow}}{\xi_{ci}^{\downarrow}}\left(\eta_i^{\downarrow} + \breve{\imath}_i^{\downarrow}\right) \tag{5.37}$$

$$S_{kj}^c = \begin{cases} \sum\limits_{i\in\mathcal{M}_{c,k}} \alpha_i\epsilon_i^{\downarrow}\gamma_i^{\downarrow}\kappa_{ci}^{\downarrow} & \text{for } k = j \\ \sum\limits_{i\in\mathcal{M}_{c,k}} \epsilon_i^{\downarrow}\gamma_i^{\downarrow}\kappa_{ci}^{\downarrow} & \text{for } k \neq j \end{cases} \tag{5.38}$$

The matrix elements S_{kj}^c form the *code allocation matrix* S_c. A_{kj}^c are the elements of the matrix A_c. With (5.35) the vector of code powers $\vec{v}_c$ can be expressed as follows:

$$\vec{v}_c = S_c\vec{v}_c + \vec{\tau}_c + \vec{b}_c + A_c\vec{p} \tag{5.39}$$

$$= (I - S_c)^{-1}\cdot\left(\vec{\tau}_c + \vec{b}_c + A_c\vec{p}\right) \tag{5.40}$$

where $\vec{\tau}_c$ is the vector of common channel powers for each scrambling code of cell c.

The total transmitted cell power is the sum of all code power levels:

$$p_c^{\downarrow} = (1\,..\,1)\cdot(I - S_c)^{-1}\cdot\left(\vec{\tau}_c + \vec{b}_c + A_c\vec{p}\right)$$
$$= \underbrace{(1\,..\,1)\cdot(I - S_c)^{-1}\cdot A_c}_{T_c^{\downarrow}}\vec{p} + \underbrace{(1\,..\,1)\cdot(I - S_c)^{-1}\cdot\left(\vec{\tau}_c + \vec{b}_c\right)}_{h_c^{\downarrow}} \tag{5.41}$$

where $T_c^{\downarrow}$ is row c of the characteristic DL traffic matrix $T^{\downarrow}$ and $h_c^{\downarrow}$ are the individual elements of the DL inhomogeneity vector $\vec{h}^{\downarrow}$. Accordingly also for the DL an equation similar to (5.25) can be obtained:

$$\vec{p} = T^{\downarrow} \cdot \vec{p} + \vec{h}^{\downarrow} \quad (5.42)$$

$$= (I - T^{\downarrow})^{-1} \cdot \vec{h}^{\downarrow} \quad (5.43)$$

Similar to $\vec{h}^{\uparrow}$, also the DL inhomogeneity $\vec{h}^{\downarrow}$ is a valuable means for improving the overall system capacity. In contrast to $\vec{h}^{\uparrow}$, which can hardly be influenced by network planning, the control of $\vec{h}^{\downarrow}$ is subject to the network design process. A large share of $\vec{h}^{\downarrow}$ are the common channel power levels, which can be reduced if a sufficient signal level is present in the whole cell area. As a consequence significantly lower cell power levels and hence a higher overall network capacity can be achieved.

5.5.3 Soft Handover

A user mapped on a DCH can be served by one or more cells (soft/softer handover), if the according pilot powers are within a certain margin. In this case the UL and DL CIR targets $\gamma^{\uparrow}$ and $\gamma^{\downarrow}$ are modified by a diversity gain factor (*SHO gain*, cf. Section 3.5.3). The amount of reduction depends on several parameters like the number of serving cells, and the difference between link gains to the serving cells. For the calculation of cell powers and cell interferences in the above equations the modified parameters are used. Additionally in DL the expression for the connection power is not only to be added to the primary serving cell but also to all other serving cells.

5.5.4 Iterative Solution

The classical method to evaluate the power levels in the system is an iterative solution, where in each iteration successively all connection powers are adapted. However, (5.25) and (5.42) allow a much more efficient calculation on cell basis.

There are two widely used first order methods for the iterative solution of a system of linear equations: Gauss-Seidel and Point Jacobi [VW00, GvL96].

The following system of linear equations shall be solved iteratively:

$$\vec{x} = C\vec{x} + h \quad (5.44)$$

With the *Point Jacobi* algorithm the value of $\vec{x}$ in iteration $m+1$, $\vec{x}^{(m+1)}$, is calculated from the value obtained in iteration m, $\vec{x}^{(m)}$, as follows:

$$\vec{x}^{(m+1)} = D^{-1}(E+F)\vec{x}^{(m)} + D^{-1}h \quad (5.45)$$

where D is the diagonal matrix of C, and E, F are the strictly lower and strictly upper triangular matrices of C.

Gauss-Seidel usually has a slightly higher rate of convergence:

$$\vec{x}^{(m+1)} = (D-E)^{-1}F\vec{x}^{(m)} + (D-E)^{-1}h \tag{5.46}$$

The proposed algorithm on cell basis has two strong advantages over the calculation on connection basis. Firstly, the rate of convergence is much higher, as the powers of all users in one cell are calculated implicitly in one step. Secondly, the computational effort in each iteration step is much lower for the calculation on cell basis. Disregarding the limits on connection basis, the individual connection powers are not of interest for the calculations, so that in each iteration only one equation per cell instead of one per mobile has to be solved. Even if connection limits are taken into account, only few "critical" connections have to be evaluated, for details see Section 5.6.

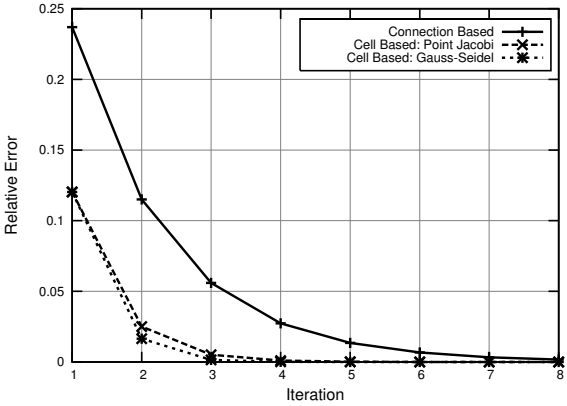


Figure 5.4: Rate of convergence for different iterative load calculation algorithms

The reduction in number of required iterations in order to yield a certain level of convergence is in Figure 5.4 illustrated for an example scenario. The scenario consists of 19 omni-directional cells. Only one traffic type (voice) was considered. The figure compares the average error of calculation the UL cell load versus the number of iterations for the three different algorithms: conventional connection based calculation, cell based Point Jacobi, and cell based Gauss Seidel. The rate of convergence for both methods on cell basis is significantly better. This is also reflected in Figure 5.5, in which the required number of iterations to yield a relative error of less than 1% is compared for varying traffic load. As cell based algorithm Gauss Seidel was chosen. The non monotonic characteristic of the curves is due to the fact, that the algorithm apart from the calculation of the pure power levels also includes the resource scheduling as described in the next section. The resource scheduling imposes non-linearities as it, depending on the traffic load, introduces changes to the system of linear equations from one step to the other (downgrading or blocking of connections).

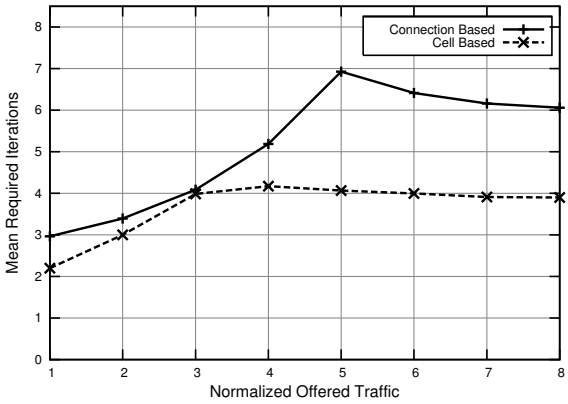


Figure 5.5: Typical number of required iterations per snapshot to yield a relative error of less than 1%

5.6 Resource Scheduling

In the previous section an efficient algorithm for a performance estimation of a WCDMA network given a resource scheduling was derived. Resource scheduling in particular means to select the appropriate bearer and channel type for a certain user. For CS traffic, e.g. speech users, this is typically quite easy, because it basically has to be decided if the user can be served or needs to be rejected. In case of multi-rate CS services, e.g. *Adaptive Multi–Rate* (AMR), and in particularly in case of PS services, multiple optional bearers are available that can be assigned to the user.

In general it is assumed that a user is tried to be served with the highest possible data rate. A downgrading to a lower data rate bearer might be necessary because of link constraints, in particular maximum link power restrictions or constraints at cell level (cf. Section 5.4). In the first case, there is no choice to be taken by the RRM. In contrast, if the data rate reduction is due to restrictions at cell level, e.g. maximum cell load levels, the RRM has potentially different possibilities to schedule resources to users, i.e. it can decide on certain actions among different options. For instance, if the cell load is too high, users could be downgraded based on service type, current bearer type, user category, etc. The particular RRM schemes are not standardized, vendor specific, and might change with software releases. Also, there is typically a large amount of possibilities to configure the radio resource management algorithms. In order to assure flexibility of the analysis algorithm and to make it adaptable to different RRM implementations and parameterizations, a flexible QoS scheme based on priorities and service type specific maximum load levels is introduced in the following. It is believed that with this scheme, a wide range of RRM mechanisms that can be found in live networks can be modeled well.

5.6.1 QoS Concept

The system is modeled to have mainly two mechanisms to control QoS. Firstly, cell resources are scheduled based on priority levels. Secondly, it is assumed that the system has the possibility to restrict the fraction of cell capacity available for certain QoS classes. Both mechanisms are performed locally at cell level.

Bearer Priorities

Resource scheduling policies are equipment vendor (and operator) specific. In order to allow a flexible parameterization, the following scheme based on priorities has been invented by the author. As described above, each service type in principle has a number of possible bearer mappings. For simulation purposes, a priority is associated with each of these bearer mappings. As an input to the simulation, the priorities have to be set adequately such that the "real system" is modeled as accurate as possible.

Bearer Combination	DL Bearer	UL Bearer	Priority
1	DCH 12.2	DCH 12.2	4
2	(unserved)	(unserved)	0

Table 5.1: Bearer table for a speech service

Table 5.1 and Table 5.2 show example bearer mappings for a CS and a PS service respectively. In order to allow a consistent representation, also the blocking state is considered as a bearer assignment (bearer combinations 2 and 5 for the CS and the PS service, respectively).

Resource scheduling is applied at cell level. Load reduction is carried out based on the priority levels: If a set of possible options to downgrade users exists, first the user is selected whose currently assigned bearer combination has the *highest priority* value. In this example first PS users with assigned bearer combination 1 are downgraded to bearer combination 2. After all PS users have been downgraded to this bearer combination, they are successively downgraded to the next lower bearer combination. As last of all possible actions, CS users are downgraded which in this case is equivalent to a blocking of the user. If there are several options with the same priority levels, either a random selection or a selection based on some other policy, e.g. based on the path loss (cf. Section 5.6.3) is carried out.

In earlier work [TPL+03] a slightly different scheme was proposed. It was suggested to to associate a "merit" with each of the bearers. The goal of the scheduling algorithm was to maximize the total merit for each of the cells individually. However, the scheme presented above, although simpler and easier to implement, matches the behavior of "real world" RRM mechanisms more closely so that the original concept has been completely replaced by the new priority level based concept.

Bearer Combination	DL Bearer	UL Bearer	Priority
1	DCH 384	DCH 128	30
2	DCH 128	DCH 128	20
3	DCH 64	DCH 64	10
4	DCH 32	DCH 32	5
5	(unserved)	(unserved)	0

Table 5.2: Bearer table for a PS data service

Load Restrictions

The proposed analysis algorithm provides an additional QoS mechanism. Services are allocated to QoS classes, the cell capacity available for the individual QoS classes can be restricted. The maximum load associated with QoS class n is denoted as l^{QoS_n}. The restriction can be formulated as follows:

$$\sum_{i=1}^{n} l^{QoS_i} < l_{max}^{QoS_n} \tag{5.47}$$

If all cell load levels are known, l^{QoS_i} can in DL easily be obtained if (5.42) is evaluated by summing up the connections of QoS class i. Accordingly, also the UL load caused by QoS class i can be calculated from (5.25).

5.6.2 Analysis Algorithm

The objective of the analysis algorithm is the scheduling of resources based on the QoS concept described above. As side-conditions, the algorithm takes into account the system constraints and limitations described in Section 5.4. Each user can be assigned a bearer combination from a discrete set of possibilities. A bearer combination is in the following considered as one possible state the user can be in. The possible states as well as the transitions between them are for an efficient processing organized in *state transition diagrams*. State transition diagrams for the example CS and PS services introduced in Section 5.6.1 are presented in Figure 5.6 and Figure 5.7. The edges between states describe which state to change to if a certain condition for a state change is fulfilled.

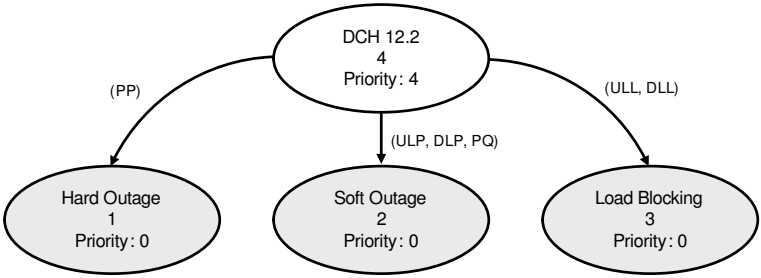


Figure 5.6: CS state diagram

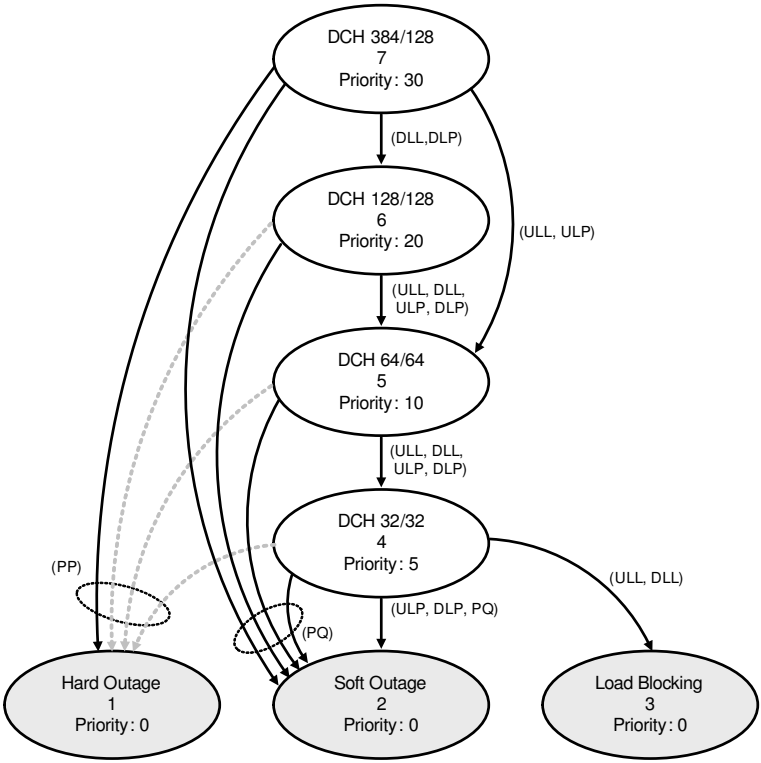


Figure 5.7: PS state diagram

In accordance with Section 5.4 the following reasons for state transitions are defined:

PP – Received **P**ilot **P**ower level below threshold
PQ – Received **P**ilot **Q**uality level below threshold
ULL – Downgrading due to too high **UL L**oad
DLL – Downgrading due to too high **DL L**oad
ULP – Required **UL P**ower too high
DLP – Required **DL P**ower too high

Code limitations are not considered as it is assumed that sufficient code trees are available. Also hardware limits are not considered here. However, the handling for these limitations is almost the same as for the UL and DL power constraints.

It is in general distinguished between *hard outage*, *soft outage*, and *load blocking*. *Hard outage* is due to insufficient pilot power (PP). In this case the user cannot be served, independent of the bearer allocation. In contrast, *soft outage* depends on either the UL or DL load situation. Reasons for soft outage are insufficient link powers (ULP, DLP) or a too low pilot quality (PQ). The events (ULP) and (DLP) do not necessarily cause soft outage as there might be a lower data rate bearer combination that requires a lower output power level. If users cannot be served due to (ULL) or (DLL), the user is denoted to be in *load blocking* state. A reduction of load levels (ULL, DLL) does not directly require users to be blocked/dropped. Typically, first users can be downgraded to bearer combinations which are less demanding in terms of radio resources.

The state transition diagrams provide a valuable guidance for the analysis algorithm. For an explanation the PS state diagram in Figure 5.7 shall be considered. The initial state is state 7. If a state transition is to be carried out, the transition to perform depends on the reason for the state transition. If the transition is due to insufficient DL link power (DLP) or too high DL load (DLL), a transition to the state with next lower priority level (State 6) is executed. If in contrast the UL power (ULP) or UL load (ULL) need to be reduced, a transition to state 5 is conducted. State 6 would be skipped because it has the same assigned UL bearer (DCH 128) and thus the same UL power would be required. If the received pilot level (PP) or pilot quality is too low (PQ), the user cannot be served – independent of the bearer allocation. Hence a direct transition to the corresponding outage state is performed.

A global overview on the snapshot evaluation algorithm is given in Figure 5.8. At first, users are allocated to their best serving cells. At the beginning it is tried to serve all users with their full desired bandwidth (highest priority level). Each connection is checked for a possible hard outage (PP). Although at this stage UL and DL cell load levels are not known, already quality related link restrictions can be checked for (PQ, ULP, DLP) assuming minimum UL and DL load levels in the cells. Based on the current state allocations in the subsequent steps the UL and DL system matrices (cf. Equations (5.42) and (5.25)) are generated and solved. With the calculated cell load levels the list $L_{overload}$ of overloaded cells is generated.

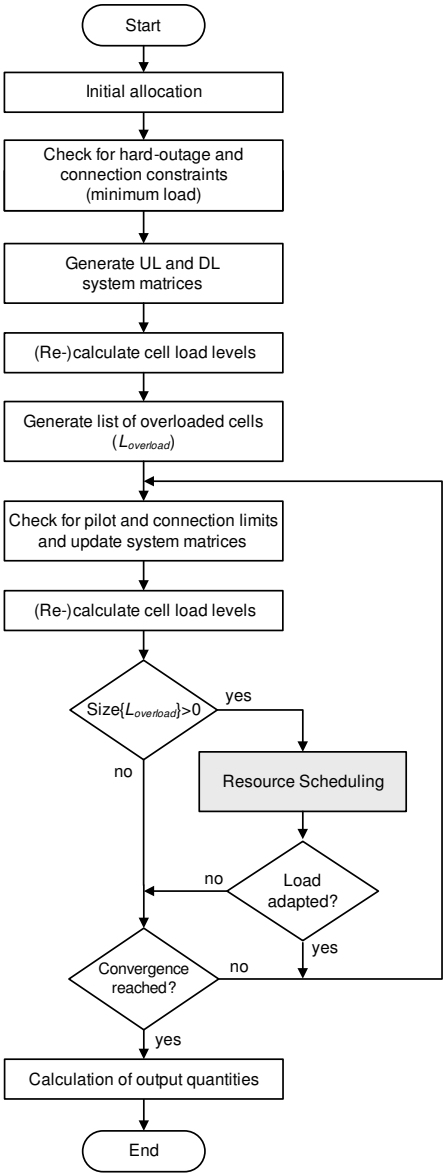


Figure 5.8: Snapshot evaluation algorithm (gray shaded block comprises randomness)

After these initializing steps, the actual evaluation loop is carried out. Based on the previously calculated load levels all users are checked for pilot outage and connection/link restrictions. It is important to note, that not all users have to be considered for this. For maximum connection power constraints only users with high path loss (i.e. at the cell border), and for minimum power constraints only users with rather low path loss (in direction of the cell origin) have to be considered. An efficient search procedure assures that only few individual connection powers have to be calculated during an iteration. The fraction of users to be checked is for typical scenarios rarely larger than 10%. For users that exceed the maximum connection power constraint or for users with too low pilot quality, the corresponding state transition is performed, and the system matrices are updated. Users with power levels lower than the minimum power in UL cause additional constant interference. This is taken account of in the way that their portion is taken off the traffic matrix in (5.25) and the constant amount according to the minimum power is added to the inhomogeneity $\vec{h}^{\uparrow}$ in the same equation. No state change is required. For the updated system equations cell load levels are recalculated. If the size of the list of overloaded cells is at least one, in the next step a load adaptation needs to be carried out. This is processed by the resource scheduling module, which is described further below. If either no load adaptation was required in one iteration of the loop, or if there was no overloaded cell at all, the convergence of the system variables is checked. Suitable variables for tracking the convergence are the cell load levels. Once a sufficiently low relative error between subsequent iterations is observed, the main loop is terminated. At this stage a resource scheduling fulfilling all side-conditions has been determined. Also, cell load levels are already available and up-to-date. In a final step various quantities of interest are calculated. These include statistics at connection and cell level as well the evaluation of quantities over the network area (per pixel results).

Resource Scheduling

The scheduling of resources in case of overload is an integral part of the snapshot evaluation algorithm. The developed algorithm is presented in Figure 5.9. Scheduling of resources is carried out locally on cell basis. Consequently for the resource scheduling only users need to be considered that are connected to one or more cells from the list of overloaded cells $L_{overload}$. The resource scheduling is carried out iteratively. In each iteration, at first a cell c from $L_{overload}$ is chosen. In order not to favor certain cells, the cell is selected randomly. The probability of each cell to be selected is proportional to the sum of the mean call arrival rates of all users in the cell. The latter takes account of the fact that load actions are in a real system triggered by certain events, here assumed at call arrivals, which take place more frequently in cells with more users. Once the cell has been determined, it is checked if this cell is currently in overload. Note, that $L_{overload}$ is not modified during the execution of the algorithm, i.e. there might be cells selected for which the overload has already been resolved. If the cell is in overload, one *load action*, i.e. the downgrading of one user, is performed. The user is selected based on the priority

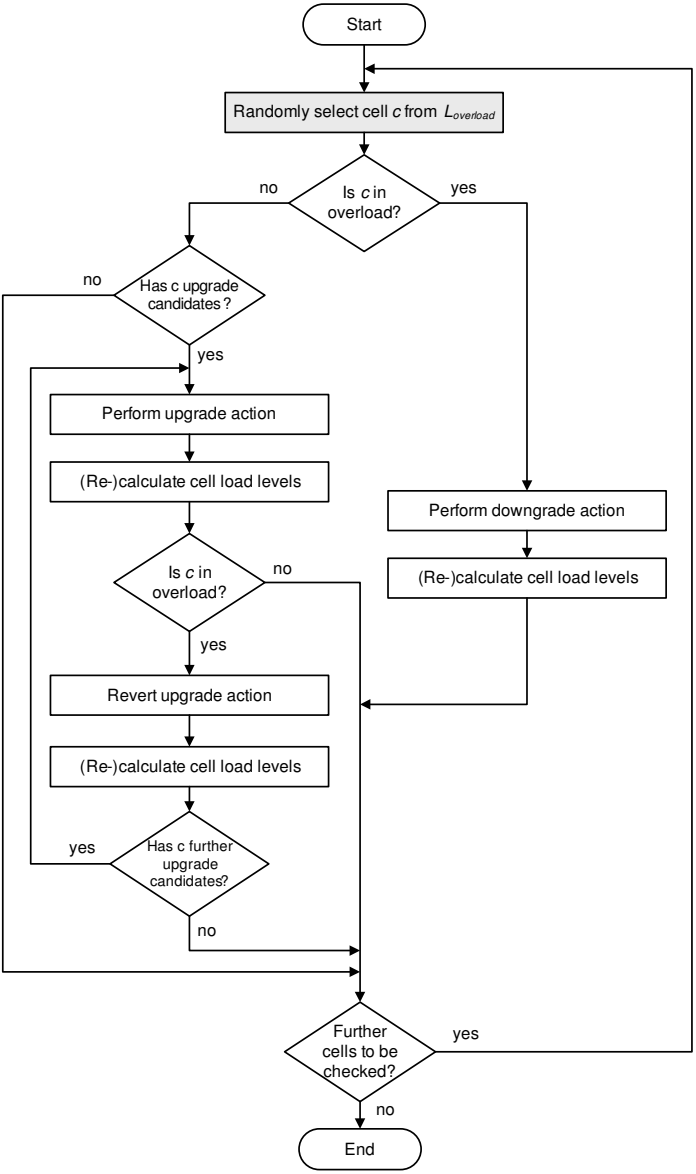


Figure 5.9: Resource scheduling algorithm (gray shaded block comprises randomness)

levels as described in Section 5.6.1. The system matrices are updated for the new resource scheduling, and the load levels are recalculated. If in contrast the cell c is not in overload, it is checked whether there are possibilities to upgrade users in this cell to higher priority bearers. This is again carried out based on the priority levels. From the possible states of all users that are higher than the priority level of the state currently allocated to the respective users, at first the one with lowest priority level is chosen for a potential *upgrade transition*. For the new allocation the cell load levels of cell c are calculated. If there is no violation of the maximum cell load targets of cell c, the new resource allocation is accepted and the resource scheduling algorithm continues. If the upgrade results in an overload, the upgrade transition with next higher priority is chosen. This is continued until all upgrade candidates have been checked. As only a recalculation of the load in the current cell is required, this process is very fast. The resource scheduling loop is terminated, once there were no state transitions in consecutive iterations in that all cells from $L_{overload}$ were checked at least once.

