

Diss. ETH No. 17549

# Modelling of Random Porous Media Using Minkowski-Functionals

A dissertation submitted to the  
SWISS FEDERAL INSTITUTE OF TECHNOLOGY  
ZURICH

for the degree of  
Doctor of Sciences ETH

presented by  
MAIK BERCHTOLD  
Dipl. Math. ETH  
born December 17, 1976  
citizen of Busswil BE

accepted on the recommendation of  
Prof. Dr. H. R. Künsch, examiner  
Prof. Dr. V. Schmidt, Prof. Dr. P. Bühlmann, co-examiner

2008

Seite Leer /  
Blank leaf

# Acknowledgements

My thank goes to Prof.Dr.Hans-Rudolf Künsch for the supervision of this thesis. With his constant support and valuable inputs he helped me circumnavigate most of the numerous cliffs along my way. His contributions are substantial, I owe him much.

My appreciation goes further to my two co-referees Prof. Dr. Volker Schmidt (University of Ulm) and Prof. Dr. Peter Bühlmann (ETHZ) for the courtesy to sacrifice a good part of their precious time to evaluate my work.

I would like to thank my parents, especially my mother, my partner Jeannine and also my little dog Nejo who gave me their love and support to bear all the setbacks and to stay the course. Without them, probably this thesis would never have been written.

Last but not least a word of thank also goes to the people of the Seminar for Statistik for providing an excellent working atmosphere and for brightening up my normal course of life during the past four years.

Seite Leer /  
Blank leaf

# Contents

<b>Abstract</b>	<b>ix</b>
<b>Zusammenfassung</b>	<b>ix</b>
<b>1 Introduction and Problem Formulation</b>	<b>1</b>
1.1 General Introduction . . . . .	1
1.2 General Problem Formulation . . . . .	1
1.3 Data . . . . .	4
<b>2 Preliminaries</b>	<b>9</b>
2.1 Definition and Properties of Minkowski -Functionals . . . . .	9
2.1.1 The Steiner Formula . . . . .	9
2.1.2 Properties of Minkowski Functionals . . . . .	11
2.1.3 The Hadwiger Characterization Theorem . . . . .	13
2.1.4 Extended Definition of Minkowski Functionals . . . . .	14
2.1.5 The Euler Characteristic . . . . .	15
2.2 The Crofton and the Kinematic-Formula . . . . .	23
2.2.1 The Crofton Formula . . . . .	23
2.2.2 Basic Consequences from the Crofton-Formula . . . . .	25
2.2.3 The Principal Kinematic Formula . . . . .	26
2.2.4 Curvature Integrals and Minkowski-Functionals . . . . .	27
2.3 Ohser-Mücklich-Estimators for Minkowski-Functionals . . . . .	29
2.3.1 Ohser-Mücklich-Estimators for $\mathbf{d} = \mathbf{2}$ . . . . .	30
2.3.2 Ohser-Mücklich-Estimators for $\mathbf{d} = \mathbf{3}$ . . . . .	33
2.4 Random Sets and Particle Processes . . . . .	37
2.4.1 Random Sets . . . . .	37
2.4.2 Point and Particle Processes . . . . .	40
2.5 Other Geometrical Characteristics . . . . .	43
2.5.1 The m-Point-Covariance Functions . . . . .	43
2.5.2 The Chord-Length-Distribution . . . . .	44
2.5.3 The Pore-Size-Distribution . . . . .	45
2.5.4 The Minkowski-Functions . . . . .	45
<b>3 The Boolean Model</b>	<b>47</b>
3.1 Definition and Relevant Properties . . . . .	47
3.2 Milcs-Formulae . . . . .	49
3.2.1 Derivation and Interpretation . . . . .	49
3.2.2 Attainability of Specific Minkowski-Values Within the Homothetic Boolean Model . . . . .	54

3.3	Simulation Procedures . . . . .	65
3.3.1	Simulation of the 2D-Circular Boolean Model . . . . .	66
3.3.2	Simulation of the 2D-Elliptical Boolean Model . . . . .	67
3.3.3	Simulation of the 3D-Spherical-Boolean Model . . . . .	70
3.3.4	Simulation of the 3D-Homothetic Ellipsoidal-Boolean Model . . . . .	72
3.3.5	The Discretization-Process . . . . .	74
3.4	2D-Results . . . . .	76
3.4.1	Typical Results for the 2D-Spherical Boolean Model . . . . .	76
3.4.2	Typical Results for the 2D-Elliptical Boolean Model . . . . .	80
3.5	3D-Results . . . . .	82
3.5.1	Typical Results for the 3D-Spherical Boolean Model . . . . .	82
3.5.2	Typical Results for the 3D-Ellipsoidal Boolean Model . . . . .	84
<b>4</b>	<b>The Thresholded Gaussian Model</b>	<b>89</b>
4.1	Definition and Relevant Properties . . . . .	89
4.2	Adler-Tomita-Analytical-Formulae . . . . .	93
4.3	Asymptotical Unbiasedness of O-M-Type-Estimators . . . . .	94
4.3.1	Integrals and Expansions . . . . .	95
4.3.2	Two Dimensional Convergence Results . . . . .	105
4.3.3	Three Dimensional Convergence Results . . . . .	107
4.3.4	Bias of the 2D-Euler-Characteristic Estimator . . . . .	113
4.4	General Surface Estimators Based on the Covariance Function . . . . .	115
4.5	Asymptotic Normality of Surface Estimators . . . . .	120
4.5.1	Variance of General Surface Estimators Based on the Covariance Function . . . . .	120
4.5.2	Asymptotic Normality for the Surface Estimator in the Gaussian Thresholded Field Model . . . . .	123
4.6	The Generalized Thresholded Field Model . . . . .	135
4.7	Simulation for the Whittle-Matérn-Covariance Model . . . . .	138
4.7.1	2D-Simulation of Generalized Thresholded Gaussian Fields with the Whittle-Matérn Covariance Model . . . . .	138
4.7.2	3D-Simulation of Generalized Thresholded Gaussian Fields with the Whittle-Matérn Covariance Model . . . . .	140
4.8	2D-Results . . . . .	141
4.9	3D-Results . . . . .	144
<b>5</b>	<b>The Gibbsian Model</b>	<b>147</b>
5.1	Definition and Properties . . . . .	148
5.2	The Hammersley and Clifford-Theorem . . . . .	150
5.3	Choice of the Gibbs-Potential . . . . .	152
5.4	The Variational Principle . . . . .	153
5.5	Julósz-Ensembles . . . . .	156
5.6	Gibbs-Sampling . . . . .	159
5.6.1	General Definition . . . . .	159
5.6.2	The Gibbs Sampler for Neighbourhood-Potentials . . . . .	160
5.7	Parameter Determination . . . . .	164
5.7.1	Pseudo-Likelihood-Logistic-Regression-Method . . . . .	164
5.7.2	Sequential Newton-Parameter Determination . . . . .	168
5.8	Simulated Annealing-Method . . . . .	171

5.9	2D-Results	176
<b>6</b>	<b>Conclusion</b>	<b>185</b>
<b>7</b>	<b>Appendix</b>	<b>187</b>
7.1	The On-Off-Markov-System	187
7.1.1	Theory of the On-Off-Markov-System	187
7.1.2	Asymptotic Normality for General Surface Estimators in the On-Off-Markov-System	190
7.2	Proofs of Chapter 2	196
7.2.1	Proof of Theorem 2.1.10	196
7.2.2	Proof of Theorem 2.1.20	197
7.2.3	Proof of Theorem 2.2.4	198
7.2.4	Proof of the Existence of a Poisson Process in $\mathbb{R}^d$	199
7.2.5	Alternative Proof of Theorem 2.5.2	200
7.3	Proofs of Chapter 4	202
7.3.1	Alternative Proof of Lemma 4.3.2	202
7.3.2	Alternative Proof of Lemma 4.3.6	203
7.3.3	Alternative Version of Lemma 4.3.7	205
7.3.4	Asymptotic Normality of the OM-Surface-Estimator	210
7.4	Proofs of Chapter 5	214
7.4.1	Converse Direction of Hammersley and Clifford (Theorem 5.2.2)	214
7.4.2	Proof of the Equivalence of Ensembles (Theorem 5.5.3)	215
	<b>Bibliography</b>	<b>218</b>
	<b>Curriculum Vitae</b>	<b>222</b>

Seite Leer /  
Blank leaf



# Abstract

Porous structures frequently arise in nature and belong to the preferred materials studied in materials science. They are encountered in a myriad of shapes and structural variations. Among the examples of natural porous media we find rocks, soils, sponges or biological tissue and the most prominent porous basic materials being concrete, ceramics or foams. It is therefore not surprising that since the seventies of the last century porous media have belonged to the core focus of interdisciplinary research. They are of interest not only to the mathematician, geo-physicist, biologist and chemist but - not least in consideration of the sheer mass of data to be handled - also for the computer science and image-processing communities. Due to the fast-paced development of synchrotron technology, high-resolution images of three-dimensional porous specimens have finally become available in recent years. The Institute for Terrestrial Ecology at the Swiss Federal Institute of Technology (ETH) and the Paul-Scherrer-Institut (PSI) in Villigen are involved in the making of such high-resolution images for sand soils of various granularity which by courtesy they made available to us for use in the present thesis.

