Technische Universität München
TUM School of Computation, Information and Technology

Wave Propagation Simulations by Bidirectional Ray-Tracing and Investigations on Ray Based Channel Modeling for Massive MIMO

Mehmet Mert Taygur

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines

– Doktors der Ingenieurswissenschaften –

genehmigten Dissertation.

Vorsitz:	Prof. Dr.-Ing. Wolfgang Kellerer
Prüfer*innen der Dissertation:	1. Prof. Dr.-Ing. Thomas F. Eibert
	2. Prof. Dr. Conor Brennan

Die Dissertation wurde am 15.11.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 26.03.2023 angenommen.

Abstract

Ray-tracing simulations are essential for the design and analysis of cellular networks which rely on multiple-input multiple-output (MIMO) systems. Compared to the other commonly utilized computational electromagnetics approaches, ray-tracing, in particular the Shooting and Bouncing Rays (SBR) method with Geometrical Optics and Uniform Theory of Diffraction (GO-UTD), is advantageous for the characterization of wave propagation within cellular MIMO systems, as the simulation time is typically much smaller in such electrically large problems. Ray-tracing simulations can effectively be utilized to model the intricate features of the channel in simple cellular links as well as sophisticated MIMO systems, such as massive MIMO, which involve a large number of antennas.

Nevertheless, the SBR method and GO-UTD approach have certain limitations and might be erroneous under certain conditions, which may arise frequently in many practically relevant propagation scenarios. The SBR method relies on the so-called reception spheres, which are typically small in size, to detect the valid ray paths. A crucial prerequisite here is to determine the size of these structures as well as the number of ray launches, which complement each other. A sub-optimal choice typically yields a decline in accuracy. In certain high-frequency problems, such as millimeter-wave propagation scenarios, the SBR approach may also result in large phase errors due to an inexact ray path calculation, thus, an according decline in simulation accuracy is usually observed. The treatment of multiple diffractions might also be problematic, since the accuracy of the UTD approach may drop considerably when it is applied successively on multiple edges.

In order to address the problems related to the SBR and GO-UTD, a bidirectional ray-tracing algorithm is proposed. The method relies on the use of large interaction surfaces, instead of small reception spheres, in order to identify the valid field contributions. The ray launching is performed not only from the transmitter, as in the traditional ray-tracing, but also from the receiver. The interaction between the antennas is then computed by evaluating a reciprocity integral. With this technique, a smaller number of ray launches is usually sufficient to achieve a good accuracy, compared to the conventional ray-tracing approach. Since the interaction surface can usually be placed in a flexible manner within the geometry, multiple diffraction effects are simulated by placing the surface above the edges and then calculating the antenna reciprocity on it. Thus, the accuracy problems of UTD related to multiple diffractions are mitigated. Furthermore, the phase errors due to inexact ray path calculations in millimeter-wave scenarios are mostly avoided by applying an asymptotic expansion on the reciprocity integral.

The differences between ray-tracing simulations and common stochastic channel models for urban massive MIMO scenarios are also addressed. A particular emphasis is given on the variation of the downlink data rate when the number of base station

antennas is changed. In order to address the problem under more realistic conditions, six different propagation-related phenomena are (separately) taken into account, i.e., pilot contamination, channel-aging, frequency selectivity, multi-antenna users, rooftop diffractions, and millimeter-wave propagation. An important difference between the two approaches is the availability of exact geometrical information in the simulation, which generally leads to distinct channel correlation characteristics. Hence, discrepancies in the variation of the data rate can be noticed when the number of the base station antennas is changed. The improvements in data rate for a larger number of base station antennas can be observed with both approaches, however, the adverse effects of the aforementioned propagation-related phenomena may typically not be mitigated by introducing more antennas into the base station.

Kurzfassung

Strahlverfolgungssimulationen ((engl.) Ray-Tracing) sind ein wichtiges Werkzeug für die Planung und Analyse von Mobilfunknetzen, welche auf Multiple-Input Multiple-Output (MIMO) Systemen basieren. Im Vergleich zu den anderen häufig verwendeten numerischen Lösungsverfahren ist die Anwendung von Ray-Tracing, vor allem in Form der Strahlabschussmethode ((engl.) Shooting and Bouncing Rays Method, SBR) zusammen mit der Geometrischen Optik (GO) und der Vereinheitlichen Beugungstheorie ((engl.) Uniform Theory of Diffraction, UTD) vorteilhaft zur Charakterisierung der Wellenausbreitung in MIMO basierten Mobilfunknetzwerken, weil die Simulationsdauer bei solch großen Problemen typischerweise deutlich geringer ist. Ray-Tracing kann in einfachen Mobilfunkverbindungen sowie in komplexen MIMO Systemen (z.B. Massive MIMO), die eine große Anzahl von Basisstationsantennen umfassen, zur Modellierung des Kommunikationskanals verwendet werden.

Dennoch können die Einschränkungen der SBR Methode und von GO-UTD zu Simulationen mit ungenauen Ergebnissen führen, die in praxisrelevanten Szenarien häufig entstehen können. Zur Erfassung der Ausbreitungspfade setzt die SBR Methode auf sogennante Empfangskugeln, die in der Regel klein sind. Die Größe der Empfangskugeln und die Strahlanzahl sollten jedoch sorgfältig eingestellt werden, da eine suboptimale Auswahl grundsätzlich zu ungenauen Ergebnissen führt. In Millimeterwellen-Szenarien kann die SBR Methode außerdem wegen ungenauer Pfadberechnungen große Phasenfehler ergeben, dadurch nimmt die Simulationsgenauigkeit typischerweise ab. Die Behandlung der Mehrfachbeugungen kann ebenfalls problematisch werden, weil die Genauigkeit der UTD abfällt, wenn sie auf mehrere Kanten nacheinander angewendet wird.

Um die Probleme in SBR und GO-UTD zu beheben, wird eine bidirektionale Strahlverfolgungsmethode ((engl.) Bidirectional Ray-Tracing) vorgestellt. Dieses Verfahren umfasst ein große Interaktionsfläche, anstelle von kleinen Empfangskugeln, wobei die Strahlen von beiden Richtungen, von der Sende- und der Empfangsantenne, gesendet werden. Die Interaktion zwischen beiden Antennen wird durch ein Reziprozitätsintegral berechnet. Im Vergleich zum klassischen Strahlverfolgungsverfahren können die Simulationen mit geringerer Strahlanzahl und ohne Genauigkeitsverluste durchgeführt werden. Da die Interaktionsfläche in der Geometrie flexibel positioniert werden kann, kann die Mehrfachbeugung mit einer Interaktionsfläche simuliert werden, die über den Kanten plaziert ist, wobei die Simulationsgenauigkeit in den Mehrfachbeugungszenarien verbessert wird. Außerdem werden Phasenfehler wegen der ungenauen Pfadberechnungen durch eine asymptotische Entwicklung des Reziprozitätsintegrals weitgehend vermieden.

Die Unterschiede zwischen Strahlverfolgung und stochastischen Methoden werden zur Kanalmodellierung von Massive MIMO Systemen in städtischen Umgebungen untersucht. Besonderes Augenmerk wird auf die Variation der Downlink-Datenrate bei

Änderung der Anzahl der Basisstationsantennen gelegt. Um realistische Bedingungen zu gewährleisten, werden sechs verschiedene ausbreitungsbezogene Phänomene (unabhängig voneinander) berücksichtigt. Darunter sind Pilotsignalstörungen, Alterungseffekte im Kanal, frequenzselektive Kanäle, Benutzerendgeräte mit mehreren Antennen, Wellenausbreitung und Beugung über den Dächern sowie Wellenausbreitung mit Millimeterwellentechnologie. Ein wesentlicher Unterschied zwischen den beiden Ansätzen ist die Verfügbarkeit der genauen Geometrieinformation, die zu verschiedenen Kanalkorrelationsergebnissen führt. Entsprechend können Abweichungen in der Datenratenvariation beobachtet werden, wenn die Anzahl der Basisstationsantennen geändert wird. Verbesserungen der Datenrate können mit beiden Ansätzen beobachtet werden, wenn die Anzahl der Basisstationsantennen vergrößert wird, jedoch kann die negative Auswirkung der obengennanten ausbreitungsbezogenen Phänomene durch eine größere Anzahl von Basisstationsantennen nicht vermindert werden.

Acknowledgement

The work presented in this dissertation was carried out at the Chair of High-Frequency Engineering, Technical University of Munich where I worked as a research assistant for almost six years, which I will always remember with a smile on my face, as I consider myself fortunate for having this opportunity and finally completing this dissertation. That being said, I cannot ignore the contributions of those who made it possible, therefore I would like to give them, at least some of the credit they deserve.

First and foremost, I want to thank my advisor Prof. Thomas F. Eibert, who not only gave me the chance to conduct this work under his supervision, but also helped me to become a better engineer through his guidance, for which I will always be grateful. Many thanks to Prof. Conor Brennan, who showed great interest in my work and provided invaluable insight to improve the dissertation.

I owe my sincere gratitude to my fellow colleagues at the university, who made my time in Germany truly enjoyable. I have been lucky for having spent a great time with many wonderful people, though some of them deserve a special mention here, namely, Dr. Alexander Paulus, Arslan Azhar, Bernd Hofmann, Björn Möhring, Carlos Lopez, Daniel Ostrzyharczik, Dario Vojvodic, Dr. Gerhard Hamberger, Han Na, Dr. Illia Sukharevsky, Dr. Jonas Kornprobst, Jonas Weindl, Dr. Josef Knapp, Dr. Kun Wang, Dr. Ole Neitz, Dr. Safiullah Khan, Dr. Simon Adrian, Dr. Uwe Siart, who were not only great colleagues but also great companions outside of work.

Special thanks to my mother Ayşe, and to my sister Elif, who are both now on their own Ph.D. quests as I am writing these lines. I hope this work would eventually be a source of motivation for them to reach their goals, as they have always supported me unconditionally, even when the distance between us became an immense burden during difficult times.

One person whom I cannot thank enough is my dear wife Yaprak, as she, besides her endless support and encouragement, accepted to sacrifice her own good for my success countless times without hesitation. Her presence in my life was a blessing for me during this journey.

Finally, I want to thank my father Ahmet, who made me love science, inspired me to become a brilliant engineer like he was, and ultimately sparked the idea in my mind to pursue a doctoral degree, thereby, this work is dedicated to his memory. Although I cannot show him the work, which he inspired me to do many years ago, I know that he proudly watches me, wherever he is, with the same smile that I have on my face…

Munich, Summer 2023

Contents

List of Common Abbreviations and Symbols

Abbreviations

BC	Broadcast Channel
CSI	Channel State Information
FDD	Frequency Division Duplexing
FDTD	Finite Difference Time Domain
FEM	Finite Element Method
GO	Geometrical Optics
GTD	Geometrical Theory of Diffraction
ISI	Intersymbol Interference
MAC	Multiple-Access Channel
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MU-MIMO	Multi-User MIMO
MoM	Method of Moments
PEC	Perfectly Electrically Conducting
PO	Physical Optics
RZF	Regularized Zero Forcing
SBR	Shooting and Bouncing Rays
SINR	Signal-to-Interference and Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SU-MIMO	Single-User MIMO
TDD	Time Division Duplexing
UTD	Uniform Theory of Diffraction
ZF	Zero Forcing

Symbols

$(.)^{-1}$	Matrix inverse operator
$(.)^{H}$	Hermitian transpose operator
$(.)^{T}$	Transpose operator
$(.)_{(i)}$	ith row of vector
$(.)_{(i,:)}$	ith row of matrix
$(.)_{(:,j)}$	jth column of matrix

$(.)_{(i,j)}$	ith row and jth column of matrix
$\mathbf{A}$	Precoding matrix
$A(s)$	Attenuation factor
$\boldsymbol{B}(\mathbf{r},t)$	Magnetic flux density (Vs/m^2)
$\mathbf{B}(\mathbf{r},\omega)$	Time-harmonic magnetic flux density (Vs/m^2)
$\mathcal{CN}$	Gaussian distributed complex random variable
C	Channel capacity for MIMO channel (bits/s/Hz)
C^{sum}	Sum capacity for multi-user MIMO channel (bits/s/Hz)
C_u	Channel capacity for user u (bits/s/Hz)
$\boldsymbol{D}(\mathbf{r},t)$	Electric flux density (As/m^2)
$\mathbf{D}(\mathbf{r},\omega)$	Time-harmonic electric flux density (As/m^2)
$\overline{\mathbf{D}}$, $\mathbf{D}$	Diffraction coefficient in dyadic and matrix form
$\boldsymbol{E}(\mathbf{r},t)$	Electric field (V/m)
$\mathbf{E}^{\{i,r,t,d\}}$	Incident, reflected, transmitted/refracted, and diffracted geometrical optics electric field (V/m)
$\mathbf{E}(\mathbf{r},\omega)$	Time-harmonic electric field (V/m)
$\mathbf{E}_A$, $\mathbf{E}_B$	Time-harmonic electric field generated by $\mathbf{J}_A$ and $\mathbf{J}_B$ (V/m)
$E[.]$	Expected value operator
$\mathbf{G}$	Channel matrix
$\mathbf{G}_{a,b}$	Channel matrix for users/receivers in ath cell and base station antennas/transmitters in bth cell
$\tilde{\mathbf{G}}_b$	Channel estimation of the base station in bth cell
$\boldsymbol{H}(\mathbf{r},t)$	Magnetic field (A/m)
$\mathbf{H}(\mathbf{r},\omega)$	Time-harmonic magnetic field (A/m)
$\mathbf{H}_A$, $\mathbf{H}_B$	Time-harmonic magnetic field generated by $\mathbf{J}_A$ and $\mathbf{J}_B$ (A/m)
$\mathbf{I}$	Identity matrix
I_A	Port current for antenna A (A)
$\boldsymbol{J}(\mathbf{r},t)$	Electric current density (A/m^2)
$\mathbf{J}_A$, $\mathbf{J}_B$	Impressed time-harmonic electric current density for antenna A and B (A/m^2)
$\boldsymbol{M}(\mathbf{r},t)$	Magnetic current density (V/m^2)
M_T, M_R	Number of transmitting and receiving antennas
P^{tx}	Total transmit power (W)
P_u	Power received or allocated for the user u (W)
$\mathbf{Q}_{\mathbf{x}}$	Covariance matrix for vector $\mathbf{x}$
$\overline{\mathbf{R}}$, $\mathbf{R}$	Reflection coefficient in dyadic and matrix form
R	Total data rate (bits/s)
R_u	Data rate for user u (bits/s)
$S_e(\mathbf{r})$	Eikonal
$\overline{\mathbf{T}}$, $\mathbf{T}$	Transmission/refraction coefficient in dyadic and matrix form
V^{gen}	Antenna generator voltage (V)
$V^{\text{gen}}_{b(v)}$	Generator voltage at the vth antenna in the bth cell (V)
V^{oc}	Induced open circuit voltage (V)
$V^{\text{oc}}_{a(u)}$	Induced open circuit voltage at the uth antenna in the ath cell (V)
W	Channel bandwidth (Hz)
Z_w	Wave impedance (Ω)
$\hat{\mathbf{d}}$	Diffraction edge vector

g	Oscillatory integral phase function
k	Wavenumber (1/m)
k_0	Wavenumber for the free-space (1/m)
l	Wedge opening angle (rad)
$\vec{\ell}_e$	Antenna vector effective length (m)
$\mathbf{n}$	Channel noise vector
$\mathbf{n}_b$	Channel noise vector for bth cell
$\hat{\mathbf{n}}$	Surface normal vector
$\mathbf{r}$	Coordinate in three-dimensional space
$\hat{\mathbf{s}}$	Propagation direction
s	Ray path length parameter
$\mathbf{y}$	Output signal vector
$\mathbf{x}$	Input signal vector
Γ	Oscillatory integral magnitude function
Ψ	Interaction surface domain
η	Refractive index
λ	Wavelength (m)
μ	Magnetic permeability (H/m)
μ_0	Magnetic permeability for the free-space ($\approx 1.25663706\times10^{-6}$ H/m)
μ_r	Relative permeability of the medium
ρ_1, ρ_2	Radii of curvature of astigmatic wavefront
σ_n^2	Noise power (W)
ω	Angular frequency (1/s)
ϵ	Dielectric permittivity (As/Vm)
ϵ_0	Dielectric permittivity for the free-space ($\approx 8.85418782\times10^{-12}$ As/Vm)
ϵ_r	Relative permittivity of the medium

Chapter 1

Introduction

Numerical solution techniques for Maxwell's equations have been an important topic in electromagnetics research for several decades already, and many different methods have been developed. A common approach is to utilize the numerically exact methods (i.e., starting from Maxwell's equations in exact form within a discretized geometry), which can further be divided into two groups as local and global methods. The local methods are based on the discretization of the entire solution domain where Maxwell's equations are given in differential form. The field expressions at a particular point is determined by the fields at the neighboring points within the discretized grid. The Finite Element Method (FEM) and the Finite Difference Time Domain (FDTD) are the notable examples of local approaches. On the other hand, global methods typically rely on an integral based representation where only the antennas and scatterers in the geometry are discretized. The field at any point is dependent on the expressions at every discretized point in the grid. The Method of Moments (MoM) can be given as a prominent example for a global method.

In general, numerically exact methods (both local and global) become computationally expensive if the simulation involves electrically large objects. For such problems, asymptotic methods, which rely on high-frequency approximations, can be employed. The major advantage of asymptotic methods is that the computational complexity in electrically large problems remains relatively small compared to the numerically exact methods (Yun and Iskander, 2015). Therefore, they are suitable for the simulation of certain problems such as wave propagation scenarios. Especially for the applications in wireless communications, asymptotic methods, and in particular the ray-tracing simulations, which are the main scope of this work, are commonly utilized to characterize the behavior of electromagnetic waves in various different propagation environments.

The use of ray-tracing simulations within wireless communication applications has been a well-known concept since several decades (McKown and Hamilton, 1991), though, the recent developments in parallel computing, especially in Graphics Processing Unit (GPU) computing domain, has yielded a breakthrough, as ray-tracing has become a widely utilized tool for channel modeling. The ray-tracing technique relies on the principles of ray optics, where the solution for Maxwell's equations is given under certain high-frequency approximations, leading to Geometrical Optics (GO) and Uniform Theory of Diffraction (UTD) formulations to describe the wave propagation (Deschamps, 1972; Kouyoumjian and Pathak, 1974). In order to achieve a good accuracy, a precise description of the propagation environment is an important prerequisite. The description

of the environment involves the geometrical attributes as well as the material properties of the surroundings. The simulation results can provide a comprehensive overview how the wave propagation takes place in a specific environment. Hence, the features of the underlying propagation channel, which are essential to characterize the performance in wireless networks, can accordingly be modeled for a particular scenario (Iskander and Zhengqing Yun, 2002). Note that ray-tracing is considered as a site-specific modeling approach, as the results correspond to the propagation behavior in a given environment. The simulation outcome is deterministic, unless a random process is introduced explicitly (Poutanen, 2011). Nevertheless, channel modeling for wireless networks does not only rely on deterministic, site-specific methods, as the literature on the topic consists of numerous stochastic approaches as well.

Although ray-tracing can effectively be utilized to analyze many real-world problems in wireless communications, various propagation scenarios may not efficiently be handled, or not accurately be simulated. A major problem, which mainly occurs due to the limitations of the GO-UTD approach, is the treatment of multiple diffractions, where the simulation complexity as well as the accuracy may deteriorate (Albani, 2005). The algorithms, which are employed to compute the actual ray paths, also play a significant role. A common choice for the scenarios with complex geometries is the Shooting and Bouncing Rays (SBR) method, where a large number of rays are launched from transmitters, traced in the geometry and collected by the so-called reception spheres in general (Yun *et al.*, 2001). The SBR method may yield a considerable increase in the computational complexity though, if the simulation parameters are selected poorly. Thus, it can be acknowledged that the traditional ray-tracing approach can significantly be improved in the context of channel modeling.

1.1 Channel Modeling

The channel modeling concept can be considered as the representation of underlying physical phenomena for a system which involves signal transmission through radio wave propagation. Especially for wireless communication systems, channel modeling has considerable relevance since the planning, optimization and analysis of such systems can effectively be accomplished by utilizing various deterministic or stochastic models, instead of performing arduous measurement campaigns which are also expensive and time-consuming in general (Gan, 2015). The progress in wireless communication technology and the extensive growth in the processing power of computers during the last decades have fueled the demand as well as the progress for the development of the models, which can handle complex environments and sophisticated network architectures. Although stochastic techniques are still useful today in many problems due to their relative simplicity, deterministic modeling approaches, such as ray-tracing, have recently gained considerable popularity as they are usually more accurate and can be used to predict various aspects of actual wave propagation phenomenon with greater detail.

1.1.1 Fundamental Properties of Wireless Communication Channels

Perhaps the most important aspect of wireless channel modeling is the representation of the variations in channel gain, in other words, the fading phenomenon. It is possible to analyze fading within two different categories, namely, large-scale and small-scale fading.

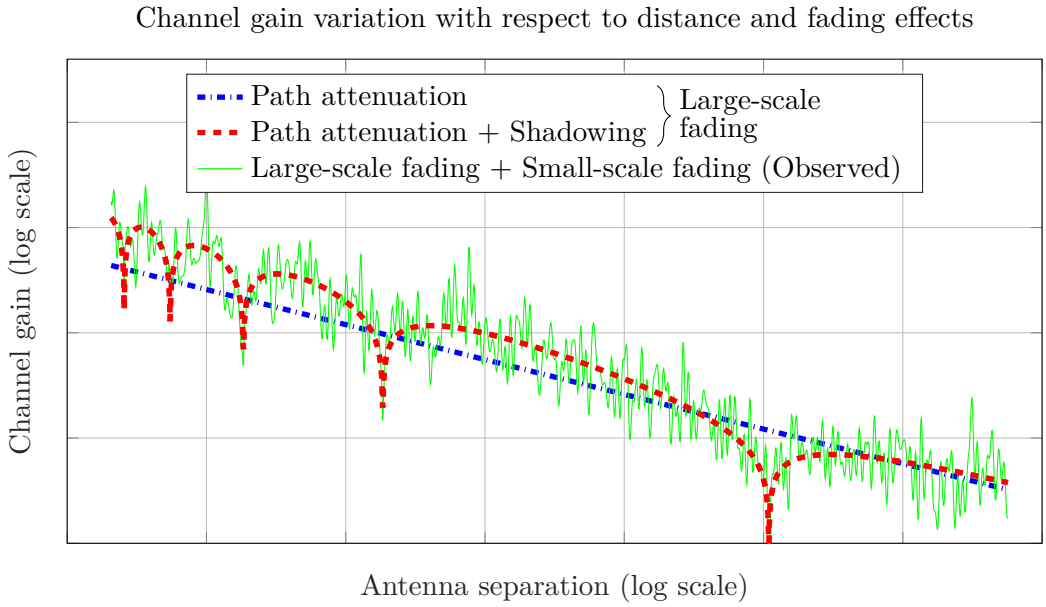


Figure 1.1: Representation of channel gain variation with respect to antenna separation for large and small-scale fading phenomena.

The former mainly describes the attenuation of a signal due to distance and shadowing effects, whereas the latter typically involves signal fluctuations from minor environmental changes and multi-paths (Özbek and Le Ruyet, 2014). It should be noted that different channel models generally represent fading phenomena with different level of detail. A simple yet very popular approach is to model large-scale fading by using several key parameters such as distance or height difference between a transceiver pair, and small-scale fading by means of statistical properties of the real channel. The observed channel gain is then considered as a combination of the large and small-scale fading effects (see Fig. 1.1).

The phenomena, which yield small-scale fading, such as multi-path propagation or environmental shifts, imply that the channel might be time-varying and frequency-selective. In other words, the channel gain can change in time and is non-uniform within a particular frequency interval (Matz and Hlawatsch, 2011). Although it is convenient to assume a static and flat-fading channel to simplify the analysis in certain cases, such effects should be taken into account to depict the wave propagation in a more realistic manner.

The frequency-selectivity is typically a result of time-dispersive channel characteristics[1], where a receiver captures multiple echoes of a single transmit signal at different time instances as the echoes presumably follow distinct propagation paths with different lengths in a rich-scattering environment. It is possible to quantify the frequency selectivity by the coherence bandwidth, which can be defined as the range of frequencies where the channel frequency response is relatively constant (Özbek and Le Ruyet, 2014). When the interval between the arrival times of the first and last echoes grow (i.e., larger delay spread), the coherence bandwidth drops (Rappaport, 1996).

[1]Note that a time-dispersive channel does not necessarily imply that the propagation medium itself is dispersive.

Time dependent fluctuations in channel gain, which occur due to environmental changes or movements of transceivers, can be quantified by coherence time. A slowly changing channel generally yields a large coherence time, where the channel characteristics may considered to be constant (Hussain, 2017). Similarly, a fast changing channel typically yields a small coherence time, which might indicate the presence of fast moving transmitters/receivers or scatterers, such as vehicles.

For systems, which involve large arrays, it is critical to characterize the directional properties of the channel. Especially in multi-antenna systems, the directional properties typically provide valuable insight about correlation, i.e., whether individual sub-channels exhibit similar fading characteristics as time and frequency dependent changes occur. Correlation has significant impact on the performance of a multi-antenna system, where a high correlation among sub-channels may imply inferior spatial multiplexing performance but typically allows coherent beamforming and vice versa (Tse and Viswanath, 2005). Note that an accurate representation of correlative properties of multi-antenna systems commonly requires knowledge about the propagation environment as well as the actual antennas (i.e., radiation pattern). Therefore, not every channel modeling approach may effectively be utilized to analyze the correlation or directional properties of a multi-antenna system.

1.1.2 Overview of Common Channel Modeling Approaches

The channel models, which are commonly utilized for the design and analysis of wireless networks, can generally be classified as either stochastic or deterministic approaches. Deterministic models typically take the wave propagation phenomena into account in an accurate way, and are site-specific, i.e., the model is based on a particular geometry and the results are generally valid only within this specified environment. Ray-tracing is a prominent example of deterministic channel modeling, where GO-UTD principles can be utilized to calculate the transfer function for a transmit and receive antenna. The stochastic models on the other hand involve the depiction of channel behavior by means of statistical approaches. The applicability of the models is generally not restricted to a particular geometry, thus, they can effectively be utilized for different scenarios.

Stochastic Channel Models

Stochastic channel models can be studied under several distinct model classes, though there are two frequently encountered classes that many well-known methods from the literature typically belong to, namely, correlation-based analytical models and geometry-based physical models. The majority of the methods which fall into either class may be utilized for point-to-point links as well as multi-antenna systems. Examples of modeling techniques for both classes are given in Table 1.1.

Correlation-based analytical models involve the representation of the channel by multivariate complex Gaussian random variables with proper correlation among the variables when multi-antenna systems with correlated channels are considered (Poutanen, 2011; Wallace and Jensen, 2002). It is usually necessary to select shadowing and fading parameters according to probabilistic distributions, which can be derived from more comprehensive models or on-site measurements. Typically used for narrowband channel simulations, correlation-based models are usually much simpler to employ compared to many other approaches.

Table 1.1: Brief overview of common stochastic channel modeling approaches.

Correlation-based analytical models	Geometry-based physical models
a) i.i.d. model	a) COST models
b) Kronecker model	b) 3GPP spatial channel models
	c) WINNER models

The i.i.d (independent and identically distributed) model is one of the simplest techniques, yet a widely used approach. In this method, the channel gain (i.e., in a multi-antenna system) is represented by statistically independent random variables, hence, it yields uncorrelated channels. Such characteristics can emerge if the environment permits rich scattering and the elements in antenna arrays have sufficient distance in-between. Nevertheless, various measurement studies indicate that it may not always be straightforward to achieve uncorrelated channels since many practical propagation scenarios involve only a small number of scatterers and antenna arrays with insufficient inter-element distance (therefore, highly correlated channels) due to spacing restrictions (Kaiser and Bourdoux, 2005).

The Kronecker model can be considered as an improvement over conventional i.i.d. model where the channel correlation is represented by two separate spatial correlation matrices for receiver and transmitter sites, respectively (Kermoal *et al.*, 2002). The complete channel matrix can be obtained by multiplying these two correlation matrices with a channel fading matrix which comprises complex i.i.d. Gaussian random variable elements.

The use of geometry based stochastic models (GBSM) has become essential for characterizing the channel behavior in modern cellular networks since they allow the incorporation of the fundamental physical features of the propagation environment into the simulations, and they provide control over a large number of parameters (Yang *et al.*, 2018). Thus, GBSM can be utilized in many different scenarios without major drawbacks in terms of accuracy.

Numerous well-known stochastic channel models have been developed by the European Cooperation in Science and Technology (COST) organization during the last decades. The COST 207 model is one of the earliest examples and it has been successfully employed to characterize the delay profile of wide-band channels for second generation cellular networks. The COST 231 project, which has provided two very useful models, the Walfisch-Ikegami and the Hata models, had an emphasis on the path loss instead of the delay profile and it is still widely utilized (Yang *et al.*, 2018). The considerations about multi-antenna systems have been included in more recent models, such as COST 259 and COST 271, which can be utilized to model delay profile, Doppler spectrum and angular dispersion (Molisch *et al.*, 2006; Asplund *et al.*, 2006). A notable feature of the COST 271 model is that both transmitter and receiver are assumed to possess multiple antennas (Roivainen, 2017). The COST 2100 model, being one of the lately developed approaches, inherits several key features from both the COST 259 and the COST 271, though it brings certain parameter and scenario updates over the older approaches (Liu *et al.*, 2012; Poutanen, 2011).

The spatial channel model (SCM) was developed by the Third Generation Partner-

ship Project (3GPP) in order to characterize the wave propagation in outdoor Multiple-Input Multiple-Output (MIMO) networks. Three different scenario configurations, namely, urban macro, suburban macro and urban micro, were considered within the model (ETSI, 2003). An operating frequency between 1-3 GHz with a bandwidth of 5 MHz is considered where the path loss expressions are based on the COST 231 Hata and the COST 231 Walfisch-Ikegami models. Multi-path components are assumed to emerge in clusters which have certain dependencies. It is possible to represent the correlation of delay profile, angular spread and shadow fading among different elements within the model (Yang *et al.*, 2018). Wide-band behavior is represented by tapped delay lines (Yuan, 2014).

The WINNER channel models have been developed by the Information Society Technologies (IST) initiative, which has introduced its first model, the WINNER I model, for the analysis of wireless networks with up to 5 GHz operating frequency and 100 MHz channel bandwidth. Seven different propagation scenarios are considered where the classification is made according to the characteristics of the propagation environment, i.e., urban, suburban or indoor environments. Similar to the SCM, it is possible to incorporate correlation among delay spread, DoA/DoD spread and shadow fading parameters. In 2007, the WINNER II model has replaced the WINNER I model, where six new propagation scenarios, such as indoor-to-outdoor and outdoor-to-indoor, have been introduced, and a distinction between line-of-sight (LOS) and non-line-of-sight (NLOS) conditions is made for relevant cases (Kyosti *et al.*, 2007). The successor of the WINNER II model, the WINNER+ model, has appeared in 2010 and came with significant improvements in path loss models, especially for indoor scenarios (Poutanen, 2011).

Deterministic Channel Modeling: Ray-Tracing

Ray-tracing simulations can effectively be utilized to characterize all the aforementioned channel parameters for a particular scenario, assuming that the propagation environment is represented meticulously within the simulation, i.e., the 3D scenario model involves sufficient geometrical detail, antennas are incorporated with their patterns accordingly and material properties are accurate. A single simulation can provide information about the instantaneous channel gain, feasible multi-paths and corresponding time delays as well as the directional properties of the transmitted and received waves for a single frequency point and a single time instance where the scenario is considered to be static (Yun and Iskander, 2015). In case a wide-band analysis or time-dependent variations of the channel (due to the movements of the antennas or objects in the environment) are needed, then separate simulations can be performed for each individual frequency and time instance. Ray-tracing can provide very accurate results but the validity is generally limited to the particular scenario which is simulated.

1.1.3 Channel Modeling for Multi-Antenna Communication Systems

An accurate channel model is essential to analyze any wireless communication system, though the availability of such models for multi-antenna networks, i.e., MIMO systems, is more important, since they are typically more complex compared to the conventional Single-Input Single-Output (SISO) systems. One of the most promising developments in MIMO technology lately has been the recognition of the benefits of large base station arrays in multi-user systems, which is nowadays known as massive MIMO (Marzetta,

2010). Although it is common practice to employ traditional MIMO channel models (i.e., mostly geometry-based physical stochastic models), for massive MIMO scenarios as well, the channel characteristics within these two schemes may differ (Gao, 2016).

Two important channel related phenomena, which may commonly be observed in massive MIMO systems when the number of base station antennas is asymptotically increased, are the declining channel correlation and fading (i.e., favorable propagation and channel hardening) (Idowu-Bismark *et al.*, 2017). The directional properties of the channel, such as angle of arrival or angle of departure, are, therefore, decisive to characterize the performance accurately. An accurate depiction of the wave propagation phenomenon is also an important requirement for massive MIMO channel modeling, since the base station is likely to be electrically large (due to many antenna elements being used), thus, the Rayleigh distance for the array might be comparable to the distance to the users. As a result of this, conventional far-field assumptions may not be valid, i.e., the influence of non-planar wavefront characteristics should be taken into account (Wu, 2015).

The aforementioned factors imply that deterministic modeling approaches, such as ray-tracing, are well suited for channel modeling in massive MIMO scenarios. Numerous studies in the literature have addressed this topic, where ray-tracing simulations have been utilized to characterize the channel properties as well as the achievable data rate in various massive MIMO scenarios (Liu *et al.*, 2020; Ademaj *et al.*, 2019). In many cases, comparisons to measurements or different channel modeling techniques are also given (Valle *et al.*, 2020; Ai *et al.*, 2017). The studies typically demonstrate how the channel characteristics and data rate performance are according to the simulations, when the number of base station antennas is fixed. However, the variation of these attributes within ray-tracing simulations, as the number of base station antennas is changed, is relatively unknown. The literature on the differences of ray-tracing and stochastic channel modeling for massive MIMO is also limited.

1.2 Improvement Strategies for Ray-Tracing Simulations

The traditional ray-tracing techniques can be described as unidirectional methods, as the rays are launched from a transmit antenna and collected at a receive antenna[2] (assuming that an SBR algorithm is utilized). An important prerequisite of the simulation process is to determine the size of the reception spheres, and of the diffraction cylinders, if wedge diffractions are also considered. Numerous studies in the literature have addressed the difficulties on finding the optimal size for such structures and, therefore, have introduced practical guidelines, i.e., wavelength-dependent size selection, though there are limitations associated with the approach in general (Iskander and Zhengqing Yun, 2002; Seidel and Rappaport, 1994). In particular, a good accuracy can be achieved by utilizing small reception spheres and an accordingly large number of ray launches. For large and complex scenarios, which might involve many ray interactions and considerable distance between the transceivers, it can be presumed that the number of ray launches should be increased so that the feasible ray paths can still be identified, i.e., relevant rays do not miss the reception spheres (Taygur *et al.*, 2018). However, increasing the number of rays typically yields a noticeable growth in the utilization of computational

[2]Unless the scenario includes certain materials which may violate the antenna reciprocity, it is feasible to launch rays from a receive antenna as well.

resources and simulation time. Enlarging the reception spheres for avoiding ray misses with a small number of ray launches may not be a viable solution either, since incorrect ray contributions, which in fact correspond to infeasible ray paths, can be captured by the reception spheres, and diminish the accuracy.

In order to solve this problem, a bidirectional ray-tracing approach can be utilized, instead of the conventional unidirectional ray-tracing where the rays are launched from one antenna and collected at the other. In the bidirectional ray-tracing method, both receiver and transmitter antennas are used for ray launching and the rays are captured on a large interaction surface instead of a small reception sphere. The interaction surface acts merely as a checkpoint for sampling the incident fields from both antennas and does not modify the ray trajectories or alter the field in any way. The field samples on the surface are then used to compute the antenna transfer function by evaluating a reciprocity integral. A crucial advantage of large interaction surfaces over reception spheres is that the size and shape can be chosen with more flexibility. Thus, large and complex scenarios can effectively be simulated by adjusting the interaction surface, rather than increasing the number of ray launches, without compromising accuracy or causing an increase in the simulation time. In general, the algorithm follows the principles of a unidirectional ray-tracer, except the ray capturing mechanism.

The use of flexible interaction surfaces and the according elimination of reception spheres is particularly advantageous in millimeter-wave scenarios, since the conventional approach with reception spheres may yield significant phase errors as well as a considerable growth in the simulation time. Here, the bidirectional ray-tracing method can be utilized to mitigate the phase errors. The calculation of the transfer function is carried out by evaluating a reciprocity integral using efficient asymptotic solution techniques (e.g., stationary phase method) (Taygur and Eibert, 2020b).

The interaction surfaces are not only utilized to replicate the functionality of reception spheres, but can also be implemented as open surfaces and be employed in diffraction scenarios, where UTD-based diffraction simulations can be avoided. Such an arrangement can be particularly advantageous in multiple diffraction scenarios, since UTD-based calculations typically yield inaccurate results if the wave propagates near optical boundaries. Furthermore, the computation time can be reduced, as no new rays are generated by means of Keller cones at each diffraction point (Taygur *et al.*, 2019a).

1.3 Scope and Outline

This dissertation mainly addresses certain shortcomings of conventional GO-UTD and SBR based ray-tracing simulations, and introduces a solution to the problem, i.e., a novel ray-tracing approach, namely, bidirectional ray-tracing, which can be utilized for channel modeling. The accuracy and computational performance of the new approach are compared to the conventional ray-tracing as well as other modeling methods where relevant. The downlink performance of two massive MIMO scenarios are simulated with bidirectional ray-tracing, where conventional ray-tracing algorithms are not suitable due to propagation conditions and general features of the scenarios.

A second contribution of the dissertation is the comparison of deterministic and stochastic channel modeling approaches for four different massive MIMO scenarios, where important properties of massive MIMO channels are investigated by means of traditional ray-tracing simulations and appropriate statistical channel models. Thus,

the similarities as well as the differences of deterministic and stochastic channel modeling techniques are investigated for massive MIMO networks. The comparisons are mainly based on the average downlink data rate according to both approaches, however, additional performance criteria are also taken into account where relevant. As the use of a large number of base station antennas is one of the major features of massive MIMO, the number of the base station antennas is generally varied in performance analyses and the effects of the array size are investigated.

The dissertation begins with a theoretical background on Geometrical Optics, Uniform Theory of Diffraction and their use in ray-tracing simulations in Chapter 2. Starting from Maxwell's equations, the formulations which describe wave propagation phenomena and essential wave-material interactions according to high-frequency approximations, are given and principles of ray-tracing simulations are presented.

In Chapter 3 the bidirectional ray-tracing approach is introduced, the simulation process is explained in detail and relevant performance comparisons to the traditional unidirectional ray-tracing as well as other relevant methods are given. Here, different use cases are considered, i.e., replacing conventional reception spheres by larger interaction surfaces, simulation of single and multiple diffraction scenarios, and treatment of millimeter-wave scenarios.

Chapter 4 gives a brief introduction to conventional MIMO as well as multi-user MIMO (MU-MIMO) first, and then explains the peculiarities of massive MIMO. Certain channel related phenomena, which are unique to massive MIMO and typically not observed in MIMO or MU-MIMO, are elaborated in detail. The benefits of utilizing a large number of base station antennas is illustrated with two numerical examples.

In Chapter 5 and 6, a performance characterization for various massive MIMO scenarios is presented, where the ray-tracing simulations were carried out with unidirectional and bidirectional ray-tracing approaches, respectively.

Finally, a summary of the contributions and an outlook on future work is given in Chapter 7.

Chapter 2

Principles of Ray-Tracing Simulations

The high-frequency assumptions which yield the Geometrical Optics (GO) and Uniform Theory of Diffraction (UTD) formulations from Maxwell's equations, will be presented in detail. The common wave interactions, which involve two different media, will be given. The ray-tracing simulation methodology, which relies on GO-UTD, will be explained. Finally, a method for verifying the accuracy of a ray-tracing simulator will be described through a numerical example.

2.1 High-Frequency Approximation for Maxwell's Equations

Maxwell's equations, which describe the behavior of electromagnetic fields, can be given by (Harrington, 2001)

$$\nabla \times \boldsymbol{E}(\mathbf{r},t) = -\frac{\partial \boldsymbol{B}(\mathbf{r},t)}{\partial t} - \boldsymbol{M}(\mathbf{r},t), \; \nabla \times \boldsymbol{H}(\mathbf{r},t) = \frac{\partial \boldsymbol{D}(\mathbf{r},t)}{\partial t} + \boldsymbol{J}(\mathbf{r},t),$$
$$\nabla \cdot \boldsymbol{D}(\mathbf{r},t) = \rho_e(\mathbf{r},t), \; \nabla \cdot \boldsymbol{B}(\mathbf{r},t) = \rho_m(\mathbf{r},t), \tag{2.1}$$

with the material relations

$$\boldsymbol{D}(\mathbf{r},t) = \epsilon(\mathbf{r})\,\boldsymbol{E}(\mathbf{r},t), \; \boldsymbol{B}(\mathbf{r},t) = \mu(\mathbf{r})\,\boldsymbol{H}(\mathbf{r},t), \tag{2.2}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic field, $\boldsymbol{D}$ and $\boldsymbol{B}$ are the electric and magnetic flux density, $\boldsymbol{J}$ and $\boldsymbol{M}$ are the electric and magnetic current density, ρ_e and ρ_m are electric and magnetic charge density, ϵ and μ are dielectric permittivity and magnetic permeability, respectively[1]. Assuming a time-harmonic dependence with a factor of $e^{j\omega t}$, and a source-free, homogeneous, isotropic medium, the Eqs. (2.1) and (2.2) can be simplified into (Balanis, 2012)

$$\nabla \times \mathbf{E}(\mathbf{r},\omega) = -j\omega\mu\mathbf{H}(\mathbf{r},\omega), \; \nabla \times \mathbf{H}(\mathbf{r},\omega) = j\omega\epsilon\mathbf{E}(\mathbf{r},\omega),$$
$$\nabla \cdot \mathbf{D}(\mathbf{r},\omega) = 0, \; \nabla \cdot \mathbf{B}(\mathbf{r},\omega) = 0, \tag{2.3}$$

where $\mathbf{r}$ and ω denote the coordinates in three-dimensional space and the angular frequency, respectively. Using Eqs. (2.3), the vector Helmholtz equation for electric and

[1] Note that the magnetic current density $\mathbf{M}$ and magnetic charge density ρ_m may not exist in reality.

magnetic fields can be derived as

$$\nabla^2\mathbf{E}(\mathbf{r},\omega)+\omega^2\mu\epsilon\mathbf{E}(\mathbf{r},\omega)=0,$$
$$\nabla^2\mathbf{H}(\mathbf{r},\omega)+\omega^2\mu\epsilon\mathbf{H}(\mathbf{r},\omega)=0, \tag{2.4}$$

with $k=\omega\sqrt{\mu\epsilon}=\frac{2\pi}{\lambda}$, where λ is the wavelength and k denotes the wavenumber. The solution for electric and magnetic fields in free-space can be then expressed by a series expansion in the form of (Kline, 1962; McNamara *et al.*, 1990)

$$\mathbf{E}(\mathbf{r},\omega)=e^{-jk_0S_e(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r},\omega)}{(j\omega)^u},$$
$$\mathbf{H}(\mathbf{r},\omega)=e^{jk_0S_e(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{H}_u(\mathbf{r},\omega)}{(j\omega)^u}, \tag{2.5}$$

where $k_0=\omega\sqrt{\mu_0\epsilon_0}$ indicates the wavenumber for free space. For the asymptotic case where the frequency tends to infinity (i.e., $\omega\to\infty$), the solution can be expressed in terms of the leading term of the series sum by

$$\mathbf{E}(\mathbf{r},\omega)\sim e^{-jk_0S_e(\mathbf{r})}\mathbf{E}_0(\mathbf{r},\omega),$$
$$\mathbf{H}(\mathbf{r},\omega)\sim e^{-jk_0S_e(\mathbf{r})}\mathbf{H}_0(\mathbf{r},\omega), \tag{2.6}$$

where $S_e(\mathbf{r})$ is often called an eikonal, which describes the phase variation along the propagation path, and satisfies

$$|\nabla S_e(\mathbf{r})|^2=\eta^2, \tag{2.7}$$

which is known as the eikonal equation, where $\eta=\sqrt{\epsilon_r\mu_r}$ is the refractive index of the medium, ϵ_r and μ_r are the relative permittivity and relative permeability of the medium, respectively[2] (Born and Wolf, 1999).

In the following sections, the field related quantities will exclusively be given under the assumptions of the asymptotic case, therefore, the subscript 0 will be omitted and the electric and magnetic field terms $\mathbf{E}_0$, $\mathbf{H}_0$ will be written as $\mathbf{E}$, $\mathbf{H}$, respectively, for the sake of brevity.

The surfaces, for which $S_e(\mathbf{r})$ has a constant value, indicate the wavefronts, and the propagation (as well as the Poynting vector) is in the direction of the normal of these equiphase surfaces. A ray may then be defined as a trajectory which is orthogonal to the wavefront (Born and Wolf, 1999). The electric and magnetic field vectors as well as the propagation direction vector are mutually orthogonal at every point along the ray. Therefore, the electromagnetic wave can be described as locally plane (McNamara *et al.*, 1990). The propagation direction, and the direction of power flow (i.e., the direction of the Poynting vector), can be given by $\hat{\mathbf{s}}=\nabla S_e(\mathbf{r})/\eta$. Note that the eikonal equation has different solutions for different wavefronts (i.e., planar, cylindrical or spherical). The relation between the electric and magnetic field can be given by

$$\hat{\mathbf{s}}\times\mathbf{E}=\mathbf{H}Z_w,\ Z_w=\sqrt{\frac{\mu}{\epsilon}}, \tag{2.8}$$

[2]Relative permittivity (ϵ_r) and relative permeability (μ_r) can in general be given as complex quantities where the imaginary part accounts for the losses i.e., due to the conductivity of the material.

where Z_w denotes the wave impedance.

In order to demonstrate how the ray trajectories can be represented in terms of the eikonal expression $S_e(\mathbf{r})$, let us define $\mathbf{r}(s)$ as a function which denotes the points on a particular ray path for an arc length parameter of s such that

$$\frac{d\mathbf{r}(s)}{ds} = \hat{\mathbf{s}} = \frac{\nabla S_e(\mathbf{r}(s))}{\eta(\mathbf{r}(s))}. \tag{2.9}$$

Using the properties of the gradient operator, the relation between $\mathbf{r}(s)$ and the eikonal can be expressed as (Keller, 2014; Lahiri, 2016)

$$\frac{dS_e(\mathbf{r}(s))}{ds} = \nabla S_e(\mathbf{r}(s)) \cdot \frac{d\mathbf{r}(s)}{ds}, \tag{2.10}$$

$$\frac{d}{ds}\left(\eta(\mathbf{r}(s))\frac{d\mathbf{r}(s)}{ds}\right) = \nabla\eta(\mathbf{r}(s)). \tag{2.11}$$

For homogeneous media, Eq. (2.11) can be simplified into

$$\frac{d^2\mathbf{r}(s)}{ds^2} = 0, \tag{2.12}$$

which implies that $\mathbf{r}$ and $S_e(\mathbf{r})$ can be written in the form of

$$\begin{aligned} \mathbf{r}(s) &= s\hat{\mathbf{s}} + \mathbf{a}, \\ S_e(\mathbf{r}(s)) &= \eta\hat{\mathbf{s}} \cdot \mathbf{r}(s) + S_0, \end{aligned} \tag{2.13}$$

where $\mathbf{a}$ and S_0 are a constant vector and a constant scalar, respectively (Keller, 2014; Yun and Iskander, 2015). Hence, the rays in homogeneous media can be represented as straight lines. The variation of the phase along such a path can be represented by

$$e^{-jk_0 S_e(s)} = e^{-jk_0 S_e(0)} e^{-jks}, \tag{2.14}$$

where $S_e(s)$ and $S_e(0)$ indicate two equiphase surfaces which are apart from each other by a distance of s. Using Eq. (2.10), the optical path length formula can be given as

$$\int_{s_1}^{s_2} \frac{dS_e(\mathbf{r}(s))}{ds} ds = \int_{s_1}^{s_2} \eta ds = S_e(s_2) - S_e(s_1), \tag{2.15}$$

where s_1 and s_2 correspond to two distinct arc lengths which are associated with a ray (Born and Wolf, 1999). Eq. (2.15) can be considered as an exemplification of the Fermat principle of least time, which states that the optical length of a ray between any two points is shorter than the optical length of any other curve which connects these points, and is stationary with respect to the variations of the ray path (Born and Wolf, 1999; Taygur and Eibert, 2020a).

Although the notion of eikonal may effectively be utilized to explain various important concepts such as the definition of a ray, power flow and the variation of the phase along the ray path, it is not sufficient to describe the variation of the amplitude of the electric and magnetic field. In order to resolve this, transport equations can be derived from the Helmholtz equations for $\mathbf{E}$ and $\mathbf{H}$ in the form of

$$\begin{aligned} \mathbf{E}(\mathbf{r}(s))\nabla^2 S_e(\mathbf{r}(s)) + 2(\nabla S_e(\mathbf{r}(s)) \cdot \nabla)\mathbf{E}(\mathbf{r}(s)) = 0, \\ \mathbf{H}(\mathbf{r}(s))\nabla^2 S_e(\mathbf{r}(s)) + 2(\nabla S_e(\mathbf{r}(s)) \cdot \nabla)\mathbf{H}(\mathbf{r}(s)) = 0. \end{aligned} \tag{2.16}$$

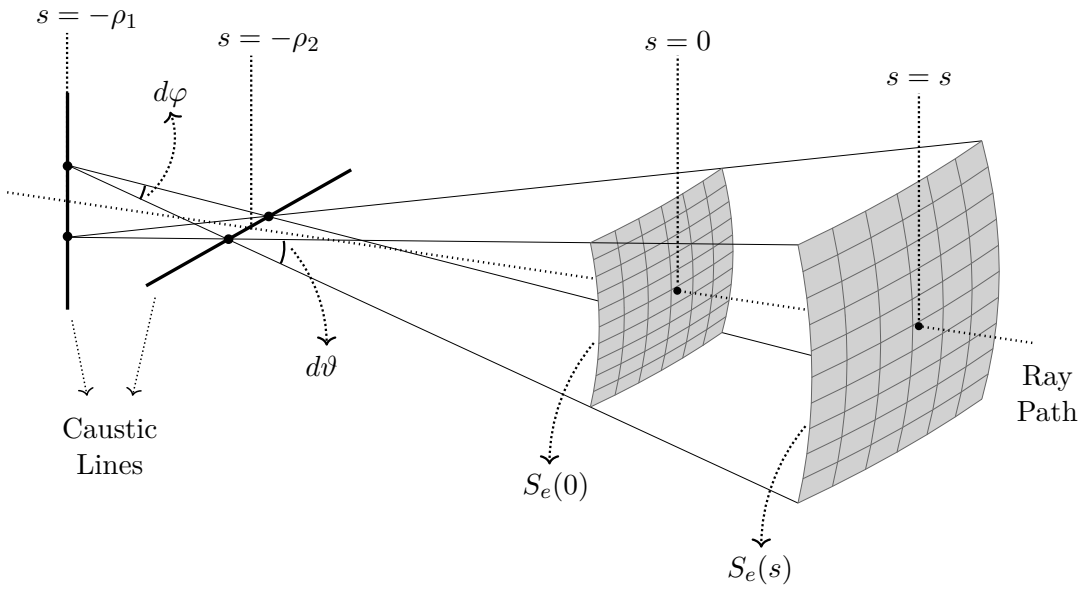


Figure 2.1: Astigmatic ray tube representation.

For the sake of brevity, let us focus on the first equation in (2.16), which involves the electric field, and express $\mathbf{E}$ as well as S solely in terms of s, omitting the $\mathbf{r}$ term. The solution of Eq. (2.16) can then be written as

$$\mathbf{E}(s) = \mathbf{E}(0)e^{\zeta(s)},$$
$$\zeta(s) = -\frac{1}{2}\int_0^s \nabla^2 S_e\left(s'\right) ds'. \tag{2.17}$$

An explicit solution can be given by means of an astigmatic ray tube[3] where two particular equiphase surfaces, $S_e\left(0\right)$ and $S_e\left(s\right)$, are apart from each other by a distance of s along the corresponding ray path. The surfaces are assumed to have radii of curvatures given by ρ_1, ρ_2 and $(\rho_1 + s)$, $(\rho_2 + s)$, respectively, as shown in Fig. 2.1. The solution can be written by means of the Gaussian curvatures of the wavefronts such that (McNamara *et al.*, 1990)

$$G(0) = \frac{1}{\rho_1\rho_2},$$
$$G(s) = \frac{1}{\left(\rho_1 + s\right)\left(\rho_2 + s\right)}, \tag{2.18}$$

with (Kouyoumjian, 1965)

$$\frac{dG}{ds} = -G\left(s\right)\nabla^2 S_e\left(s\right)\frac{1}{\eta},$$
$$\frac{G\left(s\right)}{G\left(0\right)} = e^{2\zeta(s)}. \tag{2.19}$$

[3]Astigmatic ray tubes can be considered as the general form to characterize the wavefronts, which occur in practically relevant geometrical optics problems, since the most commonly encountered waveforms, in particular, the spherical, cylindrical and plane waves, may in fact be regarded as special cases of astigmatic waves (McNamara *et al.*, 1990).

Hence, the relation between $\mathbf{E}(0)$ and $\mathbf{E}(s)$, taking both amplitude and phase variations into account, can be expressed by

$$\mathbf{E}(s) = \mathbf{E}(0)e^{-jk_0 S_e(0)}\sqrt{\frac{\rho_1\rho_2}{(\rho_1+s)(\rho_2+s)}}e^{-jks}, \tag{2.20}$$

which is commonly known as the geometrical optics field. Note that the polarization remains constant during propagation, hence (Born and Wolf, 1999)

$$\hat{\mathbf{e}}(s) = \frac{\mathbf{E}(s)}{||\mathbf{E}(s)||},\ \frac{d\hat{\mathbf{e}}(s)}{ds} = 0 \tag{2.21}$$

is valid. The conservation of power within the ray tube can be demonstrated as (Balanis, 2012)

$$\begin{gathered} dA_0 = \rho_1\rho_2 d\vartheta d\varphi,\ dA_s = (\rho_1+s)(\rho_2+s)d\vartheta d\varphi, \\ \frac{||\mathbf{E}(0)||^2}{||\mathbf{E}(s)||^2} = \frac{dA_s}{dA_0}, \end{gathered} \tag{2.22}$$

where dA_0 and dA_s denote infinitesimally small areas on $S_e(0)$ and $S_e(s)$, respectively. Note that the field expression given in Eq. (2.20) is invalid for $s = -\rho_1$ or $s = -\rho_2$, as the amplitude becomes infinite (Dooren, 1994). This causes so-called caustics to emerge at $s = -\rho_1$ and $s = -\rho_2$, as shown in Fig. 2.1.

2.2 Wave Interactions Under High-Frequency Assumptions

The interactions of geometrical optical fields can be considered as local phenomena, where the field expressions mainly depend on the local properties of the incident field (i.e., magnitude, phase and polarization), as well as the geometrical and electrical features of the different media at the interaction point (Balanis, 2012; Borovikov and Kinber, 1994). Here, three interactions, which may be regarded as fundamentally important, can be given, namely, reflection, refraction, and diffraction (Silver, 1984). The points, where interactions occur, are generally assumed to be a part of an infinite interface (i.e., an infinite planar surface for reflection/refraction and an infinitely long edge for diffraction), in accordance with the locally plane wave behavior of geometrical optics fields (James, 2007).

Another interaction, which has practical use in ray-tracing simulations, is diffusive scattering. In particular, diffusive scattering is utilized to characterize the behavior of a wave when it is incident on a rough surface.

2.2.1 Reflection and Refraction

Consider a planar interface between two homogeneous lossless dielectric media, which exhibit a distinct dielectric permittivity (for the sake of simplicity, the magnetic permeability of practically relevant dielectric materials might be assumed $\mu_r \approx 1$ (Marchenko *et al.*, 2003)), and specified as medium 1 and 2 as shown in Fig. 2.2. An electromagnetic wave, which is incident on this interface within medium 1 and propagates towards medium 2, will partially be reflected back and partially be transmitted into medium 2.

The relation between the incident electric field $\mathbf{E}^i$ on the interface and the reflected electric field $\mathbf{E}^r$ along the corresponding reflected ray path can be given by

$$\mathbf{E}^r(s) = \mathbf{E}^i(\mathbf{r}^r) \cdot \overline{\mathbf{R}} \sqrt{\frac{\rho_1^r \rho_2^r}{(\rho_1^r + s)(\rho_2^r + s)}}\, e^{-jk_1 s}, \tag{2.23}$$

where $\overline{\mathbf{R}}$ is the dyadic reflection coefficient, $\mathbf{r}^r$ denotes the point of reflection, k_1 is the wavenumber in medium 1, ρ_1^r and ρ_2^r are the radii of curvature of the reflected wave at $\mathbf{r}^r$. A similar expression can be given for the refracted/transmitted wave $\mathbf{E}^t$ by

$$\mathbf{E}^t(s) = \mathbf{E}^i(\mathbf{r}^r) \cdot \overline{\mathbf{T}} \sqrt{\frac{\rho_1^r \rho_2^r}{(\rho_1^r + s)(\rho_2^r + s)}}\, e^{-jk_2 s}, \tag{2.24}$$

where $\overline{\mathbf{T}}$ is the dyadic refraction/transmission coefficient, and k_2 is the wavenumber in medium 2. Hence, the main concern is to find appropriate expressions for $\overline{\mathbf{R}}$ and $\overline{\mathbf{T}}$. If the electric field terms are decomposed as

$$\begin{aligned} \mathbf{E}^{\{r,t,i\}} &= \mathbf{E}_{\parallel}^{\{r,t,i\}} + \mathbf{E}_{\perp}^{\{r,t,i\}} = \hat{\mathbf{e}}_{\parallel}^{\{r,t,i\}} E_{\parallel}^{\{r,t,i\}} + \hat{\mathbf{e}}_{\perp}^{\{r,t,i\}} E_{\perp}^{\{r,t,i\}}, \\ E_{\parallel}^{\{r,t,i\}} &= \mathbf{E}^{\{r,t,i\}} \cdot \hat{\mathbf{e}}_{\parallel}^{\{r,t,i\}}, \quad E_{\perp}^{\{r,t,i\}} = \mathbf{E}^{\{r,t,i\}} \cdot \hat{\mathbf{e}}_{\perp}^{\{r,t,i\}}, \end{aligned} \tag{2.25}$$

then Eq. (2.23) and (2.24) for $s = 0$ can be written in matrix form as

$$\begin{aligned} \mathbf{E}^r(0) = \mathbf{E}^i(\mathbf{r}^r)\, \mathbf{R}, \;\; \mathbf{E}^t(0) = \mathbf{E}^i(\mathbf{r}^r)\, \mathbf{T}, \\ \begin{bmatrix} E_{\parallel}^r \\ E_{\perp}^r \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} E_{\parallel}^i \\ E_{\perp}^i \end{bmatrix}, \quad \begin{bmatrix} E_{\parallel}^t \\ E_{\perp}^t \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} E_{\parallel}^i \\ E_{\perp}^i \end{bmatrix}, \end{aligned} \tag{2.26}$$

where $\hat{\mathbf{e}}_{\parallel}^{\{r,t,i\}}$ and $\hat{\mathbf{e}}_{\perp}^{\{r,t,i\}}$ denote the unit vectors which are parallel and perpendicular to the plane of incidence for the corresponding fields (i.e., reflected, refracted, incident), respectively, and $E_{\parallel}^{\{r,t,i\}}$ and $E_{\perp}^{\{r,t,i\}}$ are the complex amplitude scalars which are polarized parallel and perpendicular with respect to the plane of incidence for the corresponding fields, respectively. Let us also define the unit vectors, which indicate the propagation direction for reflected, refracted and incident rays by $\hat{\mathbf{s}}^r$, $\hat{\mathbf{s}}^t$ and $\hat{\mathbf{s}}^i$, and the unit surface normal for the interface between the two media by $\hat{\mathbf{n}}$ (pointing towards medium 1). The relation between these vectors can be given by (Hussain, 2017)

$$\begin{aligned} \hat{\mathbf{e}}_{\perp}^r = \hat{\mathbf{e}}_{\perp}^t = \hat{\mathbf{e}}_{\perp}^i = \frac{\hat{\mathbf{n}} \times \hat{\mathbf{s}}^{\{r,t,i\}}}{\left|\left|\hat{\mathbf{n}} \times \hat{\mathbf{s}}^{\{r,t,i\}}\right|\right|}, \\ \hat{\mathbf{e}}_{\parallel}^{\{r,t,i\}} = \hat{\mathbf{s}}^{\{r,t,i\}} \times \hat{\mathbf{e}}_{\perp}^{\{r,t,i\}}, \end{aligned} \tag{2.27}$$

with

$$\cos^{-1}\left(-\hat{\mathbf{s}}^i \cdot \hat{\mathbf{n}}\right) = \cos^{-1}\left(\hat{\mathbf{s}}^r \cdot \hat{\mathbf{n}}\right) \quad \longrightarrow \quad \theta^i = \theta^r, \tag{2.28}$$

$$k_1 \cos^{-1}\left(-\hat{\mathbf{s}}^i \cdot \hat{\mathbf{n}}\right) = k_2 \cos^{-1}\left(-\hat{\mathbf{s}}^t \cdot \hat{\mathbf{n}}\right) \quad \longrightarrow \quad k_1 \theta^i = k_2 \theta^t, \tag{2.29}$$

where Eq. (2.28) and Eq. (2.29) are known as Snell's law of reflection and refraction, respectively (Marchenko *et al.*, 2003). The angles $\theta^{\{i,r,t\}}$ denote the incidence angles

of the corresponding rays (incident, reflected, refracted/transmitted) with respect to the incidence plane. The tangential components of the electric and magnetic field vectors should be continuous on the interface, i.e.,

$$\begin{aligned} E_x^i + E_x^r &= E_x^t, \\ E_y^i + E_y^r &= E_y^t, \\ H_x^i + H_x^r &= H_x^t, \\ H_y^i + H_y^r &= H_y^t \end{aligned} \tag{2.30}$$

holds true, where $E_{\{x,y,z\}}$, $H_{\{x,y,z\}}$ denote the components of the electric and magnetic field in x, y, z directions, respectively. Hence, using the identity in Eq. (2.8)

$$\begin{aligned} E_\perp^i + E_\perp^r &= E_\perp^t, \\ \sqrt{\frac{\epsilon_1}{\mu_1}} \cos\theta^i \left(E_\perp^i - E_\perp^r\right) &= \sqrt{\frac{\epsilon_2}{\mu_2}} \cos\theta^t E_\perp^t, \\ \cos\theta^i \left(E_\parallel^i - E_\parallel^r\right) &= \cos\theta^t E_\parallel^t, \\ \sqrt{\frac{\epsilon_1}{\mu_1}} \left(E_\parallel^i + E_\parallel^r\right) &= \sqrt{\frac{\epsilon_2}{\mu_2}} E_\parallel^t \end{aligned} \tag{2.31}$$

can be written where ϵ_1, ϵ_2 and μ_1, μ_2 denote the permittivity and permeability for medium 1 and 2, respectively. Consequently, these equations can be solved to obtain the reflection and transmission coefficients in Eq. (2.26) such that (Balanis, 2012)

$$\begin{aligned} R_{12} = R_{21} = T_{12} = T_{21} = 0, \\ R_{11} = R_\parallel = \frac{\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^t - \sqrt{\frac{\mu_1}{\epsilon_1}} \cos\theta^i}{\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^t + \sqrt{\frac{\mu_1}{\epsilon_1}} \cos\theta^i}, \quad R_{22} = R_\perp = \frac{\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^i - \sqrt{\frac{\mu_1}{\epsilon_1}} \cos\theta^t}{\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^i + \sqrt{\frac{\mu_1}{\epsilon_1}} \cos\theta^t}, \\ T_{11} = T_\parallel = \frac{2\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^i}{\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^i + \sqrt{\frac{\mu_1}{\epsilon_1}} \cos\theta^t}, \quad T_{22} = T_\perp = \frac{2\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^i}{\sqrt{\frac{\mu_2}{\epsilon_2}} \cos\theta^t + \sqrt{\frac{\mu_1}{\epsilon_1}} \cos\theta^i}. \end{aligned} \tag{2.32}$$

It should be noted that Eq. (2.32) is valid when the physical size of the interface is substantially larger than the wavelength (Didascalou, 2000).