The process of upgrading users requires special attention. It has been described that it does not strictly follow the policy of priority levels. If an upgrade to a state with next higher priority level (with reference to the currently highest allocated priority level in the cell) is not possible due to a violation of load constraints, the next higher is checked and so forth. This procedure is applied out because the priority level does not necessarily reflect the demand in terms of resources, and the network always tries to fully exploit its available resources. Unfortunately, in the iterative procedure of the resource scheduling algorithm a problematic situation might arise. In order to illustrate the problem, it shall be assumed that in one iteration of cell c an upgrade to a state s would in principle be possible following the proposed QoS scheme. Assuming furthermore that in a previous iteration this upgrade was not possible due to the instantaneous load situation, instead other upgrades of higher priority were chosen. The additional load by these upgrades, however, might make it impossible to add the upgrade to s to the current scheduling. This effect can only be avoided by a rather complex rescheduling of resources prior to each upgrading cycle. Such an algorithm has optionally been integrated into the snapshot analysis, but is not presented in this slightly simplified description of the algorithm. Nevertheless, although not fully accurate, the error which is made by not applying the rescheduling is very low. Hence it is also not considered in practical operations of the developed snapshot evaluation module.

5.6.3 Application Example

In the following an example application of the concept described above is given. Considered is a 19 site hexagonal scenario with three sectors per site, such that the network consists of 57 cells (Figure 5.10). For all investigations the inner 3 cells of the network are evaluated. All other cells are fully simulated but are considered as border cells that contribute to the interference only.

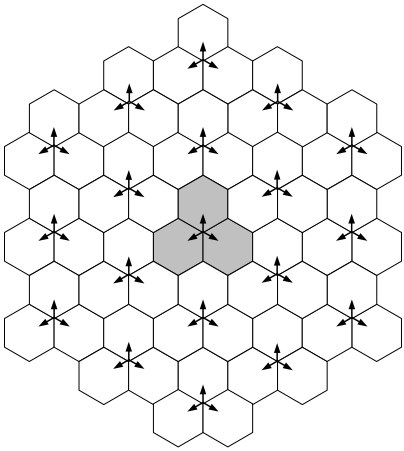


Figure 5.10: Considered hexagonal scenario (gray shaded cells are evaluated)

One CS (voice) and three different PS services are considered. The priority level for the CS service is very low (1), such that it is always served and only contributes as background load. Each of the PS services can in DL be allocated with three different bearers. The PS services are asymmetric, such that the UL load is very low. It can thus be focused on the DL. The priority levels chosen for the PS services are shown in Table 5.3. The priorities are set up in such a way that the "High Priority" service is always slightly prioritized above the "Low Priority" service. Only resources that are not needed by the latter two services are scheduled to the "Best Effort" service.

Bearer #	DL Bit Rate [kbps]	$\gamma^{\downarrow}$ [dB]	$\epsilon^{\downarrow}$	Priority Level		
				High Priority	Low Priority	Best Effort
1	384	-9	1	8	9	25
2	128	-14.3	1	6	7	15
3	64	-16.8	1	4	5	10

Table 5.3: Cost allocation table for the three used PS services

Two policies for the selection among different actions of equal priority are compared: "Random" and "Path Loss Based". In case of *Random Downgrading* (RDG), downgrade actions are randomly chosen if the scheduler has different options with the same priority level. With the *Path loss Based Downgrading* (PBDG) policy first actions with higher path loss to the serving cell are chosen, thus users with good radio conditions are favored. The latter policy can be easily integrated into the concept described above by adding the inverse of one plus the serving cell's link gain ($1/(1+\xi_{ci}^{\downarrow})$) to the priority level of all states of the users i. As the basic priority levels in Table 5.3 are integer numbers and the additional contribution is strictly lower than one, the basic order of states is not changed.

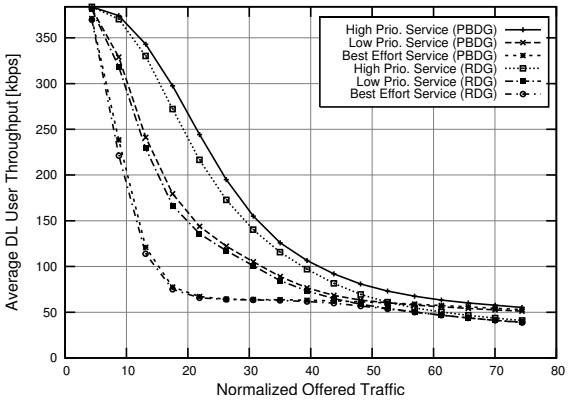


Figure 5.11: Average DL user throughput for different services and resource allocation schemes

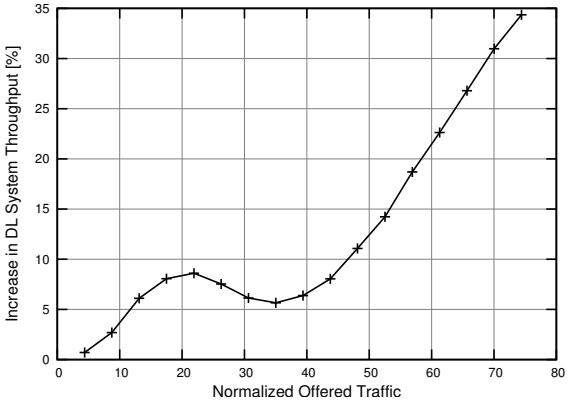


Figure 5.12: Relative increase in capacity applying a path loss based downgrading (PBDG) compared with random downgrading (RDG)

Figure 5.11 depicts the mean available data rate for the three PS services depending on the offered traffic. The traffic is scaled relative to a reference traffic mix for which no downgrading occurs. The plot shows the intended prioritization among the three services. The plot illustrates that an appropriate setup of priorities allows a very flexible modeling of different QoS schemes. In Figure 5.11 also the mean data rates obtained for the two different downgrading policies are compared. As expected the mean data rate is higher for PBDG. The overall increase in cell throughput is presented in Figure 5.12. The gain strongly depends on the load situation. For a high load, gains in capacity of 25% and above can be achieved.

5.7 Method Extensions

Focus in this thesis is on the analysis of one distinct Release 99 WCDMA cell layer, which is still the dominating configuration in UMTS networks. Traffic is carried by dedicated channels in both UL and DL direction. However, the general concept presented in this chapter can easily be extended with respect to new system features like HSPA and multiple cell layers operating on different frequencies.

5.7.1 Multi Carrier Analysis

If the network consists of multiple cell layers that are operated in different frequency bands, denoted as *carriers*, traffic needs to be scheduled not only within one cell layer but is also distributed among the cell layers. The concept of states and state transitions can be extended to consider also the allocation to bearers on different carriers.

If only one carrier is available, the serving cell – or cells in case of SHO – are determined by the pilot strength only. There is basically one possible allocation to cells for each user. If several carriers are available, different options for allocating users to cells exist. At the start of the snapshot evaluation, a best server analysis is carried out on all carriers, and all feasible serving carriers are determined taking into account the received signal strength levels on all carriers and the capabilities of the user equipment. Once the possible servers have been obtained, an extended set of states is generated for each of the users.

The set comprises distinct states for the possible serving bearer combinations on all carriers. By an adequate selection of priority levels, also preferential cell layers can be accommodated into the concept. If, for example, a two carrier WCDMA network is considered, all states from the one carrier state diagram are duplicated: one state for carrier A and an equivalent state for carrier B. If carrier A is the preferred carrier, the states corresponding to this carrier can be given a higher priority level. A combined resource scheduling based on the set of states and state transitions is applied. The actual performance evaluation for each of the carriers is carried out independently according to the WCDMA performance analysis described in Section 5.5. Inter-frequency interference can optionally be taken into account.

Also the analysis of different technologies can be incorporated. If for example the combined analysis of GSM and UMTS networks is to be carried out, only the WCDMA performance analysis needs to be replaced by a GSM performance analysis for the respective carriers. The concept of resource allocation states can be kept. Of course the states are different for the different technologies, as the available bearers are generally technology dependent.

If the network supports load balancing between different carriers, an additional load scheduling algorithm is required. This successively reschedules users from higher loaded to lower loaded carriers. The implementation depends on the specific load balancing algo-

rithm. A generic approach is possible by extending the state diagrams with additional state transitions for the load balancing algorithm. It is in this case proposed to use additional priorities associated with each state to support the guidance of the algorithm.

5.7.2 HSPA Analysis

HSDPA imposes new requirements on the snapshot analysis. With HSDPA a time multiplexed channel with *Adaptive Modulation and Coding* (AMC) and fast scheduling among users is introduced (cf. Section 2.2.4). User data rates depend on the location of the individual users in the cell as well as the current load and interference situation. In [TKSG04] an extension of the snapshot analysis approach presented in this chapter towards HSDPA is presented. The basic idea is to carry out the HSDPA analysis in two stages. First, a "macroscopic" analysis of HSDPA, integrated into the snapshot analysis, is carried out. From this already maximum and average data rates per user as well as the load and interference caused by the HSDPA channels can be estimated.

In the snapshot analysis fast fading is not explicitly considered. Rather the influence of fast fading is taken into account in terms of average values (cf. Section 3.5 and Section 5.3). Due to the fast scheduling of resources, a detailed HSDPA analysis requires the consideration of link dynamics, in particular fast fading. Hence, a second optional analysis stage is introduced focusing on the scheduling. The analysis module takes the power and interference situation calculated by the snapshot analysis as input and simulates individual HSDPA channels in a dynamic way. From this analysis additional important output quantities, e.g. packet delays, can be extracted. For a more detailed description of the HSDPA extension refer to [TKSG04].

HSUPA has strong analogies with HSDPA. Also in HSUPA, the data transmission is scheduled among users and individual data rates are adapted to the current link conditions. HSUPA can thus be considered similarly to HSDPA.

5.8 Application of Variance Reduction Techniques

Although variance reduction techniques haven proven their applicability and value in various applications, almost no research work seems to have been spent on the application of variance reduction techniques in the Monte-Carlo snapshot analysis for network planning purposes. None of the commercial tools known to the author applies variance reduction techniques. Furthermore only one scientific publication on this topic was found [FR05].

In [FR05] *importance sampling* and optimally stratified survey, a variant of *stratified sampling* (cf. Section 4.3.2), are applied for variance reduction of the Monte-Carlo snapshot analysis. As quantity of interest the admission probability per pixel is chosen, which in the study is mainly influenced by the DL load. The variance reduction is quantified for

a realistic planning scenario. The basic idea is to increase the mean of the Poisson distributed traffic considered in the snapshots, and as a consequence to focus on higher load situations. With this approach a reduction in number of required samples of up to 30% was observed by the authors.

A general observation in the basic experiments carried out in Section 4.3 is a great sensitivity of IS to its input parameters. Only a slight variation or error in the parameters can easily lead to a drastic performance degradation. This is one reason for not further considering IS (and StrS) in this work. The other even more important reason is, that IS and StrS, though in different ways, select the set of samples to analyze such that fewer samples for estimating a particular quantity of interest are required. However, in network analysis typically a large amount of quantities are evaluated. It is hence impractical to select the set of snapshots in such a way that an overall reduction in number of required snapshots is achieved. The above mentioned sensitivity to the input parameters gives rise to this.

5.8.1 Control Variates

Control Variates (CV) are introduced in Section 4.3.3. The application of CV is promising, because it does not influence the snapshot generation itself. By applying different control variates for the individual quantities of interest, variance reduction in estimating multiple quantities can be achieved within one Monte-Carlo analysis run. The benefit of CV strongly depends on the choice of the control variate itself. In this section three different control variates for the estimation of cell TX power levels are developed. They slightly differ in their computational complexity. Results are given for an example scenario.

The control variates in this section are based on the analytical method for estimating the mean network performance developed in Chapter 6. As is explained in detail in Chapter 6, the method allows an estimation of the mean cell TX power levels $\vec{p}^*$ based on an average DL traffic matrix $T^{*\downarrow}$ and an average inhomogeneity $\vec{h}^{*\downarrow}$:

$$\vec{p}^* = T^{*\downarrow}\vec{p}^* + \vec{h}^{*\downarrow} \tag{5.48}$$

Similar to (5.42) also this "average" system is iteratively solved to yield the mean cell TX power estimates. If in the iterative process cell TX power levels exceed their maximum levels, they are clipped to the latter. This procedure provides very good estimates for the mean power levels. An application as control variate, however, requires estimates of the cell TX power that are correlated with the cell TX power levels actually calculated in each snapshot, i.e. the realization of the control variate in snapshot k needs to consider the particular traffic realization of that snapshot. The basic idea of the three methods presented below is to scale the matrix $T^{*\downarrow}$ and the inhomogeneity $\vec{h}^{*\downarrow}$ based on the traffic realization in snapshot k. For the resulting modified set of equations applicable for snapshot k the following notation is used:

$$\vec{p}'_k = T'^{\downarrow}_k \vec{p}'_k + \vec{h}'^{\downarrow}_k \tag{5.49}$$

As realization of the control variate C (cf. equation (4.33)) in snapshot k, consequently the cell TX power estimate $p'^{\downarrow}_{c,k}$ is used:

$$C_k = p'^{\downarrow}_{c,k} \tag{5.50}$$

Service	Share	Type	Possible Bearers (UL/DL)
Speech	50%	CS	12.2/12.2
Video Telephony	20%	CS	64/64
File Download	20%	PS	64/384, 64/128, 64/64
E-Mail	10%	PS	64/384, 64/128, 64/64

Table 5.4: Considered traffic mix

The control variates derived in this section are applied for a cell power estimation of the center cells in the 57 cell hexagonal scenario that was also considered in Section 5.6. For the analysis carried out in this section the realistic traffic mix presented in Table 5.4 is considered. It comprises two CS and two PS traffic types. For the PS services the DL bearer is chosen based on the current traffic load from a set of three different bearers. Due to the asymmetry of the PS traffic it is focused on the DL.

Method 1: Own Cell Coefficient Scaling (OCCS)

This method is based on the assumption that the cell TX power is correlated with the current amount of traffic in the cell. As there are in general multiple traffic types with varying bandwidth demand, the own cell coefficient $T^{\downarrow}_{cc,k}$ of the traffic matrix is chosen for a quantification of the traffic amount. This coefficient depends on both the number of users in the cell and their respective quality demand (cf. Section 5.5.2). As row c of the DL traffic matrix is an aggregation of the contributions of all users in cell c, this row is scaled with the fraction of own cell coefficient $T^{\downarrow}_{cc,k}$ and mean coefficient $T^{*\downarrow}_{cc}$. The traffic dependent part of row c of the inhomogeneity is handled equivalently, whereas the static part, i.e. the static cell power τ_c, is kept unscaled. For the individual elements it yields:

$$T'^{\downarrow}_{ij,k} = \begin{cases} \Upsilon_{c,k} T^{*\downarrow}_{ij} & \text{for } i = c \\ T^{*\downarrow}_{ij} & \text{otherwise} \end{cases} \tag{5.51}$$

$$h'^{\downarrow}_{i,k} = \begin{cases} \Upsilon_{c,k} \left(h^{\downarrow}_{c} - \tau_c \right) + \tau_c & \text{for } i = c \\ h^{\downarrow}_{i} & \text{otherwise} \end{cases} \tag{5.52}$$

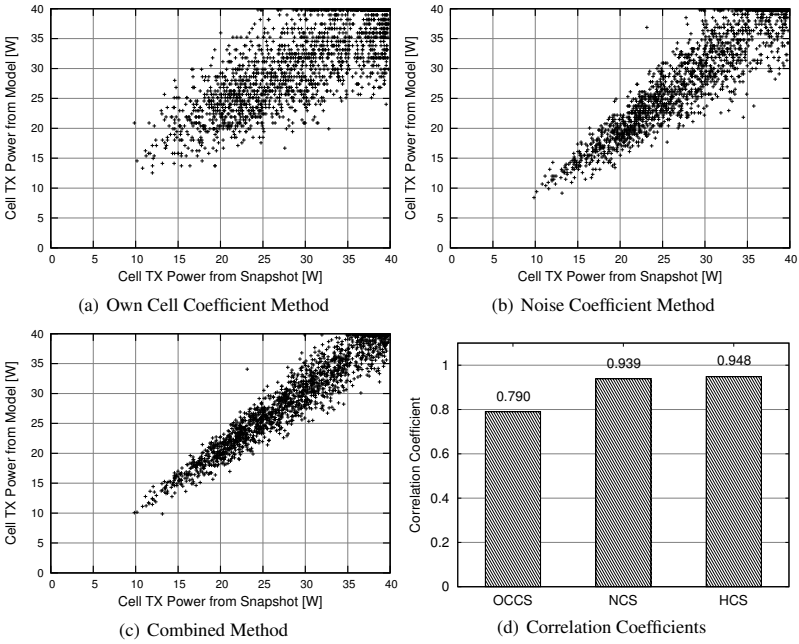
$$\text{where} \quad \Upsilon_{c,k} = \frac{T^{\downarrow}_{cc,k}}{T^{*\downarrow}_{cc}} \tag{5.53}$$

The correlation between the cell TX power approximated by the model and the reference power levels calculated by the snapshot evaluation is presented in Figure 5.13(a). The

traffic was chosen so that the evaluated cells have an average DL traffic load of 60%. In order to quantify the correlation, the *standard correlation coefficient* r is evaluated:

$$r = \frac{\text{cov}\left\{p_c^{\downarrow}, p_c'^{\downarrow}\right\}}{\sqrt{\text{var}\left\{p_c^{\downarrow}\right\} \text{var}\left\{p_c'^{\downarrow}\right\}}} \tag{5.54}$$

The correlation coefficient in this example is 0.79.



(a) Own Cell Coefficient Method (b) Noise Coefficient Method

(c) Combined Method (d) Correlation Coefficients

Figure 5.13: Correlation between control variate and results from snapshot analysis (60% load)

The calculation of the expected value μ_C of C requires a summation over all possible values of C weighted with according probabilities. For this, first the (discrete) distribution of Υ_c needs to be calculated. For each of the possible values the corresponding value of C is calculated. The values are then multiplied with their probability of occurrence and summed-up. In practice this procedure would require a too large amount of memory and computational time, as the discrete probability functions, representing the contribution of the individual pixels, would need to be successively convoluted. These successive convolutions lead to a resulting discrete probability function with a very high number of values. Hence, the distributions of Υ_c and C are approximated. This can either be done analytically or experimentally. Both approaches are sketched in the following.

An analytic approach is to first determine the distribution functions for the contribution of each pixel to the scaling factor Υ_c for cell c. Due to shadow fading, pixels are generally served by a cell only with a certain probability. Hence pixels also only contribute to the scaling parameter with a certain probability. Conditional probabilities need to be evaluated for each pixel. In order to limit the required memory for storing the distributions and in order to speed-up convolutions, the distribution functions are discretized into bins with fixed width w_1. The number of bins does not necessarily need to be the same for all distributions, but should be either even or odd for all distributions, such that a convolution results in an even number of bins.

At first it is assumed that the number of pixels P that have a contribution to the scaling factor is a power of two:

$$P = 2^s \tag{5.55}$$

The convolution is carried out step-wise. In each step distribution functions are convoluted pairwise. The bin width of the resulting distributions is in a consecutive step doubled by joining neighboring bins. The resulting distributions with bin width $w_i = 2 \cdot w_{i-1}$ are taken as input to the next step. After $s = \mathrm{ld}(P)$ steps (and $P - 1$ convolutions) the resulting distribution function with bin width $w_{s+1} = \mathrm{ld}(P) \cdot w_1$ is obtained.

Obviously relation (5.55) is typically not fulfilled. The number of pixels can generally be expressed as:

$$P = 2^s + r \tag{5.56}$$

where strictly $r < 2^s$. In a preprocessing step r pairwise convolutions are carried out to yield a resulting number of distribution functions which equals 2^s. In order not to loose accuracy in the preprocessing step, a higher bin resolution $w_0 = w_1/2$ is used. Hence also after the preprocessing, bins need to be joined pairwise.

The distribution of Υ_c, and as a consequence also μ_C, can also be determined experimentally. For this a Monte-Carlo analysis is carried out. A large number of traffic realizations is drawn basically following the procedure described in Section 5.2. The main difference is, that parameters do not need to be generated for all interfering links, as only the relation to the best server is of interest here. This leads to a significant speed-up. As solving the set of equations (5.49) for each generated Monte-Carlo sample would be too time consuming (in fact of similar complexity as evaluating the snapshot itself), equivalently to the analytical approach first the distribution of the Υ_c is approximated. For each bin of the resulting distribution the system (5.49) is solved, weighted with the probability of the bin, and summed-up.

Both approaches estimate μ_C with sufficient accuracy in reasonable time. Advantage of the first approach is that it is deterministic. On the other hand, it includes a large number of complex operations, i.e. convolutions and the calculation of conditional probabilities. For the evaluation of the conditional probabilities the same look-up table (cf. Section 6.5) can

be used that is anyway required for the analytical method, which serves as basis for the methods presented here. The second approach gains from its simplicity. It is extremely fast and mainly involves standard operations that are also carried out in the snapshot analysis. Experiments have shown that the estimation of μ_C with 10,000 samples for each of the roughly 1,000 cells in a large network with more than 1,000,000 pixels is possible within a total time of 10 sec on a standard *Personal Computer* (PC). A very beneficial property of the Monte-Carlo approach is that the speed vs. accuracy trade-off can be easily biased by changing the number of samples. As a consequence, although both approaches could in principle be used, the second was chosen for the analysis in this section.

Method 2: Noise Coefficient Scaling (NCS)

The first method is based on the own cell coefficient. The own cell coefficient reflects the total amount of traffic in the cell but does not include any information about the spatial distribution of users. On the other hand it is obvious that users located at the cell edge require higher power levels than users near the base station antenna. In the *noise coefficient method* the scaling is thus based on $h_c^{\downarrow}$. The latter incorporates traffic volume and spatial distribution (see equations (5.37) and (5.41)). It is larger when more traffic is present in the cell, but also increases if the traffic has a lower link gain to the serving cell.

Accordingly, the components of the system equation are constructed as follows:

$$T'^{\downarrow}_{ij,k} = \begin{cases} \Gamma_{c,k} T^{*\downarrow}_{ij} & \text{for } i = c \\ T^{*\downarrow}_{ij} & \text{otherwise} \end{cases} \tag{5.57}$$

$$h'^{\downarrow}_{i,k} = \begin{cases} \Gamma_{c,k} \cdot \left(h_c^{\downarrow} - \tau_c\right) + \tau_c & \text{for } i = c \\ h_i^{\downarrow} & \text{otherwise} \end{cases} \tag{5.58}$$

$$\text{where} \quad \Gamma_{c,k} = \frac{h^{\downarrow}_{c,k}}{h^{*\downarrow}_{c}} \tag{5.59}$$

The correlation between model and cell TX power from snapshot analysis is shown as an example in Figure 5.13(b). The apparently better correlation compared with the results from method 1 is also reflected by the correlation coefficient, which for this example is 0.94.

The calculation of μ_C for this method is slightly more complex than in case of the first method. Calculating the distribution function of C also involves consideration of the link gain. However, if the Monte-Carlo approach is chosen, the estimation of μ_C can be carried out without significant difference in performance compared to the estimation carried out for the first method described above.

Method 3: Hybrid Coefficient Scaling (HCS)

This method combines the advantages of the two approaches described above. For scaling of the own cell coefficient, method 1 is used, whereas for scaling the other elements, method 2 is used:

$$T'^{\downarrow}_{ij,k} = \begin{cases} \Gamma_{c,k} T^{*\downarrow}_{ij} & \text{for } i = c, j \neq c \\ \Upsilon_{c,k} T^{*\downarrow}_{cc} & \text{for } i = c, j = c \\ T^{*\downarrow}_{ij} & \text{otherwise} \end{cases} \tag{5.60}$$

$$h'^{\downarrow}_{i,k} = \begin{cases} \Gamma_{c,k} \cdot \left(h^{\downarrow}_{c} - \tau_c \right) + \tau_c & \text{for } i = c \\ h^{\downarrow}_{i} & \text{otherwise} \end{cases} \tag{5.61}$$

The main advantage of this method is, that for $h'^{\downarrow}_{c,k}$ and $T'^{\downarrow}_{cc,k}$ the exact values from the snapshot are used. So in case the cell power is mainly required to compensate the receiver noise as well as in case the interference in the cell is dominated by the intra cell power, very good results can be obtained.

Application Example

The three control variates developed above have been applied to the hexagonal reference scenario. The variance reduction was evaluated for four different load scenarios: 50%, 60%, 70%, and 80% average DL cell load. In Table 5.5 the results are presented. Results are given for a straightforward selection of $\beta = 1.0$ and for an optimum setting according to (4.37). The results obtained for optimum β are are additionally illustrated in Figure 5.14. Considerable variance reductions are achieved. In particular the third method (HCS) provides very high variance reduction coefficients (up to 20). Moreover the difference between variance reduction for optimum β compared with $\beta = 1$ is quite low, which is due to the very high correlation between model and results from snapshots. The variance reduction becomes the lower the higher the cell load. This tendency is caused by the fact that the variance itself becomes lower for high load situations, as the cell TX power in these situations is more likely clipped to its maximum value. In fact, if the offered traffic were further increased, the variance would eventually become zero (as the cell TX power were equal to the maximum value for all snapshots).

Equivalently to the cell TX power case, in Table 5.6 and Figure 5.15 results for the variance reduction in estimating the blocking probability are presented. The blocking probability is for this example defined as the probability that in a snapshot at least one user in the cell of interest is not served at its desired data rate. The according control variate is based on the control variate for the power estimation. Instead of directly taking the power estimate, a value of 1 is chosen if the cell TX power is equal to its maximum value, and 0 otherwise. The variance reduction is lower compared to the estimation of cell TX power, however, still variance reduction coefficients greater than 3 are obtained.