The geometrical structure of the pore space is known to have a major impact on the flow- and transport-properties in porous media such as permeability. Unfortunately neither the specific nature of this impact is presently known nor which exactly are the decisive characteristics of the pore space responsible for this impact. In this thesis we mainly concentrate on a certain simple class of geometrical characteristics, the so-called Minkowski-functionals. They comprise well-known elementary geometrical quantities such as volume (of the pore and also the solid phase), surface (of the boundary between pore and solid phase), the integral of mean curvature and the Euler-characteristic, the latter being an important connectivity-measure well-known in differential topology. The use of Minkowski-functionals to summarize the information content of a porous specimen can be theoretically justified by the famous Hadwiger-Theorem.

For the modelling of porous media in this thesis we avail ourselves of the methods of stochastics. Hence the goal we pursue cannot be the exact artificial reconstruction of a natural porous structure but rather the analysis of its stochastic properties. This will allow us to quantitatively compare different realizations of the same stochastic model with respect to their geometrical properties and analyse their variability. Our basic procedure is as follows: We first measure the properties of interest (mainly the Minkowski-functionals) and then fit a stochastic model such that in average the realizations of the model reproduce the measured values for those characteristics. In this thesis, we basically concentrate on three simple stochastic models, which are the germ-grain type Boolean Model, the Truncated Gaussian Fields and the Gibbsian Model. Our focus regarding the selection of stochastic models to consider was more on analytical tractability of a specific model than on its power to generate realistic artificial reproductions of porous structures. This is the main reason why we did not consider in this thesis the more complicated realistic models of the hard-core type (such as the cherry-pit model).

The main goals of our project were to investigate whether we can generate artificial random structures with predetermined Minkowski functionals and to examine to which extent the Minkowski functionals are able to summarize the geometrical information inherent in a porous structure. We had to develop algorithms to fit these simple models to real data which can cope with the enormous data sizes and at the same time keep processing time reasonably low. Together with the Institute of Terrestrial Ecology of ETH we have further investigated whether our artificially generated random media exhibit liquid transport- and flow properties comparable to their real world counterparts. The present thesis is structured as follows:

An introductory chapter deals with the abstract model-independent description of the basic modelling process we use to generate artificial random porous media with predetermined Minkowski-functionals. This procedure is then adapted in later chapters to the specific models. Furthermore in this introductory chapter we present the synchrotron sand data we have used all the time to test our algorithms on.

The second chapter is devoted to the theoretical foundations needed to understand this thesis. We give an introduction to Brunn- Minkowski-Theory where we introduce Minkowski-functionals on the convex ring and discuss their basic properties. Special attention is given to the Euler-characteristic from which all the other Minkowski functionals can be derived by means of the Crofton-formulae. To discuss these formulae we also give a brief introduction into the subject of integral geometry. Furthermore the chapter deals with the question how to measure or estimate the specific Minkowski functionals of a porous structure from a pixel image. To do this we use the so-called Ohser-Mücklich estimators which arise from discretizing the integral geometric expressions of the Crofton formulae. So far all the concepts introduced are purely deterministic. Towards the end of the chapter we demonstrate how stochastics can be incorporated into this geometrical framework. This leads us to the basic concepts of stochastic geometry which we also introduce briefly. The focus here lies on point and particle processes and also random sets in general. The chapter is concluded by listing some of the most frequently used geometrical characteristics for modelling porous media discussed in literature.

The third chapter deals with the Boolean germ-grain model. We interpret the Boolean Model as a particle process and demonstrate how the results of point process theory combined with integral geometry can be combined to derive explicit expressions for the specific Minkowski functionals for the Boolean Model. We discuss which values for the specific Minkowski functionals are attainable within the limits of the Boolean Model and describe algorithms to generate two and three dimensional Boolean structures with spherical or ellipsoidal grains and predetermined Minkowski functionals. In the results section we display two- and three-dimensional images generated according to these algorithms and conclude the chapter with a discussion of sensitivity of the Boolean structures with respect to variation of a single Minkowski characteristic while keeping the others constant.

In the fourth chapter we treat the Thresholded Gaussian Field Model. Binary images can be obtained from a Gaussian Random Field by painting a pixel black whenever the Gaussian random variable associated with this pixel exceeds a certain threshold and keep it white otherwise. Also for the Gaussian Model explicit formulae for the specific Minkowski functionals are known. We show that in the Gaussian setting the Ohser-Mücklich estimators are asymptotically unbiased if we let the spacing of the pixel grid tend to zero and provide explicit expressions for the bias of these estimators. Furthermore we consider

a certain class of general surface estimators based on the two-point covariance function. We prove asymptotical unbiasedness in the Gaussian Model also for estimators of this more general class and again give explicit expressions for the bias. In the one-dimensional case we additionally consider the variance of these estimators and prove asymptotic normality. We provide algorithms how to generate two- and three-dimensional artificial structures with predetermined specific Minkowski functionals in the Thresholded Gaussian Model and show results from our experiments which share their Minkowski-functionals with the synchrotron sand-data shown in the introduction. Recent experiments at the Institute of Terrestrial Ecology have proved that our artificial Gaussian structures are able to mimic the flow- and transport properties of their real world counterparts quite well.

The fifth chapter is devoted to the Gibbsian Model known from statistical mechanics. We discuss the equivalence of neighbourhood-Gibbs-Fields with Markovian Random Fields given by the Hammersley-Clifford Theorem and describe how the Gibbs-potential must be chosen to generate artificial structures with predetermined Minkowski functionals. Next we deal with the variational principle of statistical mechanics which characterizes the Gibbsian model as the one maximizing entropy for a given energy. This makes the Gibbsian Model a natural choice to generate porous structures arising in nature. We also describe the equivalence of ensembles which states that in the thermodynamic limit (on large lattices) the Gibbsian distribution with a suitably chosen potential is in fact equivalent to the uniform distribution on the set of images with the same specific Minkowski functionals. This result will be useful later on to simulate from the Gibbsian model without having to estimate any parameters at all! The rest of the chapter deals with simulation within the Gibbsian Model. We discuss the well-known Gibbs sampler and describe how it can be used to efficiently simulate from the Gibbsian distribution even on large lattices due to the Markov property. We describe methods how those parameters of the Gibbsian distribution which correspond to the predetermined values for the Minkowski functionals can be estimated and provide a method to simulate from the Gibbsian Model with predetermined Minkowski functionals which gets along without any parameter-estimation at all. This works because due to the equivalence of ensembles we can simulate from the uniform distribution of images with the desired specific Minkowski functionals instead of the Gibbsian distribution. To simulate from this uniform distribution we use a combination of the Gibbs-Sampler with the Simulated-Annealing-technique for optimization. With this Simulated-Annealing-method we are able to generate artificial random structures which not only share the same specific Minkowski functionals with their real world counterparts but also (at least in principle) the same values for arbitrarily many additional geometrical characteristics. At the end of the chapter we show two-dimensional Gibbsian realizations whose Minkowski functionals agree with those of selected cross-sections taken from the synchrotronized sand-images.

Finally some concluding remarks are given in chapter six. The appendix contains the proof of asymptotic normality for general surface estimators in the continuous Markovian on-off-system. This example is meant to be an introductory example for the same proof for Thresholded Gaussian Fields presented in chapter four. It can be used as a first approach to the concepts of chapter four in an easily tractable setting. The appendix contains further the proofs of some fundamental theorems used to outline the basic theory which were omitted in the main text and also contains alternative proofs to some of our own results.

Seite Leer /  
Blank leaf

# Zusammenfassung

Poröse Strukturen treten in der Natur und den Materialwissenschaften sehr häufig und in den mannigfaltigsten Ausprägungen auf. Beispiele aus der Natur sind etwa Gesteine, Bodenstrukturen, Schwämme, oder biologisches Gewebe. Zu den porösen Werkstoffen zählen Zement, Keramik und Schäume. Es ist daher kaum verwunderlich, dass das Studium der porösen Medien schon seit geraumer Zeit, insbesondere aber seit den 70er Jahren des letzten Jahrhunderts grosses interdisziplinäres Interesse genießt. Nicht nur für Mathematiker, Geophysiker, Ingenieurwissenschaftler (Material- und Werkstoffwissenschaften) Biologen und Chemiker ist die Modellierung poröser Medien von Interesse, sondern nicht zuletzt aufgrund der gewaltigen Flut zu bearbeitender Datenmengen auch für Informatiker und die Community der Bildverarbeitung. Dank der rasanten Entwicklung der Synchrotron-Technologie sind in den letzten paar Jahren hochaufgelöste Darstellungen von dreidimensionalen Proben solcher Medien verfügbar geworden. Das Institut für terrestrische Ökologie der ETH Zürich und das Paul Scherrer Institut (PSI) in Villigen sind in die Herstellung solcher Bilder für Sandbodenstrukturen verschiedener Granularität involviert und haben uns diese, für die vorliegende Doktorarbeit freundlicherweise zur Verfügung gestellt.