It is important to briefly mention about a special material configuration here, which is encountered very frequently, in particular, dielectric-PEC material interface. In this case, the incident wave, which propagates within the dielectric medium, is completely reflected back at the interface with an according phase change, as no transmission occurs. Thus, the reflection and transmission coefficients can be given as

$$\begin{aligned} R_{12} = R_{21} = T_{12} = T_{21} = 0, \\ R_{11} = R_{22} = -1, \\ T_{11} = T_{22} = 0. \end{aligned} \tag{2.33}$$

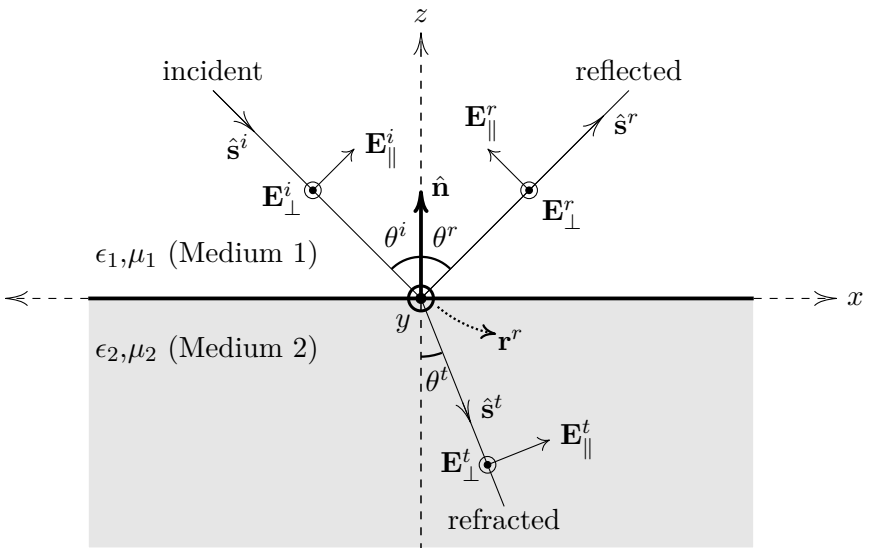


Figure 2.2: Illustration of incident, reflected, and refracted (transmitted) rays on a planar interface composed of two different media.

2.2.2 Diffraction

The propagation of a wave into shadow regions may not be handled within geometrical optics, hence the fields in these regions are predicted to be zero (with an according discontinuity at the shadow boundary), which is not physical (Dooren, 1994). The addition of diffracted rays to the geometrical optics solution might, therefore, be acknowledged as a complementary extension, yielding more accurate results in good agreement with measurements (Balanis, 2012). The foundations of a diffraction theory, which can be incorporated with the geometrical optics approach, has first been conceived by Keller who introduced the Geometrical Theory of Diffraction (GTD) (Keller, 1962). The diffracted ray concept has arose from the asymptotic approximations of certain scattering problems involving basic geometries, such as wedges or half-planes, where an exact solution is known. Nevertheless, the GTD approximation may not be utilized to obtain a complete solution in the entire problem domain as it generally becomes invalid at so-called transition regions (James, 2007). This limitation has later been addressed by Kouyoumjian and Pathak as they introduced the Uniform Theory of Diffraction (UTD), which is an improvement over Keller's theory since the total field is continuous across the transition boundaries (Kouyoumjian and Pathak, 1974).

The behavior of wedge diffracted rays within GTD/UTD formulations can be explained with the Fermat principle of least time where the geometrical optics laws may still be applied, i.e., an eikonal function can be introduced to characterize the propagating wavefronts, and the amplitude expressions might be deduced from the principles of conservation of energy along ray tubes. Similar to the reflection and refraction phenomena, the amplitude expressions for diffracted rays can be obtained by multiplying the incident field expressions with proper diffraction coefficients, which are determined by the wavelength, the angle of incidence and diffraction, and the properties of the medium at the diffraction point (Keller, 1962; McNamara *et al.*, 1990).

As only straight edge diffractions will be considered in this study, let us now consider

a perfectly conducting wedge, as shown in Fig. 2.3, where a ray traveling in $\hat{\mathbf{s}}^i$ direction is incident upon it and hits the point $\mathbf{r}^d$ on the edge. Keller's law of diffraction states that a diffracted ray and the corresponding incident ray have equal angles with the edge at the point of diffraction, provided that they are in the same medium, i.e.,

$$\begin{aligned} \beta^i &= \beta^d, \\ \hat{\mathbf{s}}^d \cdot \hat{\mathbf{d}} &= \hat{\mathbf{s}}^i \cdot \hat{\mathbf{d}}, \end{aligned} \tag{2.34}$$

where $\hat{\mathbf{s}}^d$ denote the propagation direction of an arbitrary diffracted ray and $\hat{\mathbf{d}}$ is the unit vector aligned with the edge. The diffracted rays form a cone around the edge, which is known as Keller cone (Keller, 1962).

The problem domain can be divided into three distinct regions according to the ray interactions observed within the region, and the transitions between the regions are denoted by reflection and shadow boundary lines, respectively. The total field in each region can then be given as the sum of individual field contributions emanating from distinct interactions. The diffracted field ($\mathbf{E}^d$), which may be observed at a distance of s from the diffraction point $\mathbf{r}^d$, can be expressed as

$$\mathbf{E}^d(s) = \mathbf{E}^i(\mathbf{r}^d) \cdot \overline{\mathbf{D}}\; A(s)\; e^{-jks}, \tag{2.35}$$

where $\mathbf{E}^i(\mathbf{r}^d)$ is the incident field at $\mathbf{r}^d$, $\overline{\mathbf{D}}$ is the dyadic diffraction coefficient, and $A(s)$ is the attenuation factor, which can be expressed as

$$A(s) = \sqrt{\frac{\rho_c}{s\,(\rho_c + s)}}, \tag{2.36}$$

where ρ_c is the distance between the diffraction point $\mathbf{r}^d$ (which represents the location of the first caustic for the diffracted wave) and the second caustic prior to the diffraction. If the electric field terms are decomposed into parallel and perpendicular components with respect to the incidence and diffraction planes, respectively, then Eq. (2.35) turns into

$$\begin{bmatrix} E^d_{\beta_0}(s) \\ E^d_{\phi_0}(s) \end{bmatrix} = \begin{bmatrix} D_s & 0 \\ 0 & D_h \end{bmatrix} \begin{bmatrix} E^i_{\beta'_0}(\mathbf{r}^d) \\ E^i_{\phi'_0}(\mathbf{r}^d) \end{bmatrix} A(s)\; e^{-jks},$$

$$E^d_{\beta_0}(s) = \mathbf{E}^d(s) \cdot \hat{\boldsymbol{\beta}}_0, \quad E^d_{\phi_0}(s) = \mathbf{E}^d(s) \cdot \hat{\boldsymbol{\phi}}_0,$$

$$E^i_{\beta'_0}(s) = \mathbf{E}^i(\mathbf{r}^d) \cdot \hat{\boldsymbol{\beta}}'_0, \quad E^i_{\phi'_0}(\mathbf{r}^d) = \mathbf{E}^i(\mathbf{r}^d) \cdot \hat{\boldsymbol{\phi}}'_0, \tag{2.37}$$

where the unit vector pairs which span the incidence and diffraction planes are given by ($\hat{\mathbf{s}}^i$, $\hat{\mathbf{d}}$) and ($\hat{\mathbf{s}}^d$, $\hat{\mathbf{d}}$), respectively. The unit vectors $\hat{\boldsymbol{\beta}}'_0$ and $\hat{\boldsymbol{\beta}}_0$ are then parallel to the incidence and diffraction planes, respectively, and $\hat{\boldsymbol{\phi}}'_0$ and $\hat{\boldsymbol{\phi}}_0$ are perpendicular to the incidence and diffraction planes, respectively. The diffraction coefficients D_s and D_h according to the UTD formulation are

$$\begin{aligned} D_s &= D^i - D^r, \\ D_h &= D^i + D^r, \end{aligned} \tag{2.38}$$

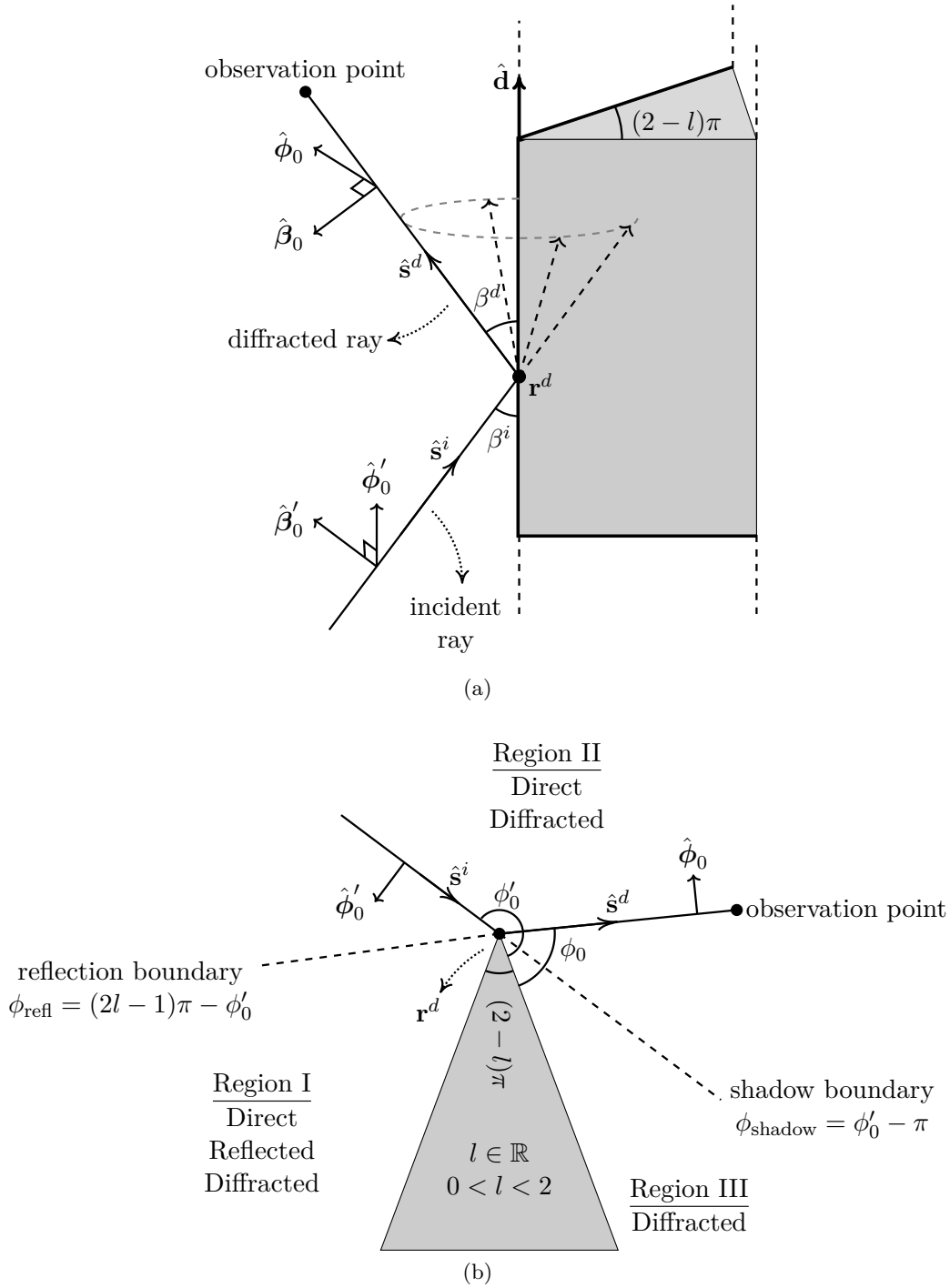


Figure 2.3: Wedge diffraction illustration with frontal (a) and side (b) views.

with the definitions

$$D^i = \frac{e^{-j\pi/4}}{2l\sqrt{2\pi k}\sin\beta^i}\left\{\cot\left[\frac{\pi+(\phi_0-\phi_0')}{2l}\right]F\left[kL\,a^+\left(\phi_0-\phi_0'\right)\right]\right.$$
$$\left.+\cot\left[\frac{\pi+(\phi_0-\phi_0')}{2l}\right]F\left[kL\,a^-\left(\phi_0-\phi_0'\right)\right]\right\},$$
$$D^r = \frac{e^{-j\pi/4}}{2l\sqrt{2\pi k}\sin\beta^i}\left\{\cot\left[\frac{\pi+(\phi_0+\phi_0')}{2l}\right]F\left[kL\,a^+\left(\phi_0+\phi_0'\right)\right]\right.$$
$$\left.+\cot\left[\frac{\pi-(\phi_0+\phi_0')}{2l}\right]F\left[kL\,a^-\left(\phi_0+\phi_0'\right)\right]\right\}. \tag{2.39}$$

The function F, known as the Fresnel transition function, is defined as

$$F(x) = 2j\sqrt{x}e^{jx}\int_{\sqrt{x}}^{\infty} e^{-jt^2}dt. \tag{2.40}$$

Numerical approximations of Eq. (2.40) for various different values of x are found in (Balanis, 2012). The function $a^{\pm}(x)$ is given by

$$a^{\pm}(x) = 2\cos^2\left(\frac{2l\pi N^{\pm}-x}{2}\right), \tag{2.41}$$

where N^+ and N^- are integers which minimize $|2\pi lN^+ - x - \pi|$ and $|2\pi lN^- - x + \pi|$, respectively (Dooren, 1994). The distance parameter L has the form

$$L = \frac{s\left(\rho_d^i+s\right)\rho_1^i\rho_2^i\sin^2\beta^i}{\rho_d^i\left(\rho_1^i+s\right)\left(\rho_2^i+s\right)}, \tag{2.42}$$

where ρ_1^i and ρ_2^i are the radii of curvature of the incident wave at $\mathbf{r}^d$, respectively, and ρ_d^i is the radius of curvature of the incident wavefront at $\mathbf{r}^d$ in the incidence plane (Kouyoumjian and Pathak, 1974). For simple wavefront types (i.e., plane, cylindrical, spherical), the distance function L and the attenuation factor $A(s)$ can be simplified into (Balanis, 2012)

$$L = \begin{cases} s\sin^2\left(\beta^i\right) & \text{plane-wave incidence,} \\ \dfrac{s\sin\left(\beta^d\right)s'\sin\left(\beta^i\right)}{s\sin\left(\beta^d\right)+s'\sin\left(\beta^i\right)} & \text{cylindrical-wave incidence,} \\ \dfrac{ss'\sin^2\left(\beta^i\right)}{s+s'} & \text{spherical-wave incidence,} \end{cases} \tag{2.43}$$

$$A(s) = \begin{cases} \dfrac{1}{\sqrt{s}} & \text{plane-wave incidence,} \\ \dfrac{1}{\sqrt{s\sin\left(\beta^i\right)}} & \text{cylindrical-wave incidence,} \\ \sqrt{\dfrac{s'}{s\left(s+s'\right)}} & \text{spherical-wave incidence,} \end{cases} \tag{2.44}$$

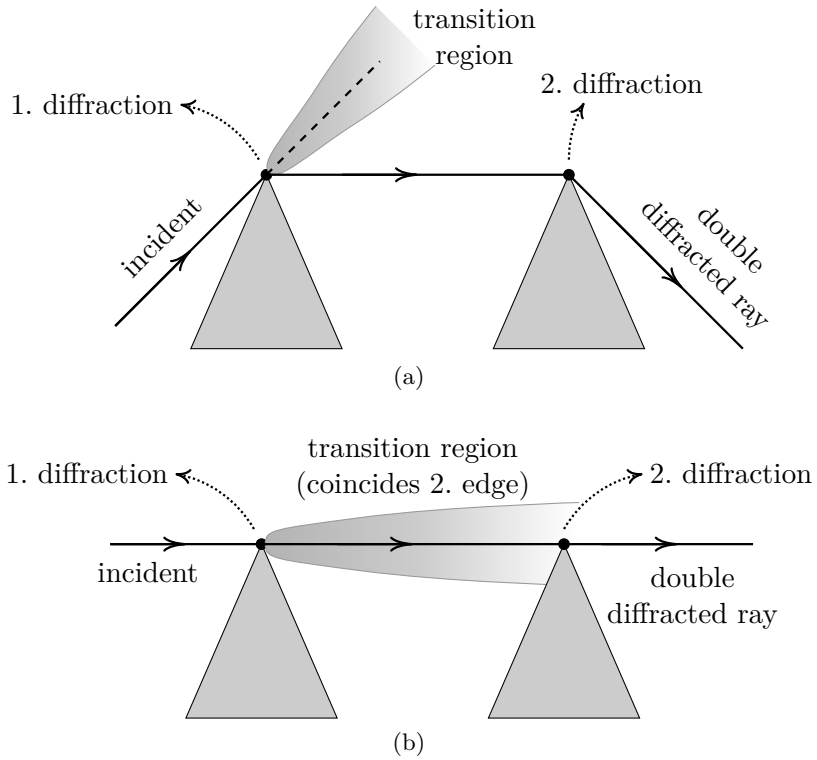


Figure 2.4: Double diffraction illustration where the transition region corresponding to the first diffraction either goes above all the edges (a) or it coincides with the second edge (b).

where s' indicates the path length between the diffraction point $\mathbf{r}^d$ and the location of the source for the incident ray.

Multiple Diffractions

The diffracted rays, which are away from the transition regions, generally comply with ray-optical wave characteristics. Hence, multiple diffractions, which may occur in many practically relevant propagation scenarios, can effectively be handled by applying UTD consecutively for every wedge along the ray path. In the particular case that an edge lies on the transition boundary of a diffracted ray (see Fig. 2.4), applying the UTD equations in a consecutive manner yields inaccurate results, since the field expressions in the transition region after a preceding diffraction may not truly be considered ray-optical (James, 2007; Pathak *et al.*, 2013). The problem has been addressed in numerous studies and various solution methods have been proposed, mainly for double and triple wedge diffraction cases (Carluccio and Albani, 2008; Schneider and Luebbers, 1991; Albani, 2005). Although it might be superfluous to consider more than three diffractions in many practical propagation scenarios (as the field would be significantly weaker compared to that from other ray contributions), certain problems, such as rooftop propagation scenarios, may require these ray paths to be taken into account since there may not be other feasible paths, which yield stronger field contributions, between two particular points. For such cases, certain empirical path-loss models (e.g., Vogler, Dey-

gout, Giovaneli), are usually employed. The actual problem is generally simplified into a basic geometrical model, which involves multiple knife-edges, and the path loss is computed according to the geometrical features of the simplified model (Deygout, 1966; Vogler, 1981; Giovaneli, 1984).

Dielectric Wedge

Diffractions at dielectric wedges are a problem of considerable interest for numerous propagation scenarios, since a perfect electric conductor may not always represent the actual material setting accurately. Therefore, the UTD coefficients, which have been presented before, may not give accurate results. In order to address this issue, several heuristic approaches have been developed. One of the earliest works on the topic, introduced by Luebbers (1984), has shown promising results, where the accuracy is improved significantly over conventional UTD based on perfectly electrically conducting (PEC) wedges, even though the methods may not properly satisfy Maxwell's equations. Later studies have introduced numerous improvements in this aspect, where even better accuracy can be achieved (Holm, 2000; Soni and Bhattacharya, 2010).

2.2.3 Diffusive Scattering

Diffusive scattering phenomena occur when a wave is incident on an object with rough or non-smooth surface, such as building facades (Yun and Iskander, 2015; Pongsilamanee and Bertoni, 2004). Diffusive scattering models are also utilized to depict the scattering behavior in ray-tracing simulations when vegetation is involved in the scenario (Mocker-Henning, 2016). Diffusive scattering can be considered as a simplification of a problem, where the aggregate outcome of many individual diffraction and reflection/refraction interactions is approximated, since it is usually impractical to create a detailed simulation model for the entities with many geometrical irregularities, and to include all the individual interactions (Maurer, 2005). Thus, diffusive scattering is usually incorporated by defining the roughness of the surface/object and selecting a scattering model which characterizes the main aspects of the wave behavior. In general, the incident wave is scattered in non-specular directions (i.e., specular reflection direction is given by Snell's law, as shown before) depending on the roughness of the considered object or surface. However, different models may yield distinct scattering directions where most of the energy goes to. In other words, specular, incident, or other directions (e.g., around the surface normal vector) can be emphasized to a greater extent in a simulation (Mani and Oestges, 2012; Lu *et al.*, 2014a).

2.3 Ray-Tracing Algorithms

A typical ray-tracing simulation involves two main steps, namely, identification of feasible ray-paths and calculation of electromagnetic field expressions. The computation of field expressions is typically based on geometrical optics and uniform theory of diffraction principles. The Fermat principle of least time describes the constraints which apply to a ray, however, it does not specify how the ray paths can practically be computed. There are two common approaches which are utilized for this purpose, namely, the Image Method and the Shooting and Bouncing Rays (SBR) method. In both approaches, there are usually a source and a target point (which typically correspond to the locations

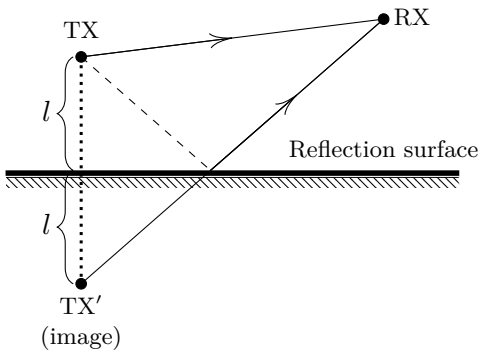


Figure 2.5: Illustration of image method.

of a transmitting and a receiving antenna, respectively), where the feasible ray paths between these two points are calculated.

2.3.1 Image Method

The image method is based on applying the image principle, as the name indicates, on planar reflection surfaces in the geometry. Here, image points for the source (or for the target, alternatively) are calculated first with respect to the visible reflecting surfaces. These points are then connected to the target with a direct ray, as shown in Fig. 2.5. If multiple reflections should be taken into account in the simulation, the process is carried out recursively, in particular, new images are generated from already existing image points (i.e., higher order images).

The major advantage of the image method is that the exact ray paths between two points can be calculated. Nevertheless, it is not a suitable method for simulating large and complicated geometries where many interactions occur (e.g., urban propagation modeling), as the complexity of the simulation may grow exponentially (McKown and Hamilton, 1991).

2.3.2 Shooting and Bouncing Rays Method

The basic premise of the shooting and bouncing rays technique is to launch a large number of rays from the source point and to determine if they reach the target location (Ling *et al.*, 1989). Compared to the image method, the shooting and bouncing rays method can conveniently be utilized for simulating wave propagation in complex environments, such as an urban landscape. Each ray is traced within the scenario, as wave interactions, such as reflection or diffraction, occur, and eventually reaches a receiver or leaves the scenario altogether. In order to ensure that every feasible ray path is discovered, the number of ray launches should generally be sufficiently large.

The reception test is generally performed within a certain tolerance around the receiver location. In particular, a so-called reception sphere is placed at the receiver location and the rays, which hit the sphere, are acknowledged as feasible ray paths. The size of the reception sphere is an important parameter to be determined, since a small sphere may yield significant errors due to valid ray paths being missed whereas a large sphere causes incorrect ray contributions to be captured, which also deteriorates the

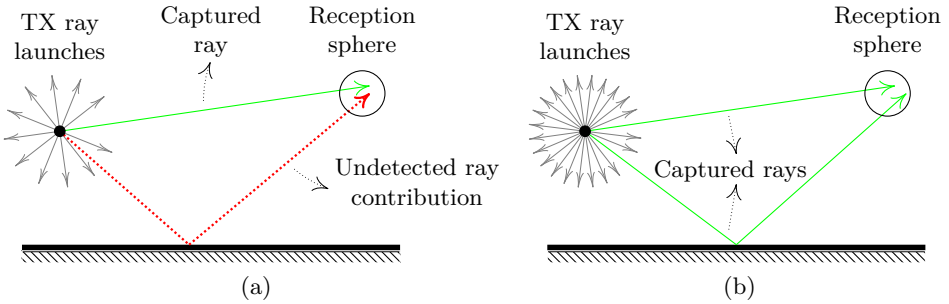


Figure 2.6: Demonstration of the ray miss phenomenon when the number of ray launches is not sufficiently large (a). The problem can be resolved by increasing the number of ray launches without changing the size of the reception sphere (b).

accuracy. Clearly, the number of ray launches is an essential part of this trade-off where the drawbacks of employing small reception spheres might be alleviated to a certain extent by increasing the number of ray launches, as illustrated in Fig. 2.6. However, such an arrangement does not come without a cost, as the computational load usually grows. A practical approach to determine both parameters is simply to pick a sphere size in the order of a wavelength, and then to adjust the number of ray launches accordingly. The number of ray launches should ideally be not more than what is needed to identify the feasible ray paths. In order to determine whether a particular choice for the number of ray launches is sufficient, a convergence analysis can be carried out. In particular, multiple ray-tracing simulations can be performed for a certain scenario where the number of ray launches is gradually increased until the simulation outcome (i.e., electric field at a certain observation point or transfer function between a transceiver pair) between consecutive runs does not change drastically. The break point, where further increase in the number of ray launches does not yield meaningful differences, can be considered as sufficient.

It should be noted that a reception sphere might be hit from multiple rays, which are associated with a single wavefront and are all feasible (see Fig. 2.7). This yields an unwanted situation known as duplicate ray problem where the correct result may actually be obtained by keeping only one ray for each unique wavefront (Iskander and Zhengqing Yun, 2002). The issue can partially be resolved by employing adaptive reception spheres where the size of the sphere is dependent on the angle between adjacent rays and the ray path length. An even more powerful technique for solving the problem is to classify the rays by certain criteria and to remove the duplicates accordingly (Yun *et al.*, 2001). The classification criteria, which may commonly be utilized for duplicate elimination, can include ray path history (i.e., the objects in the geometry that a particular ray interacts with), number of interactions, or angle of transmission/reception.

A structure, which is similar to reception spheres, is utilized for realizing diffraction interactions in order to detect the incident rays coming towards the edge, i.e., thin cylinders placed along the diffracting edge (Mocker-Henning, 2016; Taygur and Eibert, 2020b). Following the hit of an incident ray on the cylinder, a large number of rays are usually generated on the Keller cone accordingly and traced further. It can be inferred that the number of rays in the simulation may grow quickly with this approach, especially if multiple diffractions are taken into account (Taygur *et al.*, 2018).

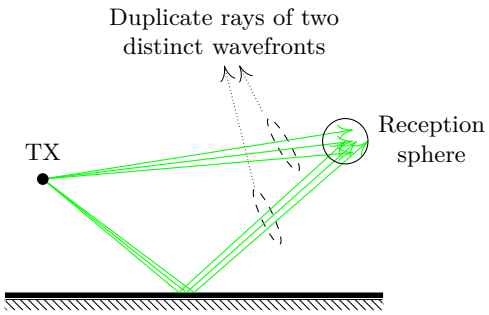


Figure 2.7: Illustration of duplicate ray problem in a scenario with two distinct wavefronts.

Computational Aspects of SBR and Parallelization

In the SBR method, a significant portion of the computation involves the decision whether the rays intersect with any object, and if it does, where they hit. This process requires careful consideration, since a naive algorithmic implementation, which involves a brute force search over every geometrical entity for every single ray, typically yields a poor computational performance. Therefore, acceleration structures are generally utilized to improve the efficiency of ray-object intersection calculations. One of the most commonly utilized schemes for ray-tracing acceleration is bounding volume hierarchy (BVH), which can briefly be described as a geometrical subdivision approach where the scene is divided into sub-volumes in different sizes and the topological relation between the sub-volumes are represented by a hierarchical tree structure (Yun and Iskander, 2015).

A notable trait of the SBR approach is its suitability for parallel computing as the tracing process can independently be carried out for every individual ray, without much need for data exchange or communication between the processes (Glassner, 2007). Hence, acceleration of ray-tracing simulations by means of various parallel computing platforms, especially GPUs, for solving electromagnetic wave propagation problems has been a popular research topic in the last decade (Brem, 2015; Mocker-Henning, 2016). The widespread use of GPUs for wave propagation simulations is a plausible occurrence rather than a coincidence, considering the fact that certain similarities exist between the solutions of image rendering problems in computer graphics and wave propagation modeling in electromagnetics by ray-tracing. As GPU hardware comprises many processing cores, which can run in parallel, a significant speed-up might be achieved (Yun and Iskander, 2015). Modern GPU computing platforms do not only provide a simple parallel programming environment, but also relevant software tools, libraries, and frameworks, such as NVIDIA OptiX™, which is a programmable ray-tracing engine and can handle ray-geometry intersections (Parker *et al.*, 2010). Although it was originally developed to solve various problems in the computer graphics domain, OptiX has successfully been utilized for radio propagation modeling by ray-tracing as well (Kasdorf *et al.*, 2021; Schiller *et al.*, 2015).

2.3.3 Computation of Antenna Transfer Functions

For many practical simulation scenarios, it is necessary to characterize the transfer function for a transmitter-receiver antenna pair, and to consider how the fields emerge initially from a transmitter and what happens when they are collected by a receiving antenna. The field expressions, which are associated with a ray at the beginning of the tracing process, are generally described by the antenna pattern of the transmitter. If a ray reaches a receiver, or in other words hits a reception sphere, the induced voltage at the ports of the receiving antenna is determined by the pattern of the receiving antenna and the incident field (associated with the ray). The antenna transfer function can then be expressed in terms of this induced voltage and the generator voltage at the transmitting antenna site (which is assumed to be already known). It is particularly convenient here to use the so-called vector effective length expression, which relates the receiving characteristics of an antenna with its far-zone field. Let $\mathbf{E}^r\left(\hat{\mathbf{s}}^r,\, r\right)$ be the far-zone electric field for the receiving antenna where $\hat{\mathbf{s}}^r$ and r denote the observation direction and observation distance, respectively. The relation between $\mathbf{E}^r\left(\hat{\mathbf{s}}^r,\, r\right)$ and the vector effective length can then be written as

$$\mathbf{E}^r\left(\hat{\mathbf{s}}^r,\, r\right) = \vec{\ell}_e\left(\hat{\mathbf{s}}^r\right) \frac{k\, I_{in}\, Z_w}{4\pi r} e^{-jkr} \tag{2.45}$$

where $\vec{\ell}_e$ is the vector effective length, Z_w is the wave impedance and I_{in} is the current on the antenna port which generates $\mathbf{E}^r\left(\hat{\mathbf{s}}^r,\, r\right)$ on transmit. Note that $\vec{\ell}_e$ is a function of the observation direction. By using the port voltage-current representation, as in Collin (1985), the induced open-circuit voltage at the receiving antenna port can be written as

$$\begin{aligned} V^{\mathrm{oc}} &= -\frac{1}{I_{in}} \iiint_{\Omega} \mathbf{J}^r \cdot \mathbf{E}^i dV \\ &= \vec{\ell}_e\left(\hat{\mathbf{s}}^r\right) \cdot \mathbf{E}^i\left(-\hat{\mathbf{s}}^r\right), \end{aligned} \tag{2.46}$$

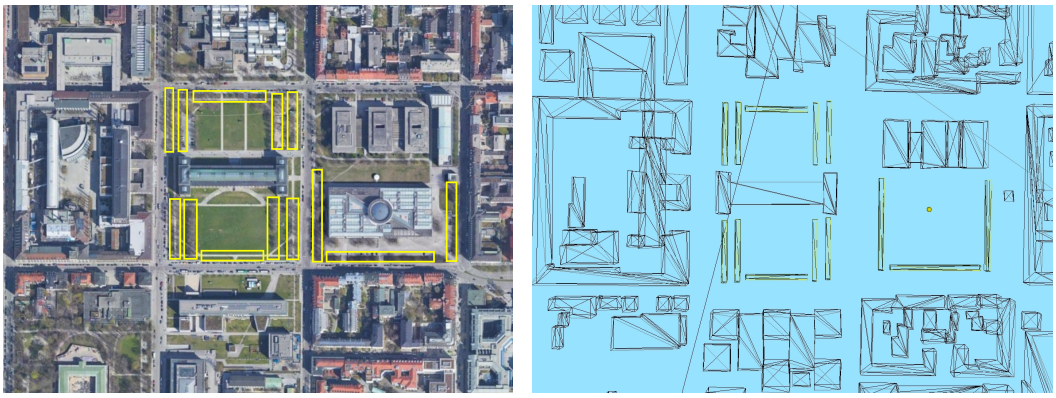
where $\mathbf{J}^r$ is the impressed current at the receiving antenna, Ω is the volume in which the impressed current $\mathbf{J}^r$ is non-zero (i.e., the volume which comprises the receive antenna) and $\mathbf{E}^i$ is the incident electric field upon the antenna. Thus, the antenna transfer function can be given as

$$G = \frac{\vec{\ell}_e\left(\hat{\mathbf{s}}^r\right) \cdot \mathbf{E}^i\left(-\hat{\mathbf{s}}^r\right)}{V^{\mathrm{gen}}}, \tag{2.47}$$

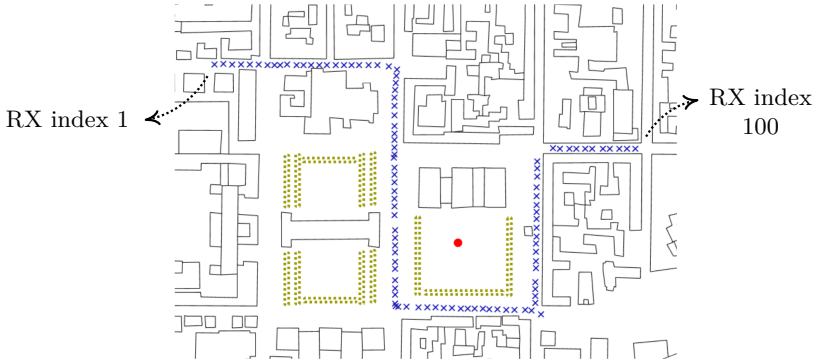
where V^{gen} denotes the generator voltage at the transmitting antenna port.

2.4 Verifying the Accuracy of a Simulation

Since the ray-tracing simulations are utilized to replace the real-world measurement campaigns in many cases, the accuracy can be verified by comparing the measurement data and simulation results for a particular scenario. The ray-tracing simulator, which is utilized in this study, has accordingly been tested with a measurement data set in order to characterize the accuracy of the simulator. The data set consists of channel gain information for a fixed transmitter and a mobile receiver moving in Munich downtown (Damasso and Correa, 1999). The considered operating frequency within the



(a) (b)



(c)

Figure 2.8: Satellite image (a), simulation model (b), and measurement locations (c) for the Munich downtown scenario. Blue crosses and the large red dot in (c) show the receiver locations for each individual measurement and the base station location, respectively. The foliage blocks are indicated by yellow boxes (Taygur and Eibert, 2021a). © 2021 IEEE.

measurement data is 947 MHz and 100 individual measurement points are utilized for the assessment.

An accurate depiction of the physical and electrical properties of the objects in the scenario is decisive for achieving a good accuracy in ray-tracing simulations, as stated previously. Nevertheless, certain simplifications (for the problem geometry in particular) are usually introduced in order to prevent the simulation model to become excessively complex. A very common practice is to consider only the geometrical features which are significantly larger than a wavelength. In this way, the buildings in a large urban scenario for instance, can be depicted as simple polyhedrons (Rizk *et al.*, 2000). Further simplifications onto the building models can be introduced, such as the elimination of windows or indoor structures, without compromising the simulation accuracy significantly (Taygur and Eibert, 2021a). The geometrical model of the considered propagation environment (i.e., Munich downtown) is designed according to these principles as well, where the buildings are implemented as so-called 2.5D polyhedrons.

Even though the buildings are the most important aspect of geometrical modeling, a

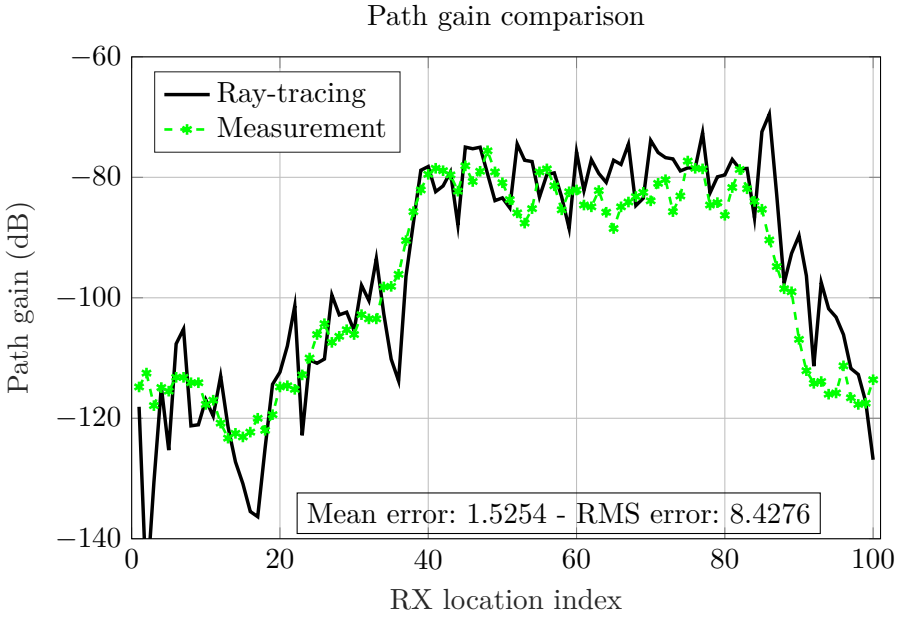


Figure 2.9: Comparison of ray-tracing simulation results to the measurement data for the Munich downtown urban scenario (Taygur and Eibert, 2021a). © 2021 IEEE.

typical urban scenario consists of many other entities, which might affect the propagation, such as foliage. In the considered urban scenario, the transmitter site is surrounded by tightly clustered trees (see Fig. 2.8), and the resulting attenuation, as the wave passes through this foliage block, is incorporated by the ITU-R early foliage model which can be given by (Parsons, 2000)

$$G_{\mathrm{dB}} = \frac{(f_{\mathrm{MHz}})^{0.3}(d)^{0.6}}{5},$$
$$G = 10^{\frac{G_{\mathrm{dB}}}{20}}, \tag{2.48}$$

where G indicates the additional attenuation factor which applies to electric and magnetic fields, f_{MHz} is the operating frequency in MHz, d is the distance that a ray propagates within the foliage block. The tree clusters are implemented as simple rectangular blocks which do not alter the ray trajectories in any way.

Finding a correct representation for the electrical properties (i.e., the permittivity) of the objects in the scenario is generally a more challenging problem, since it may not be possible to determine an exact value or to perform measurements in many cases. Nevertheless, the influence of an imprecise choice for the permittivity may not be as critical as the effect of large objects or foliage in the geometry (Taygur and Eibert, 2021a). Hence, it is usually sufficient to pick an acceptable value which reflects the electrical properties of the considered material reasonably well. Uniform material configurations may also be utilized to reduce the model complexity without causing much decline in accuracy, if the wave interactions mostly occur on the objects with similar electrical properties, i.e., buildings with concrete or brick walls and asphalt ground. Considering the general characteristics of the urban test scenario, a uniform material setting can

be applied here, thereby a relative permittivity of $\epsilon_r = 4$ (similar to commonly utilized materials in urban infrastructure such as concrete, brick, hardwood) is used for every object in the model (Zhekov *et al.*, 2020; Pinhasi *et al.*, 2008).

The antenna patterns for the receiver and transmitter were assumed to be isotropic. Since the measurements have been taken outdoors (i.e., no outdoor-to-indoor propagation), only reflections and diffractions on the vertical edges of the buildings were considered with up to a maximum limit of 7 and 1, respectively. The radii of the reception spheres as well as the diffraction cylinders were chosen as one wavelength, and the number of ray launches was set to 100 million. The simulation results are shown in Fig. 2.9.

The results demonstrate that an RMS error of 8.4 dB occurs between the measurement and the simulation. It can be observed that the deviations from the reference are more evident for the measurement locations outside of the index range of 40-80, where there is supposedly no dominant LOS link between the transmitter and receiver. Nevertheless, the small mean error of 1.5 dB indicates that channel analysis studies, which involve aggregate results over an entire network rather than individual links, can provide considerably accurate results.

Chapter 3

Bidirectional Ray-Tracing

Bidirectional ray-tracing based on reciprocity addresses certain important problems of the conventional unidirectional ray-tracing technique, in particular, the accuracy problems with multiple UTD, reception spheres, and the treatment of electrically very large scenarios, i.e., urban millimeter-wave propagation. In the following, the theoretical foundations of the method will be introduced first, then comparisons to the unidirectional ray-tracing approach and relevant multiple diffraction models will be given by numerical examples. Afterwards, the treatment of millimeter-wave scenarios via an asymptotic expansion of the reciprocity integration will be explained, and relevant numerical examples will be given.

3.1 Computation of Antenna Transfer Functions by the Reciprocity Theorem

Let us consider two antennas, A and B, which are excited with current densities $\mathbf{J}_A$ and $\mathbf{J}_B$, respectively. The radiated fields should then satisfy

$$\nabla \times \mathbf{E}_A = -j\omega\mu\mathbf{H}_A, \tag{3.1a}$$

$$\nabla \times \mathbf{H}_A = \mathbf{J}_A + j\omega\epsilon\mathbf{E}_A, \tag{3.1b}$$

$$\nabla \times \mathbf{E}_B = -j\omega\mu\mathbf{H}_B, \tag{3.1c}$$

$$\nabla \times \mathbf{H}_B = \mathbf{J}_B + j\omega\epsilon\mathbf{E}_B \tag{3.1d}$$

according to Maxwell's equations. If the equations (3.1a)-(3.1d) are dot multiplied by $\mathbf{H}_B$, $\mathbf{E}_B$, $\mathbf{H}_A$, and $\mathbf{E}_A$ respectively, then

$$\mathbf{H}_B \cdot (\nabla \times \mathbf{E}_A) = -j\omega\mu\mathbf{H}_B \cdot \mathbf{H}_A, \tag{3.2a}$$

$$\mathbf{E}_B \cdot (\nabla \times \mathbf{H}_A) = \mathbf{E}_B \cdot \mathbf{J}_A + j\omega\epsilon\mathbf{E}_B \cdot \mathbf{E}_A, \tag{3.2b}$$

$$\mathbf{H}_A \cdot (\nabla \times \mathbf{E}_B) = -j\omega\mu\mathbf{H}_A \cdot \mathbf{H}_B, \tag{3.2c}$$

$$\mathbf{E}_A \cdot (\nabla \times \mathbf{H}_B) = \mathbf{E}_A \cdot \mathbf{J}_B + j\omega\epsilon\mathbf{E}_A \cdot \mathbf{E}_B, \tag{3.2d}$$

is obtained. Subtracting Eq. (3.2a) from (3.2d) and Eq. (3.2c) from (3.2b) yields

$$\mathbf{E}_A \cdot (\nabla \times \mathbf{H}_B) - \mathbf{H}_B \cdot (\nabla \times \mathbf{E}_A) = \mathbf{E}_A \cdot \mathbf{J}_B + j\omega\epsilon \mathbf{E}_A \cdot \mathbf{E}_B + j\omega\mu \mathbf{H}_B \cdot \mathbf{H}_A, \tag{3.3a}$$

$$\mathbf{E}_B \cdot (\nabla \times \mathbf{H}_A) - \mathbf{H}_A \cdot (\nabla \times \mathbf{E}_B) = \mathbf{E}_B \cdot \mathbf{J}_A + j\omega\epsilon \mathbf{E}_B \cdot \mathbf{E}_A + j\omega\mu \mathbf{H}_A \cdot \mathbf{H}_B. \tag{3.3b}$$

Using the vector identity

$$\mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}) = \nabla \cdot (\mathbf{A} \times \mathbf{B}), \tag{3.4}$$

Eq. (3.3a) and (3.3b) become

$$\nabla \cdot (\mathbf{H}_B \times \mathbf{E}_A) = \mathbf{E}_A \cdot \mathbf{J}_B + j\omega\epsilon \mathbf{E}_A \cdot \mathbf{E}_B + j\omega\mu \mathbf{H}_B \cdot \mathbf{H}_A, \tag{3.5a}$$

$$\nabla \cdot (\mathbf{H}_A \times \mathbf{E}_B) = \mathbf{E}_B \cdot \mathbf{J}_A + j\omega\epsilon \mathbf{E}_B \cdot \mathbf{E}_A + j\omega\mu \mathbf{H}_A \cdot \mathbf{H}_B. \tag{3.5b}$$

Finally, subtracting Eq. (3.5b) from (3.5a) results in

$$\nabla \cdot (\mathbf{H}_A \times \mathbf{E}_B - \mathbf{H}_B \times \mathbf{E}_A) = \mathbf{E}_B \cdot \mathbf{J}_A - \mathbf{E}_A \cdot \mathbf{J}_B. \tag{3.6}$$

which is known as the Lorentz reciprocity theorem in differential form. By taking a volume integral on both sides and then applying the divergence theorem on the left-hand side, Eq. (3.6) can be written as

$$\oiint_{\Psi'} [(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A)] \cdot d\mathbf{S} = \iiint_{\Omega'} [\mathbf{E}_B \cdot \mathbf{J}_A - \mathbf{E}_A \cdot \mathbf{J}_B] \, dV, \tag{3.7}$$

where Ω' indicates the volume, i.e., the problem domain, which includes A and B, and Ψ' is the surface boundary of Ω'. Eq. (3.7) is known as the reciprocity theorem in integral form (Balanis, 2012). Let us now assume that the volume Ω' is modified such that it only contains antenna A, and antennas A and B are designated as receiver and transmitter, respectively. Thus, Eq. (3.7) can be modified as

$$\oiint_{\Psi} [(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A)] \cdot d\mathbf{S} = \iiint_{\Omega} [\mathbf{E}_B \cdot \mathbf{J}_A] \, dV, \tag{3.8}$$

where Ω denotes the new restricted volume and Ψ is the corresponding surface boundary for Ω (see Fig. 3.1). The volume integral at the right-hand side can be reduced into the terminal region as the tangential component of the electric field is zero on the antenna A (assuming PEC). Hence, Eq. (3.8) can be written as (Neitz, 2020; Harrington, 2001)

$$\oiint_{\Psi} [(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A)] \cdot d\mathbf{S} = \int_{\mathcal{L}_A} \mathbf{E}_B \cdot d\mathbf{l} \iint_{\mathcal{A}_A} \mathbf{J}_A \cdot d\mathbf{a} = V^{\mathrm{oc}} I_A, \tag{3.9}$$

where V^{oc} is the induced open-circuit voltage at antenna A on receive, I_A is the port current of antenna A on transmit, $\mathcal{L}_A$ and $\mathcal{A}_A$ are the terminal path and terminal cross section, respectively. Thus, the transfer function can be obtained by evaluating the surface integral in Eq. (3.9), assuming that the generator voltage at antenna B and I_A are already known. The evaluation is carried out as follows. First, both receiver and transmitter antennas launch rays, which are then collected on the interaction surface. Once the valid ray paths between the antennas and the interaction surface have been determined, the field expressions are calculated at ray-surface intersections. Since the intersection points on the surface are distributed arbitrarily in general, an interpolation routine is applied in order to obtain the field expressions on specific points on the surface which are relevant for the numerical integration algorithm.

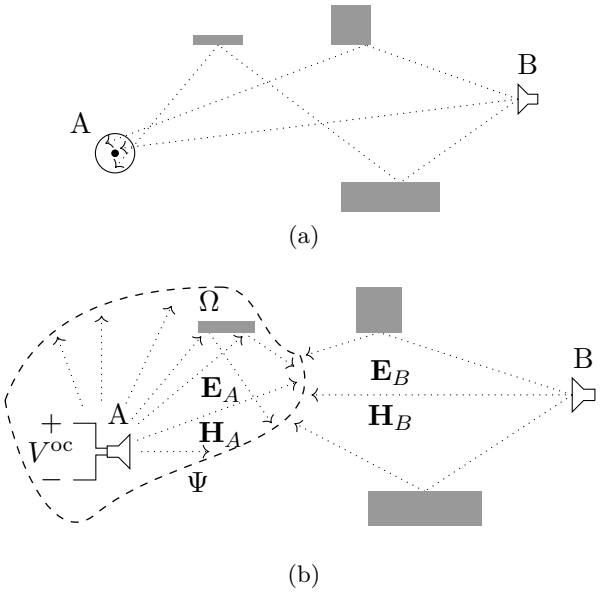


Figure 3.1: Illustrative comparison of unidirectional (a) and bidirectional ray-tracing (b) approaches (Taygur *et al.*, 2018). © 2018 IEEE.

3.1.1 Evaluation of Oscillatory Surface Integral

Since the surface integral in Eq. (3.9) involves oscillatory terms, which depend on the wavelength λ, its evaluation requires careful consideration. For such a problem, it may not be practical to utilize well-known conventional integration algorithms, such as quadrature or trapezoidal rule, since a large number of samples is usually needed. Therefore, an alternative technique, which can be used with a smaller number of samples, is preferred here. Let us re-express the integral in Eq. (3.9) by

$$\oiint_{\Psi} [(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A)] \cdot d\mathbf{S} = \oiint_{\Psi} \Theta(\mathbf{r}) dS \; = \oiint_{\Psi} \Gamma(\mathbf{r}) e^{jkg(\mathbf{r})} dS \tag{3.10}$$

with

$$\Theta(\mathbf{r}) = [(\mathbf{H}_A(\mathbf{r}) \times \mathbf{E}_B(\mathbf{r})) - (\mathbf{H}_B(\mathbf{r}) \times \mathbf{E}_A(\mathbf{r}))] \cdot \hat{\mathbf{n}},$$

$$\Gamma(\mathbf{r}) = |\Theta(\mathbf{r})|, \; g(\mathbf{r}) = \frac{\arg(\Theta(\mathbf{r}))}{k}, \tag{3.11}$$

where $\hat{\mathbf{n}}$ is the unit surface normal for Ψ, $\Gamma(\mathbf{r})$ and $g(\mathbf{r})$ are magnitude and phase functions, respectively[1]. The considered approach mainly involves the linearization of Γ and g and the evaluation of the resulting integral accordingly. The stationary phase points and critical points, which may not easily be treated by other techniques such as Levin or Filon (Levin, 1982; Gamkrelidze, 1989), can be handled without any difficulties. The accuracy of the integration process is heavily dependent on how good the linear approximation is, therefore, the integration domain should generally be subdivided into smaller sub-domains, in which the approximation can be considered as sufficiently

[1]The phase function is mostly determined by the ray path lengths, which can be represented as a smooth function.

good. For practical purposes, the subdomains can be assumed as triangles. The integral in Eq. (3.10) can then be approximated by

$$\oiint_{\Psi} \Gamma(\mathbf{r}) e^{jkg(\mathbf{r})} dS \approx \sum_{t=1}^{T} \iint_{\Psi_t} \tilde{\Gamma}_t(\mathbf{r}) e^{jk\tilde{g}_t(\mathbf{r})} dS = \sum_{t=1}^{T} \tilde{\varrho}_t(\mathbf{r}), \tag{3.12}$$

where T is the number of all triangles (i.e., number of subdomains), Ψ_t is the domain of the triangle t, $\tilde{\Gamma}_t$ and $\tilde{g}_t$ are the linear approximations for Γ and g in triangle t, respectively. Assuming that the coordinates of any point in a particular triangle are given by barycentric coordinates w_1, w_2, w_3, the individual integral terms $\tilde{\varrho}_t(\mathbf{r})$ can be expressed as

$$\tilde{\varrho}_t = 2\Delta_t \oiint_{\Psi_t} \tilde{\Gamma}_t(w_1, w_2, w_3) e^{jk\tilde{g}_t(w_1,w_2,w_3)} dS. \tag{3.13}$$

Defining A_{100}, A_{010}, and A_{001} as the values of $\tilde{\Gamma}_t$, and Ω_{100}, Ω_{010}, and Ω_{001} as the values of $\tilde{g}_t$ at the triangle corners, and introducing

$$J_1 = \int_0^1 w_1 \int_0^{1-w_1} e^{j(pw_1+qw_2)} dw_2 dw_1, \quad J_2 = \int_0^1 w_2 \int_0^{1-w_2} e^{j(pw_1+qw_2)} dw_1 dw_2,$$

$$J_3 = \int_0^1 \int_0^{1-w_1} e^{j(pw_1+qw_2)} dw_2 dw_1,$$

$$p = k(\Omega_{100} - \Omega_{001}),\ q = k(\Omega_{010} - \Omega_{001}), \tag{3.14}$$

$\tilde{\varrho}_t$ can be written as (Sukharevsky, 2015)

$$\tilde{\varrho}_t = 2\Delta_t e^{jk\Omega_{001}} \left[(A_{100} - A_{001})\, J_1 - (A_{010} - A_{001})\, J_2 - A_{001} J_3 \right]. \tag{3.15}$$

Even though the expansion presented in Eq. (3.15) yields an exact solution when Γ and g are linear, the nature of the reciprocity problem implies that g is almost exclusively non-linear since it is determined by the distances between the points on the triangle and the antennas. The accuracy is generally influenced by the relative size of the triangles with respect to the wavelength, i.e., the smaller the wavelength is, the smaller the triangles should be in order to achieve a certain accuracy (Taygur *et al.*, 2018). Thereby, the computational complexity grows with the frequency. Nevertheless, fast oscillations due to the complex exponential term do not have to be sampled as frequently as it is typically needed in the conventional methods (i.e., multiple samples within a single wavelength), thus, it is possible to achieve a certain accuracy by using a smaller number of samples.

3.1.2 Construction of Field Expressions

Since the previously described integration approach requires the knowledge about the magnitude and phase functions on certain points in small triangular domains, it is usually necessary to perform interpolation. This is because the field expressions on the surface can be obtained by ray hits, but the hits generally do not occur at the exact corner points. The computation of the field expressions is performed individually for each unique wavefront and can be summarized as follows. Let $\mathbf{r}_0$ be an arbitrary point

in the triangular subdomain t, on which the field expression should be found, $\mathbf{r}^r_{1,\dots,n}$ be n ray hit points on t, and $\hat{\mathbf{s}}_{1,\dots,n}$ be the direction rays which are incident on the interaction surface and yield intersections at $\mathbf{r}^r_{1,\dots,n}$. The computation steps can then be given as:

1. By using the properties of the incident wavefront (in particular, the radii of curvature ρ_1, ρ_2 and $\hat{\mathbf{s}}_{1,\dots,n}$), the ray path length as well as the direction of the incident ray at $\mathbf{r}_0$ is obtained.
2. The difference of the ray path length at $\mathbf{r}_0$ and any one of the hit points $\mathbf{r}^r_{1,\dots,n}$ is utilized to compute the variation of the spreading factor and phase.
3. The amplitude and polarization terms for electric field at $\mathbf{r}_0$ are obtained by means of a truncated Fourier series.

In the ideal case, n should be the same as the degrees of freedom in the interpolation algorithm. If n is smaller than the total number of ray hits on the triangle t, then it is possible to choose a particular set of hit points ($\mathbf{r}^r_{1,\dots,n}$) to improve the interpolation accuracy. In this study, $n = 5$ is assumed and $\mathbf{r}^r_{1,\dots,5}$ are chosen to be as close as possible to the (1/3, 1/3, 1/3), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1/2, 1/2, 0) in terms of the barycentric coordinates. It should be noted that the selection is empirical and can further be optimized.

Low Order Interpolation by Truncated Fourier Series

The interpolation by a truncated Fourier series is utilized for the treatment of eight scalar variables, two orthogonal polarization vectors (which consist of six scalars in total), and two amplitude terms of the electric field (associated with the two orthogonal polarization vectors). Let us assume that the quantity to interpolate is given by $\xi(\vartheta, \varphi)$, where ϑ and φ denote the ray propagation direction. $\xi(\vartheta, \varphi)$ can then be approximated according to

$$\xi(\vartheta, \varphi) \approx w_1 + w_2 \cos(\vartheta) + w_3 \sin(\vartheta) + w_4 \cos(\varphi) + w_5 \sin(\varphi), \tag{3.16}$$

where $w_{1,..,5}$ denote the interpolation weights. The weights can be obtained by solving the minimization problem

$$\min_{\mathbf{w}\in\mathbb{R}^5} ||\mathbf{A}\mathbf{w} - \mathbf{q}||^2, \tag{3.17}$$

with

$$\mathbf{A} = \begin{pmatrix} 1 & \cos(\vartheta_1) & \sin(\vartheta_1) & \cos(\varphi_1) & \sin(\varphi_1) \\ 1 & \cos(\vartheta_2) & \sin(\vartheta_2) & \cos(\varphi_2) & \sin(\varphi_2) \\ 1 & \cos(\vartheta_3) & \sin(\vartheta_3) & \cos(\varphi_3) & \sin(\varphi_3) \\ 1 & \cos(\vartheta_4) & \sin(\vartheta_4) & \cos(\varphi_4) & \sin(\varphi_4) \\ 1 & \cos(\vartheta_5) & \sin(\vartheta_5) & \cos(\varphi_5) & \sin(\varphi_5) \end{pmatrix}, \ \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{pmatrix}, \ \mathbf{q} = \begin{pmatrix} \xi(\vartheta_1, \varphi_1) \\ \xi(\vartheta_2, \varphi_2) \\ \xi(\vartheta_3, \varphi_3) \\ \xi(\vartheta_4, \varphi_4) \\ \xi(\vartheta_5, \varphi_5) \end{pmatrix}, \tag{3.18}$$

where ϑ_u, φ_u, $1 \leqslant u \leqslant 5$ correspond to the directions of the rays incident on the triangle which can be derived from $\hat{\mathbf{s}}_{1,\dots,5}$ and yield intersections at $\mathbf{r}^r_{1,\dots,5}$. The solution of the minimization problem is given by

$$\mathbf{w} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^+\mathbf{q}, \tag{3.19}$$

where $(.)^+$ is the pseudoinverse operator. It should be noted that the solution holds for $\text{rank}(\mathbf{A}) < 5$ as well (i.e., less than 5 ray hits on the triangle).

An implicit assumption made within the interpolation algorithm is that the same triangular subdivision of the interaction surface is utilized for both integration and interpolation processes. Similar to the case in the integration algorithm, smaller triangles yield better accuracy for the interpolation. Nevertheless, the size of the triangles utilized for the integration should usually be much smaller than those utilized for the interpolation process and be scaled with respect to the wavelength in order to achieve good accuracy. Therefore, two different mesh structures are utilized for interpolation and integration. For the integration routine, good results can be achieved with triangles having $\lambda^2/2$ area, i.e., 6 samples per λ^2 (Taygur *et al.*, 2018). The interpolation triangles should on the other hand be small enough, such that the variations of the incident direction and magnitude of the incident fields on the triangle are accurately represented with the proposed interpolation approach. The number of triangles should generally be larger if, for instance, the surface is in the vicinity of a highly directive antenna than that the surface is illuminated from a large distance by an antenna with smaller directivity. It was shown by Taygur *et al.* (2018) that a good accuracy can be achieved by utilizing (right isosceles) triangles as large as 0.125 m^2 in a scenario which involved a dipole antenna near the interaction surface at a 1 m distance.

3.2 Replacement of Reception Spheres with Interaction Surfaces

In order to assess the performance of the bidirectional ray-tracing against traditional unidirectional ray-tracing, a path gain analysis is carried out with both methods in two different scenarios where the interaction surfaces are mainly utilized as a substitute for the reception spheres. In both scenarios, an operating frequency of 2.45 GHz is assumed. A cubic interaction surface is located around the receiver in the bidirectional ray-tracing simulations and the surface was subdivided into 192 triangles (i.e., 32 triangles at each face) for the interpolation process. This mesh structure was further optimized during the numerical integration where the area of a triangle was approximately $\lambda^2/2$. Note that the size of the interaction surface is changed within the scenario in order to keep the number of ray hits on the surface relatively constant, as the distance between the receiver and transmitter is increased. The radius of the reception spheres in unidirectional ray-tracing simulations was 0.12 m, which corresponds to approximately one wavelength. The patterns of the transmitter and receiver antennas were assumed to be half-wavelength dipoles, which are oriented vertically, i.e., perpendicular to the ground plane.

3.2.1 Two-Ray Ground Reflection Scenario

The problem geometry consists of a receiver and transmitter, which are 1 m and 2 m above a flat ground plane made of PEC material, respectively. The distance between the receiver and transmitter was changed from 5 m to 83 m with 2 m steps. The cubic interaction surface, which was placed around the receiver, has a side length varying from 2 m to 20 m. The number of ray launches for the transmitter and receiver sites were set to 160×10^4 and 5×10^4, respectively. For the unidirectional ray-tracing, the number of ray launches was 165×10^4. Additionally, a reference unidirectional ray-tracing

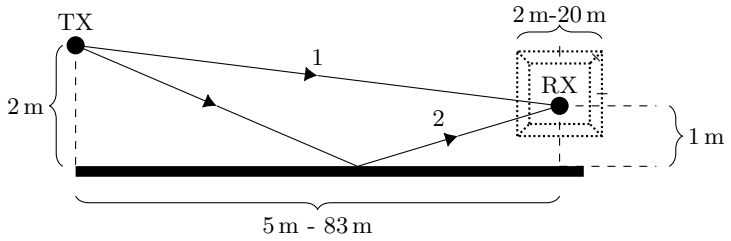


Figure 3.2: Two-ray ground reflection scenario illustration (Taygur *et al.*, 2018). © 2018 IEEE.

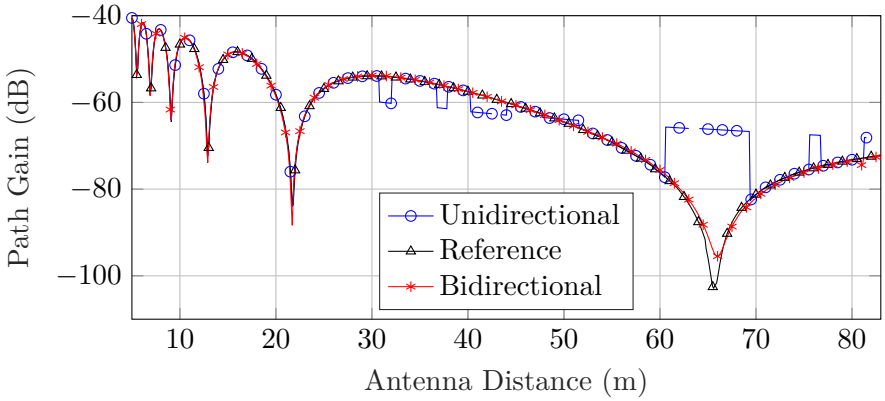


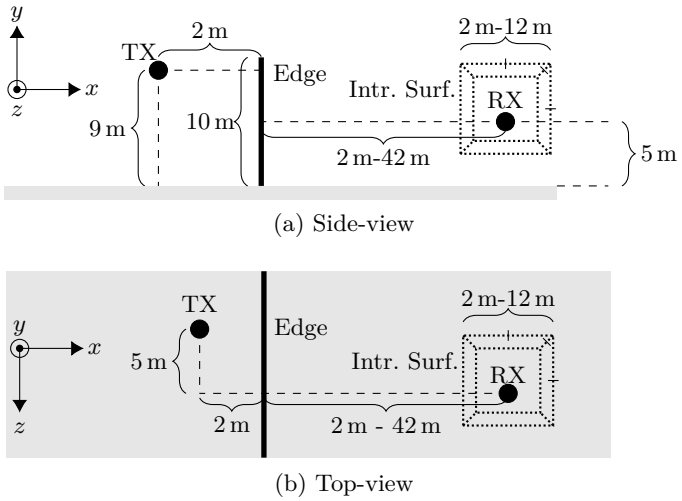
Figure 3.3: Two-ray ground reflection scenario simulation results (Taygur *et al.*, 2018). © 2018 IEEE.

simulation was performed with 600×10^4 ray launches. The illustration of the scenario and the numerical results are shown in Fig. 3.2 and Fig. 3.3, respectively.