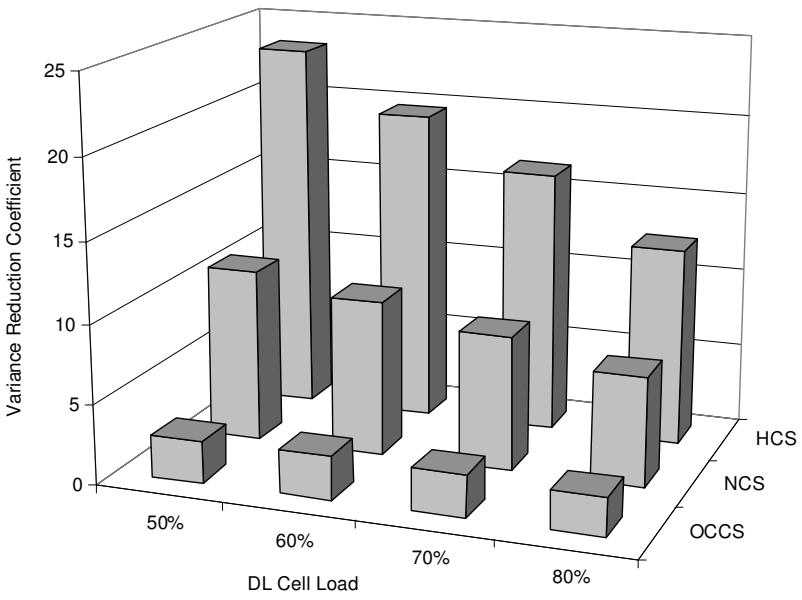


Figure 5.14: Variance reduction in estimating the mean cell TX power achieved by applying control variates (optimum β)

Applied Method	DL cell load 50%		DL cell load 60%		DL cell load 70%		DL cell load 80%	
	β	ψ	β	ψ	β	ψ	β	ψ
OCCS	1	2.60	1	2.67	1	2.58	1	2.36
	0.97	2.60	0.92	2.70	0.88	2.66	0.92	2.39
NCS	1	7.61	1	7.54	1	6.78	1	5.67
	0.83	10.86	0.84	9.79	0.85	8.45	0.84	6.88
HCS	1	23.29	1	19.24	1	15.80	1	12.32
	0.99	23.36	0.97	19.67	0.95	16.53	0.97	12.47

Table 5.5: Variance reduction in estimating the mean cell TX power applying control variates

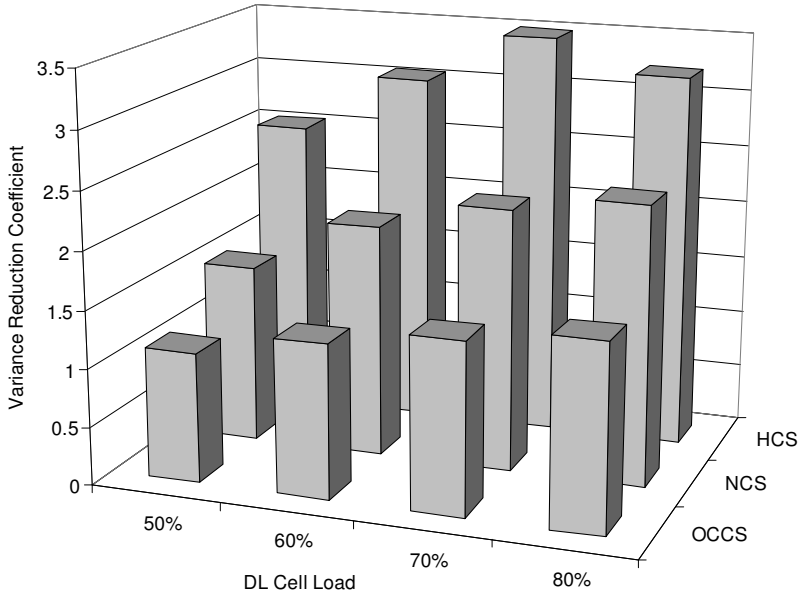


Figure 5.15: Variance reduction in estimating the blocking probability achieved by applying control variates (optimum β)

Applied Method	DL cell load 50%		DL cell load 60%		DL cell load 70%		DL cell load 80%	
	β	ψ	β	ψ	β	ψ	β	ψ
OCCS	1	1.06	1	1.18	1	1.19	1	1.27
	0.59	1.12	0.62	1.33	0.59	1.48	0.60	1.60
NCS	1	0.69	1	1.38	1	1.81	1	2.08
	0.39	1.54	0.60	2.01	0.69	2.26	0.75	2.40
HCS	1	2.44	1	2.86	1	3.16	1	2.90
	0.86	2.54	0.85	3.05	0.83	3.49	0.82	3.22

Table 5.6: Variance reduction in estimating the blocking probability applying control variates

5.8.2 Correlated Sampling

Correlated sampling has been introduced in Section 4.3.4. It is a valuable means for variance reduction if two or more systems need to be compared. It can be very simply implemented, as it only requires to use the same series of random numbers for the analysis of all systems. In this section the benefit of correlated sampling is illustrated for the scenario already considered in the previous section. Although also blocking probability or cell TX power could be chosen, for the analysis in this section the DL cell throughput is selected. The cell throughput is typically put considerable focus on in network planning. In order to trigger a substantial number of data rate changes and consequently a visible influence of the network configuration on the cell throughput, the offered traffic is increased by roughly 40% compared to the 80% DL load case considered in the previous section.

The variance reduction coefficient is evaluated for a comparison between the total network throughput in the original network and various modified network configurations. The network configuration is modified by a change of the pilot power in a sub-set of the inner 21 cells. The results are shown in Figure 5.16 for a varying number of modified cells and a varying difference in pilot power. The exact quantities are in addition presented in Table 5.7. The smaller the changes in the configuration, the higher the achieved variance reduction. If only the pilot power of one cell is reduced by 1 dB, the variance reduction is the greatest: An average variance reduction coefficient of 253.8 is observed. For larger changes in the network configuration, the variance reduction coefficient takes values as low as 2.25, which is obtained when the pilot power is reduced for 15 cells by 9 dB. Typical use case for correlated sampling is the comparison of different configurations in network optimization, e.g. local search techniques. For this, typically the effect of small changes needs to be evaluated, for which according to the analysis carried out in this section, very high variance reduction gains can be expected

Number of Modified Cells	Pilot Power Reduction [dB] 1	2	3	4	5
1	253.82	46.31	38.41	37.05	36.83
3	51.11	14.55	12.43	12.23	11.86
5	35.28	8.88	7.73	7.38	7.06
7	26.21	6.44	5.44	5.23	4.83
9	21.76	5.22	4.91	4.32	3.70
11	18.17	4.70	4.12	3.10	3.10
13	13.89	3.65	3.44	2.82	2.68
15	12.33	3.40	3.00	2.54	2.25

Table 5.7: Variance reduction in estimation of the mean DL network throughput applying correlated sampling

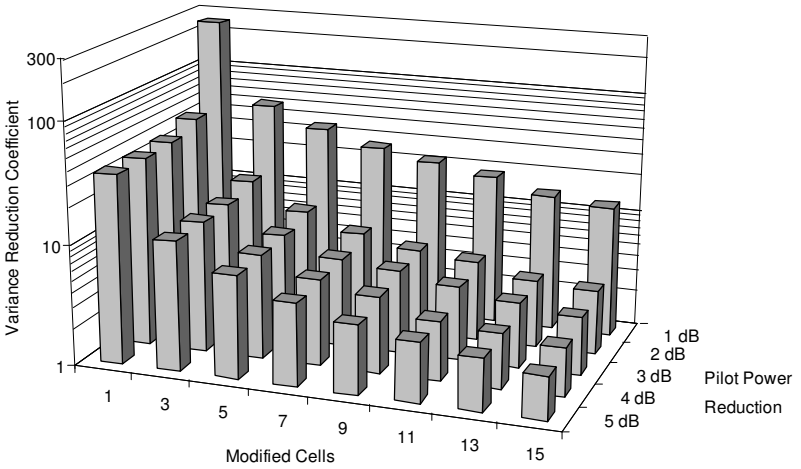


Figure 5.16: Variance reduction in estimation of the mean DL network throughput applying correlated sampling

5.9 Validation of Results

The validation of results is of course of great importance for any kind of simulation approach, in particular if it is applied in operational business. Ideally the results are to be compared with results from the live networks. Such a comparison, however, is very costly in both time and monetary cost and can hardly be taken. Also, it would be desirable to verify the analysis method and not the accuracy of its input parameters, e.g. traffic distributions and path loss values. The latter however is difficult to extract from the live network. Thus, a comparison with other simulations is the only feasible option.

The presented method has been compared against different simulation approaches. First comparisons were carried out within the EU project MOMENTUM. In that project besides the snapshot analysis method, a dynamic system level simulator was developed. Both methods were applied to the public and internal reference scenarios defined within the same project. Although, due to some small limitations of the developed dynamic simulator, no detailed quantitative comparison was possible, the results showed a good qualitative agreement [GLL+03].

Further studies have been carried out internally. The method was compared with a commercial snapshot simulator and with the results obtained from a dynamic system level simulator. The study is confidential and exact results cannot be presented here. However, it can be stated that comparable results with a relative error in cell capacity of less than 10% between results from dynamic simulations and from the snapshot analysis method developed in this thesis were observed. It needs to be stressed, that this requires to set the

input parameters accurately. In particular the findings and derivations in Section 3.5 with respect to the influence of fast fading shall be recalled.

5.9.1 Limitations

The snapshot approach has natural inherent limitations with respect to traffic handling, which are due to the consideration of independent snapshots that do not include any information about the recent history. These limitations with respect to blocking probabilities and the modeling of PS traffic are discussed in the following.

Notes on Blocking Probabilities

Real blocking probabilities, e.g. in terms of Erlang B blocking, are difficult to obtain from snapshot simulations. The reason is that a snapshot comprises the offered traffic in a point in time including traffic which might have been blocked in the previous history. The snapshot does not comprise any information about previously blocked traffic.

In order to illustrate this effect, the blocking probability in an M/M/N queuing system, e.g. [TG96], shall be estimated by means of a snapshot analysis. The blocking probability p_p of an M/M/N system is given by the well known Erlang B formula:

$$p_b = \frac{A^N}{N! \sum_{i=0}^{N} \frac{A^i}{i!}} \tag{5.62}$$

where A is the average amount of offered traffic.

If the system is to be analyzed by a snapshot approach similar to the one described in this chapter, the offered traffic in a snapshot is the number of users drawn from a Poisson distributed random variable X. The probability of an offered traffic of i consequently is:

$$\mathrm{P}\{X = i\} = \frac{\mathrm{e}^{-A} A^i}{i!} \tag{5.63}$$

A straightforward definition of a blocking probability would be the probability, that the offered amount of traffic is equal or greater than the system capacity N:

$$p_H = \sum_{i=N}^{\infty} \frac{\mathrm{e}^{-A} A^i}{i!} \tag{5.64}$$

Another definition of the blocking probability is the expected amount of users that can not be served in one snapshot:

$$p_L = \sum_{i=N+1}^{\infty} (i - N) \cdot \frac{\mathrm{e}^{-A} A^i}{i!} \tag{5.65}$$

Obviously both estimates do not match the Erlang B blocking probability, which is due to the fact, that the traffic in the snapshots is overestimated because previous blocking can not be considered. In fact, it can be quite easily shown that $p_H > p_b$. For low blocking probabilities p_L is a lower bound for p_b.

The error made in estimating blocking probabilities by snapshot simulations of course needs to be considered. It also impacts the estimation of the other system parameters, e.g. cell load levels. However, firstly for low blocking rates, which are the typical working points for mobile networks, the error is rather low. Secondly, given the inaccuracy of input variables in the planning process, like path loss and traffic predictions, mostly relative comparisons of network configurations and different parameter settings are required. Hence the absolute values for blocking probabilities are not of such significant importance. Nevertheless, if accurate blocking probabilities are to be evaluated, the STD approach can be effectively applied [LPTG03].

The Delay Feedback Problem

Tightly coupled to the estimation of blocking probabilities, also the data rate adaptation for PS services introduces some inaccuracy. PS sources have typically an amount of traffic to be transmitted in a packet call. The time required for the transmission of this data, i.e. the call holding time, depends on the available data rate. This data rate, however, depends on the network load, as PS users are assumed to be served by different bearers depending on the cell load. Hence, there is a feedback of the network load on the traffic model. With a pure static snapshot approach this phenomenon, in [TPL+03] called *delay feedback problem*, can hardly be resolved. Similar to the problem of estimating accurate blocking probabilities, the history of the network state needs to be considered. A possible solution is again the STD simulation approach, but also for the delay feedback problem it shall be stated that the error made is only significant if the network load is very high.

5.10 Conclusions

In this chapter a Monte-Carlo method for the analysis of WCDMA networks has been presented. The developed method significantly outperforms current state-of-the-art solutions in terms of both, performance and flexibility. Although exact figures are confidential, the application of the developed method in everyday planning shows typical performance gains of a factor of ten in speed, and a factor of five in memory consumption compared with available commercial snapshot tools - even without an application of variance reduction techniques. This speed-up is mainly due to the very fast radio performance evaluation. For solving the power control equations, power control is performed on cell basis rather than on connection basis, such that the size of the original problem is significantly reduced.

The scheduling of resources in a snapshot is carried out by a new algorithm, which builds on a QoS scheme that can be very flexibly parameterized and thus allows to model a large variety of real world RRM algorithms. The method can be easily extended with respect to new system features and different technologies. Such extensions have partly already been introduced. Work on this topic will be further extended in future, especially with respect to new technologies and combined planning of technologies, e.g. GSM and UMTS networks.

In order to further improve the performance of the snapshot analysis, the application of variance reduction techniques has been investigated. Three different control variates for the estimation of the total cell TX power and blocking probabilities have been derived. Significant variance reductions were observed in the example study. This study will in the future need to be further extended with respect to different scenarios and the estimation of additional parameters. For these further control variates have to be derived. Once control variates are available for the most important output quantities, this method can be applied in everyday planning to significantly reduce the number of snapshots to be analyzed and hence to cut down the simulation time. In alignment with the results in Chapter 4 very good results in variance reduction were also obtained by applying *Correlated Sampling* (CorS) for the performance comparison between multiple network configurations. Due to its simplicity this approach can be directly applied in network planning simulations without considerable effort.

Chapter 6

Analytical Performance Evaluation

The advanced snapshot analysis method presented in the previous section allows for a very detailed, flexible, and accurate analysis of WCDMA networks. Although the developed method is significantly more efficient compared to current state-of-the-art methods in terms of computational time and memory requirements, there are applications for which a snapshot analysis is not well suited. In particular for an application in search based optimization techniques, which require a large number of different network configurations to be compared in a short time frame, a snapshot analysis is not applicable for reasonable network sizes.

A crucial part of the performance assessment is the estimation of cell load levels, which due to power control and the resulting mutual dependency between users strongly depend on the traffic distribution. First approaches for a load estimating by means of an "average snapshot", rather than averaging the results of a large number of snapshots have been presented in [DKKN99]. In [TTGH99] the concept of *demand nodes* is introduced. That is, the traffic given by pixel based average traffic maps is discretized into discrete demand nodes, e.g. representing users. In this chapter the following three methods for a load estimation are distinguished:

Static Method: Originally in [SLW+01] and taken up within the framework of the *European Union* (EU) project MOMENTUM, [mom], to which the author contributed, a different approach was developed. Instead of considering discrete users, the average contribution of the individual pixels to the power and interference budget is evaluated. Significant additional research focus has recently been spent on this approach, in the following referred to as *static method*, [EG05, EGK+06] and [SM04, Sta05, Sta04]. In Section 6.1 this basic approach is recalled and numerically compared with the results obtained from snapshot simulations.

Basic Statistical Method: It is shown that the static load estimation, by disregarding shadow fading, results in significant inaccuracies in terms of estimated cell load levels. In order to overcome these limitations, a new statistical method which considers the statistics of shadow fading is proposed. The derivation of the full method is split into two steps. In

a first step shadow fading is taken account of by means of serving probabilities per pixel. As a result a more accurate estimation of the load distribution between cells is obtained. For this method the term *basic statistical method* is used (Section 6.2).

Extended Statistical Method: In a second step an *extended statistical method* for load estimation is derived that fully considers shadow fading also in the derived cell load equations (Section 6.3).

As shown by the results, the statistical methods considerably outperform the static load estimation in terms of accuracy. In Section 6.4 the estimation of the most relevant quantities per pixel based on the results from the load estimation are discussed.

6.1 Static Load Estimation

In the following the static analysis approach is presented. The derivation is focused on the DL, the UL equations can generally be obtained equivalently. Nevertheless, there are a few substantial differences in the consideration of UL and DL. These will be pointed out in the corresponding sections, e.g. Section 6.1.5.

6.1.1 Average Required Link Power

A snapshot as considered in the previous chapters represents a possible traffic constellation of users. Snapshots are drawn from average traffic maps, which represent the average amount of users of a particular service within the area of one pixel. In other words, the number of drawn users in a pixel, averaged over all snapshots, converges to the mean value given by the traffic maps.

Besides the traffic demand, most of the other relevant parameters are also given on pixel basis or are mapped on pixels based on land use information, like orthogonality factors and users' velocity distribution parameters. The basic idea of the static load estimation is to calculate an average required capacity per pixel. That is, for the DL the average required DL cell power, and for the UL the amount of average interference at all cells is calculated for each pixel.

The required transmit power for a certain user has been derived in Section 5.5.2 and is given by equation (5.33). In Section 5.5.2 multiple scrambling codes per cell are considered. For the sake of simplicity in this section, only one scrambling code per cell is assumed. For most practical applications this assumption holds, otherwise the error made is not very large (additional intra-cell interference due to the lack of orthogonality between scrambling codes). Nevertheless, the concept can easily be extended in the direction of multiple scrambling codes. Without loss of generality outdoor and indoor traffic per service type is not distinguished (cf. Section 5.2), so that for the sake of readability no additional penetration loss is considered in the calculation of link gains. However,

multiple service types are considered. Indoor traffic can be integrated into the presented concept as users of an additional service type with higher path loss. Furthermore, the effect of SHO on the network performance is not considered, but the model can be extended in this direction.

For a user i of service type s located in pixel p the following notation is used: $i(p,s)$. Assuming one scrambling code per cell, hence $p_j^{\downarrow} = v_{c,k}$ and $v_{c,j} = 0 \quad for\ j \neq k$ in (5.33), the DL power required for user i served by cell c is

$$t_{i(p,s),c}^{\downarrow} = \alpha_{i(p,s)} \epsilon_{i(p,s)}^{\downarrow} \gamma_{i(p,s)}^{\downarrow} \kappa_{ci(p,s)}^{\downarrow} p_c^{\downarrow} + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \epsilon_{i(p,s)}^{\downarrow} \gamma_{i(p,s)}^{\downarrow} \kappa_{ji(p,s)}^{\downarrow} \frac{\xi_{ji(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} p_j^{\downarrow} + \frac{\epsilon_{i(p,s)}^{\downarrow} \gamma_{i(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} \left(\eta_{i(p,s)}^{\downarrow} + \breve{\imath}_{i(p,s)}^{\downarrow} \right) \tag{6.1}$$

$$= \zeta_{ci(p,s)}^{\downarrow} p_c^{\downarrow} + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \zeta_{ji(p,s)}^{\downarrow} \frac{\xi_{ji(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} p_j^{\downarrow} + \frac{\nu_{i(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} \tag{6.2}$$

where

$$\zeta_{ji(p,s)}^{\downarrow} = \begin{cases} \alpha_{i(p,s)} \epsilon_{i(p,s)}^{\downarrow} \gamma_{i(p,s)}^{\downarrow} \kappa_{ji(p,s)}^{\downarrow} & \text{if j serving cell} \\ \epsilon_{i(p,s)}^{\downarrow} \gamma_{i(p,s)}^{\downarrow} \kappa_{ji(p,s)}^{\downarrow} & \text{otherwise} \end{cases} \tag{6.3}$$

$$\nu_{i(p,s)}^{\downarrow} = \epsilon_{i(p,s)}^{\downarrow} \gamma_{i(p,s)}^{\downarrow} \left(\eta_{i(p,s)}^{\downarrow} + \breve{\imath}_{i(p,s)}^{\downarrow} \right) \tag{6.4}$$

Assuming that all users in pixel p are served by cell c, the average transmission power $t_{c,p,s}^{*\downarrow}$ for a user of service type s in pixel p yields:

$$t_{c,p,s}^{*\downarrow} = \mathrm{E}\left\{ t_{i(p,s)}^{\downarrow} \right\} = \mathrm{E}\left\{ \zeta_{ci(p,s)}^{\downarrow} p_c^{\downarrow} \right\} + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \mathrm{E}\left\{ \zeta_{ji(p,s)}^{\downarrow} \frac{\xi_{ji(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} p_j^{\downarrow} \right\} + \mathrm{E}\left\{ \frac{\nu_{i(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} \right\} \tag{6.5}$$

The last term in (6.5) can be split, as $\nu_{i(p,s)}^{\downarrow}$ and $\xi_{ci(p,s)}^{\downarrow}$ are independent:

$$\mathrm{E}\left\{ \frac{\nu_{i(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} \right\} = \mathrm{E}\left\{ \nu_{i(p,s)}^{\downarrow} \right\} \mathrm{E}\left\{ \frac{1}{\xi_{ci(p,s)}^{\downarrow}} \right\} \tag{6.6}$$

For the other two sub-terms in (6.5) the following approximation is made:

$$\mathrm{E}\left\{ \zeta_{ci(p,s)}^{\downarrow} p_c^{\downarrow} \right\} \approx \mathrm{E}\left\{ \zeta_{ci(p,s)}^{\downarrow} \right\} \cdot \mathrm{E}\left\{ p_c^{\downarrow} \right\} \tag{6.7}$$

$$\mathrm{E}\left\{ \zeta_{ji(p,s)}^{\downarrow} \frac{\xi_{ji(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} p_j^{\downarrow} \right\} \approx \mathrm{E}\left\{ \zeta_{ji(p,s)}^{\downarrow} \right\} \cdot \mathrm{E}\left\{ \frac{\xi_{ji(p,s)}^{\downarrow}}{\xi_{ci(p,s)}^{\downarrow}} \right\} \cdot \mathrm{E}\left\{ p_j^{\downarrow} \right\} \tag{6.8}$$

This approximation is based on the assumption that the correlation between cell power levels and the specifics of a particular user in the cell, e.g. service type, speed, and location, is low. That is, the individual user is assumed to have no significant impact on the

overall cell performance. Obviously this assumption only holds for low data rate users which only demand a limited share of the resources. This will be made evident when presenting results on the method.

Note, that $\zeta^{\downarrow}_{ji(p,\,s)}$ and $\nu^{\downarrow}_{i(p,\,s)}$ only depend on service type, location (pixel), and user's velocity, so that they can be considered as location and service dependent random variables with mean $\zeta^{*\downarrow}_{j,p,s}$ and $\nu^{*\downarrow}_{p,s}$

$$\zeta^{*\downarrow}_{j,p,s} = \mathrm{E}\left\{\zeta^{\downarrow}_{ji(p,\,s)}\right\} \tag{6.9}$$

$$\nu^{*\downarrow}_{p,s} = \mathrm{E}\left\{\nu^{\downarrow}_{i(p,\,s)}\right\} \tag{6.10}$$

which can be calculated simply by numerical integration over the speed distribution for each pixel (type).

The average cell power levels are denoted as

$$\mathrm{E}\left\{p^{\downarrow}_{i}\right\} = p^{*\downarrow}_{i} \tag{6.11}$$

The link gain $\xi^{\downarrow}_{ji(p,\,s)}$ from cell j to mobile i can be split into a median link gain $\xi^{*\downarrow}_{c,p,s}$, which is constant for a given pixel p, and a location dependent random shadowing component $\omega_{cp} = 10^{\Sigma_{c,p}/10}$ (cf. Section 3.2):

$$\xi^{\downarrow}_{ji(p,\,s)} = \xi^{*\downarrow}_{j,p,s} \cdot \omega_{jp} \tag{6.12}$$

The main approximation of the static method is that shadow fading is neglected. Hence, the link gain is approximated by its median link gain, i.e.

$$\xi^{\downarrow}_{ji(p,\,s)} \approx \xi^{*\downarrow}_{j,p,s} \tag{6.13}$$

The reasoning behind disregarding shadow fading is the assumption that shadowing effects are averaged out in the load estimation when considering a large number of pixels. As is shown below when results on the method are presented, this is the main reason for the inaccuracy of the static method.

With (6.13), equations (6.6) and (6.8) can be further simplified:

$$\mathrm{E}\left\{\frac{1}{\xi^{\downarrow}_{ci(p,\,s)}}\right\} \approx \frac{1}{\xi^{*\downarrow}_{c,p,s}} \tag{6.14}$$

$$\mathrm{E}\left\{\frac{\xi^{\downarrow}_{ji(p,\,s)}}{\xi^{\downarrow}_{ci(p,\,s)}}\right\} \approx \frac{\xi^{*\downarrow}_{j,p,s}}{\xi^{*\downarrow}_{c,p,s}} \tag{6.15}$$

Consequently, (6.5) is approximated to

$$t^{*\downarrow}_{c,p,s} \approx \zeta^{*\downarrow}_{c,p,s} p^{*\downarrow}_{c} + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \zeta^{*\downarrow}_{j,p,s} \frac{\xi^{*\downarrow}_{j,p,s}}{\xi^{*\downarrow}_{c,p,s}} p^{*\downarrow}_{j} + \frac{\nu^{*\downarrow}_{p,s}}{\xi^{*\downarrow}_{c,p,s}} \tag{6.16}$$

6.1.2 Cell Assignment

As described in Section 5.5, users are served by the cell with highest pilot power reception level. Since the static method described in this section does not consider shadow fading, the serving cell of a particular pixel p is the cell c for which the following expression holds:

$$\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \geq \hat{\tau}_j \xi^{*\downarrow}_{j,p,s} \quad \forall j \in \mathcal{C},\ j \neq c \tag{6.17}$$

In case there are multiple cells with the same received pilot power, the best serving cell is randomly chosen. This cell is in the following denoted R^*_p.

6.1.3 Serving Probability

The serving probability $\delta_{c,p,s}$ is defined as the probability that pixel p is served by cell c. For the simple model described in this section it yields:

$$\delta_{c,p,s} = \begin{cases} 1 & \text{if } R^*_p = c \\ 0 & \text{otherwise} \end{cases} \tag{6.18}$$

6.1.4 Coverage Probability

A user is only covered, i.e. can be served by the system, if none of the various outage restrictions described in Section 5.4 is violated. Here it is assumed that coverage is described by sufficient pilot coverage, i.e. the received pilot power level of the best serving cell exceeds the minimum required receive level:

$$\upsilon_{c,p,s} = \begin{cases} 1 & \text{if } \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} > \hat{r}_{p,\,min} \\ 0 & \text{otherwise} \end{cases} \tag{6.19}$$

where $\upsilon_{c,p,s}$ denotes the *coverage probability*.

For a more detailed discussion on the calculation and incorporation of outage probabilities into the method refer to Section 6.2.1.