Die geometrische Struktur des Porenraums beeinflusst wesentlich die Fließ- und Transporteigenschaften von Flüssigkeiten, wie etwa die Permeabilität in porösen Medien. Dabei ist es allerdings noch weitgehend unklar, welcher Art diese Beeinflussung ist und welche geometrischen Charakteristika des Porenraums entscheidend sind. Wir konzentrieren uns in dieser Arbeit hauptsächlich auf eine bestimmte einfache Klasse solcher Charakteristika, die sog. Minkowski-Funktionale. Zu ihnen gehören die elementaren geometrischen Grössen Volumen (des Poren- bzw. Komplementär-raums), Oberfläche (der Grenze zwischen Poren- und Komplementär-raum), das mittlere Krümmungsintegral (dieser Grenzfläche) und die Euler-Charakteristik. Letztere ist eine wichtiges Mass aus der Differentialtopologie, welches den Grad des Zusammenhangs einer Struktur misst. Es gibt eine theoretische Rechtfertigung dafür, dass Minkowski-Funktionale gut dafür geeignet sind, die in einem Bild vorhandene geometrische Information zusammenzufassen. Diese wird durch den fundamentalen Satz von Hadwiger geliefert.

Zur Modellierung der porösen Strukturen bedienen wir uns in dieser Arbeit der Stochastik. Unser Ziel ist es daher nicht, eine gegebene natürliche Struktur zu reproduzieren, sondern nur ihre stochastischen Eigenschaften. Dies hat den Vorteil, dass wir verschiedene Realisierungen eines Modells bezüglich ihrer geometrischen Eigenschaften miteinander vergleichen und ihre Variabilität untersuchen können. Wir gehen dabei wie folgt vor: Wir messen die geometrischen Eigenschaften (vorwiegend die Minkowski-Funktionale) einer gegebenen Struktur und bestimmen dann die Parameter eines statistischen Modells so, dass dessen zufällige Realisierungen im Mittel dieselben geometrischen Eigenschaften haben wie die gegebene poröse Struktur. Wir betrachten hauptsächlich drei einfache stochastische Modelle: das Boolesche Keim-Korn Modell, trunkierte Gaußsche Zu-

fallsfelder und das Gibbs-Modell aus der statistischen Mechanik. Bei der Auswahl der Modelle war uns wichtig, dass die betrachteten Modelle bis zu einem gewissen Grad analytisch zugänglich sind, wohlwissend, dass zur Modellierung poröser Medien oft realistischere Modelle, wie etwa das Kirschkerne-Modell (cherry-pit, Hardcore-Type) vorgezogen werden, die für analytische Betrachtungen aber weit weniger geeignet sind. Die Hauptziele der vorliegenden Arbeit waren herauszufinden, in welchem Mass wir mit diesen einfachen Modellen künstliche Strukturen mit vorgegebenen geometrischen Charakteristiken erzeugen können und inwieweit diese künstlichen Realisierungen sich trotz übereinstimmender Charakteristiken von der gegebenen natürlichen Struktur unterscheiden. Desweiteren mussten Algorithmen entwickelt werden, die mit der beträchtlichen Grösse der zu verarbeitenden Datenmengen umgehen können, um diese einfachen Modelle den Daten anzupassen, und wir haben diese Algorithmen anhand der hochauflösenden Synchrotron-Sandbilder getestet. Zusammen mit dem Institut für terrestrische Ökologie der ETH Zürich haben wir ferner untersucht, ob unsere künstlichen Realisierungen ähnliche Fluss- und Transporteigenschaften aufweisen, wie die realen Sandstrukturen.

Im Detail gliedert sich die vorliegende Arbeit wie folgt:

In einem Einführungskapitel wird kurz der allgemeine Modellierungsprozess beschrieben, den wir im Rest der Arbeit für die spezifischen Modelle adaptieren und es werden die Synchrotron-Sanddaten, die wir zum Testen unserer Algorithmen verwenden, vorgestellt und analysiert.

Das zweite Kapitel beschäftigt sich mit den theoretischen Grundlagen, die wir in dieser Doktorarbeit benötigen. Es beginnt mit einer Einführung in die Brunn-Minkowski-Theorie, wo Minkowski-Funktionale auf dem Konvexring eingeführt und ihre Eigenschaften diskutiert werden. Ein besonderer Fokus liegt dabei auf der Eulercharakteristik, aus welcher sich die anderen Minkowski-Funktionale ableiten lassen. Dies ist im wesentlichen die Bedeutung der Crofton-Formeln, die wir in der anschliessenden Einführung in die Integralgeometrie diskutieren. Desweiteren beschreiben wir, wie sich die spezifischen Minkowski-Funktionale einer porösen Struktur aus einem Pixelbild approximieren lassen. In der ganzen Arbeit verwenden wir dabei die sogenannten Ohser-Mücklich-Schätzer, welche sich direkt aus der Diskretisierung der Integrale in den Crofton-Formeln ableiten lassen. Nach diesen rein deterministisch-geometrischen Betrachtungen folgt eine kurze Einführung in die stochastische Geometrie, wo vor allem die Konzepte der Punkt- und Partikelprozesse und allgemeiner der zufälligen Mengen erläutert werden. Das Kapitel schliesst mit einer Zusammenstellung derjenigen geometrischen Charakteristiken, die nebst den Minkowski-Funktionalen in der Literatur zur Beschreibung poröser Strukturen am häufigsten verwendet werden.

Das dritte Kapitel ist dem Booleschen Keim-Korn-Modell gewidmet. Wir fassen das Boolesche Modell als einen Partikelprozess auf und zeigen, wie sich die spezifischen Minkowski-Funktionale im Booleschen Modell aus den Beziehungen der Integralgeometrie ableiten lassen. Wir diskutieren, welche spezifischen Minkowski-Funktionale überhaupt mit dem Booleschen Modell erreichbar sind. Dann widmen wir uns der Simulation von zwei und dreidimensionalen Booleschen Strukturen mit vorgegebenen spezifischen Minkowski-Funktionalen. Wir geben Algorithmen an für das Boolesche Modell mit sphärischen und ellipsoidalen Körnern und zeigen zwei und dreidimensionale künstliche Boolesche Strukturen aus unseren Experimenten mit denselben spezifischen Minkowski-Funktionalen wie die Synchrotron-Sandbilder. Zum Schluss dieses Kapitels diskutieren

wir die Sensitivität der Booleschen Strukturen, wenn einzelne spezifische Minkowski-Funktionale variiert werden.

Das vierte Kapitel widmet sich hauptsächlich dem Gaußschen Zufallsfeld-Modell. Wir können binäre Bilder erzeugen, indem wir bei einer beliebigen Schwelle trunkieren. Auch im Gaußschen Modell sind explizite Formeln für die spezifischen Minkowski-Funktionale bekannt. Wir zeigen, dass die Ohser-Mücklich-Schätzer asymptotisch erwartungstreu sind, wenn die Gitterkonstante gegen null geht. Wir geben auch explizite Ausdrücke für den Bias dieser Schätzer an. Desweiteren betrachten wir eine allgemeinere Klasse von Oberflächenenschätzern, die auf der Kovarianzfunktion des Zufallsfeldes basieren. Auch hier beweisen wir asymptotische Erwartungstreue und geben den Bias explizit an. Im eindimensionalen Fall berechnen wir auch die Varianz solcher Oberflächenestimator und beweisen asymptotische Normalität für das trunkierte Gaußsche Zufallsfeld-Modell. Desweiteren geben wir Algorithmen an, wie man in zwei und drei Dimensionen im Gaußschen Zufallsfeld-Modell Realisierungen mit vorgegebenen spezifischen Minkowski-Funktionalen erzeugen kann und zeigen solche zwei- und dreidimensionale Realisierungen aus unseren Experimenten mit denselben spezifischen Minkowski-Funktionalen wie die Synchrotron-Sandbilder. Neue Experimente am Institut für terrestrische Ökologie haben gezeigt, dass unsere Gauss-Realisierungen, die Fluss- und Transporteigenschaften des realen Mediums sehr gut wiedergeben können.

Im fünften Kapitel verwenden wir das Gibbsche Modell aus der statistischen Mechanik, um zufällige Strukturen mit vorgegebenen Minkowski-Funktionalen zu erzeugen. Am Anfang des Kapitels arbeiten wir die Äquivalenz von Gibbschen Feldern mit Nachbarschaftspotentialen und Markovfeldern heraus (Hammersley-Clifford-Theorem). Wir beschreiben, wie das Gibbspotential gewählt werden muss, um vorgegebene spezifische Minkowski-Funktionale zu erzeugen. Dann beschäftigen wir uns mit dem Variationsprinzip der statistischen Mechanik, welches das Gibbs-Modell als dasjenige mit der maximalen Entropie auszeichnet und es deshalb als Modell für natürliche Porenstrukturen geeignet erscheinen lässt. Als nächstes gehen wir auf die Äquivalenz von Teilchensembles ein, die besagt, dass sich im thermodynamischen Limes (grosse Gitter) die Gibbsverteilung mit geeignetem Potential als die Gleichverteilung auf der Menge aller Bilder mit den vorgegebenen Minkowski-Funktionalen herausstellt. Diese Erkenntnis kann später zum Simulieren, ohne vorher Parameter schätzen zu müssen, verwendet werden. Als nächstes wenden wir uns der Simulation von Gibbs-Modellen zu. Wir diskutieren den Gibbs-Sampler, der die Simulation von der Gibbs-Verteilung dank der Markov-Eigenschaft sehr einfach macht. Dann wenden wir uns dem Problem der Parameterschätzung zu und diskutieren verschiedene Methoden, wie man diejenigen Parameter im Gibbs-Modell schätzen kann, die den vorgegebenen Minkowski-Funktionalen entsprechen. Schliesslich diskutieren wir auch eine völlig parameterfreie Methode. Dank der Äquivalenz der Teilchensembles können wir statt von der Gibbs-Verteilung von der Gleichverteilung auf der Menge aller Bilder mit den vorgegebenen Minkowski-Funktionalen simulieren. Dies erreichen wir durch eine Kombination des Gibbs-Samplers mit der Simulated-Annealing-Optimierungstechnik. Diese Simulated-Annealing-Methode erlaubt es uns, Gibbs-Realisierungen zu erzeugen, die nicht nur vorgegebene Minkowski-Funktionale aufweisen, sondern zusätzlich vorgegebene Werte für (zumindest im Prinzip) beliebig viele andere geometrische Charakteristiken. Wir schliessen das Kapitel ab mit zweidimensionalen Bildern aus unseren Experimenten, welche in verschiedenen geometrischen Charakteristiken mit den Synchrotron-Sandbildern übereinstimmen.