The results indicate that the unidirectional ray-tracing is in good agreement with the reference until the distance between the antennas becomes larger than 30 m. Beyond this distance, deviations can clearly be seen as ray misses occur frequently. It was observed that the ray misses might occur with both reflection and LOS rays. The bidirectional ray-tracing results do not show significant differences and are in consistently good agreement with the reference. This outcome can be attributed to the fact that the interaction surface was enlarged to reduce the likelihood of ray misses as the antenna distance grew. It should be noted that the size of the reception sphere was kept fixed throughout the simulation whereas the interaction surface became larger. Although enlarging the reception sphere would result in no significant drawbacks in this particular scenario, such an adjustment may not always yield improvements in more complex scenarios where incorrect ray contributions might easily be captured (relevant example presented in Section 3.4.3).

3.2.2 Knife-Edge Diffraction Scenario

In this scenario, the transmitter and receiver antennas are separated by a half-plane located above a ground plane made of PEC material. Since there is no direct LOS link between the antennas, there are 4 feasible paths, all of which involve a diffraction (i.e., diffraction, reflection-diffraction, diffraction-reflection, reflection-diffraction-reflection).



(a) Side-view

(b) Top-view

Figure 3.4: Knife-edge diffraction scenario illustrations from side (a) and top (b) view perspective (Taygur *et al.*, 2018). © 2018 IEEE.

Similar to the previous scenario, the receiver antenna was progressively moved further away from the half-plane (2 m-42 m) and the cubic interaction surface was enlarged accordingly (side length change 2 m-12 m). The scenario illustration is shown in Fig. 3.4.

The diffraction contributions were simulated according to the UTD principles. Here, the number of ray launches may not directly dictate whether the reception sphere or interaction surface can be hit, rather, whether the diffraction interactions can be detected or not. In case a ray hits the diffraction cylinder lying along the edge (with a radius of $\lambda/2$), then the number of the new rays, which emerge on the Keller cone, determines the actual ray hits on the reception sphere or interaction surface. In the bidirectional ray-tracing simulation, the number of ray launches from the transmitter and receiver sites were 50×10^5 and 1×10^5, respectively. In order to establish fair conditions, the number of ray launches in the unidirectional ray-tracing simulation was set to 51×10^5. The diffracted ray density on the Keller cone was 100 rays/rad for both simulations. A reference simulation (using the unidirectional approach) was also carried out to compare the results, where the number of ray launches was kept the same (51×10^5), but the diffraction ray density was increased to 1000 rays/rad. The simulation results are given in Fig. 3.5.

The simulation results show similar features as in the previous scenario, i.e., the unidirectional ray-tracing results deviate from the reference as reception spheres may not be hit beyond 30 m distance whereas the bidirectional ray-tracing is in good agreement with the reference due to the flexible interaction surface. In certain exceptional cases, no ray hits occurred on the reception sphere, hence no prediction about the path gain could be made with the unidirectional approach.

3.3 Reciprocity Integration Over Open Surfaces

Following the definition of the reciprocity integral given in Eq. (3.9), the interaction surface was designated as a closed surface in the previously presented numerical examples. It is possible to work with an open surface in many scenarios though, as the integrand

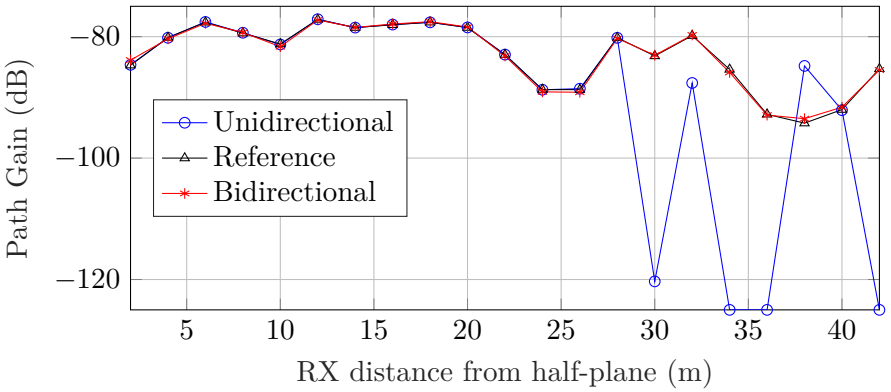


Figure 3.5: Knife-edge diffraction scenario simulation results. Note that the path gain drops abruptly (down to minus infinity in logarithmic scale) at certain points for unidirectional ray-tracing, since no ray hits occur (Taygur *et al.*, 2018). © 2018 IEEE.

is strongly oscillatory and its magnitude decays rapidly outside certain regions on the surface, thus, the result of the integral is mostly determined by the contributions within these regions. In particular, the values of Γ and g in the vicinity of stationary phase points (where $\nabla g = 0$) and critical points of second kind (generally occur at truncation boundaries) constitute the majority of the contributions to the result of the integral.

An open interaction surface can mainly be utilized in two different configurations. First, a sufficiently large open surface can be placed between the receiver and transmitter. Second, the surface can be placed at the boundary of a geometry object such as an edge or a wall. The latter configuration has considerable advantages for simulating diffraction effects, where the boundary of the interaction surface coincides with a diffraction edge, as it is possible to avoid UTD-based computations. Diffraction effects are thus incorporated by an approach which resembles to physical optics (PO) integration. Although a PO-like method is typically less accurate than UTD for a single diffraction, it might be practical to avoid the Keller cones and diffraction cylinders in various scenarios. More importantly though, for complex geometries, which may involve multiple wedge diffractions, the bidirectional ray-tracing with a large open interaction surface might be more efficient than unidirectional ray-tracing with multiple UTD calculations, and it can also provide considerable improvements in accuracy. This is because successive UTD operations usually yield an exponential growth in the number of rays and produce inaccurate numerical results as well, if diffractions occur near the optical boundaries (Holm, 1996; Schneider and Luebbers, 1991).

3.3.1 Open Interaction Surfaces in Single Knife-Edge Diffraction Scenarios

Characterization of Numerical Differences with UTD

Consider a knife-edge diffraction scenario, where the edge of the interaction surface coincides with the half-plane, as shown in Fig. 3.6. Under the assumptions of physical optics, it is possible to solve the reciprocity integral analytically in this problem, which can then be used to characterize the differences between the UTD and bidirectional ray-tracing approaches for single knife-edge scenarios. Note that a two dimensional problem

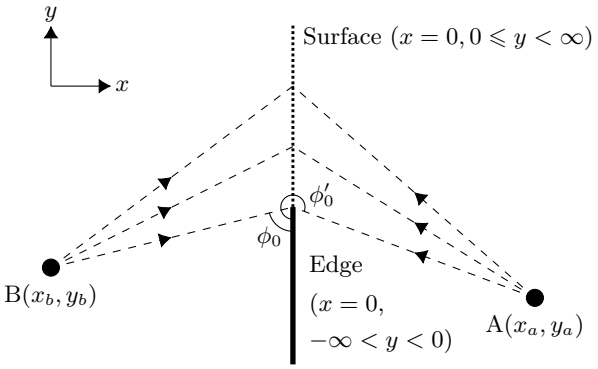


Figure 3.6: A single knife-edge diffraction scenario with an infinitely large, open interaction surface above the edge (Taygur *et al.*, 2018). © 2018 IEEE.

will be considered, thereby, the reciprocity surface integral is reduced into a line integral.

The interaction between two infinitely long line sources, A and B (which are on receive and transmit, respectively), can be computed by evaluating the oscillatory integral

$$\varrho = \int_0^{\zeta} \left[(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A) \right] \cdot d\mathbf{y} = \int_0^{\zeta} \Gamma(0,y) e^{jkg(0,y)} dy, \tag{3.20}$$

where $\zeta \to \infty$. The electric and magnetic field expressions on the interaction surface are given as

$$\begin{gathered} \mathbf{H}_A = e^{-jk\sqrt{x_a^2+(y-y_a)^2}} H_A(y)\hat{\mathbf{z}},\ \mathbf{H}_B = e^{-jk\sqrt{x_b^2+(y-y_b)^2}} H_B(y)\hat{\mathbf{z}}, \\ \mathbf{E}_A = Z_w \mathbf{H}_A \times \begin{bmatrix} \sin\phi_0' \\ \cos\phi_0' \\ 0 \end{bmatrix},\ \mathbf{E}_B = Z_w \mathbf{H}_B \times \begin{bmatrix} \sin\phi_0 \\ \cos\phi_0 \\ 0 \end{bmatrix}, \end{gathered} \tag{3.21}$$

with the definitions

$$\tan\phi_0' = -\frac{y_a}{x_a},\ \ \tan\phi_0 = \frac{y_b}{x_b}, \tag{3.22}$$

where $H_A(y)$ and $H_B(y)$ are the amplitudes of the magnetic fields of A and B on the surface, ϕ_0' and ϕ_0 indicate the angles between the incident rays and the edge plane, respectively. The magnitude and phase functions in Eq. (3.20), i.e., Γ and g, can then be written as

$$\begin{aligned} \Gamma &= -Z_w H_A(y) H_B(y) \left(\sin\phi_0' - \sin\phi_0 \right), \\ g &= \sqrt{x_a^2 + (y-y_a)^2} + \sqrt{x_b^2 + (y-y_b)^2} \end{aligned} \tag{3.23}$$

according to the definitions Eq. (3.11). Since $\nabla g \neq 0$ is valid in the entire integration domain, there are no stationary phase points. Here, an asymptotic expansion (with respect to k) is obtained by applying integration by parts, which is given as

$$\int_b^a u(y)v'(y)dy = u(a)v(a) - u(b)v(b) - \int_b^a u'(y)v(y)dy. \tag{3.24}$$

Rewriting Eq. (3.20) as

$$\int_0^{\zeta} \Gamma(0,y) e^{jkg(0,y)} dy = \int_0^{\zeta} \frac{\Gamma(0,y)}{jk\dfrac{\partial g(0,y)}{\partial y}} \frac{\partial}{\partial y} e^{jkg(0,y)} dy$$

$$= \frac{1}{jk} \int_0^{\zeta} \frac{\Gamma(0,y)}{g_y(0,y)} \frac{\partial}{\partial y} e^{jkg(0,y)} dy, \tag{3.25}$$

and applying integration by parts results in

$$u(y) = \frac{\Gamma(0,y)}{jkg_y(0,y)}, \; \frac{dv(y)}{dy} dy = \frac{\partial}{\partial y} e^{jkg(0,y)} dy,$$

$$\varrho = \frac{1}{jk} \left[\frac{\Gamma(0,\zeta)}{g_y(0,\zeta)} e^{jkg(0,\zeta)} - \frac{\Gamma(0,0)}{g_y(0,0)} e^{jkg(0,0)} \right] + \int_0^{\zeta} \frac{\partial}{\partial y} \left[\frac{\Gamma(0,0)}{jkg_y(0,y)} \right] e^{jkg(0,y)} dy. \tag{3.26}$$

The integral term at right-hand side in Eq. (3.26) can further be expanded by applying integration by parts, which then yields another difference term with a coefficient of k^{-2}. Thus, the asymptotic expansion for Eq. (3.20) as $k \to \infty$ is found as (Olver, 2008)

$$\varrho = \int_0^{\zeta} \Gamma(0,y) e^{jkg(0,y)} dy = \frac{1}{jk} \left[\frac{\Gamma(0,\zeta)}{g_y(0,\zeta)} e^{jkg(0,\zeta)} - \frac{\Gamma(0,0)}{g_y(0,0)} e^{jkg(0,0)} \right] + \mathcal{O}\left(k^{-2}\right) \tag{3.27}$$

with

$$g_y = \frac{\partial g}{\partial y}, \; \forall y \geqslant 0 : g_y(0,y) \neq 0. \tag{3.28}$$

Note that $\lim_{\zeta \to \infty} \Gamma(0,\zeta) = 0$ holds, therefore, Eq. (3.27) can be simplified into

$$\varrho \approx \frac{-1}{jk} \frac{\Gamma(0,0)}{g_y(0,0)} e^{jkg(0)} = \frac{jZ_w H_A(0) H_B(0) \left(\sin\phi_0' - \sin\phi_0\right)}{k\left(\cos\phi_0' + \cos\phi_0\right)} e^{jkg(0)}. \tag{3.29}$$

In order to establish a comparison between the UTD-based solution and Eq. (3.29), let us now consider the bidirectional ray-tracing scenario depicted in Fig. 3.7, for which an analytical path gain expression will be derived[2].

The oscillatory integral which describes the interaction between A and B can be written as

$$\varrho = \int_{-\infty}^{\infty} \left[(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A) \right] \cdot d\mathbf{y} = \int_{-\infty}^{\infty} \Gamma(x_s, y) e^{jkg(x_s,y)} dy. \tag{3.30}$$

Here, a stationary phase point occurs at $(x_s,\, y_s)$ since $\nabla g_{xy}(x_s, y_s) = 0$, and the sum of the ray path lengths is minimum, assuming a single wavefront from A and B are incident on the surface[3]. The stationary phase point can also be illustrated as the intersection point between the surface and the ray, which emerges at the diffraction point

[2]Based on the discussions on the use of open interaction surfaces, it can be inferred that the path gain results would be very similar, if bidirectional ray-tracing with a closed surface around A or unidirectional ray-tracing techniques were employed to simulate this particular scenario. However, bidirectional ray-tracing with an open surface between the edge and the antenna is preferred to compute an analytical path gain expression here, since the derivation is relatively straightforward in this way.

[3]Note that this relates to the Fermat principle of least time.

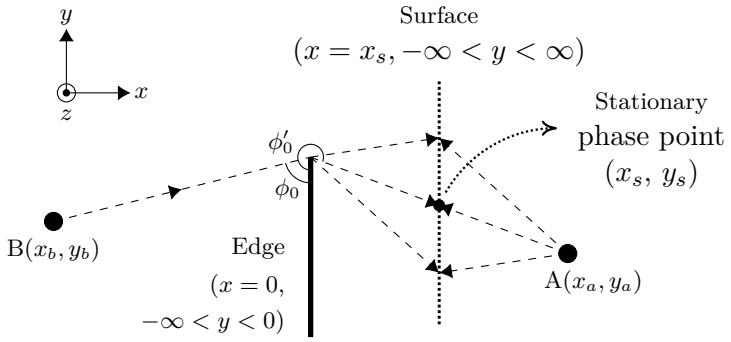


Figure 3.7: A single knife-edge diffraction scenario based on UTD and bidirectional ray-tracing with an infinitely large, open interaction surface.

$(x = 0, y = 0)$ and reaches A, if unidirectional ray-tracing was performed. An asymptotic expansion for the integral is found as

$$\varrho = \Gamma(x_s, y_s) e^{jkg(x_s,y_s)} \sqrt{\frac{2\pi}{k|g_{yy}(x_s, y_s)|}}\, e^{j\,\mathrm{sgn}(g_{yy}(x_s,y_s))\pi/4}. \tag{3.31}$$

Thus, an explicit result can be obtained once the field expressions are known. The electric and magnetic fields on the interaction surface can be given by

$$\begin{aligned}
\mathbf{H}_A &= H_{0,A}\, e^{-jk\sqrt{(x_s-x_a)^2+(y-y_a)^2}} \sqrt{\frac{R_1+R_2}{(x_s-x_a)^2+(y-y_a)^2}}\, \hat{\mathbf{z}}, \\
\mathbf{H}_B &= H_{0,B}\, e^{-jk\sqrt{x_s^2+y^2}} \frac{D_m}{\sqrt{x_s^2+y^2}}\, \hat{\mathbf{z}}, \\
\mathbf{E}_A &= Z_w \mathbf{H}_A \times \begin{bmatrix} -\sin\phi_0' \\ -\cos\phi_0' \\ 0 \end{bmatrix},\ \mathbf{E}_B = Z_w \mathbf{H}_B \times \begin{bmatrix} \sin\phi_0 \\ \cos\phi_0 \\ 0 \end{bmatrix},
\end{aligned} \tag{3.32}$$

with the definitions

$$\begin{aligned}
\tan\phi_0' &= -\frac{y_a}{x_a} = -\frac{y_s}{x_s}, \\
R_1 = \sqrt{x_s^2+y_s^2},\ R_2 &= \sqrt{(x_s-x_a)^2+(y_s-y_a)^2},
\end{aligned} \tag{3.33}$$

where $H_{0,A}$ and $H_{0,B}$ are the amplitudes of the magnetic fields of A and B on the tip of the diffraction edge at (0, 0), respectively[4]. Thus, Γ and g can be given by

$$\begin{aligned}
\Gamma(x_s, y) &= 2Z_w H_{0,A} H_{0,B} \frac{D_m(\phi_0', \phi_0)\sin\phi_0'}{\sqrt{x_s^2+y^2}} \sqrt{\frac{R_1+R_2}{(x_s-x_a)^2+(y-y_a)^2}}, \\
g(x_s, y) &= \sqrt{(x_s-x_a)^2+(y-y_a)^2} + \sqrt{x_s^2+y^2},
\end{aligned} \tag{3.34}$$

[4]Note that $H_{0,A}$, $H_{0,B}$ in Eq. (3.32) and $H_A(y=0)$, $H_B(y=0)$ in Eq. (3.21) are identical.

according to Eq. (3.11), where D_m denotes the diffraction coefficient. The integral result can then be written as

$$\varrho = 2Z_w H_{0,A} H_{0,B} D_m\left(\phi_0', \phi_0\right) \sin\phi_0' \sqrt{\frac{R_1+R_2}{R_1 R_2}} \sqrt{\frac{2\pi}{k\left|\frac{\sin^2\phi_0'}{R_1}+\frac{\sin^2\phi_0'}{R_2}\right|}}\, e^{-jk(R_1+R_2)+j\pi/4} \tag{3.35}$$

with

$$D_m = \frac{-4\cos\left(\frac{\phi_0'}{2}\right)\cos\left(\frac{\phi_0}{2}\right)}{\sqrt{8jk\pi}\left(\cos\phi_0'+\cos\phi_0\right)}, \tag{3.36}$$

for $|\phi_0' - \phi_0| \neq \pi$. By inserting the diffraction expression[5], the integral result can then be simplified into

$$\begin{aligned}
\varrho &= 2Z_w H_{0,A} H_{0,B} D_m\left(\phi_0', \phi_0\right) \sin\phi_0' \sqrt{\frac{R_1+R_2}{R_1 R_2}} \sqrt{\frac{R_1 R_2 2\pi}{k\sin^2\phi_0'\left(R_1+R_2\right)}}\, e^{-jk(R_1+R_2)+j\pi/4} \\
&= 2Z_w H_{0,A} H_{0,B} \left(\frac{-4\cos\left(\frac{\phi_0'}{2}\right)\cos\left(\frac{\phi_0}{2}\right)}{\sqrt{8jk\pi}\left(\cos\phi_0'+\cos\phi_0\right)}\right) \sqrt{\frac{2\pi}{k}}\, e^{-jk(R_1+R_2)+j\pi/4} \\
&= \frac{-4Z_w H_{0,A} H_{0,B}\cos\left(\frac{\phi_0'}{2}\right)\cos\left(\frac{\phi_0}{2}\right)}{k\left(\cos\phi_0'+\cos\phi_0\right)} e^{-jk(R_1+R_2)}.
\end{aligned} \tag{3.37}$$

If the phase related terms in Eqs. (3.29) and (3.37) are omitted, and the common terms are accordingly eliminated, two trigonometric terms given by

$$K^{\text{interaction}} = \frac{\left(\sin\phi_0' - \sin\phi_0\right)}{\left(\cos\phi_0'+\cos\phi_0\right)},\ K^{\text{diffraction}} = \frac{-4\cos\left(\frac{\phi_0'}{2}\right)\cos\left(\frac{\phi_0}{2}\right)}{\cos\phi_0'+\cos\phi_0}, \tag{3.38}$$

can be utilized to demonstrate the numerical differences between two simulation approaches. Thus, a correction factor for the bidirectional ray-tracing approach can be defined by

$$K^{\text{correction}} = \frac{K^{\text{diffraction}}}{K^{\text{interaction}}} = \frac{-4\cos\left(\frac{\phi_0'}{2}\right)\cos\left(\frac{\phi_0}{2}\right)}{\left(\sin\phi_0' - \sin\phi_0\right)}, \tag{3.39}$$

where a surface plot of the function for different ϕ_0' and ϕ_0 is shown in Fig. 3.8.

The surface plot shows that the difference between open surface integration and UTD/GTD methods is relatively large when the antennas are in the deep shadow ($\phi_0' \to 360$, $\phi_0 \to 0$). As the antennas are moved towards the transition region, the difference diminishes. It should be noted that neither $K^{\text{diffraction}}$ nor $K^{\text{interaction}}$ are valid at the optical boundary, therefore the above presented formulations should be used with caution, if the antennas are near the boundary (e.g., $\phi_0' \to 270$, $\phi_0 \to 90$).

[5]Since the given expression does not involve any correction terms for transition regions, it should be considered as a GTD based solution rather than UTD.

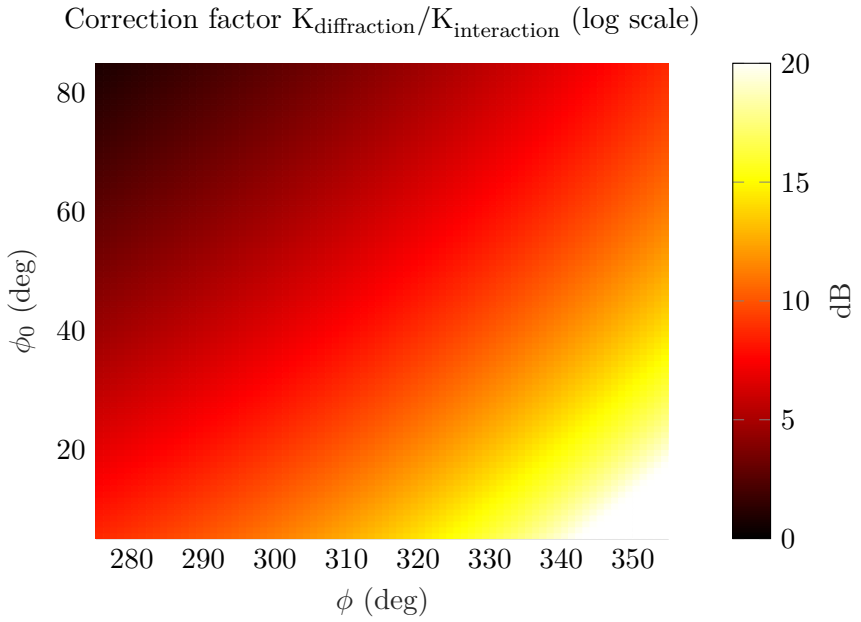


Figure 3.8: Variation of the correction factor $K^{\text{diffraction}}/K^{\text{interaction}}$ with respect to ϕ_0' and ϕ_0 (Taygur *et al.*, 2018). © 2018 IEEE.

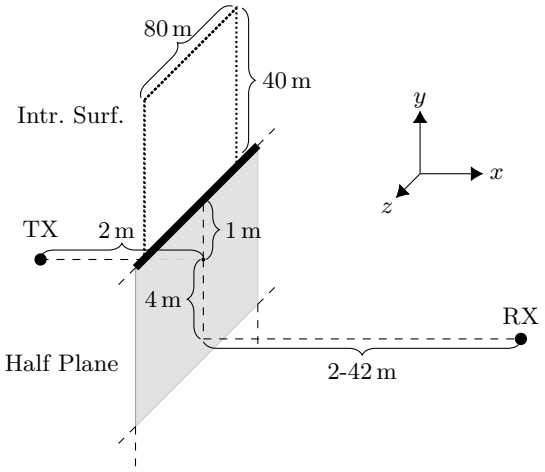


Figure 3.9: Single knife-edge scenario illustration (Taygur *et al.*, 2018). © 2018 IEEE.

Validating the Accuracy of the Correction Factor

Consider a single-knife edge scenario, as shown in Fig. 3.9, which is simulated by unidirectional and bidirectional ray-tracing methods. In order to verify that the correction factor given in Eq. (3.39) accurately represents the differences between integration on open surface and UTD/GTD formulation, the simulation result for the bidirectional ray tracing simulation is scaled accordingly and compared to the unidirectional ray-tracing results as well as the original simulation results (without scaling).

The scenario involves a fixed transmitter and a receiver which is then gradually moved away from the half-plane, where both antennas were assumed to be isotropic

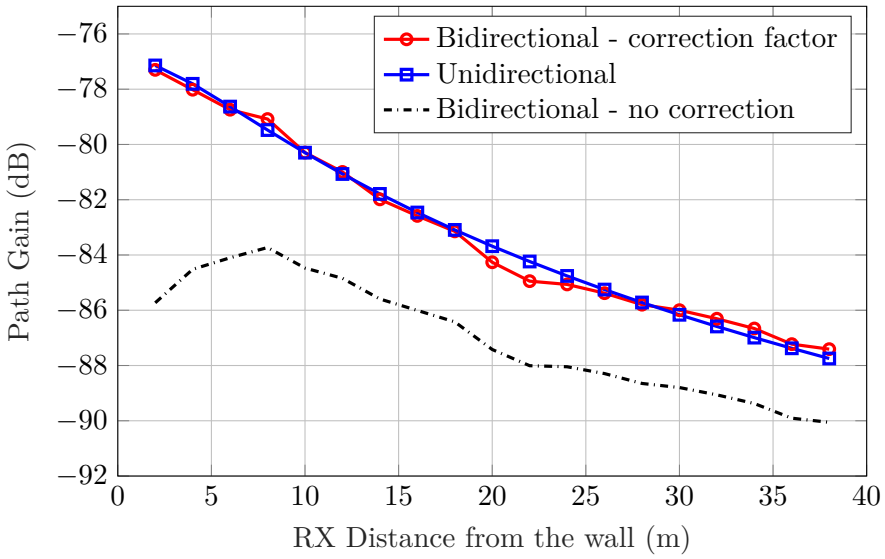


Figure 3.10: Validating the accuracy of correction factor in a single knife-edge diffraction scenario (Taygur *et al.*, 2018). © 2018 IEEE.

radiators, generating magnetically polarized fields, i.e., magnetic field is in z direction along the diffraction edge. Note that the positional change of the receiver represents a shift away from the deep shadow region, as the difference between the simulation results for unidirectional and bidirectional ray-tracing simulations would diminish. In the bidirectional ray-tracing simulation, an interaction surface with 4096 triangles and a size of 40 m × 80 m was utilized. The simulation results are given in Fig. 3.10.

The results demonstrate that the correction factor in Eq. (3.39) accurately represents the numerical differences between two simulation methodologies, as the disparity vanishes when the correction is applied to the bidirectional ray-tracing. Hence, it can be concluded that the correction factor accurately depicts the error of the bidirectional ray-tracing compared to the reference (i.e., the UTD/GTD solution) in the given knife-edge diffraction scenario. The difference between the unidirectional and unscaled bidirectional ray-tracing results is relatively larger when the receiver is close to the half-plane, as predicted.

The consistency between the unidirectional and the corrected bidirectional ray-tracing results implies that it is possible to express the integration result by an algebraic expression, i.e., an asymptotic expansion, under certain conditions. The asymptotic expansion yields accurate results for a sufficiently large integration domain where the contributions near the surface boundaries are relatively much smaller than the contributions near the stationary phase or critical points. The extent of the interaction surface might be limited though (i.e., due to certain problem constraints), thus the accuracy of the asymptotic approach declines. In such a case, a full-integration might be preferable as it is usually difficult to estimate whether the interaction surface is sufficiently large for asymptotic expansion to be valid.

Note that the derivation of the correction factor was based on a 2-dimensional problem whereas the considered scenario here is 3-dimensional. The scenario represents a special case though, since a direct incidence (i.e., the incident ray direction vectors and

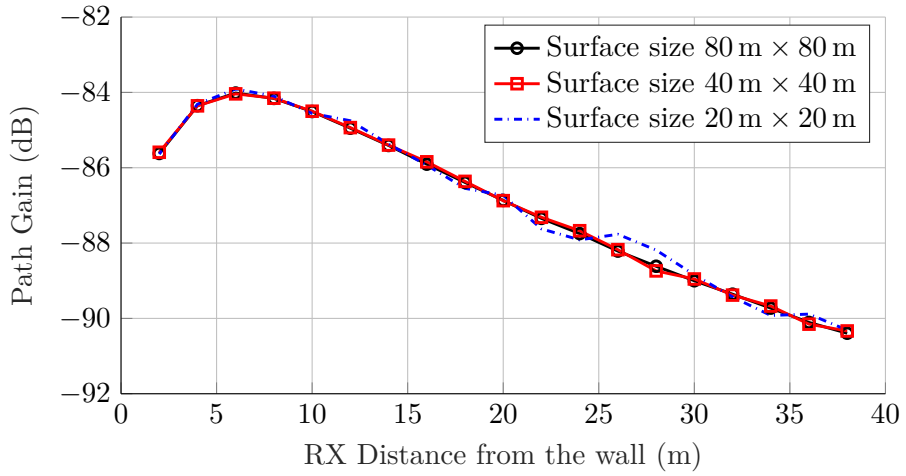


Figure 3.11: Characterization of the effects utilizing different interaction surfaces with different sizes.

the edge vector are perpendicular) occurs. The applicability of the correction factor in the case of an oblique incidence might be limited, since the diffraction coefficients as well as the asymptotic expansion take different forms than they were previously presented.

Investigating the Influence of the Surface Size

An important consideration regarding the selection of the interaction surface is its size. The analytical solution for the simplified knife-edge problem shown in Fig. 3.6 implies that an infinitely large surface is asymptotically optimal as the magnitude of the integrand reduces to zero beyond a certain region. However, in a realistic scenario, such as the one given in Fig. 3.9, the size of the interaction surface should indeed remain within practical limits, therefore, it is important to characterize how the variations in the physical size of the surface affect the simulation accuracy. The problem geometry in Fig. 3.9 is therefore accordingly modified, i.e., three different square interaction surfaces with edge lengths of 20 m, 40 m, and 80 m were utilized to compute the path gain.

The results shown in Fig. 3.11 indicate that the path gain results are very similar, especially for the surfaces with size 80 m × 80 m and 40 m × 40 m. The third surface, which is 20 m × 20 m large, yields slight deviations compared to the other two cases, hence, it can be inferred that further reductions in surface size may cause an even larger disparity and deterioration in the accuracy. Nevertheless, the results are in accordance with the expectation that the majority of the contributions are coming from the critical point at the boundary between the half-plane and interaction surface, rather than the distant parts of the surface (i.e., uppermost boundary of the interaction surface).

3.3.2 Open Interaction Surfaces in Multiple Knife-Edge Diffraction Scenarios

A notable drawback of the UTD-based diffraction computations are the imminent accuracy problems when multiple diffractions occur near the optical boundaries, as stated previously. In such scenarios, the bidirectional ray-tracing method can be useful to tackle

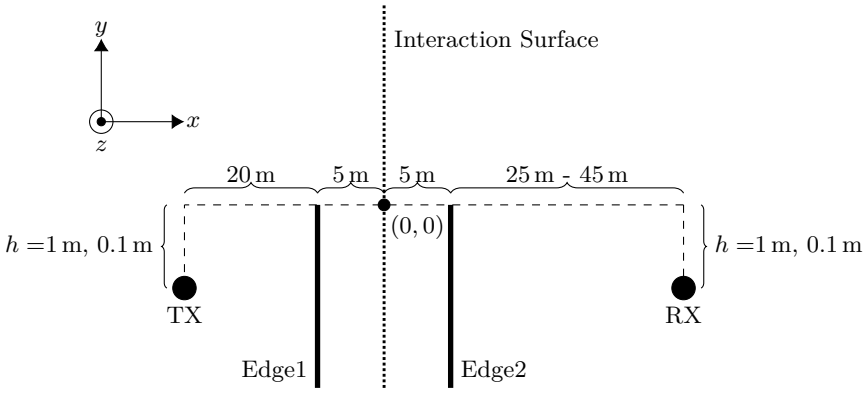


Figure 3.12: Double knife edge diffraction scenario illustration (Taygur *et al.*, 2018). © 2018 IEEE

the accuracy problems where a large and open interaction surface is used in a similar fashion as in the single knife-edge diffraction scenario, i.e., the surface is extended above the edges and placed between the receiver and transmitter antennas. The approach to validate the simulation accuracy for UTD and open surface integration is an important consideration here, thereby, various analytical techniques, such as Vogler or Deygout method, can be utilized as reference (Vogler, 1981; Deygout, 1966). Compared to the alternatives though, the Vogler method is typically considered to be a more accurate approach, as reported by various studies in the literature (Tzaras and Saunders, 2000; Bibb *et al.*, 2014).

Double Knife-Edge Diffraction Near Optical Boundary

The path gain in a double knife-edge diffraction scenario, which is illustrated in Fig. 3.12, is investigated with four different approaches:

1. Vogler method (reference).
2. Unidirectional ray-tracing with two consecutive UTD diffractions.
3. Bidirectional ray-tracing with single UTD on each edge where diffracted wavefronts are incident on the interaction surface in between the edges.
4. Bidirectional ray-tracing with no UTD where transmitter and receiver rays graze over the respective edges and hit the interaction surface directly.

The Vogler method was originally developed to approximate the attenuation of radio signals as they travel over irregular terrains. Thereby, a multiple knife-edge diffraction geometry was considered in order to model such geometries and the attenuation over this terrain profile was given in terms of a series of integrals (Fresnel integrals). Since an analytical solution is usually not available, the numerical solution for each integral is obtained by a power series expansion (Vogler, 1981).

The receiver antenna was progressively shifted away from the edge along the x-axis with 2 m steps. Two different cases were considered regarding the distance between the antennas and the optical boundaries, i.e., $h = 1\,\text{m}$ and $h = 0.1\,\text{m}$. Note that the receiving antenna approaches the optical boundaries as it moves away from the edge. Evidently, the antenna is closest to the optical boundary at $h = 0.1\,\text{m}$, $x = 50\,\text{m}$, and farthest at $h = 1\,\text{m}$, $x = 30\,\text{m}$.

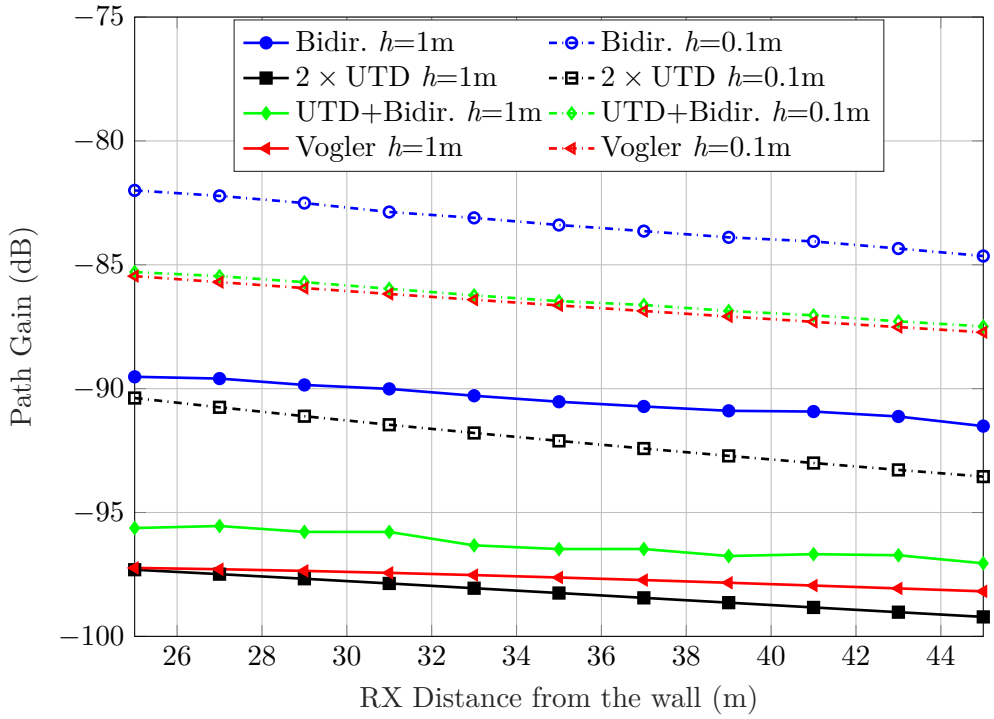


Figure 3.13: Double knife-edge simulation comparison (Taygur *et al.*, 2018). © 2018 IEEE.

The path gain characteristics presented in Fig. 3.13 reveal that the unidirectional ray-tracing simulation with consecutive UTD computations yields notable errors, especially when the antennas are near the optical boundary. The error diminishes as the receiver gradually moves away from the boundary and approaches the shadow region, i.e., $h = 1\,\text{m}$ and $x = 30\,\text{m}$. The bidirectional ray-tracing simulations on the other hand yield better results near the optical boundary, especially for the combination with single UTD where two singly diffracted wavefronts from both sites hit the interaction surface. The same approach (with single UTD from both sides) however does not yield a similar accuracy when the antennas are slightly shifted into the shadow region $h = 1\,\text{m}$. Here, it should be noted that the bidirectional ray-tracing with single UTD from both sides involves also direct ray contributions on the surface, which reduces the error of multiple UTD when the antennas are near the optical boundary. As the antennas are gradually shifted into the shadow region, the multiple UTD approach becomes more accurate, thus, the additional contributions from the direct rays on the surface cause a deviation from the reference. On the other hand, the accuracy of bidirectional ray-tracing with direct rays on the interaction surface (i.e., without UTD) declines, when the receiver moves into the shadow region.

Multiple Knife-Edge Diffraction at Optical Boundary

A path gain analysis, similar to that in the previous scenario, is performed for a multiple knife-edge diffraction scenario where the number of knife-edges varies from 2 to 5 and the antennas are located on the optical boundary, i.e., no height difference between the

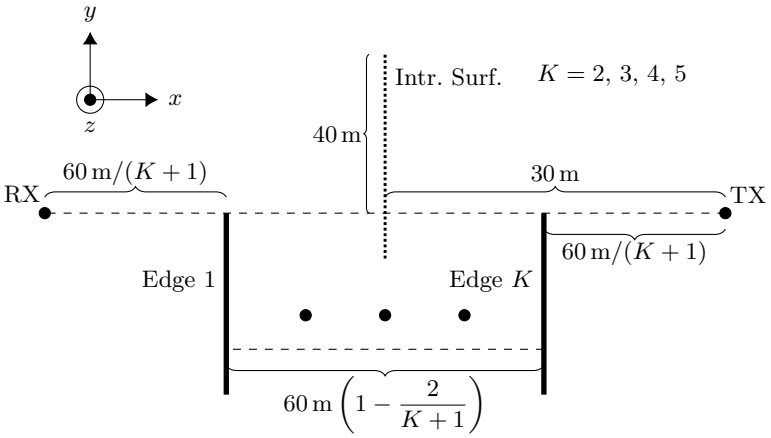


Figure 3.14: Multiple knife-edge diffraction scenario illustration (Taygur *et al.*, 2019a). © 2019 IEEE

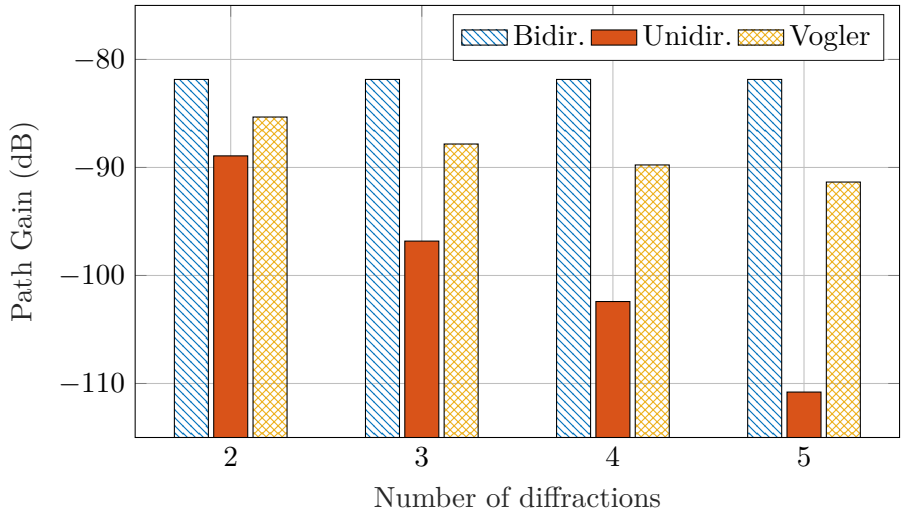


Figure 3.15: Comparison of bidirectional and unidirectional ray-tracing techniques with Vogler method for a multiple diffraction scenario at optical boundaries.

antenna and the tip of the edges. The distances between the objects (i.e., antennas and edges) are adjusted to be identical for each configuration and the separation between receiver and transmitter antennas was fixed to 60 m. The illustration of the scenario is shown in Fig. 3.14. The bidirectional ray-tracing approach, which is performed with only direct rays incident on the interaction surface and no UTD, is compared to the Vogler method. Additionally, a unidirectional ray-tracing simulation with multiple UTD computations is performed to demonstrate the accuracy issues of this approach. The results are given in Fig. 3.15.

The results show the accuracy issues of the UTD approach, especially when the number of consecutive diffractions grows. The bidirectional ray-tracing simulation results show a notable deviation from the reference (Vogler) as well, however, the difference does not grow as drastically as in the UTD case, therefore, it is advantageous to utilize the bidirectional approach.

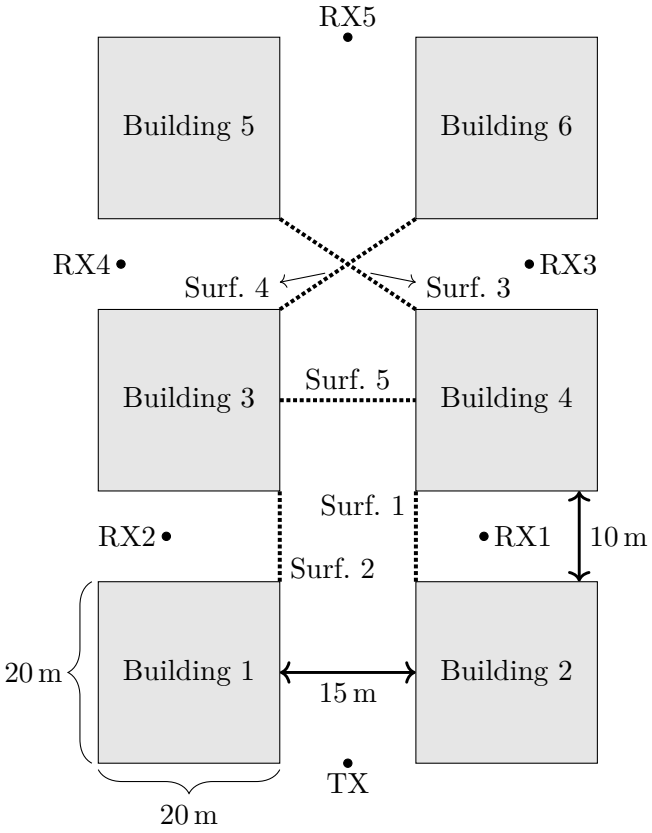


Figure 3.16: Urban building grid scenario illustration (Taygur *et al.*, 2018).

3.3.3 Use of Open Interaction Surfaces in an Urban-Like Environment

In order to compare the accuracy of bidirectional and unidirectional ray-tracing techniques in more complex scenarios, a 3×2 urban-like building grid is simulated with both methods. Five receivers and one transmitter were placed in the geometry 1 m above the ground level. Each receiver is associated with a different interaction surface[6], which are placed in a way that dominant propagation effects, such as diffractions on building corners, can properly be treated. The interaction surfaces extend up to 20 m height above the ground. It can be noticed that the third and fourth surfaces were placed diagonally, which is beneficial for capturing the rays from the transmitter more effectively (a similar configuration could be used for first and second surfaces as well). With both unidirectional and bidirectional ray-tracing methods, up to 6 reflections were taken into account. The diffractions in the unidirectional simulations were handled according to a heuristic UTD approach (Soni and Bhattacharya, 2010). The total number of ray launches was 450 million and 44 million for unidirectional and bidirectional ray-tracing, respectively. The illustration of the scenario and simulation results are shown in Fig. 3.16 and Fig. 3.17, respectively.

[6]Note that the simulation could theoretically be performed with a single interaction surface, which was placed, for instance, between building 1 and 2 near the transmitter. However, such a configuration does not yield any benefits in terms of simulation complexity, since the field contributions from different transmitter-receiver pairs are considered separately.

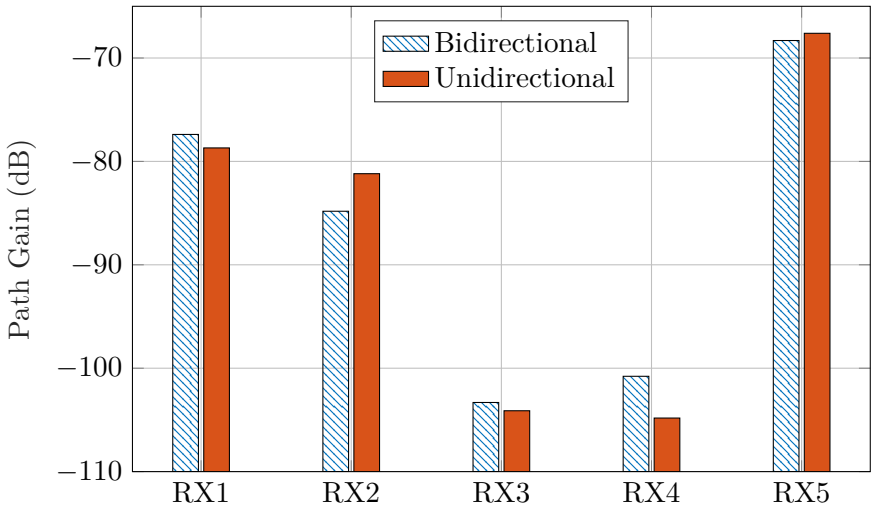


Figure 3.17: Comparison of bidirectional and unidirectional ray-tracing simulations for a 2×3 urban grid (Taygur *et al.*, 2018).

The results show that there is a good agreement between unidirectional and bidirectional ray-tracing approaches for all receivers. The diffractions are relatively insignificant in this scenario, as reflections from the ground plane and building walls constitute most of the contributions. Hence, the differences in the evaluation of diffraction effects by UTD or reciprocity integration has little effect on the final result.

3.4 Identification of Exact Ray Paths

The parameter selection in SBR based simulations, in particular the selection of the size of reception spheres or diffraction cylinders as well as the number of ray launches, involves certain challenges, as discussed before. In general, the size of these structures should be small enough to prevent incorrect rays to be captured whereas the number of ray launches should be accordingly large. Nevertheless, such a condition may not be adequate to achieve a good accuracy in millimeter-wave scenarios, as minor errors in the computation of ray path length may yield significant phase errors due to small wavelength. Here, it can be inferred that the size of the reception spheres and diffraction cylinders can be selected proportional to the wavelength in order to avoid inaccurate ray path length calculations and corresponding phase errors. However, this arrangement would generally yield an increase in simulation complexity as the frequency grows, since the number of ray launches should then accordingly be increased in order to compensate for the diminishing sizes of these structures.

The phase error problem can be resolved by means of bidirectional ray-tracing as the properties of exact rays (i.e., the rays which start and end exactly at the transmitter and receiver locations, respectively) are calculated on the interaction surface. The transfer function is obtained by applying a stationary phase point approximation within the reciprocity integral. It is possible to reduce phase errors as well as incorrect ray contributions with this approach, and also the computation time is typically lower compared to the conventional bidirectional ray-tracing since the integration is performed in a much more efficient manner (Taygur and Eibert, 2020b).

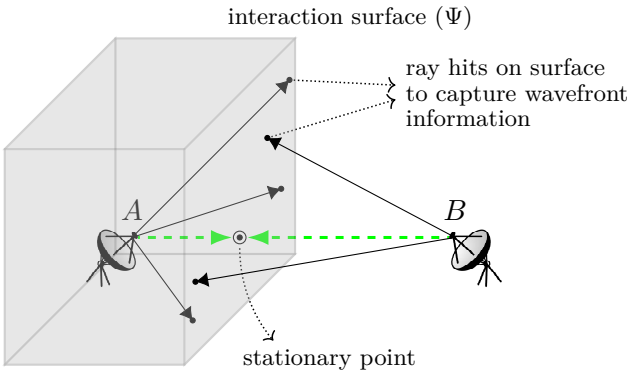


Figure 3.18: Identification of the stationary point on the interaction surface by using the bidirectional ray-tracing technique (Taygur and Eibert, 2020b). © IEEE 2020

3.4.1 Identifying Exact Ray Paths

The exact ray paths satisfy the Fermat principle of least time which can be expressed as (Born and Wolf, 1999)

$$S^{\text{path}} = \int \eta(\mathbf{r}) ds, \tag{3.40}$$

where S^{path} denotes the ray path length, $\mathbf{r} \in \mathbb{R}^3$ is the coordinate vector and $\eta(\mathbf{r})$ is the refractive index of the medium in which the wave propagates. The length of a ray path is stationary with respect to variations of the path, therefore, identifying the exact ray paths between two points can be considered as a minimization problem where S^{path} represents the cost function. The solution of this optimization problem for a homogeneous medium thus yields ray trajectories which can be illustrated as straight lines.

Let us now consider a scenario, which consists of two antennas, namely A and B, located in a homogeneous medium (e.g., free space), and an interaction surface Ψ placed between A and B (see Fig. 3.18). Note that there could be only one ray, which connects A and B, in this scenario, and the ray intersects with Ψ, given that Ψ is sufficiently large. If both antennas are operated in transmit mode and the propagating waves (note that there could be only a single unique wavefront which may occur from both antennas, which effectively yields a single wavefront pair) from both sites are incident upon Ψ, the intersection point on Ψ can be determined by means of certain properties of the incident waves on the surface, such as the radii of curvature.

In order to collect the wave-related information on Ψ, a bidirectional ray-tracing approach can be utilized, where the antennas launch rays towards this interaction surface. Consider two rays, which originate from A and B, respectively, and hit the same point on Ψ. The point, where the sum of the path lengths of these two rays is smallest, is the intersection point of Ψ and the exact ray, which connects A and B, and the sum value corresponds to the path length of this exact ray. In case multiple wavefronts occur in the problem, the minimization should be performed individually for each wavefront pair.

The intersection point (i.e., the minimum), which is obtained by the above mentioned minimization process, may be called the stationary phase point. The directions of the two rays, which come from two antennas and meet at the stationary phase point,

are typically opposite, such that the inner product of the normalized direction vectors equals -1. If the interaction surface is bounded by the objects in the problem geometry, the minima may occur at the boundary of the surface where the inner product might be different than -1 (for example, a diffraction scenario where the antennas are located in the shadow regions and the interaction surface is placed directly above the edge). The minimization of the path length can then be considered as a constrained optimization problem and the minima itself is called a critical point of second kind, rather than a stationary phase point.

The minimization of the ray path length can effectively be carried out by a line search algorithm. Note that the solution domain is constrained, i.e., the minima should remain within the boundaries of Ψ. The path lengths of the rays, which hit the interaction surface from both sites, are computed by using the information on the propagation direction and radii of curvature of the rays at the hit points. In case diffractions occur on the propagation path, the relevant information, such as the edge orientation and the directions of incident and outgoing waves on the edge, are also utilized. The pseudocode of the method is shown in Algorithm 3.1.

It should be noted that the exact ray paths may also be found by using an image method, or a hybrid approach, which combines the SBR and image method (Hussain, 2017). A hybrid approach may especially be useful in large, complex scenarios, where the approximate ray paths, which likely yield a ray hit, are obtained by SBR method first, then the image method is utilized to refine the approximate ray paths. Although such an approach is very effective to address the limitations of reception tests in conventional SBR technique, certain UTD-related issues should still be handled separately, e.g., by a reciprocity integration, as demonstrated previously.

3.4.2 Calculating the Antenna Transfer Function

The antenna transfer function is computed by applying the reciprocity theorem and evaluating the integral given in Eq. (3.9). Since the integral is highly oscillatory, computing the result by means of conventional techniques might be cumbersome if the wavelength is much smaller than the problem geometry, i.e., in millimeter-wave propagation scenarios. Nevertheless, asymptotic methods can be utilized to solve such problems efficiently.

Let us now re-consider the transfer function, which is given according to the reciprocity integral

$$\frac{V^{\text{oc}}}{V^{\text{gen}}} = \frac{1}{I_A} \oiint_{\Psi} [(\mathbf{H}_A \times \mathbf{E}_B) - (\mathbf{H}_B \times \mathbf{E}_A)] \cdot d\mathbf{S} = \frac{1}{I_A} \oiint_{\Psi} \Gamma(\mathbf{r}) e^{jkg(\mathbf{r})} dS, \tag{3.41}$$

with

$$\Theta(\mathbf{r}) = [(\mathbf{H}_A(\mathbf{r}) \times \mathbf{E}_B(\mathbf{r})) - (\mathbf{H}_B(\mathbf{r}) \times \mathbf{E}_A(\mathbf{r}))] \cdot \hat{\mathbf{n}},$$

$$\Gamma(\mathbf{r}) = |\Theta(\mathbf{r})|, \; g(\mathbf{r}) = \frac{\arg(\Theta(\mathbf{r}))}{k}, \tag{3.42}$$

where Γ and g are known as the magnitude and phase functions, respectively, as stated before. The result of this integral can be written in terms of Γ and g and their derivatives at a particular point, where $\nabla g = 0$ (Olver, 2008). The phase function g is strongly dependent on the ray path length in general, and can be expressed as

$$g = g^{\text{tx-path}}(\mathbf{r}) + g^{\text{rx-path}}(\mathbf{r}) + \frac{g_c(\mathbf{r})}{k}, \tag{3.43}$$

Algorithm 3.1 Ray path length computation and minimization (Taygur and Eibert, 2020b). © IEEE 2020.

procedure MINIMIZELENGTH
  $\mathbf{d}_1, \mathbf{d}_2 \leftarrow$ orthonormal vectors on the surface
  $\mathbf{c} \leftarrow$ surface center point
  repeat
    $l_0 \leftarrow \text{pathLength}(\mathbf{c})$
    $\alpha \leftarrow \text{pathLength}(\mathbf{c} + \epsilon\mathbf{d}_1) - l_0$ ▷ $0 < \epsilon \ll \lambda$
    repeat
      $\alpha \leftarrow k\alpha$ ▷ faster convergence often for $k > 1$
      $l_1 \leftarrow \text{pathLength}(\mathbf{c} - \alpha\mathbf{d}_1)$
    until $l_1 > l_0$ or max. iteration limit
    $\mathbf{c} \leftarrow \mathbf{c} - \mathbf{d}_1\alpha/k$
    $l_1 \leftarrow \text{pathLength}(\mathbf{c})$
    $\alpha \leftarrow \text{pathLength}(\mathbf{c} + \epsilon\mathbf{d}_2) - l_1$
    repeat
      $\alpha \leftarrow k\alpha$
      $l_2 \leftarrow \text{pathLength}(\mathbf{c} - \alpha\mathbf{d}_2)$
    until $l_2 > l_1$ or max. iteration limit
    $\mathbf{c} \leftarrow \mathbf{c} - \mathbf{d}_2\alpha/k$
    $l_2 \leftarrow \text{pathLength}(\mathbf{c})$
  until $l_2 \geq l_0$
  if $\mathbf{c}$ outside the surface **then**
    $\mathbf{c} \leftarrow \arg\min_{\partial\Psi} \|\partial\Psi - \mathbf{c}\|$ ▷ $\partial\Psi$: boundary of Ψ
  end if
  $l_2 \leftarrow \text{pathLength}(\mathbf{c})$
  return $\mathbf{c}, l_2$
end procedure

procedure PATHLENGTH($\mathbf{c}$)
  for both wavefronts from A and B **do**
  ▷ path length is sought at $\mathbf{c}$
  ▷ let r be a ray associated with a wavefront whose properties are known.
    $\rho_{1,\mathrm{r}},\ \rho_{2,\mathrm{r}}$ ▷ radii of curvature
    $\mathbf{P}_r,\ \mathbf{d}_r$ ▷ hit point and incident direction
    $\mathbf{S} \leftarrow \mathbf{P}_r - \rho_{2,\mathrm{r}}\mathbf{d}_r - (\rho_{1,\mathrm{r}} - \rho_{2,\mathrm{r}})\,\mathbf{T}_r\mathbf{d}_r$
  ▷ $\mathbf{T}_r$ represents rotation of the propagation direction vector at a diffraction edge (identity matrix, if no diffraction).
    if $\rho_{1,\mathrm{r}} = \rho_{2,\mathrm{r}}$ **then** ▷ no diffraction
      $l_{\{A,B\}} \leftarrow \|\mathbf{S} - \mathbf{c}\|$
    else
      $\mathbf{M} \leftarrow$ arbitrary point on edge
      $l_{\{A,B\}} \leftarrow \min_{\mathbf{M}} \|\mathbf{S} - \mathbf{M}\| + \|\mathbf{c} - \mathbf{M}\|$
    end if
  end for
  return $l_A + l_B$
end procedure

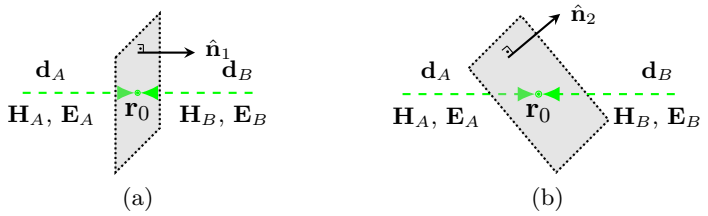


Figure 3.19: Illustration of two distinct surface configurations where the rays and the surface normal have a direct (a) and oblique (b) incidence (Taygur and Eibert, 2020b). © IEEE 2020.

where $g^{\text{tx-path}}$ and $g^{\text{rx-path}}$ are the path lengths of the transmitter and receiver rays (assuming a bidirectional ray-tracing procedure was performed in prior), which hit the interaction surface, respectively, and g_c denotes an additional phase term which usually depends on antenna patterns and diffraction coefficients. Here, it can be noticed that the point, where $\nabla g = 0$, corresponds to the stationary phase point. The result of the integral in Eq. (3.42) for $k \to \infty$ can be then written as (Fedoryuk, 1971; Wong, 2001)

$$\varrho \sim \frac{2\pi\Gamma(\mathbf{r}_0)}{k\sqrt{\det\left(\text{Hess}\left(g(\mathbf{r}_0)\right)\right)}} e^{jkg(\mathbf{r}_0)+j\pi/4} + \mathcal{O}(k^{-1}), \tag{3.44}$$

where $\mathbf{r}_0$ is the stationary phase point, Hess(.) and det(.) denote Hessian and determinant operators, respectively.

In order to evaluate the asymptotic expansion given in Eq. (3.44), the second order derivatives of g should be known (Hessian operation). However, traditional numerical differentiation algorithms, such as the finite difference method, are generally unstable, therefore, they will not be preferred for solving this problem (Conte and De Boor, 1980). The Hessian can be calculated indirectly, by utilizing the equivalence of two different geometrical configurations. Let us now consider two scenarios, as depicted in Fig. 3.19. The problem in general can be viewed as the calculation of the asymptotic expansion in two different scenarios where the incident rays reaching the surface make arbitrary angles with the surface normal, as in Fig. 3.19(b). On the other hand in Fig. 3.19(a), the incident ray direction vectors are aligned with the surface normal, i.e., $|\mathbf{d}_A \cdot \hat{\mathbf{n}}_1| = 1$, $|\mathbf{d}_B \cdot \hat{\mathbf{n}}_1| = 1$. Assuming that the propagation environment (which is not explicitly shown here) around the surface is identical in both cases, the transfer function as well as the asymptotic expansion should be identical, regardless of how the interaction surface is oriented with respect to the incident ray directions. Thus, the relation between the asymptotic expansion expressions can be given as

$$\begin{aligned} |\varrho| &= \frac{2\pi\Gamma_1(\mathbf{r}_0)}{k\sqrt{\det\left(\mathbf{q}_1(\mathbf{r}_0)\right)}} = \frac{2\pi\Gamma_2(\mathbf{r}_0)}{k\sqrt{\det\left(\mathbf{q}_2(\mathbf{r}_0)\right)}}, \\ &\to \frac{\Gamma_1(\mathbf{r}_0)}{\sqrt{\det\left(\mathbf{q}_1(\mathbf{r}_0)\right)}} = \frac{\Gamma_2(\mathbf{r}_0)}{\sqrt{\det\left(\mathbf{q}_2(\mathbf{r}_0)\right)}}, \\ &\to \sqrt{\det\left(\mathbf{q}_2(\mathbf{r}_0)\right)} = \sqrt{\det\left(\mathbf{q}_1(\mathbf{r}_0)\right)}\,\frac{\Gamma_2(\mathbf{r}_0)}{\Gamma_1(\mathbf{r}_0)}, \end{aligned} \tag{3.45}$$

where Γ_1, Γ_2 and $\mathbf{q}_1$, $\mathbf{q}_2$ denote the magnitude function and the Hessians for the first and second case, respectively. Note that the goal here is to find $\mathbf{q}_2$ (which involves a complicated Hessian expression) without evaluating any numerical derivatives. The computation of the magnitude functions Γ_1 and Γ_2 is relatively straightforward and can be given according to Eq. (3.42) as

$$\Gamma_{\{1,2\}} = \Big| \big[\left(\mathbf{H}_A(\mathbf{r}_0) \times \mathbf{E}_B(\mathbf{r}_0)\right) - \left(\mathbf{H}_B(\mathbf{r}_0) \times \mathbf{E}_A(\mathbf{r}_0)\right) \big] \cdot \hat{\mathbf{n}}_{\{1,2\}} \Big|. \tag{3.46}$$

Finally, $\mathbf{q}_1$ can be written as

$$\mathbf{q}_1(\mathbf{r}_0) = \begin{bmatrix} \dfrac{1}{\rho_{1,A}(\mathbf{r}_0)} + \dfrac{1}{\rho_{1,B}(\mathbf{r}_0)} & 0 \\ 0 & \dfrac{1}{\rho_{2,A}(\mathbf{r}_0)} + \dfrac{1}{\rho_{2,B}(\mathbf{r}_0)} \end{bmatrix}, \tag{3.47}$$

where $\rho_{1,A}$, $\rho_{2,A}$ and $\rho_{1,B}$, $\rho_{2,B}$ denote the radii of curvature of two incident wavefronts, coming from antennas A and B, respectively. The asymptotic expansion can then be given by

$$\varrho \sim \frac{2\pi\Gamma_1 e^{jkg(\mathbf{r}_0)+j\pi/4}}{k\sqrt{\left(\dfrac{1}{\rho_{1,A}(\mathbf{r}_0)} + \dfrac{1}{\rho_{1,B}(\mathbf{r}_0)}\right)\left(\dfrac{1}{\rho_{2,A}(\mathbf{r}_0)} + \dfrac{1}{\rho_{2,B}(\mathbf{r}_0)}\right)}}, \tag{3.48}$$

with

$$g(\mathbf{r}_0) = \arg\Big(\big[\left(\mathbf{H}_A(\mathbf{r}_0) \times \mathbf{E}_B(\mathbf{r}_0)\right) - \left(\mathbf{H}_B(\mathbf{r}_0) \times \mathbf{E}_A(\mathbf{r}_0)\right) \big] \cdot \hat{\mathbf{n}}_2 \Big)/k. \tag{3.49}$$

Thus, the asymptotic expansion can be calculated according to Eq. (3.48) and (3.49) for an arbitrarily oriented interaction surface whose normal is given by $\hat{\mathbf{n}}_2$.