6.1.5 Characteristic Cell Power Equations

The average cell transmit power required to serve a user in a pixel p is approximated by (6.16). The average cell transmit power to serve all users of service type s in pixel p consequently is the mean transmit power required for one user $t^{*\downarrow}_{c,p,s}$ multiplied with the average number of users in that pixel $\theta_{p,s}$ and the coverage probability $\upsilon_{c,p,s}$. The total cell power of cell c is the sum of the contributions of all pixels that are served by the particular

cell. Assuming S different services and a total amount of P pixels, for the cell power of cell c it yields

$$p_c^{*\downarrow} = \sum_{s=1}^{S}\sum_{p=1}^{P} \theta_{p,s}\delta_{c,p,s}\upsilon_{c,p,s}t_{c,p,s}^{*\downarrow} + \tau_c \tag{6.20}$$

$$\begin{aligned} &= \sum_{s=1}^{S}\sum_{p=1}^{P} \theta_{p,s}\delta_{c,p,s}\upsilon_{c,p,s}\left(\zeta_{c,p,s}^{*\downarrow}p_c^{*\downarrow} + \sum_{\substack{j\in\mathcal{C}\\ j\neq c}} \zeta_{j,p,s}^{*\downarrow}\frac{\xi_{j,p,s}^{*\downarrow}}{\xi_{c,p,s}^{*\downarrow}}p_j^{*\downarrow} + \frac{\nu_{p,s}^{*\downarrow}}{\xi_{c,p,s}^{*\downarrow}}\right) + \tau_c \\ &= \underbrace{\sum_{s=1}^{S}\sum_{p=1}^{P} \theta_{p,s}\delta_{c,p,s}\upsilon_{c,p,s}\zeta_{c,p,s}^{*\downarrow}}_{T_{cc}^{*\downarrow}} p_c^{*\downarrow} + \sum_{\substack{j\in\mathcal{C}\\ j\neq c}} \underbrace{\sum_{s=1}^{S}\sum_{p=1}^{P} \theta_{p,s}\delta_{c,p,s}\upsilon_{c,p,s}\zeta_{j,p,s}^{*\downarrow}\frac{\xi_{j,p,s}^{*\downarrow}}{\xi_{c,p,s}^{*\downarrow}}}_{T_{cj}^{*\downarrow}} p_j^{*\downarrow} \\ &\quad + \underbrace{\tau_c + \sum_{s=1}^{S}\sum_{p=1}^{P} \theta_{p,s}\delta_{c,p,s}\upsilon_{c,p,s}\frac{\nu_{p,s}^{*\downarrow}}{\xi_{c,p,s}^{*\downarrow}}}_{h_c^{*\downarrow}} \end{aligned} \tag{6.21}$$

With (6.21) a linear set of equations can be formulated:

$$\vec{p}^{\,*} = T^{*\downarrow}\vec{p}^{\,*} + \vec{h}^{\,*\downarrow} \tag{6.22}$$

This equation approximates the average cell power of each cell in the system. It is similar to (5.42), which determines the cell power levels for a certain snapshot. The matrix $T^{*\downarrow}$ describes the average coupling between cells.

Equation (6.22) is solved iteratively (cf. Section 5.5.4). An important characteristic of $T^{*\downarrow}$ and $\vec{h}^{\,*\downarrow}$ is that the traffic served by a cell c only influences the entries of rows c. Hence, the amount of traffic in cell c in one iteration only influences the cell power of cell c. Thus, while clipping cell power levels to their maximum in case of calculated cell power levels exceeding their maximum, the influence of an overloaded cell on the other cells is not overestimated.

Also for the UL a set of equations for calculating the UL cell interference levels equivalent to (6.22) can be obtained (cf. (5.25)). Solving the UL set of equations however is considerably more complicated than solving the DL set of equations, because the traffic served by a cell c, rather than impacting the entries of row c as in the DL case, impact the entries of column c of the coupling matrix. Thus, while iteratively solving the system, it is not sufficient to clip interference levels to the respective maximum levels for overloaded cells. Instead, the respective matrix entries also need to be reduced, in order not to overestimate the impact of the overloaded cell on the other cells in the system. More details on this topic and methods to efficiently solve this problem can be found in [EGR05, Maj06].

6.1.6 Application Example: Static Load Estimation

The static load estimation presented in this section is in the following applied to a large realistic planning scenario, referred to as "Metropolis", consisting of 921 cells, comprising urban, sub-urban, and rural areas (cf. Figure 6.1). The estimation of mean cell power levels obtained from the static method is compared with the results from snapshot simulations for four different traffic types: Speech (CS12.2), CS64 Streaming, PS128, and PS384. The applied CIR target and activity rates are presented in Table 6.1. Traffic is inhomogeneously distributed with clutter dependent weights. For all traffic types the amount of traffic was adjusted, such that the average cell load in the network is roughly 40 % (8 W). A maximum cell TX power $p^{\downarrow}_{j,\,max}$ of 20 W is used, total static cell power τ_j and pilot power $\hat{\tau}_j$ are set to 4 W and 2 W, respectively.

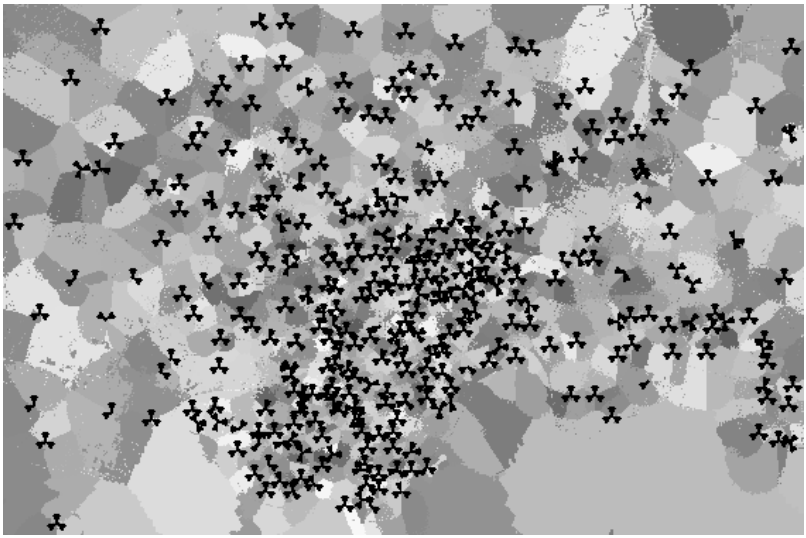


Figure 6.1: Realistic planning scenario "Metropolis" considered for results obtained in this chapter (Depicted: site location, cell orientation, and best server area)

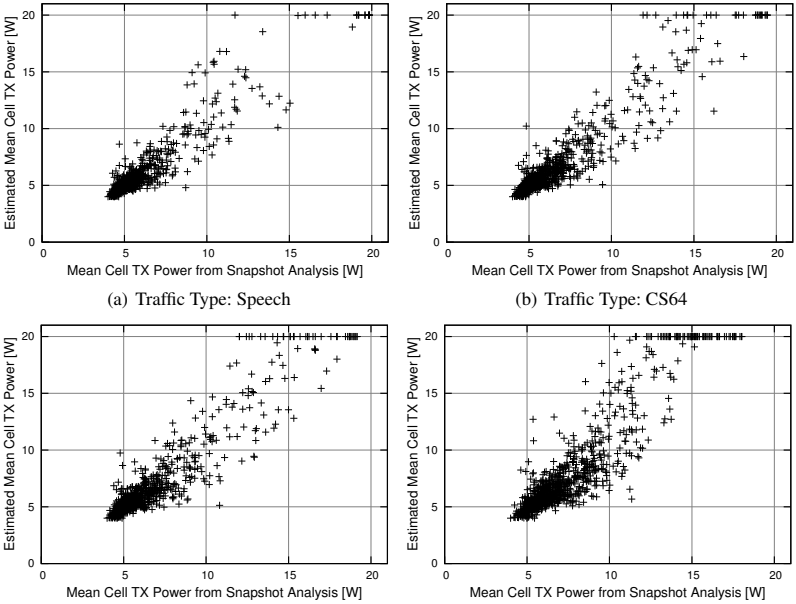
Bearer	DL Bit Rate	$\gamma^{\downarrow}$	$\epsilon^{\downarrow}$
Speech	12.2 kbps	-17.7 dB	0.5
CS64	64 kbps	-12.9 dB	1.0
PS128	128 kbps	-11.8 dB	1.0
PS384	384 kbps	-6.8 dB	1.0

Table 6.1: Traffic parameters used for example results in this chapter

Shadow fading is in the snapshot simulations modeled according to the standard shadow fading model (3.15). The applied shadow fading parameters, presented in Table 6.2, are chosen according to the default values given in Section 3.2 (cf. Table 3.2).

Parameter	Value
σ	8 dB
ρ_L	0.5

Table 6.2: Shadow fading parameters used for example results in this chapter



(a) Traffic Type: Speech

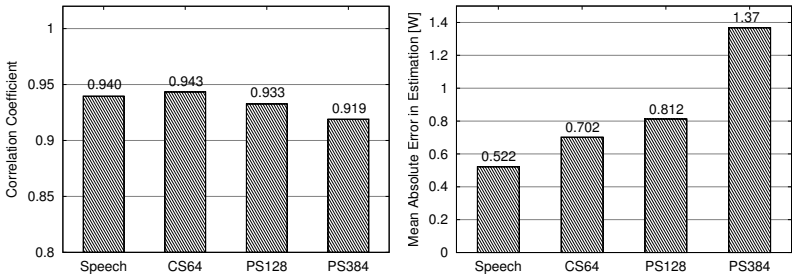
(b) Traffic Type: CS64

(c) Traffic Type: PS128

(d) Traffic Type: PS384

Figure 6.2: Correlation between the mean cell TX power obtained from the static method and snapshot analysis for different traffic types

The correlation between the cell load levels obtained from the static method and the snapshot analysis is illustrated in Figure 6.2. The plots clearly show a positive correlation between the two quantities, but also a considerable error in estimation. The plots also illustrate the clipping to the maximum cell load levels for the analytical load estimation in case of overloaded cells. For the results obtained from snapshots, this effect is not so clearly visible. This is mainly due to the statistical variation of offered traffic between snapshots. As a consequence, even though the cell might be in overload in the majority of snapshots, the other snapshots cause an average cell power which is below the maximum.



(a) Correlation Coefficient (r) (b) Mean Absolute Error (e)

Figure 6.3: Correlation coefficient and mean absolute error in cell power estimation for the static method

But even for cells that are overloaded in each snapshot, the mean cell power obtained from the snapshot analysis does not exactly equal the maximum power. The reason is that discrete users are considered in snapshots. In case of overload, individual users are dropped, and cell power levels are recalculated. Hence, due to this granularity of power levels, the cell load levels is never exactly equal to the maximum power.

In Figure 6.3 the correlation between static load estimation and the reference values taken from the snapshot analysis is quantified. The *correlation coefficient* r and the *mean absolute error* e depicted in the charts are defined as:

$$r = \frac{\operatorname{cov}\left\{\mathrm{E}\left\{p_c^{\downarrow}\right\}, p_c^{*\downarrow}\right\}}{\sqrt{\operatorname{var}\left\{\mathrm{E}\left\{p_c^{\downarrow}\right\}\right\} \operatorname{var}\left\{p_c^{*\downarrow}\right\}}} \tag{6.23}$$

$$e = \frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} \left| \mathrm{E}\left\{p_c^{\downarrow}\right\} - p_c^{*\downarrow} \right| \tag{6.24}$$

where $\mathrm{E}\left\{p_c^{\downarrow}\right\}$ refers to the reference value obtained from snapshot simulations.

The charts in Figure 6.3 clearly reflect what intuitively can be obtained from the correlation plots in Figure 6.2. Both bar charts show a clear tendency of a decreased performance if higher data rate services are considered. Only the correlation factor for the CS64 slightly opposes this tendency. The correlation factor for CS64 is slightly higher than the correlation factor for the Speech traffic, although as expected the relative error is higher. This effect, however, is very small and an exception that is not supported by any other results obtained from the method. It is hence concluded that this is a statistical effect which is caused by the empirical type of analysis.

6.2 Statistical Load Estimation

Although the results of the static load estimation presented above show a fairly strong positive correlation between estimated cell load levels and reference values, a large error in estimation is also made evident. In this section and in the next section it is shown that this error is mainly due to the effects of shadow fading. The effects of shadow fading are twofold. Firstly, the strongest server of a pixel is not constant, but a random variable that can take different cells as value. Consequently, the DL contribution of a certain pixel to the DL cell power needs to be split between several servers according to their best server probabilities. Secondly, shadow fading also impacts the link gain statistics involved in the power calculations. The first effect is investigated in this section, the second is taken account of in the derivation of the extended statistical method in Section 6.3.

6.2.1 Serving Probabilities

In the static load estimation it is assumed that each pixel is served by exactly one serving cell. Due to shadow fading, however, pixels are served with certain probabilities by different cells. The best serving cell of pixel p, in the static load estimation denoted R_p^*, consequently is a random variable $R_p\{\mathcal{S}_p\}$, where $\mathcal{S}_p$ denotes the *set of potential servers* of pixel p. $\mathcal{S}_p$ contains all cells that have a considerable probability to be best serving cell of p and is introduced in order to limit the computational complexity of the developed method. However, $\mathcal{S}_p$ can also be chosen identical to $\mathcal{C}$.

For the shadow fading component associated with the link from a cell j to pixel p the following notations are used:

$$(\omega_{jp})_{[dB]} = \sigma_p \cdot \left(\rho_{Lp} X_p + \sqrt{1-\rho_{Lp}^2}\, Y_{jp}\right) \tag{6.25}$$

$$= \sigma'_p X_p + \sigma''_p Y_{jp} \tag{6.26}$$

where $(...)_{[dB]}$ indicates a conversion from linear scale to decibels, and

$$\sigma'_p = \sigma_p \rho_{Lp} \tag{6.27}$$

$$\sigma''_p = \sigma_p \sqrt{1-\rho_{Lp}^2} \tag{6.28}$$

The random variables X_p and Y_{jp} are standard normal distributed random variables. Their probability density functions and distribution are denoted as

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} \tag{6.29}$$

$$\Phi(x) = \int_{-\infty}^{x} \phi(y)\, dy \tag{6.30}$$

The probability that cell c is best serving cell of pixel p is the probability that the received pilot signal level of cell c is above the received pilot power level of all other cells:

$$\begin{aligned}
\delta_{c,p,s} &= \mathrm{P}\left\{R_i\{\mathcal{S}_i\} = c\right\} \\
&= \mathrm{P}\left\{\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp}\right)_{[dB]} > \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s} \omega_{jp}\right)_{[dB]} \;\forall j \in \mathcal{S}_p : j \neq c\right\} \\
&= \mathrm{P}\left\{\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} + \left(\omega_{cp}\right)_{[dB]} > \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + \left(\omega_{jp}\right)_{[dB]} \forall j \in \mathcal{S}_p : j \neq c\right\} \\
&= \mathrm{P}\left\{\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} + \sigma''_p Y_{cp} > \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + \sigma''_p Y_{jp} \;\forall j \in \mathcal{S}_p : j \neq c\right\}
\end{aligned} \tag{6.31}$$

As the random variables Y_{jp} are independent, (6.31) can be solved by integrating over the probability density of Y_{cp}, evaluating for each value the probability that the received pilot strength of cell c is greater than the received pilot power levels of all other cells:

$$\delta_{c,p,s} = \int_{-\infty}^{+\infty} \phi\left(x\right) \mathrm{P}\left\{R_p\{\mathcal{S}_p\} = c \,|\, Y_{cp} = x\right\} dx \tag{6.32}$$

where

$$\begin{aligned}
&\mathrm{P}\left\{R_p\{\mathcal{S}_p\} = c \mid\; Y_{cp} = x\right\} \\
&\quad= \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \mathrm{P}\left\{\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} + \sigma''_p x > \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + \sigma''_p Y_{jp}\right\} \\
&\quad= \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \mathrm{P}\left\{Y_{jp} < x + \frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]}\right\} \\
&\quad= \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \Phi\left(x + \frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]}\right)
\end{aligned} \tag{6.33}$$

Substituting (6.33) in (6.32), the final result for the serving probability is obtained:

$$\delta_{c,p,s} = \int_{-\infty}^{+\infty} \phi\left(x\right) \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \Phi\left(x + \frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]}\right) dx \tag{6.34}$$

6.2.2 Coverage Probability

In the previous sub-section the probability that a particular cell is best serving cell of a pixel has been derived. It is, however, only served if also the pilot coverage constraint introduced in Section 6.1.4 is fulfilled.

As the product of serving probability $\delta_{c,p,s}$ and $\upsilon_{c,p,s}$ is required for the calculation of the characteristic system matrices (cf. (6.20)), in the following directly this product is evaluated. It is obtained by multiplying the integrand in (6.34) with the probability, that the received pilot signal of cell c is below the minimum level for a particular value of $Y_{cp} = x$:

$$\begin{aligned}\delta_{c,p,s} \cdot \upsilon_{c,p,s} = &\int_{-\infty}^{+\infty} \phi(x)\, \mathrm{P}\left\{R_p\{\mathcal{S}_p\} = c \mid Y_{cp} = x\right\} \\ &\cdot \mathrm{P}\left\{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \mid Y_{cp} = x\right\} dx \end{aligned} \tag{6.35}$$

where

$$\begin{aligned}&\mathrm{P}\left\{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \mid Y_{cp} = x\right\} \\ &\quad = \mathrm{P}\left\{\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} + \sigma'_p X_p + \sigma''_p x > \left(\hat{r}_{p,\,min}\right)_{[dB]}\right\} \\ &\quad = \mathrm{P}\left\{X_p > \frac{1}{\sigma'_p}\left(\left(\frac{\hat{r}_{p,\,min}}{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}\right)_{[dB]} - \sigma''_p x\right)\right\} \\ &\quad = 1 - \Phi\left(\frac{1}{\sigma'_p}\left(\left(\frac{\hat{r}_{p,\,min}}{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}\right)_{[dB]} - \sigma''_p x\right)\right)\end{aligned} \tag{6.36}$$

Equations (6.36) and (6.33) in (6.35) finally yields

$$\begin{aligned}\delta_{c,p,s} \cdot \upsilon_{c,p,s} = &\int_{-\infty}^{+\infty} \phi(x)\left(1 - \Phi\left(\frac{1}{\sigma'_p}\left(\left(\frac{\hat{r}_{p,\,min}}{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}\right)_{[dB]} - \sigma''_p x\right)\right)\right) \\ &\cdot \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \Phi\left(x + \frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]}\right) dx \end{aligned} \tag{6.37}$$

6.2.3 Characteristic Cell Power Equations

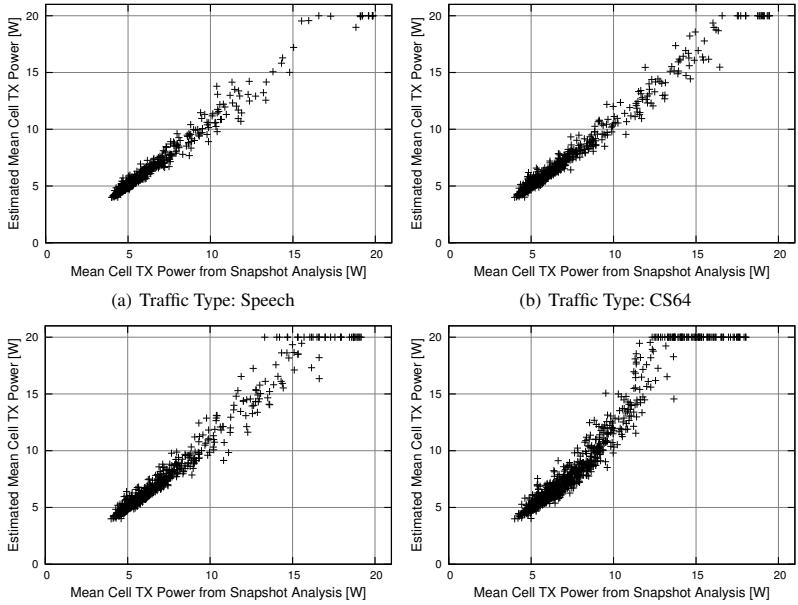
The characteristic system equation is obtained equivalently to the description in Section 6.1.5 with the difference that a pixel might contribute to the cell power of multiple cells:

$$p^{*\downarrow}_c = \sum_{s=1}^{S} \sum_{p=1}^{P} \theta_{p,s} \delta_{c,p,s} \upsilon_{c,p,s} t^{*\downarrow}_{R^*_p,p,s} + \tau_c \tag{6.38}$$

where $\delta_{c,p,s}$ and $\upsilon_{c,p,s}$ are given by equations (6.34) and (6.37). Note that the link power $t^{*\downarrow}_{R^*_p,p,s}$ is calculated according to (6.16) for the "static" server R^*_p, although the contribution might be added to a different serving cell j. The reason is that shadow fading is not considered in the link gain statistics at this point. The contribution of pixel p to the cell power j would be significantly overestimated if calculated based on the median link gain of cell j which is for equal pilot power levels by definition less (or equal) the median link gain of cell R^*_p. From (6.38) a matrix equation similar to (6.22) can be obtained.

6.2.4 Application Example: Statistical Load Estimation

In analogy with the analysis in Section 6.1.6 the improved method developed in this section is also applied to the example planning scenario. Numerical results are presented in Figure 6.5. Both the error of estimation and the correlation coefficient are considerably improved for all traffic types compared with the results for the static method in Section 6.1.6. This is also visually reflected in the correlation plots in Figure 6.4. For the traffic types with lower data rate (Speech and CS64), the average error has been reduced by a factor of approximately 2.



(a) Traffic Type: Speech

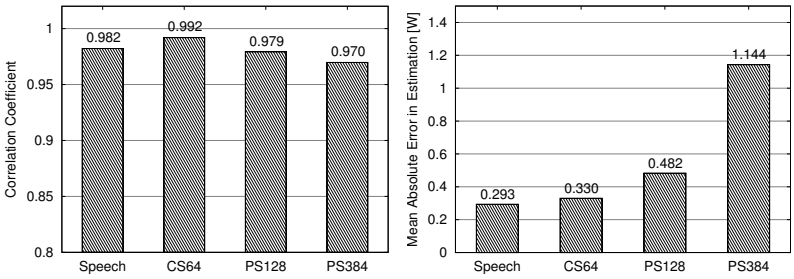
(b) Traffic Type: CS64

(c) Traffic Type: PS128

(d) Traffic Type: PS384

Figure 6.4: Correlation between the mean cell TX power obtained from basic statistical method and snapshot analysis for different traffic types

Furthermore, the results make evident that the error in estimation significantly increases the greater the demand of the services, i.e. for higher data rate traffic. A general observation made from the plots in Figure 6.4 is that in particular for the high data rate traffic cases not only the variance of the estimation error is increased, but that there is also a tendency to overestimate the cell load levels by the analytical method. This tendency needs to be further investigated but is expected to be due to the higher (relative) statistical variance in offered traffic for the high data rate services, as the average number of offered users yielding the same average load is lower.



(a) Correlation Coefficient (r) (b) Mean Absolute Error (e)

Figure 6.5: Correlation coefficient and mean absolute error in cell power estimation for the basic statistical method

6.3 Extended Statistical Load Estimation

The static method presented first in this chapter does not consider shadow fading at all. The basic statistical method presented in the previous section considers shadow fading in terms of serving probabilities for each pixel. Shadow fading however also has an impact on the average link power required for the individual pixels, as the corresponding equations contain fractions of link gains and inverse link gains. The expected value of these quantities is influenced by shadow fading.