Das abschliessende sechste Kapitel enthält einige Schlussbemerkungen. Im Appendix findet man den Beweis der asymptotischen Normalität der verallgemeinerten Oberflächenschätzer für das stetige Ein-/Ausschalt-Markovmodell. Dieses Beispiel ist als Einführungsbeispiel für denselben Beweis aus Kapitel vier für das Gaussche Zufallfeld-Modell gedacht und eignet sich für eine erste Annäherung an die im Kapitel vier verwendeten Konzepte in einem einfach behandelbaren Rahmen. Weiter befinden sich im Appendix die Beweise zu einigen der im Aufbau der Theorie benötigten Sätze, die wir im Haupttext ausgelassen haben. Der Leser findet dort auch alternative Beweise zu einigen unserer eigenen Resultate.

Maik Berchtold

Zürich, im November 2007



# Chapter 1

## Introduction and Problem Formulation

### 1.1 General Introduction

### 1.2 General Problem Formulation

In this section we try to formulate the general problem one faces when trying to model two-phase porous structures by means of Minkowski functionals. Any details and exact definitions can be found in later chapters. The *general modelling problem* can be summarized into three basic steps:

#### **Step 1: Measuring Minkowski Functionals for the Structure to Simulate: The Ohser Mücklich-Estimators**

Usually one starts out with the analysis of a  $d = 2$  or  $d = 3$  -dimensional real data sample  $D$ . Because the porous structures we are interested in consist of two phases, we must always specify whether we calculate the Minkowski-functionals for the solid phase  $P$  (black or 1-values) or for the pore phase (white or 0-values). When we talk about Minkowski-functionals  $V_j(P)$ ,  $j \in \{0, \dots, d\}$  associated with a sample  $D$  in the following, we always mean the Minkowski-functionals of its associated solid phase  $P$ . Of course one could also concentrate on the pore phase instead, different choices merely represent different research questions of interest. For example in soil flow physics where one is interested in modelling the fluid flow through a porous medium or in percolation theory the focus is more on the pore phase. As far as the choice of models is concerned, it turns out that this distinction is more than just a matter of taste. Sometimes models that work well for the solid phase might perform poorly when modelling the pore phase and vice versa although the (specific) Minkowski functionals of a set and its complement are closely related if both are defined.

The main goal for us is to generate an *artificial* porous structure  $S$  which has the same Minkowski functionals as the solid phase  $P$  of the given sample  $D$ , hence our *target values*  $v_j$  the Minkowski-functionals of  $S$  should have are  $v_j := V_j(P)$ . To work with such a sample it has to be available in a digitized version as a binarized pixel image. For such a binarized image representation Minkowski-functionals must be measured. For this purpose usually the cuboidal graph  $\Gamma$  whose vertices are the image pixels is considered. Using  $\Gamma$ , a

continuous graph approximation  $P^\Gamma$  of the solid phase  $P$  is constructed according to the following construction rules:

**Algorithm 1.2.1. (Construction of the Graph Approximation Set)**

- A cubic voxel of  $\Gamma$  belongs to  $P^\Gamma$  if and only if all its 8 pixel vertices belong to the solid phase  $P$  of  $D$ .
- A voxel face belongs to  $P^\Gamma$  if and only if all its four pixel vertices belong to  $P$ .
- A graph edge belongs to  $P^\Gamma$  only if both its pixel vertices belong to  $P$ .
- All isolated pixel vertices belong to  $P^\Gamma$ .

Of course the resulting set  $P^\Gamma$  is only a rough approximation of the solid phase  $P$  of  $D$ . Especially if  $P$  is not convex, the line connecting two pixel vertices in the solid phase belongs entirely to  $P^\Gamma$  but not necessarily to the solid phase  $P$  itself. It is therefore obvious that  $V_j(P^\Gamma) \neq V_j(P)$  in general and the differences may be substantial. Another reason for this discrepancy is the fact that Minkowski functionals of  $P$  are closely related to curvature quantities for the interface between the two phases in  $D$ , while the boundary of  $P^\Gamma$  is following the voxel boundaries in the cuboidal graph  $\Gamma$  and therefore is rather rough. Thus instead of really measuring  $V_j(P)$  all we can hope is being able to *estimate*  $V_j(P)$  as  $V_j(P^\Gamma)$ . Ohser and Mücklich have developed estimators which make it possible to estimate the  $V_j(P^\Gamma)$  accurately only by counting pixel configurations in local neighborhoods of the vertices in  $P^\Gamma$  (see sections 2.3.1 and 2.3.2 for a detailed discussion of these estimators.) This fast and easy estimation is possible by discretizing the so-called *Crofton formulae* (discussed in section 2.2.1) from integral geometry. The Crofton formulae give representations of the Minkowski functionals  $V_j(P)$  as certain spatial and rotational averages of simple quantities which can locally be evaluated by pixel counting. The continuous spatial and rotational averages are replaced by averaging over directions that are defined by the voxel vertices in a pixel lattice. Hence a further approximation error comes in, which obviously doesn't need to vanish if the grid-spacing becomes small.

Some models such as the *Thresholded Gaussian Field Model on a Lattice* and the *Gibbsian Model* are *discrete* which means that they don't attempt to model the (continuous) solid phase of  $P$  but only its digitized version  $P^L$  directly. Of course the images  $S$  generated with these models are nevertheless interpreted as *continuous* images according to the above construction rules.

Sometimes the target Minkowski values  $v_j$  are not given by real data samples but set artificially. This is mainly the case if one wants to investigate whether a specific set of predetermined Minkowski values  $v_j$  can be achieved within a specific model or when one is interested in how the properties of an artificially generated structure change with respect to changes of their Minkowski functionals. Such an analysis is performed in (Lehmann, Berchtold, Ahrenholz, Tölke, Krafczyk, Flühler and Künsch n.d.) where together with soil-physicists we investigated how the permeability and fluid transport properties of a porous medium depend on the Minkowski functionals within a stationary and isotropic Boolean Model with elliptical grains.

### Step 2: Connection between Data and Stochastic Models: The Basic Modelling Equation

To create such an artificial structure which has predefined Minkowski-values  $v_j$  per unit volume, we use *stochastic* models. Concretely we construct an (infinitely extended) stationary and isotropic random set  $\Theta$  in Euclidean space  $\mathbb{R}^d$  and choose an extract  $\Theta^W$  of it which has the same size and form as the given data sample  $D$ . The stochastic interpretation is as follows: The random set  $\Theta$  is observed through a so-called *observation window*  $W$  where the set  $W$  is a compact (and for technical reasons also convex) subset of  $\mathbb{R}^d$  of the same form and size as  $D$ .  $\Theta^W$  is then given by  $\Theta^W = \Theta \cap W$ . Because  $\Theta$  and hence also  $\Theta^W$  are random, every realization  $\Theta^W(\omega)$  will have different Minkowski-functionals  $V_j(\Theta^W(\omega))$ . Because of this stochastic modelling approach we cannot expect to find a realization which *exactly* fulfills  $v_j = V_j(\Theta^W(\omega))$ , we must rather find a way to express the Minkowski functionals per unit volume of the random set  $\Theta$  as a *whole* independent of both specific realizations and the observation window  $W$  we chose. This is provided by the *specific Minkowski-functionals*  $\bar{V}_j(\Theta)$  which involve taking expectation of the Minkowski functionals per unit volume over all realizations of  $\Theta^W$ . The connection between the measured or predefined Minkowski values per unit volume  $v_j$  and the specific Minkowski functionals  $\bar{V}_j(\Theta)$  of the stochastic model is given by the following requirement which we call **Basic Modelling Equations**:

$$v_j \stackrel{!}{=} \bar{V}_j(\Theta), \quad j \in \{0, \dots, d\}. \quad (1.1)$$

This means that because of the stochastic nature of our modelling approach we will only be able to match the predefined Minkowski values  $v_j$  in expectation and for a specific realization the specific Minkowski functionals will slightly differ from the preset values. For most models the right hand side of equation 1.1 can not analytically be calculated and nothing about the variance of the Minkowski functionals per unit volume is known. However there are exceptions such as the *Boolean Model* or the *Thresholded Gaussian Field Model* which we investigate in the following chapters.