Simulation of Diffraction Scenarios

Single wedge diffraction scenarios can be simulated with bidirectional ray-tracing where the interaction surface is placed above the edge and the transfer function is computed by evaluating the reciprocity integral over the entire surface, as shown before. The principles of asymptotic expansion may be applied for such scenarios as well, even though the exact ray path does not strictly yield a stationary phase point, but a critical point of second kind. The advantages of this approach would be twofold. First, Keller cones, which are used in unidirectional ray-tracing, can be avoided (as in the conventional bidirectional ray-tracing), second, computation of the transfer function can essentially be reduced to the evaluation of an algebraic expression. Let us now consider a single wedge diffraction scenario as depicted in Fig. 3.20(a). It can be noticed that the problem involves UTD diffraction, where the diffracted rays hit the interaction surface located at a certain distance from the wedge. Furthermore, the transfer function can be written according to the formulation given in Eq. (3.48) such that

$$\Theta = \left[\mathbf{H}_A \times \left(\mathbf{E}_B \cdot \bar{\mathbf{D}}\sqrt{\frac{\rho_t}{d(d+\rho_t)}}\right) - \left(\mathbf{H}_B \cdot \bar{\mathbf{D}}\sqrt{\frac{\rho_t}{d(d+\rho_t)}}\right) \times \mathbf{E}_A\right] \cdot \mathbf{d},\ g = \arg\left(\Theta\right)/k,$$

$$\varrho \sim \frac{2\pi\ \left|\Theta\right| e^{jkg+j\pi/4}}{k\sqrt{\left(\dfrac{1}{\rho_r} + \dfrac{1}{\rho_t + d}\right)\left(\dfrac{1}{\rho_r} + \dfrac{1}{d}\right)}}. \tag{3.50}$$

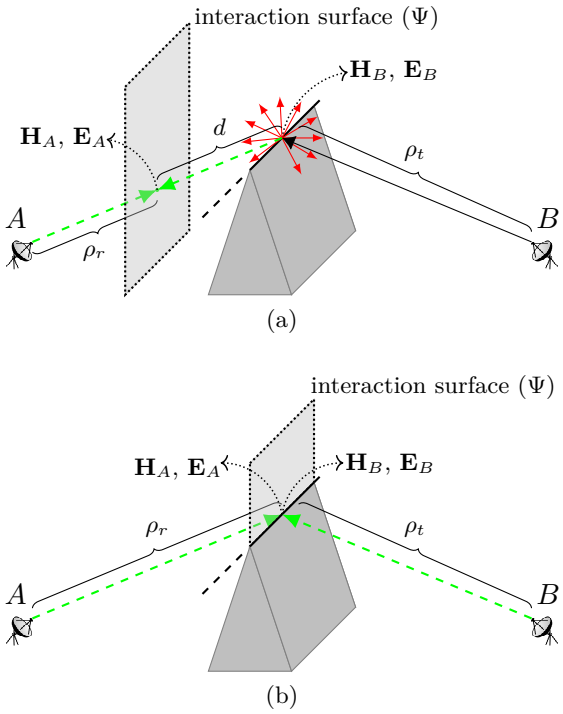


Figure 3.20: Illustration of two different interaction surface configurations for diffraction scenarios where the interaction surface is either between antenna A and the wedge (a) or above the edge between antennas A and B (b). Note that the first case involves Keller cones whereas the second case does not (Taygur and Eibert, 2020b). © IEEE 2020.

where $\bar{\mathbf{D}}$ denotes the dyadic diffraction coefficient based on the UTD. If the distance between the wedge and the interaction surface is gradually reduced such that the surface eventually coincides with the surface as shown in Fig. 3.20(b) (in other words, $d \to 0$), the expression ϱ given in Eq. (3.50) can be modified as

$$\varrho \sim \frac{2\pi\sqrt{\frac{1}{d}\left(\frac{\rho_t}{(d+\rho_t)}\right)}\left|\left[\mathbf{H}_A \times \left(\mathbf{E}_B \cdot \bar{\mathbf{D}}\right) - \left(\mathbf{H}_B \cdot \bar{\mathbf{D}}\right) \times \mathbf{E}_A\right] \cdot \mathbf{d}\right|}{k\sqrt{\frac{1}{d}\left(\frac{1}{\rho_r} + \frac{1}{\rho_t + d} + \frac{d}{\rho_r^2} + \frac{d}{\rho_r\left(\rho_t + d\right)}\right)}} e^{jkg + j\pi/4},$$

$$\lim_{d\to 0} \varrho \sim \frac{2\pi\left|\left[\mathbf{H}_A \times \left(\mathbf{E}_B \cdot \bar{\mathbf{D}}\right) - \left(\mathbf{H}_B \cdot \bar{\mathbf{D}}\right) \times \mathbf{E}_A\right] \cdot \mathbf{d}\right|}{k\sqrt{\left(\frac{1}{\rho_r} + \frac{1}{\rho_t}\right)}} e^{jkg + j\pi/4}. \tag{3.51}$$

Note that the resulting expression involves UTD coefficients and yields the same numerical result as if the problem was simulated with unidirectional ray-tracing, though the Keller cones are eliminated. The solution is thus obtained by identifying the hit point on the surface, which yields the smallest path length first (i.e., the sum of receiver and transmitter ray path lengths at a particular point), and then by applying Eq. (3.51) on this point.

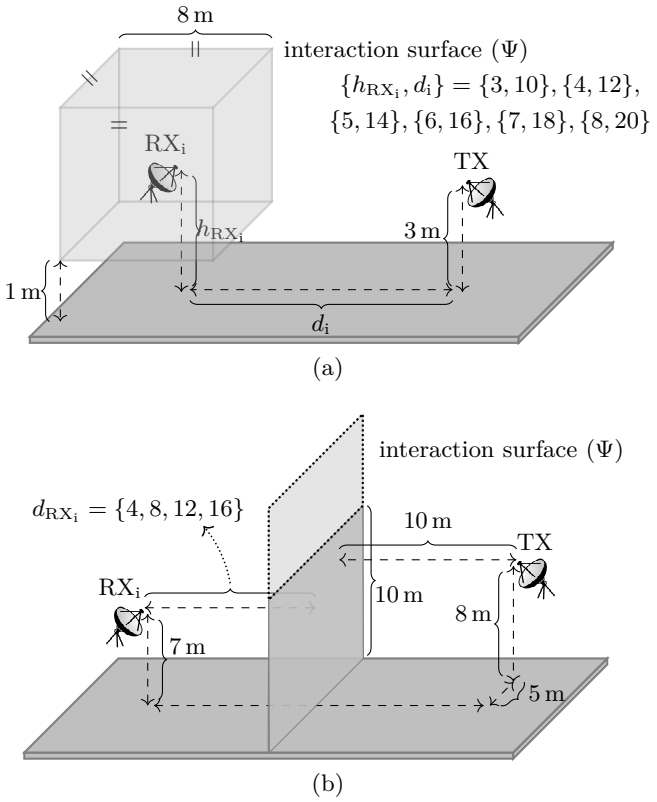


Figure 3.21: Illustrations of two-ray ground reflection (a) and single knife-edge diffraction (b) scenarios (Taygur and Eibert, 2020b). © IEEE 2020

3.4.3 Numerical Results

Two-Ray Ground Reflection & Single Knife-Edge Diffraction Scenarios

Two simple test scenarios, which are illustrated in Fig. 3.21, are simulated with the proposed technique as well as the conventional unidirectional ray-tracing, and the accuracy of both approaches are compared for 10 different frequency points between 2.5 and 25 GHz. The accuracy comparison is made according to a reference simulation which utilizes the unidirectional ray-tracing approach with a large number of rays and small reception spheres/diffraction cylinders. The comparison metric was chosen as the coupling between the receiver and the transmitter. The position of the receivers was changed in both scenarios, where 6 and 4 different position configurations were considered in ground reflection and knife-edge diffraction scenarios, respectively. The reference solution for the diffraction scenario involves also ground reflections where 4 unique ray paths occur in total (diffraction, diffraction-reflection, reflection-diffraction, reflection-diffraction-reflection). The environment in both scenarios was modeled with PEC material. The parameter selection has been carried out as follows: the smallest number of ray launches and diffracted rays, which yield less than 1 dB error in a unidirectional ray-tracing simulation for 2.5 GHz, is determined, then, the same number of ray launches (i.e., sum of receiver and transmitter ray launches) is utilized in the bidirectional case. The radius of the reception spheres and diffraction cylinders in the unidirectional ray-tracing were chosen as 0.36 m and 0.1 m (approximately corresponding to 3λ and 0.8λ

at 2.5 GHz), respectively. The results are shown in Fig. 3.22 and 3.23.

It can be observed that the unidirectional ray-tracing simulation yields up to 3.5 and 6 dB error for ground reflection and knife-edge diffraction scenarios, respectively. The causes of larger error in the diffraction scenario can be given as the presence of the diffraction cylinders as an additional source of phase error and the larger number of unique ray paths compared to the ground reflection scenario. Note that the error does not exhibit a regular pattern with respect to the change in frequency, though, the largest error typically occurs beyond 7.5 GHz. Thus, it can be inferred that the likelihood of phase errors increase with frequency, when the reception spheres and/or diffraction cylinders are kept constant in size. On the other hand, the error of the proposed method does not exceed 0.5 dB, which demonstrates its advantage in terms of accuracy.

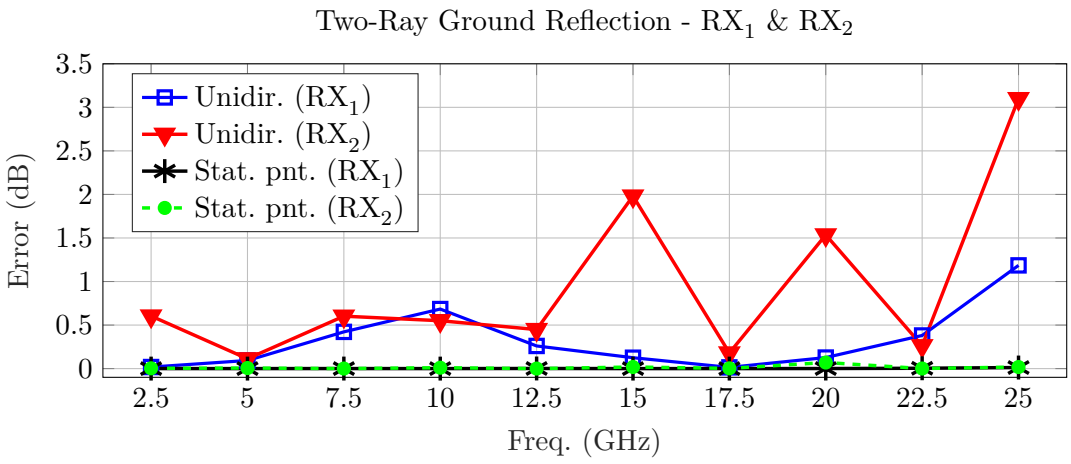
Urban Scenario

An urban environment model, which involves 5 receivers and 12 buildings in a 3×4 grid arrangement (shown in Fig. 3.24), is simulated with bidirectional ray-tracing based on exact ray path computation and unidirectional ray-tracing, where the accuracy of both methods (with respect to a reference simulation based on unidirectional ray-tracing with large number of ray launches and small reception spheres) is compared in a more realistic propagation environment. Three different simulation cases are considered:

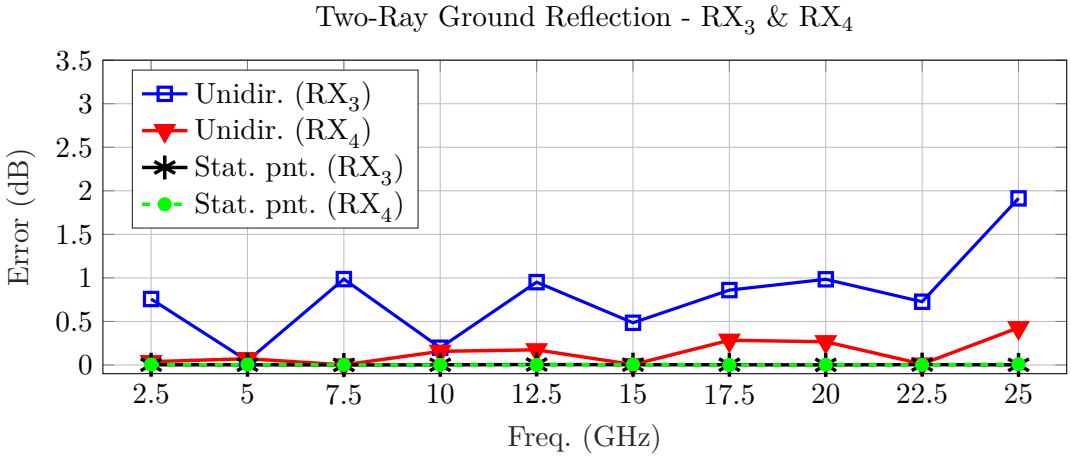
Case I: The number of ray launches, which is necessary to achieve a certain simulation accuracy, is evaluated for both methods. The accuracy criterion was defined as no more than 0.5 dB deviation in coupling for any receiver with respect to the reference simulation. The reception spheres in the unidirectional ray-tracing simulation were adjusted in a way that they are large enough to capture the correct ray contributions with as little ray launches as possible, but small enough to prevent phase errors and incorrect ray contributions (identified according to the reference simulation data.).

Case II: The accuracy in both approaches are compared when the same number of ray launches is utilized. The number of ray launches is chosen according to the configuration in Case I. The simulations were performed for both 2.5 and 25 GHz. Since the size of the reception spheres is adjusted to prevent incorrect ray contributions, it is expected that the variation of phase errors can clearly be observed between 2.5 and 25 GHz.

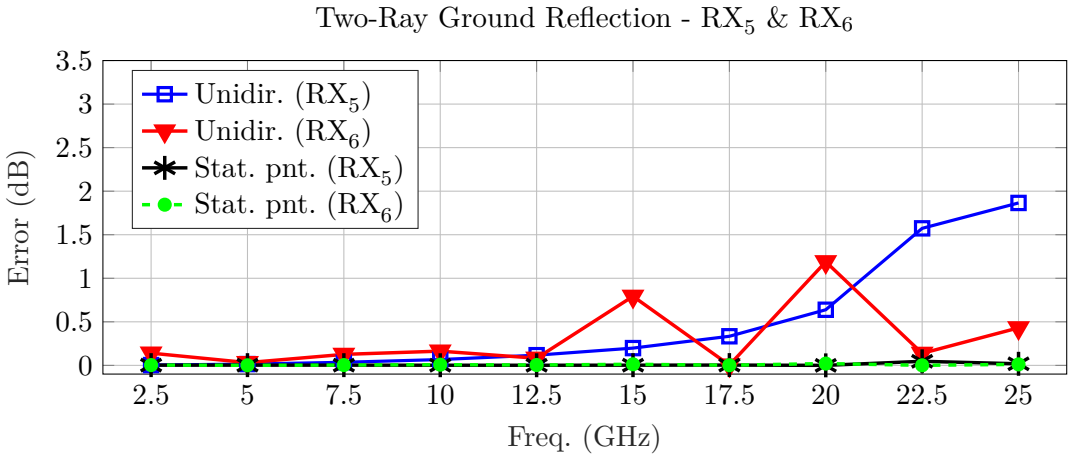
Case III: Assuming that reference data, which can be used to adjust the simulation parameters, is typically not available, a convergence analysis was carried out, i.e., multiple simulations were performed where the number of ray launches was increased progressively until the simulation output (i.e., path gain or coupling) converged to a certain value. In this case, the number of ray launches is increased by 1 million (starting at also 1 million) at each successive step until the difference in coupling between two consecutive runs remains below 0.5 dB. Three distinct scenarios are considered, a bidirectional ray-tracing simulation with the same parameters as in Case I and II, a unidirectional ray-tracing simulation with large reception spheres having a diameter of 4 m (the same as the side length of the cubic interaction surfaces utilized in the bidirectional ray-tracing approach), and a second unidirectional ray-tracing simulation with smaller reception spheres (radius of 0.48 m) where incorrect ray contributions are avoided. Note that incorrect rays may easily be captured with large reception spheres, thus, the error is likely to be larger compared to the latter case where the incorrect rays were avoided. All the simulations are performed at 25 GHz.



(a)

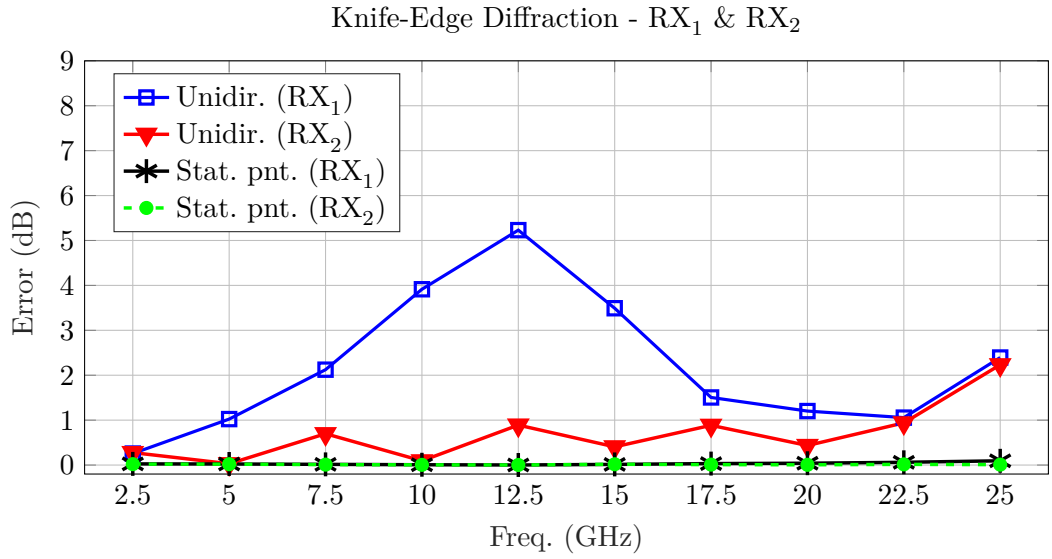


(b)

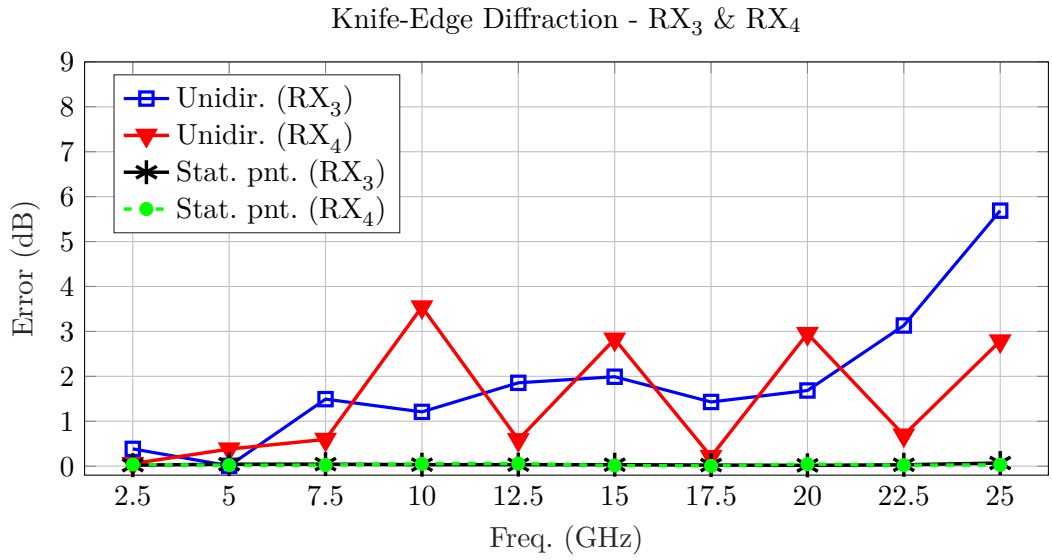


(c)

Figure 3.22: The error characteristics of unidirectional and bidirectional ray-tracing with asymptotic expansion with respect to frequency in a two-ray ground reflection scenario for receiver positions 1, 2, (a) 3, 4, (b) 5 and 6 (c) (Taygur and Eibert, 2020b). © IEEE 2020.



(a)



(b)

Figure 3.23: The error characteristics of unidirectional and bidirectional ray-tracing with asymptotic expansion with respect to frequency in a knife-edge diffraction scenario for receiver positions 1, 2, (a) 3 and 4 (b) (Taygur and Eibert, 2020b). © IEEE 2020.

The simulation parameters for all three cases as well as the reference simulation, and the numerical results are shown in Table 3.1 and Fig. 3.25, respectively.

The results for Case I indicate that the number of ray launches, which is required to achieve a certain accuracy, is larger for the unidirectional ray-tracing. It can be noticed that the discrepancy between two methods is particularly large for RX_5. Although the difference is generally smaller for other receivers, it can be noticed that the number of ray launches can easily be 2-3 times larger in unidirectional ray-tracing than that is needed in bidirectional ray-tracing.

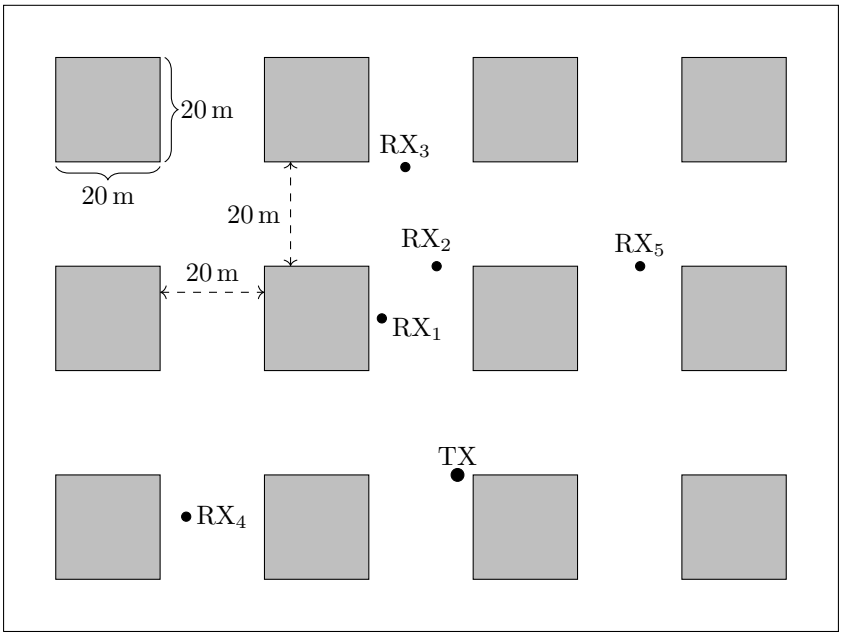
Table 3.1: Parameters utilized for the urban scenario simulations (Taygur and Eibert, 2020b). © IEEE 2020

Case	Simulation	f (GHz)	Total ray launches ($\times 10^6$)	Interaction surface or reception sphere size (m)
I	Bidirectional	25	1.05 / 1.05 / 1.05 / 2.05 / 2.05*	Cubic intr. surf. side len. 4 for all RX
	Unidirectional		3 / 4 / 3 / 5 / 14	Sph. radius 0.36 / 0.48 / 0.36 / 0.24 / 0.12
II	Bidirectional	25	1.05 / 1.05 / 1.05 / 2.05 / 2.05*	Cubic intr. surf. side len. 4 for all RX
	Unidirectional			Sph. radius 0.48 for all RX
	Bidirectional	2.5		Cubic intr. surf. side len. 4 for all RX
	Unidirectional			Sph. radius 0.48 for all RX
III	Bidirectional	25	1.05 / 1.05 / 1.05 / 2.05 / 2.05*	Cubic intr. surf. side len. 4 for all RX
	Unidirectional		4 / 3 / 4 / 1 / 4	Sph. radius 2 for all RX
	Unidirectional		1 / 1 / 1 / 3 / 3	Sph. radius 0.48 for all RX
Ref.	Unidirectional	2.5 & 25	120	Sph. radius 0.048 for all RX

The data in each cell denote the corresponding parameter for 5 receivers separated by slash marks.
The number of the correct multi-paths is found as 6 / 6 / 4 / 6 / 6 for $RX_{1,\ldots,5}$.
$*$: In bidirectional ray-tracing, 50,000 rays are launched from the receiver site.

In Case II, the error in unidirectional ray-tracing at 25 GHz is the most notable outcome, as a clear increase compared to the 2.5 GHz case can be observed. Since the only changing parameter is frequency here, it can be inferred that the difference stems from an excessive phase error at 25 GHz, where the reception sphere is much larger than the wavelength. Note that the error remains below 0.5 dB in the other simulations, including the bidirectional ray-tracing at 25 GHz, where no large phase error occurs.

Lastly, in Case III, the unidirectional ray-tracing simulation with large spheres yields an error up to 17 dB, which includes both incorrect ray contributions and phase error. The small sphere case does not yield an error of more than 2 dB (note that the reception spheres here are still very large compared to the wavelength at 25 GHz). Even though the number of ray launches was determined according to a well-defined convergence criterion in both cases, it can be noticed that an improper choice regarding the size of a reception sphere may introduce a significant error. On the other hand, the interaction surfaces do not yield an error larger than 0.5 dB in the bidirectional ray-tracing case.



Coordinates: TX:(0,0), RX_1:(-14.5,30), RX_2:(-4,40), RX_3:(-10,59), RX_4:(-52,-8), RX_5:(35,40)

Figure 3.24: Urban scenario illustration (Taygur and Eibert, 2020b). © IEEE 2020

Performance Analysis

The prominent advantage of the exact ray identification compared to the conventional bidirectional ray-tracing approach is that the reciprocity integral can be evaluated much quicker. Nevertheless, a non-asymptotic integration technique may still be needed during the simulation process, for instance, if a wavefront pair upon the surface does not yield a stationary phase point due to shadowing effects (Taygur and Eibert, 2020b).

In such a scenario, the simulation time scales with frequency in general. However, the total simulation time may still remain small compared to that is needed with the conventional integration approach, since the contribution of the wavefronts, which yield stationary phase points, are computed very fast, by only evaluating the asymptotic expansion. In order to demonstrate the advantages of the proposed approach, a comparison concerning computation time and accuracy was carried out by evaluating the transfer function for RX_1 from the previously shown urban scenario with both integration methods at 10 different frequency points between 2.5 and 25 GHz.

As it can be seen in Fig. 3.26, the integration time scales with frequency in both methods, hence, it can be concluded that not every wavefront pair on the surface yields a stationary phase point. However, the evaluation time can be 4 times larger with the conventional integration than that is with the exact-ray approach. On the other hand, the numerical results, in particular the received power and the phase of the induced voltage at the receiver terminals, are relatively similar with both approaches.

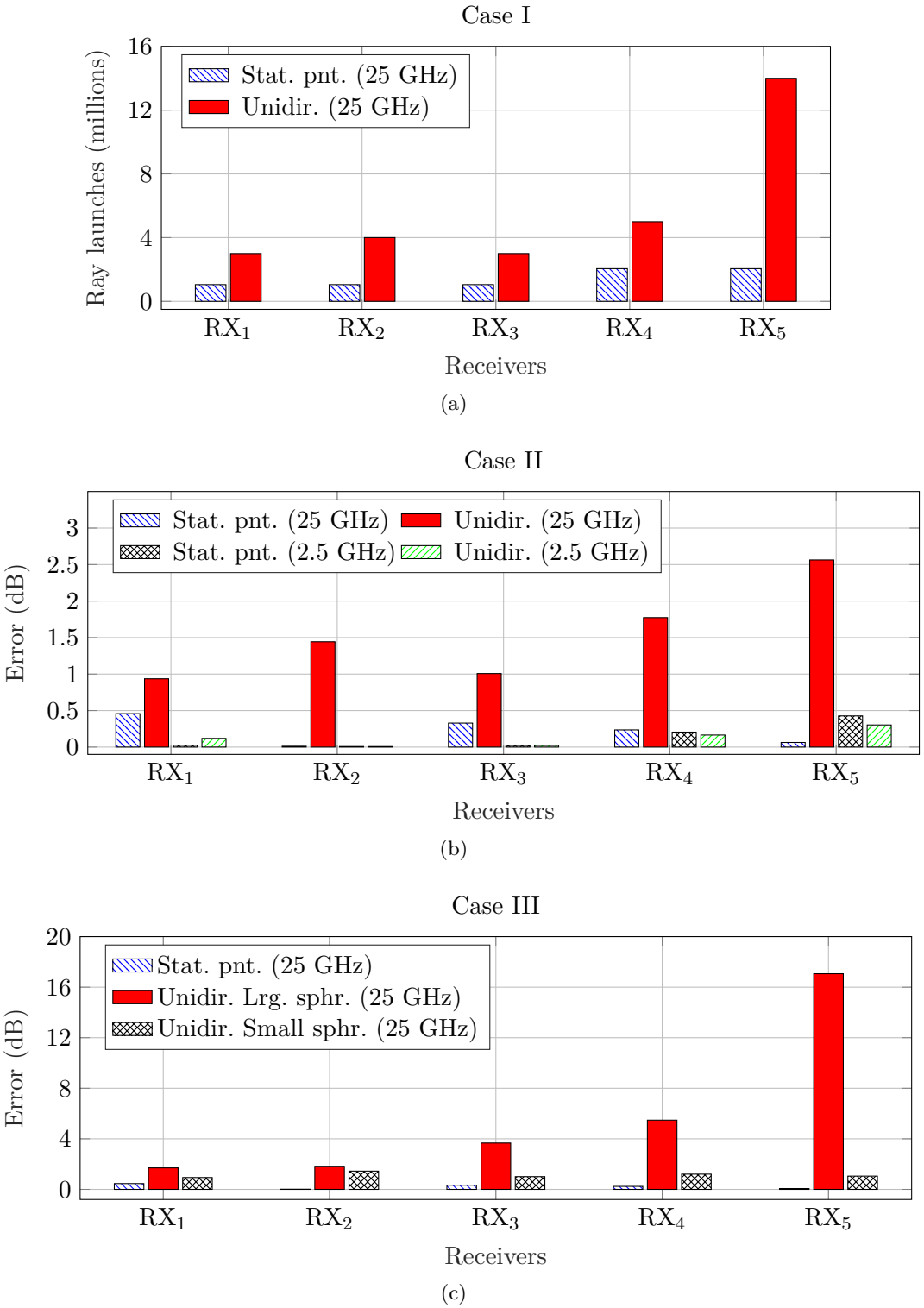


Figure 3.25: The error characteristics of unidirectional and bidirectional ray-tracing with asymptotic expansion in three different urban scenario simulations i.e., for Case I (a), Case II (b) and Case III (c) (Taygur and Eibert, 2020b). © IEEE 2020.

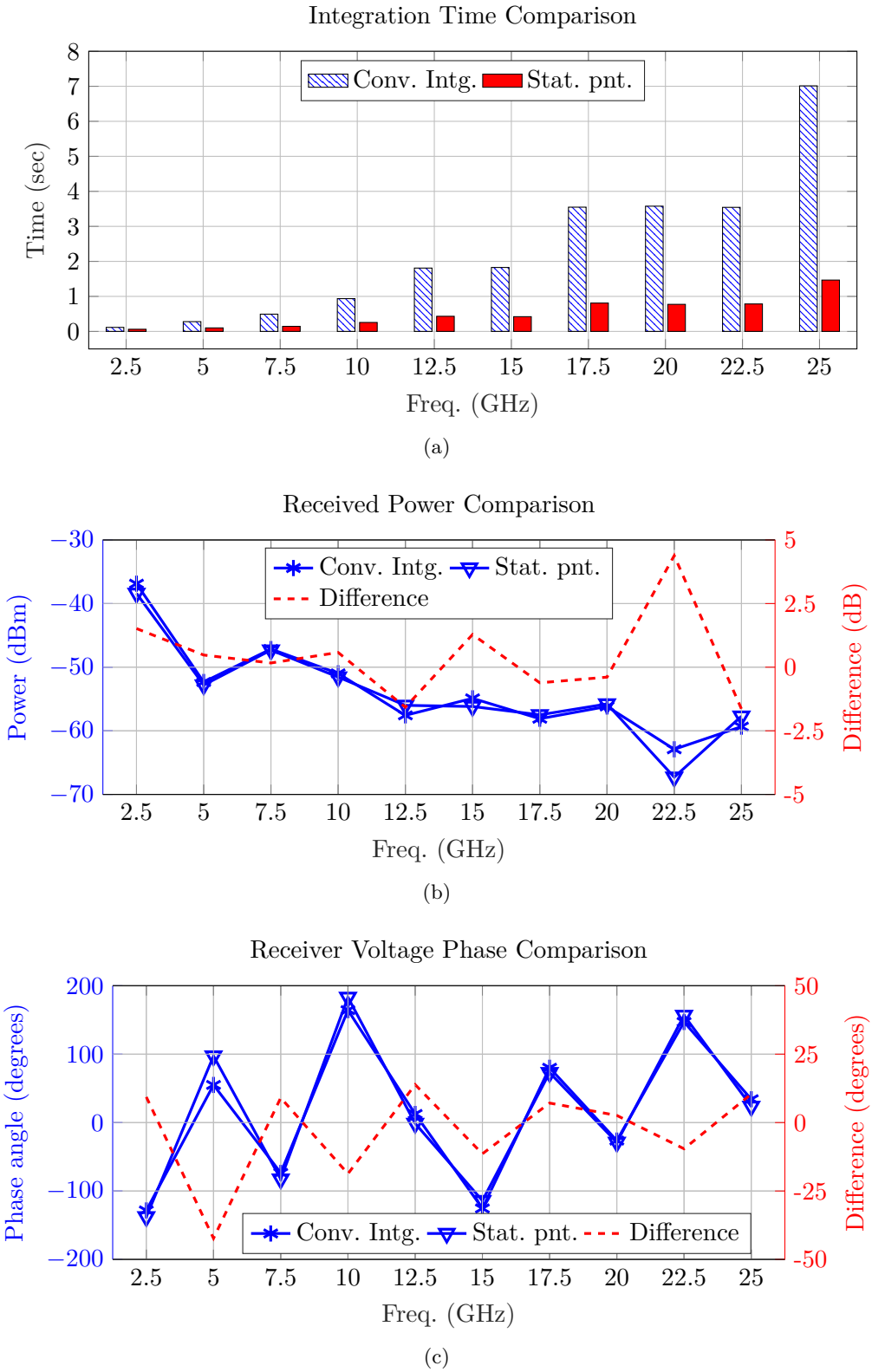


Figure 3.26: Comparison of unidirectional and bidirectional ray-tracing with asymptotic expansion in terms of integration time (a), received power (b) and phase of receive voltage (c) (Taygur and Eibert, 2020b). © IEEE 2020.

3.5 Iterative Shooting and Bouncing Rays Method for Bidirectional Ray-Tracing

The SBR algorithm is frequently utilized to simulate urban-like propagation scenarios where the environment consists of many scatterers, such as buildings, and yield accordingly many reflections. The method typically relies on a large number of ray launches in order to ensure that the number of rays which hit the interaction surface is adequate to evaluate the reciprocity integral. However, most of these launches usually do not yield any ray hits, therefore, the process may become very inefficient. In order to address this issue, an iterative approach might be adopted, i.e., first a small number of ray launches can be used to determine the launch directions which yield ray hits on the interaction surface, then a subsequent launch process is carried out in order to determine the new launch according to the outcome of the previous step (Taygur *et al.*, 2019b). This procedure can be repeated until a convergence criterion or an upper limit on the iteration count is reached. The launch directions, which yield a ray hit at any iteration, are saved, and the new launch directions are chosen in the vicinity of these saved directions within a certain tolerance (here, it can be assumed that the directions are defined according to the spherical coordinates on a unit sphere, thus, ϑ and φ defines a launch direction). Utilizing a large tolerance interval (i.e., $[-\pi/2, \pi/2]$) usually ensures that the feasible ray paths in the scenario can be correctly identified in bidirectional ray-tracing simulations. The pseudocode of the method is given in Algorithm 3.2.

Since the launch directions, which have previously yielded a ray hit, are permanently stored, the number of hits increases at each iteration step, given that the number of ray launches remains constant at each iteration step. Based on the notation in Algorithm 3.2, the probability of a ray hit on the interaction surface (p_0) for the initial ray launch process can be given as

$$\begin{aligned} p_0 &= \frac{K_0}{N},\ N \geq K, \\ M_0 &= N - K_0 = N(1 - p_0),\ 0 < p_0 \leq 1, \end{aligned} \tag{3.52}$$

where $N \in \mathbb{Z}^+$, $K_0 \in \mathbb{Z}^+$ and $M_0 \in \mathbb{Z}^+$ denote the number of ray launches, ray hits and misses, respectively.

Note that the launch directions of N rays during the first iteration are uniformly distributed over a unit sphere, since it is not known initially which launch directions yield ray hits. In the second iteration, M_0 rays are launched in new directions, which are calculated according to the directions of the previous rays. Hence, the total number of ray hits can be given as $K_1 = K_0 + p_1 M_0$ (note that the total number of ray launches is still N). In general, the number of ray hits (K_i) and misses (M_i) at the ith iteration step can be written as

$$K_i = K_{i-1} + p_i M_{i-1},\ M_i = N - K_i, \tag{3.53}$$

where p_i denote the probability of a hit at the ith iteration. The two equalities in Eq. (3.53) yield

$$\begin{aligned} M_i &= N - K_{i-1} - p_i M_{i-1} = N - (N - M_{i-1}) - p_i M_{i-1}, \\ \frac{M_i}{M_{i-1}} &= (1 - p_i) < 1,\ M_i < M_{i-1}. \end{aligned} \tag{3.54}$$

Algorithm 3.2 Iterative SBR pseudocode (TAYGUR *et al.*, 2019b). Reproduced courtesy of The Electromagnetics Academy.

```
1: procedure ITERATIVESBR(d_0, ε)
   ▷ d_0[1 : N][1 : 2]: initial ray directions in (ϑ,φ) in spherical coordinates, N: Number
   of ray launches, ε: iteration stop criteria
2:     d ← d_0;
3:     K_{i−1} ← 0;                    ▷ K_{i−1}:Number of rays which hit in previous iteration
4:     while (true) do
5:         h ← rayTrace(d);
   ▷ h[1 : K_i][1 : 2]: Launch directions of rays which yield a hit, K: Number of ray hits
   at current iteration
6:         if ((K_i − K_{i−1})/K_i < ε) then
7:             break;
8:         end if
9:         if (K < N) then
10:            d[1 : K_i][:] ← h[:][:];
   ▷ The ray directions, which previously yield hits, are re-committed into the buffer
   where the rays are re-traced. The performance can further be improved by avoiding
   re-tracing
11:            for j ← (K : N) do
12:                h[mod(j, K_i)][0] ← h[mod(j, K_i)][0] + Δϑ;
13:                h[mod(j, K_i)][1] ← h[mod(j, K_i)][1] + Δφ;
   ▷ Δϑ,Δφ: Uniformly distributed random variables in [−π/2, π/2]
14:                d[j][:] ← h[mod(j, K_i)][:];
15:            end for
16:        else
17:            break;
18:        end if
19:        K_{i−1} ← K_i;
20:    end while
21: end procedure
```

Thus, the number of missed rays (out of N ray launches) decreases at each iteration. However, there is no strong guarantee that all the feasible ray paths can correctly be identified according to the preceding analysis. The probability of a ray hit on the surface typically improves at each step (value also dependent on the tolerance interval for the new launch directions), as the rays become more focused. In general, the outcome of the first iteration may considerably affect the final result, especially if the tolerance interval for the new launch directions is relatively small, i.e., the new launch directions become usually similar to the launch directions from the previous steps which resulted in a hit. Using a larger interval may alleviate this problem at the expense of slower convergence, as the feasible ray paths, which went undetected in previous iterations, might be discovered.

Note that this algorithm can also be utilized for unidirectional ray-tracing simulations with slight modifications, i.e., reception sphere size and the tolerance interval for the new launch directions may have to be adjusted.

Chapter 4

Multiple Antenna Communication Systems

Multiple antenna communication systems, in particular, multiple-input multiple-output (MIMO) communication systems and its variants for multi-user scenarios, namely multi-user MIMO (MU-MIMO) and massive MIMO, are widely utilized in wireless communications. Since this dissertation involves the characterization of channel properties and data rate performances of various urban massive MIMO scenarios by ray-tracing simulations, a theoretical introduction on multiple antenna communication systems, with an emphasis on massive MIMO, is presented here. The main purpose is to establish a background on the topic and to demonstrate certain important concepts for the reader. This is essential for the subsequent chapters, where the simulations of various MIMO scenarios based on ray-tracing are presented, which are a major contribution of the dissertation. The given material in this chapter is based on the fundamental concepts from the state-of-the-art, and does not introduce any novel contributions.

4.1 Multiple-Input Multiple-Output (MIMO) Systems

The classical MIMO communication concept can be described as the use of multiple antennas at both transmitter and receiver sites in a point-to-point wireless network (Goldsmith, 2005). The principles of multi-antenna communications have already been recognized since several decades, though, the notion of MIMO and its merits have mostly been acknowledged after the pioneering works in the 1990s (Foschini and Gans, 1998; Telatar, 1999). Although conventional multi-antenna transmission/reception strategies, such as beamforming or diversity, may enhance the overall performance or reliability of the network, much more significant improvements in channel capacity can be achieved by spatial multiplexing in a MIMO system (Biglieri *et al.*, 2007).

Spatial multiplexing mainly relies on the transmission of multiple independent data streams over the spatial dimension, which is established by the multiple antennas at both receiver and transmitter sites. The data streams can thus share the same time-frequency resources, enabling relatively large data rates. As multiple antennas are utilized to transmit multiple data streams, each receiver antenna typically picks up a linear combination of the transmitted signals. The number of the independent data streams, which can be supported in a MIMO system, can be found by $\min(M_T, M_R)$ where M_T is the number of the transmitting antennas and M_R is the number of the receiving antennas.

4.1.1 System Model

Consider a narrow-band, quasi-static (i.e., both coherence bandwidth and coherence time are sufficiently large) point-to-point wireless communication link where the transmitter and receiver are equipped with M_T and M_R antennas, respectively. The system model can be given by

$$\begin{bmatrix} \mathbf{y}_{(1)} \\ \vdots \\ \mathbf{y}_{(M_R-1)} \\ \mathbf{y}_{(M_R)} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{(1,1)} & \dots & \dots & \mathbf{G}_{(1,M_T)} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{G}_{(M_R,1)} & \dots & \dots & \mathbf{G}_{(M_R,M_T)} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{(1)} \\ \vdots \\ \mathbf{x}_{(M_T-1)} \\ \mathbf{x}_{(M_T)} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_{(1)} \\ \vdots \\ \mathbf{n}_{(M_R-1)} \\ \mathbf{n}_{(M_R)} \end{bmatrix}, \tag{4.1}$$

or briefly

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n}, \tag{4.2}$$

where $\mathbf{n} \sim \mathcal{CN}\left(0, \sigma_n^2 \mathbf{I}\right)$ denotes white Gaussian noise terms with power σ_n^2, $\mathbf{G} \in \mathbb{C}^{M_R \times M_T}$ is the channel matrix, $\mathbf{x} \in \mathbb{C}^{M_T}$ and $\mathbf{y} \in \mathbb{C}^{M_R}$ are transmitted and received symbol vectors, respectively. The signal which is received by the uth receiver antenna can be written as

$$\mathbf{y}_{(u)} = \sum_{v=1}^{M_T} \mathbf{G}_{(u,v)} \mathbf{x}_{(v)} + \mathbf{n}_{(u)}, \ 1 \leqslant u \leqslant M_R. \tag{4.3}$$

The individual elements of $\mathbf{G}$ denote the channel gain for a specific transmitter and receiver antenna, i.e., $\mathbf{G}_{(u,v)}$ denotes the uth row and the vth column of the matrix and represents the channel coefficient for the uth receiver and the vth transmitter antenna. Here, it should be noted that the channel matrix is assumed to be deterministic (i.e., it can be also considered as an instantaneous state of a time-varying channel) and the time dependence is suppressed within the equation (Tse and Viswanath, 2005). The power of the transmitted signal is typically limited (i.e., to a value of P^{tx}), therefore, trace $\left(E\left[\mathbf{x}\mathbf{x}^H\right]\right) \leqslant P^{\text{tx}}$, where $E[.]$ and $(.)^H$ are expected value and Hermitian transpose operators, respectively.

4.1.2 Channel Capacity

The maximum data transmission rate at which the information can reliably be transferred, in other words the channel capacity, can be computed by Shannon's well-known formula such that (Shannon, 1948)

$$C = \max_{\mathbf{x}} \left(\log_2 \left(\det \left(\mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{G}\mathbf{Q}_{\mathbf{x}}\mathbf{G}^H \right) \right) \right), \text{ s.t. } \text{trace}(\mathbf{Q}_{\mathbf{x}}) \leqslant P^{\text{tx}}. \tag{4.4}$$

The achievable data rate can then be written as

$$R = \max_{\mathbf{x}} \left(W \log_2 \det \left(\mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{G}\mathbf{Q}_{\mathbf{x}}\mathbf{G}^H \right) \right), \text{ s.t. } \text{trace}(\mathbf{Q}_{\mathbf{x}}) \leqslant P^{\text{tx}}, \tag{4.5}$$

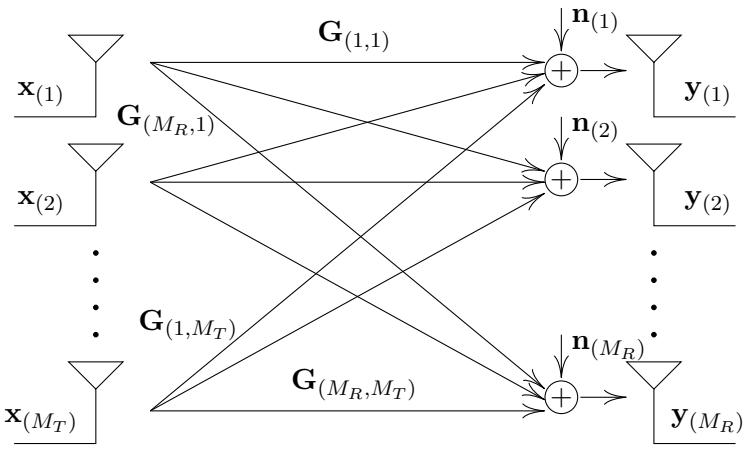


Figure 4.1: MIMO system representation.

where W denotes the channel bandwidth, $\mathbf{Q}_\mathbf{x}$ is the covariance matrix for $\mathbf{x}$ (i.e., $E\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{Q}_\mathbf{x}$) and σ_n^2 is the noise power. It is assumed that $\mathbf{x}$ is Gaussian distributed, as it maximizes the capacity (Cover and Thomas, 2005). The term trace($\mathbf{Q}_\mathbf{x}$) can be implemented as the power allocation scheme, which specifies how the total power P^{tx} is distributed for individual transmitted symbols $\mathbf{x}_{(1)}, \cdots, \mathbf{x}_{(M_T)}$, i.e., the elements of vector $\mathbf{x}$. The optimal power allocation, which maximizes the mutual information (hence, achieving the capacity), can be computed by applying the waterfilling algorithm. In general, the algorithm yields an allocation where the power is allocated predominantly to the symbols which are transmitted through channels with relatively high gains (Goldsmith, 2005). Note that the transmitter needs channel state information (CSI) in order to apply waterfilling, however, in many practical cases where the CSI is unknown to the transmitter (perfect CSI at receiver is still assumed), it is not possible for the transmitter to perform an optimization with respect to $\mathbf{Q}_\mathbf{x}$. As a result, it may not be possible to calculate the optimal power allocation. In such a case, $\mathbf{Q}_\mathbf{x} = \mathbf{I}\, P^{\text{tx}}/M_T$ is generally the conventional choice (Goldsmith, 2005).

4.1.3 Characterizing the Degrees of Freedom

Introducing multiple antennas into the receiver and transmitter devices in a MIMO network should ideally provide a degree of freedom for the transmission of multiple independent data streams (Jafar and Fakhereddin, 2007). In order to identify whether a channel provides sufficient degrees of freedom, and to establish a performance comparison between different MIMO systems, the channel can be decomposed into multiple equivalent SISO channels, and the capacity analysis can be carried out over these equivalent channels (Patil, 2017). Let the channel matrix $\mathbf{G}$ be factorized by a singular value decomposition (SVD), which is given by

$$\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \tag{4.6}$$

where $\mathbf{U} \in \mathbb{C}^{M_R \times M_R}$ and $\mathbf{V} \in \mathbb{C}^{M_T \times M_T}$ are unitary matrices (i.e., $\mathbf{U}^H\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^H\mathbf{V} = \mathbf{I}$) and $\mathbf{\Lambda} \in \mathbb{R}^{M_R \times M_T}$ is a rectangular matrix whose off-diagonal elements are zero and diagonal elements are non-negative real numbers which denote the singular

values of matrix $\mathbf{G}$ (Tse and Viswanath, 2005). The system model equation can then be expressed as

$$\mathbf{y} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H\mathbf{x} + \mathbf{n}. \tag{4.7}$$

It is possible to modify the system such that the signal is multiplied by $\mathbf{V}$ at the transmitter and by $\mathbf{U}^H$ at the receiver, respectively. The multiplications can be considered as transmit precoding and receive beamforming processes (Tsoulos, 2006), respectively, and the system model equation can be given as

$$\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H\mathbf{V}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}, \tag{4.8}$$

with

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{U}^H\mathbf{y}, \\ \tilde{\mathbf{n}} &= \mathbf{U}^H\mathbf{n}, \\ \tilde{\mathbf{x}} &= \mathbf{V}^{-1}\mathbf{x} = \mathbf{V}^H\mathbf{x}, \end{aligned} \tag{4.9}$$

where $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ denote the signals before transmit precoding and after receive beamforming, respectively. Note that neither signal transformation introduces any signal gain or loss, i.e., energy is preserved. Furthermore, the distribution of $\mathbf{n}$ does not change under a unitary transformation, in other words, $\tilde{\mathbf{n}}$ possesses the same statistical properties as $\mathbf{n}$ (Sandmann *et al.*, 2015). Therefore, the capacity and data rate expressions presented in Eq. (4.4) and (4.5) remain valid. Using the properties of unitary matrices

$$\mathbf{U}^H\mathbf{U} = \mathbf{I},\ \mathbf{V}^H\mathbf{V} = \mathbf{I}, \tag{4.10}$$

Eq. (4.8) can be reduced into

$$\tilde{\mathbf{y}} = \boldsymbol{\Lambda}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}. \tag{4.11}$$

Since $\boldsymbol{\Lambda}_{(u,v)} = 0,\ u \neq v$, Eq. (4.11) can be expressed as

$$\tilde{\mathbf{y}}_{(u)} = \Lambda_u\tilde{\mathbf{x}}_{(u)} + \tilde{\mathbf{n}}_{(u)},\ 1 \leqslant u \leqslant \min\left(M_R, M_T\right), \tag{4.12}$$

where $\Lambda_u = \boldsymbol{\Lambda}_{(u,u)}$. It can be noticed that every term in Eq. (4.12) is a scalar, hence, the complete system is represented in terms of multiple parallel scalar (i.e., can be considered as SISO) channels. The number of independent data streams, which can be transmitted in this system, is given by $\min\left(M_R, M_T\right)$, assuming $\forall u$: $\Lambda_u > 0$. Consequently, the channel capacity and data rate expressions can be written as

$$\begin{aligned} C = \sum_{u=1}^{\operatorname{rank}(\mathbf{G})} \log_2\left(1 + \frac{P_u\Lambda_u^2}{\sigma_n^2}\right),\ R = W \sum_{u=1}^{\operatorname{rank}(\mathbf{G})} \log_2\left(1 + \frac{P_u\Lambda_u^2}{\sigma_n^2}\right), \\ \text{s.t.} \sum_{u=1}^{\operatorname{rank}(\mathbf{G})} P_u \leqslant P^{\text{tx}}, \end{aligned} \tag{4.13}$$

where P_u denotes the power allocation coefficients. Based on this representation, the optimal power allocation, in other words, the waterfilling scheme can be expressed as

$$P_u = \begin{cases} \chi - \dfrac{\sigma_n^2}{\Lambda_u^2}, & \dfrac{\sigma_n^2}{\Lambda_u^2} < \chi, \\ 0, & \dfrac{\sigma_n^2}{\Lambda_u^2} \geqslant \chi, \end{cases} \tag{4.14}$$

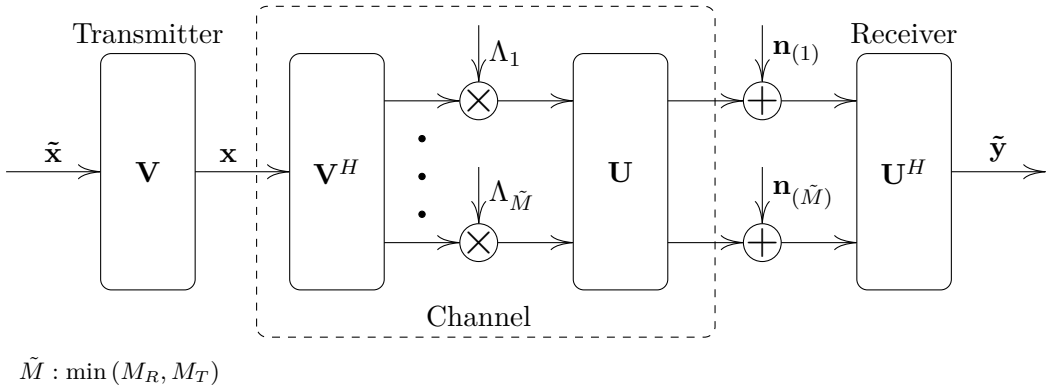


Figure 4.2: MIMO system illustration based on channel decomposition.

where χ is selected accordingly to satisfy the total power constraint. The waterfilling scheme is visualized in Fig. 4.3. The channels, which yield a small value for $\frac{\sigma_n^2}{\Lambda_u^2}$, i.e., a large SNR (signal-to-noise ratio), have a large margin from the threshold χ, and therefore a larger portion of the transmission power can be allocated to the channels with good SNR. As the difference between the threshold and $\frac{\sigma_n^2}{\Lambda_u^2}$ diminishes, the power allocated to the channel decreases accordingly. Certain channels, where the SNR value is small, i.e., $\frac{\sigma_n^2}{\Lambda_u^2}$ exceeds χ, no power can be allocated. Note that the gray shaded area in Fig. 4.3 indicates the total transmission power as a whole. Hence, the waterfilling algorithm ensures that the power is mostly allocated to the channels with favorable conditions, and prevents the waste of power on the channels with a poor SNR (Tsoulos, 2006). In practice, the optimal power allocation is estimated by iteration where the individual allocations and the constant χ are evaluated and modified, until the power constraint is satisfied (Du and Swamy, 2010).

It can be acknowledged that the spatial multiplexing performance is strongly dependent on the singular values of the channel matrix. A noticeable discrepancy in the singular values indicates that the channel matrix is ill-conditioned, as the common cause is usually the similarity among certain channels (i.e., high channel correlation), or the disparity of the channel gains. In either case, the system performance with spatial multiplexing usually declines. A commonly utilized criterion to characterize the similarity/discrepancy of the singular values is the condition number of the channel matrix where large values usually imply a substantial discrepancy while small values indicate similarity (Patil, 2017).

4.2 Multi-User MIMO

The MIMO concept, which has previously been presented, demonstrates the benefits of utilizing more than one antenna at the receiver and transmitter. Nevertheless, its use is limited to point-to-point links where a base station can serve only a single user.

A straightforward method to serve multiple users is to employ time or frequency division techniques where each user can make use of a certain portion of the available time and frequency resources (see Fig. 4.4). However, such a use of time-frequency resources is generally not optimal. It is advantageous to carry out the data transmission

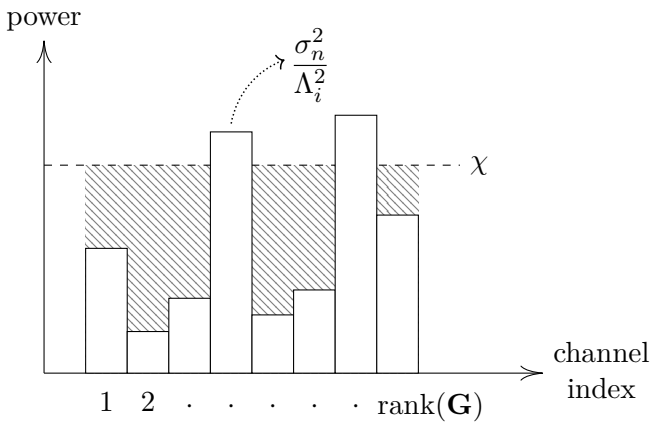


Figure 4.3: Waterfilling algorithm illustration.

simultaneously over the same frequency band for each user by exploiting the spatial signatures of the users, which defines the basic principles of multi-user MIMO (MU-MIMO) (Paulraj *et al.*, 2003). Since the signals are received/transmitted over the same time-frequency resources, significant co-channel interference may arise. Therefore, the transmission strategies and transceiver architectures are generally different than in single-user MIMO (SU-MIMO), since the users are independent and usually may not cooperate to mitigate this interference (Clerckx and Oestges, 2013). Furthermore, the channel characteristics and signal models are different for downlink and uplink transmissions in MU-MIMO due to lack of cooperation among the users. Here, the downlink is performed over a broadcast channel (BC) whereas the uplink over a multiple-access channel (MAC) (see Fig. 4.5) (Heath and Lozano, 2019). Since simultaneous uplink and downlink transmission over the same frequency band is not feasible, these two processes are separated via duplexing, which is commonly realized by either a time division duplexing (TDD) or frequency division duplexing (FDD) scheme. In a TDD-based system uplink and downlink transmissions take place at different time slots/instances whereas the FDD approach requires separate frequency bands to be allocated for uplink and downlink.

The individual user equipment does not have to necessarily possess multiple antennas in a MU-MIMO network, i.e., single-antenna users may also be considered for the sake simplicity. An important difference compared to the SU-MIMO is that the capacity in multi-user channels may not be expressed by a single number, it is rather characterized by a so-called capacity region (i.e., a set in N-dimensional space), where the users are typically associated with different rates (Biglieri *et al.*, 2007). Scalar performance measures based on the capacity region can still be defined though, for instance, the sum capacity, which describes the sum of the achievable spectral efficiency values for all the users in the network (Huang and Papadias, 2008).

It should be noted that the downlink (BC) and uplink (MAC) transmissions in MU-MIMO have a duality relationship. An important outcome of the BC-MAC duality is that the achievable sum capacities are equal under particular power constraints. Such an equality can be demonstrated not only for optimal non-linear coding/decoding schemes, but also for sub-optimal linear approaches (Heath and Lozano, 2019). The analyses in this study are primarily focused on downlink transmission scenarios, noting that

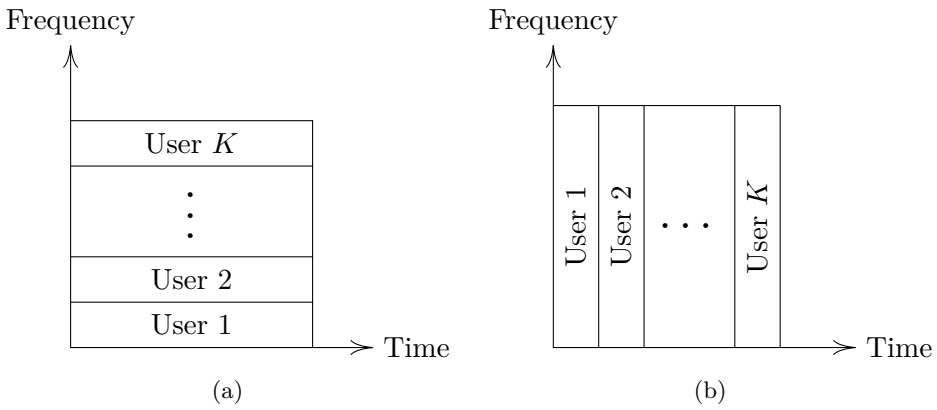


Figure 4.4: Illustration of frequency (a) and time (b) division techniques to serve multiple users in a network.

Figure 4.5: MU-MIMO uplink (a) and downlink (b) illustrations.

the uplink performance would be comparable if particular constraints are met. In the following sections, a time-invariant, narrow-band (i.e., flat-fading) channel and a network with single-antenna users is assumed.

Downlink Broadcast Channel

The broadcast channel (BC) in MU-MIMO defines a downlink channel where multiple downlink data streams are sent from a base station towards multiple users. The base station is responsible for coordinating the signal transmission in a way that the users (which do not cooperate) experience as little interference as possible. Assuming that the transmitted signals are subject to a common total power constraint, the transmission power for each distinct signal should also be determined and allocated accordingly by the base station. This is commonly carried out via a precoder/beamformer, which relies on the availability of accurate CSI. Here, the base station obtains the CSI by means of uplink pilot signals (transmitted from users to the base station) in order to perform precoding. The signal model for a BC downlink system can be written as

$$\mathbf{y}^{\mathrm{dl}} = \mathbf{G}^{\mathrm{dl}}\mathbf{x}^{\mathrm{dl}} + \mathbf{n}^{\mathrm{dl}}, \tag{4.15}$$

where $\mathbf{y}^{\mathrm{dl}}$ denotes the received signals by the users, $\mathbf{G}^{\mathrm{dl}} \in \mathbb{C}^{M_R \times M_T}$ is the channel matrix, $\mathbf{x}^{\mathrm{dl}} \in \mathbb{C}^{M_T}$ is the transmitted signal from the base station and $\mathbf{n}^{\mathrm{dl}} \sim \mathcal{CN}\left(0, \sigma_n^2\,\mathbf{I}\right)$ is the noise term. Note that $\mathbf{x}^{\mathrm{dl}}$ denotes here the signal after the precoder, and the

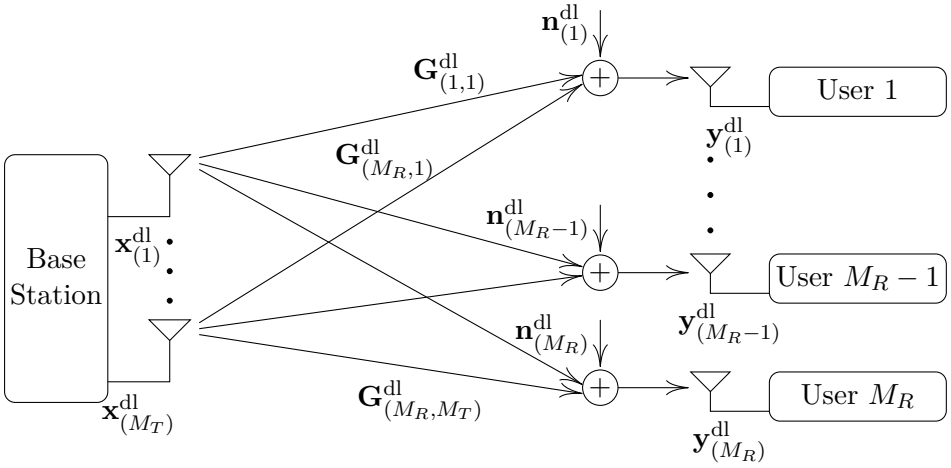


Figure 4.6: MU-MIMO BC system representation.

transmission of individual user signals is assumed to be synchronous. An illustration of the system is shown in Fig. 4.6.

Sum Capacity

The optimal performance can be achieved by utilizing the *dirty-paper coding* (DPC) technique at the base station. The premise of DPC relies on the fact that the interference at the receiver sites can be mitigated by a pre-subtraction at the base station, if the channel is perfectly known to the transmitter. The BC sum capacity expression can be given as (Costa, 1983)

$$\begin{aligned}
\sum_{u\in\mathcal{U}} C_u &\leqslant \log_2 \det\left(\mathbf{I} + \frac{1}{\sigma_n^2}\sum_{u\in\mathcal{U}} P_u \left(\mathbf{G}^{\mathrm{dl}}_{(u,:)}\right)^H \mathbf{G}^{\mathrm{dl}}_{(u,:)}\right) \\
\sum_{u\in\mathcal{U}} C_u &\leqslant \log_2 \det\left(\mathbf{I} + \frac{1}{\sigma_n^2}\left(\mathbf{G}^{\mathrm{dl}}\right)^H \mathbf{Q}_{\mathbf{x}}\mathbf{G}^{\mathrm{dl}}\right) \\
\mathbf{Q}_{\mathbf{x}} &= \mathrm{diag}\left(P_1, \cdots, P_{M_R}\right),\ \mathrm{trace}\left(\mathbf{Q}_{\mathbf{x}}\right) \leqslant P^{\mathrm{tx}}, \\
\mathcal{U} &\subseteq \{1, \cdots, M_R\},\ 1 \leqslant u \leqslant M_R,
\end{aligned} \tag{4.16}$$

where the covariance matrix $\mathbf{Q}_{\mathbf{x}}$ is to be optimized in order to maximize the sum capacity. A detailed elaboration for the derivation of Eq. (4.16) based on BC-MAC duality can be found in (Clerckx and Oestges, 2013). Various examples concerning the implementation of DPC-based precoding can be found in (Tse and Viswanath, 2005; Erez and ten Brink, 2005; Erez *et al.*, 2005; Elliott and Krzymien, 2009).