6.3.1 Average Required Link Power

In the derivation of the average required DL power for a user in a certain pixel in Section 6.1.1, it is assumed that a pixel is served by exactly one cell R_p^*. In the previous section it has been discussed, however, that shadow fading causes the best serving cell to be a random variable $R_p\{\mathcal{S}_p\}$. Furthermore, in case a cell c is known to be the best serving cell of a pixel, the link gains between the other cells to the same pixel are not described by independent normal distributed random variables anymore. In fact, there is a correlation between the link gain of best serving cell c and the link gains of the other cells. It is in particular known that the realization of the link gains of all other cells are so low that they are not best serving cell. Equation (6.5) therefore has to be extended by conditional expressions when multiple servers per pixel are considered:

$$\begin{aligned} t_{c,p,s}^{*\downarrow} &= \mathrm{E}\left\{\zeta_{ci(p,\,s)}^{\downarrow} p_c^{\downarrow}\right\} \\ &\quad + \sum_{\substack{j\in\mathcal{C}\\ j\neq c}} \mathrm{E}\left\{ \zeta_{ji(p,\,s)}^{\downarrow} \frac{\xi_{ji(p,\,s)}^{\downarrow}}{\xi_{ci(p,\,s)}^{\downarrow}} p_j^{\downarrow} \;\middle|\; R_p\{\mathcal{S}_p\} = c \,\cap\, \hat{\tau}_c \xi_{c,p,s}^{*\downarrow} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\ &\quad + \mathrm{E}\left\{ \frac{\nu_{i(p,\,s)}^{\downarrow}}{\xi_{ci(p,\,s)}^{\downarrow}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \,\cap\, \hat{\tau}_c \xi_{c,p,s}^{*\downarrow} \omega_{cp} > \hat{r}_{p,\,min} \right\} \end{aligned} \tag{6.39}$$

The last sub-term can be evaluated in analogy with (6.6):

$$\begin{aligned}
&\mathrm{E}\left\{ \frac{\nu^{\downarrow}_{i(p,\,s)}}{\xi^{\downarrow}_{ci(p,\,s)}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\
&\qquad = \mathrm{E}\left\{ \nu^{\downarrow}_{i(p,\,s)} \right\} \cdot \mathrm{E}\left\{ \frac{1}{\xi^{\downarrow}_{ci(p,\,s)}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\
&\qquad = \nu^{*\downarrow}_{p,s} \frac{1}{\xi^{*\downarrow}_{c,p,s}} \cdot \mathrm{E}\left\{ \frac{1}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\}
\end{aligned} \tag{6.40}$$

The other two sub-terms of (6.39) are with the same reasoning as given in Section 6.1.1 approximated as follows:

$$\begin{aligned}
&\mathrm{E}\left\{ \zeta^{\downarrow}_{ji(p,\,s)} \frac{\xi^{\downarrow}_{ji(p,\,s)}}{\xi^{\downarrow}_{ci(p,\,s)}} p^{\downarrow}_{j} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\
&\qquad \approx \zeta^{*\downarrow}_{j,p,s} \cdot \mathrm{E}\left\{ \frac{\xi^{\downarrow}_{ji(p,\,s)}}{\xi^{\downarrow}_{ci(p,\,s)}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \cdot p^{*\downarrow}_{j} \\
&\qquad = \zeta^{*\downarrow}_{j,p,s} \frac{\xi^{*\downarrow}_{j,p,s}}{\xi^{*\downarrow}_{c,p,s}} \cdot \mathrm{E}\left\{ \frac{\omega_{jp}}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \cdot p^{*\downarrow}_{j}
\end{aligned} \tag{6.41}$$

$$\mathrm{E}\left\{ \zeta^{\downarrow}_{ci(p,\,s)} p^{\downarrow}_{c} \right\} \approx \mathrm{E}\left\{ \zeta^{\downarrow}_{ci(p,\,s)} \right\} \cdot \mathrm{E}\left\{ p^{\downarrow}_{c} \right\} = \zeta^{*\downarrow}_{c,p,s} \cdot p^{*\downarrow}_{c} \tag{6.42}$$

With (6.40), (6.41), and (6.42), equation (6.39) can be approximated:

$$\begin{aligned}
t^{*\downarrow}_{c,p,s} &= \zeta^{*\downarrow}_{c,p,s} p^{*\downarrow}_{c} \\
&\quad + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \zeta^{*\downarrow}_{j,p,s} \frac{\xi^{*\downarrow}_{j,p,s}}{\xi^{*\downarrow}_{c,p,s}} \cdot \underbrace{\mathrm{E}\left\{ \frac{\omega_{jp}}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\}}_{\Xi_{c,p,s}} \cdot p^{*\downarrow}_{j} \\
&\quad + \nu^{*\downarrow}_{p,s} \frac{1}{\xi^{*\downarrow}_{c,p,s}} \cdot \underbrace{\mathrm{E}\left\{ \frac{1}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\}}_{\Psi_{c,p,s}} \\
&= \zeta^{*\downarrow}_{c,p,s} p^{*\downarrow}_{c} + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \zeta^{*\downarrow}_{j,p,s} \frac{\xi^{*\downarrow}_{j,p,s}}{\xi^{*\downarrow}_{c,p,s}} \cdot \Xi_{c,p,s} \cdot p^{*\downarrow}_{j} + \nu^{*\downarrow}_{p,s} \frac{1}{\xi^{*\downarrow}_{c,p,s}} \cdot \Psi_{c,p,s}
\end{aligned} \tag{6.43}$$

Inverse Shadow Fading Component Ψ

Instead of evaluating $\Psi_{c,p,s}$ directly, for a simplified derivation the following expression is considered:

$$\delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \Psi_{c,p,s} = \delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \mathrm{E}\left\{ \frac{1}{\omega_{cp}} \;\middle|\; R_i\{\mathcal{S}_i\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \quad (6.44)$$

where $\delta_{c,p,s} \cdot \upsilon_{c,p,s}$ is given by (6.37).

The inverse shadow fading component may be rewritten:

$$\frac{1}{\omega_{cp}} = 10^{-\frac{1}{10}\sigma'_p X_p} \cdot 10^{-\frac{1}{10}\sigma''_p Y_{cp}}$$

For the coverage power constraint it yields:

$$\begin{aligned} & \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \\ \Leftrightarrow \quad & \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} \cdot \left(\sigma'_p X_p + \sigma''_p Y_{cp}\right) > \left(\hat{r}_{p,\,min}\right)_{[dBm]} \\ \Leftrightarrow \quad & X_p > -\frac{\sigma''_p Y_{cp} + \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \left(\hat{r}_{c,\,min}\right)_{[dB]}}{\sigma'_p} \end{aligned} \quad (6.45)$$

Equation (6.44) is calculated by integrating over the shadow fading component Y_{cp}. For each possible value of Y_{cp} the expected value of the inverse shadow fading component and the probability that c is best serving cell and that the pixel is covered by c is evaluated. The coverage probability depends for a given Y_{cp} on X_p only:

$$\begin{aligned} & \delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \mathrm{E}\left\{ \frac{1}{\omega_{cp}} \;\middle|\; R_i\{\mathcal{S}_i\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\ & \quad = \int_{-\infty}^{+\infty} 10^{-\frac{\sigma''_p}{10} x} \cdot \mathrm{E}\left\{ 10^{-\frac{\sigma''_p}{10} X_p} \;\middle|\; X_p > -\frac{\sigma''_p Y_{cp} + \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \left(\hat{r}_{c,\,min}\right)_{[dB]}}{\sigma'_p} \right\} \\ & \qquad \cdot \mathrm{P}\left\{ X_p > -\frac{\sigma''_p Y_{cp} + \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \left(\hat{r}_{c,\,min}\right)_{[dB]}}{\sigma'_p} \right\} \\ & \qquad \cdot \mathrm{P}\left\{ R_i\{\mathcal{S}_i\} = c \mid Y_{cp} = x \right\} \phi(x)\, dx \end{aligned} \quad (6.46)$$

By substituting

$$a = \frac{\sigma''_p x + \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \left(\hat{r}_{c,\,min}\right)_{[dB]}}{\sigma'_p}$$

the expected value in (6.45) is evaluated as:

$$\begin{aligned}
&\mathrm{P}\left\{X_p > -a\right\} \cdot \mathrm{E}\left\{ 10^{-\frac{\sigma'_p}{10}X_p} \;\middle|\; X_p > -a \right\} \\
&\quad = \int_{-a}^{\infty} 10^{-\frac{\sigma'_p}{10}x} \phi\left(x\right) dx \\
&\quad = \int_{-a}^{\infty} \mathrm{e}^{-\frac{\ln 10}{10}\sigma'_p x} \phi\left(x\right) dx \\
&\qquad\qquad \text{subst.: } b = \frac{\ln 10}{10}\sigma'_p, \quad y = -x \\
&\quad = \int_{-\infty}^{a} \mathrm{e}^{bx} \phi\left(y\right) dy \qquad (6.47) \\
&\quad = \frac{1}{\sqrt{2\pi}} \int_{-a}^{\infty} \mathrm{e}^{-by} \mathrm{e}^{-\frac{y^2}{2}} dy \\
&\quad = \frac{1}{\sqrt{2\pi}} \int_{-a}^{\infty} \mathrm{e}^{-\frac{(y-b)^2}{2} + \frac{b^2}{2}} dy \\
&\quad = \mathrm{e}^{\frac{b^2}{2}} \cdot \Phi\left(a - b\right) \qquad (6.48) \\
&\quad = \mathrm{e}^{\frac{1}{2}\left(\frac{\ln 10}{10}\right)^2 {\sigma'_p}^2} \cdot \Phi\left(\frac{\sigma''_p x + \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \left(\hat{r}_{c,\,min}\right)_{[dB]}}{\sigma'_p} - \frac{\ln 10}{10}\sigma'_p \right) \qquad (6.49)
\end{aligned}$$

Substituting (6.49) and (6.33) in (6.46), the following final expression for Ψ is obtained:

$$\begin{aligned}
&\delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \Psi_{c,p,s} \\
&\quad = \delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \mathrm{E}\left\{ \frac{1}{\omega_{cp}} \;\middle|\; R_i\{\mathcal{S}_i\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\
&\quad = \int_{-\infty}^{+\infty} 10^{-\frac{\sigma''_p}{10}x} \cdot \phi\left(x\right) \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \Phi\left(x + \frac{1}{\sigma''_p} \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}} \right)_{[dB]} \right) \\
&\qquad \cdot \mathrm{e}^{\frac{1}{2}\left(\frac{\ln 10}{10}\right)^2 {\sigma'_p}^2} \cdot \Phi\left(\frac{\sigma''_p x + \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \left(\hat{r}_{c,\,min}\right)_{[dB]}}{\sigma'_p} - \frac{\ln 10}{10}\sigma'_p \right) dx
\end{aligned} \qquad (6.50)$$

Shadow Fading Quotient Ξ

For evaluating Ξ, the following expression shall be considered:

$$\delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \Xi_{c,p,s} = \delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \mathrm{E}\left\{ \frac{\omega_{jp}}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \qquad (6.51)$$

The quotient of the two shadow fading components may be rewritten:

$$\frac{\omega_{jp}}{\omega_{cp}} = \frac{10^{\sigma'_p X_p + \sigma''_p Y_{jp}}}{10^{\sigma'_p X_p + \sigma''_p Y_{cp}}} = \frac{10^{\sigma''_p Y_{jp}}}{10^{\sigma''_p Y_{cp}}} \tag{6.52}$$

Hence, the location dependent shadowing component X_p only impacts the coverage probability, but does not directly influence the quotient itself. Taking into consideration, that (6.51) is evaluated for the case that cell c is best server, Y_{jp} can only take values so that j does not become best serving cell. Hence, (6.51) is calculated as follows:

$$\begin{aligned}
&\delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \mathrm{E}\left\{ \frac{\omega_{jp}}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \,\cap\, \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,min} \right\} \\
&\quad = \int_{-\infty}^{+\infty} 10^{-\frac{\sigma''_p}{10} x} \cdot \mathrm{E}\left\{ 10^{\frac{\sigma''_p}{10} Y_{jp}} \;\middle|\; \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + Y_{jp} < \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} + x \right\} \\
&\qquad \cdot \mathrm{P}\left\{R_p\{\mathcal{S}_p\} = c \mid Y_{cp} = x\right\} \cdot \mathrm{P}\left\{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,min} \mid Y_{cp} = x\right\} \phi(x)\, dx
\end{aligned} \tag{6.53}$$

where

$$\begin{aligned}
&\mathrm{E}\left\{ 10^{\frac{\sigma''_p}{10} Y_{jp}} \;\middle|\; \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + \sigma''_p Y_{jp} < \left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} + \sigma''_p x \right\} \\
&\quad = \frac{\mathrm{E}\left\{ \mathrm{e}^{\frac{\ln 10}{10} \sigma''_p Y_{jp}} \;\middle|\; Y_{jp} < \frac{1}{\sigma''_p}\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \frac{1}{\sigma''_p}\left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + x \right\}}{\mathrm{P}\left\{ Y_{jp} < \frac{1}{\sigma''_p}\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \frac{1}{\sigma''_p}\left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + x \right\}} \\
&\qquad \text{subst.: } b = \frac{\ln 10}{10}\sigma''_p, \quad a = \frac{1}{\sigma''_p}\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}\right)_{[dB]} - \frac{1}{\sigma''_p}\left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}\right)_{[dB]} + x \\
&\quad = \frac{1}{\Phi(a)} \cdot \int_{-\infty}^{a} \mathrm{e}^{bx} \phi(x)\, dx = \mathrm{e}^{\frac{b^2}{2}} \cdot \frac{\Phi(a-b)}{\Phi(a)} \qquad \text{(cf. (6.47) and (6.48))} \\
&\quad = \mathrm{e}^{\frac{1}{2}\left(\frac{\ln 10}{10}\sigma''_p\right)^2} \cdot \frac{\Phi\left(\frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]} + x - \frac{\ln 10}{10}\sigma''_p \right)}{\Phi\left(\frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]} + x \right)}
\end{aligned} \tag{6.54}$$

and finally with (6.33) and (6.36) it yields

$$
\begin{aligned}
&\delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \Xi_{c,p,s} \\
&\quad = \delta_{c,p,s} \cdot \upsilon_{c,p,s} \cdot \mathrm{E}\left\{ \frac{\omega_{jp}}{\omega_{cp}} \;\middle|\; R_p\{\mathcal{S}_p\} = c \;\cap\; \hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp} > \hat{r}_{p,\,min} \right\} \\
&\quad = \int_{-\infty}^{+\infty} 10^{-\frac{\sigma''_p}{10} x} \cdot \left(1 - \Phi\left(\frac{1}{\sigma'_p} \left(\left(\frac{\hat{r}_{p,\,min}}{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}} \right)_{[dB]} - \sigma''_p x \right) \right) \right) \\
&\qquad \cdot \mathrm{e}^{\frac{1}{2}\left(\frac{\ln 10}{10}\sigma''_p\right)^2} \cdot \frac{\Phi\left(\frac{1}{\sigma''_p} \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}} \right)_{[dB]} + x - \frac{\ln 10}{10}\sigma''_p \right)}{\Phi\left(\frac{1}{\sigma''_p} \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}} \right)_{[dB]} + x \right)} \\
&\qquad \cdot \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \Phi\left(x + \frac{1}{\sigma''_p} \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}} \right)_{[dB]} \right) \phi(x)\, dx
\end{aligned}
\tag{6.55}
$$

6.3.2 Characteristic Cell Power Equations

The total power per cell is calculated equivalently to the derivation for the other two methods described in this chapter by summing up the average contributions of each of the pixels p:

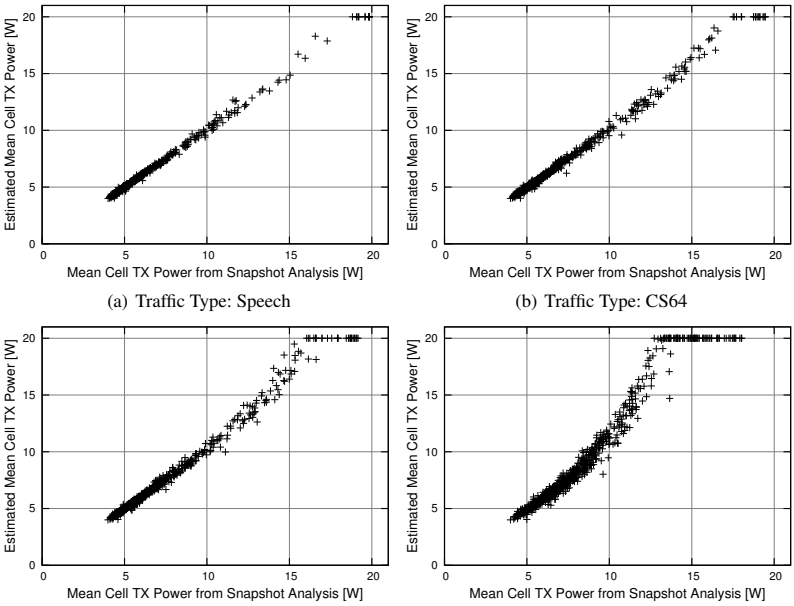
$$
p^{*\downarrow}_c = \sum_{s=1}^{S} \sum_{p=1}^{P} \theta_{p,s} \delta_{c,p,s} \upsilon_{c,p,s} t^{*\downarrow}_{c,p,s} + \tau_c \tag{6.56}
$$

$$
\begin{aligned}
&= \underbrace{\sum_{s=1}^{S} \sum_{p=1}^{P} \theta_{p,s} \delta_{c,p,s} \upsilon_{c,p,s} \zeta^{*\downarrow}_{c,p,s}}_{T^{*\downarrow}_{cc}} \cdot p^{*\downarrow}_c \\
&\quad + \sum_{\substack{j \in \mathcal{C} \\ j \neq c}} \underbrace{\sum_{s=1}^{S} \sum_{p=1}^{P} \theta_{p,s} \delta_{c,p,s} \upsilon_{c,p,s} \zeta^{*\downarrow}_{j,p,s} \frac{\xi^{*\downarrow}_{j,p,s}}{\xi^{*\downarrow}_{c,p,s}} \cdot \Xi_{c,p,s}}_{T^{*\downarrow}_{cj}} \cdot p^{*\downarrow}_j \\
&\quad + \underbrace{\tau_c + \sum_{s=1}^{S} \sum_{p=1}^{P} \theta_{p,s} \delta_{c,p,s} \upsilon_{c,p,s} \nu^{*\downarrow}_{p,s} \frac{1}{\xi^{*\downarrow}_{c,p,s}} \cdot \Psi_{c,p,s}}_{h^{*\downarrow}_c}
\end{aligned}
\tag{6.57}
$$

where $\delta_{c,p,s}$ and $\upsilon_{c,p,s}$ are given by equations (6.34) and (6.37), and $\Psi_{c,p,s}$ and $\Xi_{c,p,s}$ are calculated according to equations (6.55) and (6.50) respectively. From (6.57) a matrix equation analogues to (6.22) is obtained.

6.3.3 Application: Extended Statistical Load Estimation

Numerical results for the application of the method derived in this section on the test scenario considered in this chapter are presented in Figure 6.6 and Figure 6.7. In particular for the lower data rate traffic the correlation between snapshot results and analytical evaluation are extremely good. The correlation coefficient is well above 0.99 for the Speech and CS64 traffic types. Even for the worst case PS384 traffic, it is still almost 0.98.



(a) Traffic Type: Speech

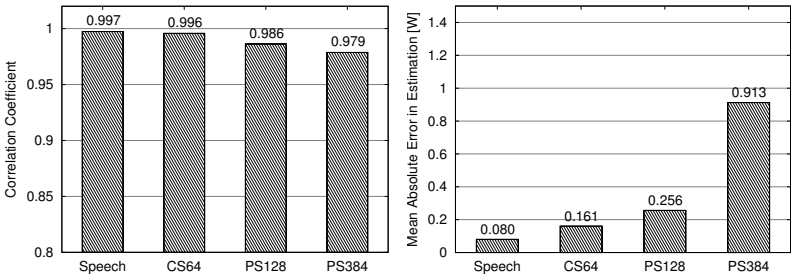
(b) Traffic Type: CS64

(c) Traffic Type: PS128

(d) Traffic Type: PS384

Figure 6.6: Correlation between the mean cell TX power obtained from extended statistical load estimation and snapshot analysis for different traffic types

The mean absolute error is also considerably reduced by the extended statistical load estimation. For the PS128 traffic, the mean absolute error is still below 0.26 W, which corresponds to a relative error of less than 5 %, as opposed to a relative error of greater than 10 % in case of the static method. From the correlation plots it can be obtained that the error for the higher data rate traffic is mainly due to the increasing tendency of underestimating the cell power the higher the service data rates. This statistical effect was already discussed in Section 6.2.4. However, it shall be noted here, that a WCDMA network is typically not designed for pure high data rate traffic, but for traffic mixtures. Hence, for typical applications of the method, extremely good estimates can be expected.



(a) Correlation Coefficient (r) (b) Mean Absolute Error (e)

Figure 6.7: Correlation coefficient and mean absolute error in cell power estimation for the extended statistical load estimation

6.4 Evaluation of Per Pixel Quantities

In the previous sections efficient analytical methods for estimating cell load levels have been derived. This constitutes the most challenging part in the performance assessment of WCDMA networks. In addition, for typical network planning and optimization applications various quantities per pixel are of interest. These include average pilot quality (E_c/I_0) and average required power levels as well as coverage probabilities per pixel.

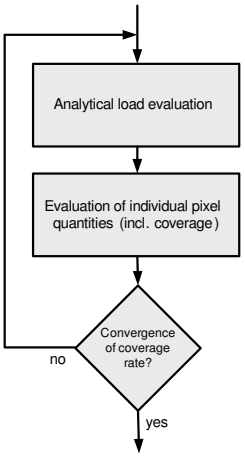


Figure 6.8: Evaluation loop to be carried out for detailed load and outage analysis

Coverage probabilities are of particular interest. In the cell load derivations in this chapter it is assumed that coverage is purely determined by a sufficient pilot reception level. In general, there are additional constraints that have an impact on the coverage, e.g. sufficient

pilot quality and required UL/DL connection power below threshold (cf. Section 5.4). The coverage probabilities however also impact the load estimation, as they reduce the offered traffic to be considered in the load estimation. In turn, the load estimates vice versa influence the coverage. Thus, if a very detailed load estimation is required and the outage rate is not as low as in the example scenario considered for the evaluations in this chapter (<1 %), the load estimation has to be carried out iteratively as presented in Figure 6.8. In each iteration first the cell loads are estimated taking the coverage rates per pixel evaluated in the previous iteration into consideration. Based on these load estimates an analysis per pixel is carried out, yielding new coverage estimates. The loop is continued until a sufficient convergence of the coverage rate is achieved. Typically, only few iterations are required: If the outage is very low, only one.

The estimation of the different quantities per pixel is not a trivial task. In the following the estimation of the most important of them is discussed.

6.4.1 Pilot RSCP

The received pilot power level, also widely referred to as *pilot RSCP (received signal code power)*, is already taken into account for calculating coverage probabilities in the statistical load estimation (Section 6.2.2). The calculation is exact, as no load or interference estimates need to be considered.

Also the average pilot RSCP of the best serving cell in a pixel can be calculated analytically:

$$\begin{aligned}
\mathrm{E}\left\{\hat{r}_{p,s}\right\} &= \mathrm{E}\left\{\hat{\tau}_{R_p\{\mathcal{S}_p\}}\xi^{\downarrow}_{R_p\{\mathcal{S}_p\}i(p,\,s)}\right\} \\
&= \sum_{c\in\mathcal{C}} \mathrm{P}\left\{R_p\{\mathcal{S}_p\}=c\right\}\cdot \mathrm{E}\left\{\hat{\tau}_c\xi^{\downarrow}_{ci(p,\,s)} \mid R_p\{\mathcal{S}_p\}=c\right\} \\
&= \sum_{c\in\mathcal{C}} \hat{\tau}_c\xi^{*\downarrow}_{c,p,s}\cdot \mathrm{P}\left\{R_p\{\mathcal{S}_p\}=c\right\}\cdot \mathrm{E}\left\{\omega_{cp} \mid R_p\{\mathcal{S}_p\}=c\right\}
\end{aligned} \tag{6.58}$$

where the addend in (6.58) can be evaluated in analogy to (6.50):

$$\begin{aligned}
&\mathrm{P}\left\{R_p\{\mathcal{S}_p\}=c\right\}\cdot \mathrm{E}\left\{\omega_{cp} \mid R_i\{\mathcal{S}_i\}=c\right\} \\
&\quad= \int_{-\infty}^{+\infty} 10^{\frac{\sigma''_p}{10}x}\cdot\phi\left(x\right)\prod_{\substack{j\in\mathcal{S}_p\\ j\neq c}}\Phi\left(x+\frac{1}{\sigma''_p}\left(\frac{\hat{\tau}_c\xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j\xi^{*\downarrow}_{j,p,s}}\right)_{[dB]}\right)\cdot \mathrm{e}^{\frac{1}{2}\left(\frac{\ln 10}{10}\right)^2\sigma'^2_p}dx
\end{aligned} \tag{6.59}$$

A typical quantity of interest used in network dimensioning is the fraction between mean received pilot level under the consideration of shadow fading and the static pilot reception level disregarding shadow fading:

$$g = \frac{\mathrm{E}\left\{\hat{r}_{p,s}\right\}}{\hat{\tau}_{R^*_p}\xi^{*\downarrow}_{R^*_p,p,s}} \tag{6.60}$$

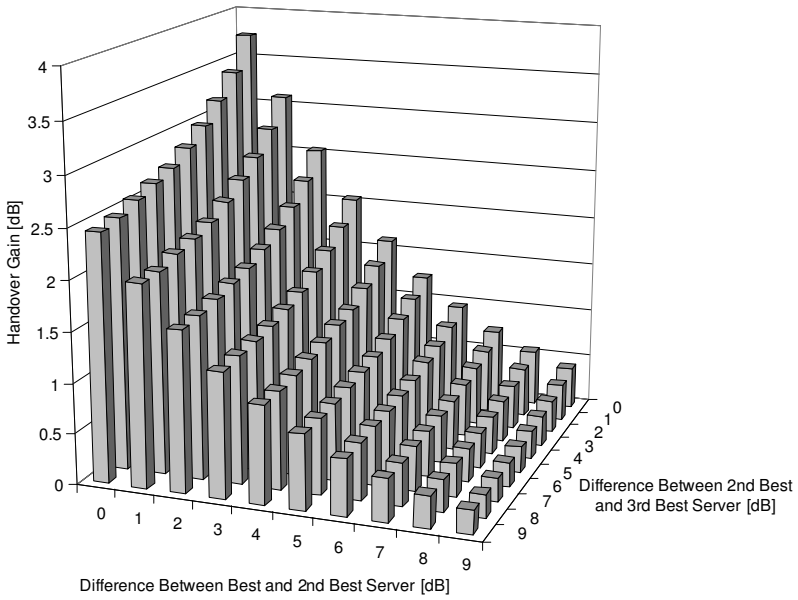


Figure 6.9: Shadowing gain (g) considering three potential servers for varying relative difference in average pilot reception level (σ=8 dB, ρ=0.5)

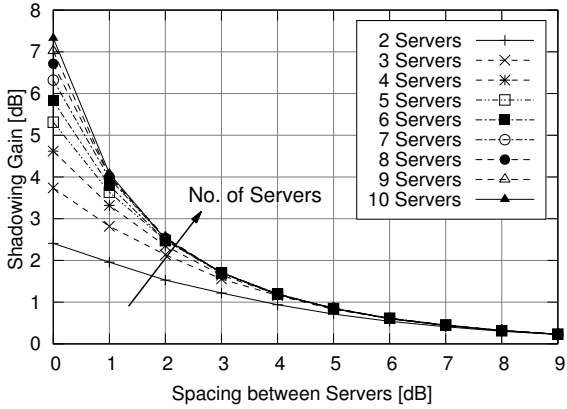


Figure 6.10: Shadowing gain (g) as a function of the number of potential servers assuming equal spacing of average pilot reception level between servers (σ=8 dB, ρ=0.5)

The quantity g is sometimes referred to as *shadowing gain*. It does not depend on the absolute value of the pilot reception levels, but on the relative difference between average received pilot levels. The shadowing gain considering three potential servers is depicted in Figure 6.9 for varying relative difference of mean pilot receive levels between the servers. Shadow fading parameters are chosen according to Table 6.2. In case all servers are equally strong, the gain can be as high as 3.7 dB. If all three servers are equally spaced by 9 dB the shadowing gain is decreased to about 0.2 dB. In Figure 6.10 g is evaluated for a varying amount of potential servers as a function of the spacing between received pilot levels. Equal spacing between the average received pilot levels of the potential servers is assumed. The shadowing gain increases with the number of servers, where the highest increase is achieved for a low number of servers.