### Step 3: Choice of Distributions and Determination of Parameters

The right hand side of the equation system 1.1 usually contains the model parameters which determine the flexibility of a specific model. If all model parameters are simply numbers or vectors and if the equations in 1.1 can be solved for these model parameters analytically, the determination of a right set of parameters leading to a random set which fulfills the basic modelling equations is straightforward, otherwise they *must* be solved numerically and another approximation error enters the scene. The parameter set may also include distributions, eg. the grain distribution  $Q$  in the case of the Boolean Model. In such a case, one usually chooses a suitable *parametric* class of distributions and uses the parameters of this class instead of the distribution itself. The final parameter sets often contains more than only  $d + 1$  parameters which means that the equation system 1.1 is under-determined. This phenomenon is referred to as *indeterminacy* and choosing different solutions makes it possible to generate a wide range of random artificial structures with rather different structural properties and which all share the same set of Minkowski functionals per unit volume within the same model.

### 1.3 Data

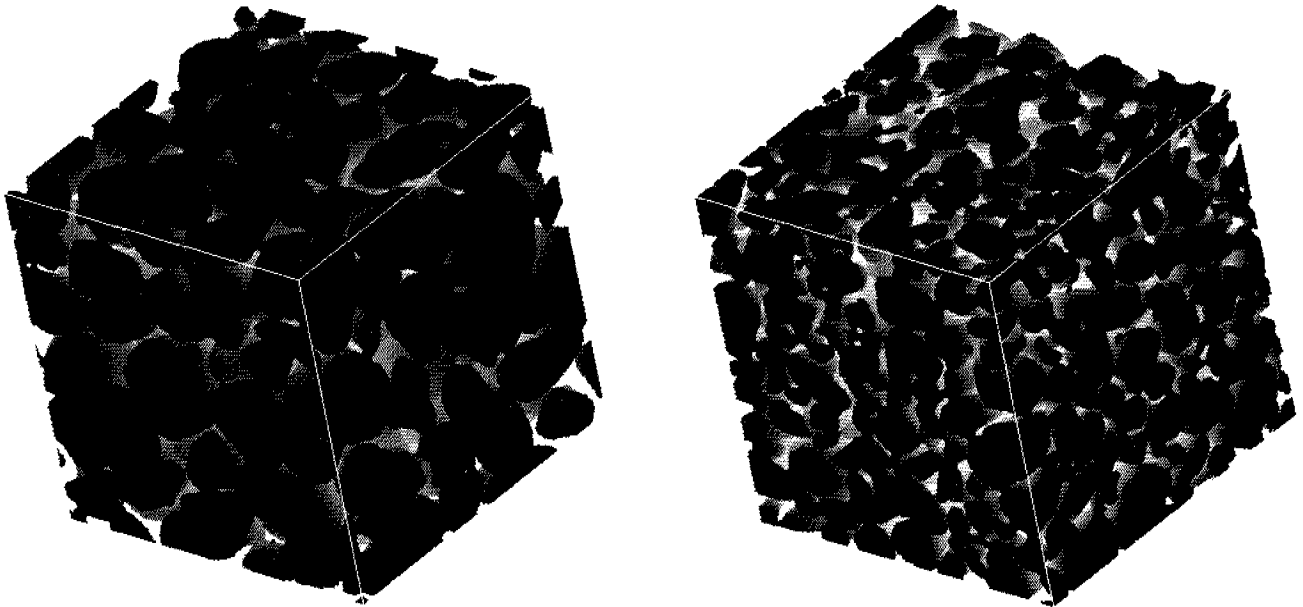
Nowadays the imaging power of modern X-ray synchrotrons is on such an advanced level that the resolution of a sample cube with voxel size as small as  $10\mu\text{m}$  may distinguish far more than  $10^9$  pixels. Thus the resulting data sets are of massive size which makes their sheer handling a non-trivial task. Especially data-storage issues may arise. Usually it is impossible to load such an image into the memory of an ordinary workstation as a whole and still be able to manipulate it in reasonable time even if a pixel is represented by a just a single bit. The consequence for displaying such 3D geometries on a screen is that one can only display relatively small extracts of the whole structure at a time or one has to reduce resolution. From the computational point of view one must always be aware that standard algorithms applied to such data sets must be adapted, e.g. the images be split up in smaller pieces and loaded into memory one by one, which may affect performance heavily. We therefore decided to write all our programs in a low-level language and chose the *C programming language*. This had the drawback that functionality from powerful statistical software packages such as *R* were available only in a reduced amount or via special interfaces. For displaying three dimensional structures we used the extremely powerful module-based AVS Software (Advanced Visual Systems 2006).

The data we use in this thesis was kindly made available to us by courtesy of the *Institute of Terrestrial Ecology (ITO)* at the *Swiss Federal Institute of Technology (ETH)*. The data consists of a collection of 3D binary images of cubic sand samples generated using X-ray technology at the *Hamburg synchrotron laboratories (HASYLAB)* in Germany. There are two types of sands present in the samples with different granularity ("coarse" and "fine"). The images are  $800^3$  pixels in size and the size of a single voxel is  $11\mu\text{m}$ . The binary coding is such that the value 1 is attributed to pixels in the solid phase whereas the pixels in the pore phase are denoted by 0. Two examples are shown in figure 1.1.

We now shortly describe the results of a basic image analysis for these data. We calculate the specific Minkowski functionals, the two-point-covariance function and the chord-length-distribution function for both sand samples. The following table shows the estimated specific Minkowski functionals for the solid phase which were estimated according to the Ohser-Mücklich-procedure presented in subsections 2.3.2. In three dimensions Minkowski functionals are up to proportionality volume  $v$ , surface area  $s$ , integral of mean curvature  $M$  and Euler-characteristic  $\chi$ . A detailed discussion of the latter two geometrical characteristics is given in subsections 2.1.5 and 2.2.4 below. We indicate here the estimated *specific* Minkowski quantities, i.e. the quantities per unit volume of  $W$  since this will allow us convenient comparison with the artificially generated structures later. For completeness also the estimated Minkowski quantities for the pore space are given.

	coarse / solid $P$	coarse / pore $P^c$	fine / solid $P$	fine / pore $P^c$
$\hat{v}(\cdot)/v(W)$	$0.6043 \cdot 10^{-0}$	$0.3968 \cdot 10^{-0}$	$0.5981 \cdot 10^{-0}$	$0.4068 \cdot 10^{-0}$
$\hat{s}(\cdot)/v(W)$	$0.6870 \cdot 10^{-1}$	$0.7089 \cdot 10^{-1}$	$0.1061 \cdot 10^{-0}$	$0.1074 \cdot 10^{-0}$
$\hat{M}(\cdot)/v(W)$	$0.1174 \cdot 10^{-2}$	$-0.1288 \cdot 10^{-2}$	$0.1204 \cdot 10^{-2}$	$-0.1338 \cdot 10^{-2}$
$\hat{\chi}(\cdot)/v(W)$	$-0.2112 \cdot 10^{-4}$	$-0.2188 \cdot 10^{-4}$	$-0.1165 \cdot 10^{-3}$	$-0.1210 \cdot 10^{-3}$

**Table 1.1:** *Minkowski functionals per unit volume for the sand sample from figure 1.1. The values for both the solid and the pore phase are provided.*



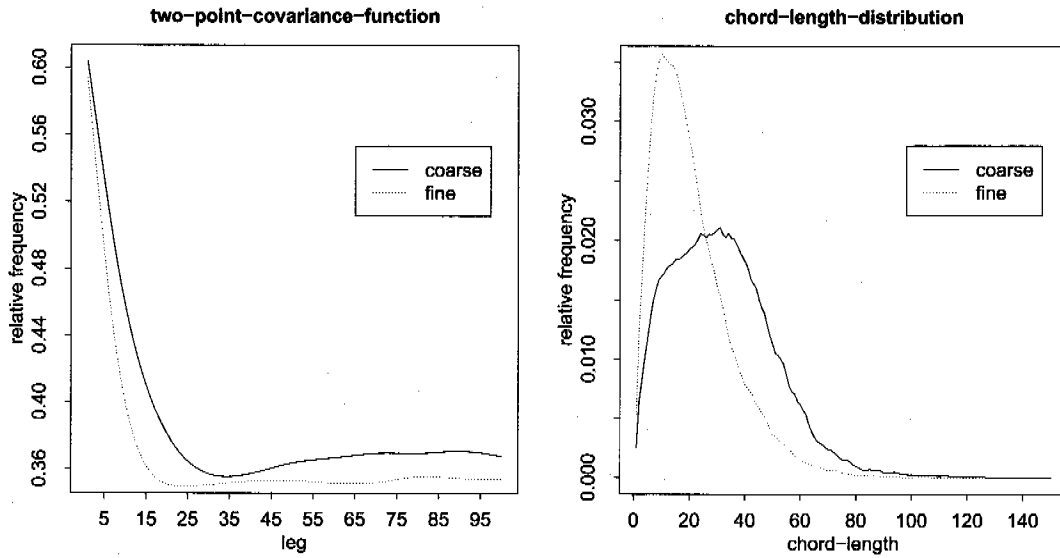
**Figure 1.1:** *Two  $256^3$ -pixel-sized cubes taken from the  $800^3$ -pixel-sized samples of the HASYLAB synchrotron sand data with side length =  $11\mu\text{m}$ . The left image shows a sand-type with coarse granularity whereas the image on the right displays a sand type with fine granularity.*

The comparison between the values for the solid and pore space in table 1.1 reveals the approximation errors when estimating Minkowski-functionals of a continuous image from a grid approximation via the Ohser-Mücklich procedure. Obvious equalities such as  $v(P)/v(W) + v(P^c)/v(W) = 1$  or  $s(P)/v(W) = s(P^c)/v(W)$  are only approximately fulfilled by the estimators for both sand types. For the integral of mean curvature we expect a sign flipping when changing from  $P$  to  $P^c$  whereas in three dimensions the Euler-characteristic of a set  $P$  and its complement  $P^c$  agree. (However be aware of Theorem 2.1.13).