The sum capacity for MAC at the uplink may also be given by Eq. (4.16), if the total transmit power from all users equals to P^{tx} and the noise covariance matrix $\mathbf{Q}_{\mathbf{n}} = \mathbf{I}\sigma_n^2$.

Although DPC-based precoding is the optimal approach concerning the achievable sum capacity, it is rarely considered as a viable option in real-world applications due to its prohibitively high complexity and sensitivity to inaccuracies in the CSI (Stankovic and Haardt, 2008; Khina and Erez, 2010). Therefore, linear precoders, which have simpler designs, are typically preferred over DPC.

Linear Precoding

Let us consider the BC representation in Eq. (4.15), where $\mathbf{x}^{\mathrm{dl}}$ denotes the transmitted signal by the base station after precoding takes place. If the transmitter employs a linear precoder, then Eq. (4.15) can be re-written as

$$\begin{aligned} \mathbf{y}^{\mathrm{dl}} &= \mathbf{G}^{\mathrm{dl}}\mathbf{A}^{\mathrm{dl}}\tilde{\mathbf{x}}^{\mathrm{dl}} + \mathbf{n}^{\mathrm{dl}}, \\ \mathbf{A}^{\mathrm{dl}}\tilde{\mathbf{x}}^{\mathrm{dl}} &= \mathbf{x}^{\mathrm{dl}}, \end{aligned} \tag{4.17}$$

where $\tilde{\mathbf{x}}^{\mathrm{dl}} \in \mathbb{C}^{M_R}$ contains the data intended for individual users and $\mathbf{A}^{\mathrm{dl}} \in \mathbb{C}^{M_T \times M_R}$ is the precoding matrix. Hence, the signal and interference, which is received by the uth user in the network, can be written as

$$\mathbf{y}^{\mathrm{dl}}_{(u)} = \underbrace{\mathbf{G}^{\mathrm{dl}}_{(u,:)}\mathbf{A}^{\mathrm{dl}}_{(:,u)}\tilde{\mathbf{x}}^{\mathrm{dl}}_{(u)}}_{\text{User signal}} + \underbrace{\sum_{\substack{v=1 \\ v\neq u}}^{M_R} \mathbf{G}^{\mathrm{dl}}_{(u,:)}\mathbf{A}^{\mathrm{dl}}_{(:,v)}\tilde{\mathbf{x}}^{\mathrm{dl}}_{(v)}}_{\text{Interference}} + \underbrace{\mathbf{n}^{\mathrm{dl}}_{(u)}}_{\text{Noise}}. \tag{4.18}$$

Thus, the SINR (signal-to-interference and noise ratio) and the individual capacity for user u can be given by

$$\begin{aligned} \mathrm{SINR}^{\mathrm{dl}}_u &= \frac{P^{\mathrm{tx}}|\mathbf{G}^{\mathrm{dl}}_{(u,:)}\mathbf{A}^{\mathrm{dl}}_{(:,u)}|^2}{P^{\mathrm{tx}}\sum\limits_{v=1,v\neq u}^{M_R}|\mathbf{G}^{\mathrm{dl}}_{(u,:)}\mathbf{A}^{\mathrm{dl}}_{(:,v)}|^2 + \sigma_n^2}, \\ C_u &= \log_2\left(1 + \mathrm{SINR}^{\mathrm{dl}}_u\right). \end{aligned} \tag{4.19}$$

The sum capacity can then be calculated as

$$C^{\mathrm{sum}} = \sum_{u=1}^{M_R} C_u. \tag{4.20}$$

The formulations of linear precoding techniques can be given by

$$\begin{aligned} &\text{Maximum ratio combining (MRC)} \rightarrow \mathbf{A}^{\text{dl-MRC}} = \left(\mathbf{G}^{\mathrm{dl}}\right)^H, \\ &\text{Zero-forcing (ZF)} \rightarrow \mathbf{A}^{\text{dl-ZF}} = \left(\mathbf{G}^{\mathrm{dl}}\right)^H \left(\mathbf{G}^{\mathrm{dl}}\left(\mathbf{G}^{\mathrm{dl}}\right)^H\right)^+, \\ &\text{Regularized zero-forcing (RZF)} \rightarrow \mathbf{A}^{\text{dl-RZF}} = \left(\mathbf{G}^{\mathrm{dl}}\right)^H \left(\mathbf{G}^{\mathrm{dl}}\left(\mathbf{G}^{\mathrm{dl}}\right)^H + \frac{\sigma_n^2}{P^{\mathrm{tx}}}\mathbf{I}\right)^{-1}. \end{aligned} \tag{4.21}$$

A power scaling factor may also be included in order to establish a more precise control for the transmit power. Maximum ratio combining is the simplest precoder among the given schemes, however, it is generally not possible to achieve the performance of a ZF or RZF precoder, unless the system operates in the low SNR regime. At a sufficiently high SNR, the performance of the MRC precoder deteriorates as the inter-user interference may not efficiently be suppressed. On the other hand, the ZF precoder yields ideally a

diagonal system matrix $\mathbf{G}^{\text{dl}}\mathbf{A}^{\text{dl-ZF}}$, which is interference-free and therefore, outperforms MRC in the high-SNR regime. A common issue of ZF precoders is the noise enhancement, which may arise if $\mathbf{G}^{\text{dl}}\left(\mathbf{G}^{\text{dl}}\right)^{H}$ is ill-conditioned (Gao, 2016). In order to address this problem, a regularization factor can be added into ZF formulation, which is then usually called RZF precoding (Heath and Lozano, 2019). RZF may generally be considered as a superior scheme, as the drawbacks of the other two approaches can be avoided.

4.3 Massive MIMO

The traditional MU-MIMO concept typically describes a setting where the number of base station antennas and the total number of user terminal antennas in the network are comparable. Nevertheless, introducing a large number of antennas into the base station (i.e., several times larger than the number of user antennas) yields considerable performance gains due to distinct channel characteristics (Larsson, 2013). This has raised important questions concerning the limits of using large arrays at the base stations in the last decade. Preliminary studies on the subject have demonstrated that considerable gains in spectral efficiency/data rate can be achieved when such large arrays are utilized as the inter-user interference can significantly be reduced with simple linear precoding techniques (Hien Quoc Ngo *et al.*, 2013). This concept, i.e., employing a large number of base station antennas for serving a much smaller number of users, is considered to be a drastic paradigm shift from conventional MU-MIMO, thereby it is commonly referred to as massive MIMO (Marzetta, 2010; Luo and Zhang, 2016).

Using certain properties of random matrices, it can be shown that inter-user interference can be reduced to zero in Gaussian channels by linear precoding schemes as the number of base station antennas is increased asymptotically. Furthermore, the performance (i.e., sum capacity) of linear techniques in this regime approaches to that of optimal non-linear methods, though the complexity of linear transceivers is generally much less (Zhao *et al.*, 2018). Let us now consider a channel described by $\mathbf{G} \in \mathbb{C}^{M_R \times M_T}$ whose entries are independent and identically distributed, circularly symmetric complex Gaussian random variables with zero mean and unit variance. According to the Marchenko-Pastur law, the Gram matrix given by

$$\mathbf{Z} = \frac{1}{M_T}\mathbf{G}\mathbf{G}^{H}, \tag{4.22}$$

has eigenvalues distributed according to (Tulino and Verdú, 2004)

$$\begin{gathered}
\Phi_{\sigma_\Lambda(\mathbf{Z})}(x) = \left[1 - \frac{1}{\beta}\right]^{+}\delta(x) + \frac{\sqrt{[x-a]^{+}\,[b-x]^{+}}}{2\pi\beta x}, \\
a = \left(1 - \frac{1}{\sqrt{\beta}}\right)^{2},\ b = \left(1 + \frac{1}{\sqrt{\beta}}\right)^{2}, \\
\beta = \frac{M_T}{M_R},\ [\alpha]^{+} = \max(\alpha, 0), \\
\delta(x) = \begin{cases} 1, & x = 0, \\ 0, & \text{otherwise}, \end{cases}
\end{gathered} \tag{4.23}$$

where $\sigma_\Lambda(\mathbf{Z})$ denotes the spectrum of $\mathbf{Z}$ (i.e., set of all the eigenvalues) and $\Phi_{\sigma_\Lambda(\mathbf{Z})}(x)$ is the probability density function for $\sigma_\Lambda(\mathbf{Z})$. If the number of the base station antennas

grows asymptotically, then the smallest and largest eigenvalues of the matrix $\mathbf{Z}$ can be given by (Rusek *et al.*, 2013)

$$\lim_{M_T \to \infty} \min\left(\sigma_\Lambda\left(\mathbf{Z}\right)\right) = 1, \ \lim_{M_T \to \infty} \max\left(\sigma_\Lambda\left(\mathbf{Z}\right)\right) = 1, \tag{4.24}$$

which also implies (Marzetta, 2010; Clerckx and Oestges, 2013)

$$\mathbf{Z} = \frac{1}{M_T}\mathbf{G}\mathbf{G}^H \xrightarrow{M_T \to \infty} \mathbf{I}. \tag{4.25}$$

Using the result in Eq. (4.25), it can be inferred that the effective channel matrix expressed by $\mathbf{GA}$ should be diagonal. Consequently, the interference terms in the effective channel can be written as

$$\lim_{M_T \to \infty} \mathbf{G}_{(u,:)}\mathbf{A}_{(:,v)} = 0, \ v \neq u. \tag{4.26}$$

The eigenvalue distribution and the convergence of the effective channel matrix to the identity matrix with respect to the number of base station antennas are shown in Fig. 4.7 and 4.8, respectively.

As the inter-user interference diminishes, so-called *favorable propagation* conditions can be observed, i.e., the effective user channels become mutually orthogonal (Ngo *et al.*, 2014). Note that a downlink transmission case with maximum-ratio combiner $\mathbf{A} = \mathbf{G}^H$ is implicitly assumed in Eqs. (4.23)-(4.26).

Besides the considerations on the decline of inter-user interference, an important implication of the asymptotic behavior in massive MIMO is that the user channels become more deterministic as the number of base station antennas is increased such that (Ngo and Larsson, 2017)

$$\frac{\left|\left|\mathbf{G}_{(v,:)}\right|\right|^2}{E\left[\left\|\mathbf{G}_{(v,:)}\right\|^2\right]} \xrightarrow{M_T \to \infty} 1, \ 1 \leqslant v \leqslant M_R. \tag{4.27}$$

This phenomenon is commonly known as *channel hardening* and it is shown to improve the link reliability as the fluctuations in the channel due to fading vanish (Gunnarsson *et al.*, 2018).

Although the asymptotic analyses as well as the channel capacity investigations for large-scale MU-MIMO networks commonly assume Rayleigh fading channels (i.e., i.i.d. complex Gaussian channel coefficients), such an assumption is not necessary to observe the benefits of large numbers of base station antennas. In practice, it is possible to observe favorable propagation as well as the channel hardening phenomenon in numerous real-world scenarios, despite the channel coefficients not being truly Gaussian. For the particular case of single-antenna users, achieving mutually orthogonal user channels is relatively simple, unless the angular separation between the users is very little such that the corresponding channel vectors are practically co-linear (Marzetta, 2010; Ngo *et al.*, 2011). Although correlation among the user channels may cripple the network performance, it might be possible to overcome this issue by adding more antennas into the base station. Studies based on realistic channel evaluations show that 80% of the optimal sum capacity performance is attainable by linear precoding even in densely populated environments (Zhao *et al.*, 2018). On the other hand, the situation changes for multi-antenna users since it is generally not possible to guarantee an adequate separation

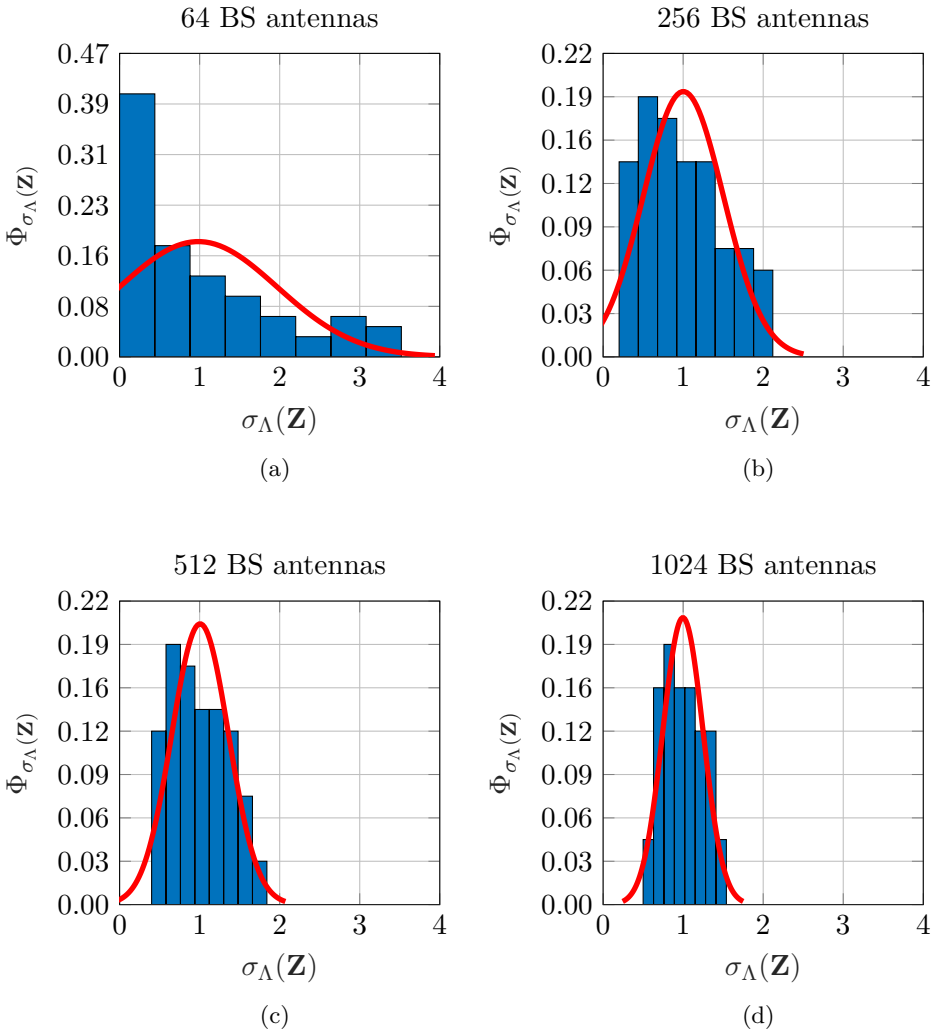


Figure 4.7: Eigenvalue distribution for the effective channel matrix ($\mathbf{Z}$) with 64 (a), 256 (b), 512 (c) and 1024 (d) antennas at the base station. The red curves in each plot show the best fitting Gaussian distribution.

between the antennas of a single terminal, thus, the channel correlation among the users remains high even with a very large number of base station antennas. In this case, more sophisticated transceiver algorithms (e.g., block diagonalization) are preferred in order to take advantage of the additional antenna(s) at the user terminals.

A crucial requirement for achieving the performance gains in the large-array regime is that the base station has accurate CSI (Luo and Zhang, 2016). In practice, the CSI acquisition process involves the estimation of channel coefficients, usually by means of pilot signals (also known as training based estimation) (Chockalingam and Rajan, 2014). Although a reliable and accurate channel estimate is an important prerequisite for achieving the optimal performance in almost every multi-antenna system, the implications of the process are generally much more profound for massive MIMO.

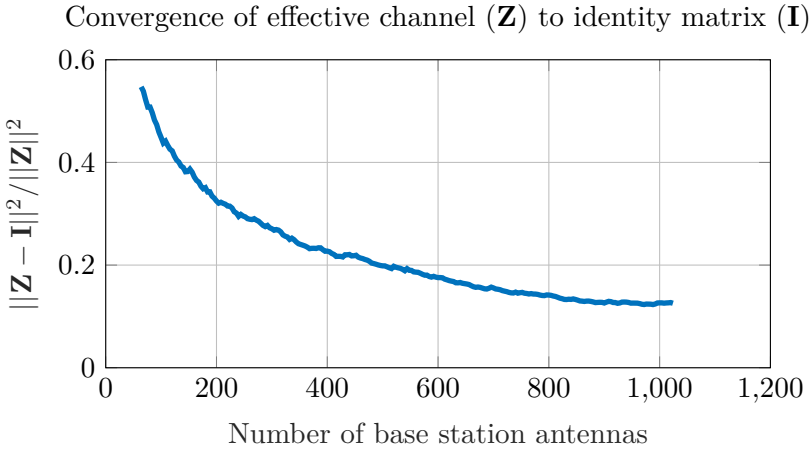


Figure 4.8: Convergence of effective channel matrix ($\mathbf{Z}$) to identity matrix ($\mathbf{I}$) as the number of base station antennas is increased.

4.3.1 CSI Acquisition

The acquisition of CSI via pilot transmission is not error-free, as the estimated channel generally differs from the actual channel. The difference can simply be expressed by

$$\mathbf{G} = \hat{\mathbf{G}} + \tilde{\mathbf{G}}, \tag{4.28}$$

where $\hat{\mathbf{G}}$ denote the estimation error, $\mathbf{G}$ and $\tilde{\mathbf{G}}$ are the actual and estimated channel matrices, respectively. Apart from hardware-related deficiencies, one of the major causes of erroneous estimation is channel noise, which corrupts the pilot signals in the same way that it affects data transmission (Heath and Lozano, 2019). The pilot signals might also be impaired due to interference-related disturbances, which can be arduous to deal with, as the interference typically stems from the pilot signals of other users in the network. In a hypothetically ideal scenario where the pilots are not corrupted by noise or interference, CSI acquisition might still be far from perfect due to the time-varying nature of the channel. Here, discrepancies may arise between the actual channel and the estimation obtained by pilots, if the channel varies rapidly but the acquisition process is not performed accordingly frequent, which yields the so-called *channel aging* problem. Regardless of the causes, an imperfect channel estimation generally yields a reduction in the achievable data rate. Although decent improvements in channel capacity can still be obtained by increasing the number of base station antennas with an imperfect CSI, the performance gains are usually limited (Taesang Yoo and Goldsmith, 2006).

A key concern regarding the implementation of CSI acquisition in massive MIMO is the feasibility of TDD and FDD based transmission. The TDD approach offers a considerable benefit in general since channel reciprocity typically holds for both uplink and downlink transmission, where the same frequency band is utilized (within a sufficiently short time frame so that the channel is assumed to be stationary). Hence, uplink pilot signals transmitted by the users can effectively be utilized to estimate the channel for both receiving (uplink) and precoding (downlink). Such a presumption about reciprocity is not possible in FDD-based communication though, since uplink and downlink channels occupy different frequency bands, as shown in Fig. 4.9. Therefore, CSI acquisition for FDD-based communication involves pilot transmission in both ways. In particular,

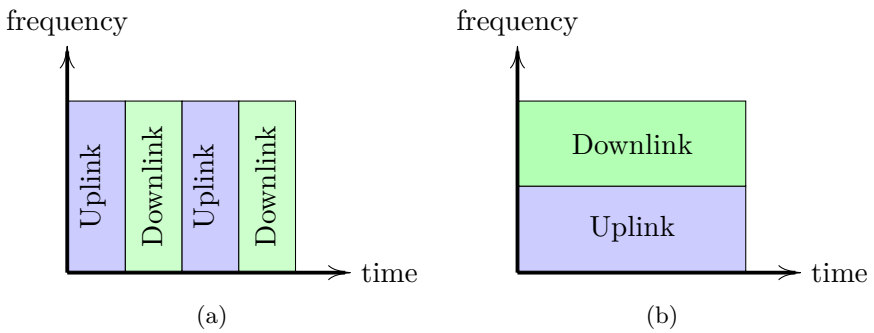


Figure 4.9: Illustrative comparison of resource allocation for uplink and downlink transmissions in time-division duplexing (a) and frequency-division duplexing (b) schemes.

the base station first transmits downlink pilots to users, which then feed this information back to the base station along with their respective uplink pilot signals (Elijah *et al.*, 2016). Here, it can be noticed that the CSI acquisition overhead in an FDD-based system becomes larger when the number of base station antennas is increased (due to the downlink feedback) whereas the overhead for TDD depends on the number of user antennas, implying a significant advantage in terms of scalability (Jiang *et al.*, 2015).

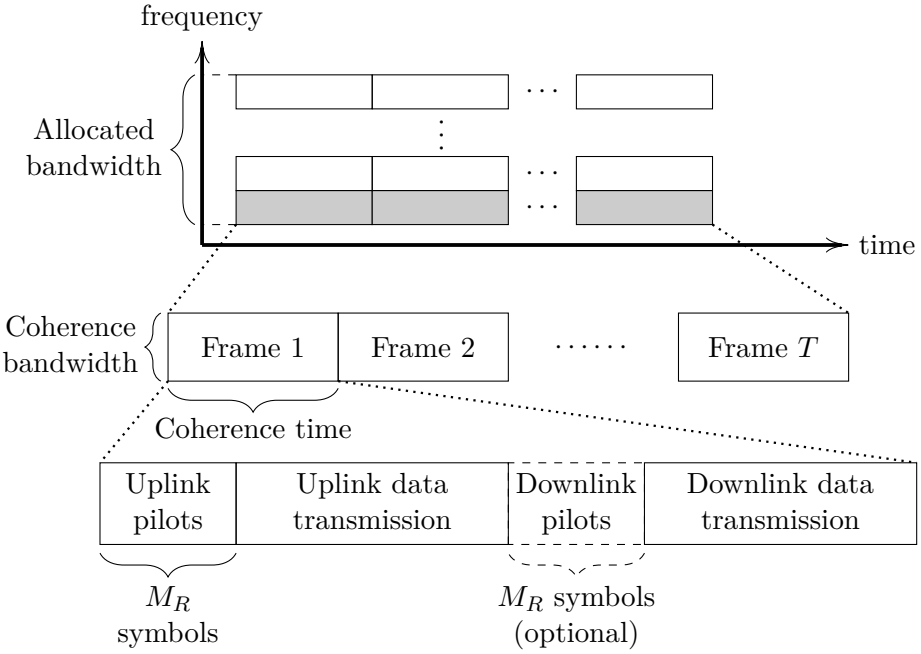


Figure 4.10: Illustration of a multi-carrier TDD-based massive MIMO frame structure.

A typical TDD-based transmission scheme can be described as follows (also see Fig. 4.10):

1. The users transmit (orthogonal) pilot signals.
2. The base station estimates the channel according to the pilots, then computes the receive filter and precoding matrix for uplink and downlink, respectively.

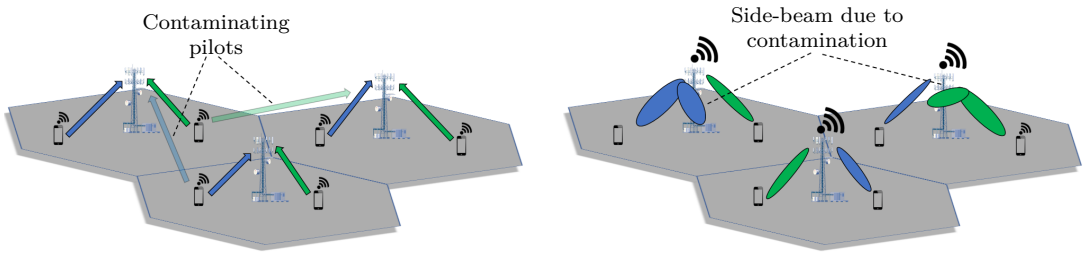
3. Uplink data transmission.
4. Downlink data transmission.

The third and fourth steps may usually be interchanged, i.e., downlink data transmission may take place before the uplink. In the specific case that user terminals contain multiple antennas, an additional operation prior to downlink data transmission might be performed, where the base station transmits downlink pilot signals through the precoded channel. Thus, users would acquire the information on the effective channel after precoding, and can adapt their own receivers accordingly (Heath and Lozano, 2019).

4.3.2 Pilot Contamination

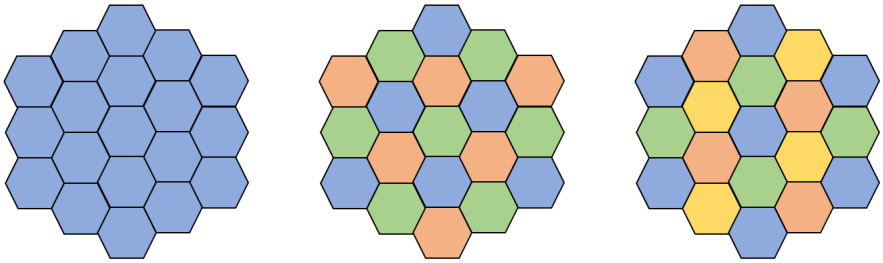
The transmission of pilot signals as well as the actual data commonly takes place within a short time frame where the channel does not vary drastically (i.e., coherence time). In general, this implies a constraint on the number of pilot signals which can be allocated in a single frame, as the users utilize distinct time slots to transmit their pilots (known as the time-multiplexed pilot scheme), in order to ensure orthogonality (Coldrey and Bohlin, 2007). Although it is generally not difficult to allocate mutually orthogonal pilots for the users in a single cell (consisting of a single base station), it is practically impossible to comply with such a requirement in a large, crowded network, comprising of many cells. Consequently, pilot signals should be reused (here, the reuse of pilots implies that more than one user send a pilot signal in the same time slot) at different parts of the network, yielding an inter-cellular interference, which is commonly known as *pilot contamination*, as depicted in Fig. 4.11. It should be noted that the pilot contamination phenomenon is not exclusive to massive MIMO, however, it is a much more relevant issue for such large networks, since pilot contamination may not be completely eliminated no matter how large the number of base station antennas is, while many other problems, such as inter-user interference or fading can effectively be mitigated. Nevertheless, it is possible to confine the adverse effects of pilot contamination, as the most common approach is to utilize distinct pilot signals among adjacent cells, such that the inter-cellular interference would affect only the second (or higher) order neighbors of a particular cell (see Fig. 4.12). Although the number of orthogonal pilots required in the network may still be significantly larger than that is necessary in a single-cell with this strategy, it can still efficiently be implemented without exhausting the available time-frequency resources in many practical scenarios. Here, an important assumption concerning the timing of CSI acquisition and data transmission is the time synchronization among the cells. One might argue that a single set of orthogonal pilot signals can be sufficient for each cell, if the CSI acquisition is performed in a particular cell while data transmission takes place in every other cell (Zhao *et al.*, 2018). Utilizing such staggered pilot sequences may in fact be helpful to eliminate the interference due to uplink pilots, even if the same set of signals are utilized in every cell. However, the inter-cellular interference may not be completely eliminated, since the data transmission in other cells might also yield significant interference (see Fig. 4.13) (Heath and Lozano, 2019).

Resolving the pilot contamination issue may not be straightforward, if the CSI acquisition overhead is already excessive due to a large number of orthogonal pilots. In certain cases, it is possible to differentiate between non-orthogonal pilot signals by exploiting certain channel related features though, hence, pilot contamination can be mitigated to a certain extent. Various examples, which demonstrate the advantages of this approach,



(a) (b)

Figure 4.11: Demonstration of pilot contamination and its effect in a 3-cell scenario. As the users first transmit the uplink pilots, an interference towards the neighboring cells may arise (a). Consequently, the precoding process is impaired where the base station may also try to serve the users in other cells which caused the pilot interference (b).

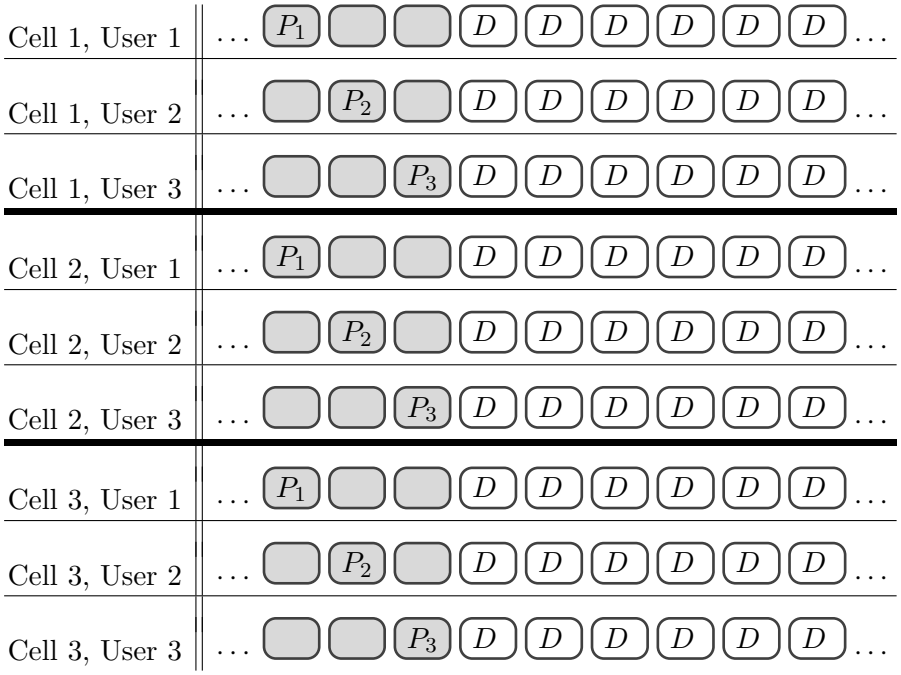


Reuse factor = 1 Reuse factor = 3 Reuse factor = 4

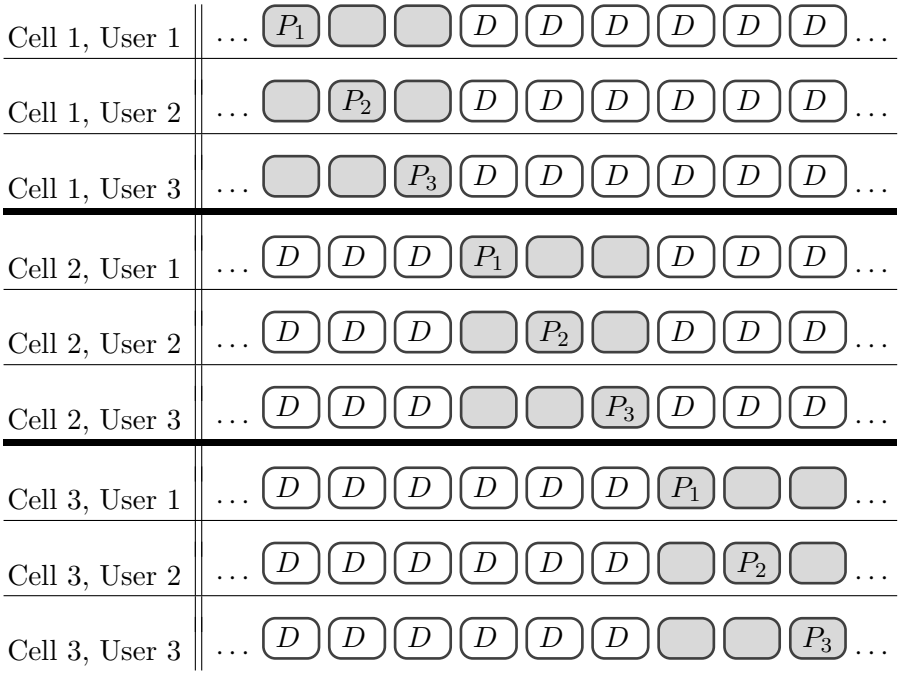
(a) (b) (c)

Figure 4.12: Illustration of pilot reuse in multi-cell (each hexagon denotes a single-cell and each color indicates a distinct set of pilot signals used in the particular cell) environments with an according reuse factor of 1 (a), 3 (b) and 4 (c).

can be found in (Yin *et al.*, 2014, 2016; Li, Mingmei *et al.*, 2013; Wang *et al.*, 2015), where the correct pilot signal and the contaminator (in other words, the interfering terminals) are distinguished according to the angular properties of the received waves, which arrive the base station from different directions and therefore interact with different antennas. Hence, the interference can be suppressed, once the spatial features of the correct pilot signal are identified. In addition to the utilization of spatial features for identifying pilot interference, it is also possible to exploit certain discrepancies in the time domain, such as propagation delay (Chen, 2016; Haghighatshoar and Caire, 2017). Here, it is reasonable to assume that the propagation distance for the contaminating users is longer than that for the in-cell users, thus, they can easily be identified and eliminated according to the differences in propagation delay, if the uplink transmission is synchronized. Note that the use of time differences and spatial discrepancies for suppressing pilot contamination may require a certain degree of coordination among the cells in the network (Heath and Lozano, 2019).



(a)



(b)

Figure 4.13: Comparison of aligned (a) and staggered (b) pilot transmission for a 3-cell (assumed to be adjacent) network with 9 users in total. P_n and D denote the nth orthogonal pilot signal and a data frame, either uplink or downlink, respectively.

4.3.3 Downlink Performance

Despite certain conceptual differences, a rigorous sum capacity analysis for a massive MIMO downlink can be given by using the relevant results from the MU-MIMO case. Assuming that linear transceivers are utilized, the downlink capacity expression (C_u^{dl}) for a single user (denoted by u) can be given as

$$\mathrm{SINR}_u^{\mathrm{dl}} = \frac{P^{\mathrm{tx}}|\mathbf{G}^{\mathrm{dl}}_{(u,:)}\mathbf{A}^{\mathrm{dl}}_{(:,u)}|^2}{P^{\mathrm{tx}}\left(\sum\limits_{v=1,v\neq u}^{M_R}|\mathbf{G}^{\mathrm{dl}}_{(u,:)}\mathbf{A}^{\mathrm{dl}}_{(:,v)}|^2\right)+\sigma_n^2},\ C_u = \log_2\left(1+\mathrm{SINR}_u^{\mathrm{dl}}\right), \tag{4.29}$$

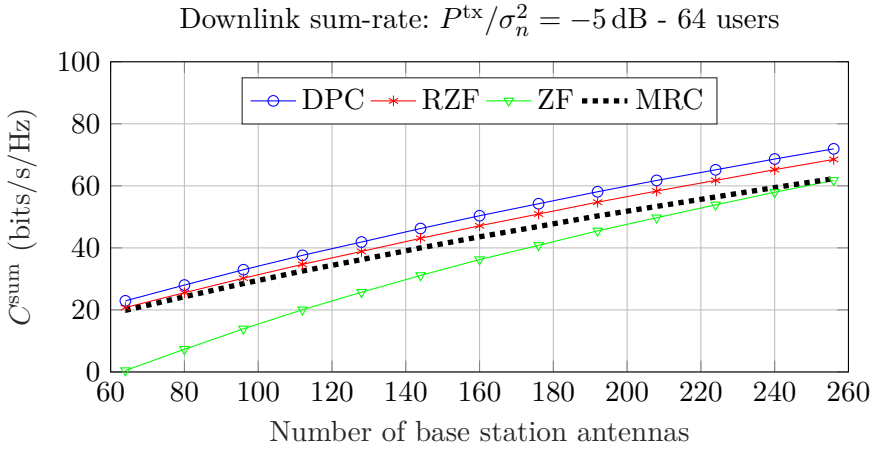
$$\mathbf{G}^{\mathrm{dl}} \in \mathbb{C}^{M_R\times M_T}, \tag{4.30}$$

where P^{tx} is the total transmission power of the base station, $\mathbf{A}^{\mathrm{dl}}$ denote the precoding matrix, which are the same as those given in Eq. (4.21). Similarly, the optimal sum capacity expression can be given according to the conventional MU-MIMO formulations such that

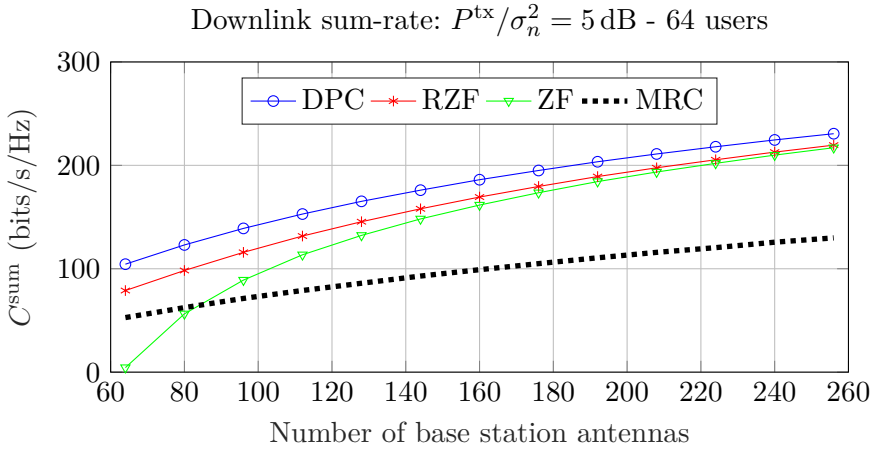
$$C^{\mathrm{sum}} = \log_2\det\left(\mathbf{I}+\frac{1}{\sigma_n^2}\left(\mathbf{G}^{\mathrm{dl}}\right)^H\mathbf{Q}_{\mathbf{x}}\mathbf{G}^{\mathrm{dl}}\right),$$
$$\mathbf{Q}_{\mathbf{x}} = \mathrm{diag}\left(P_1,\cdots,P_{M_R}\right),\ \mathrm{trace}\left(\mathbf{Q}_{\mathbf{x}}\right)\leqslant P^{\mathrm{tx}}. \tag{4.31}$$

An important consideration here is the performance of the linear precoding methods compared to that of DPC when the number of base station antennas is increased. In order to make a comparison, a downlink sum capacity analysis will be presented for a single-cell scenario with 3 different power settings and 64 users in the following, where a simple stochastic channel model is utilized. The channel coefficients, i.e., the individual elements of the channel matrix $\mathbf{G}^{\mathrm{dl}}$, were generated by circularly symmetric complex Gaussian random variables with zero mean and unit variance. The sum capacity performance under linear precoding is given by Eq. (4.30) where the matrix $\mathbf{A}^{\mathrm{dl}}$ depends on the choice of precoder, as shown in Eq. (4.21). The DPC performance is given by Eq. (4.31) where an equal power allocation is assumed, hence $\mathbf{Q}_{\mathbf{x}} = P^{\mathrm{tx}}\mathbf{I}$. The three different power settings are realized by changing the $P^{\mathrm{tx}}/\sigma_n^2$ ratio. The number of base station antennas is changed from 64 to 1024 within two different domains, i.e., small-array (number of antennas 64-256) and large-array (number of antennas 256-1024), and the results are shown in Fig. 4.14 and 4.15, respectively.

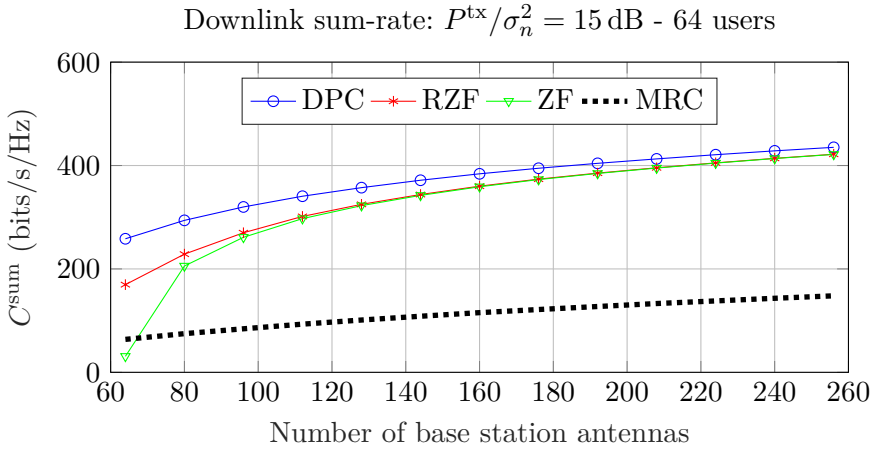
The results represent two different array configurations where the ratio of the number of base station antennas and the number of users (i.e., M_T/M_R) varies from 1 to 4 and 4 to 16 in Fig. 4.14 and 4.15, respectively. The benefits of increasing the number of base station antennas, can be observed for each individual case, while the number of the users in the network remains constant. The rate of change is particularly high for the low-SNR regime ($P^{\mathrm{tx}}/\sigma_n^2 = -5\,\mathrm{dB}$) according to both figures. The performance gains, which can be obtained by introducing more antennas into the base station, tend to diminish when the $P^{\mathrm{tx}}/\sigma_n^2$ is large. The advantages of RZF over ZF and MRC can clearly be seen, as it consistently yields a performance similar to DPC (which is the optimal capacity achieving approach) for every individual transmission case.



(a)

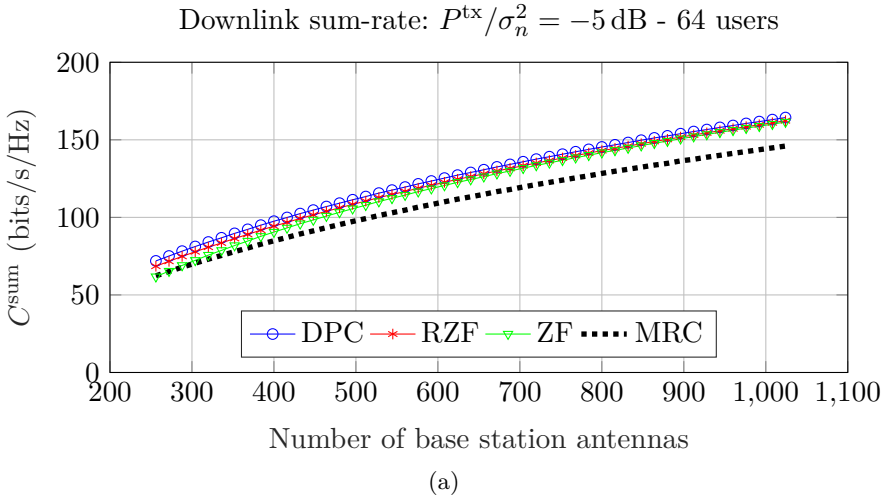


(b)

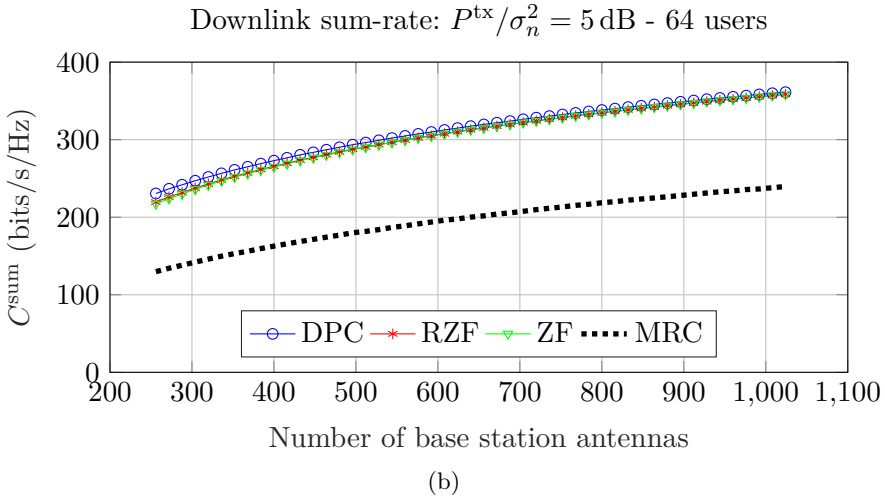


(c)

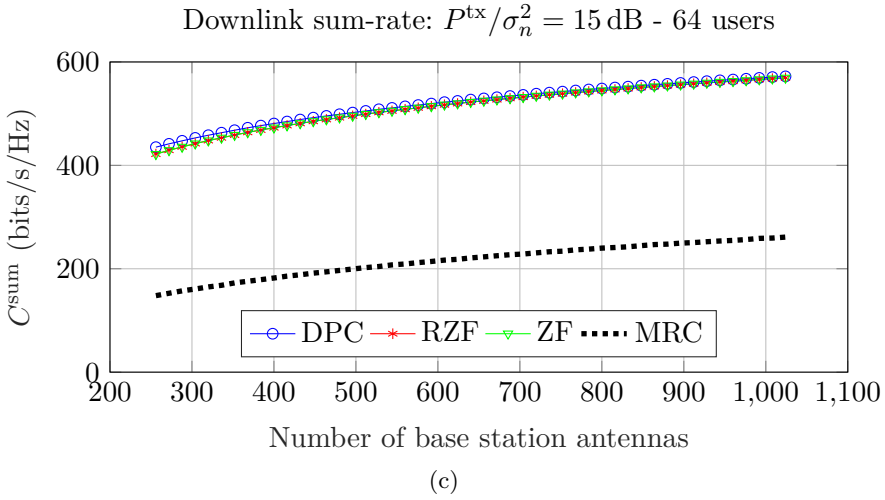
Figure 4.14: Downlink sum capacity variation with respect to the number of the base station antennas for 64 users and a small transmit array ($1 \leqslant M_T/M_R \leqslant 4$) with three different SNR settings -5 dB (a), 5 dB (b) and 15 dB (c).



(a)



(b)



(c)

Figure 4.15: Downlink sum capacity variation with respect to the number of the base station antennas for 64 users and a large transmit array ($4 \leqslant M_T/M_R \leqslant 16$) with three different SNR settings $-5\,\text{dB}$ (a), $5\,\text{dB}$ (b) and $15\,\text{dB}$ (c).

4.3.4 Practical Aspects

Channel Correlation

Channel correlation has a considerable impact on the capacity of a massive MIMO network, as the general goal is to reduce the correlation among the user channels (i.e., rows of the channel matrix $\mathbf{G}$) in order to achieve a better performance. Correlation is influenced by many factors, for instance, the geometry of the propagation environment or the properties of the antenna arrays. As the density of the scatterers in the environment and the spacing between the base station antennas increases, the correlation typically tends to decline (Biglieri *et al.*, 2007). The presence of line-of-sight components and small angular spread in the direction of arrival and departure (DoA and DoD) are also prominent factors which can yield a high correlation (Molisch, 2011). Although the positions of the users also affect the channel correlation significantly, they are typically assumed to be mobile and cannot be controlled. Therefore, relevant countermeasures against the correlation are implemented at the base station, i.e., distributing the array elements or employing directive antennas (Rusek *et al.*, 2013).

Identifying the factors, which affect the channel correlation, and estimating the capacity accordingly in a certain environment may not always be straightforward. In particular, degenerate channel phenomena may arise in rare cases where the receiving and transmitting arrays can yield very small correlation and the propagation environment consist of many scatterers, though the channel does not provide degrees of freedom and the network performance is poor. Such a situation might emerge when so-called *keyholes* exist in the environment, i.e., the link between the transmitter and receiver is established by means of only a few propagation paths (Almers *et al.*, 2003; Shin and Lee, 2003). Relay channels, indoor propagation through hallways, outdoor scenarios with rooftop propagation are some of the cases where keyhole channel phenomenon can be observed (Du and Swamy, 2010). Nevertheless, the keyhole effect is typically a rare phenomenon, and can be effectively dealt with in certain cases, e.g., by employing horizontally oriented arrays in rooftop diffraction scenarios, instead of vertical ones (Molisch, 2011; Chizhik *et al.*, 2002).

Channel Estimation

The acquisition of CSI is typically carried out by training sequences which are mutually orthogonal in some form, i.e., time or frequency. Acquisition of the CSI at the receiver site is usually straightforward (assuming downlink). On the other hand, at the transmitter site, either a feedback from the receiver based on the receiver's channel estimation, or a training process based on uplink pilot signals is usually required.

Note that perfect CSI is commonly assumed in system analysis (which is assumed in this study as well for the most part, unless stated otherwise), however, such an assumption generally depicts an ideal situation which may not hold for many real world scenarios. The errors in the CSI may stem from impairments in the acquisition process (e.g., estimation errors, limited feedback) as well as feedback delays (Clerckx and Oestges, 2013). In certain extreme cases, CSI may not be available at all, for instance, due to rapidly changing channel characteristics, where the base station would possess only the information about the statistical properties of the channel coefficients. As a result, the achievable sum-rates may diminish drastically and the conventional linear

receiving/precoding schemes, which have previously been presented, usually become ineffective (Heath and Lozano, 2019). In order to avoid such problems, proper measures, such as increasing the acquisition frequency, should be taken, hence, accurate channel information would always be available at the base station.

Computational Overhead

One of the major problems concerning the computational overhead in massive MIMO is the complexity of signal processing tasks, such as linear precoding and receiving, which would require the inversion of large matrices. The computational overhead would then grow, if the number of base station antennas is increased. Nevertheless, in (Björnson *et al.*, 2016) it is shown that the combined computational overhead for several signal processing tasks, including modulation/demodulation and channel estimation, is tolerable for most practically relevant systems (i.e., the number of base station antennas in the order of several hundreds) with modern hardware. Furthermore, various numerical matrix inversion methods, which are typically more efficient than naive algorithms, might be utilized for precoding and receiving tasks, thus, the overhead can significantly be reduced (Nagy *et al.*, 2018; Qin *et al.*, 2016). In the exceptional cases where the number of base station antennas is extremely large, utilizing the ZF/RZF methods can potentially be problematic due to computational overhead. However, a satisfactory performance in this regime can still be achieved by MRC precoding (which is much simpler and its complexity scales much more gradually compared to the other linear schemes) (Heath and Lozano, 2019).

Hardware Impairments

Another important consideration for such large-scale MIMO systems is the impact of non-ideal hardware characteristics, such as phase noise in oscillators, amplifier distortions or quantization noise. The impairments, which are additive and vary for each antenna/RF chain, typically vanish as the number of base station antennas grow, whereas multiplicative impairments, such as phase noise, may not decline but are not augmented either (Björnson and Larsson, 2015).

Hardware impairments require also certain measures to be taken regarding the CSI acquisition process as reciprocity of downlink and uplink channels may be violated. Here, the uplink pilot signal transmission by the users and the downlink data transmission by the base station are subject to distinct impairments due to the differences in the RF chains (e.g., deviations in manufacturing, voltage supply, temperature) (Luo and Zhang, 2016). In order to prevent such variations to hinder the reciprocity between uplink and downlink channels in TDD, a calibration/compensation scheme can be utilized (Jiang and Kaltenberger, 2018).

Chapter 5

Massive MIMO Downlink Performance Analysis by Unidirectional Ray-Tracing

The downlink performance analyses for 4 different massive MIMO scenarios are presented, where unidirectional ray-tracing simulations as well as statistical channel modeling techniques (which are applicable to the particular scenario), are utilized in order to characterize the channel. The analyses are based on an urban-like propagation environment and address the following scenarios:

1. The adverse effects of pilot contamination and uplink channel noise.
2. Characterization of channel aging effect with mobile users.
3. Downlink performance with frequency-selective fading.
4. Feasible transmission strategies with multi-antenna terminals.

The primary network performance metric in each scenario is the average user downlink data rate. Additional performance data regarding channel characteristics are also provided when applicable.

In the following, the urban scenario will be described first and a parametric convergence study, which clarifies how the essential simulation parameters (e.g., number of ray launches or number of interactions) were selected, will be presented. Afterwards, the downlink performance analyses for the aforementioned 4 problems will be given.

5.1 Urban Environment & Simulation Parameters

The performance analyses are based on an urban geometry, which consists of 16 buildings arranged as a 4×4 grid and cover a $160\,\text{m} \times 160\,\text{m}$ area. The buildings have different heights varying between 33 m-72 m and were assumed to have usable indoor spaces ($20\,\text{m} \times 20\,\text{m}$ area) where users can be located. The building walls have a thickness of 10 cm and do not include any windows. A dielectric constant of $\epsilon_r = 4$ (i.e., similar to concrete) was assumed everywhere in the scenario and an operating frequency of 2.45 GHz with a channel bandwidth of 10 MHz was used. A base station, which comprises 256 vertically oriented half-wavelength dipole antennas arranged as a cylindrical array, is located on top of one of the buildings near the center of the geometry (Note that a single base station implies a single-cell network here. A multi-cell scenario can accordingly be created by bringing several single-cell blocks together). The total transmission power from the base station was assumed as 1 Watt. The number of ray launches from each

transmitter antenna was set to 20 million.

As the problems involve both outdoor and indoor users, different interaction settings were utilized for these two cases, i.e., up to 1 diffraction for outdoor users and 4 transmissions for indoors users as well as 10 reflections for both groups were considered. The user terminal antennas are represented by reception spheres with a diameter of 1λ ($\approx 12\,\mathrm{cm}$), and are assumed to have isotropic patterns with a vertical polarization (with respect to the ground). The simulation parameters, which differ among the scenarios, are given in Table 5.1.

Table 5.1: Differences between certain simulation parameters for the urban massive MIMO simulations.

<table>
<tr><th>Scenario</th><th>Number of cells</th><th>Outdoor/Indoor users (per cell)</th><th>Channel frequency response</th><th>Number of user antennas</th></tr>
<tr><td>1) Pilot contamination</td><td>4</td><td>48/16</td><td rowspan="2">Flat</td><td rowspan="3">1</td></tr>
<tr><td>2) Channel aging</td><td rowspan="3">1</td><td>64/0</td></tr>
<tr><td>3) Selective fading</td><td rowspan="2">48/16</td><td>Frequency selective</td></tr>
<tr><td>4) Multi-antenna users</td><td>Flat</td><td>2</td></tr>
</table>

Parametric Convergence Study

In order to ensure that the selected ray launch and interaction (reflection, refraction, diffraction) parameters yield sufficiently accurate results, a convergence study has been carried out for the indoor and the outdoor simulation cases separately. In the first case, 25 users were placed at the first 10 floors of 16 buildings in the scenario (which yield 4000 user locations in total) whereas the latter case consists of a total of 1281 users, all at the ground level outside of the buildings. The users are considered to receive a signal from a single transmitting antenna (at the location of the base station) and a channel gain value is obtained for each. The average channel gain (among all users) is computed for different parameter configurations as the number of ray launches and number of interactions has been varied progressively, and the relative change in the channel gain from the previous step is calculated. In total 9 different configurations were investigated where the parameters for each case are given in Table 5.2 and 5.3 (for indoor and outdoor users, respectively).

The selected parameter configurations for the indoor and the outdoor simulations (20 million ray launches, 10 reflections, 4 transmissions/1 diffraction) yield a relative difference of smaller than 4×10^{-3} at the last step for both cases, as shown in Fig. 5.1 and 5.2, respectively.

Table 5.2: Simulation parameters for the indoor convergence analysis of urban massive MIMO scenario.

Cases	Number of ray launches	Number of reflections	Number of transmissions
Case 1	4,000,000	2	2
Case 2	6,000,000	3	2
Case 3	8,000,000	4	2
Case 4	10,000,000	5	3
Case 5	12,000,000	6	3
Case 6	14,000,000	7	4
Case 7	16,000,000	8	4
Case 8	18,000,000	9	4
Case 9	20,000,000	10	4

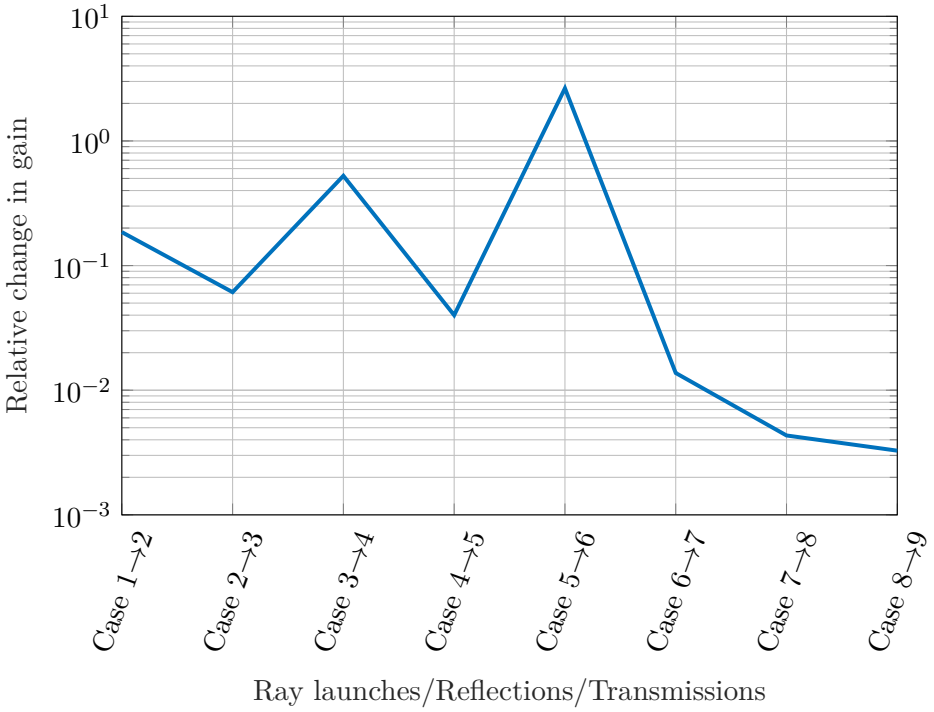


Figure 5.1: Variation of the relative change in channel gain for indoor simulation.

Table 5.3: Simulation parameters for the outdoor convergence analysis of urban massive MIMO scenario.

Cases	Number of ray launches	Number of reflections	Number of diffractions
Case 1	4,000,000	2	0
Case 2	6,000,000	3	0
Case 3	8,000,000	4	0
Case 4	10,000,000	5	0
Case 5	12,000,000	6	1
Case 6	14,000,000	7	1
Case 7	16,000,000	8	1
Case 8	18,000,000	9	1
Case 9	20,000,000	10	1

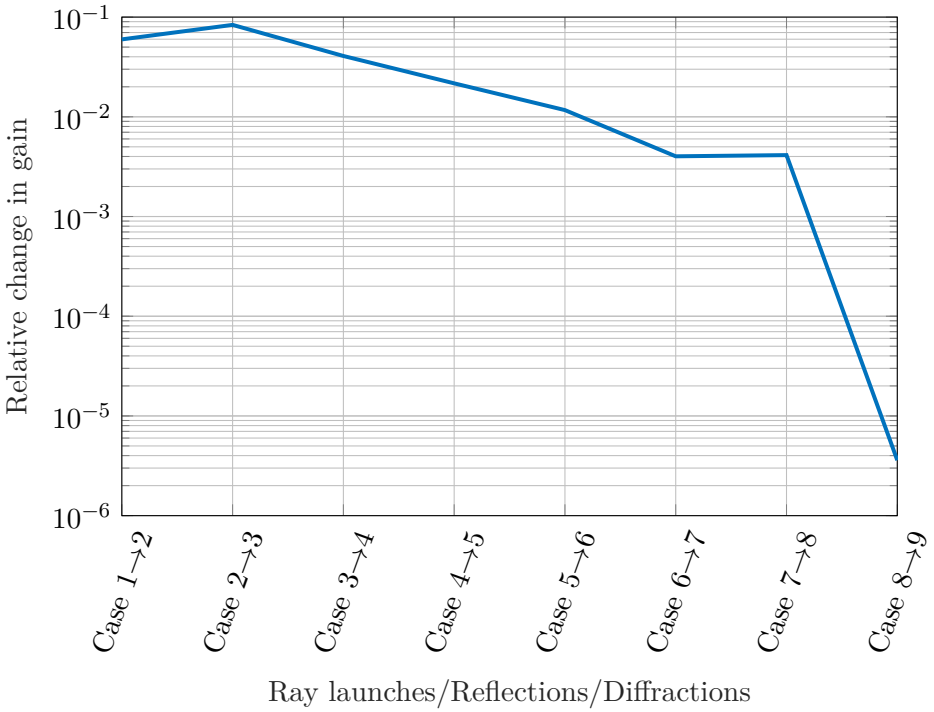


Figure 5.2: Variation of the relative change in channel gain for outdoor simulation.

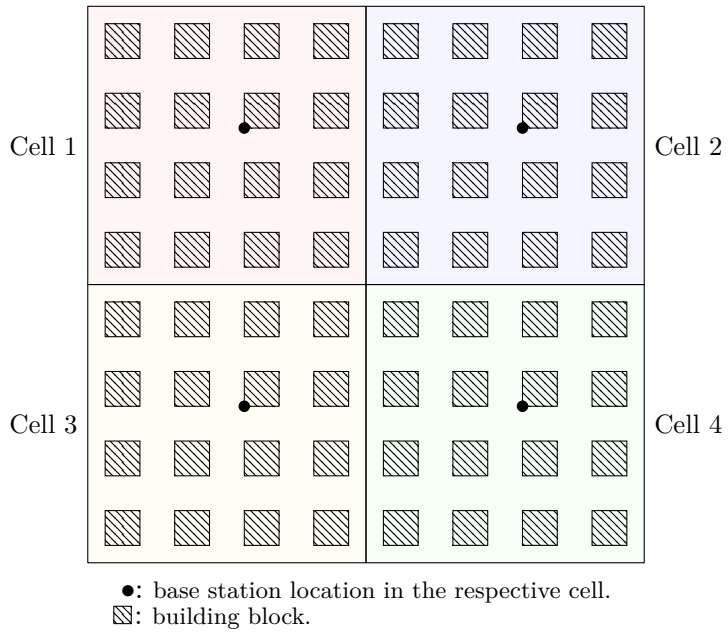


Figure 5.3: Illustration of 2×2 multi-cell network with 4×4 building grid in each cell.

5.2 Characterizing the Effects of Pilot Contamination on Downlink Data Rate

One of the fundamental requirements to achieve the optimal performance in massive MIMO is to carry out the precoding process in a precise manner. This is generally dependent on the accuracy of the CSI, which is acquired by the base station via uplink pilot signals (transmitted by the users). Although it is common practice to assume that the acquisition process is perfect, this is typically not achievable due to channel noise and inter-cell interference (Lu *et al.*, 2014b; Elijah *et al.*, 2016). In particular, the uplink pilot signals, which are transmitted by the users to the base station, might interfere with each other if the same time-frequency resources are utilized by multiple users. This phenomenon is commonly known as pilot contamination. In general, the users, which share the same resources for pilot signal transmission, should ideally be located in distinct cells. Nevertheless, the impact of the interference might still be significant, if, for instance, the users are at the neighboring cells. Even if the inter-cell interference may be eliminated in the network, channel noise may not completely be avoided. Consequently, the data rate performance declines due to such imperfections in the CSI acquisition process (Shen *et al.*, 2015).

In order to demonstrate the impact of inter-cell interference and channel noise during CSI acquisition on the downlink performance, an urban scenario consisting of 4 cells has been analyzed by ray-tracing simulations (Taygur and Eibert, 2018b). The cells were arranged in a 2×2 grid, which are composed of the previously described single-cell urban geometry, as shown in Fig. 5.3. Each cell comprises 16 outdoor and 48 indoor user terminals. The analysis has been repeated with the WINNER II channel model (Kyosti *et al.*, 2007), where a second set of channel data is obtained, in order to compare the ray-tracing results.

5.2.1 Methodology

The channel coefficients for ray-tracing and WINNER II simulations can be given by

$$\underline{\text{Ray-Tracing}} \qquad \underline{\text{WINNER II}}$$

$$\mathbf{G}_{a,b_{(u,v)}} = \frac{V^{\text{oc}}_{a(u)}}{V^{\text{gen}}_{b(v)}}, \qquad \mathbf{G}_{a,b_{(u,v)}} = \xi G\left(d_{a(u),b(v)}\right),$$

$$G(x) = 10^{-(22.7\log_{10}(x)+34.8+\chi)0.05},$$

$$\xi \sim \mathcal{CN}(0,1),\ \chi \sim \mathcal{N}(0,4), \tag{5.1}$$

where $\mathbf{G}_{a,b_{(u,v)}}$ denotes the channel between the uth user in the ath cell ($u = 1, \ldots, 64$, $a = 1, \ldots, 4$) and the vth base station antenna in the bth cell ($v = 1, \ldots, 256$, $b = 1, \ldots, 4$), $V^{\text{oc}}_{a(u)}$ is the induced open-circuit voltage at the terminals of the uth user antenna in the ath cell, $V^{\text{gen}}_{b(v)}$ is the generator voltage of the vth base station antenna in the bth cell, χ is the slow-fading coefficient, $G(x)$ is path-gain function depending on the distance, and $d_{a(u),b(v)}$ denotes the distance between the uth user in the ath cell and vth base station antenna in the bth cell.