6.4.2 Soft-Handover Probability

The evaluation of SHO probabilities can be carried out similarly to the calculation of coverage probabilities in Section 6.2.2. SHO probabilities are of interest for various applications, e.g. in network dimensioning and network optimization.

A user is considered to be in SHO if the difference in pilot power reception between best serving cell c and any other cell is below a certain margin, denoted SHO margin Δ_{SHO}:

$$\left(\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp}\right)_{[dB]} - \left(\hat{\tau}_j \xi^{*\downarrow}_{j,p,s} \omega_{jp}\right)_{[dB]} = \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s} \omega_{jp}}\right)_{[dB]} < \Delta_{\text{SHO}}$$

Accordingly, the total SHO probability for a user of service s in pixel p is one minus the probability that the difference in reception level between best serving cell and all other cells is greater Δ_{SHO}:

$$\delta^{\text{SHO}}_{p,s} = 1 - \text{P}\left\{ \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s} \omega_{jp}}\right)_{[dB]} > \Delta_{\text{SHO}} \; \forall j \in \mathcal{S}_p : j \neq c \right\} \tag{6.61}$$

The SHO relation in (6.61) can be rewritten:

$$\begin{aligned} & \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s} \omega_{cp}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s} \omega_{jp}}\right)_{[dB]} > \Delta_{\text{SHO}} \\ \Leftrightarrow \quad & \left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]} + \sigma''_p \left(Y_{cp} - Y_{jp}\right) > \Delta_{\text{SHO}} \\ \Leftrightarrow \quad & Y_{jp} < \frac{1}{\sigma''_p} \left(\left(\frac{\hat{\tau}_c \xi^{*\downarrow}_{c,p,s}}{\hat{\tau}_j \xi^{*\downarrow}_{j,p,s}}\right)_{[dB]} - \Delta_{\text{SHO}} \right) + Y_{cp} \end{aligned} \tag{6.62}$$

(6.62) in (6.61) yields:

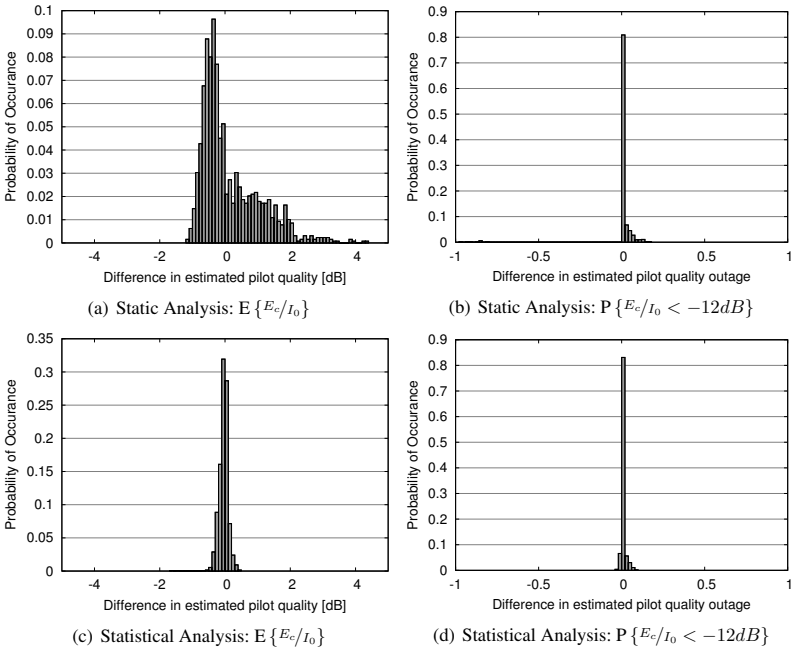
$$\begin{aligned}
\delta_{p,s}^{\text{SHO}} &= 1 - \mathrm{P}\left\{ Y_{jp} < \frac{1}{\sigma_p''} \left(\left(\frac{\hat{\tau}_c \xi_{c,p,s}^{*\downarrow}}{\hat{\tau}_j \xi_{j,p,s}^{*\downarrow}} \right)_{[dB]} - \Delta_{\text{SHO}} \right) + Y_{cp} \; \forall j \in \mathcal{S}_p : j \neq c \right\} \\
&= 1 - \int_{-\infty}^{+\infty} \mathrm{P}\left\{ Y_{jp} < x + \frac{1}{\sigma_p''} \left(\left(\frac{\hat{\tau}_c \xi_{c,p,s}^{*\downarrow}}{\hat{\tau}_j \xi_{j,p,s}^{*\downarrow}} \right)_{[dB]} - \Delta_{\text{SHO}} \right) \; \forall j \in \mathcal{S}_p : j \neq c \right\} \\
&\qquad \cdot \phi(x)\, dx \\
&= 1 - \int_{-\infty}^{+\infty} \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \mathrm{P}\left\{ Y_{jp} < x + \frac{1}{\sigma_p''} \left(\left(\frac{\hat{\tau}_c \xi_{c,p,s}^{*\downarrow}}{\hat{\tau}_j \xi_{j,p,s}^{*\downarrow}} \right)_{[dB]} - \Delta_{\text{SHO}} \right) \right\} \phi(x)\, dx \\
&= 1 - \int_{-\infty}^{+\infty} \prod_{\substack{j \in \mathcal{S}_p \\ j \neq c}} \Phi\left(x + \frac{1}{\sigma_p''} \left(\left(\frac{\hat{\tau}_c \xi_{c,p,s}^{*\downarrow}}{\hat{\tau}_j \xi_{j,p,s}^{*\downarrow}} \right)_{[dB]} - \Delta_{\text{SHO}} \right) \right) \phi(x)\, dx
\end{aligned} \tag{6.63}$$

6.4.3 Pilot Quality

The pilot quality, denoted $^{E_c}/_{I_0}$, is defined as the received pilot strength of the best server over the total received interference. Both, average $^{E_c}/_{I_0}$ and the probability that the $^{E_c}/_{I_0}$ is below a certain target level are vital parameters in network planning.

Disregarding shadow fading, the $^{E_c}/_{I_0}$ can be evaluated easily by considering static median link gains and constant cell power levels as given by the load estimation. When considering shadow fading, the evaluation is significantly more difficult. For the distribution function of $^{E_c}/_{I_0}$ no reasonable analytical approximation is known yet. The main problem is the correlation between E_c and I_0. Estimating the individual statistics of both quantities can be carried out quite well. Estimating the first has been discussed in the previous subsection. The total received interference I_0 is basically a sum of log-normal distributed random variables which can be approximated by a log-normal distribution [YS84, Lin92, Ho95, BADM95, BX04]. As no analytical approach is known, the proposed solution for the evaluation of the fraction of the both quantities is an experimental evaluation, i.e. a simple Monte-Carlo analysis. For this it is assumed that the variance of the DL cell power levels can be neglected compared to the significantly larger variance of the shadow fading component. Hence, the cell transmit power levels are assumed to be constants as given by the load estimation. The experimental Monte-Carlo analysis is carried out for the shadowing components of the different link gains only. Such an analysis is very fast, as it is carried out locally, individually for each pixel.

The estimation of the pilot quality per pixel has been evaluated for the example scenario considered in this chapter. The mean $^{E_c}/_{I_0}$ was estimated by a static analysis, considering only median link gains as well as by a experimental analysis as proposed above. For both analyses, the most accurate cell power estimates obtained from the extended statistical



(a) Static Analysis: $\mathrm{E}\{E_c/I_0\}$

(b) Static Analysis: $\mathrm{P}\{E_c/I_0 < -12dB\}$

(c) Statistical Analysis: $\mathrm{E}\{E_c/I_0\}$

(d) Statistical Analysis: $\mathrm{P}\{E_c/I_0 < -12dB\}$

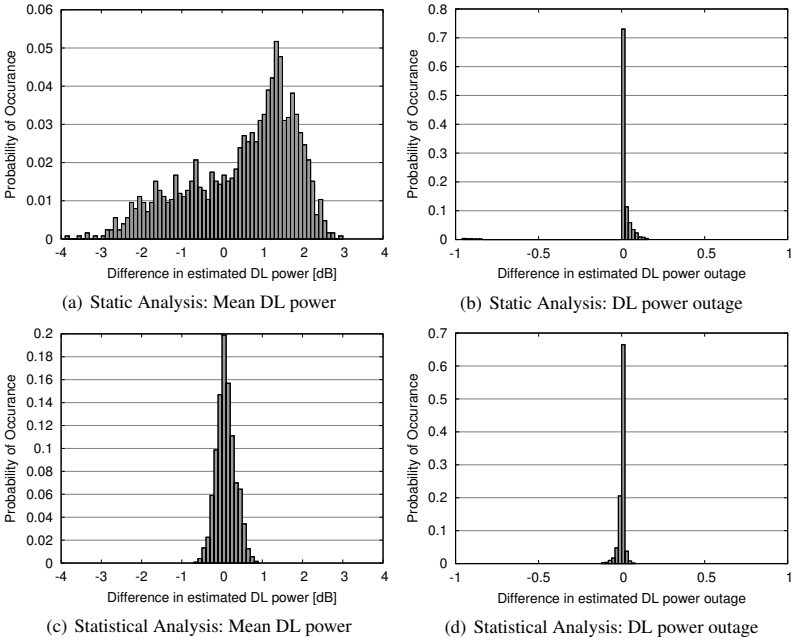
Figure 6.11: Comparison of the error (w.r.t. results from the snapshot analysis) in estimating mean E_c/I_0 and probability of E_c/I_0 below threshold of -12 dB per pixel for the static analysis and the statistical analysis. Scenario: "Metropolis", traffic type: CS64.

load analysis were used. The results are shown in Figure 6.11(a) and Figure 6.11(c). It is clearly visible that the consideration of shadow fading results in more accurate estimation.

The estimation of outage rates due to insufficient pilot quality has also been investigated. In order to trigger a considerable outage rate the required E_c/I_0 was increased to -12 dB (initially considered value -15 dB). The error in estimating the outage rate was evaluated for those pixels for that outage was obtained by either the snapshot analysis or the analyzed method. The resulting histograms in Figure 6.11(b) and Figure 6.11(d) do not clearly make the advantage of an application of the statistical method evident. Apart from the small probability of obtaining very high values of wrong estimates for the static estimation (difference of nearly -1), both methods at this level seem to estimate the outage rates almost equally well. This impression also arises when considering the overall estimated network outage due to pilot quality as gathered by the static and the analytical analysis in Figure 6.13(a) and Figure 6.13(c). The overall error compared with the snapshot analysis is fairly similar for both methods of estimation. However, the numerical evalua-

tion of the correlation coefficient between the outage values per pixel from estimation and reference value taken from the snapshot analysis (cf. Figure 6.13(b) and Figure 6.13(d)), the advantage of the statistical method is clearly made evident. The correlation is significantly higher for the statistical analysis (0.975) in comparison to the correlation coefficient obtained from the static analysis (0.498). The correlation is in particular of importance for network optimization, where typically relative improvements are to be achieved.

6.4.4 Downlink Connection Power



(a) Static Analysis: Mean DL power

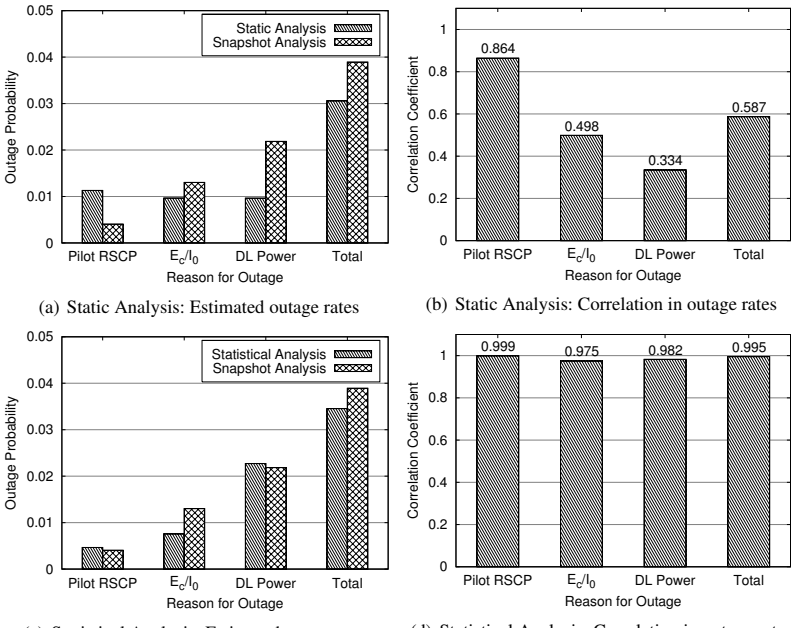
(b) Static Analysis: DL power outage

(c) Statistical Analysis: Mean DL power

(d) Statistical Analysis: DL power outage

Figure 6.12: Comparison of the error (w.r.t. results from the snapshot analysis) in estimating mean DL connection power and DL connection power outage ($\hat{t}^{\downarrow}_{max}$=33 dBm) per pixel for the static analysis and the statistical analysis. Scenario: "Metropolis", traffic type: CS64.

The evaluation of the mean required DL power for serving a user in a certain pixel can be estimated analytically by equations (6.16) for a static analysis and (6.39) for the statistical analysis. The advantage of the statistical method was already pointed out when presenting results on the cell power estimation in the first three sections of this chapter. The advantage of a statistical analysis for estimating DL connection power levels on pixel basis is even larger (cf. Figure 6.12(a) and Figure 6.12(c)).



(a) Static Analysis: Estimated outage rates

(b) Static Analysis: Correlation in outage rates

(c) Statistical Analysis: Estimated outage rates

(d) Statistical Analysis: Correlation in outage rates

Figure 6.13: Error and correlation (w.r.t. results from the snapshot analysis) in estimating overall outage rates for the "Metropolis" scenario based on the static and the statistical analysis for traffic type CS64 (Outage limits: $\hat{r}_{min}$=-115 dBm, $\hat{\gamma}_{min}$=-12 dB, $t^{\downarrow}_{max}$=33 dBm)

For the estimation of outage rates due to too high DL connection power levels a similar problem as for the estimation of $^{E_c}/_{I_0}$ values arises. Although the mean values can be well approximated analytically, the distribution function required for an estimation of outage rates is not known. Therefore, an experimental Monte-Carlo analysis as described above needs to be carried out also in this case. The example results shown in Figure 6.12(b) have been obtained by a straight forward application of a static method, purely considering median link gains and average cell power levels. The results from the experimental statistical analysis are presented in and Figure 6.12(d). For both analyses the allowed maximum DL connection power was reduced to 33 dBm (originally 36 dBm) in order to yield considerable outage. Similar to the pilot quality outage estimation, the advantage of the statistical analysis becomes obvious when considering the correlation coefficients in Figure 6.13(b) and Figure 6.13(d) respectively.

Parameter	Analytical Evaluation	Experimental Evaluation
Pilot RSCP Mean	•	
Pilot RSCP Outage	•	
Pilot E_c/I_0 Mean		•
Pilot E_c/I_0 Outage		•
DL Link Power Mean	•	
DL Link Power Outage		•
UL Link Power Mean	•	
UL Link Power Outage	•	
SHO Probability	•	

Table 6.3: Means for evaluating the most important quantities per pixel

6.4.5 Uplink Connection Power

This chapter focuses on the analysis of the DL. Nevertheless, as has been pointed out at the beginning of the chapter, the UL load can be estimated similarly to the DL. This in particular includes an analytical evaluation of the mean required UL power per pixel. In contrast to the DL connection power outage, the UL connection outage can also be estimated by means of an analytical evaluation. The required UL connection power depends on the UL interference level at the serving cell(s) and the respective link gains. The mean interference level is given by the UL cell load estimation. The variance of the UL cell load can be neglected with respect to the variance of the shadow fading components. Moreover, the UL interference level is approximately independent of the link gain of a particular mobile. Due to this independence, which is not given for the DL interference, the UL power distribution and thus the outage rates can be well approximated analytically.

6.5 Practical Implementation

As long as shadow fading is disregarded, i.e. for the static methods, the implementation of the performance analysis is straight forward. The calculations involved only require basic operations. This does not hold for the statistical analysis, which considers shadow fading statistics.

Firstly, although already simplified as much as possible, the calculation of most coefficients and quantities derived in this chapter require a numerical integration. For a full evaluation several numerical integrations would in principle need to be carried out per pixel. Obviously this is infeasible for an application in the everyday planning process. In fact, the computational complexity would likely not be significantly lower than the computational complexity of a snapshot analysis. Fortunately, most of the quantities and coefficients to be determined only depend on the relative difference of the median link gains of the potential servers (and the relative difference w.r.t. the minimum required sig-

nal strength). The computational complexity of the evaluation can thus be dramatically reduced by pre-processing the values and using look-up tables within the actual calculation. Of course, depending on the desired accuracy, the amount of required memory for these look-up tables can become very large. However, in network planning typically the time required to yield certain results is a more valuable resource than the required memory for these operations.

Secondly, as described above and summarized in Table 6.3 not for all quantities required at pixel level a good analytical approximation is known. For these, either a crude approximation can be applied, e.g. based on the results from the static analysis, or an experimental analysis needs to be carried out. The choice among these options depends on the particular application. If the proposed experimental analysis is applied, the number of Monte-Carlo trials needs to be decided on very carefully. In fact, it might be beneficial to decide on the number of trials dynamically depending on the relevance of the particular pixels.

6.6 Conclusions

In particular automatic network optimization techniques require fast methods for network performance assessment. The most challenging part in the performance analysis of WCDMA networks is the estimation of cell load levels. In this chapter methods for estimating the network load of WCDMA networks have been presented and numerically compared. The current state-of-the-art approach (in scientific publications) for estimating the cell load is a static analysis of the network, which does not consider shadow fading statistics in the load estimation. The static method results in considerable inaccuracies when comparing the results with the reference results obtained from snapshot simulations. This deficiency is mainly due to the disregard of shadow fading.

Two new methods, which extend the static method by considering the statistical effects of shadow fading on the cell load, have been developed and proposed. The methods significantly improve the accuracy of the cell load estimation. If implemented efficiently, the methods can provide very good results in terms of estimated cell load levels with low computational complexity.

Based on the estimated cell load levels, further quantities of interest can be estimated on pixel level. For this it is also important to consider shadow fading. Unfortunately, not all of the quantities of interest can be estimated analytically (cf. Table 6.3). For the quantities which do not allow an analytical evaluation, it is proposed to carry out a simple experimental (Monte-Carlo) analysis on pixel level. The results presented in this chapter show that the resulting analytical method very well approximates the results from snapshot simulations.

Chapter 7

Automated Network Optimization

In the previous chapters different performance evaluation methods for WCDMA networks have been presented. The methods differ in the level of detail in which the network is analyzed and as a consequence in the computational complexity. In this chapter the application of the methods, in particular the analytical methods described in Chapter 6, for an automated network optimization is presented.

After a short introduction to optimization parameters and targets, two different approaches for an automated radio network design and optimization are presented. The first approach is a path loss based approach that applies fast heuristics. It used is in particular for very fast initial network planning and in situations, where not the complete set of input data is available, e.g. only site locations. In the scope of this thesis it is used to generate a good starting point for the second more detailed optimization technique. The second approach is a (local) search based approach that uses search heuristics for optimizing the network. An inherent part of this is a fast network performance evaluation based on the models described in Chapter 6. Both methods have also been presented by the author in [TK05].

7.1 Optimization Parameters and Targets

The targets of the radio network optimization are mainly twofold. The first target is to minimize the interference caused by the individual cells, while a sufficient coverage over the planning area is maintained. This in general is a trade-off and needs to be balanced, e.g. tilting down the antenna causes lower coverage, but also lower interference in neighboring cells and thus a potentially higher network capacity. The second target is the traffic distribution between cells. It is desirable to maintain similar cell loading of neighboring cells in order to minimize blocking probabilities and to maximize spare capacity for traffic fluctuations and future traffic evolution.

The most effective parameter in network optimization is the antenna tilt. Antenna tilts need to be set so that the traffic within the “own” cell is served with maximum link gain,

but at the same time the interference in neighboring cells is minimized. The possible tilt angles are typically restricted because of technical and civil engineering reasons. Especially in case of collocated sites with multi-band antennas, strong restrictions on the possible tilt angles to be taken into account during optimization could occur.

The transmitted pilot channel power and the other common channel powers, which are typically coupled by a fixed offset, are also vital parameters of network optimization. It needs to be assured that these channels are received with sufficient quality by all users in the serving cell. At the same time a minimization of the common channel powers yields significant capacity gains: Firstly, additional power becomes available for other (user traffic) channels, and secondly, interference is reduced. The gains obtained from reducing the pilot power are often underestimated. It is important to note that in a capacity limited WCDMA network (e.g. in urban areas) the reduction of pilot power levels by a certain factor also reduces the total transmit power of cells and as a consequence the cell loading by up to the same factor.

Optimization of azimuth angles of sectored sites is also of great importance in particular in case of antennas with rather small horizontal beam-width (e.g. 65 degree vs. 90 degree in case of three-sectored sites). In this case the difference between antenna gains in direction of the main lobe and the half-angle between neighboring sectors is comparatively large, and cells of neighboring sites might need to be adjusted so that maximum coverage is achieved. It is observed that during optimization azimuth changes are in particular introduced in order to reduce coverage problems. For possible azimuth angles typically even stronger restrictions apply than for the tilt angles.

The antenna height is also often a degree of freedom for the optimization. Higher antennas can provide better coverage, but also cause more interference in neighboring cells. Additional important parameters are the antenna type and the number of deployed sectors at a site. Both parameters are closely coupled, as a larger number of sectors also suggests the use of an antenna pattern with a smaller horizontal beam-width. The choice of sectorization is typically a trade-off between increased network capacity and higher monetary cost.

7.2 Optimization based on Fast Heuristic

Network planning requires an optimization to be carried out in a very short time frame. The method presented in this section allows a complete network configuration including tilt, azimuth, and pilot power selection within a few minutes even for very large networks ($N \times 1000$ sites). It builds on a path loss based network analysis, in the following referred to as *ray analysis*.

Ray Analysis

The ray analysis constructs "rays" from the cell origin to the (potential) cell border of each cell. Along the individual rays the total path loss is the macro path loss as given by a simple prediction model (e.g. Cost 231-Hata model), superimposed with the antenna gain. The use of a simple prediction model guarantees "well defined" cell areas but in particular neglects the influence of height data, i.e. diffraction losses. The terminus of a ray, i.e. the cell border, is defined as the first point starting from the cell origin to which another cell has the same total path loss as the path loss given by the ray. The only degrees of freedom in the ray analysis are the tilt angles. A simple heuristic for selecting tilt angles based on the cell radius R, i.e. the length of the ray in direction of the main antenna beam, is applied. A commonly used simple rule for setting the antenna tilt φ based on the cell radius R is:

$$\varphi = \arctan\left(\frac{h_{BS} - h_{MS}}{R}\right) + \varepsilon\vartheta \tag{7.1}$$

where h_{BS} and h_{MS} are the heights of the BS antenna and the MS antenna respectively, and ϑ is the vertical opening angle of the antenna. The scaling parameter ε is typically chosen between 0.2 and 1.

In order to obtain similar cell sizes in the vicinity of a cell i, the tilt angles of all neighbor cells are during the ray analysis of i assumed to be chosen so that the radius of those cells equals the radius of cell i. The ray analysis of cell i starts by assuming a low (potential) cell radius R' and according tilt angle. The potential cell radius R' is increased until it matches the cell radius R obtained from the ray analysis. In a practical implementation this process can of course be sped up by the application of a more efficient search, e.g. a binary or Fibonacci search [CLRS01].

In Figure 7.1 the results of the ray analysis are shown for an example network. It consists of more than 900 cells and is used for illustration purposes throughout this chapter. The planning area is restricted by a bounding polygon.

Tilt Optimization

In the ray analysis tilt angles are chosen based on a simple rule, e.g. (7.1). The tilt angles are used to find a suitable cell layout, but are typically not optimal in terms of provided network coverage and capacity. Thus, in a second step tilt angles are adjusted based on heuristics which take additional characteristic parameters of the cells into account: antenna height, main interferers, etc. The heuristics were derived from extensive simulations and have been verified with measurements from live networks. Obviously, cell areas might after the second step (slightly) differ from the ones obtained during ray analysis.

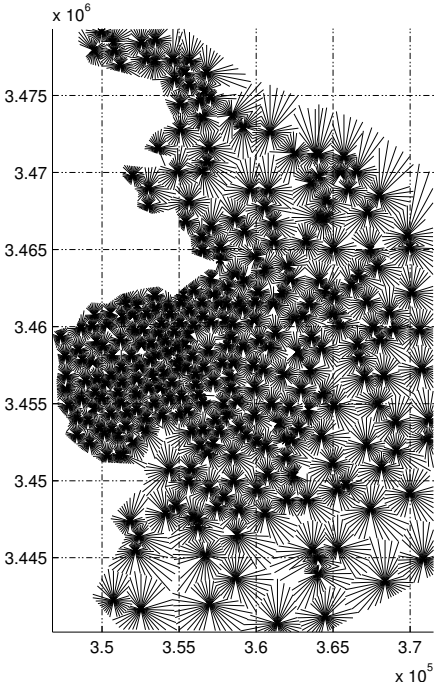
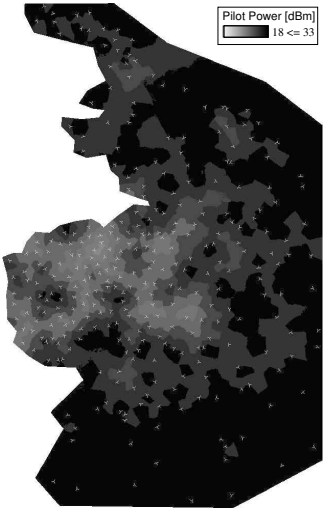


Figure 7.1: Illustration of "ray analysis"

Pilot Optimization

The pilot optimization is carried out prior to the tilt optimization. The objective of the pilot optimization is to minimize the common channel power levels while maintaining a minimum pilot reception (RSCP) level at the cell border. Pilot power levels are reduced successively for all cells until the required RSCP level cannot be achieved with the desired probability anymore. Large differences between pilot levels of neighbor cells might not be acceptable as they can for example degrade the SHO performance. Thus, an additional parameter, in the following referred to as *Maximum Neighbor Offset Factor* (NOF), is taken into account. The NOF controls the maximum difference between pilot levels of neighboring cells. In Figure 7.2 the results of the pilot power optimization are depicted for three different settings of NOF. The (idealized) best server areas of the individual cells are colored according to their assigned pilot power levels. In Figure 7.2(a) the results of the pilot power optimization are depicted for an unrestricted NOF. A quite heterogeneous setting of pilot power levels is observed. In contrast, the plots for restrictions of NOF =1 dB (Figure 7.2(b) and NOF =2 dB (Figure 7.2(c) show a smooth gradient from dense-urban with small cell sizes to rural areas with large cell sizes.