Figure 1.2 shows the empirical *two-point-correlation-function* and the empirical *chord-length-distribution* for the solid phase of the two data samples in figure 1.1. For a definition of these functions see subsection 2.5.

The data we use in our 2D experiments are square  $800^2$ -cross-sections from the HASYLAB samples in figure 1.1. These were taken parallel to sides of the sample cubes, eg. parallel to the  $xy$ -,  $xz$ - or  $yz$ -planes. Figure 1.3 shows an example for each of the three directions. These cross-sections were taken from the center of the sample.

Figure 1.3 shows that the data samples exhibit a high degree of homogeneity and isotropy, hence the approach to model these structures with stationary and isotropic models seems to be quite reasonable. Another justification for stationary modelling can be given by investigating how Minkowski-functionals in equally-sized subcubes of the  $800^3$  samples are distributed. If the variation in Minkowski-values among these subcubes is sufficiently small, this constitutes another indication for a high degree of stationarity. The boxplots in figures 1.4 and 1.5 show the distributions of Minkowski-functionals



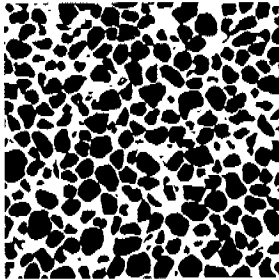
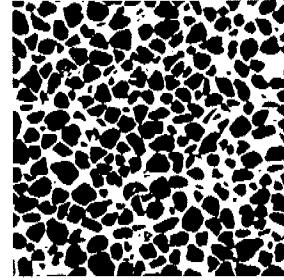
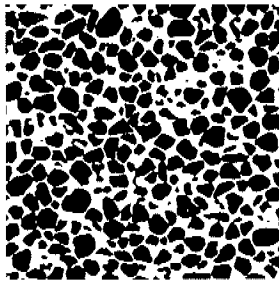
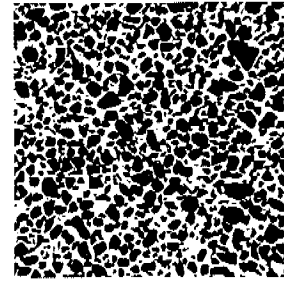
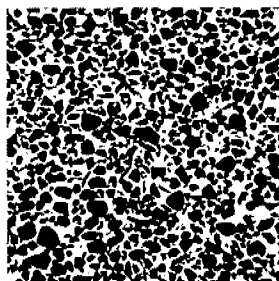
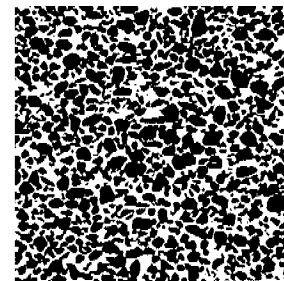
**Figure 1.2:** The two-point-covariance-function (left) and the chordlength-distribution function (right) for the coarse and fine data sample in figure 1.1.

among 1000 partly overlapping subcubes of size  $256^3$  for the coarse as well as for the fine data sample. Examples for such subcubes for each sand type are displayed in figure 1.1.

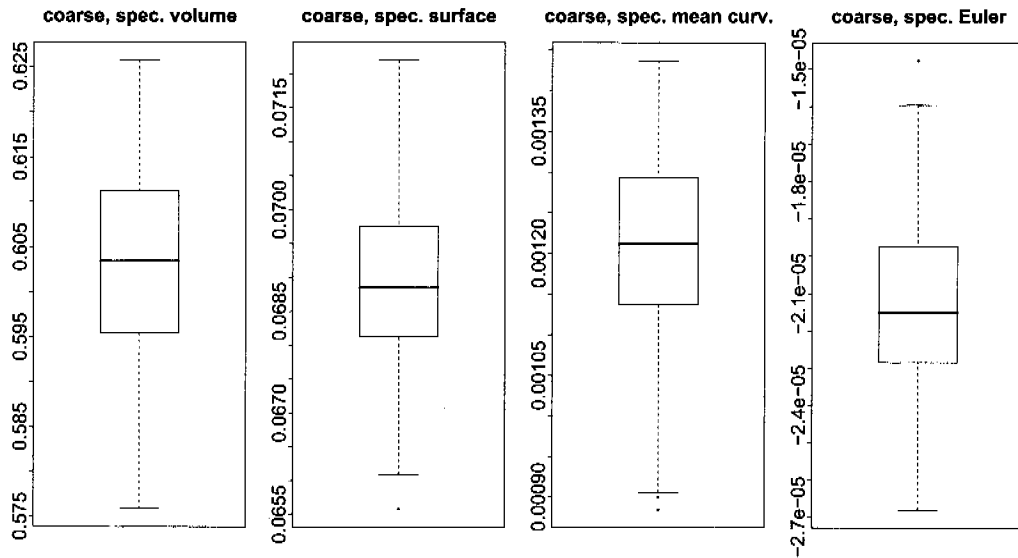
Mean value and standard deviation from the subcube-analysis are collected in table 1.2:

	$\bar{v}$	$\bar{s}$	$\bar{M}$	$\bar{\chi}$
coarse, mean	$0.6032 \cdot 10^{-0}$	$0.6897 \cdot 10^{-1}$	$0.1212 \cdot 10^{-2}$	$-0.2134 \cdot 10^{-4}$
coarse, sd	$0.1081 \cdot 10^{-1}$	$0.1150 \cdot 10^{-2}$	$0.1016 \cdot 10^{-3}$	$0.2262 \cdot 10^{-5}$
fine, mean	$0.5983 \cdot 10^{-0}$	$0.1061 \cdot 10^{-0}$	$0.1235 \cdot 10^{-2}$	$-0.1167 \cdot 10^{-3}$
fine, sd	$0.4193 \cdot 10^{-2}$	$0.1267 \cdot 10^{-2}$	$0.1321 \cdot 10^{-3}$	$0.5041 \cdot 10^{-5}$

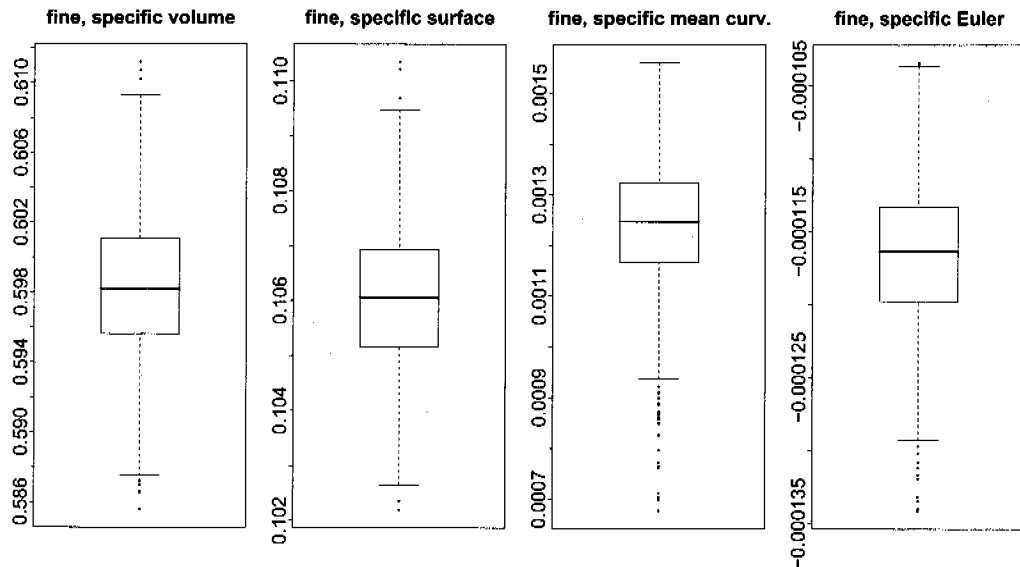
**Table 1.2:** Mean and standard deviation (sd) for the specific Minkowski functionals  $\bar{v}, \bar{s}, \bar{M}, \bar{\chi}$  of 1000 partly overlapping  $256^3$ -pixel sized extracts of the coarse and fine sand samples in 1.1

cross-section coarse,  $z = 400$ cross-section coarse,  $y = 400$ cross-section coarse,  $x = 400$ cross-section fine,  $z = 400$ cross-section fine,  $y = 400$ cross-section fine,  $x = 400$ 

**Figure 1.3:** The first three cross-sections are taken from the full  $800^3$ -pixel-sized coarse-sand-sample in figure 1.1. They represent the mid planes of the coarse-sand-sample-cube in  $z$ -,  $x$ - and  $y$ -directions. The last three images display the corresponding cross sections for the  $800^3$ -pixel-sized fine-sand-sample-cube from figure 1.1.