5.2.2 Linear Precoding with Inaccurate Channel State Information

Assuming a least-squares estimator is utilized, the channel estimate of the base station in the bth cell ($\tilde{\mathbf{G}}^b$) can be written as

$$\tilde{\mathbf{G}}_b = \sqrt{P^{\text{pilot}}}\left[\mathbf{G}_{b,b} + \left(\sum_{a=1,a\neq b}^{4} \mathbf{G}_{a,b}\right)\right] + \mathbf{n}_b,$$

$$\tilde{\mathbf{G}}_b,\ \mathbf{G}_{b,b},\ \mathbf{G}_{a,b},\ \mathbf{n}_b \in \mathbb{C}^{64\times 256},$$

$$\mathbf{n}_b \sim \mathcal{CN}(0,\sigma_n^2),\ \left(\sigma_n^2\right)_{\text{dB}} = -204 + 10\log_{10}(W), \tag{5.2}$$

where W is the channel bandwidth (which was specified as 10 MHz previously), P^{pilot} is the transmission power of the uplink pilot signals (assumed to be uniform among every user), $\mathbf{n}_b$ denotes the channel noise vector, $\mathbf{G}_{b,b}$ is the channel matrix between the base station and users of the bth cell, $\mathbf{G}_{a,b}$ is the channel matrix between the base station of the bth cell and the users of the ath cell, which represents the interference from other cells.

Note that the same orthogonal pilot set is utilized in each cell according to Eq. (5.2). It is possible to mitigate the adverse effects of pilot contamination by allocating mutually orthogonal pilot signals for users in different cells, as mentioned previously, and the number of distinct orthogonal pilot signal sets in the network can be represented by a pilot reuse factor (PRF). For a reuse factor of 1, 2, and 4, the channel estimate of the base station ($\tilde{\mathbf{G}}_b$) can then be given as

$$\text{PRF} = 1 \rightarrow \tilde{\mathbf{G}}_b = \sqrt{P^{\text{pilot}}}\left[\mathbf{G}_{b,b} + \left(\sum_{a=1,a\neq b}^{4} \mathbf{G}_{a,b}\right)\right] + \mathbf{n}_b,\ b = 1,\ 2,\ 3,\ 4,$$

$$\text{PRF} = 2 \rightarrow \tilde{\mathbf{G}}_b = \sqrt{P^{\text{pilot}}}\left(\mathbf{G}_{b,b} + \mathbf{G}_{a,b}\right) + \mathbf{n}_b,\ (a,\ b) = (1,\ 4),\ (2,\ 3),\ (3,\ 2),\ (4,\ 1),$$

$$\text{PRF} = 4 \rightarrow \tilde{\mathbf{G}}_b = \sqrt{P^{\text{pilot}}}\,\mathbf{G}_{b,b} + \mathbf{n}_b,\ b = 1,\ 2,\ 3,\ 4. \tag{5.3}$$

Assuming that the base station employs the regularized zero-forcing technique, the precoding matrix for the base station in the bth cell can be written as

$$\mathbf{A}_b = (\tilde{\mathbf{G}}_b)^H \left(\tilde{\mathbf{G}}_b(\tilde{\mathbf{G}}_b)^H + \alpha\mathbf{I}\right)^{-1}, \ \alpha > 0, \tag{5.4}$$

where α is a regularization term. Using the precoding matrix, it is possible to obtain the SINR for each user in the network. The SINR expression consists of two distinct interference terms, which can be described as:

1. *Intra-cellular Interference:* The interference which arises within the cell, as the base station does not fully suppress the interfering signals of different users.
2. *Inter-cellular Interference:* Caused by non-orthogonal uplink pilot signals which are received by the base stations in different cells. Consequently, the base stations then try to serve the users in different cells during the downlink data transmission phase.

The SINR expression for the uth user in the bth cell can be given by

$$\mathrm{SINR}_{b(u)} = \frac{P_b^{\mathrm{tx}} |\mathbf{G}_{b,b_{(u,:)}} \mathbf{A}_{b(:,u)}|^2}{I_{b(u)}^{\mathrm{intra}} + I_{b(u)}^{\mathrm{inter}} + \sigma_n^2},$$

$$I_{b(u)}^{\mathrm{intra}} = P_b^{\mathrm{tx}} \left(\sum_{\substack{v=1 \\ v \neq u}}^{64} |\mathbf{G}_{b,b_{(u,:)}} \mathbf{A}_{b(:,v)}|^2 \right), \ I_{b(u)}^{\mathrm{inter}} = \sum_{\substack{a=1 \\ a \neq b}}^{4} P_a^{\mathrm{tx}} \sum_{v=1}^{64} |\mathbf{G}_{b,a_{(u,:)}} \mathbf{A}_{a(:,v)}|^2, \tag{5.5}$$

where P_b^{tx} is the transmission power from the base station in the bth cell, $I_{b(u)}^{\mathrm{intra}}$ and $I_{b(u)}^{\mathrm{inter}}$ denote intra-cell and inter-cell interference terms, which apply to the uth user in bth cell, respectively. The average downlink data rate for all the users in the bth cell can be written as

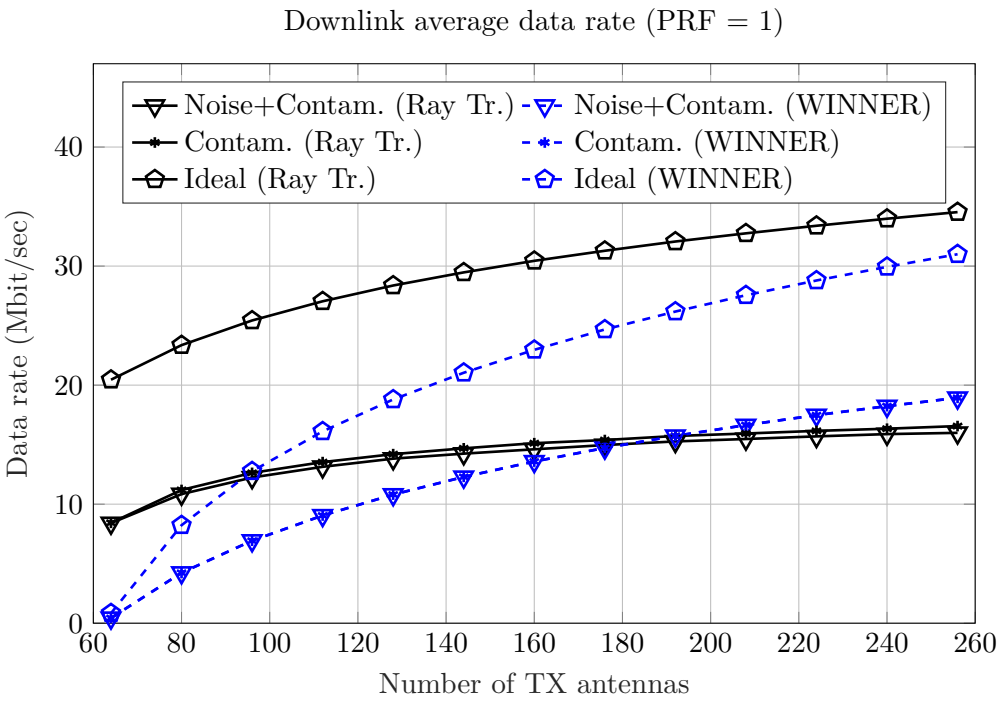
$$R_b^{\mathrm{avg}} = \frac{W}{64} \sum_{u=1}^{64} \log_2 \left(1 + \mathrm{SINR}_{b(u)}\right). \tag{5.6}$$

5.2.3 Downlink Data Rate Simulation Results

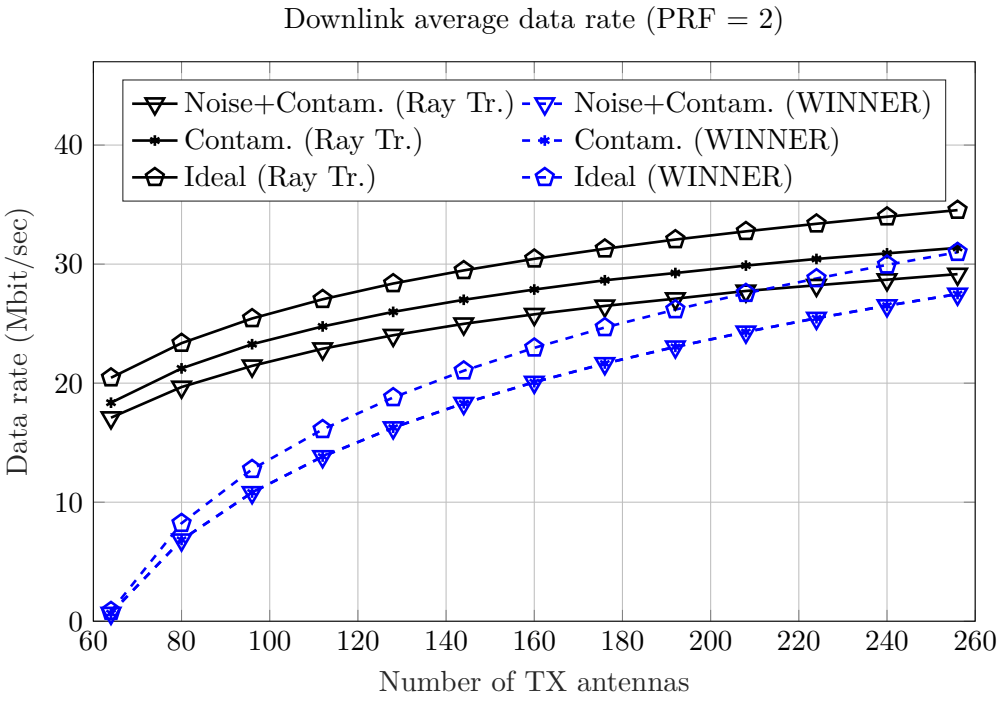
A downlink data rate analysis was performed for three different pilot reuse schemes with PRF = 1, 2, 4 by means of ray-tracing and the WINNER II channel model. The results are compared for two cases, first, no pilot contamination or uplink channel noise are present (which yields perfect CSI), second, pilot contamination is present but uplink channel noise is not. The number of antennas at the base station is varied from 64 to 256 in order to characterize the influence of the size of transmit array. The results are shown in Fig. 5.4.

It can be observed that the use of non-orthogonal pilot signals may cause a considerable drop in downlink data rate performance, especially for reuse factor PRF = 1. On the other hand, the effect of uplink channel noise is relatively limited and does not affect the performance in any significant way for both ray-tracing and the WINNER II model. This indicates that the primary limiting factor of downlink performance is interference (both inter-cellular and intra-cellular) rather than uplink/downlink channel noise. It can be noticed that the variation of data rate for ray-tracing and the WINNER II model is

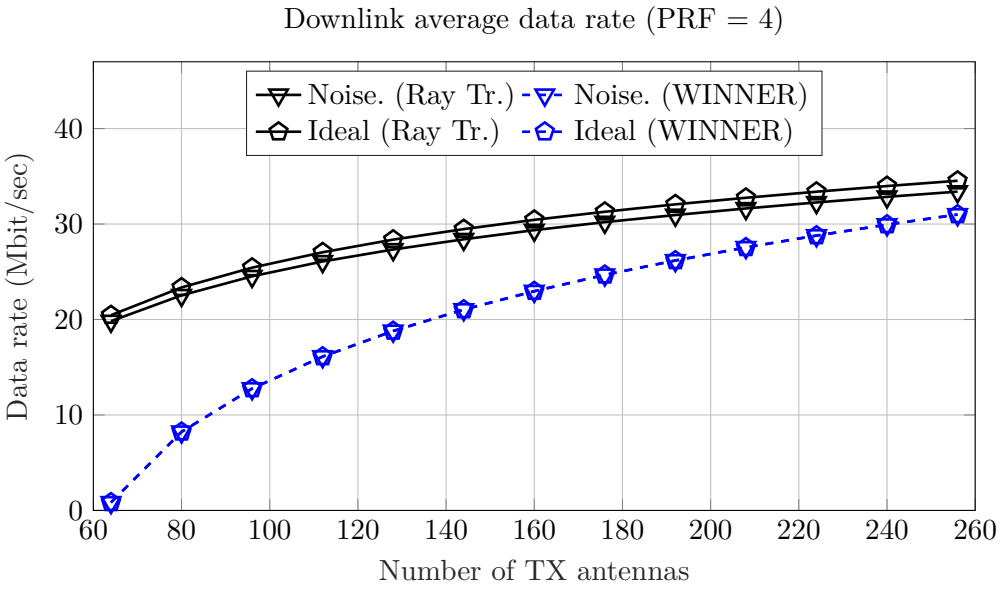
distinctive. In particular, the WINNER II model predicts significantly lower data rates compared to the ray-tracing simulations for a small number of base station antennas for every reuse case. As the number of antennas approach the upper limit of 256, both models produce similar results.



(a)



(b)



(c)

Figure 5.4: Downlink average data rate under pilot contamination for reuse factor PRF = 1 (a), 2 (b) and 4 (c).

5.2.4 Characterizing the Properties of Contaminating Channels

An important fact regarding pilot contamination is that the interfering signals, which are captured by a particular base station, come from the users in different cells. Since these out-of-cell users are physically isolated compared to the users in the cell (i.e., the users which are actually served by the base station), it can be inferred that the channel properties for these two sets of users can be significantly different. Thus, the contaminating signals and the actual pilot can be distinguished by using certain channel properties. In the following, three different parameters, namely, signal delay, channel gain and channel orthogonality will be characterized and comparisons between in-cell and out-of-cell users will be given for the considered scenario.

Signal Delay

The signal delay is characterized only for the ray-tracing model and computed according to the average path length of the rays for each individual channel. Note that only the channels, which yield a received signal power above the noise threshold (calculated as -134 dBm according to Eq. (5.2)), are considered. Thus, the delay expressions for in-cell users and out-of cell users (i.e., the contaminating users) can be given as

$$
\begin{aligned}
T^{\text{in-cell}} &: \frac{\mathcal{D}_{b(u),b(v)}}{\nu_0},\ ||\mathbf{G}_{b,b_{(u,v)}}||^2 > \sigma_n^2, \\
T^{\text{contam}} &: \frac{\mathcal{D}_{a(u),b(v)}}{\nu_0},\ ||\mathbf{G}_{a,b_{(u,v)}}||^2 > \sigma_n^2,\ a \neq b, \\
&\mathbf{G}_{b,b_{(u,v)}},\ \mathbf{G}_{a,b_{(u,v)}} \in \mathbb{C},
\end{aligned}
\tag{5.7}
$$

where $\mathcal{D}_{a(u),b(v)}$ is the average ray-path length between the uth user in the ath cell and the vth transmitter antenna in the bth cell and ν_0 is the speed of light in vacuum.

Channel Gain

The channel gain expressions can be written as

$$\begin{aligned} G^{\text{in-cell}} &: ||\mathbf{G}_{b,b_{(u,v)}}||^2,\ ||\mathbf{G}_{b,b_{(u,v)}}||^2 > \sigma_n^2, \\ G^{\text{contam}} &: ||\mathbf{G}_{a,b_{(u,v)}}||^2,\ ||\mathbf{G}_{a,b_{(u,v)}}||^2 > \sigma_n^2,\ a \neq b, \\ &\mathbf{G}_{b,b_{(u,v)}},\ \mathbf{G}_{a,b_{(u,v)}} \in \mathbb{C}. \end{aligned} \tag{5.8}$$

The calculation was performed for both ray-tracing and the WINNER II model. Note that the noise threshold criterion from the signal delay computation is applied here as well.

Channel Orthogonality

The channel orthogonality parameter represents a measure of similarity where the inner product of two channel vectors (after normalization), which are associated with two users in different cells and share the same pilot, are calculated. In particular, the orthogonality parameter $\mathcal{U}_{a(u),b(v)}$ is expressed as

$$\begin{aligned} \mathbf{v}_{b(u),b} = \frac{\mathbf{G}_{b,b_{(u,:)}}}{||\mathbf{G}_{b,b_{(u,:)}}||},\ \mathbf{v}_{a(v),b} &= \frac{\mathbf{G}_{a,b_{(v,:)}}}{||\mathbf{G}_{a,b_{(v,:)}}||}, \\ \mathbf{G}_{b,b_{(u,:)}},\ \mathbf{G}_{a,b_{(v,:)}} &\in \mathbb{C}^{1\times 256}, \\ \mathcal{U}_{a(u),b(v)} &: \left\langle \mathbf{v}_{b(u),b}, \mathbf{v}_{a(v),b} \right\rangle. \end{aligned} \tag{5.9}$$

It should be noted that the cell indices a and b in each parameter calculation indicate two distinct cells where the same pilot signal set is utilized. Thus, the possible values for a and b depend on the reuse factor (see Eq. (5.3)). The case PRF = 4 does not apply here though, since no pilot contamination exists in that case.

Numerical Results

The results in Fig. 5.5 show that in-cell and contaminating user channels can be distinguished more effectively for PRF = 2 than for PRF = 1, as the difference between the CDF curves of in-cell and contaminating user channels is generally larger for PRF = 2 with each channel parameter. The signal delay characteristics are relatively similar for PRF = 1, 2, though the curve for in-cell users is noticeably different, hence it is possible to differentiate the contaminating signals to a certain extent. It can be noticed that the WINNER model overestimates the channel gain for both in-cell and contaminating users. Relatively high channel gains also explains why the uplink noise had almost no adverse effects on the data rate according to the WINNER model in Fig. 5.4. Despite stronger channel gain, the data rate performance in WINNER model remains poorer since the downlink interference is also estimated considerably stronger.

On the other hand, the channel orthogonality curve for ray-tracing reaches the maximum much slower than for the WINNER model, which indicates that the disparity

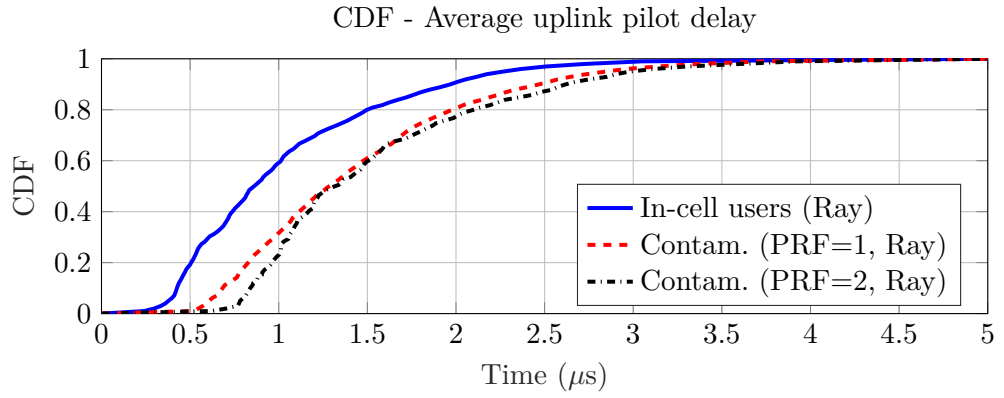
between the in-cell and contaminating user channels is generally smaller according to the ray-tracing simulation. However, applying different pilot reuse schemes do not affect the channel similarity according to the WINNER model, as the curves almost completely overlap. This can be attributed to the fact that the channel matrices in the WINNER model were created without precise geometrical considerations, i.e., the channel vectors of the contaminating users in different cells do not vary significantly although certain base station antennas may in fact have a relatively stronger link with certain users due to the geometrical features of the scenario. It should be noted that the presented parameters might be correlated to a certain extent, i.e., the channels, which yield a lower gain, may introduce a larger signal delay with higher probability.

5.3 User Mobility and Channel Aging

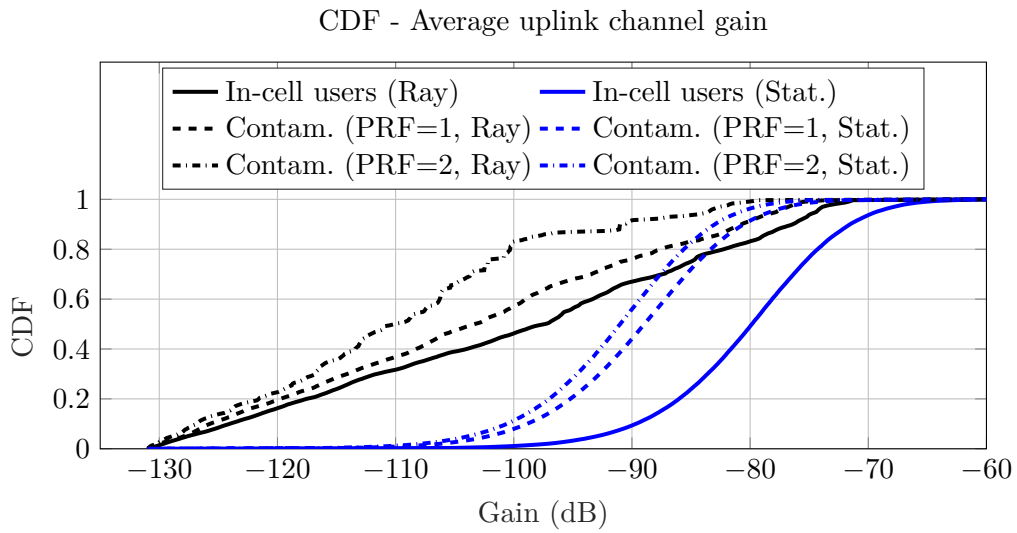
Massive MIMO precoding typically relies on the instantaneous CSI, which comes from the user terminals via uplink pilot signals. As long as the channel remains relatively constant (i.e., the channel impulse response does not vary) within a certain period of time, the precoding process can effectively be carried out without acquiring CSI repeatedly. In case the channel changes in time though, the base station should update the CSI accordingly, in order to avoid the so-called channel aging, where the actual channel is not consistent with the CSI that the base station possesses. Unless the base station updates the CSI sufficiently often, the data rate performance generally deteriorates due to sub-optimal beamforming (Papazafeiropoulos, 2017). Although it seems reasonable to acquire CSI at the base station as often as possible, the CSI acquisition typically creates an overhead, which limits the time for actual data transmission.

In order to better illustrate the problem, let us consider a TDD transmission frame which consists of CSI acquisition as well as uplink/downlink data transfer. The base station is assumed to utilize the same CSI for uplink and downlink data transmission until a new acquisition process is carried out. The duration of CSI acquisition is typically determined by the length of the pilot sequences transmitted by the users (since the pilot signals are mutually orthogonal, the length of the sequence is ultimately proportional to the number of users served by the base station). Thus, increasing the acquisition frequency limits the share of uplink and downlink data transmission within a certain period of time, as shown in Fig. 5.6. Therefore, the frequency of CSI acquisition should be selected carefully in order to avoid the adverse effects of channel aging as well as excessive acquisition overhead (Taygur and Eibert, 2018a). Although there are several factors which cause the channel aging phenomenon, movements of the user terminals (i.e., user mobility) can be considered as the most prominent. In general, the faster the users move, the quicker the CSI at the base station becomes obsolete and channel aging emerges.

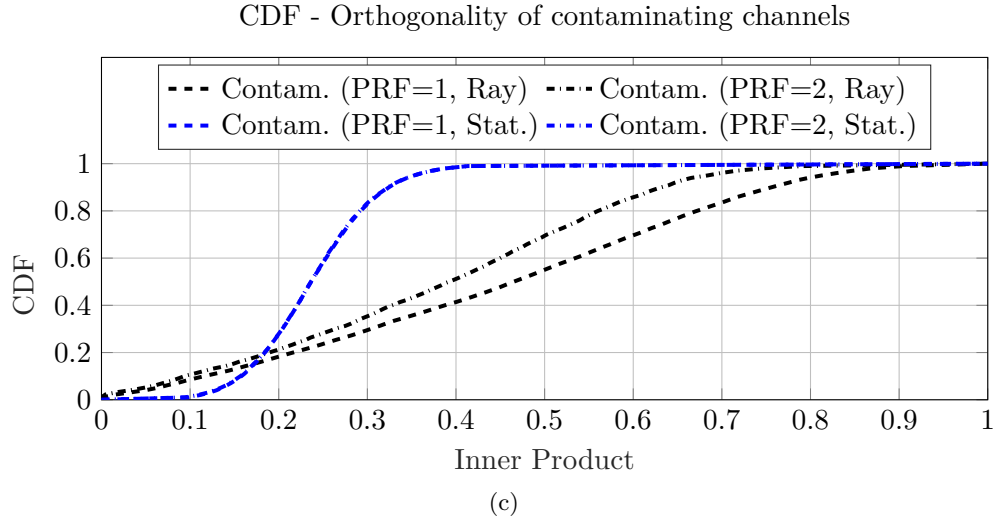
In order to demonstrate the variation of the data rate due to user mobility, a single-cell urban scenario is simulated by ray-tracing where the users were assumed to be vehicles. The ray-tracing simulation was utilized to characterize the channel at every $100\,\mu$s for a 10 ms time span. The base station has the perfect CSI at the beginning of this time frame. Afterwards, the movements of the users yield discrepancies between the actual channel and the CSI that the base station possesses. For each time instance, the average downlink data rate is computed. The analysis is repeated with a statistical channel aging model, namely the Jakes' model (Truong and Heath, 2013), and the differences between the two approaches are compared (Taygur and Eibert, 2018a).



(a)

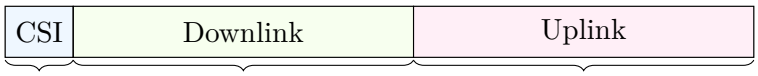


(b)



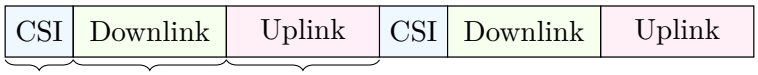
(c)

Figure 5.5: Properties of the contaminating channels with respect to signal delay (a), channel gain (b) and channel orthogonality (c).



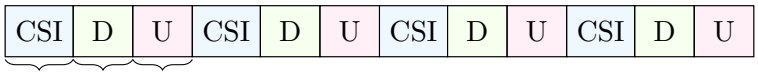
$t^{\mathrm{dl}} + t^{\mathrm{ul}} = 10\, t^{\mathrm{CSI}} \rightarrow 9.09\%$ overhead.

(a) Acquisition frequency of 1x



$t^{\mathrm{dl}} + t^{\mathrm{ul}} = 4.5\, t^{\mathrm{CSI}} \rightarrow 18.18\%$ overhead.

(b) Acquisition frequency of 2x



$t^{\mathrm{dl}} + t^{\mathrm{ul}} = 1.75\, t^{\mathrm{CSI}} \rightarrow 36.36\%$ overhead.

(c) Acquisition frequency of 4x

Figure 5.6: CSI acquisition overhead for different acquisition frequencies, 1x (a), 2x (b) and 4x (c).

5.3.1 Methodology

The channel matrix at the time instance of $t = 0$ is defined as

$$\mathbf{G}_{0(u,v)} = \frac{V_{0,u}^{\mathrm{oc}}}{V_{0,v}^{\mathrm{gen}}},\ \mathbf{G}_0 \in \mathbb{C}^{64\times 256}, \tag{5.10}$$

where $\mathbf{G}_{0(u,v)}$ denotes the channel between the uth user and the vth antenna of the base station, $V_{0,u}^{\mathrm{oc}}$ is the induced open-circuit voltage at the terminals of the uth user antenna and $V_{0,v}^{\mathrm{gen}}$ is the generator voltage of the vth base station antenna. The CSI, that the base station has, is also equal to $\mathbf{G}_0$. Note that the initial conditions for the ray-tracing and Jakes' model are identical, i.e., the same channel matrix, which is obtained by ray-tracing simulation is used. The channel matrices for other time instances ($\mathbf{G}_t,\ t > 0$) can be expressed as

$$\underline{\text{Ray-Tracing}} \qquad \underline{\text{Jakes' Model}}$$

$$\mathbf{G}_{t(u,v)} = \frac{V_{t,u}^{\mathrm{oc}}}{V_{t,v}^{\mathrm{gen}}}, \qquad \mathbf{G}_{t(u,v)} = p\,\mathbf{G}_{0(u,v)} + \xi\sqrt{1-p^2}|\mathbf{G}_{0(u,v)}|,$$

$$p = J_0(2\pi F_D t),$$

$$\xi = \mathcal{CN}(0,1), \tag{5.11}$$

where $J_0(.)$ is the Bessel function of first-kind and F_D denotes the absolute maximum Doppler frequency, which occurs when the relative velocity vector and the propagation direction vector of the incident wave are aligned. It should be noted that these two vectors are unlikely to be aligned in this considered urban scenario since the base station is located on top of a building and is much higher than the vehicles on the ground.

Therefore, it is reasonable to scale down F_D for this scenario, in order to represent the propagation conditions more accurately. Hence, F_D can be given as

$$F_D = \frac{v_r\, f_c}{\nu_0} \frac{1}{\kappa}, \tag{5.12}$$

where κ is the scaling factor. All the 64 users (i.e., the vehicles) are assumed to be moving with the same speed, where three different speed settings were considered i.e., $v_r = 20,\ 40,\ 60\,\mathrm{km/h}$.

5.3.2 Beamforming with Obsolete CSI

The regularized zero-forcing method is utilized by the base station for precoding. The precoding matrix is obtained according to the channel matrix at time instance $t = 0$ by

$$\mathbf{A} = (\mathbf{G}_0)^H \left(\mathbf{G}_0 \, (\mathbf{G}_0)^H + \alpha \mathbf{I} \right)^{-1}. \tag{5.13}$$

Note that the CSI acquisition is assumed to be perfect in this case, i.e., the channel estimation of the base station is identical to the actual channel at $t = 0$. The SINR expression for the uth user in the cell and the average downlink data rate in the cell at time instance t can be given as

$$\mathrm{SINR}_{t,u} = \frac{P^{\mathrm{tx}} |\mathbf{G}_{t(u,:)} \mathbf{A}_{(:,u)}|^2}{P^{\mathrm{tx}} \left[\sum\limits_{v=1, v\neq u}^{64} |\mathbf{G}_{t(u,:)} \mathbf{A}_{(:,v)}|^2 \right] + \sigma_n^2},$$

$$R_t^{\mathrm{avg}} = \frac{W}{64} \sum_{u=1}^{64} \log_2 \left(1 + \mathrm{SINR}_{t,u} \right). \tag{5.14}$$

5.3.3 Numerical Results

Influence of Doppler Scale Factor

The Doppler parameter in Jakes' model is scaled by factor κ in order to depict the propagation conditions of the scenario more accurately, as stated previously. Three different values for a scaling factor $\kappa = 1,\ 4,\ 8$ are utilized as the variation of the average downlink data rate and channel aging effect are analyzed for a 10 ms time interval. A comparison to the ray-tracing method was made where the scenario consists of 256 base station antennas and the vehicles were assumed to move with 20 km/h. The results are shown in Fig. 5.7.

The results demonstrate that the variation of the data rate for $\kappa = 8$ is in good agreement with the ray-tracing simulation. For $\kappa = 1$ and 4 on the other hand, a significant difference in data rate exists compared to the ray-tracing results. Based on these results, a more comprehensive data rate performance analysis for ray-tracing and Jakes' model with $\kappa = 8$ is performed.

Downlink Data Rate Performance

The channel aging effect is analyzed for ray-tracing and Jakes' model with three different user speed settings, 20, 40 and 60 km/h. The analysis was repeated for three different

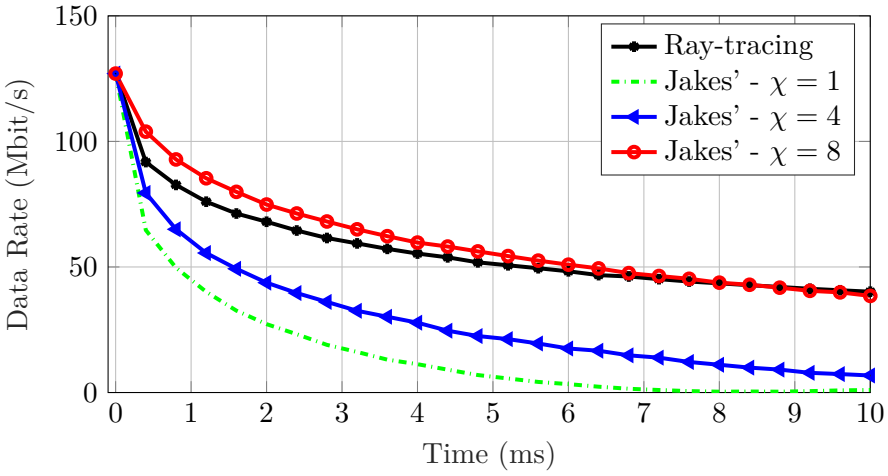
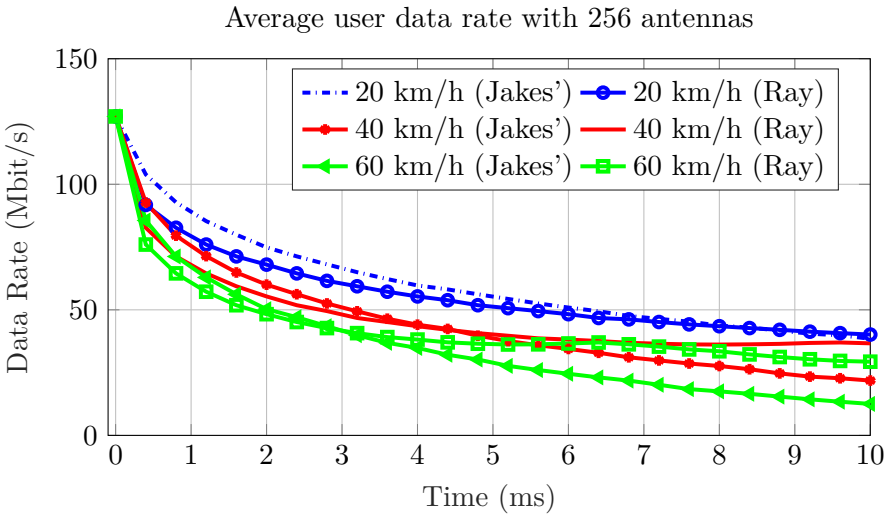


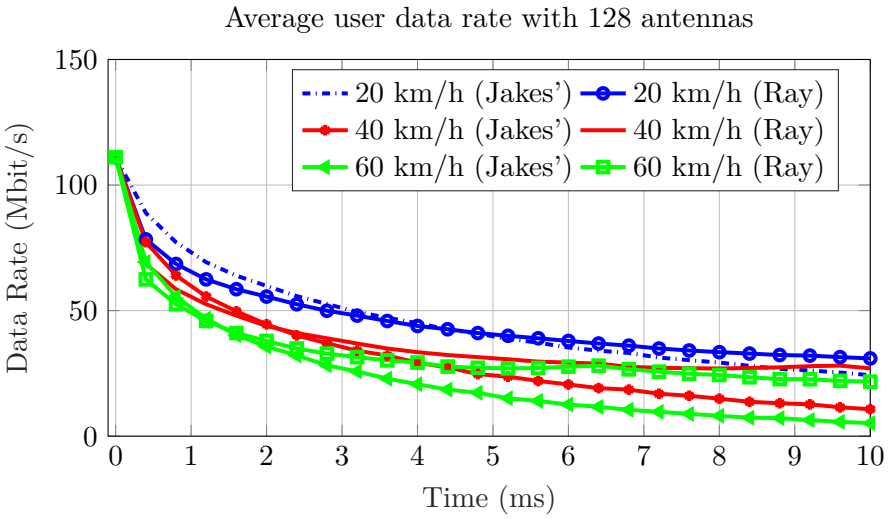
Figure 5.7: Characterization of the effects of spread factor parameter on Jakes' model and comparison to ray-tracing simulation.

base station configurations where the number of antennas at the base station was 64, 128 and 256. The results are shown in Fig. 5.8.

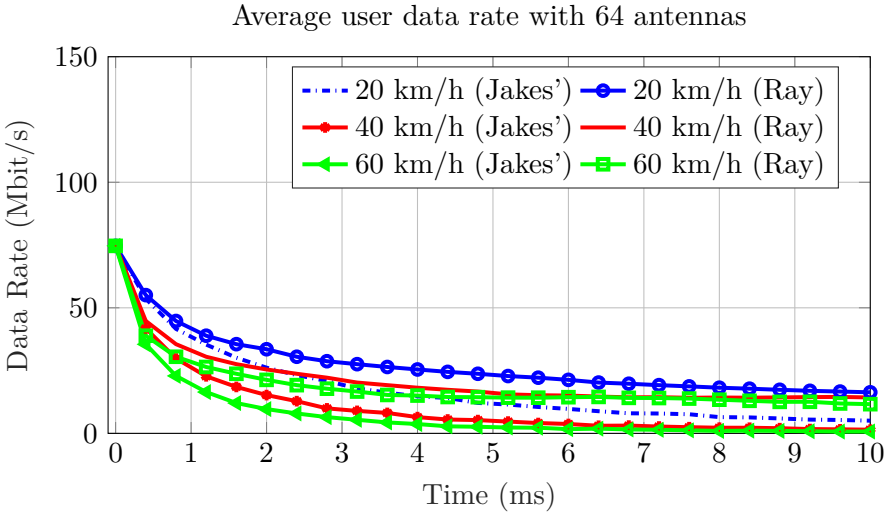
The results indicate that the user movement yields drastic changes in data rate, especially within the first 2 ms. The average downlink data rate after 10 ms can be drastically smaller than the rate at the beginning (i.e., at $t = 0$ where the CSI is acquired). In general, the faster the users move, the more the data rate performance deteriorates. There is a good agreement between Jakes' model and ray-tracing results, though the differences between two approaches become more noticeable around 10 milliseconds. The benefits of utilizing a large number of base station antennas is evident, as the data rate is generally better when the base station consists of more antennas. Nevertheless, a large number of base station antennas may not effectively eliminate the drastic deterioration in performance due to channel aging.



(a)



(b)



(c)

Figure 5.8: Data rate variation with obsolete CSI for 256 (a), 128 (b), and 64 (c) base station antennas.

5.4 Characterization of Frequency-Selectivity in Massive MIMO Channels

Frequency-selectivity is a commonly encountered phenomenon in wireless networks where multi-paths occur frequently, for instance in dense urban environments. The channel frequency response is usually described as non-flat, and (multiple) echoes of the transmitted signal reach the receiver with a certain delay. The frequency-selectivity of a channel is generally quantified by the coherence bandwidth, and it is defined as the band over which the frequency response of the channel remains relatively unchanged, whereas the delay spread can be described as the time difference between the arrival of the first and the last echo to the receiver (Tse and Viswanath, 2005).

An important problem observed in frequency selective channels is the so-called inter-symbol interference (ISI), where the delayed echoes of a transmitted symbol interfere with the subsequent symbols, thereby yielding a distortion in the signal. Consequently, data transmission might be impaired if ISI is not properly dealt with. There are two common techniques to mitigate ISI (Taygur and Eibert, 2019a):

1. *Single-Carrier Equalization:* It can be realized by a frequency-domain equalizing filter, commonly at the receiver site, where CSI should accordingly be acquired in order to adjust the filter coefficients (Pancaldi *et al.*, 2008).
2. *Orthogonal Frequency Division Multiplexing:* The data transmission is performed via multiple sub-carriers, which occupy a small portion of the allocated channel bandwidth. The main goal is to set the sub-carrier bandwidth to a level where the frequency response of the sub-channel is constant. In other words, the sub-carrier bandwidth should be smaller than the coherence bandwidth (Goldsmith, 2005).

Here, an equalization based solution will be considered for mitigating the ISI since the performance of an orthogonal frequency-division multiplexing (OFDM) system is typically sensitive to the impairments in the RF-chain, especially when the number of the chains is large (Li and Torlak, 2018; Torres *et al.*, 2014). Therefore, the use of single-carrier frequency-domain equalization can be advantageous in massive MIMO, considering that the base station comprises of a large number of RF-chains.

Linear equalizers, such as Minimum Mean Square Error (MMSE) or ZF equalizer, are commonly utilized to carry out frequency-domain equalization at the receivers. An MMSE equalizer is generally superior to ZF in terms of preventing noise enhancement, nevertheless, both approaches suffer from this problem to a certain extent. In order to better clarify the situation, let us consider a frequency dependent scalar channel (i.e., a SISO system) represented by $H(f)$. The ZF and MMSE equalizers ($L(f)$), which can be applied at the receiver site, can then be expressed as

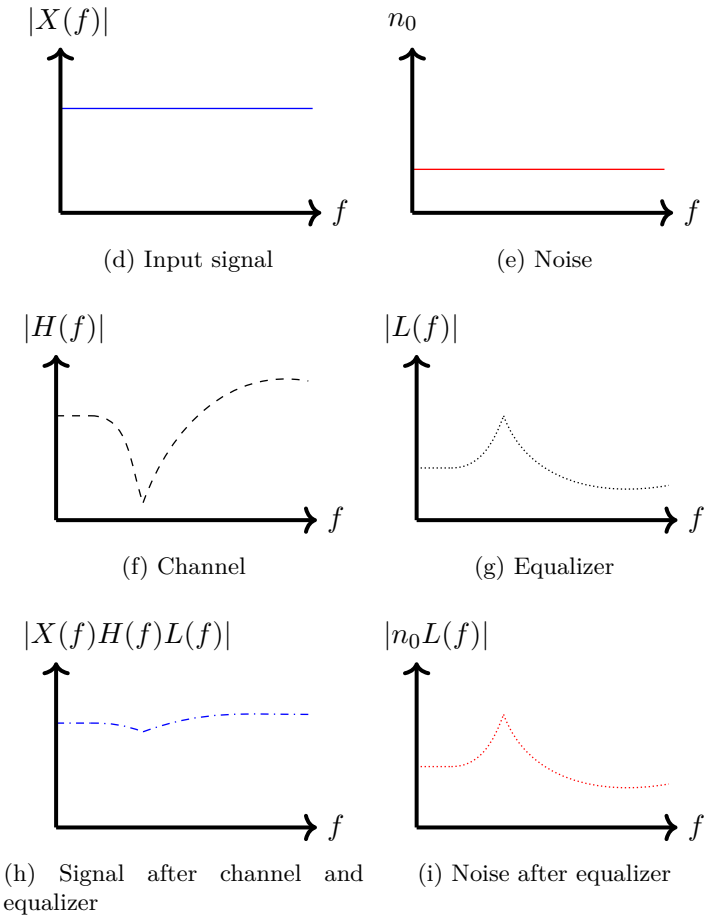
$$L(f) = \frac{H^*(f)}{|H(f)|^2 + \zeta}\,, \quad \zeta = \begin{cases} \sigma_n^2/\sigma_s^2 & \text{for MMSE}\,, \\ 0 & \text{for ZF}, \end{cases} \tag{5.15}$$

where σ_n^2 and σ_s^2 denote the noise and signal powers, respectively. Given that the signal $X(f)$ is sent from the transmitter, the received signal after equalization ($Y(f)$) can be written as

$$\begin{aligned} Y(f) &= X(f)H(f)L(f) + n_0 L(f), \\ \mathrm{Var}(X) &= \sigma_s^2,\ \mathrm{Var}(n_0) = \sigma_n^2, \end{aligned} \tag{5.16}$$

where n_0 denotes the additive white Gaussian noise. After the equalizer at the receiver, the noise cannot be considered as white anymore (i.e., as in $n_0 L(f)$), since certain frequency components may excessively be attenuated or amplified (see Fig. 5.9).

It is possible to mitigate the ISI by introducing frequency dependent precoding at the base station in massive MIMO, which helps to reduce the complexity of the receiver, since equalization is effectively carried out at the transmitter (which may also be called as pre-equalization) (Payami and Tufvesson, 2013; Pitarokoilis *et al.*, 2012). Furthermore, such an arrangement can be beneficial to avoid noise enhancement issues in certain cases (Wen *et al.*, 2010; Liu *et al.*, 2016). Assuming that an RZF precoder is utilized and



(d) Input signal (e) Noise

(f) Channel (g) Equalizer

(h) Signal after channel and equalizer (i) Noise after equalizer

Figure 5.9: Illustration of the effects of channel equalization. An input signal (a) which goes through the channel (c) and equalizer (d) can afterwards be recovered without significant distortion (e) but the characteristics of the noise (b) may considerably be altered (f) (Taygur and Eibert, 2019a). © 2019 IEEE.

CSI is acquired perfectly at the base station, the ideal precoding matrix can be given as

$$\mathbf{A}(f) = (\mathbf{G}(f))^H \left(\mathbf{G}(f)(\mathbf{G}(f))^H + \alpha\mathbf{I}\right)^{-1},$$
$$f^{\text{low}} \leqslant f \leqslant f^{\text{up}}, \tag{5.17}$$

where α is a regularization constant, f^{low} and f^{up} are the lower and upper frequency bounds of the allocated channel bandwidth, respectively. Here, the elements of the precoding matrix $\mathbf{A}$ are frequency dependent complex coefficients. As the precoding is carried out in the digital domain, the matrix $\mathbf{A}$ is computed for certain discrete frequency points. Thus, the expression for the precoding matrix can be re-written as

$$\mathbf{A}_n = (\mathbf{G}(f_n))^H \left(\mathbf{G}(f_n)(\mathbf{G}(f_n))^H + \alpha\mathbf{I}\right)^{-1},$$
$$n = 1, \cdots, N_F,$$
$$f_n = f^{\text{low}} + \frac{(f^{\text{up}} - f^{\text{low}})(2n-1)}{2N_F}, \tag{5.18}$$

where f_n denotes a frequency point for which the channel frequency response is known and a corresponding precoding matrix is calculated, and N_F is the total number of frequency points. It should be emphasized that the precoding is performed over each frequency point individually, thereby N_F different precoding matrices should be calculated, which implies that the complexity of the approach scales linearly with the number of frequency points. The downlink SINR and data rate for the uth user in the network can then be written as

$$\mathrm{SINR}_u(f) = \frac{P^{\mathrm{tx}}|\mathbf{G}_{(u,:)}(f)\ \mathbf{A}_{n(:,u)}|^2}{P^{\mathrm{tx}}\sum\limits_{\substack{v=1\\v\neq u}}^{64}|\mathbf{G}_{(u,:)}(f)\ \mathbf{A}_{n(:,v)}|^2+\sigma_n^2},\ R_u = \int_{f^{\mathrm{low}}}^{f^{\mathrm{up}}}\log_2\left(1+\mathrm{SINR}_u(f)\right)df,$$

$$n = \left\lceil \frac{N_F\left(f-f^{\mathrm{low}}\right)}{f^{\mathrm{up}}-f^{\mathrm{low}}} \right\rceil. \tag{5.19}$$

The SINR expression is a function of the frequency here, and the precoder to be used ($\mathbf{A}_n$) is determined according to the frequency. The average downlink data rate in the entire cell can be given by

$$R^{\mathrm{avg}} = \frac{W}{64}\sum_{u=1}^{64} R_u. \tag{5.20}$$

5.4.1 Creating Frequency-Selective Channel Matrices

In order to demonstrate the effects of frequency selective channel behavior under different channel modeling approaches, ray-tracing simulations and the WINNER II (Kyosti *et al.*, 2007) model are utilized to generate two different channel matrices. In both approaches, a frequency dependent channel matrix $\mathbf{G}(f)$ is obtained for a 10 MHz bandwidth with a 10 kHz resolution (i.e., 1001 frequency points in total[1]). The channel matrix, which is created by ray-tracing simulations, can be expressed as

$$\mathbf{G}_{(u,v)}(f) = \sum_{m=1}^{M}\frac{V_{m,u}^{\mathrm{oc}}(f)}{V_v^{\mathrm{gen}}}, \tag{5.21}$$

where V_v^{gen} is the generator voltage at the vth transmitter, $V_{m,u}^{\mathrm{oc}}(f)$ is the open circuit voltage at the terminals of uth receiver which is induced by the mth multi-path contribution and M is the total number of unique multi-paths. The generator voltage can be considered as frequency-independent for simplicity. A tapped delay line model is utilized in WINNER II in order to create a delayed echo effect (Kyosti *et al.*, 2007). A 20-tap model is used to express the impulse response of the channel with

$$\mathbf{g}_{(u,v)}(t) = \sum_{n=1}^{20} A_{n,u,v}\delta(t-\tau_n) \tag{5.22}$$

[1]This implies $N_F < 1001$ in general. For the sake of brevity, discretization of the frequency domain for the channel matrix computation is omitted in the notation and simply given by e.g., $\mathbf{G}(f)$ or $\mathrm{SINR}_u(f)$.

where $\tau_{n,u,v}$ is the time delay for the nth tap and $A_{n,u,v}$ is the amplitude for the nth tap. The tap delays and amplitudes can be expressed as

$$\tau_{n,u,v} = \text{sort}(-3.2\ln(\delta)10^{-7.12}),\ \delta \sim \mathcal{U}(0,1),$$
$$\xi_{u,v} = \sum_{n=1}^{20} e^{-\tau_{n,u,v}/(10^{-7.12})},\ \gamma(d) = 10^{-(31.1+41.1\log_{10}(0.707d)+\kappa)/20},\ \kappa \sim \mathcal{N}(0,4),$$
$$A_{n,u,v} = \gamma(d_{u,v})\frac{\chi e^{-\tau_{n,u,v}}}{\xi_{u,v}\left(10^{-7.12}\right)},\ \chi \sim \mathcal{CN}(0,1), \tag{5.23}$$

where $d_{u,v}$ is the distance between the uth receiver and vth transmitter. The amplitudes of the channel taps decrease exponentially with the delay. The impulse response $\mathbf{g}_{(u,v)}(t)$ is converted into frequency domain representation via Fourier transform such that[2]

$$\mathbf{G}_{(u,v)}(f) = \mathfrak{F}\left\{\mathbf{g}_{(u,v)}(t)\right\}. \tag{5.24}$$

5.4.2 Numerical Results

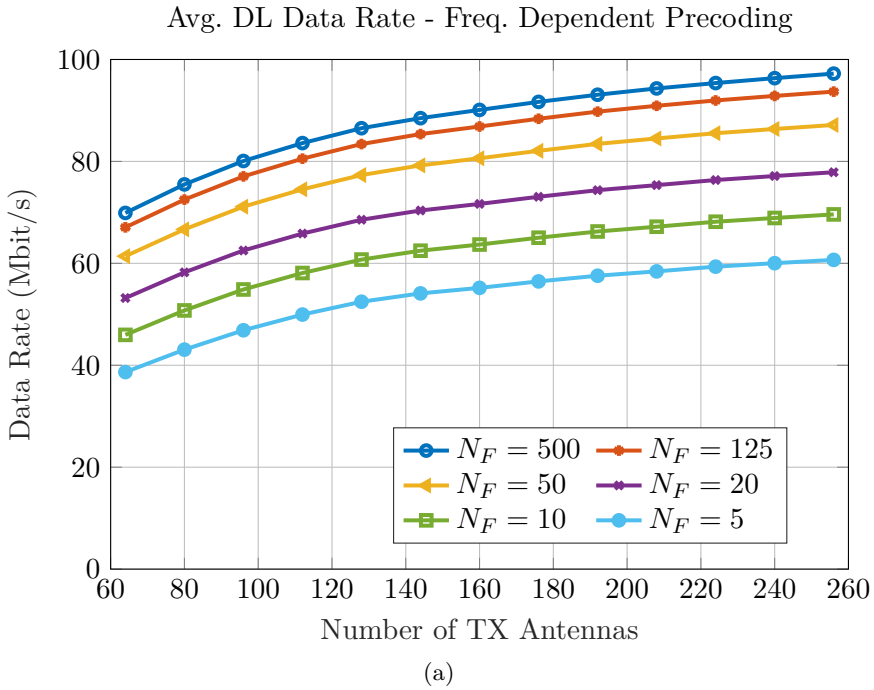
A downlink data rate analysis is performed according to both channel models in order to characterize the effects of frequency-selective fading in massive MIMO. The total number of discrete frequency points in precoding (N_F) is changed from 5 to 500 and the number of base station antennas is changed from 64 to 256. The number of frequency points determines whether the variations in the channel frequency response can be captured by the equalizer. The improvement in performance, which can be obtained by improving the discretization, diminishes as the channel frequency response at adjacent frequency bins typically become similar. Hence, the coherence bandwidth of the channel can be estimated by using the spacing between adjacent frequency bins. The results are shown in Fig. 5.10.

The improvement in data rate performance is evident in both figures, as the number of the frequency points is increased. For the ray-tracing model, a gradual improvement in performance can be observed at each increment in the number of frequency points. On the other hand, the improvement is very limited in the WINNER II model beyond 50 frequency points, which indicates that the coherence bandwidth according to the WINNER II model is larger than that in the ray-tracing simulations. Once the number of frequency points is sufficiently large, such that the coherence bandwidth is larger than the spacing between adjacent frequency bins, an improvement in performance may only be obtained by increasing the number of base station antennas. It can be noticed that the system performance for 256 antennas is similar in both models but a considerable difference exists for the 64 antenna case.

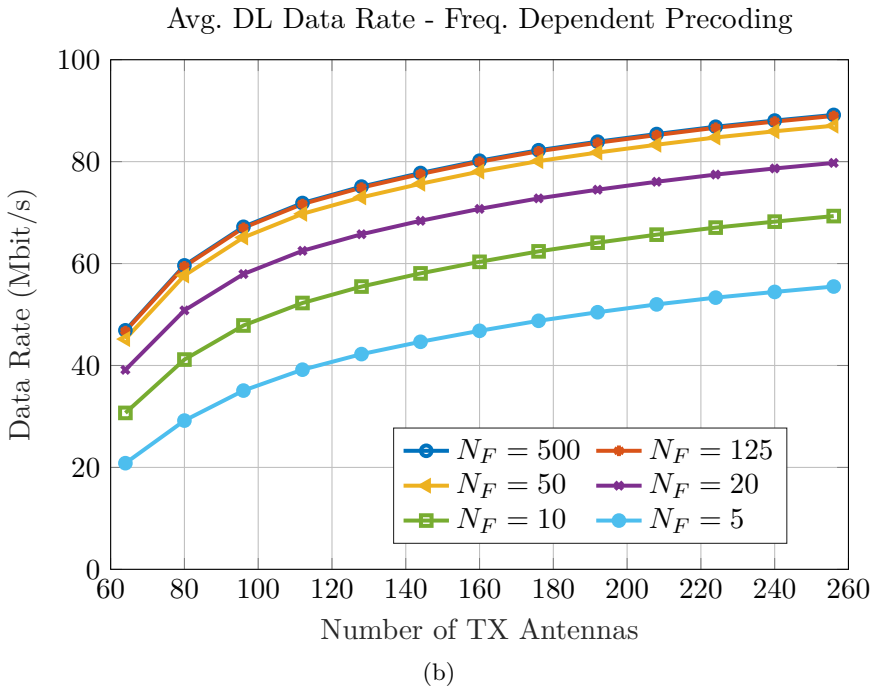
5.5 Multi-Antenna User Terminals in Massive MIMO

The use of multiple antennas at the user terminals enables various downlink precoding strategies, which are different than those in the single antenna user case (Li *et al.*, 2016; Bengtsson *et al.*, 2016). The applicability of these approaches generally depends on

[2]In order to obtain a frequency-domain representation with 10 kHz resolution over a 10 MHz bandwidth after the Fourier transform, the impulse response of the channel was discretized accordingly where the tap delays were quantized with 10^{-7} s resolution.



(a)



(b)

Figure 5.10: Average downlink data rate based on ray-tracing simulations (a) and the WINNER II model (b) with frequency selective-fading. The number of discrete frequency points utilized in the precoder is denoted by N_F.

the receiver architecture as well as the availability of CSI. There are three common downlink transmission schemes which can be utilized in a network with multi-antenna user terminals (Taygur and Eibert, 2019b).

Regularized Zero Forcing: The base station transmits independent data streams for each receive antenna in the network, i.e., a single user terminal may receive as many independent data streams as the number of antennas it has. The base station must ensure that the interference among the data streams is properly suppressed, since every user antenna, including those which are located in the same terminal, are treated as a separate, independent entity.
Regularized Block Diagonalization: The base station transmits independent data streams for each receive antenna in the network (as in the Regularized Zero Forcing case), however, the user terminals can jointly process the signals from their own antennas in this case. Hence, the entire channel may effectively be decomposed into several single-user MIMO channels where the user terminals employ common MIMO decoding techniques in order to extract the data streams. The requirements concerning interference suppression is less restrictive compared to the first case where the base station mitigates only the interference among different user terminals.
Regularized Zero Forcing & Receive Combining: Each user terminal receives only a single data stream where the user antennas are used to improve the signal power by means of coherent combining.

In the following, all three approaches will be discussed in more detail where precoding and detection techniques will be presented, and a downlink data rate expression will be given for each case. The discussions are based on a single-cell system which consists of 64 user terminals, each having 2 receive antennas, and a base station with 256 transmit antennas. Thus, the channel matrix can can be written as

$$\mathbf{G} = \left[\mathbf{G}_1^T \;\; \mathbf{G}_2^T \;\; \cdots \;\; \mathbf{G}_u^T \;\; \cdots \;\; \mathbf{G}_{64}^T\right]^T, \; \mathbf{G}_u \in \mathbb{C}^{2\times 256}, \tag{5.25}$$

where $\mathbf{G}_u$ represent the channel for the uth user.

5.5.1 Regularized Zero-Forcing

The regularized zero forcing approach implies that every user antenna is treated as if it belongs to an independent user terminal and receives an independent data stream by the base station (Yoo and Goldsmith, 2006; Björnson *et al.*, 2013). Each user terminal consists of two receiver chains, which are independent of each other and cannot cooperate. Even though the receivers have a relatively simple design and CSI at the base station is generally sufficient for normal operation, the base station has to prevent interference not only among different user terminals, but also between the antennas which belong to the same terminal (Spencer *et al.*, 2004). This might be a challenging task, considering that the channel vectors for the antennas in the same user terminal are similar due to very limited physical separation. The precoding matrix can be given as

$$\begin{aligned} \mathbf{A}^{\mathrm{RZF}} &= \mathbf{G}^H \left(\mathbf{G}\mathbf{G}^H + \alpha\mathbf{I}\right)^{-1}, \\ \mathbf{A} &= \left[\mathbf{A}_1 \;\; \mathbf{A}_2 \;\; \cdots \;\; \mathbf{A}_u \;\; \cdots \;\; \mathbf{A}_{64}\right], \; \mathbf{A}_u \in \mathbb{C}^{256\times 2}. \end{aligned} \tag{5.26}$$

The effective channel after precoding ($\mathbf{G}\mathbf{A}^{\mathrm{RZF}}$) for a 4 user scenario is depicted in Fig. 5.11. The downlink data rate for the uth user in the network can be given as

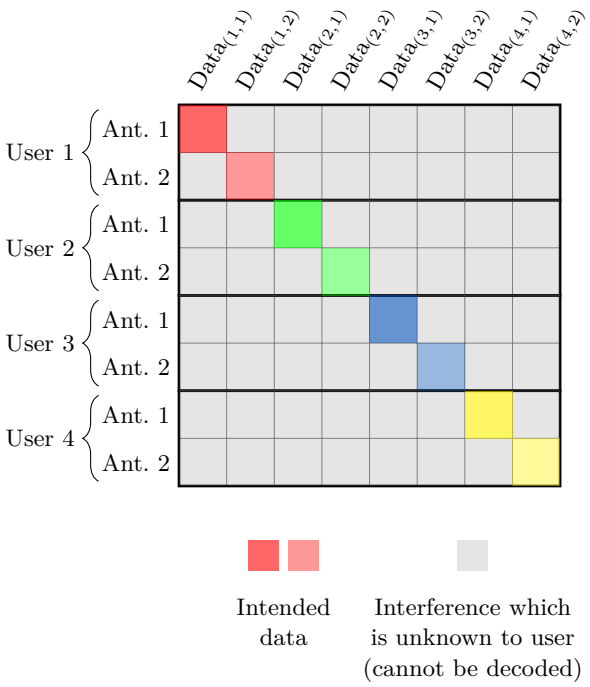


Figure 5.11: Illustration of effective channel after RZF precoding with no cooperation between receiver blocks. $\text{Data}_{(u,a)}$ denotes the data for the ath receiver of uth user (Taygur and Eibert, 2019a). © 2019 IEEE.

$$R_u = W \log_2 \left(\left(1 + \frac{P^{\text{tx}} \, |\mathbf{G}_{u(1,:)} \mathbf{A}_{u(:,1)}|^2}{P^{\text{tx}} \sum\limits_{v=1, v \neq 2u-1}^{128} |\mathbf{G}_{u(1,:)} \mathbf{A}_{(:,v)}|^2 + \sigma_n^2} \right) \times \left(1 + \frac{P^{\text{tx}} \, |\mathbf{G}_{u(2,:)} \mathbf{A}_{u(:,2)}|^2}{P^{\text{tx}} \sum\limits_{v=1, v \neq 2u}^{128} |\mathbf{G}_{u(2,:)} \mathbf{A}_{(:,v)}|^2 + \sigma_n^2} \right) \right). \tag{5.27}$$

5.5.2 Regularized Block Diagonalization

The strict interference suppression requirements in the regularized zero forcing approach may yield sub-optimal performance in certain cases. This problem can be resolved by changing the receiver architecture such that the received signals are jointly processed. The base station then suppresses the interference among user terminals as in the previous case, but does not attempt to suppress the interference among the antennas which are located in the same user terminal (i.e., inter-antenna interference). Thus, the entire channel is effectively decomposed into multiple SU-MIMO channel blocks (hence the name block diagonalization) where the user terminals then employ appropriate decoding techniques in order to eliminate inter-antenna interference and to extract the data (Stankovic and Haardt, 2008). Note that the user terminals need to acquire the CSI for their own effective channels in order to employ SU-MIMO decoding.

Regularized block diagonalization can be considered as a generalization of the regularized zero forcing. The data rate performance is typically better than that of the regularized zero forcing method due to less stringent interference suppression requirements. An explicit expression for the precoding matrix $\mathbf{A}^{\mathrm{RBD}}$ can be derived as follows. Let us define $\overline{\mathbf{G}}_u$, which is the complementary channel matrix for user u, as

$$\begin{gathered}\overline{\mathbf{G}}_u = \begin{bmatrix}\mathbf{G}_1^T & \cdots & \mathbf{G}_{u-1}^T & \mathbf{G}_{u+1}^T & \cdots & \mathbf{G}_{64}^T\end{bmatrix}^T, \\ \overline{\mathbf{G}}_u \in \mathbb{C}^{126\times 256}.\end{gathered} \tag{5.28}$$

Applying singular value decomposition on $\overline{\mathbf{G}}_u$ yields

$$\begin{gathered}\overline{\mathbf{G}}_u = \overline{\mathbf{U}}_u \overline{\mathbf{\Sigma}}_u \overline{\mathbf{V}}_u^H, \\ \overline{\mathbf{U}}_u \in \mathbb{C}^{126\times 126},\ \overline{\mathbf{\Sigma}}_u \in \mathbb{C}^{126\times 256},\ \overline{\mathbf{V}}_u \in \mathbb{C}^{256\times 256},\end{gathered} \tag{5.29}$$

where $\overline{\mathbf{\Sigma}}_u$ is a diagonal matrix with the singular values of $\overline{\mathbf{G}}_u$, $\overline{\mathbf{V}}_u$ and $\overline{\mathbf{U}}_u$ are unitary matrices which contain right and left singular vectors of $\overline{\mathbf{G}}_u$, respectively (Zu *et al.*, 2013). The precoder matrix $\mathbf{A}^{\mathrm{RBD}}$ can be written as

$$\begin{gathered}\mathbf{A}^{\mathrm{RBD}} = \begin{bmatrix}\mathbf{A}_1^{\mathrm{RBD}} & \mathbf{A}_2^{\mathrm{RBD}} & \cdots & \mathbf{A}_u^{\mathrm{RBD}} & \cdots & \mathbf{A}_{64}^{\mathrm{RBD}}\end{bmatrix},\ \mathbf{A}_u^{\mathrm{RBD}} \in \mathbb{C}^{256\times 2}, \\ \mathbf{A}_u^{\mathrm{RBD}} = \mathbf{A}_u^a \mathbf{A}_u^b, \\ \mathbf{A}_u^a = \overline{\mathbf{V}}_u \left(\overline{\mathbf{\Sigma}}_u^T \overline{\mathbf{\Sigma}}_u + \alpha \mathbf{I}_{256\times 256}\right)^{-1/2}, \\ \mathbf{G}_u^{\mathrm{Ga}} = \mathbf{G}_u \mathbf{A}_u^a = \mathbf{U}_u^{\mathrm{Ga}} \mathbf{\Sigma}_u^{\mathrm{Ga}} \begin{bmatrix}\mathbf{V}_u^{\mathrm{Ga},(1)} & \mathbf{V}_u^{\mathrm{Ga},(0)}\end{bmatrix}^H, \\ \mathbf{V}_u^{\mathrm{Ga},(1)} \in \mathbb{C}^{2\times 256},\ \mathbf{V}_u^{\mathrm{Ga},(0)} \in \mathbb{C}^{264\times 256}, \\ \mathbf{A}_u^b = \mathbf{V}_u^{\mathrm{Ga},(1)},\end{gathered} \tag{5.30}$$

where $\mathbf{V}_u^{\mathrm{Ga},(0)}$ and $\mathbf{V}_u^{\mathrm{Ga},(1)}$ are the right singular vectors of $\mathbf{G}_u^{\mathrm{Ga}}$, which correspond to zero and non-zero singular values, respectively (Sung *et al.*, 2009). An illustration of the effective channel for a 4 user scenario is shown in Fig. 5.12.