(a) NOF unrestricted



(b) NOF=1dB

(c) NOF=2dB

Figure 7.2: Results of the pilot power optimization for different settings for the maximum difference in pilot power between neighboring cells (Neighbor Offset Factor (NOF))

Azimuth Optimization

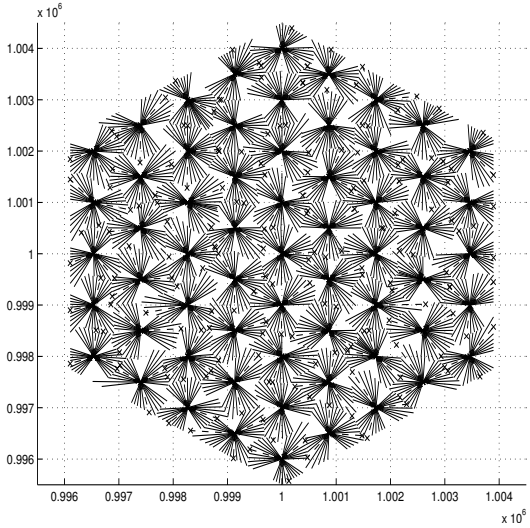
While tilt optimization and pilot optimization can be carried out in extremely short time (typically few seconds for even very large networks), the azimuth optimization is slightly more complex. For the azimuth optimization a greedy heuristic is applied, which successively changes the azimuth angles for each cell and selects the "best" azimuth based on a cost function. The cost function is calculated based on the ray analysis. The cost function itself is a weighted combination of cell size and pilot quality (Ec/I0) at the cell border. In this way in particular the network coverage is improved. In Figure 7.3 results obtained from the azimuth optimization for a hexagonal cell layout with 3 sectored sites are presented. Antennas with 65 degree horizontal beam-width are used. For this scenario the optimum cell layout is well known. Starting from a random setting of azimuth angles the algorithm finds the constellation known to be optimal within less than 3 minutes (standard PC) and 4 iterations over all cells. Only some cells in the outermost ring differ a few degrees from their optimum, which is due to border effects. The considered area is restricted by a bounding hexagon.

7.3 Search Based Optimization

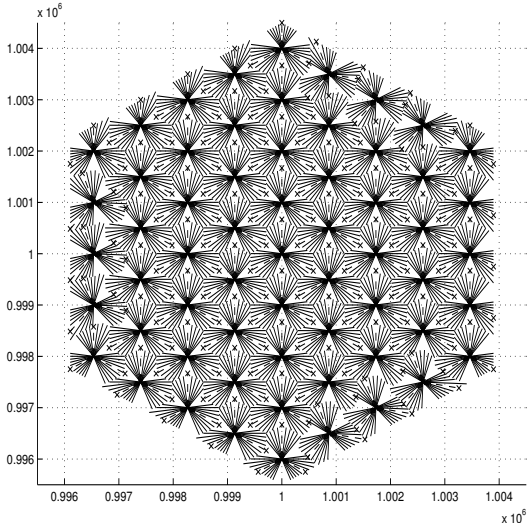
The optimization method described in the previous section is well suited for an initial planning and in cases where the level of completeness or accuracy of input data is limited. For a detailed optimization that takes the full set of input data into account, a *search approach* is proposed. That is, the space of possible configurations denoted as *search space* is explored in order to find the point in the search space that is optimal with respect to a certain criterion. This point is denoted as *global optimum*. *Exhaustive search* traverses the complete search space in a systematic manner. As all points in the search space are visited, with exhaustive search it is guaranteed to find the global optimum. The search space is very large for typical applications in network planning. For each site there can easily be several hundreds of possible configurations. Furthermore, configurations of different sites cannot be considered independently, so that the amount of possible network configurations grows exponentially with the number of sites. Hence targeting this area, an exhaustive search is too time consuming and *local search algorithms* are commonly used for network optimization purposes.

7.3.1 A Brief Introduction to Local Search

Local search algorithms start at some point in the search space denoted as *initial solution* and subsequently move from the present to neighboring solutions, if they fulfill some criterion, i.e. appear to be better or more promising. Local search algorithms cannot guarantee to find the global optimum. The objective of the search algorithm – developed or tailored for a particular problem – is to find a solution that is at least close to the global



(a) before optimization



(b) after optimization

Figure 7.3: Example application of fast heuristics for azimuth optimization

optimum. Local search algorithms are thus often classified as *heuristics*, where a heuristic is in [Ree95] defined as:

"A heuristic is a technique which seeks good (i.e. near-optimal) solutions at a reasonable computational cost without being able to guarantee either feasibility or optimality, or even in many cases to state how close to optimality a particular feasible solution is."

The basic procedure of local search is independent of the actual search algorithm applied. Starting from the initial solution in each search step first a *search neighborhood* is generated. The search neighborhood is a sub-set of points from the search space that are "close" to the current solution, i.e. that have some attributes in common with the current solution. The point that is most appropriate with respect to some criterion is selected from the search neighborhood and accepted as the new initial solution for the next search step. If no appropriate new solution is found (or some other stop criterion is fulfilled) the search is terminated. The comparison of points from the search space is carried out by means of cost values associated with them. The cost values are generated from a *cost function*, in the literature often also referred to as *objective function*. The objective function maps a given point in the search space to a cost value. The cost value can be a scalar but could also be represented by a vector. In the latter case, an appropriate function to compare cost values needs to be defined.

Local search algorithms are very much application specific. However, several search paradigms have been developed in the last three decades. The simplest search paradigm is the *descent method*. This method always selects the solution from the neighborhood that has the lowest cost. If this value is lower than the lowest value in the last search step, the solution is accepted as a new solution, otherwise the algorithm is terminated. The algorithm consequently explores the search space by always moving in the direction of the greatest improvement, so that it typically gets trapped in a *local minimum*. A local minimum is a solution that is optimal with respect to the search neighborhood but which generally is worse than the global optimum. In order to escape from local minima, among several others, one widely applied approach is to carry out restarts, that is, the local search is restarted from a new solution that is selected from a different area of the search space. The new start solutions are often selected randomly. If restarts are applied, the algorithm strictly speaking is not a local search algorithm anymore.

Another option for escaping from local minima is to also accept cost deteriorating neighbors under certain conditions. The most prominent local search paradigms that apply this strategy are *simulated annealing* and *tabu search* [Ree95, AL97].

Simulated annealing is based on an analogy with the physical annealing process. In simulated annealing improving points from the neighborhood are always selected when exploring the search neighborhood, non-improving points are accepted as new solutions with a certain probability. The probability of acceptance is a function of the level of deterioration, but also gradually decreases during the algorithm execution. The reduction of the probability of acceptance is determined by the *cooling scheme* [AL97].

In contrast to the simulated annealing which comprises randomness, classical tabu search is deterministic. The basic operation is equivalent to the descent method with the difference that the best point from the neighborhood is also accepted if it is worse than the current solution. In this way the search is directed away from local minima. In order to avoid a move back to already visited solutions, a *tabu list* is introduced. The tabu list typically contains sets of attributes of solutions that have already been visited. If a point from the neighborhood exhibits one of the sets of attributes stored in the tabu list, the point is only accepted as new solution if its quality, i.e. cost, exceeds a certain *aspiration level*. The tabu list is updated, keeping the individual entries only for a number of iterations. The size of the tabu list is a very important design parameter of the tabu search, it in particular needs to be chosen large enough to prevent cycling, but a too large list might introduce too many restrictions. Several enhancements to the basic operation of the tabu search have been introduced, most of which modify the handling of the tabu list. These include intensification and diversification schemes [AL97].

7.3.2 Application of Local Search to Radio Network Planning

The basic structure of the local search algorithm that has been developed for an optimization of WCDMA networks is depicted in Figure 7.4. The algorithm comprises the basic elements of a local search method that have been presented in the previous section. The local search starts from an initial solution, which for example can be the current configuration of the network to be optimized, a manually planned solution, or a solution suggested by the fast heuristics presented in the previous section.

At the beginning of each search step the search neighborhood is generated. At first a cluster of cells is selected for which parameter changes are considered. Based on the selected cells the search neighborhood is generated and explored to yield a new solution.

The quality of a certain solution is assessed by a performance analysis, which applies either of the methods described in Chapter 6. The choice of the method depends on the particular application and is a trade-off between accuracy and speed of the optimization process. From the results of the performance analysis, a cost value is generated by means of a cost function. The cost function is basically a linear combination of the evaluated quantities. In addition, penalty components are added if certain thresholds are exceeded for any of the different cost function components (e.g. coverage probability below some design target). The search process can be guided by appropriately setting weights for the different cost function components.

As search paradigm either a descend method or a tabu search can be applied. The tabu list is maintained independent from the selected search paradigm. The search paradigm only influences the way in which the search is terminated. In case of the descend method the search process is terminated if no improvement could be found. For the tabu search, non-improving moves are accepted in order to escape from local minima. The tabu search is terminated once no improving moves were found for a certain number of search steps.

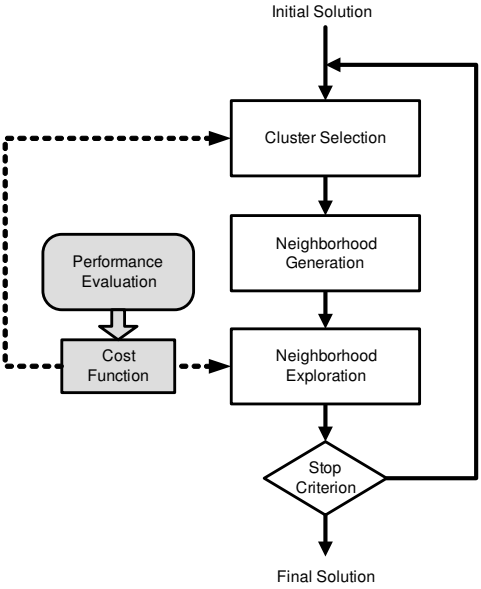


Figure 7.4: Basic local search algorithm

Cluster Selection

The considered networks are typically very large. It is thus beneficial to limit the number of cells for which changes are considered in one search step. This process is called *cluster selection*. The cluster selection only impacts the neighborhood generation, for the performance evaluation always the complete network is considered. The cluster selection uses information on the performance of the cells given by cost function and performance evaluation to identify the cluster of cells with worst performance. These are considered at first in the optimization.

In a new search step a new cluster is selected if either of the following conditions is fulfilled:

1. A new cluster has become the cluster with the worst performance
2. There was no improving move found for the current cluster, respectively for a number of search steps in case of tabu search

The size of a cluster is again a trade-off. If chosen too small, the algorithm might fail to sufficiently improve the performance in a certain problem area, because not all relevant cells might be considered in the neighborhood. On the other hand, a large cluster also impacts the time required to explore a neighborhood.

Neighborhood Generation

During the development of the algorithm it has proven to be efficient to organize the search neighborhood as a search tree. This is due to the observation that some configuration changes typically entail subsequent changes of different parameters of the same cell to yield good configurations. For example, if the antenna azimuth is changed, typically also tilt changes are beneficial. Also, it is for instance useful to first change the tilt, before pilot power levels are adapted. On the first level of the constructed tree there are the cells that are subject to a change, i.e. the cells from the selected cluster. Each of the other levels of the search tree contain changes of one type of parameter. If following the example given above the azimuth and the tilt can be changed, the second level of the tree contains the azimuth modifications as children of the according cell (cf. Figure 7.5). Children of each of these nodes are the possible tilt angle modifications for the same cell. The depth of the tree increases with the number of parameters that can be changed.

Neighborhood Exploration

The neighborhood is evaluated by systematically exploring the neighborhood tree. Each leaf of the tree represents one potential network configuration, and each node on the path from the root node to the leaf represents one parameter change to be introduced to yield the candidate solution.

If the degrees of freedom per cell are very large, the tree can also become very large. Hence, the children of nodes are sorted. If for example the children represent tilt changes, the exploration is carried out in an ordered sequence of tilt angles. That is, solutions are evaluated first in descending and then in ascending order of tilt angles (or vice versa) with respect to the tilt angle in the current solution. This is illustrated in Figure 7.5. The advantage of this procedure is that the exploration of such a sequence can be interrupted if only ascending cost values are found for a certain number of angles, as it can be assumed that in such a case the probability of finding an improving move is quite low. Following this strategy, the number of configurations to be evaluated can be dramatically reduced. The experience shows that in many cases the number of evaluations can be reduced by up to 90%.

7.3.3 Advanced Search Algorithm

The performance of the local search method strongly depends on the applied performance evaluation. The methods presented in Chapter 6 are all good candidates for an application. The choice between them is a trade-off between accuracy and running time. The basic static method is the fastest but as was shown has some weaknesses in terms of accuracy of results. The presented statistical methods significantly outperform the latter method in accuracy of results but even if implemented efficiently are of higher computa-

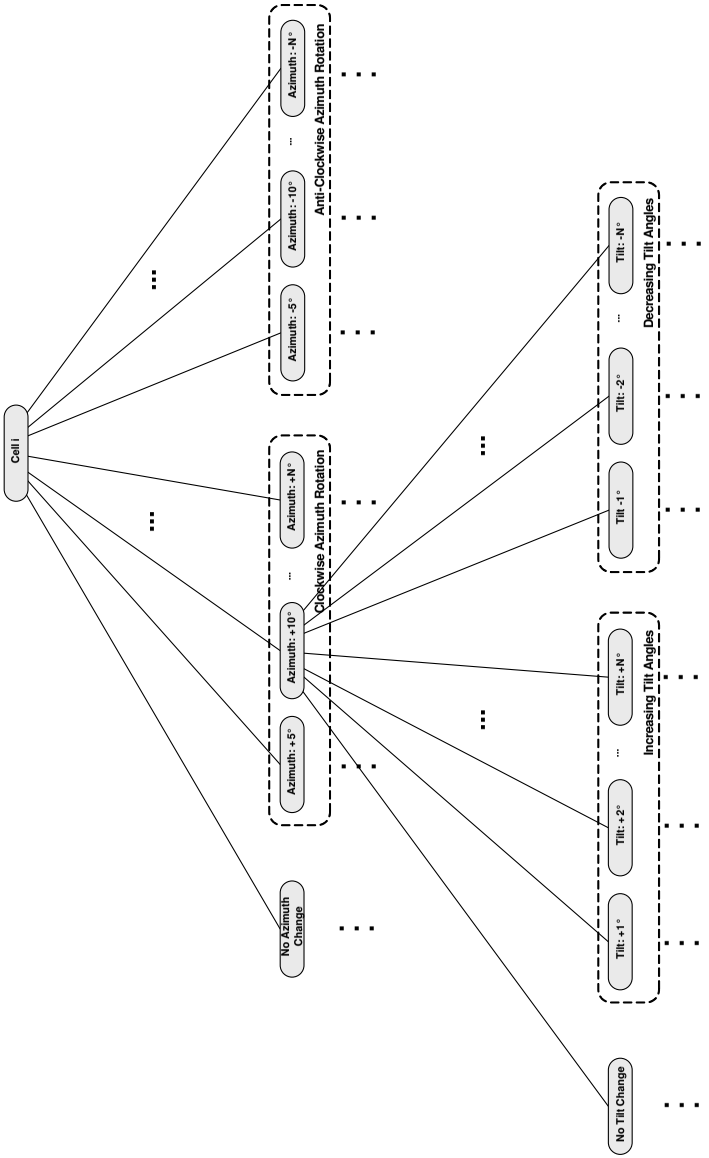


Figure 7.5: Illustration of the neighborhood tree

tional complexity, especially if the experimental analysis is applied for evaluating quantities per pixel (cf. Section 6.4). The local search optimization presented in this section can be extended to yield a hybrid method which makes use of two methods for performance evaluation (cf. Figure 7.6). The exploration of the neighborhood is split into two parts. In a first step the neighborhood is explored by the use of a simple and fast *basic performance evaluation* method. As a result a list of candidate solution is generated. The list is sorted with respect to cost values. The next candidate solution is selected from this list using a more accurate but also more time consuming *advanced performance evaluation*. Either the first improving solution from the list or the best solution from a subset of most promising solutions ("short list") is selected.

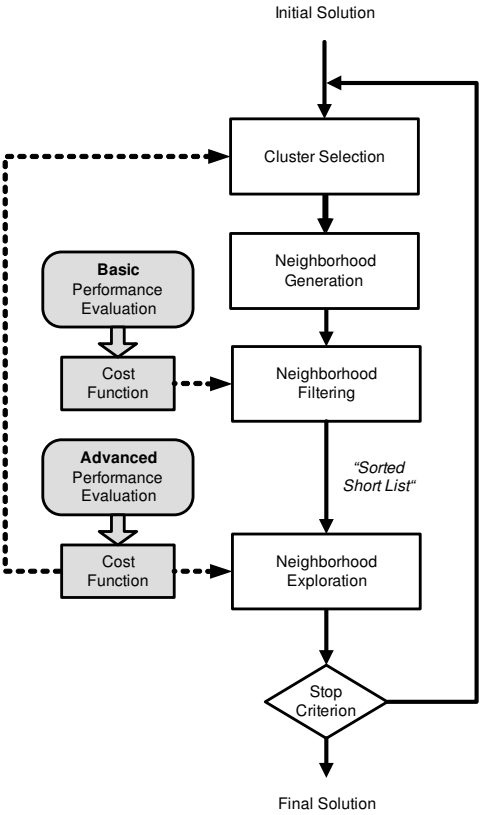


Figure 7.6: Hybrid local search algorithm

7.4 Application Example

In the following, results from an application of the developed methods to a realistic planning scenario are presented. The considered UMTS network located in a large Asian city consists of more than 900 cells. The initial antenna configuration is taken from the existing co-located GSM network. The original GSM antennas were replaced by standard UMTS antennas.

Different optimization scenarios have been studied:

1. Optimization of tilt angles
2. Optimization of tilt and azimuth angles
3. Optimization of tilt angles and pilot power levels
4. Optimization of tilt angles, pilot power levels, and azimuth angles

The optimization was carried out by first applying the fast heuristics presented in Section 7.2. The result was taken as input for an application of the local search optimization.

All optimization results have been evaluated by means of snapshot simulations and compared with the results for the original network. A high load traffic scenario, comprising a traffic mix with an average of about 20,000 active users demanding both PS and CS services, is considered. In order to guarantee a sufficient reliability of results for each scenario, more than 2000 snapshots have been analyzed.

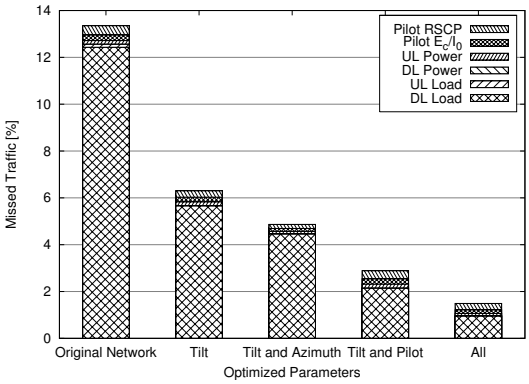


Figure 7.7: Comparison of missed traffic per reason before and after optimization

In Figure 7.7 the unserved traffic for the four different optimization scenarios is compared to the original network. The tilt optimization already halved the unserved traffic. The complete optimization results in a reduction of unserved traffic from more than 13% to about 1.4%.

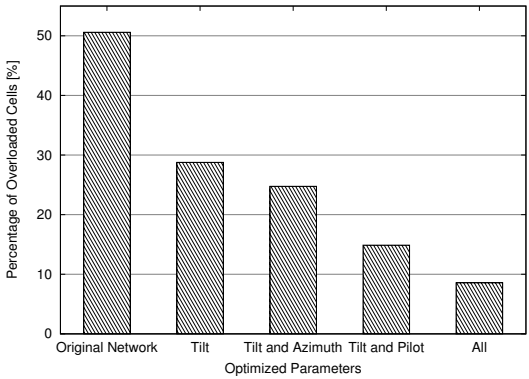


Figure 7.8: Comparison of percentage of cells with more than 2% missed traffic before and after optimization

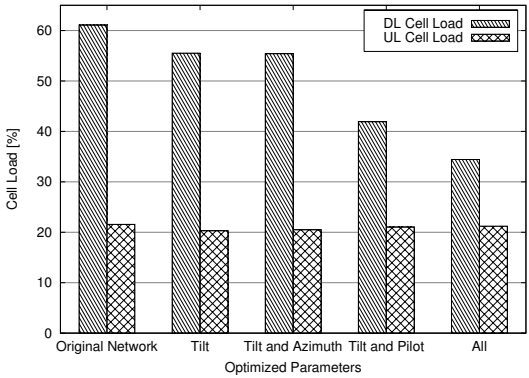


Figure 7.9: Comparison of average cell load before and after optimization

The same tendency as for the unserved traffic can be observed for the percentage of overloaded cells (Figure 7.8). An overloaded cell is here defined as a cell for which a blocking probability of more than a threshold of 2% is observed. The fraction of overloaded cells is reduced from more than 50% to less than 10%.

The downlink cell load is also significantly reduced in all optimized scenarios, although the amount of served traffic is increased at the same time (cf. Figure 7.9). As expected the pilot power optimization has the greatest impact on the DL cell load. An interesting observation is the slight increase in UL cell load indicating the general phenomena that a reduction of the DL cell load not necessarily also results in a reduction of the UL cell load and vice versa. The latter observation gives rise to the fact that during optimization, both UL and DL, need to be considered. However, in the study presented here, the UL

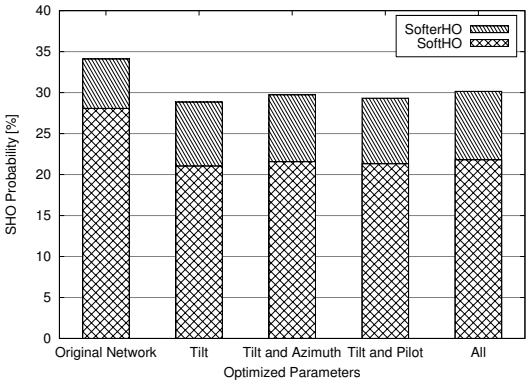


Figure 7.10: Comparison of SHO probability before and after optimization

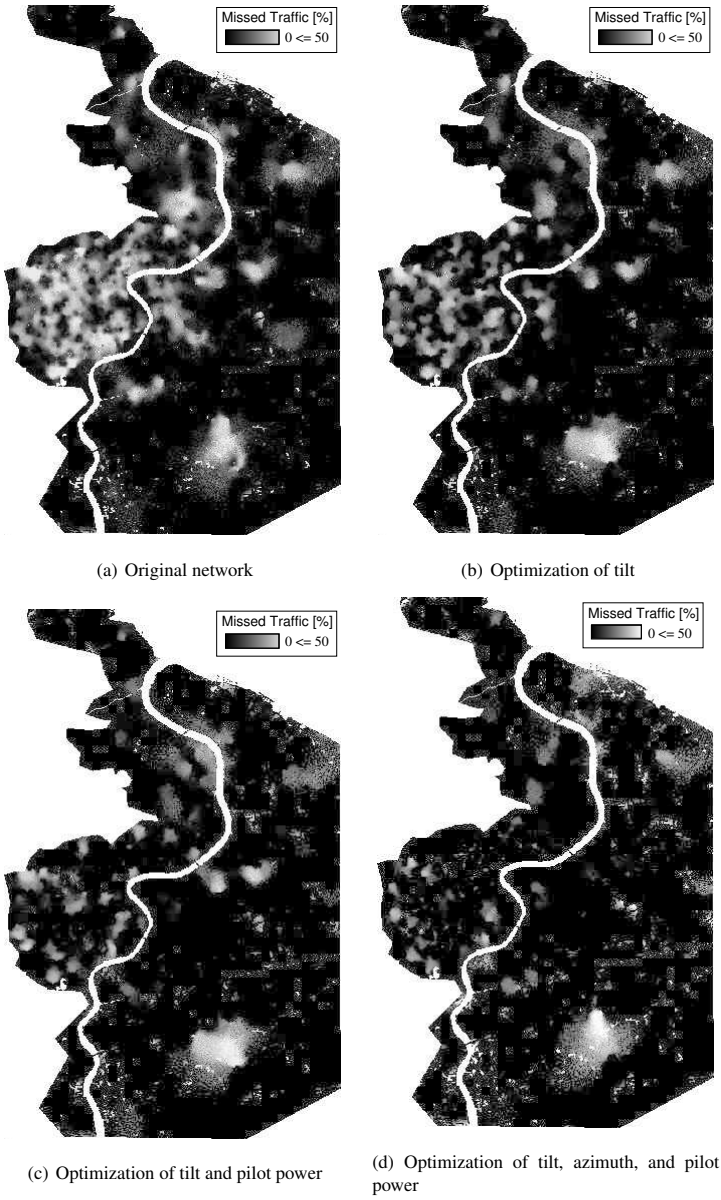
cell load is low (<25%) and moreover the increase in UL load is lower than the increase in carried traffic due to decreased blocking rates.

As a result of the optimization the SHO probability is also reduced from about 35% to about 30% (cf. Figure 7.10). It is interesting to note that although the SHO probability was not an objective of the optimization, the optimization also resulted in SHO rates expected for a decent network configuration.