**Figure 1.4:** Boxplots of Minkowski functionals from 1000 partly overlapping subcubes of size  $256^3$  taken from the coarse sand sample in 1.1



**Figure 1.5:** Boxplots of Minkowski functionals from 1000 partly overlapping subcubes of size  $256^3$  pixels taken from the fine sand sample in figure 1.1.



# Chapter 2

## Preliminaries

### 2.1 Definition and Properties of Minkowski -Functionals

This first chapter is dedicated to the exposition of the theoretical background for this thesis and also to fix the notation. The material collected in this chapter concerns facts from various disciplines in geometry such as convex geometry, integral geometry and stochastic geometry. In our presentation we closely follow standard works of reference from these disciplines among which the most important are (Schneider and Weil 1992, Schneider and Weil 2000) and (Stoyan, Kendall and Mecke 1985).

#### 2.1.1 The Steiner Formula

The Minkowski-functionals arise quite naturally from the study of a basic question in convex geometry. Let  $\mathcal{K}$  be the collection of all convex bodies, i.e sets that are convex and compact in  $\mathbb{R}^d$ . The collection  $\mathcal{K}$  can be equipped with the *Hausdorff-metric*  $d_H$  which leads to the metric space  $(\mathcal{K}, d_H)$ . For each  $K \in \mathcal{K}$  and  $\rho > 0$  we define the *parallel set*  $K_\rho = \{x + y \mid x \in K, y \in B_\rho^d(0)\}$  where  $B_\rho^d(0)$  is the  $d$ -dimensional ball centered at 0 with radius  $\rho$ . Is it possible to determine the *volume*  $V_d(K_\rho)$  of the parallel set  $K_\rho$  for arbitrary  $\rho > 0$  using only information about the convex set  $K$  itself? The answer is yes. The volume  $V_d(K_\rho)$  can even be expressed as a polynomial of degree  $d$  whose coefficients depend on the original convex body  $K$  only. This assertion is provided by the famous *Steiner-Formula*:

**Theorem 2.1.1. (Steiner-Formula)** *There exist unique functionals  $V_m : \mathcal{K} \rightarrow \mathbb{R}$ ,  $m = 0, \dots, d$  such that for  $K \in \mathcal{K}$  and  $\rho \geq 0$*

$$V_d(K_\rho) = \sum_{m=0}^d \rho^{d-m} \kappa_{d-m} \cdot V_m(K),$$

where  $\kappa_d := V_d(B^d)$  is the volume of the unit ball in dimension  $d$ .

The Steiner-Formula gives rise to a first definition of the Minkowski-functionals as follows:

**Definition 2.1.2. (intrinsic volumes and Minkowski-functionals)**

- *i) For  $m \in 0, \dots, d$  the coefficients  $V_m(K) : \mathcal{K} \rightarrow \mathbb{R}$  in the Steiner-Formula are called **intrinsic volumes** of the convex body  $K$ .*
- *ii) The functionals  $W_m : \mathcal{K} \rightarrow \mathbb{R}$ ,  $m = 0, \dots, d$  which are scaled versions of the intrinsic volumes  $V_m(K)$  such that  $\kappa_m \cdot V_{d-m}(K) = \binom{d}{m} \cdot W_m(K)$ , are called the **Minkowski-functionals** of a convex body  $K$ .*

To which of the two sets of functionals preference is to be given is just a matter of taste. The choice is often motivated by the aesthetics of relevant formulac. We will use both sets of functionals in the following, whenever the one is more convenient than the other, under the tacit assumption that all theorems, properties and techniques could also be stated for the other functional-set respectively.

Because the basic definition of Minkowski-functionals is essentially based on Theorem 2.1.1 a proof of the Steiner-Formula following (Schneider 2004) will now be given. The same reference was consulted for most of the integral-geometry related theorems of this chapter.

To motivate the idea of the proof, look at the case where  $d = 2$  and the set  $K$  is sufficiently simple, eg. a triangle  $\Delta$ . The area of the parallel set  $\Delta_\rho$  for this triangle then consists of the area of the triangle  $\Delta$  itself plus the area of three rectangles along the three sides of the triangle (area = length of the triangle side times  $\rho$ ) plus the area of the circle with radius  $\rho$  (split into three sectors at each of the three vertices of  $\Delta$ ). This exactly reflects the expansion for the volume of the parallel set  $\Delta_\rho$  as a polynomial of order  $d = 2$  given by the Steiner-Formula. The following proof generalizes this idea for arbitrary convex polytopes and shows that Theorem 2.1.1 holds for general  $K \in \mathcal{K}$  by approximation of  $K$  by means of convex polytopes.

**Proof. (Theorem 2.1.1)** We first assume that  $K \in \mathcal{K}$  is a *convex polytope*  $P$  and  $\rho > 0$ . For  $m \in \{0, \dots, d\}$  we denote by  $\mathcal{F}_m(P)$  the set of all  $m$ -dimensional *faces* of  $P$  and  $\mathcal{F}(P) = \cup_{m=0}^d \mathcal{F}_m(P)$  is the collection of all faces of  $P$ . Further  $\forall m \in \{0, \dots, d-1\}$ ,  $F \in \mathcal{F}_m$  and  $x \in \text{relint}F$  let  $N(P, F)$  stand for the *normal cone* of  $P$  at  $F$  which is the closed convex cone of outer normal vectors corresponding to supporting hyperplanes to  $P$  supporting at the point  $x$ . Note that  $N(P, F)$  is independent of the choice of  $x$ . We now define the *external angle* of a face  $F$  of  $P$  as

$$\alpha(F, P) := \frac{\lambda_{d-m}(N(P, F) \cap B^d(0))}{\kappa_{d-m}},$$

where  $\lambda_{d-m}$  is the  $d-m$  dimensional Lebesgue measure. As a convention we use  $\alpha(P, P) = 1$  and  $\alpha(\emptyset, P) = 0$ . Because for every  $x \in \mathbb{R}^d$  the nearest polytope point  $p(P, x)$  lies in the relative interior of a unique face of  $P$  we can decompose the parallel sets  $P_\rho$  as follows into disjoint components:

$$P_\rho = \bigcup_{F \in \mathcal{F}(P)} P_\rho \cap p(P, \cdot)^{-1}(\text{relint}F).$$

For  $m \in \{0, \dots, d-1\}$  and  $F \in \mathcal{F}_m$  every such component can be written as a direct orthogonal sum:

$$P_\rho \cap p(P, \cdot)^{-1}(\text{relint}F) = \text{relint}F \bigoplus (N(P, F) \cap \rho B^d(0)).$$

and an application of Fubini's Theorem leads to:

$$\lambda_d(P_\rho \cap p(P, \cdot)^{-1}(\text{relint}F)) = \lambda_m(F) \cdot \lambda_{d-m}(N(P, F) \cap \rho B^d(0)) = \lambda_m(F) \rho^{d-m} \kappa_{d-m} \alpha(F, P).$$

Defining the intrinsic volumes for convex polytopes as

$$V_m(P) := \sum_{F \in \mathcal{F}_m(P)} \lambda_m(F) \alpha(F, P) \tag{2.1}$$

we can write:

$$V_d(P_\rho) := \sum_{m=0}^d \rho^{d-m} \kappa_{d-m} V_m(P), \quad (2.2)$$

which proves the Steiner-Formula for polytopes.

For a general  $K \in \mathcal{K}$  and the uniqueness part we use equation (2.2) for  $\rho = 1, \dots, d+1$  which produces a system of  $d+1$  linear equations for  $d+1$  variables  $V_m(P), m \in \{0, \dots, d\}$ . The determinant of this linear system is Vandermonde and therefore a solution exists which can be written as

$$V_m(P) = \sum_{\nu=1}^{d+1} a_{m\nu} V_d(P_\nu), \text{ for } m = 0, \dots, d.$$

Because the coefficients  $a_{m\nu}$  are independent of the polytope  $P$  the following definition of the intrinsic volumes for a general convex body  $K \in \mathcal{K}$  makes sense:

$$V_m(K) := \sum_{\nu=1}^{d+1} a_{m\nu} V_d(K_\nu) \quad (2.3)$$

Because the volume mapping  $V_d(\cdot)$  is continuous with respect to the Hausdorff metric  $d_H$ , the functional  $V_m(\cdot)$  is also continuous with respect to  $d_H$  on  $\mathcal{K}$  for  $m \in \{0, \dots, d\}$ . Approximation of  $K$  by convex polytopes proves the theorem.  $\square$

Note that whereas the expression for general  $K$  in equation (2.3) is less useful, for polytopes the intrinsic volumes are expressed quite explicitly by equation (2.1). We can get a hint for a more intuitive interpretation of the intrinsic volumes (and thus the Minkowski functionals) if we consider equation (2.1) for some instructive instances which are  $m = d, d - 1$  and  $m = 0$ . We rediscover that  $V_d(P)$  is the *volume functional* for convex polytopes but furthermore we see that  $2 \cdot V_{d-1}(P)$  corresponds to the *surface area* and  $V_0(P) \equiv 1$  which we call the *Euler characteristic*  $\chi$  of the convex polytope  $P$ . (The Euler characteristic is defined to be 1 for every convex body. For a definition of  $\chi$  on more general set classes see Definition 2.1.11 below.) It is true that these correspondencies between Minkowski functionals and intuitive geometrical quantities still hold for general convex bodies  $K \in \mathcal{K}$ . We will come back to this point later in section 2.2.2.