The downlink data rate for the uth user can be written as

$$\begin{gathered}R_u = W \log_2 \left(\det\left(\mathbf{I} + P^{\mathrm{tx}} \left(\mathbf{\Delta}_u\right)^{-1} \mathbf{G}_u \mathbf{A}_u \mathbf{A}_u^H \mathbf{G}_u^H\right)\right), \\ \mathbf{\Delta}_u = P^{\mathrm{tx}} \left(\sum_{v=1, v\neq u}^{64} \mathbf{G}_u \mathbf{A}_v \mathbf{A}_v^H \mathbf{G}_u^H\right) + \mathbf{I}\sigma_n^2.\end{gathered} \tag{5.31}$$

5.5.3 Regularized Zero Forcing & Receive Combining

In this scheme, the terminals utilize their antennas to combine the incoming signal coherently in order to improve the SINR. The base station transmits a single data stream towards each user, and generally does not know how many antennas the users have (Jindal, 2008). The prominent advantages of this transmission scheme are the relative simplicity of the receiver system and the possibility to accommodate a larger number of users in the network compared to the other two schemes (Bengtsson *et al.*, 2016). Combining the signal at the receiver can be carried out by equal gain combining (EGC) method, which may not always yield the optimal result in terms of SINR (compared to

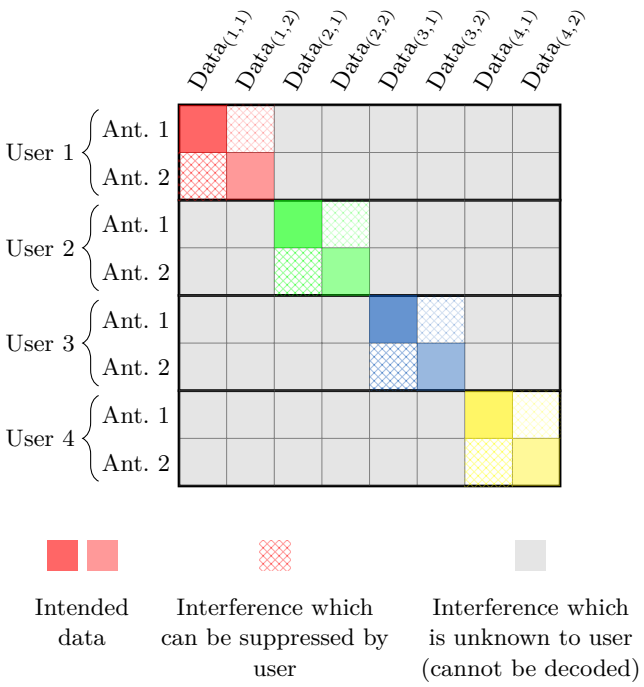


Figure 5.12: Illustration of effective channel after BRD precoding. $\text{Data}_{(u,a)}$ denotes the data for the ath receiver of uth user (Taygur and Eibert, 2019a). © 2019 IEEE.

maximum ratio combining), nevertheless, it is straightforward to implement since the user terminals do not need CSI (Ngo *et al.*, 2013). During the uplink training phase, the users employ both of their antennas to transmit the pilot signal, which creates an effective channel denoted by $\widehat{\mathbf{G}}$ such that

$$\widehat{\mathbf{G}}_u = \mathbf{p}\mathbf{G}_u, \quad \mathbf{p} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \widehat{\mathbf{G}}_u \in \mathbb{C}^{1\times 256},$$
$$\widehat{\mathbf{G}} = \begin{bmatrix} \widehat{\mathbf{G}}_1^T & \widehat{\mathbf{G}}_2^T & \cdots & \widehat{\mathbf{G}}_u^T & \cdots & \widehat{\mathbf{G}}_{64}^T \end{bmatrix}^T, \quad \widehat{\mathbf{G}} \in \mathbb{C}^{64\times 256}, \tag{5.32}$$

where $\mathbf{p}$ denotes the EGC at the receiver. Assuming that the CSI acquisition and the channel estimation of the base station are perfect, the regularized zero forcing precoder $\mathbf{A}^{\text{RZF-comb}}$ can be given by

$$\mathbf{A}^{\text{RZF-comb}} = \left(\widehat{\mathbf{G}}\right)^H \left(\widehat{\mathbf{G}}\widehat{\mathbf{G}}^H + \alpha\mathbf{I}\right)^{-1}, \quad \mathbf{A}^{\text{RZF-comb}} \in \mathbb{C}^{256\times 64},$$
$$\mathbf{A}^{\text{RZF-comb}} = \begin{bmatrix} \widehat{\mathbf{A}}_1 & \widehat{\mathbf{A}}_2 & \cdots & \widehat{\mathbf{A}}_u & \cdots & \widehat{\mathbf{A}}_{64} \end{bmatrix}, \quad \widehat{\mathbf{A}}_u \in \mathbb{C}^{256\times 1}. \tag{5.33}$$

The illustration of the effective channel after precoding (combining at the receiver not considered) is shown in Fig. 5.13. The downlink data rate for the uth user can then be written as

$$R_u = W \log_2 \left(1 + \frac{P^{\text{tx}} \left|\sqrt{2}\mathbf{p}\mathbf{G}_u\widehat{\mathbf{A}}_u\right|^2}{P^{\text{tx}} \left(\sum_{v=1, v\neq u}^{64} \left|\sqrt{2}\mathbf{p}\mathbf{G}_u\widehat{\mathbf{A}}_v\right|^2 \right) + \sigma_n^2} \right). \tag{5.34}$$

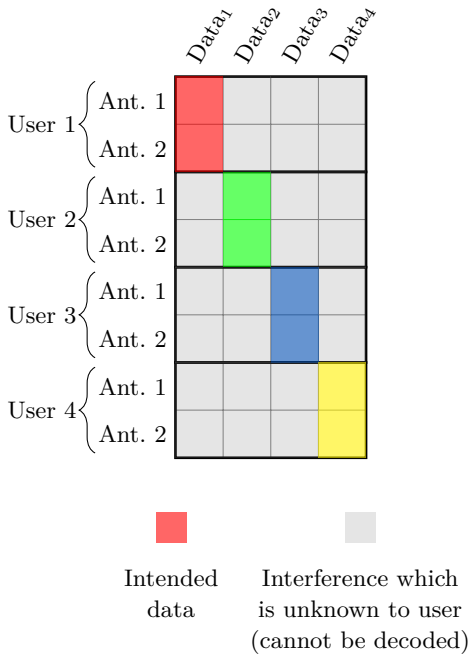


Figure 5.13: Illustration of effective channel after RZF precoding where users employ receive combining. Data_u denotes the data for the uth user (Taygur and Eibert, 2019a). © 2019 IEEE.

Note that the factor of $\sqrt{2}$ in front of $\mathbf{p}$ occurs due to the fact that the received signals are directly combined to achieve an array gain at the user terminals.

5.5.4 Channel Similarity Characterization for User Antennas

Since the antennas in the same user terminal are physically very close to each other in general, the channel vectors associated with these antennas can be similar, as stated previously. As the data rate is commonly influenced by this similarity, it is essential to characterize it in order to estimate the performance. The similarity of the channel vectors, which correspond to two antennas of the uth user, can be given by

$$\mathbf{v}_{1,u} = \frac{\sqrt{\mathbf{G}_{u(1,:)}\left(\mathbf{G}_{u(1,:)}\right)^H}}{||\mathbf{G}_{u(1,:)}||}, \; \mathbf{v}_{2,u} = \frac{\sqrt{\mathbf{G}_{u(2,:)}\left(\mathbf{G}_{u(2,:)}\right)^H}}{||\mathbf{G}_{u(2,:)}||},$$
$$\mathbf{v}_{1,u}, \mathbf{v}_{2,u} \in \mathbb{R}^{256},$$
$$\mathcal{U}_u = \langle \mathbf{v}_{1,u}, \mathbf{v}_{2,u} \rangle, \tag{5.35}$$

where $\mathcal{U}_u : 0 \leqslant \mathcal{U}_u \leqslant 1$ is the similarity coefficient ($\mathcal{U}_u = 1$ indicates complete alignment of two vectors i.e., maximum similarity and $\mathcal{U}_u = 0$ indicates complete orthogonality, i.e., minimum similarity), $\langle ., . \rangle$ denotes the inner product of two vectors, and $\mathbf{G}_{u(1,:)}$ and $\mathbf{G}_{u(2,:)}$ are the first and second rows of matrix $\mathbf{G}_u$, respectively.

5.5.5 Methodology

A performance comparison for the previously described transmission schemes is carried out under two different channel models, namely, ray-tracing and WINNER II (Kyosti *et al.*, 2007). The coefficients of the channel matrix for both models can be given by

$$\underline{\text{Ray-Tracing}} \qquad \underline{\text{WINNER II}}$$

$$\mathbf{G}_{(u,v)} = \frac{V_u^{\text{oc}}}{V_v^{\text{gen}}}, \qquad \mathbf{G}_{(u,v)} = \xi G^{\{\text{out,in}\}}\left(d_{u,v}\right),$$

$$G^{\text{out}}(x) = 10^{-(41.1\log_{10}(x\sqrt{2})+30.8+\chi^{\text{out}})\times 0.05},$$

$$G^{\text{in}}(x) = 10^{-(41.1\log_{10}(x\sqrt{2})+41+\chi^{\text{in}})\times 0.05},$$

$$\xi \sim \mathcal{CN}(0,1),$$

$$\chi^{\text{out}} \sim \mathcal{N}(0,4),\ \chi^{\text{in}} \sim \mathcal{N}(0,7), \tag{5.36}$$

where G^{out} and G^{in} are path gain coefficients which apply to outdoor and indoor users, respectively, and $d_{u,v}$ denotes the distance between the uth receive antenna in the cell (located in a user terminal) and vth base station antenna.

As the downlink data rate for each user is computed according to Eqs. (5.27), (5.31) and (5.34) by using the respective channel matrices, the average downlink data rate R^{avg} can be calculated by

$$R^{\text{avg}} = \frac{1}{64}\sum_{u=1}^{64} R_u. \tag{5.37}$$

5.5.6 Simulation Results

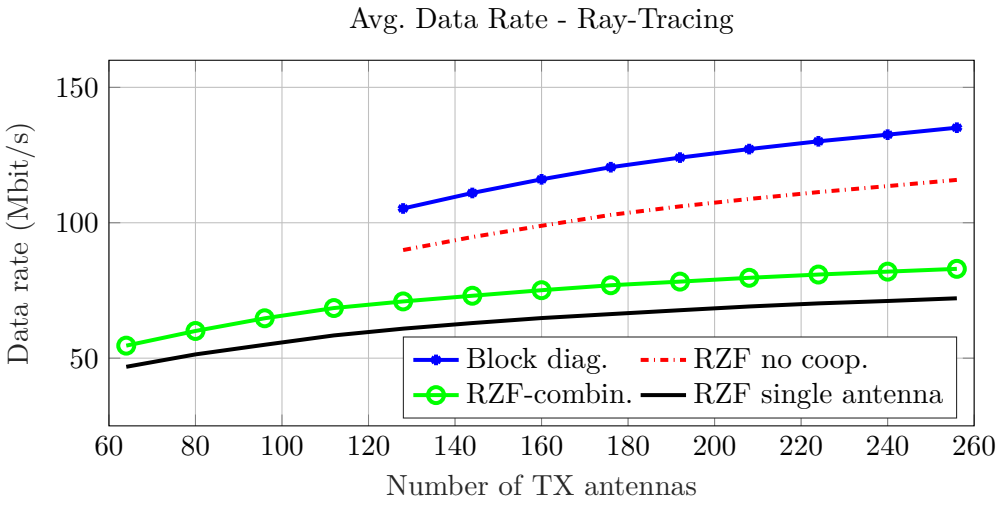
The average downlink data rate performance for the three transmission schemes under two different channel modeling approaches is investigated. In order to demonstrate the influence of the use of multiple antennas at the user terminals, an additional scenario, which involves single antenna user terminals and regularized zero forcing precoding, is also considered. The number of base station antennas was gradually changed from 64 to 256. Note that the regularized zero forcing with non-cooperative receivers and block diagonalization schemes may not be utilized with less than 128 base station antennas, given that there are 64 user terminals with 2 antennas at each.

The downlink data rate curves in Fig. 5.14 show that the best performance is achieved by block diagonalization, which is followed by non-cooperative RZF. The discrepancy between the performances of two methods imply that the channel vectors, which correspond to the receive antennas in the same user terminal, may have a significant similarity. When the interference between these channels may not effectively be suppressed due to such a channel similarity, the block diagonalization technique outperforms non-cooperative RZF. On the other hand, RZF with the receive combining approach yields only a slight improvement in data rate over the single-antenna RZF case, and generally lags behind the other two multi-antenna approaches.

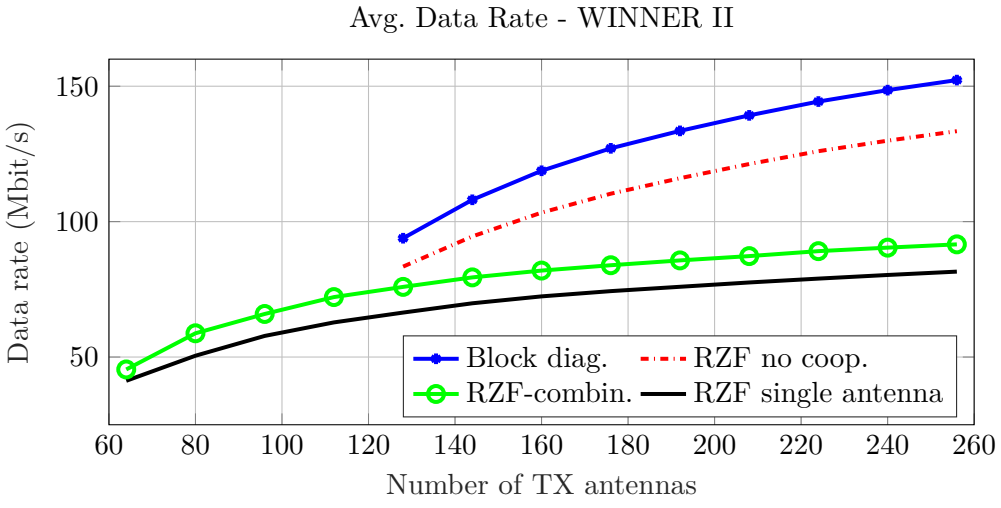
It can be observed that both ray-tracing and the WINNER II model yield similar performance characteristics. The WINNER model predicts a slightly higher data rate for the 256 base station antenna case and the improvement of the performance with respect to the number of base station antennas is more noticeable. In addition to the downlink data rate performance study, the channel similarity of terminal antennas of 64 users is

investigated according to Eq. (5.35). A distribution is presented for the ray-tracing and WINNER II models in Fig. 5.15.

The ray-tracing and WINNER models predict significantly different channel similarity characteristics, according to Fig. 5.15, where the ray-tracing results imply that the similarity is very high. Although, neither model yields a similarity factor below 0.2, the distribution for the WINNER II model is closer to the left-hand side of the plot, indicating a larger channel discrepancy. This also explains why the WINNER model shows a notable increase in data rate (compared to the ray-tracing results) as the number of base station antennas grows, i.e., new antennas added into the base station yield distinct, uncorrelated channels, thus, the data rate improves. Based on these results, it can be concluded that the correlation between the antennas in the same user terminal should be introduced into the WINNER model in order to depict the channel similarity accurately.



(a)



(b)

Figure 5.14: Average downlink data rate based on ray-tracing simulations (a) and WINNER II model (b) with multi-antenna users.

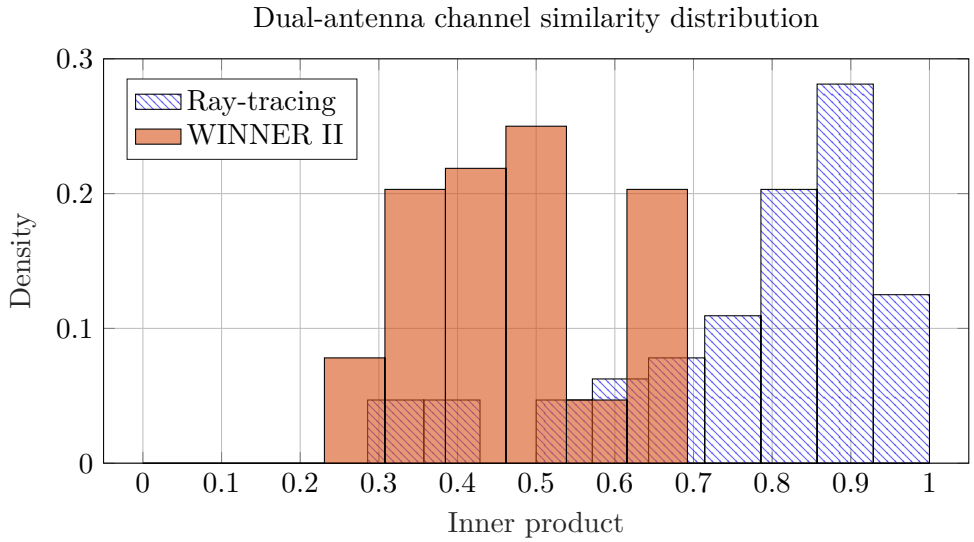


Figure 5.15: Distribution of channel similarity of dual antenna arrangement in user terminals according to ray-tracing and the WINNER II model.

Chapter 6

Massive MIMO Downlink Performance Analysis by Bidirectional Ray-Tracing

The downlink performance analysis of two massive MIMO scenarios is presented where the scenarios involve particular propagation conditions, which favor the use of the bidirectional ray-tracing approach. In particular, the following scenarios are considered:

1. Propagation over rooftops with users on ground level and base station above the buildings.
2. Millimeter-wave communication in an urban environment & comparisons to lower carrier frequencies.

The scenarios involve distinct geometries and propagation conditions, therefore, the essential simulation parameters differ significantly. As in the previous chapter, the average user downlink data rate was considered as the primary performance metric. Additional data regarding the important features of the channel, such as signal-interference characteristics, is also given.

6.1 Analysis of Scenarios with Rooftop Diffractions

A commonly encountered propagation phenomenon in urban environments are rooftop diffractions, where the wave propagates over building roofs via diffractions. The channel transfer function in such a geometry can be obtained by bidirectional ray-tracing simulations, where a large interaction surface, which goes beyond the roof level, is placed between the receivers and the base station (Taygur *et al.*, 2019b).

6.1.1 Methodology

Urban Scenario

The simulated urban environment has 25 buildings arranged in a grid-like structure and covers an area of 250 m×250 m. The heights of the buildings change between 22 m-25 m. The buildings as well as the ground plane are assumed to have a relative permittivity of $\epsilon_r = 4$. A base station, which consists of 100 antennas in a 10×10 arrangement, is placed on top of a building, and 10 single antenna terminals are distributed randomly on the ground level where all the antennas are assumed as half-wavelength dipoles and the operating frequency is 2.45 GHz. The building arrangement has been adjusted to prevent

propagation paths through street canyons, thus, rooftop propagation is the only viable option for establishing a link between the receivers and the base station. An interaction surface with a size of 240 m×80 m (divided into 8192 triangles) is located between the receivers and the base station. An illustration of the scenario is shown in Fig. 6.1.

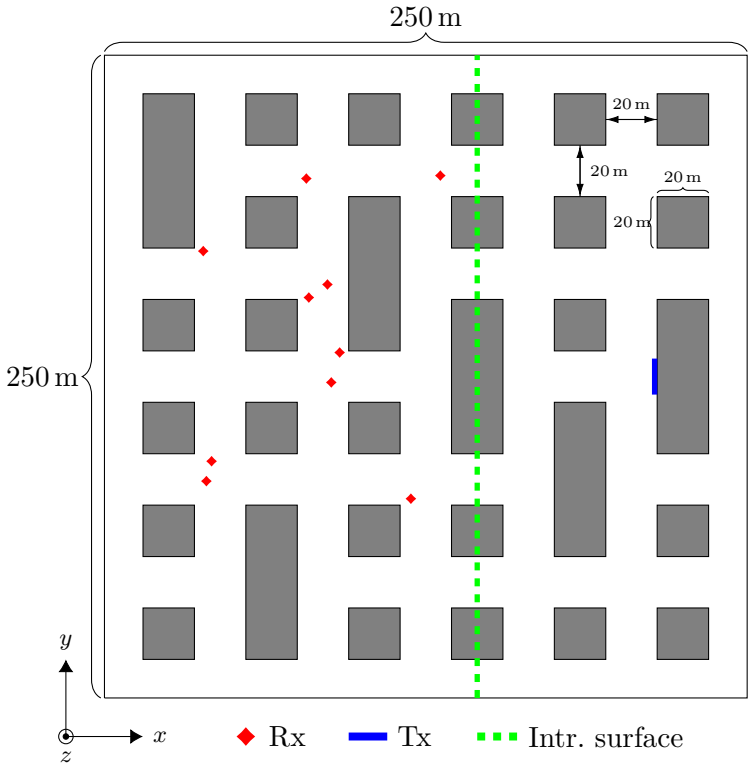


Figure 6.1: Urban scenario 2D top view illustration (Taygur *et al.*, 2019b). Reproduced courtesy of The Electromagnetics Academy.

6.1.2 Numerical Results

Convergence Analysis

A parameter convergence analysis has been performed for the urban scenario by means of the iterative SBR approach described in Chapter 3. A single transmitting antenna from the base station and 10 user terminals are utilized where three different settings are used for the number of ray launches from each antenna, i.e., 100,000, 200,000 and 300,000. In total, 11 iterations are carried out (including the initial ray launch), and the following parameters were analyzed:

1. Total unique ray hit count at each iteration, averaged according to the number of receivers.
2. The relative change in the unique ray hit count at each iteration, averaged with respect to the number of receivers.
3. Average channel gain (i.e., coupling for all receiver-transmitter pairs) at the last iteration.

The results are shown in Fig. 6.2.

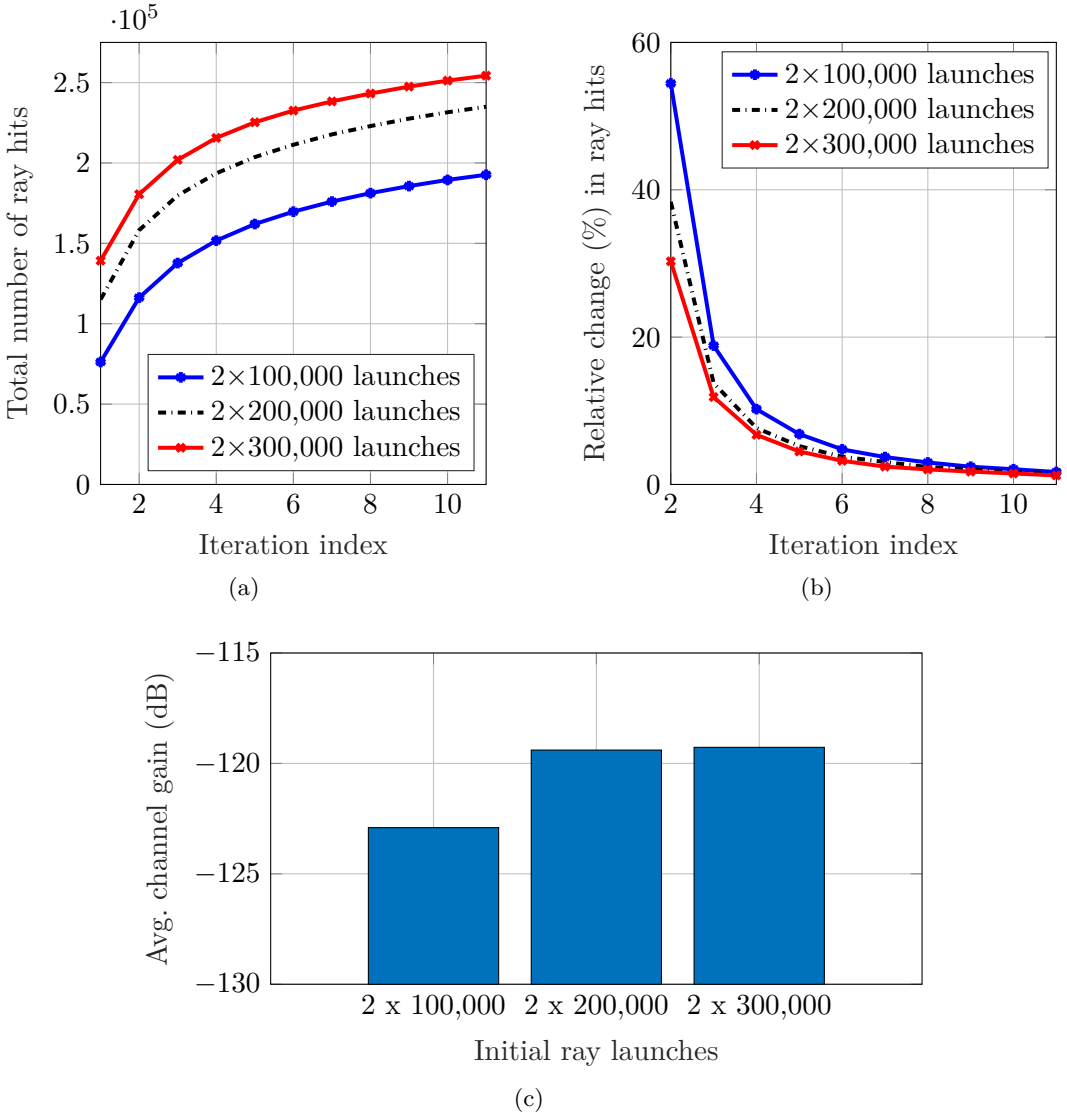


Figure 6.2: Variation of total number of ray hits (a), relative change in number of ray hits (b) and average channel gain (c) for three different ray launch configurations in the urban scenario with iterative SBR.

It can be observed that the (unique) ray hit count increases monotonically at each iteration step, and the relative improvement beyond 10 iterations is rather limited. The ray hit count is correlated with the number of ray launches. In Fig. 6.2(c), it can be noticed that 2×300,000 ray launches (factor of 2 is due to receiver and transmitter launches) yields convergence in terms of average path gain after 11 iterations, as the difference compared to the 2×200,000 launches is found to be less than 0.5 dB.

Simulation Process

Based on the outcome of the convergence analysis, the bidirectional ray-tracing simulations were performed with 2×300,000 ray launches (for each individual receiver-transmitter pair) and 11 iterations. Up to 4 reflections as well as 1 diffraction on building

rooftops were considered at both sites, and a channel matrix is obtained accordingly. A flat channel response was assumed for a 10 MHz bandwidth and the total transmission power from the base station was 1 W. In order to compare the bidirectional ray-tracing to a statistical channel model, another channel matrix is computed according to the ITU-R P.1411 rooftop propagation model (ITU, 2003). Three different performance parameters are investigated for both channel models:

1. Average downlink data rate with RZF precoding, which is calculated according to

$$\mathbf{A} = \mathbf{G}^H \left(\mathbf{G}\mathbf{G}^H + \alpha \mathbf{I}\right)^{-1},$$
$$\mathrm{SINR}_u = \frac{|\mathbf{G}_{(u,:)}\mathbf{A}_{(:,u)}|^2}{\sum\limits_{v=1,v\neq u}^{10} |\mathbf{G}_{(u,:)}\mathbf{A}_{(:,v)}|^2 + \sigma_n^2},\ u = 1, \cdots, 16,$$
$$R^{\mathrm{avg}} = \frac{W}{10}\sum_{u=1}^{10} \log_2\left(1 + \mathrm{SINR}_u\right), \tag{6.1}$$

   where σ denotes the noise power, $\mathbf{G}$ and $\mathbf{A}$ are the channel and precoding matrices, respectively.
2. Average signal ($\mathcal{S}$), interference ($\mathcal{I}$) and noise power (σ_n^2) (at the receiving antenna ports) given by

$$\mathcal{S} = \frac{1}{10}\sum_{u=1}^{10} |\mathbf{G}_{(u,:)}\mathbf{A}_{(:,u)}|^2,\ \mathcal{I} = \frac{1}{10}\sum_{u=1}^{10}\left[\sum_{v=1,v\neq u}^{10} |\mathbf{G}_{(u,:)}\mathbf{A}_{(:,v)}|^2\right],$$
$$\sigma_n^2 = 10^{-20.4+\log_{10}(W)}, \tag{6.2}$$

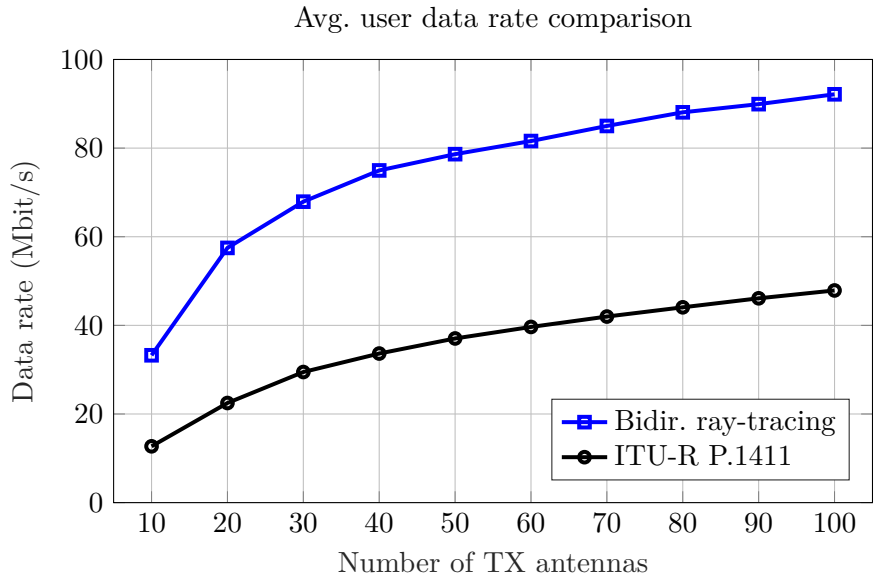
   where W is the channel bandwidth (defined as 10 MHz).
3. Channel similarity factor ($\mathcal{U}$) defined by

$$\mathbf{v}_u = \frac{\mathbf{G}_{(u,:)}}{||\mathbf{G}_{(u,:)}||},\ \mathbf{v}_v = \frac{\mathbf{G}_{(v,:)}}{||\mathbf{G}_{(v,:)}||},$$
$$\mathcal{U} = \frac{1}{90}\sum_{u=1}^{10}\sum_{v=1,v\neq u}^{10} \langle \mathbf{v}_u, \mathbf{v}_v \rangle. \tag{6.3}$$

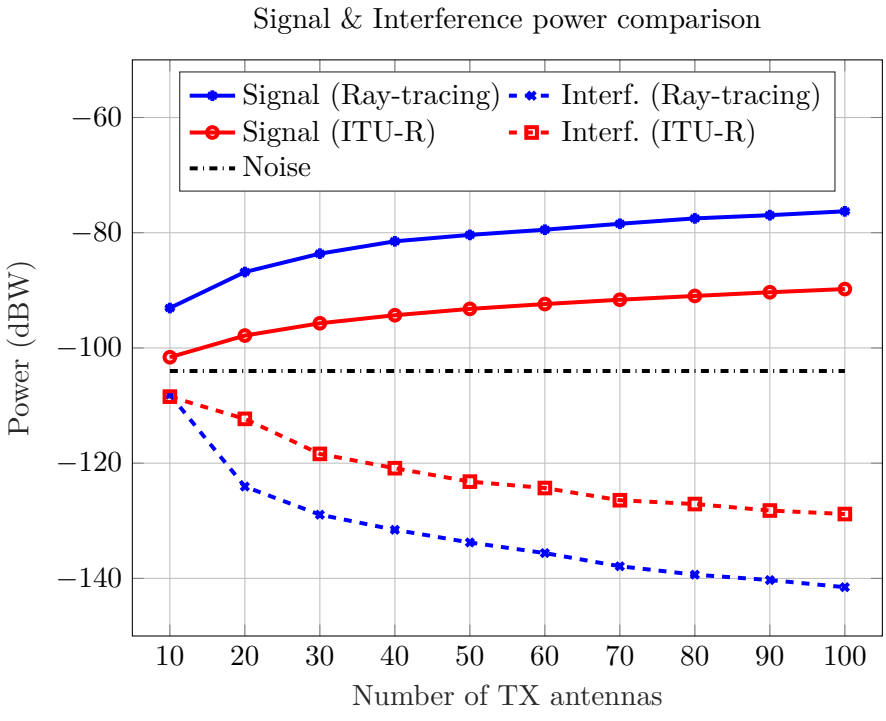
The variations of the parameters is investigated as the number of the transmitting antennas at the base station was changed from 10 to 100. The results are shown in Fig. 6.3.

The data rate results indicate that there is a significant difference between the ITU model and bidirectional ray-tracing approach. The improvements in data rate (as the number of base station antennas is increased) is relatively limited according to the ITU model, compared to bidirectional ray-tracing. The channel similarity/orthogonality characteristics are comparable for both models (see Fig. 6.3(c)), as the directional properties of the propagation environment are represented sufficiently well with both approaches. On the other hand, signal and interference power are considerably distinct where the difference in signal power easily exceeds 10 dB in certain cases. Thus, it can be inferred that the disparity in data rate performance arises from the difference in signal power where the channel gain is likely to be overestimated with the bidirectional ray-tracing,

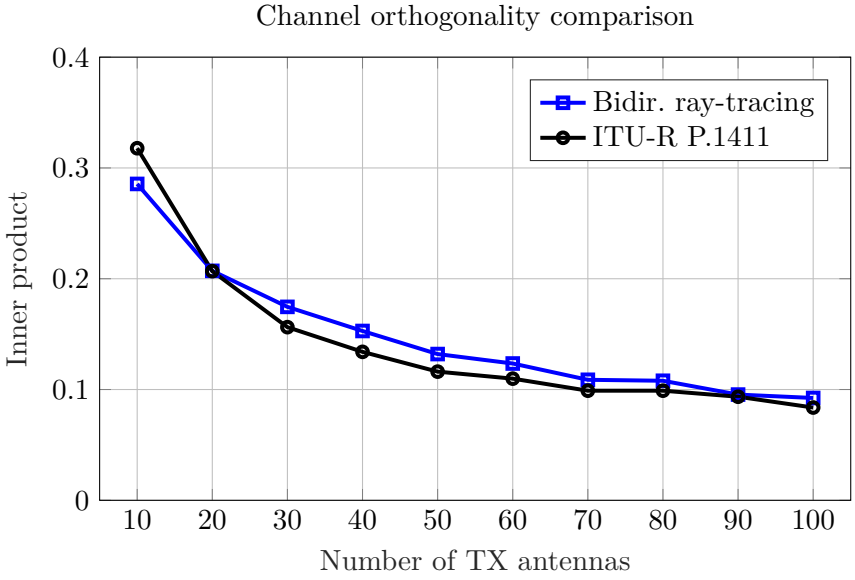
as the path losses in multiple diffraction scenarios are typically underestimated when the number of diffractions becomes larger (see Chapter 3 for the comparisons between the bidirectional ray-tracing and the Vogler method). Note that the difference in the interference power is not relevant for the data rate calculations since the noise is the dominant factor, as it can be seen in Fig. 6.3(b).



(a)



(b)



(c)

Figure 6.3: Comparison of data rate (a), signal-interference levels (b) and channel orthogonality (c) under bidirectional ray-tracing and ITU channel models for the rooftop diffraction scenario (Taygur *et al.*, 2019b). Reproduced courtesy of The Electromagnetics Academy.

6.2 Investigations on Millimeter-Wave Massive MIMO Scenarios

Millimeter-wave communication systems generally operate at carrier frequencies above 20 GHz. Since the frequency spectrum in this region is not as congested as at lower frequencies (i.e., sub-6 GHz band), relatively large bandwidths can be allocated, which yield high data rates (Busari *et al.*, 2018). Nevertheless, a typical problem in millimeter-wave communication links, the so-called path loss issue, affects the performance as the frequency becomes higher[1]. The problem can be mitigated by utilizing large arrays, which provide a high gain and are coherent with the massive MIMO concept, where the base station consists of a large number of antennas (Taygur and Eibert, 2021b).

As it was demonstrated previously, mmWave scenarios can be simulated efficiently by using bidirectional ray-tracing with asymptotic expansion, where a closed interaction surface is placed around the receivers.

6.2.1 Methodology

Urban Scenario

The urban scenario consists of 16 buildings arranged as a 4×4 grid. A permittivity of $\epsilon_r = 4$ is utilized everywhere in the geometry. A base station, which consists of 256

[1]A more accurate interpretation of this phenomenon is that the effective antenna aperture typically shrinks at higher frequencies (rather than an explicit increase in path attenuation) as the size of antennas is generally proportional to the wavelength.

antennas in 16×16 cylindrical array formation, is placed in the middle of the geometry, 5 m above the ground level. The number of user terminals (having a single antenna) is 16, and they are placed randomly at 1 m above the ground level where the user terminals and the base station have a distance of minimum 20 m. All the antennas in the scenario are assumed to have isotropic patterns with vertical polarization. Cubic interaction surfaces with side length of 2 m, are placed around the user terminals. Three different carrier frequency and array sizes (i.e., spacing of the antenna elements in the array) configurations are considered:

1. 28 GHz carrier with λ distance between antenna elements.
2. 7 GHz carrier with λ distance between antenna elements.
3. 7 GHz carrier with $\lambda/4$ distance between antenna elements.

Here, λ denotes the wavelength at the corresponding carrier frequency. Note that the array sizes are identical for cases 1 and 3, whereas the case 2 involves a larger array. Furthermore, two distinct channel bandwidth configurations are considered, 100 MHz bandwidth at 7 GHz and 100-400 MHz bandwidth at 28 GHz, where the array has an inter-element spacing of λ in both cases. An illustration of the scenario is given in Fig. 6.4.

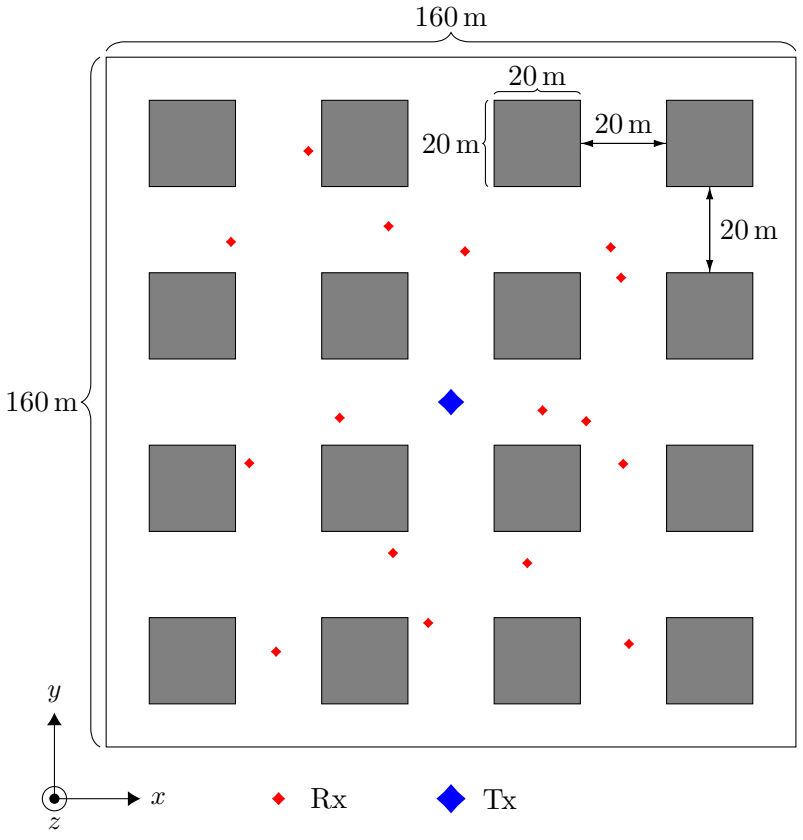


Figure 6.4: Urban scenario 2D top view illustration.

Table 6.1: Simulation parameters for the convergence analysis of millimeter-wave urban massive MIMO scenario.

Cases	Number of ray launches	Number of reflections	Cases	Number of ray launches	Number of reflections
Case 1	50,000	1	Case 9	12,000,000	3
Case 2	100,000	2	Case 10	15,000,000	5
Case 3	200,000	2	Case 11	18,000,000	5
Case 4	500,000	2	Case 12	21,000,000	5
Case 5	1,000,000	3	Case 13	24,000,000	7
Case 6	3,000,000	3	Case 14	27,000,000	7
Case 7	6,000,000	3	Case 15	30,000,000	7
Case 8	9,000,000	3			

6.2.2 Numerical Results

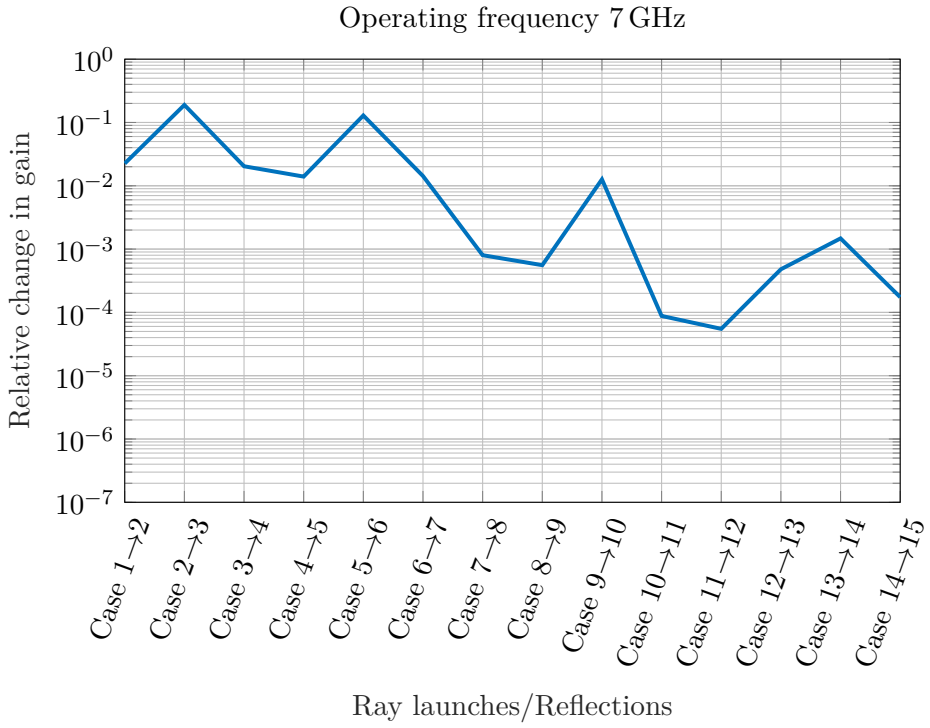
Convergence Study

In order to find the ray launch and interaction-related (i.e., maximum number of reflections) parameters, which yield convergence, a parametric sweep study is carried out. The variation of the average channel gain (channels between a single base station antenna and 16 user terminals) is used as the convergence criterion at 7 GHz and 28 GHz. The ray launches from the user terminals are kept constant at 50,000 whereas the ray launches from the transmitter site are increased from 50,000 to 30,000,000, and the reflections are simultaneously changed from 1 to 7. Note that there are no UTD based diffraction calculations involved, as the diffraction effects are taken into account implicitly via bidirectional ray-tracing where certain wavefront pairs yield a critical point of second kind on the interaction surface. In total 15 different configurations were considered where the simulation parameters for each case is given in Table 6.1. The variation of the average channel gain with respect to the previous step, is shown in Fig. 6.5.

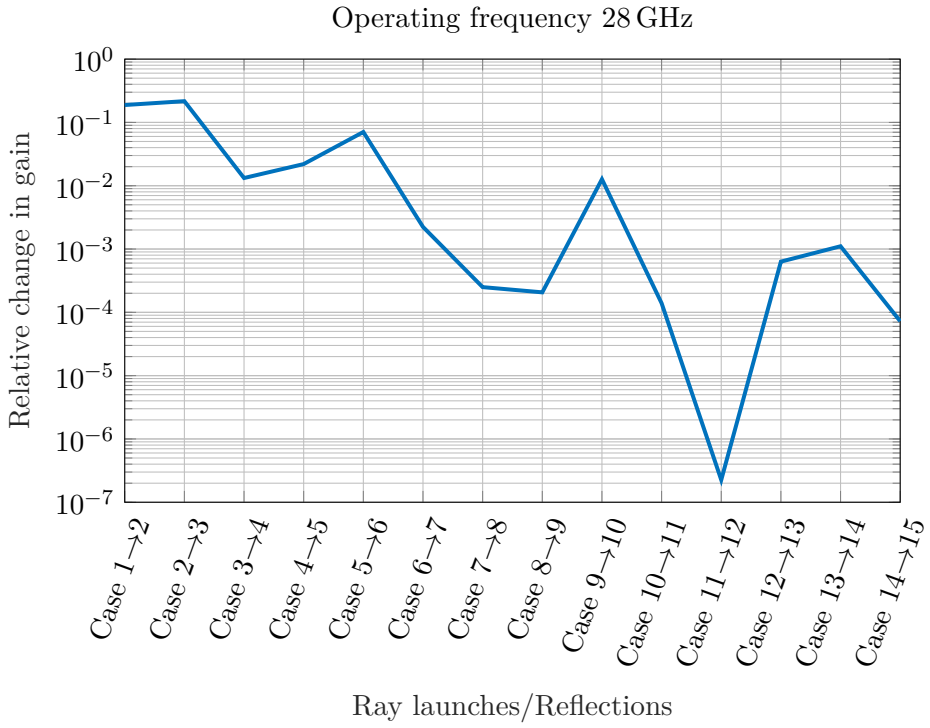
The results indicate that the relative change in channel gain becomes relatively small beyond 15 million ray launches and 5 reflections (Case 10), where the difference is below 2% for both 7 GHz and 28 GHz. Nevertheless, a more stringent convergence criterion can be attained with 30 million ray launches and 7 reflections, where the relative change is less than 0.02% for both frequencies.

Simulation Process

The bidirectional ray-tracing simulations are utilized to obtain the channel matrix and to compute the average downlink data rate accordingly. The performance under two different precoding approaches, RZF and MRC, is investigated where the number of base station antennas is changed from 16 to 256. Three distinct analyses were performed:



(a)



(b)

Figure 6.5: Variation of relative difference in channel gain for 7 GHz (a) and 28 GHz (b) operating frequencies.

1. Average downlink data rate with MRC and RZF precoding using 100 MHz channel bandwidth and following carrier frequency/array configurations:
   Configuration A: 28 GHz carrier with λ array element spacing.
   Configuration B: 7 GHz carrier with λ array element spacing.
   Configuration C: 7 GHz carrier with $\lambda/4$ array element spacing.
2. Average signal, interference and noise power characterization of each case from the first analysis.
3. Average downlink data rate with MRC and RZF precoding using 100 MHz channel bandwidth with Configuration B and 100-400 MHz with Configuration A.

The data rates (R^{avg}) under both precoding schemes are given as

$$\mathbf{A} = \begin{cases} \mathbf{G}^H \left(\mathbf{G}\mathbf{G}^H + \alpha \mathbf{I}\right)^{-1}, & \text{RZF}, \\ \mathbf{G}^H, & \text{MRC}, \end{cases}$$

$$\text{SINR}_u = \frac{|\mathbf{G}_{(u,:)}\mathbf{A}_{(:,u)}|^2}{\sum\limits_{v=1, v\neq u}^{16} |\mathbf{G}_{(u,:)}\mathbf{A}_{(:,v)}|^2 + \sigma_n^2}, \; u = 1, \cdots, 16,$$

$$R^{\text{avg}} = \frac{W}{16} \sum_{u=1}^{16} \log_2 \left(1 + \text{SINR}_u\right), \tag{6.4}$$

where σ_n^2 denotes the noise power, W is the allocated channel bandwidth, $\mathbf{G}$ and $\mathbf{A}$ are the channel and precoding matrices, respectively. The signal ($\mathcal{S}$), interference ($\mathcal{I}$) and noise power (σ_n^2) expressions can be written as

$$\mathcal{S} = \frac{1}{16} \sum_{u=1}^{16} |\mathbf{G}_{(u,:)}\mathbf{A}_{(:,u)}|^2, \; \mathcal{I} = \frac{1}{16} \sum_{u=1}^{16} \left[\sum_{v=1, v\neq u}^{16} |\mathbf{G}_{(u,:)}\mathbf{A}_{(:,v)}|^2 \right],$$

$$\sigma_n^2 = 10^{-20.4 + \log_{10}(W)}. \tag{6.5}$$

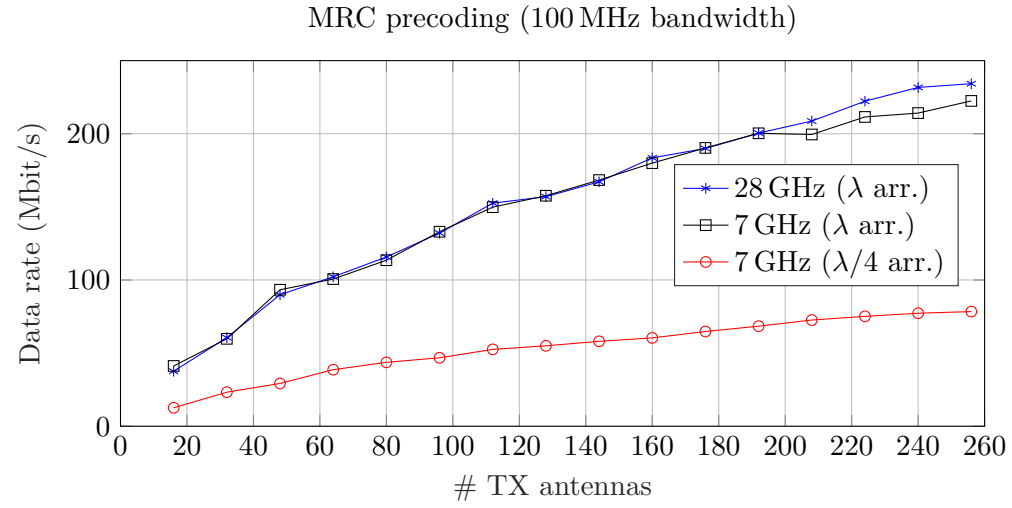
The results are shown in Figs. 6.6-6.8.

The data rate analysis in Fig. 6.6 shows that RZF precoding yields a much better data rate performance compared to MRC. The Configuration B shows the best performance with both precoders, whereas the Configuration A and C show similar characteristics under RZF but the latter is noticeably outperformed under MRC.

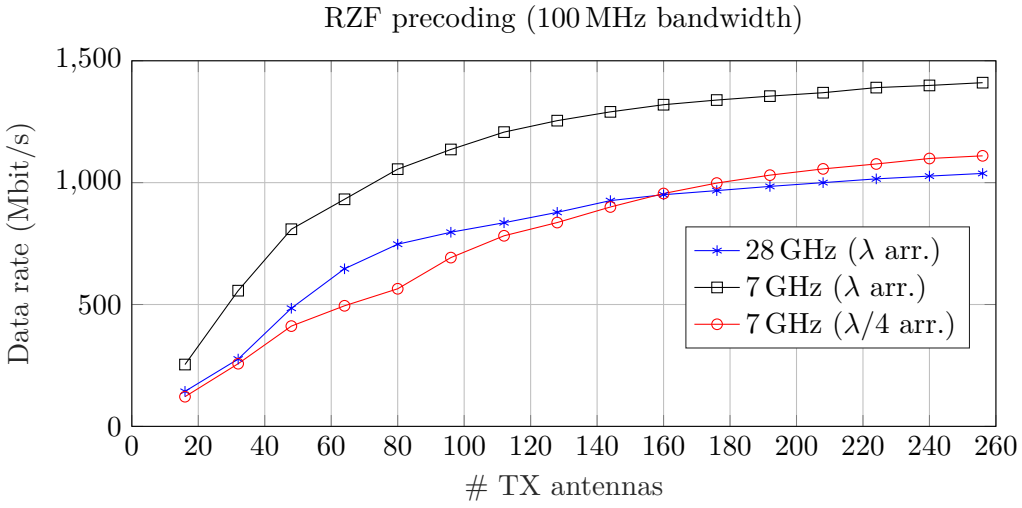
The differences in performance can be attributed to distinct SINR characteristics, which are presented in Fig. 6.7. It can be noticed that the Configurations A and B yield very similar signal and interference characteristics under MRC precoding, although the power level is considerably smaller in Configuration A due to the path loss effect. The interference cannot be suppressed well enough when the element spacing is reduced to $\lambda/4$ and MRC precoding is employed, therefore a poor SINR performance and an accordingly low data rate is observed. The RZF precoder yields drastically different SINR characteristics, where the noise power (constant noise spectral density is presumed) is dominant compared to the interference, hence, the data rate performance is mainly determined by the signal power. A particularly intriguing situation occurs with Configuration C, as the signal power is comparable to that in Configuration A (note that the actual physical dimensions of both arrays are identical). However, the sub-optimal array spacing of $\lambda/4$

results into a poor performance here, even though the path loss effect is not as strong as in 28 GHz.

The use of larger bandwidths yields significant improvements in data rate, as shown in Fig. 6.8, as the performance scales almost linearly with both precoding methods. An important observation is that the linear scaling characteristics of the data rate (with respect to the allocated channel bandwidth) persists independent of the number of base station antennas. Similar to the previously presented results, the RZF precoder yields much better data rate performance compared to the MRC. However, it should be noted that RZF precoding is computationally more complex than MRC in general, therefore, the latter approach might be a better option if there are many more users (and an accordingly large base station array with many more antennas) in the scenario.

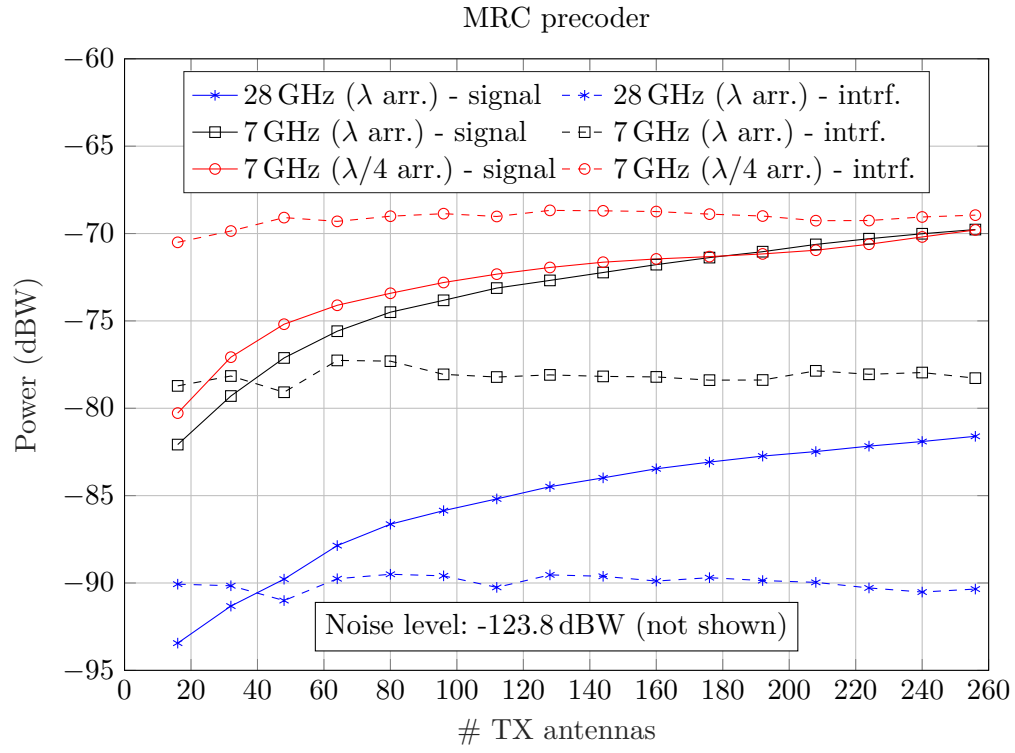


(a)

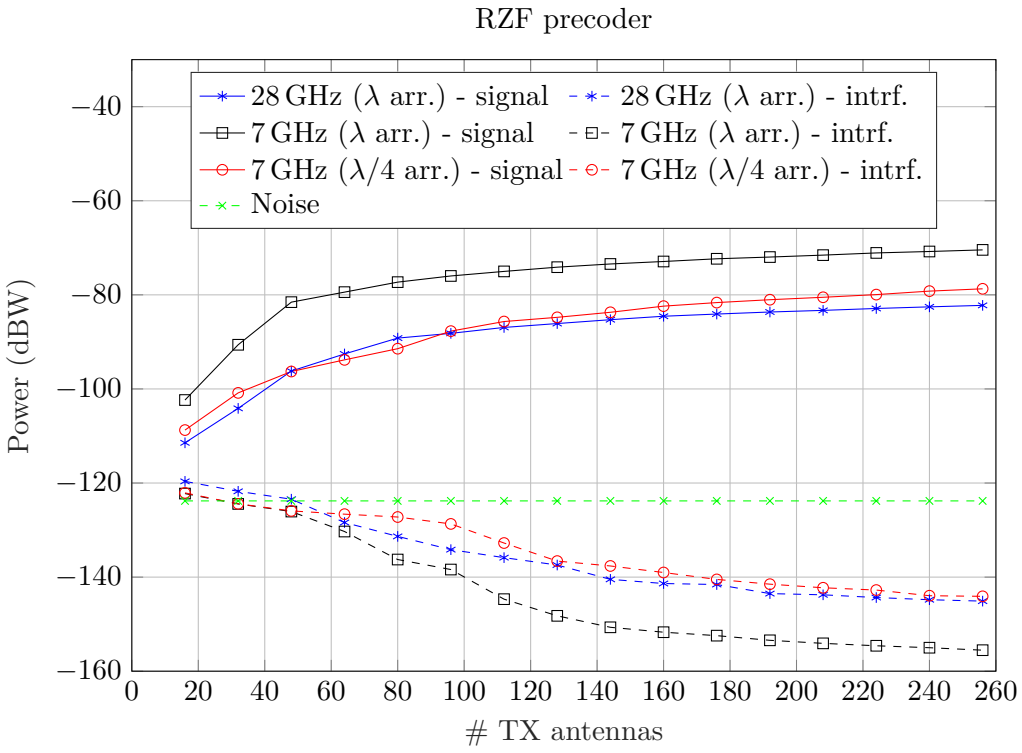


(b)

Figure 6.6: Comparison of data rate for different array configurations with 100 MHz bandwidth under MRC (a) and RZF (b) precoding schemes (Taygur and Eibert, 2021b). © IEEE 2021.

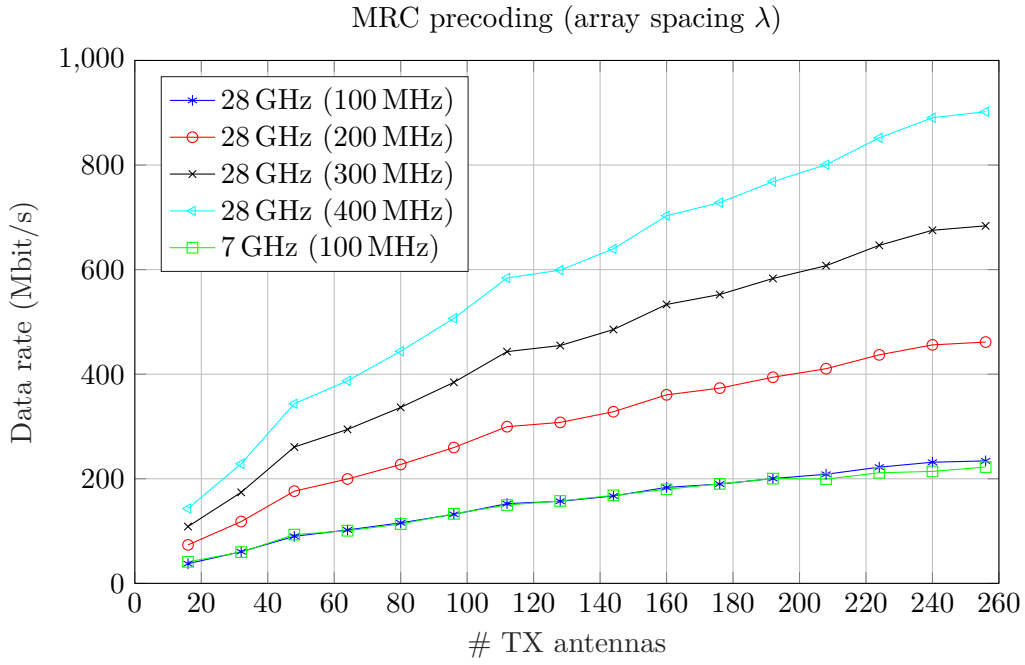


(a)

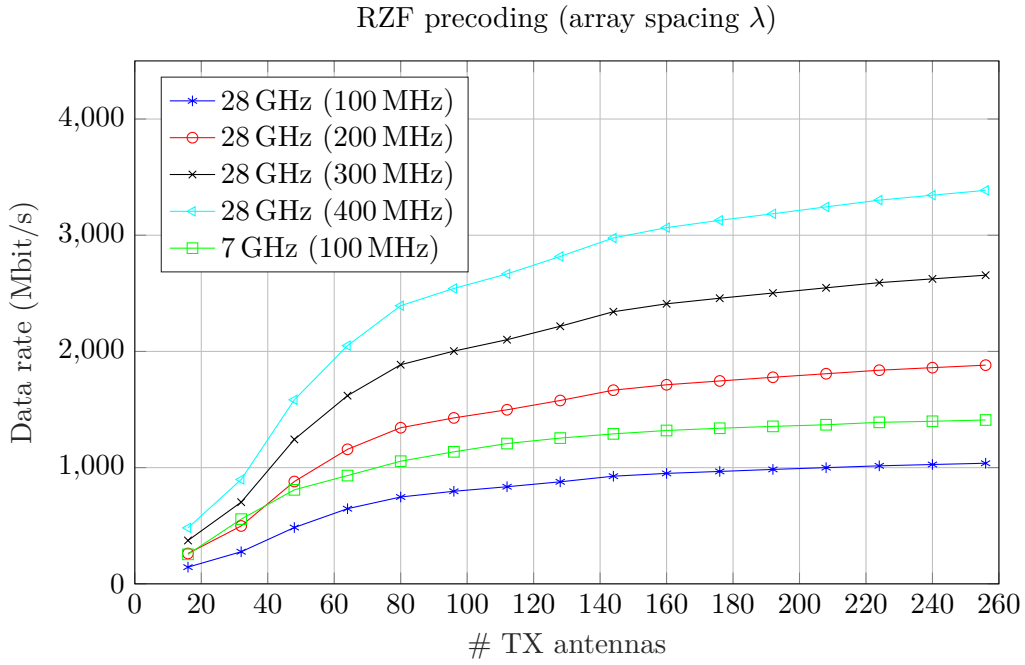


(b)

Figure 6.7: Comparison of signal and interference power levels under (a) MRC and (b) RZF precoding schemes.



(a)



(b)

Figure 6.8: Comparison of data rate for different array configurations with λ spacing under MRC (a) and RZF (b) precoding schemes (Taygur and Eibert, 2021b). © IEEE 2021.

Chapter 7

Conclusion

Investigations on a novel ray-tracing approach, namely, the bidirectional ray-tracing, and channel analyses of various massive MIMO scenarios based on bidirectional ray-tracing as well as conventional unidirectional ray-tracing were presented. The new approach addresses certain shortcomings of traditional unidirectional ray-tracing approach, which may yield inaccurate results or become computationally inefficient in certain cases.