Finally, in Figure 7.11 the spatial distribution of missed traffic is illustrated. The planning area is restricted to a polygon, outside of which no traffic is distributed. On the river dividing the area also no traffic is spread. The black areas indicate locations where traffic is distributed and served, gray areas indicate high blocking rates. It is observed that the optimization in particular was able to significantly reduce the missed traffic in the urban and dense-urban areas (gray dotted regions in Figure 7.11(a)). The problem area in the lower part of the planning area which is still persistent in the plot representing the fully optimized network (Figure 7.11(d)) is due to a missing site, which would need to be installed.

7.5 Conclusions

In this chapter two different methods for an automated optimization of WCDMA networks have been presented. The first is a very fast method which allows an optimization of very large networks in a very short time frame (within minutes). The method bases on a path loss analysis and applies heuristics for the selection of the most relevant cell configuration parameters, i.e. tilt, azimuth, and pilot power. The result of the fast method can in particular be used as an initial solution for the second method, which carries out a detailed local search analysis that takes all relevant planning parameters into account.



(a) Original network

(b) Optimization of tilt

(c) Optimization of tilt and pilot power

(d) Optimization of tilt, azimuth, and pilot power

Figure 7.11: Graphical illustration of the probability of missed traffic for the original network compared with three optimized networks

The most crucial design parameter of a local search optimization is the accuracy and time efficiency of the network performance evaluation that is applied for the comparison of different configurations. It is proposed to use the analytical methods developed and discussed in Chapter 6. In addition to the performance evaluation, the generation and exploration of the search neighborhood is also of significant importance. It is proposed to efficiently organize the search neighborhood by means of a search tree. The optimization results presented for a large example network prove the applicability of the proposed methods for everyday planning purposes.

Chapter 8

Conclusions and Outlook

For the planning and optimization of mobile communication networks, a fast and accurate performance evaluation is of essential importance. In this thesis several advanced novel methods for the performance analysis of WCDMA networks are developed and discussed. They significantly out-perform state-of-the-art methods in terms of the flexibility of modeling the air interface, accuracy of results, and computational complexity.

For the detailed and accurate performance evaluation of WCDMA networks the Monte-Carlo snapshot analysis is best suited. The performance of Monte-Carlo snapshot methods is greatly influenced by two main factors. These are, firstly the performance of the method for evaluating the performance of a single snapshot and secondly the quality of the estimator applied for obtaining the quantities of interest from the individual snapshot results. Both topics are comprehensively addressed in this thesis. The developed snapshot evaluation is very flexible in terms of modeling of system features, e.g. different QoS schemes, and is of significantly lower computational complexity compared with any currently available solution known to the author. The performance gain is mainly due to the calculation of the characteristic power control equations on cell basis rather than on connection basis and a highly linear modeling of the system. The number of total snapshots that need to be evaluated in order to obtain an output quantity with certain statistical significance is proportional to the variance of the quantity's Monte-Carlo estimator. In case of the straight forward application of a *sample mean estimation* the variance of the estimator is equal to the variance of the quantity of interest itself, divided by the number of samples. The variance, and as a consequence the number of required snapshots, can be significantly reduced by the application of variance reduction schemes. While various variance reduction schemes are known from literature, they have hardly been considered in the context of radio network planning simulations yet. In this thesis different variance reduction techniques are analyzed with respect to their applicability in network planning. Two of these techniques, namely *Control Variates* (CV) and *Correlated Sampling* (CorS), are successfully adopted to the radio network planning problem. It is shown that the application of the developed control variates can provide a considerable variance reduction in estimating both cell power levels and blocking probabilities. Although these two output quantities

constitute very important figures in the analysis of WCDMA networks, the deduction of control variates for additional output quantities is of great interest. This will require some further research. However, a similar approach as for the control variates developed in this thesis can be taken. If the performance of different network configurations is to be compared it is proposed to apply CorS. CorS can provide a significant variance reduction, in particular if the differences between the compared networks are small. In contrast to CV, CorS can be applied in the snapshot analysis without significant additional implementation effort.

While providing detailed and accurate results for various applications in network planning, the snapshot analysis is a too time-consuming technique to be applied in radio network optimization, where typically a large number of different network configurations - though at a lower level of detail - need to be compared in a very short time-frame. In particular for this purpose three fast analytical methods are presented. Based on an existing static method, two statistical methods incorporating the impact of shadow fading on the performance evaluation are developed in this thesis. The example results clearly show the advantage of the statistical methods compared with the static method. Furthermore, the results show a very good agreement with those obtained from the snapshot analysis. This holds particularly for low to medium data rate services (user data rate $\leq$128 kbps). It is important to note, that for an application in everyday planning the statistical methods need to be implemented efficiently, e.g. by using look-up tables.

In automated network optimization, a fast and accurate performance analysis is of essential importance. This fact seems to have been taken care of only to some extent in the design of many optimization tools. One common approach is to apply a detailed and too time consuming (snapshot analysis) method for evaluating different candidate solutions, resulting in an unacceptable run-time of the algorithms. The other extreme, which is even less satisfactory, is to build on a fast but inaccurate performance analysis on the one hand and on the other hand to apply a very complex sophisticated optimization algorithm, resulting in (near) optimal results with respect to a cost function which does not (or not sufficiently) correlate with reality. The local search optimization method developed in this thesis makes use of the statistical methods discussed above and applies a local search algorithm which has been developed specifically for the special requirements of an air interface optimization. The presented results make the benefits of an automated network optimization evident.

Work in this area of research will continue in future. Major challenges will be to extend the existing models and algorithms towards the different emerging mobile communication technologies, e.g. *Worldwide Interoperability for Microwave Access* (WiMAX) and 3GPP *(3GPP) Long Term Evolution* (LTE). Due to the co-existence of multiple technologies, also the combined analysis and optimization of multiple technologies and their mutual dependency will be of increasing interest. Furthermore, the constantly improving performance of computer hardware might in the near future allow a more detailed analysis of networks, e.g. explicit incorporation of user mobility and traffic dynamics into the analysis. First research in this direction has started.

Appendix A

Generating Correlated Random Variables

For the modeling of shadowing in simulations depending on the applied model (cf. Section 3.2) a number of correlated random variables have to be generated. In the following a methodology for this is derived [JBS00].

Assume that n normal random variates

$$X = (X_1, X_2, \cdots, X_n)^T \tag{A.1}$$

are to be drawn from a normal distribution with expected values

$$\mu = (\mu_1, \mu_2, \cdots, \mu_n)^T \tag{A.2}$$

and variance-covariance matrix given by

$$\text{var}\{X\} = V = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{pmatrix} \tag{A.3}$$

Apart from being symmetric, an important property of the latter matrix is that in practical cases it also is positive definite. This will be investigated in further detail in the following subsection.

Positive Definiteness of the Covariance Matrix

A matrix $A \in R^{n \times n}$ is positive definite if $a^T A a > 0$ for all nonzero $a \in R^n$ [GvL96]. This property shall in the following be investigated for V.

Considered shall be the scalar $Y = a^T X = \sum_{i=1}^{K} a_i X_i$. The variance of Y can be expressed

as follows:

$$\begin{aligned} \text{var}\{Y\} &= \text{var}\left\{\sum_{i=1}^{n} a_i X_i\right\} \\ &= \sum_{i=1}^{n} a_i^2\sigma_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} a_i a_j \sigma_i \sigma_j \rho_{ij} \\ &= a^T V a \end{aligned} \tag{A.4}$$

On the other hand var $\{Y\}$ is greater than zero if at least one of the random variates X_i has non zero variance. Hence, in practical cases the following expression holds:

$$a^T V a > 0 \tag{A.5}$$

Consequently positive definiteness of V can be assumed.

Generation Algorithm

The property of positive definiteness of V will be made use of in this section for an efficient generation of values of X.

At first a vector of independent univariate standard normal random variables is generated:

$$K = (K_1, K_2, \cdots, K_n)^T \tag{A.6}$$

where

$$\text{E}\{K\} = 0 \tag{A.7}$$

and the variance-covariance matrix is equal to the identity matrix:

$$\text{var}\{K\} = I \tag{A.8}$$

In order to generate X from K, the matrix S in

$$X^* = \mu + SK \tag{A.9}$$

shall be calculated such that $X^* = X$. As X^* and X are normal distributed, it is sufficient if the first two moments are equal:

$$\text{E}\{X^*\} = \mu + S\text{E}\{K\} = \mu \tag{A.10}$$

and

$$\begin{aligned} \text{var}\{X^*\} &= \text{E}\left\{(X^* - \mu)(X^* - \mu)^T\right\} = \text{E}\left\{SK(SK)^T\right\} \\ &= \text{E}\left\{SKK^TS^T\right\} = S\text{E}\left\{KK^T\right\}S^T \\ &= SS^T \end{aligned} \tag{A.11}$$

Thus, the matrix S is to be determined such that:

$$SS^T = V \tag{A.12}$$

As V is symmetric the decomposition in (A.12) can be carried out by the *Cholesky factorisation*. As V is positive definite (cf. Section A) very stable and efficient algorithms are known [GvL96].

Bibliography

[3GP99a] 3GPP. *Broadcast/Multicast Control Protocol (BMC) Specification*. Technical Specification 25.324, 3GPP, Sophia Antipolis, France, December 1999.

[3GP99b] 3GPP. *MAC Protocol Specification*. Technical Specification 25.321, 3GPP, Sophia Antipolis, France, December 1999.

[3GP99c] 3GPP. *Packet Data Convergence Protocol (PDCP) Specification*. Technical Specification 25.323, 3GPP, Sophia Antipolis, France, December 1999.

[3GP99d] 3GPP. *Physical layer procedures (FDD)*. Technical Specification 25.214, 3GPP, Sophia Antipolis, France, December 1999.

[3GP99e] 3GPP. *Radio Interface Protocol Architecture*. Technical Specification 25.301, 3GPP, Sophia Antipolis, France, December 1999.

[3GP99f] 3GPP. *RLC Protocol Specification*. Technical Specification 25.322, 3GPP, Sophia Antipolis, France, December 1999.

[3GP99g] 3GPP. *RRC Protocol Specification*. Technical Specification 25.331, 3GPP, Sophia Antipolis, France, October 1999.

[3GP99h] 3GPP. *Spreading and modulation (FDD)*. Technical Specification 25.213, 3GPP, Sophia Antipolis, France, October 1999.

[3GP01] 3GPP. *Physical Layer Aspects of UTRA High Speed Downlink Packet Access*. Technical Specification 25.848, 3GPP, Sophia Antipolis, France, March 2001.

[3GP04a] 3GPP. *Introduction of the Multimedia Broadcast Multicast Service (MBMS) in the Radio Access Network (RAN)*. Technical Specification 25.346, 3GPP, Sophia Antipolis, France, March 2004.

[3GP04b] 3GPP. *Radio resource management strategies*. Technical Specification 25.922, 3GPP, Sophia Antipolis, France, March 2004.

[3GP05] 3GPP. *FDD Enhanced Uplink; Physical Layer Aspects*. Technical Specification 25.808, 3GPP, Sophia Antipolis, France, March 2005.

[Air] Aircom International Ltd. *ASSET3G: Radio Network Planning Software*. Redhill, UK. `http://www.aircom.co.uk/Radio_Planning_Tool.html`.

[AL97] E. Aarts, J. K. Lenstra, editors. *Local Search in Combinatorial Optimization*. John Wiley & Sons, Inc., New York, NY, USA, 1997.

[BA02] N. Blaunstein, J.B. Andersen. *Multipath Phenomena in Cellular Networks*. Artech House, Norwood, MA, USA, 2002.

[BADM95] N. C. Beaulieu, A. A. Abu-Dayya, P. J. McLane. *Estimating the Distribution of a Sum of Independent Lognormal Random Variables*. IEEE Transactions on Communications, Vol. 43, No. 12, pp. 2869 – 2873, December 1995.

[Bal76] P. Balaban. *Statistical evaluation of the error rate of the fiberguide repeater using importance sampling*. Bell System Technical Journal, Vol. 55, No. 6, pp. 745–766, August 1976.

[Böt06] H. Böttner. *Entwicklung und Analyse verschiedener Verfahren zur Antennenmodellierung in der Funknetzplanung*. Diplomarbeit (In cooperation with Siemens), Fachhochschule Südwestfalen, Meschede, Germany, April 2006.

[BX04] N. C. Beaulieu, Q. Xie. *An Optimal Lognormal Approximation to Lognormal Sum Distributions*. IEEE Transactions Vehicular Technology, Vol. 53, No. 2, pp. 479 – 489, March 2004.

[CCK+02] F. D. Cardoso, L. M. Correia, T. Kürner, P. Lourenco, G. Marques, D. Matic. *Evaluation of Propagation Models*. Technical Report D4.1 (Project Internal), IST-2000-28088 MOMENTUM, 2002.

[CLRS01] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein. *Introduction to Algorithms, Second Edition*. The MIT Press, September 2001.

[DKKN99] Z. Dziong, K. Krishnan, S. Kumar, S. Nanda. *Statistical "Snap-Shot" for Multi-Cell CDMA Sytsem Capacity Analysis*. In *Proc. of the IEEE Wireless Communications and Networking Conference (WCNC'99)*, pp. 1234–1237, New Orleans, LA, USA, September 1999.

[DS94] J. Dunlop, D. G. Smith. *Telecommunications Engineering*. Stanley Thornes, Cheltenham, UK, 1994.

[Dub99] A. Dubi. *Monte Carlo Applications in System Engineering*. John Wiley & Sons, Inc., New York, NY, USA, 1999.

[EFF+03] A. Eisenblätter, A. Fügenschuh, E. R. Fledderus, H.-F. Geerdes, B. Heideck, D. Junglas, T. Koch, T. Kürner, A. Martin. *Mathematical Methods for Automatic Optimization of UMTS Radio Networks*. Technical Report D4.3, IST-2000-28088 MOMENTUM, 2003.

[EG05] A. Eisenblätter, H.-F. Geerdes. *A Novel View on Cell Coverage and Coupling for UMTS Radio Network Evaluation and Design*. In *Proc. of the International Network Optimization Conference INOC'05*, pp. B2.307–14, Lisbon, Portugal, March 2005. ENOG.

[EGK+06] A. Eisenblätter, H.-F. Geerdes, T. Koch, A. Martin, R. Wessäly. *UMTS Radio Network Evaluation and Optimization Beyond Snapshots*. Math. Meth. Oper. Res., Vol. 63, No. 1, 2006.

[EGR05] A. Eisenblätter, H.-F. Geerdes, N. Rochau. *Analytical Approximate Load Control in WCDMA Radio Networks*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-05 Fall)*, Dallas, TX, USA, Sep. 2005. IEEE.

[EGT05] A. Eisenblätter, H.-F. Geerdes, U. Türke. *Public UMTS radio network evaluation and planning scenarios*. Int. J. Mobile Network Design and Innovation, Vol. 1, No. 1, pp. 40–53, 2005.

[For] Forsk. *ATOLL: Global RF Planning Solution*. Blagnac, France. `http://www.forsk.com/htm/products/atoll_overview.htm`.

[FR05] B. Fourestié, S. Renou. *On the Enhancement of Monte Carlo 3G Network Modelling Tools for QoS prediction*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-05 Spring)*, Stockholm, Sweden, May 2005.

[GCF+01] F. Gil, A. Claro, J. Ferreira, C. Pardelinha, L. Correia. *A 3D Interpolation Method for Base - Station - Antenna Radiation Patterns*. IEEE Antennas and Propagation Magazine, Vol. 43, No. 2, pp. 132–137, 2001.

[GLL+03] H.-F. Geerdes, E. Lamers, P. Lourenço, E. Meijerink, U. Türke, S. Verwijmeren, T. Kürner. *Evaluation of Reference and Public Scenarios*. Technical Report D5.3, IST-2000-28088 MOMENTUM, 2003.

[GvL96] G. H. Golub, C. F. van Loan. *Matrix Computations*. Johns Hopkins, 1996.

[Hat80] M. Hata. *Empirical formula for propagation loss in land mobile radio services*. IEEE Transactions on Vehicular Technology, Vol. 29, pp. 317–325, 1980.

[HCLS01] C.-J. Ho, J.A. Copeland, C.-T. Lea, G.L. Stüber. *On Call Admission Control in DS/CDMA Cellular Networks*. IEEE Transactions on Vehicular Technology, Vol. 50, No. 6, pp. 1328–1343, May 2001.

[Ho95] C.-L. Ho. *Calculating Mean and Variance of Power Sums with Two Log-Normal Components*. IEEE Transactions Vehicular Technology, Vol. 44, No. 4, pp. 756 – 762, November 1995.

[Hop03] R. Hoppe. *Comparison and Evalution of Algorithms for the Interpolation of 3D Antenna Patterns Based on 2D Horizontal and 2D Vertical Patterns*. Technical Report, AWE Communications GmbH, March 2003.

[HT04] H. Holma, A. Toskala, editors. *WCDMA for UMTS*. John Wiley & Sons, Inc., Chichester, UK, 3rd edition, 2004.

[HT06] H. Holma, A. Toskala. *HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications*. John Wiley & Sons, Chichester, UK, 2006.

[HY96] C. Huang, R. Yates. *Call Admission in Power Controlled CDMA Systems*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-96 Fall)*, pp. 1665–1669, May 1996.

[IT97] ITU-T. *Guidelines For Evaluation of Radio Transmission Technologies For IMT-2000*. Recommendation M.1225, International Telecommunication Union, 1997.

[JBS00] M. C. Jeruchim, P. Balaban, K. S. Shanmugan, editors. *Simulation of Communication Systems: Modeling, Methodology and Techniques*. Kluwer Academic Publishers, Norwell, MA, USA, 2000.

[Jer84] M. C. Jeruchim. *Techniques for estimating the bit error rate in the simulation of digital communication*. IEEE Journal on Selected Areas in Communications, Vol. 2, No. 1, pp. 153–170, January 1984.

[Kal86] M. H. Kalos. *Monte Carlo Methods, Volume I: Basics*. John Wiley & Sons, Inc., New York, NY, USA, 1986.

[Kat00] *Technical Information and New Products*. Number 3. KATHREIN-Werke KG, September 2000. `http://www.kathrein.de`.

[Kat06] *790–2500 MHz Base Station Antennas for Mobile Communictions (catalogue)*. KATHREIN-Werke KG, Rosenheim, Germany, 2006. `http://www.kathrein.de`.

[KFW96] Th. Kürner, R. Fauß, A. Wäsch. *A hybrid propagation modelling approach for DCS1800 macro cells*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-96 Fall)*, pp. 1628–1632, May 1996.

[Kür99] Th. Kürner. *Propagation Models for Macro Cells*. In *Digital Mobile Radio Towards Future Generation Systems, COST 231 Final Report*. Commission of the European Communities and COST Telecommunications, Brussels, Belgium, 1999.

[LE94] Z. Liu, M. Elzarki. *SIR-Based Call Admission Control for DS-CDMA Cellular Systems*. IEEE Journal on Selected Areas in Communications, Vol. 12, No. 4, pp. 638–644, May 1994.

[Lin92] J.-P. M. G. Linnartz. *Exact Analysis of the Outage Probability in Mobile-User Mobile Radio*. IEEE Transactions on Communications, Vol. 40, No. 1, pp. 20 – 23, January 1992.

[LM98] J. S. Lee, L. E. Miller. *CDMA Systems Engineering Handbook*. Artech House, Inc., Norwood, MA, USA, 1998.

[LPTG03] E. Lamers, R. Perera, U. Türke, C. Görg. *Short Term Dynamic System Level Simulation Concepts for UMTS Network Planning*. In *Proc. of the International Teletraffic Congress (ITC 18)*, pp. 101–110, Berlin, Germany, September 2003.

[Maj06] K. Majewski. *Interference Powers at the Antennas of UMTS Radio Networks.* Technical Report CT SE 6-05/06-KM-2, Siemens AG, CT SE 6, Munich, Germany, July 2006.

[Mar] Marconi Communications Ltd. *PlanetEV: RF Planning and Optimization Software*. London, UK. `http://www.marconi.com`.

[MH01] L. Mendo, J. M. Hernando. *On Dimension Reduction for the Power Control Problem*. IEEE Transactions on Communications, Vol. 49, pp. 243–248, 2001.

[mom] MOMENTUM *Project, IST-2000-28088*. http://momentum.zib.de.

[OOKF68] Y. Okumura, E. Ohmori, T. Kawano, K. Fukuda. *Field strength and its variability in VHF and UHF land-mobile service*. Review of the Electrical Communication Laboratory, Vol. 16, No. 9-10, pp. 825–873, 1968.

[Pro95] J. G. Proakis. *Digital Communications*. McGraw-Hill, Singapore, 1995.

[Rap02] T. S. Rappaport. *Wireless Communications*. Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2002.

[Ree95] C. R. Reeves, editor. *Modern heuristic techniques for combinatorial problems*. McGraw-Hill, London, UK, 1995.

[Rub81] Y. R. Rubinstein. *Simulation and the Monte Carlo Method*. John Wiley & Sons, Inc., New York, NY, USA, 1981.

[Rub86] Y. R. Rubinstein. *Monte Carlo Optimization, Simulation and Sensitivity of Queueing Networks*. John Wiley & Sons, Inc., New York, NY, USA, 1986.

[Sau99] S. R. Saunders. *Antennas and Propagation for Wireless Communication Systems*. John Wiley & Sons, Inc., New York, NY, USA, 1999.

[SB80] K. S. Shanmugan, P Balaban. *A modified Monte-Carlo simulation technique for the evaluation of error rates in digital communication systems*. IEEE Transactions on Communications, Vol. 28, No. 11, pp. 745–766, November 1980.

[Sha48] C.E. Shannon. *A mathematical theory of communication*. Bell System Technical Journal, Vol. 27, pp. 379–423 and 623–656, July and October 1948.

[SJLSW99] K. Sipilä, M. Jäsberg, J. Laiho-Steffens, A. Wacker. *Soft Handover Gains in a Fast Power Controlled WCDMA Uplink*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-99 Spring)*, pp. 1594–1598, Houston, TX, USA, May 1999.

[SLSWJ99] K. Sipilä, J. Laiho-Steffens, A. Wacker, M. Jäsberg. *Modeling the Impact of the Fast Power Control on the WCDMA Uplink*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-99 Spring)*, pp. 1266–1270, Houston, TX, USA, May 1999.

[SLW+01] B. Schröder, B. Liesenfeld, A. Weller, K. Leibnitz, D. Staehle, P. Tran-Gia. *An analytical approach for determining coverage probabilities in large UMTS networks*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-01 Fall)*, Atlantic City, NJ, USA, October 2001.

[SM04] D. Staehle, A. Mäder. *An Analytic Model for Deriving the Node-B Transmit Power in Heterogeneous UMTS Networks*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-04 Spring)*, Milano, Italy, May 2004.

[Sta04] D. Staehle. *Analytic Methods for UMTS Radio Network Planning*. Phd Thesis, University of Würzburg, November 2004.

[Sta05] D. Staehle. *An Analytic Method for Coverage Prediction in the UMTS Radio Network Planning Process*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-05 Spring)*, Stockholm, Sweden, June 2005.

[TG96] P. Tran-Gia. *Analytische Leistungsbewertung verteilter Systeme*. Springer, Berlin and Heidelberg and New York, 1996.

[TK05] U. Türke, M. Koonert. *Advanced Site Configuration Techniques for Automatic UMTS Radio Network Design*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-05 Spring)*, Stockholm, Sweden, May 2005.

[TKSG04] U. Türke, M. Koonert, R. Schelb, C. Görg. *HSDPA Performance Analysis in UMTS Radio Network Planning Simulations*. In *Proc. of the IEEE Vehicular Technology Conference (VTC-04 Spring)*, Milan, Italy, May 2004.

[TPL+03] U. Türke, R. Perera, E. Lamers, T. Winter, C. Görg. *An Advanced Approach for QoS Analysis in UMTS Radio Network Simulations*. In *Proc. of the International Teletraffic Congress (ITC 18)*, pp. 91–100, Berlin, Germany, September 2003.

[TTGH99] K. Tutschku, P. Tran-Gia, M. Heuler. *A framework for spatial traffic estimation and characterization in mobile communication network design*. In *16th International Teletraffic Congress - ITC16*, pp. 139–148, Edinburgh, UK, June 1999.

[Vit95] A. J. Viterbi. *CDMA: Principles of Spread Spectrum Communication*. Addison-Wesley, Redwood City, CA, USA, 1995.

[VW00] R. S. Varga, P. A. Whitlock. *Matrix Iterative Analysis*. Springer, Berlin and Heidelberg, Germany, 2000.

[WDT84] A. A.-Q. Wael, M. Devetsikiotis, J. K. Townsend. *Importance Sampling Methodologies for Simulation of Communication Systems with Time-Varying Channels and Adaptive Equalizers*. IEEE Journal on Selected Areas in Communications, Vol. 2, No. 1, pp. 153–170, January 1984.

[WLSSJ99] A. Wacker, J. Laiho-Steffens, K. Sipilä, M. Jäsberg. *Static simulator for studying WCDMA radio network planning issues*. In *Proc. of the IEEE Ve-*

hicular Technology Conference (VTC-99 Spring), pp. 2436–2440, Houston, TX, USA, May 1999.

[WM02] R. E. Walpole, R. H. Myers. *Probability and Statistics for Engineers and Scientists*. Prentice Hall, Upper Saddle River, NJ, USA, 7th edition, 2002.

[WTK04] T. Winter, U. Türke, M. Koonert. *A generic approach for including live measurements and traffic forecasts in the generation of realistic traffic scenarios in mobile radio networks.* In *Proc. of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems (MSWim'04)*, pp. 83–86, 2004.

[WTL+03] T. Winter, U. Türke, E. Lamers, R. Perera, A. Serrador, L. Correia. *Advanced simulation approach for integrated static and short-term dynamic UMTS performance evaluation*. Technical Report D2.7, IST-2000-28088 MOMENTUM, 2003.

[YS84] Y.-S. Yeh, S.C. Schwartz. *Outage Probability in Mobile Telephony Due To Multiple Log-Normal Interferers*. IEEE Transactions on Communications, Vol. 32, No. 4, pp. 380 – 388, 1984.

[ZK05] A. A. Zavala, K. Koulinas. *Antenna Radiation Pattern Estimation in 3D for Indoor Environments*. In *Proc. of the 15th International Conference on Electronics, Communications and Computers*, Los Alamitos, CA, USA, 2005.