The advantages of bidirectional ray-tracing over conventional unidirectional ray-tracing were clearly seen, especially in the scenarios with dominant diffraction paths. Since reception spheres and diffraction cylinders (placed on wedges) were eliminated, these effects were incorporated by means of a PO-like integration procedure, i.e., numerical evaluation of the reciprocity integral on an interaction surface, instead of UTD-based calculations. The benefits of this approach are twofold. First, a smaller number of rays are utilized in the simulations compared to the unidirectional ray-tracing, since interaction surfaces can typically be created more flexibly in terms of size and shape, compared to the reception spheres and diffraction cylinders. Second, the accuracy problems of UTD in certain scenarios, i.e., multiple diffraction problems where the wave propagation occurs near the optical boundaries, were avoided. In such cases, the interaction surface was set up as a large, open surface, which separates the receiver and transmitter.

It is possible to utilize an interaction surface in closed form as well, where the surface is typically placed around a receiving antenna. Such a configuration is generally feasible when the diffraction paths are not dominant in the scenario, i.e., line of sight or reflection/refraction paths constitute the majority of the contributions. A closed surface around the receiving antenna can be considered as a significant improvement over the reception spheres, as a good accuracy was achieved by a smaller number of rays compared to the unidirectional ray-tracing.

The bidirectional ray-tracing was also utilized in millimeter-wave scenarios, where the reciprocity integral was evaluated by means of a stationary phase approximation. Hence, the integral result was computed with a single algebraic expression, as the majority of the contributions come from a small region near the stationary phase point. This approach helps to deal with the phase errors, which may frequently occur with unidirectional ray-tracing due to inexact ray path length calculations. The numerical results have shown that the ray path lengths (corresponding to exact rays) as well as the phase variations were identified accurately by means of this bidirectional ray-tracing approach. The accuracy problems due to missed or incorrect ray contributions, which emerge from

a poor selection of the reception sphere size, are also avoided.

Both unidirectional and bidirectional ray-tracing approaches were utilized to characterize the downlink data rate in various urban massive MIMO scenarios, which involved building grids. The buildings had usable indoor spaces in certain scenarios where users could be placed. Thus, separate analyses could be performed for indoor and outdoor users. The unidirectional ray-tracing was utilized for the investigation of pilot contamination, channel aging, frequency-selective fading, and multiple antenna user terminal scenarios, whereas the bidirectional ray-tracing was utilized to characterize the performance in rooftop diffraction and millimeter-wave scenarios. Comparisons with applicable statistical channel models were given where relevant, and the fundamental differences of ray-tracing and statistical models were highlighted.

The results of massive MIMO analyses generally indicate that an improvement in downlink data rate can be achieved by increasing the number of base station antennas. However, the improvements in performance gradually diminish as the number of user antennas becomes significantly smaller than the number of base station antennas. The effect is more noticeable with ray-tracing compared to the statistical models, as the directional properties of the channel are represented more accurately with ray-tracing. Here, a notable problem of many statistical models is that the individual channel coefficients are defined as independent random variables, unless a channel correlation factor is explicitly introduced. Hence, adding new antennas into the base station commonly yields uncorrelated sub-channels and an accordingly high data rate, which is generally not realistic.

The adverse effects of pilot contamination, channel aging or frequency selective fading may not completely be avoided by increasing the number of base station antennas, i.e., the performance remains limited regardless of the size of the base station array. A similar situation was also observed in the multi-antenna user terminal and millimeter-wave propagation scenarios where a better data rate was achieved by adding more antennas into the base station but the performance was still hindered in certain cases, i.e., if the precoding scheme is sub-optimal or the antenna spacing in the array is significantly smaller than a wavelength.

The bidirectional ray-tracing method can be improved in several ways. Considering the similarities with the conventional physical optics approach, the simulation accuracy in single diffraction scenarios can be improved by introducing equivalent edge currents, in a similar way as Physical Theory of Diffraction (PTD) is utilized to improve the accuracy of physical optics simulations. Such an approach can eliminate the need for a correction factor, which was introduced to compensate the numerical differences between the UTD and the bidirectional ray-tracing algorithm.

For multiple diffraction problems, the method can be modified such that a surface is placed above every edge where they are utilized as Huygens surfaces, i.e., equivalent sources are computed by means of the incident rays on each surface, then new rays are launched according to this equivalent source representation. The process continues iteratively until an interaction surface is reached. This approach can be helpful to simulate complex multiple diffraction scenarios where a single interaction surface cannot directly be hit by the receiver or transmitter, especially if they are in deep shadow region. Here, the Huygens surfaces are utilized to improve the visibility between the sources and the interaction surface.

Although the massive MIMO analyses in this dissertation provide a good insight on

the achievable downlink data rate in a simple urban-like environment as well as on the differences between ray-tracing simulations and statistical channel modeling approaches, they can be improved in many ways. It should be noted that the channel-related phenomena were investigated in an isolated manner here. In other words, no two phenomena were simultaneously considered in the scenarios, i.e., investigations on pilot contamination did not involve any channel aging effect or frequency selective fading. Thus, it can be inferred that more realistic results can be obtained by taking several channel-related phenomena into account simultaneously. Another important factor, which may improve the simulation accuracy, is the level of detail of the geometry. Even though a simple building grid may give a good idea about the general propagation characteristics in similar urban-like environments, the problem geometry should be depicted in a more detailed and precise manner if accurate results are needed for a particular scenario. Note that the simulation of such detailed geometries might be computationally expensive, however, it can be predicted that the developments in the hardware technology and new computing paradigms will enable shorter simulation times in the future.

Appendix A

Derivation of Geometrical Optics Postulates

A.1 Eikonal Equation

The solution of the vector Helmholtz equation, which is given in the form of

$$\begin{aligned}\nabla^2\mathbf{E}(\mathbf{r},\omega)+\omega^2\mu\epsilon\mathbf{E}(\mathbf{r},\omega)=0,\\ \nabla^2\mathbf{H}(\mathbf{r},\omega)+\omega^2\mu\epsilon\mathbf{H}(\mathbf{r},\omega)=0,\\ k=\omega\sqrt{\mu\epsilon}=\frac{2\pi}{\lambda},\end{aligned} \tag{A.1}$$

for a source-free region and a homogeneous, isotropic medium, can be given by the Luneburg-Kline approximation where the electric and magnetic field is expressed as a series sum by (Kline, 1962)

$$\begin{aligned}\mathbf{E}(\mathbf{r},\omega)=e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u},\\ \mathbf{H}(\mathbf{r},\omega)=e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^u},\\ k_0=\omega\sqrt{\mu_0\epsilon_0},\end{aligned} \tag{A.2}$$

where $\epsilon_0 = 8.854\times10^{-12}$ and $\mu_0 = 4\pi\times10^{-7}$ are the permittivity and permeability in free space, respectively. Plugging the electric and magnetic field expressions into Maxwell's equations yields

$$\begin{aligned}\nabla\times\left(e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u}\right)=-j\omega\mu e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^u},\\ \nabla\times\left(e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^u}\right)=j\omega\epsilon e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u},\end{aligned} \tag{A.3}$$

with

$$\nabla \cdot \left(e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u} \right) = 0,$$
$$\nabla \cdot \left(e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^u} \right) = 0. \tag{A.4}$$

Using the identities

$$\begin{aligned} \nabla \times (f\mathbf{F}) &= \nabla f \times \mathbf{F} + f(\nabla \times \mathbf{F}), \\ \nabla \cdot (f\mathbf{F}) &= \mathbf{F} \cdot \nabla f + f \nabla \cdot \mathbf{F}, \\ \nabla e^{-jk_0 S(\mathbf{r})} &= -jk_0 e^{-jk_0 S(\mathbf{r})} \nabla S(\mathbf{r}), \end{aligned} \tag{A.5}$$

Eq. (A.4) can be re-written as (McNamara *et al.*, 1990)

$$e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \left[\frac{\nabla \times \mathbf{E}_u(\mathbf{r})}{(j\omega)^u} - \frac{(1/\nu)(\nabla S(\mathbf{r}) \times \mathbf{E}_u(\mathbf{r}))}{(j\omega)^{u-1}} \right] = -\mu e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^{u-1}},$$
$$e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \left[\frac{\nabla \times \mathbf{H}_u(\mathbf{r})}{(j\omega)^u} - \frac{(1/\nu)(\nabla S(\mathbf{r}) \times \mathbf{H}_u(\mathbf{r}))}{(j\omega)^{u-1}} \right] = \epsilon e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^{u-1}},$$
$$e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \left[\frac{(1/\nu)(\nabla S(\mathbf{r}) \cdot \mathbf{E}_u(\mathbf{r}))}{(j\omega)^{u-1}} - \frac{\nabla \cdot \mathbf{E}_u(\mathbf{r})}{(j\omega)^u} \right] = 0,$$
$$e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \left[\frac{(1/\nu)(\nabla S(\mathbf{r}) \cdot \mathbf{H}_u(\mathbf{r}))}{(j\omega)^{u-1}} - \frac{\nabla \cdot \mathbf{H}_u(\mathbf{r})}{(j\omega)^u} \right] = 0, \tag{A.6}$$

with

$$\nu = \frac{\omega}{k_0} = \frac{\omega}{\omega\sqrt{\mu_0 \epsilon_0}} = \frac{1}{\sqrt{\mu_0 \epsilon_0}},$$
$$\epsilon = \epsilon_0 \epsilon_r, \ \mu = \mu_0 \mu_r. \tag{A.7}$$

For the asymptotic case where $\omega \to \infty$, Eq. (A.6) can be simplified into (Born and Wolf, 1999)

$$\begin{aligned} \frac{1}{\nu\mu}(\nabla S(\mathbf{r}) \times \mathbf{E}_0(\mathbf{r})) &= \mathbf{H}_0(\mathbf{r}), \\ \frac{-1}{\nu\epsilon}(\nabla S(\mathbf{r}) \times \mathbf{H}_0(\mathbf{r})) &= \mathbf{E}_0(\mathbf{r}), \\ \nabla S(\mathbf{r}) \cdot \mathbf{E}_0(\mathbf{r}) &= 0, \\ \nabla S(\mathbf{r}) \cdot \mathbf{H}_0(\mathbf{r}) &= 0. \end{aligned} \tag{A.8}$$

Inserting the cross product terms in Eq. (A.8) within each other yields

$$\frac{1}{\nu\mu}\left[\nabla S(\mathbf{r}) \times \left(\left(\frac{-1}{\nu\epsilon} \right) \nabla S(\mathbf{r}) \times \mathbf{H}_0(\mathbf{r}) \right) \right] = \mathbf{H}_0(\mathbf{r}),$$
$$\frac{-1}{\nu\epsilon}\left[\nabla S(\mathbf{r}) \times \left(\left(\frac{1}{\nu\mu} \right) \nabla S(\mathbf{r}) \times \mathbf{E}_0(\mathbf{r}) \right) \right] = \mathbf{E}_0(\mathbf{r}). \tag{A.9}$$

Using the identity

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B} (\mathbf{A} \cdot \mathbf{C}) - \mathbf{C} (\mathbf{A} \cdot \mathbf{B}), \tag{A.10}$$

Eq. (A.9) can be re-written as

$$\begin{aligned} \frac{-1}{\nu^2 \epsilon \mu} \left[\nabla S (\mathbf{r}) (\nabla S (\mathbf{r}) \cdot \mathbf{H}_0 (\mathbf{r})) - \mathbf{H}_0 (\mathbf{r}) (\nabla S (\mathbf{r}) \cdot \nabla S (\mathbf{r}))\right] &= \mathbf{H}_0 (\mathbf{r}) \\ = \mathbf{H}_0 (\mathbf{r}) |\nabla S (\mathbf{r})|^2 &= \epsilon_r \mu_r \mathbf{H}_0 (\mathbf{r}), \\ \frac{-1}{\nu^2 \epsilon \mu} \left[\nabla S (\mathbf{r}) (\nabla S (\mathbf{r}) \cdot \mathbf{E}_0 (\mathbf{r})) - \mathbf{E}_0 (\mathbf{r}) (\nabla S (\mathbf{r}) \cdot \nabla S (\mathbf{r}))\right] &= \mathbf{E}_0 (\mathbf{r}) \\ = \mathbf{E}_0 (\mathbf{r}) |\nabla S (\mathbf{r})|^2 &= \epsilon_r \mu_r \mathbf{E}_0 (\mathbf{r}), \end{aligned} \tag{A.11}$$

which implies

$$|\nabla S (\mathbf{r})|^2 = \epsilon_r \mu_r = \eta^2. \tag{A.12}$$

where η denotes the refractive index of the medium.

A.2 Direction of Power Flow

The Poynting vector, which indicates the direction of power flow, can be found by

$$\mathbf{S} = \frac{1}{2} (\mathbf{E}_0 \times \mathbf{H}_0^*). \tag{A.13}$$

Using Eq. (A.8) and vector identity in (A.10) , it is possible to express $\mathbf{S}$ as

$$\begin{aligned} \mathbf{S} &= \mathbf{E}_0 (\mathbf{r}) \times \left(\frac{1}{2\nu\mu} (\nabla S (\mathbf{r}) \times \mathbf{E}_0 (\mathbf{r})) \right)^* \\ &= \frac{1}{2\nu\mu} (\mathbf{E}_0 (\mathbf{r}) \times (\nabla S (\mathbf{r}) \times \mathbf{E}_0 (\mathbf{r}))^*) \\ &= \frac{1}{2\nu\mu} \left[(\nabla S (\mathbf{r})) (\mathbf{E}_0 (\mathbf{r}) \cdot \mathbf{E}_0^* (\mathbf{r})) - \mathbf{E}_0^* (\mathbf{r}) (\mathbf{E}_0 (\mathbf{r}) \cdot \nabla S (\mathbf{r})) \right] \\ &= \frac{1}{2\nu\mu} |\mathbf{E}_0 (\mathbf{r})|^2 \nabla S (\mathbf{r}), \end{aligned} \tag{A.14}$$

which indicates that the direction of power flow is given by $\hat{\mathbf{s}} = \nabla S (\mathbf{r}) / \eta$, and it is orthogonal to both electric and magnetic field vectors according to Eq. (A.8). It is also possible to state the explicit relation between electric and magnetic fields using Eq. (A.8) by

$$\begin{aligned} \frac{\nabla S (\mathbf{r})}{\eta} \times \mathbf{E}_0 (\mathbf{r}) &= \nu\mu \frac{\mathbf{H}_0 (\mathbf{r})}{\eta}, \\ \hat{\mathbf{s}} \times \mathbf{E}_0 (\mathbf{r}) &= \frac{\omega}{\omega \sqrt{\mu_0 \epsilon_0}} \mu_0 \mu_r \frac{\mathbf{H}_0 (\mathbf{r})}{\sqrt{\mu_r \epsilon_r}}, \\ \hat{\mathbf{s}} \times \mathbf{E}_0 (\mathbf{r}) &= \mathbf{H}_0 (\mathbf{r}) \frac{\sqrt{\mu_0 \mu_r}}{\sqrt{\epsilon_0 \epsilon_r}} = \mathbf{H}_0 (\mathbf{r}) \sqrt{\frac{\mu}{\epsilon}} = \mathbf{H}_0 (\mathbf{r}) Z_w, \end{aligned} \tag{A.15}$$

where Z_w denotes the wave impedance.

A.3 Transport Equation

Inserting the electric and magnetic field expressions given in Eq. (A.2) into the Helmholtz equation yields

$$\nabla^2\left(e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u}\right)+k^2e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u}=0,$$
$$\nabla^2\left(e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^u}\right)+k^2e^{-jk_0S(\mathbf{r})}\sum_{u=0}^{\infty}\frac{\mathbf{H}_u(\mathbf{r})}{(j\omega)^u}=0. \tag{A.16}$$

Let us concentrate solely on the first equation of (A.16) which involves the electric field, for the sake of brevity, as the intermediary derivation steps do not differ for electric and magnetic field. Using the identities

$$\nabla^2(f\mathbf{F})=\mathbf{F}\nabla^2f+2(\nabla f\cdot\nabla)\mathbf{F}+f\nabla^2\mathbf{F},$$
$$\nabla\cdot(f\mathbf{F})=f\nabla\cdot\mathbf{F}+\mathbf{F}\cdot\nabla f,$$
$$\nabla^2f=\nabla\cdot(\nabla f), \tag{A.17}$$

and defining

$$\tilde{\mathbf{E}}_u=\sum_{u=0}^{\infty}\frac{\mathbf{E}_u(\mathbf{r})}{(j\omega)^u}, \tag{A.18}$$

Eq. (A.16) can be expressed as

$$\tilde{\mathbf{E}}_u\left(\nabla\cdot\nabla e^{-jk_0S(\mathbf{r})}\right)+2\left(\nabla e^{-jk_0S(\mathbf{r})}\cdot\nabla\right)\tilde{\mathbf{E}}_u+e^{-jk_0S(\mathbf{r})}\left(\nabla^2\tilde{\mathbf{E}}_u\right)+k^2e^{-jk_0S(\mathbf{r})}\tilde{\mathbf{E}}_u=0$$

$$=\tilde{\mathbf{E}}_u\left(\nabla\cdot\left(-jk_0e^{-jk_0S(\mathbf{r})}\nabla S(\mathbf{r})\right)\right)+2\left(\left(-jk_0e^{-jk_0S(\mathbf{r})}\nabla S(\mathbf{r})\right)\cdot\nabla\right)\tilde{\mathbf{E}}_u$$
$$+e^{-jk_0S(\mathbf{r})}\left(\nabla^2\tilde{\mathbf{E}}_u\right)+k^2e^{-jk_0S(\mathbf{r})}\tilde{\mathbf{E}}_u=0$$

$$=(-jk_0)\tilde{\mathbf{E}}_u\left(e^{-jkS(\mathbf{r})}\nabla\cdot\nabla S(\mathbf{r})+\nabla S(\mathbf{r})\cdot\nabla e^{-jkS(\mathbf{r})}\right)$$
$$+2(-jk)e^{-jkS(\mathbf{r})}(\nabla S(\mathbf{r})\cdot\nabla)\tilde{\mathbf{E}}_u+e^{-jkS(\mathbf{r})}\left(\nabla^2\tilde{\mathbf{E}}_u\right)+k^2e^{-jkS(\mathbf{r})}\tilde{\mathbf{E}}_u=0$$

$$=(-jk_0)\tilde{\mathbf{E}}_u\left(e^{-jk_0S(\mathbf{r})}\nabla^2S(\mathbf{r})+(-jk_0)e^{-jk_0S(\mathbf{r})}|\nabla S(\mathbf{r})|^2\right)$$
$$+2(-jk_0)e^{-jk_0S(\mathbf{r})}(\nabla S(\mathbf{r})\cdot\nabla)\tilde{\mathbf{E}}_u+e^{-jk_0S(\mathbf{r})}\left(\nabla^2\tilde{\mathbf{E}}_u+k^2\tilde{\mathbf{E}}_u\right)=0$$

$$=e^{-jk_0S(\mathbf{r})}\left[k_0^2\left(\eta^2\tilde{\mathbf{E}}_u-\tilde{\mathbf{E}}_u|\nabla S(\mathbf{r})|^2\right)+(-jk_0)\left(\tilde{\mathbf{E}}_u\nabla^2S(\mathbf{r})+2(\nabla S(\mathbf{r})\cdot\nabla)\tilde{\mathbf{E}}_u\right)\right.$$
$$\left.+\nabla^2\tilde{\mathbf{E}}_u\right]=0$$

$$= e^{-jk_0 S(\mathbf{r})} \sum_{u=0}^{\infty} \left[\frac{\left(1/\nu^2\right) \mathbf{E}_u(\mathbf{r}) \left(|\nabla S(\mathbf{r})|^2 - \eta^2 \right)}{(j\omega)^{u-2}} \right.$$
$$\left. - \frac{(1/\nu) \left(\mathbf{E}_u(\mathbf{r}) \nabla^2 S(\mathbf{r}) + 2 (\nabla S(\mathbf{r}) \cdot \nabla) \mathbf{E}_u(\mathbf{r}) \right)}{(j\omega)^{u-1}} + \frac{\nabla^2 \mathbf{E}_u(\mathbf{r})}{(j\omega)^l} \right] = 0. \quad \text{(A.19)}$$

The last equation implies that the coefficients of various powers of $(j\omega)$ should individually be equal to zero, since ω is not a parameter of $\mathbf{E}_u$ or $S(\mathbf{r})$ (McNamara *et al.*, 1990). The term with highest order, which involves $(j\omega)^{u-2}$ at the denominator, will be ignored here since it yields the eikonal equation for $u = 0$. The next highest order term with $(j\omega)^{u-1}$ at the denominator yields

$$\mathbf{E}_0(\mathbf{r}) \nabla^2 S(\mathbf{r}) + 2 (\nabla S(\mathbf{r}) \cdot \nabla) \mathbf{E}_0(\mathbf{r}) = 0, \quad \text{(A.20)}$$

for $l = 0$, and it is commonly known as zeroth-order transport equation. Equivalently, Eq. (A.20) can easily be modified for the magnetic field by

$$\mathbf{H}_0(\mathbf{r}) \nabla^2 S(\mathbf{r}) + 2 (\nabla S(\mathbf{r}) \cdot \nabla) \mathbf{H}_0(\mathbf{r}) = 0. \quad \text{(A.21)}$$

APPENDIX B

Linear Minimum Mean Square Error Estimation

The linear Minimum Mean Square Error (MMSE) estimation problem can be represented by (KAILATH *et al.*, 2000)

$$\mathbf{A}^{\text{MMSE}} = \arg \min_{\mathbf{A}^{\text{MMSE}}} E\left[\left|\left|\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right|\right|^2\right], \tag{B.1}$$

with

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n}, \tag{B.2}$$

where $\mathbf{A}^{\text{MMSE}}$ denotes the linear estimator, $\mathbf{x} \in \mathbb{C}^{M_T}$ is the parameter to be estimated, $\mathbf{y} \in \mathbb{C}^{M_R}$ is the observation, $\mathbf{G} \in \mathbb{C}^{M_R \times M_T}$ is a known matrix and $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}\left(0, \sigma_n^2 \mathbf{I}\right)$ is a noise term. The cost function in the minimization problem can also be expressed as

$$\begin{aligned} E\left[\left|\left|\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right|\right|^2\right] &= E\left[\text{trace}\left(\left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right)\left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right)^H\right)\right] \\ &= \text{trace}\left(E\left[\left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right)\left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right)^H\right]\right) \\ &= \text{trace}\left(\text{cov}\left[\left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right), \left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right)\right]\right). \end{aligned} \tag{B.3}$$

The covariance term can shortly be represented by

$$\mathbf{Q}_{\mathbf{M}} = \text{cov}\left[\left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right), \left(\mathbf{x} - \mathbf{A}^{\text{MMSE}}\mathbf{y}\right)\right]. \tag{B.4}$$

The minimization problem in Eq. (B.1) can be then given by (VERDÚ, 1998)

$$\mathbf{A}^{\text{MMSE}} = \arg \min_{\mathbf{A}^{\text{MMSE}}} \text{trace}\left(\mathbf{Q}_{\mathbf{M}}\right). \tag{B.5}$$

In order to find an explicit solution, the covariance matrix $\mathbf{Q_M}$ can be expanded as

$$\begin{aligned}
\mathbf{Q_M} &= E\left[\left(\mathbf{x}-\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\right)\left(\mathbf{x}-\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\right)^H\right] \\
&= E\left[\left(\mathbf{x}-\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\right)\left(\mathbf{x}^H-\mathbf{y}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H\right)\right] \\
&= E\left[\mathbf{x}\mathbf{x}^H-\mathbf{x}\mathbf{y}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H-\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\mathbf{x}^H+\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\mathbf{y}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H\right] \\
&= E\left[\mathbf{x}\mathbf{x}^H\right]-E\left[\mathbf{x}\mathbf{y}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H\right]-E\left[\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\mathbf{x}^H\right] \\
&\quad +E\left[\mathbf{A}^{\mathrm{MMSE}}\mathbf{y}\mathbf{y}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H\right]. \qquad \text{(B.6)}
\end{aligned}$$

Note that the matrix $\mathbf{A}^{\mathrm{MMSE}}$ is deterministic, hence, can be removed from the expected value operator. Furthermore, $E\left[\mathbf{x}\mathbf{x}^H\right]=\mathbf{Q_x}$, thus, Eq. (B.6) can written as (Cover and Thomas, 2005)

$$\begin{aligned}
\mathbf{Q_M} &= \mathbf{Q_x}-E\left[\mathbf{x}\mathbf{y}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H-\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{y}\mathbf{x}^H\right]+\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{y}\mathbf{y}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H \\
&= \mathbf{Q_x}-E\left[\mathbf{x}\left(\mathbf{x}^H\mathbf{G}^H+\mathbf{n}^H\right)\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H-\mathbf{A}^{\mathrm{MMSE}}E\left[\left(\mathbf{G}\mathbf{x}+\mathbf{n}\right)\mathbf{x}^H\right] \\
&\quad +\mathbf{A}^{\mathrm{MMSE}}E\left[\left(\mathbf{G}\mathbf{x}+\mathbf{n}\right)\left(\mathbf{x}^H\mathbf{G}^H+\mathbf{n}^H\right)\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H. \\
&= \mathbf{Q_x}-E\left[\mathbf{x}\mathbf{x}^H\mathbf{G}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H-E\left[\mathbf{x}\mathbf{n}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H-\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{G}\mathbf{x}\mathbf{x}^H\right] \\
&\quad -\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{n}\mathbf{x}^H\right]+\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{n}\mathbf{n}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H \\
&\quad +\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{G}\mathbf{x}\mathbf{x}^H\mathbf{G}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H+\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{G}\mathbf{x}\mathbf{n}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H \\
&\quad +\mathbf{A}^{\mathrm{MMSE}}E\left[\mathbf{n}\mathbf{x}^H\mathbf{G}^H\right]\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H. \qquad \text{(B.7)}
\end{aligned}$$

Since $\mathbf{G}$ is deterministic, it can be removed from the expected value operator, similar to $\mathbf{A}^{\mathrm{MMSE}}$. Additionally, it should be remembered that $E\left[\mathbf{n}\right]=\mathbf{0}$, $E\left[\mathbf{n}\mathbf{n}^H\right]=\mathbf{I}\sigma_n^2$ and $\mathbf{n}$ is independent of $\mathbf{x}$, i.e., $E\left[\mathbf{n}\mathbf{x}\right]=E\left[\mathbf{n}\right]E\left[\mathbf{x}\right]$, hence, Eq. (B.7) can be simplified into (Bertsekas and Tsitsiklis, 2008)

$$\begin{aligned}
\mathbf{Q_M} &= \mathbf{Q_x}-\mathbf{Q_x}\mathbf{G}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H-\mathbf{A}^{\mathrm{MMSE}}\mathbf{G}\mathbf{Q_x}+\sigma_n^2\mathbf{A}^{\mathrm{MMSE}}\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H \\
&\quad +\mathbf{A}^{\mathrm{MMSE}}\mathbf{G}\mathbf{Q_x}\mathbf{G}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H. \qquad \text{(B.8)}
\end{aligned}$$

The minimum of $\mathrm{trace}\left(\mathbf{Q_M}\right)$ can be computed by solving

$$\left.\frac{\partial\,\mathrm{trace}\left(\mathbf{Q_M}\right)}{\partial\mathbf{A}^{\mathrm{MMSE}}}\right|_{\mathbf{A}^{\mathrm{MMSE}}}=\mathbf{0}, \qquad \text{(B.9)}$$

with respect to $\mathbf{A}^{\mathrm{MMSE}}$ by (Abadir and Magnus, 2005)

$$\frac{\partial}{\partial \mathbf{A}^{\mathrm{MMSE}}} \operatorname{trace}\Big(\mathbf{Q_x} - \mathbf{Q_x}\mathbf{G}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H - \mathbf{A}^{\mathrm{MMSE}}\mathbf{G}\mathbf{Q_x} + \sigma_n^2\mathbf{A}^{\mathrm{MMSE}}\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H + \mathbf{A}^{\mathrm{MMSE}}\mathbf{G}\mathbf{Q_x}\mathbf{G}^H\left(\mathbf{A}^{\mathrm{MMSE}}\right)^H\Big) = \mathbf{0},$$

$$-\mathbf{Q_x}\mathbf{G}^H - \mathbf{Q_x}^H\mathbf{G}^H + 2\sigma_n^2\mathbf{A}^{\mathrm{MMSE}} + \mathbf{A}^{\mathrm{MMSE}}\left(\left(\mathbf{G}\mathbf{Q_x}\mathbf{G}^H\right)^H + \mathbf{G}\mathbf{Q_x}\mathbf{G}^H\right) = \mathbf{0}. \tag{B.10}$$

Since the transmitted data streams, i.e., the individual elements of $\mathbf{x}$, are independent, $\mathbf{Q_x}$ is merely a diagonal matrix whose off-diagonal elements are zero ($\mathbf{Q_x}_{ii} \in \mathbb{R}^{M_T}$, $1 \leqslant i \leqslant M_T$, $\mathbf{Q_x}_{ij} = 0,\ i \neq j,\ 1 \leqslant i,j \leqslant M_T$). Hence, the Eq. (B.10) can be reduced into

$$-2\mathbf{Q_x}\mathbf{G}^H + 2\sigma_n^2\mathbf{A}^{\mathrm{MMSE}} + 2\mathbf{A}^{\mathrm{MMSE}}\mathbf{G}\mathbf{Q_x}\mathbf{G}^H = 0, \tag{B.11}$$

with

$$\mathbf{Q_x} = \mathbf{Q_x}^H. \tag{B.12}$$

The solution can be then written as

$$\mathbf{A}^{\mathrm{MMSE}} = \mathbf{Q_x}\mathbf{G}^H\left(\mathbf{G}\mathbf{Q_x}\mathbf{G}^H + \sigma_n^2\mathbf{I}\right)^{-1}, \tag{B.13}$$

or equivalently (Harville, 1997)

$$\mathbf{A}^{\mathrm{MMSE}} = \left(\mathbf{G}^H\mathbf{G}\frac{1}{\sigma_n^2} + \mathbf{Q_x}^{-1}\right)^{-1}\mathbf{G}^H\frac{1}{\sigma_n^2}. \tag{B.14}$$

In practical problems, $\mathbf{Q_x} = \mathbf{I}P^{\mathrm{tx}}$ may commonly be encountered where the Eq. (B.13) and (B.14) can be reduced into

$$\mathbf{A}^{\mathrm{MMSE}} = \mathbf{G}^H\left(\mathbf{G}\mathbf{G}^H + \frac{\sigma_n^2}{P^{\mathrm{tx}}}\mathbf{I}\right)^{-1}, \tag{B.15}$$

$$\mathbf{A}^{\mathrm{MMSE}} = \left(\mathbf{G}^H\mathbf{G} + \frac{\sigma_n^2}{P^{\mathrm{tx}}}\mathbf{I}\right)^{-1}\mathbf{G}^H. \tag{B.16}$$

Bibliography

Abadir K. M. and Magnus J. R. (2005) *Matrix Algebra*, Cambridge University Press, Cambridge, UK.

Ademaj F., Schwarz S., Berisha T. and Rupp M. (2019) A Spatial Consistency Model for Geometry-Based Stochastic Channels, *IEEE Access* 7, 183414–183427.

Ai B., Guan K., He R., Li J., Li G., He D., Zhong Z. and Huq K. M. S. (2017) On Indoor Millimeter Wave Massive MIMO Channels: Measurement and Simulation, *IEEE Journal on Selected Areas in Communications* 35(7), 1678–1690.

Albani M. (2005) A uniform double diffraction coefficient for a pair of wedges in arbitrary configuration, *IEEE Transactions on Antennas and Propagation* 53(2), 702–710.

Almers P., Tufvesson F. and Molisch A. (2003) Measurement of keyhole effect in a wireless multiple-input multiple-output (MIMO) channel, *IEEE Communications Letters* 7(8), 373–375.

Asplund H., Glazunov A. A., Molisch A. F., Pedersen K. I. and Steinbauer M. (2006) The COST 259 Directional Channel Model-Part II: Macrocells, *IEEE Transactions on Wireless Communications* 5(12), 3434–3450.

Balanis C. A. (2012) *Advanced Engineering Electromagnetics*, Wiley, Hoboken, NJ, 2nd edn.

Bengtsson E. L., Karlsson P. C., Tufvesson F., Vieira J., Malkowsky S., Liu L., Rusek F. and Edfors O. (2016) Transmission Schemes for Multiple Antenna Terminals in Real Massive MIMO Systems, in *IEEE Global Communications Conference (GLOBECOM)*, IEEE, Washington, DC.

Bertsekas D. P. and Tsitsiklis J. N. (2008) *Introduction to Probability*, Athena Scientific, Belmont, MA, 2nd edn.

Bibb D. A., Dang J., Yun Z. and Iskander M. F. (2014) Computational accuracy and speed of some knife-edge diffraction models, in *IEEE Antennas and Propagation Society International Symposium (APSURSI)*, pp. 705–706, Memphis, TN.

Biglieri E., Calderbank R., Constantinides A., Goldsmith A., Paulraj A. and Poor H. V. (2007) *MIMO Wireless Communications*, Cambridge University Press, Cambridge, UK.

Björnson E., Kountouris M., Bengtsson M. and Ottersten B. (2013) Receive Combining vs. Multi-Stream Multiplexing in Downlink Systems With Multi-Antenna Users, *IEEE Transactions on Signal Processing* 61(13), 3431–3446.

Björnson E. and Larsson E. G. (2015) Three Practical Aspects of Massive MIMO: Intermittent User Activity, Pilot Synchronism, and Asymmetric Deployment, in *IEEE Globecom Workshops*, San Diego, CA.

Björnson E., Larsson E. G. and Marzetta T. L. (2016) Massive MIMO: Ten Myths and One Critical Question, *IEEE Communications Magazine* 54(2), 114–123.

Born M. and Wolf E. (1999) *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Cambridge University Press, Cambridge, UK, 7th edn.

Borovikov V. A. and Kinber B. E. (1994) *Geometrical Theory of Diffraction*, Institution of Electrical Engineers, London, UK, 1st edn.

Brem R. M. (2015) Highly parallel ray based wave propagation modeling and improved accuracy concepts for radiation and scattering problems involving composite metallic and dielectric objects, Ph.D. thesis, Technische Universität München, München.

Busari S. A., Huq K. M. S., Mumtaz S., Dai L. and Rodriguez J. (2018) Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey, *IEEE Communications Surveys & Tutorials* 20(2), 836–869.

Carluccio G. and Albani M. (2008) An Efficient Ray Tracing Algorithm for Multiple Straight Wedge Diffraction, *IEEE Transactions on Antennas and Propagation* 56(11), 3534–3542.

Chen K. (2016) *Principles of Synthetic Aperture Radar Imaging: A System Simulation Approach*, CRC Press, Boca Raton, FL, 1st edn.

Chizhik D., Foschini G., Gans M. and Valenzuela R. (2002) Keyholes, correlations, and capacities of multielement transmit and receive antennas, *IEEE Transactions on Wireless Communications* 1(2), 361–368.

Chockalingam A. and Rajan B. S. (2014) *Large MIMO Systems*, Cambridge University Press, Cambridge, UK, 1st edn.

Clerckx B. and Oestges C. (2013) *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*, Elsevier, Waltham, MA, 2nd edn.

Coldrey M. and Bohlin P. (2007) Training-Based MIMO Systems—Part I: Performance Comparison, *IEEE Transactions on Signal Processing* 55(11), 5464–5476.

Collin R. E. (1985) *Antennas and Radiowave Propagation*, McGraw-Hill, New York, NY, 1st edn.

Conte S. D. and De Boor C. (1980) *Elementary Numerical Analysis: An Algorithmic Approach*, McGraw-Hill, New York, 3rd edn.

Costa M. (1983) Writing on dirty paper (Corresp.), *IEEE Transactions on Information Theory* 29(3), 439–441.

Cover T. M. and Thomas J. A. (2005) *Elements of Information Theory*, Wiley, Hoboken, NJ, 1st edn.

Damasso E. and Correa L. (1999) Digital Mobile Radio Towards Future Generation Systems - COST 231 Final Report, Tech. rep., Luxembourg.

Deschamps G. (1972) Ray techniques in electromagnetics, *Proceedings of the IEEE* 60(9), 1022–1035.

Deygout J. (1966) Multiple knife-edge diffraction of microwaves, *IEEE Transactions on Antennas and Propagation* 14(4), 480–489.

Didascalou D. (2000) Ray-optical wave propagation modelling in arbitrarily shaped tunnels, Ph.D. thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany.

Dooren V. G. A. J. (1994) A deterministic approach to the modelling of electromagnetic wave propagation in urban environments, Ph.D. thesis, Eindhoven University of Technology, Eindhoven, Netherlands.

Du K.-L. and Swamy M. N. S. (2010) *Wireless Communication Systems: From RF Subsystems to 4G Enabling Technologies*, Cambridge University Press, Cambridge, UK, 1st edn.

Elijah O., Leow C. Y., Rahman T. A., Nunoo S. and Iliya S. Z. (2016) A Comprehensive Survey of Pilot Contamination in Massive MIMO—5G System, *IEEE Communications Surveys & Tutorials* 18(2), 905–923.

Elliott R. C. and Krzymien W. A. (2009) Downlink Scheduling via Genetic Algorithms for Multiuser Single-Carrier and Multicarrier MIMO Systems With Dirty Paper Coding, *IEEE Transactions on Vehicular Technology* 58(7), 3247–3262.

Erez U., Shamai S. and Zamir R. (2005) Capacity and lattice strategies for canceling known interference, *IEEE Transactions on Information Theory* 51(11), 3820–3833.

Erez U. and ten Brink S. (2005) A close-to-capacity dirty paper coding scheme, *IEEE Transactions on Information Theory* 51(10), 3417–3432.

ETSI (2003) Spatial channel model for multiple input multiple output (MIMO) simulations, Tech. rep.

Fedoryuk M. V. (1971) The stationary phase method and pseudodifferential operators, *Russian Mathematical Surveys* 26(1), 65–115.

Foschini G. and Gans M. (1998) On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas, *Wireless Personal Communications* 6(3), 311–335.

Gamkrelidze R. V. (1989) *Analysis I: Integral Representations and Asymptotic Methods*, Springer, Berlin, Germany, 1st edn.

Gan M. (2015) Accurate and low-complexity ray tracing channel modeling, Ph.D. thesis, Vienna University of Technology, Wien, Austria.

Gao X. (2016) Massive MIMO in real propagation environments, Ph.D. thesis, Lund University, Lund, Sweden.

Giovaneli C. (1984) An analysis of simplified solutions for multiple knife-edge diffraction, *IEEE Transactions on Antennas and Propagation* 32(3), 297–301.

Glassner A. S. (2007) *An Introduction to Ray Tracing*, Kaufmann, San Francisco, CA.

Goldsmith A. (2005) *Wireless Communications*, Cambridge University Press, Cambridge, UK.

Gunnarsson S., Flordelis J., der Perre L. V. and Tufvesson F. (2018) Channel Hardening in Massive MIMO-A Measurement Based Analysis, in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Greece.

Haghighatshoar S. and Caire G. (2017) Massive MIMO Pilot Decontamination and Channel Interpolation via Wideband Sparse Channel Estimation, *IEEE Transactions on Wireless Communications* 16(12), 8316–8332.

Harrington R. F. (2001) *Time-Harmonic Electromagnetic Fields*, IEEE Press, New York, NY, 1st edn.

Harville D. A. (1997) *Matrix Algebra from a Statistician's Perspective*, Springer, New York, NY.

Heath R. W. and Lozano A. (2019) *Foundations of MIMO Communication*, Cambridge University Press, Cambridge, UK.

Hien Quoc Ngo, Larsson E. G. and Marzetta T. L. (2013) Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems, *IEEE Transactions on Communications* 61(4), 1436–1449.

Holm P. (1996) UTD-diffraction coefficients for higher order wedge diffracted fields, *IEEE Transactions on Antennas and Propagation* 44(6), 879–888.

Holm P. (2000) A new heuristic UTD diffraction coefficient for nonperfectly conducting wedges, *IEEE Transactions on Antennas and Propagation* 48(8), 1211–1219.

Huang H. C. and Papadias C. (2008) *MIMO for Multiuser Wireless Systems: Theory and Applications*, Springer, New York, NY.

Hussain S. (2017) Efficient ray-tracing algorithms for radio wave propagation in urban environments, Ph.D. thesis, Dublin City University, Dublin, Ireland.

Idowu-Bismark O., Idachaba F. E. and Atayero A. A. (2017) Massive MIMO Channel Characterization and Modeling: The Present and the Future, *International Journal of Applied Engineering Research* 12(23), 13742–13754.

Iskander M. and Zhengqing Yun (2002) Propagation prediction models for wireless communication systems, *IEEE Transactions on Microwave Theory and Techniques* 50(3), 662–673.

ITU (2003) ITU-R P.1411-2 - Propagation data and prediction methods for the planning of short-range outdoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz, Tech. rep., International Telecommunication Union.

Jafar S. and Fakhereddin M. (2007) Degrees of Freedom for the MIMO Interference Channel, *IEEE Transactions on Information Theory* 53(7), 2637–2642.

James G. L. (2007) *Geometrical Theory of Diffraction for Electromagnetic Waves*, The Institution of Engineering and Technology, Stevenage, UK, 3rd edn.

Jiang X. and Kaltenberger F. (2018) Channel Reciprocity Calibration in TDD Hybrid Beamforming Massive MIMO Systems, *IEEE Journal of Selected Topics in Signal Processing* 12(3), 422–431.

Jiang Z., Molisch A. F., Caire G. and Niu Z. (2015) Achievable Rates of FDD Massive MIMO Systems With Spatial Channel Correlation, *IEEE Transactions on Wireless Communications* 14(5), 2868–2882.

Jindal N. (2008) Antenna combining for the MIMO downlink channel, *IEEE Transactions on Wireless Communications* 7(10), 3834–3844.

Kailath T., Sayed A. H. and Hassibi B. (2000) *Linear Estimation*, Prentice Hall, Upper Saddle River, NJ.

Kaiser T. and Bourdoux A. (Eds.) (2005) *Smart Antennas - State of the Art*, Hindawi, New York, NY.

Kasdorf S., Troksa B., Harmon J., Key C. and Notaros B. (2021) Shooting-Bouncing-Rays Technique to Model Mine Tunnels: Algorithm Acceleration, *Applied Computational Electromagnetics Society* 35(11), 1330–1331.

Keller J. B. (1962) Geometrical Theory of Diffraction, *Journal of the Optical Society of America* 52(2), 116.

Keller O. (2014) *Light: The Physics of the Photon*, CRC Press, Boca Raton, FL.

Kermoal J., Schumacher L., Pedersen K., Mogensen P. and Frederiksen F. (2002) A stochastic MIMO radio channel model with experimental validation, *IEEE Journal on Selected Areas in Communications* 20(6), 1211–1226.

Khina A. and Erez U. (2010) On the robustness of dirty paper coding, *IEEE Transactions on Communications* 58(5), 1437–1446.

Kline M. (1962) Electromagnetic theory and geometrical optics, Tech. rep., New York University.

Kouyoumjian R. and Pathak P. (1974) A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface, *Proceedings of the IEEE* 62(11), 1448–1461.

Kouyoumjian R. G. (1965) Asymptotic high-frequency methods, *Proceedings of the IEEE* 53(8), 864–876.

Kyosti P., Meinila J., Zhao X., Schneider C., Narandzic M., Hong A., Ylitalo J., Holappa V.-M., Alatossava M. and Rautiainen T. (2007) WINNER II Channel Models, Tech. rep., Information Society Technologies.

Lahiri A. (2016) *Basic Optics: Principles and Concepts*, Elsevier, Amsterdam, Netherlands.

Larsson E. G. (2013) Massive MIMO: Fundamentals, Opportunities and Challenges.

Levin D. (1982) Procedures for computing one- and two-dimensional integrals of functions with rapid irregular oscillations, *Mathematics of Computation* 38(158), 531–531.

Li T. and Torlak M. (2018) Performance of ZF Linear Equalizers for Single Carrier Massive MIMO Uplink Systems, *IEEE Access* 6, 32156–32172.

Li X., Bjornson E., Zhou S. and Wang J. (2016) Massive MIMO with multi-antenna users: When are additional user antennas beneficial?, in *International Conference on Telecommunications (ICT)*, pp. 1–6, IEEE, Thessaloniki, Greece.

Li, Mingmei, Jin, Shi and Gao, Xiqi (2013) Spatial orthogonality-based pilot reuse for multi-cell massive MIMO transmission, in *International Conference on Wireless Communications and Signal Processing*, pp. 1–6.

Ling H., Chou R.-C. and Lee S.-W. (1989) Shooting and bouncing rays: Calculating the RCS of an arbitrarily shaped cavity, *IEEE Transactions on Antennas and Propagation* 37(2), 194–205.

Liu G., She J., Lu W., Zhang M. and Bo Y. (2020) 3D deterministic ray tracing method for massive MIMO channel modelling and parameters extraction, *IET Communications* 14(18), 3169–3174.

Liu L., Oestges C., Poutanen J., Haneda K., Vainikainen P., Quitin F., Tufvesson F. and Doncker P. (2012) The COST 2100 MIMO channel model, *IEEE Wireless Communications* 19(6), 92–99.

Liu Y., Li G. Y. and Han W. (2016) Comparison of OFDM and Single-Carrier for Large-Scale Antenna Systems, *arXiv:1606.08751 [cs, math]* .

Lu J. S., Bertoni H. L. and Degli-Esposti V. (2014a) Scale Model Investigation of Mechanisms for Scattering From Office Buildings at 2 GHz, *IEEE Transactions on Antennas and Propagation* 62(12), 6435–6442.

Lu L., Li G. Y., Swindlehurst A. L., Ashikhmin A. and Zhang R. (2014b) An Overview of Massive MIMO: Benefits and Challenges, *IEEE Journal of Selected Topics in Signal Processing* 8(5), 742–758.

Luebbers R. (1984) Finite conductivity uniform GTD versus knife edge diffraction in prediction of propagation path loss, *IEEE Transactions on Antennas and Propagation* 32(1), 70–76.

Luo F.-L. and Zhang C. (2016) *Signal Processing for 5G: Algorithms and Implementations*, IEEE Press, Chichester, UK, 1st edn.

Mani F. and Oestges C. (2012) A Ray Based Method to Evaluate Scattering by Vegetation Elements, *IEEE Transactions on Antennas and Propagation* 60(8), 4006–4009.

Marchenko O., Kazantsev S. and Windholz L. (2003) *Demonstrational Optics*, Springer, Boston, MA, 1st edn.

Marzetta T. L. (2010) Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas, *IEEE Transactions on Wireless Communications* 9(11), 3590–3600.

Matz G. and Hlawatsch F. (2011) Fundamentals of Time-Varying Communication Channels, in *Wireless Communications Over Rapidly Time-Varying Channels*, Elsevier.

Maurer J. (2005) Strahlenoptisches Kanalmodell für die Fahrzeug-Fahrzeug-Funkkommunikation, Ph.D. thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany.

McKown J. and Hamilton R. (1991) Ray tracing as a design tool for radio networks, *IEEE Network* 5(6), 27–30.

McNamara D. A., Pistorius C. W. I. and Malherbe J. A. G. (1990) *Introduction to the Uniform Geometrical Theory of Diffraction*, Artech House, Boston, MA, 1st edn.

Mocker-Henning M. S. L. (2016) Kombinierte elektromagnetische Vollwellen- und strahlbasierte Simulationen von Fahrzeugantennen in virtuellen Testfahrten, Ph.D. thesis, Technische Universität München, Munich, Germany.

Molisch A. F. (2011) *Wireless Communications*, Wiley, Chichester, UK, 2nd edn.

Molisch A. F., Asplund H., Heddergott R., Steinbauer M. and Zwick T. (2006) The COST259 Directional Channel Model-Part I: Overview and Methodology, *IEEE Transactions on Wireless Communications* 5(12), 3421–3433.

Nagy B., Elsabrouty M. and Elramly S. (2018) Fast Converging Weighted Neumann Series Precoding for Massive MIMO Systems, *IEEE Wireless Communications Letters* 7(2), 154–157.

Neitz O. (2020) Propagating Plane-Wave Expansions for Microwave Near-Field Imaging and Radar Cross Section Determination, Ph.D. thesis, Technische Universität München, Munich, Germany.

Ngo H. Q. and Larsson E. G. (2017) No Downlink Pilots Are Needed in TDD Massive MIMO, *IEEE Transactions on Wireless Communications* 16(5), 2921–2935.

Ngo H. Q., Larsson E. G. and Marzetta T. L. (2013) Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots, in *Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL.

Ngo H. Q., Larsson E. G. and Marzetta T. L. (2014) Aspects of favorable propagation in Massive MIMO, in *European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal.

Ngo H. Q., Marzetta T. L. and Larsson E. G. (2011) Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic.

Olver S. (2008) Numerical Approximation of Highly Oscillatory Integrals, Ph.D. thesis, University of Cambridge, Cambridge, UK.

Özbek B. and Le Ruyet D. (2014) *Feedback Strategies for Wireless Communication*, Springer, New York, NY.

Pancaldi F., Vitetta G., Kalbasi R., Al-Dhahir N., Uysal M. and Mheidat H. (2008) Single-carrier frequency domain equalization, *IEEE Signal Processing Magazine* 25(5), 37–56.

Papazafeiropoulos A. K. (2017) Impact of General Channel Aging Conditions on the Downlink Performance of Massive MIMO, *IEEE Transactions on Vehicular Technology* 66(2), 1428–1442.

Parker S. G., Bigler J., Dietrich A., Friedrich H., Hoberock J., Luebke D., McAllister D., McGuire M., Morley K., Robison A. and Stich M. (2010) OptiX: A general purpose ray tracing engine, in *ACM SIGGRAPH*, Los Angeles, CA.

Parsons J. D. (2000) *The Mobile Radio Propagation Channel*, Wiley, Chichester, UK, 2nd edn.

Pathak P. H., Carluccio G. and Albani M. (2013) The Uniform Geometrical Theory of Diffraction and Some of Its Applications, *IEEE Antennas and Propagation Magazine* 55(4), 41–69.

Patil D. A. (2017) Block Diagonalization Based Beamforming, Ph.D. thesis, KTH Royal Institute of Technology, Stockholm, Sweden.

Paulraj A., Nabar R. and Gore D. (2003) *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, UK.

Payami S. and Tufvesson F. (2013) Delay spread properties in a measured massive MIMO system at 2.6 GHz, in *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, London, UK.

Pinhasi Y., Yahalom A. and Petnev S. (2008) Propagation of ultra wide-band signals in lossy dispersive media, in *IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems*, Tel Aviv, Israel.

Pitarokoilis A., Mohammed S. K. and Larsson E. G. (2012) On the Optimality of Single-Carrier Transmission in Large-Scale Antenna Systems, *IEEE Wireless Communications Letters* 1(4), 276–279.

Pongsilamanee P. and Bertoni H. (2004) Specular and nonspecular scattering from building facades, *IEEE Transactions on Antennas and Propagation* 52(7), 1879–1889.

Poutanen J. (2011) Geometry-based radio channel modeling: Propagation analysis and concept development, Ph.D. thesis, Aalto University, Espoo, Finland.

Qin X., Yan Z. and He G. (2016) A Near-Optimal Detection Scheme Based on Joint Steepest Descent and Jacobi Method for Uplink Massive MIMO Systems, *IEEE Communications Letters* 20(2), 276–279.

Rappaport T. S. (1996) *Wireless Communications: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, 1st edn.

Rizk K., Wagen J.-F. and Gardiol F. (2000) Influence of database accuracy on two-dimensional ray-tracing-based predictions in urban microcells, *IEEE Transactions on Vehicular Technology* 49(2), 631–642.

Roivainen A. (2017) Three-dimensional geometry-based radio channel model: Parametrization and validation at 10 GHz, Ph.D. thesis, University of Oulu, Oulu, Finland.

Rusek F., Persson D., Buon Kiong Lau, Larsson E. G., Marzetta T. L. and Tufvesson F. (2013) Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays, *IEEE Signal Processing Magazine* 30(1), 40–60.

Sandmann A., Ahrens A. and Lochmann S. (2015) Resource Allocation in SVD-Assisted Optical MIMO Systems using Polynomial Matrix Factorization, in *ITG Conference on Systems, Communications and Coding*, Hamburg, Germany.

Schiller M., Knoll A., Mocker M. and Eibert T. (2015) GPU accelerated ray launching for high-fidelity virtual test drives of VANET applications, in *International Conference on High Performance Computing & Simulation (HPCS)*, Amsterdam, Netherlands.

Schneider M. and Luebbers R. (1991) A general, uniform double wedge diffraction coefficient, *IEEE Transactions on Antennas and Propagation* 39(1), 8–14.

Seidel S. and Rappaport T. (1994) Site-specific propagation prediction for wireless in-building personal communication system design, *IEEE Transactions on Vehicular Technology* 43(4), 879–891.

Shannon C. E. (1948) A mathematical theory of communication, *The Bell System Technical Journal* 27(3), 379–423.

Shen J.-C., Zhang J. and Letaief K. B. (2015) Downlink User Capacity of Massive MIMO Under Pilot Contamination, *IEEE Transactions on Wireless Communications* 14(6), 3183–3193.

Shin H. and Lee J. H. (2003) Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole, *IEEE Transactions on Information Theory* 49(10), 2636–2647.

Silver S. (1984) *Microwave Antenna Theory and Design*, Institution of Electrical Engineers, London, UK, 1st edn.

Soni S. K. and Bhattacharya A. (2010) New Heuristic Diffraction Coefficient for Modeling of Wireless Channel, *Progress In Electromagnetics Research C* 12, 125–137.

Spencer Q., Peel C., Swindlehurst A. and Haardt M. (2004) An introduction to the multi-user MIMO downlink, *IEEE Communications Magazine* 42(10), 60–67.

Stankovic V. and Haardt M. (2008) Generalized Design of Multi-User MIMO Precoding Matrices, *IEEE Transactions on Wireless Communications* 7(3), 953–961.

Sukharevsky O. I. (2015) *Electromagnetic Wave Scattering by Aerial and Ground Radar Objects*, CRC Press, Boca Raton, FL.

Sung H., Lee S.-R. and Lee I. (2009) Generalized Channel Inversion Methods for Multiuser MIMO Systems, *IEEE Transactions on Communications* 57(11), 3489–3499.

Taesang Yoo and Goldsmith A. (2006) Capacity and power allocation for fading MIMO channels with channel estimation error, *IEEE Transactions on Information Theory* 52(5), 2203–2214.

Taygur M. M. and Eibert T. F. (2018a) Analyzing the Channel Aging Effects on Massive MIMO Downlink by Ray-Tracing, in *IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy.

Taygur M. M. and Eibert T. F. (2018b) Investigating the Effects of Uplink Noise and Pilot Signal Interference by Ray-Tracing Simulations, in *IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy.

Taygur M. M. and Eibert T. F. (2019a) Characterization of Frequency-Selective Massive MIMO Channels by Ray-Tracing, in *European Conference on Antennas and Propagation (EuCAP)*, Krakow, Poland.

Taygur M. M. and Eibert T. F. (2019b) Investigations on Massive MIMO Performance with Multi-Antenna Users by Ray-Tracing, in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Istanbul, Turkey.

Taygur M. M. and Eibert T. F. (2020a) Determination of Exact Ray Paths by Bidirectional Ray-Tracing, in *General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Rome, Italy.

Taygur M. M. and Eibert T. F. (2020b) A Ray-Tracing Algorithm Based on the Computation of (Exact) Ray Paths With Bidirectional Ray-Tracing, *IEEE Transactions on Antennas and Propagation* 68(8), 6277–6286.

Taygur M. M. and Eibert T. F. (2021a) Foliage Modeling for Urban Ray-Tracing Simulations Using Satellite Images, in *European Conference on Antennas and Propagation (EuCAP)*, Düsseldorf, Germany.

Taygur M. M. and Eibert T. F. (2021b) Investigations on Millimeter-Wave Massive MIMO Scenarios by Bidirectional Ray-Tracing Using Stationary-Phase Point Approximation, in *General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Rome, Italy.

Taygur M. M., Sukharevsky I. and Eibert T. F. (2019a) Investigating Multiple Diffractions Near Optical Boundaries with Bidirectional Ray-Tracing and Antenna Reciprocity, in *URSI International Symposium on Electromagnetic Theory (EMTS)*, San Diego, CA.

Taygur M. M., Sukharevsky I. O. and Eibert a. T. F. (2019b) Investigation of Massive MIMO Scenarios Involving Rooftop Propagation by Bidirectional Ray-Tracing, *Progress In Electromagnetics Research C* 91, 129–142.

Taygur M. M., Sukharevsky I. O. and Eibert T. F. (2018) A Bidirectional Ray-Tracing Method for Antenna Coupling Evaluation Based on the Reciprocity Theorem, *IEEE Transactions on Antennas and Propagation* 66(12), 6654–6664.

Telatar E. (1999) Capacity of Multi-antenna Gaussian Channels: Capacity of Multi-antenna Gaussian Channels, *European Transactions on Telecommunications* 10(6), 585–595.

Torres P., Charrua L. and Gusmao A. (2014) On the SC/FDE Uplink Alternative to OFDM in a Massive MU-MIMO Context, in *International Conference on Wireless and Mobile Communications*, Seville, Spain.

Truong K. T. and Heath R. W. (2013) Effects of channel aging in massive MIMO systems, *Journal of Communications and Networks* 15(4), 338–351.

Tse D. and Viswanath P. (2005) *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, UK, 1st edn.

Tsoulos G. V. (2006) *MIMO System Technology for Wireless Communications*, CRC Press, Boca Raton, FL, 1st edn.

Tulino A. M. and Verdú S. (2004) *Random Matrix Theory and Wireless Communications*, Now Publishers, Hanover, MA, 1st edn.

Tzaras C. and Saunders S. (2000) Comparison of multiple-diffraction models for digital broadcasting coverage prediction, *IEEE Transactions on Broadcasting* 46(3), 221–226.

Valle L., Pérez J. R. and Torres R. P. (2020) Characterisation of Indoor Massive MIMO Channels Using Ray-Tracing: A Case Study in the 3.2–4.0 GHz 5G Band, *Electronics* 9(8), 1250.

Verdú S. (1998) *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1st edn.

Vogler L. E. (1981) The attenuation of electromagnetic waves by multiple knife-edge diffraction, Tech. rep.

Wallace J. and Jensen M. (2002) Modeling the indoor MIMO wireless channel, *IEEE Transactions on Antennas and Propagation* 50(5), 591–599.

Wang F., Eibert T. F. and Jin Y.-Q. (2015) Simulation of ISAR Imaging for a Space Target and Reconstruction Under Sparse Sampling via Compressed Sensing, *IEEE Transactions on Geoscience and Remote Sensing* 53(6), 3432–3441.

Wen W., Xia M. and Wu Y.-C. (2010) Low Complexity Pre-Equalization Algorithms for Zero-Padded Block Transmission, *IEEE Transactions on Wireless Communications* 9(8), 2498–2504.

Wong R. (2001) *Asymptotic Approximations of Integrals*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Wu S. (2015) Massive MIMO channel modelling for 5G wireless communication systems, Ph.D. thesis, Heriot-Watt University.

Yang Y., Xu J., Shi G. and Wang C.-X. (2018) *5G Wireless Systems: Simulation and Evaluation Techniques*, Wireless Networks, Springer, Cham, Switzerland, 1st edn.

Yin H., Cottatellucci L., Gesbert D., Müller R. R. and He G. (2016) Robust Pilot Decontamination Based on Joint Angle and Power Domain Discrimination, *IEEE Transactions on Signal Processing* 64(11), 2990–3003.

Yin H., Gesbert D. and Cottatellucci L. (2014) Dealing With Interference in Distributed Large-Scale MIMO Systems: A Statistical Approach, *IEEE Journal of Selected Topics in Signal Processing* 8(5), 942–953.

Yoo T. and Goldsmith A. (2006) On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming, *IEEE Journal on Selected Areas in Communications* 24(3), 528–541.

Yuan Y. (2014) Realistic geometry-based stochastic channel models for advanced wireless MIMO systems, Ph.D. thesis, Heriot-Watt University.

Yun Z., Iskander M. and Zhang Z. (2001) Development of a new shooting-and-bouncing ray (SBR) tracing method that avoids ray double counting, in *IEEE Antennas and Propagation Society International Symposium.*

Yun Z. and Iskander M. F. (2015) Ray Tracing for Radio Propagation Modeling: Principles and Applications, *IEEE Access* 3, 1089–1100.

Zhao L., Zhao H., Zheng K. and Xiang W. (2018) *Massive MIMO in 5G Networks: Selected Applications*, Springer, Cham, Switzerland.

Zhekov S. S., Franek O. and Pedersen G. F. (2020) Dielectric Properties of Common Building Materials for Ultrawideband Propagation Studies [Measurements Corner], *IEEE Antennas and Propagation Magazine* 62(1), 72–81.

Zu K., de Lamare R. C. and Haardt M. (2013) Generalized Design of Low-Complexity Block Diagonalization Type Precoding Algorithms for Multiuser MIMO Systems, *IEEE Transactions on Communications* 61(10), 4232–4242.

Publications of the Author

Taygur M. M. and Eibert T. F. (2017) Investigation of distributed and collocated base stations in a large urban massive MIMO scenario, in *European Conference on Antennas and Propagation (EUCAP)*, pp. 1577–1581.

Taygur M. M. and Eibert T. F. (2018a) Analyzing the Channel Aging Effects on Massive MIMO Downlink by Ray-Tracing, in *IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy.

Taygur M. M. and Eibert T. F. (2018b) Investigating the Effects of Uplink Noise and Pilot Signal Interference by Ray-Tracing Simulations, in *IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy.

Taygur M. M. and Eibert T. F. (2019a) Characterization of Frequency-Selective Massive MIMO Channels by Ray-Tracing, in *European Conference on Antennas and Propagation (EuCAP)*, Krakow, Poland.

Taygur M. M. and Eibert T. F. (2019b) Investigations on Massive MIMO Performance with Multi-Antenna Users by Ray-Tracing, in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Istanbul, Turkey.

Taygur M. M. and Eibert T. F. (2020a) Asymptotic Expansion of the Reciprocity Integral in a Bidirectional Ray-Tracing Approach, in *European Conference on Antennas and Propagation (EuCAP)*, Copenhagen, Denmark.

Taygur M. M. and Eibert T. F. (2020b) Determination of Exact Ray Paths by Bidirectional Ray-Tracing, in *General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Rome, Italy.

Taygur M. M. and Eibert T. F. (2020c) A Ray-Tracing Algorithm Based on the Computation of (Exact) Ray Paths With Bidirectional Ray-Tracing, *IEEE Transactions on Antennas and Propagation* 68(8), 6277–6286.

Taygur M. M. and Eibert T. F. (2021a) Foliage Modeling for Urban Ray-Tracing Simulations Using Satellite Images, in *European Conference on Antennas and Propagation (EuCAP)*, Düsseldorf, Germany.

Taygur M. M. and Eibert T. F. (2021b) Investigations on Millimeter-Wave Massive MIMO Scenarios by Bidirectional Ray-Tracing Using Stationary-Phase Point Approximation, in *General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Rome, Italy.

Taygur M. M., Sukharevsky I. and Eibert T. F. (2019a) Investigating Multiple Diffractions Near Optical Boundaries with Bidirectional Ray-Tracing and Antenna Reciprocity, in *URSI International Symposium on Electromagnetic Theory (EMTS)*, San Diego, CA.

Taygur M. M., Sukharevsky I. O. and Eibert a. T. F. (2019b) Investigation of Massive MIMO Scenarios Involving Rooftop Propagation by Bidirectional Ray-Tracing, *Progress In Electromagnetics Research C* 91, 129–142.

Taygur M. M., Sukharevsky I. O. and Eibert T. F. (2018a) A Bidirectional Ray-Tracing Method for Antenna Coupling Evaluation Based on the Reciprocity Theorem, *IEEE Transactions on Antennas and Propagation* 66(12), 6654–6664.

Taygur M. M., Sukharevsky I. O. and Eibert T. F. (2018b) Computation of Antenna Transfer Functions with a Bidirectional Ray-Tracing Algorithm Utilizing Antenna Reciprocity, in *URSI Atlantic Radio Science Meeting (AT-RASC)*, Gran Canaria, Spain.