### 2.1.2 Properties of Minkowski Functionals

From Theorem 2.1.1 and its proof a whole list of properties for intrinsic volumes can be derived. We summarize the most important ones in the following theorem.

**Theorem 2.1.3. (Properties of Intrinsic Volumes)** *The intrinsic volumes  $V_m(\cdot) : \mathcal{K} \rightarrow \mathbb{R}, m \in \{0, \dots, d\}$  have the following properties:*

- i)  $V_m$  is invariant under rigid motions.
- ii)  $V_m$  is Hausdorff-continuous.
- iii)  $V_m$  is homogeneous of degree  $m$ , i.e.  $V_m(c \cdot K) = c^m V_m(K)$  for  $c \geq 0$ .
- iv)  $V_m$  is independent of the dimension of the surrounding space.

- v)  $V_m$  is increasing under set inclusion, i.e.  $K_1 \subset K_2 \Rightarrow V_m(K_1) \leq V_m(K_2)$ .
- vi)  $V_m$  is additive, i.e.  $V_m(K_1 \cup K_2) = V_m(K_1) + V_m(K_2) - V_m(K_1 \cap K_2) \quad \forall K_1, K_2 \in \mathcal{K}$  with  $K_1 \cup K_2 \in \mathcal{K}$ .
- vii)  $V_m$  is non-negative, i.e.  $V_m(K) \geq 0, K \in \mathcal{K}$ .

**Proof. (Theorem 2.1.3)**

- i),ii). Parts i) and ii) follow directly from the corresponding properties for  $V_d(\cdot)$  and equation (2.3).
- iii). First note that  $(\alpha \cdot K)_\nu = \alpha \cdot K_{\frac{\nu}{\alpha}}$ . Applying equation (2.3) and the Steiner-Formula we have:

$$V_m(\alpha \cdot K) = \alpha^d \sum_{\nu=1}^{d+1} a_{m\nu} V_d(K_{\frac{\nu}{\alpha}}) = \alpha^d \sum_{\nu=1}^{d+1} \sum_{s=0}^d a_{m\nu} \left(\frac{\nu}{\alpha}\right)^{d-s} \kappa_{d-s} V_s(K) = \alpha^m V_m(K),$$

where d-homogeneity of the volume  $V_d(\cdot)$  is assumed to be obvious.

- iv). This follows essentially from (2.1). Let  $P$  a convex polytope  $P$  in a  $k$ -dimensional affine subspace  $E_k$  of  $\mathbb{R}^d$  and  $F \in \mathcal{F}_m$  an  $m$ -face of  $P$ . ( $m \leq k$ ). Then we have that the  $k-m$ -dimensional normal cone  $N^{(k)}(F, P)$  corresponds exactly to the projection of the  $d-m$ -dimensional normal cone  $N^{(d)}(F, P)$  onto the  $k$ -dimensional linear subspace  $V_k$  associated with  $E_k$  and therefore the external angle is independent of the dimension of the surrounding space:

$$\alpha^{(d)}(F, P) = \frac{\lambda_{d-m}(N^{(d)}(F, P) \cap B^d(0))}{\kappa_{d-m}} = \frac{\lambda_{k-m}(N^{(k)}(F, P) \cap B^k(0))}{\kappa_{k-m}} = \alpha^{(k)}(F, P).$$

Assertion iv) trivially holds for Lebesgue-measure  $\lambda_m(\cdot)$  and thus formula (2.1) establishes assertion iv) for polytopes. Then the usual approximation argument is used for general  $K \in \mathcal{K}$ .

- v),vii). For these points we need the theory of section 2.2. The statements are most easily seen from a so-called *Crofton-formula* (Theorem 2.2.2). However, v) for polytopes also follows from equation 2.1 because external angles are non-negative. For general  $K$  the result follows from polytope-approximation.
- vi). Let  $K_1, K_2 \in \mathcal{K}$ . First one proves the following equality for indicator-functions:

$$\mathbf{1}_{((K_1 \cup K_2)_\rho)} + \mathbf{1}_{((K_1 \cap K_2)_\rho)} = \mathbf{1}_{(K_1)_\rho} + \mathbf{1}_{(K_2)_\rho}.$$

The only case where the statement is non-trivial is for those  $x \in \mathbb{R}^d$  where  $\mathbf{1}_{((K_1 \cup K_2)_\rho)} = 1$  and  $\mathbf{1}_{((K_1 \cap K_2)_\rho)} = 0$ . Then without loss of generality there exists a  $y \in K_1 \cap K_2^c$  and  $z \in \rho B^d(0)$  such that  $x = y + z$ . Assume now that  $\mathbf{1}_{(K_2)_\rho} = 1$ . Then we find similarly  $\tilde{y} \in K_2 \cap K_1^c$  and  $\tilde{z} \in \rho B^d(0)$  such that  $x = \tilde{y} + \tilde{z}$  and  $y \neq \tilde{y}$ . But because  $K_1 \cup K_2$  is convex, the whole line  $[y, \tilde{y}] \subset K_1 \cup K_2$ . Because  $K_1, K_2$  are closed there must exist  $\hat{y} \in [y, \tilde{y}]$  with  $\hat{y} \in K_1 \cap K_2$  and  $\hat{z} := |x - \hat{y}| < \rho$  which is a contradiction. Thus we must have  $\mathbf{1}_{(K_2)_\rho} = 0$ .

Integration with respect to Lebesgue measure shows that the above equation also holds for the volume functional  $V_d(\cdot)$ . Finally applying the Steiner-Formula and comparing coefficients of  $\rho^m$  establishes the claim.  $\square$

It is needless to say that these properties (except of course iii) and iv)) also hold for the Minkowski-functionals  $W_m(\cdot)$ . These properties are all most welcome for a functional used to describe information stored in an image of a porous structure, we would even expect these properties from a reasonably natural and intuitive summary statistic. But unfortunately porous structures can be very complicated and in general certainly do not belong to  $\mathcal{K}$ . If we want to use the Minkowski functionals for our purposes, their definition needs to be extended to a broader range of sets. In fact we will see that an extension to sufficiently general sets to describe complex porous structures is possible in such a way that the above properties are conserved (see 2.1.4 for details).

### 2.1.3 The Hadwiger Characterization Theorem

The second important question we can ask in this context is how large the class of functionals on  $\mathcal{K}$  which share the properties of Theorem 2.1.3 might be and what these functionals look like. The surprising answer is that properties *i*) (motion-invariance), *ii*) (continuity) and *vi*) (additivity) already suffice to characterize such functionals as linear combinations of Minkowski-functionals. If one considers the properties of Theorem 2.1.3 as indispensable for a summary statistic of an image of a porous medium, then there is no need to go beyond the  $d+1$  Minkowski-functionals. Therefore the following famous theorem due to Hadwiger is the main justification for describing images of porous media using Minkowski functionals.

**Theorem 2.1.4. (*Hadwiger's Characterization Theorem for Intrinsic Volumes*)**

Let  $\phi : \mathcal{K} \rightarrow \mathbb{R}$  an additive, motion-invariant and Hausdorff-continuous functional, then there exist constants  $c_0, \dots, c_d$  such that:

$$\phi(K) = \sum_{m=0}^d c_m \cdot V_m(K) \quad \forall K \in \mathcal{K}.$$

The proof of Theorem 2.1.4 is simple as soon as one has the following quite involved lemma (which is proved in (Klain 1995)) at hand:

**Lemma 2.1.5.** Let  $\phi : \mathcal{K} \rightarrow \mathbb{R}$  an additive, Hausdorff-continuous and motion-invariant functional that vanishes on lower dimensional sets and on unit cubes. Then  $\phi \equiv 0$ .

The proof of this Lemma is done inductively with respect to dimension  $d$ . It is so interesting that we provide here at least its basic skeleton. The case  $d = 1$  is simple because convex bodies are just closed segments. For  $d > 1$  the induction hypothesis is used to prove that  $\phi$  vanishes on right (convex) cylinders, next cutting an arbitrary cylinder in two halves in such a way that they can be glued together again to form a right cylinder shows that  $\phi$  vanishes on arbitrary cylinders. Next one considers sums of polytopes  $P$  and segments  $S$ . A decomposition of  $P + S$  into  $P$  and cylinders can always be found which reveals  $\phi(P + S) = \phi(P)$ . Then an induction argument is used to prove this equality for sums of polytopes and *finitely many* segments, i.e.  $\phi(P) = \phi(P + S_1 + \dots + S_m)$ . Sums of finitely many segments are called *zonotopes*. The same equality therefore holds for Hausdorff-limits of sequences of zonotopes. Such sets are called *zonoids* and again by continuity  $P$  can be replaced by an arbitrary convex body  $K$  to find  $\phi(K + Z) = \phi(K)$  for  $Z$  a zonoid. Next one uses that for a (sufficiently smooth) centrally symmetric convex body  $K$  one always finds two zonoids  $Z_1, Z_2$  such that  $K + Z_1 = Z_2$ , which establishes  $\phi(K) = 0$  for (sufficiently smooth) centrally symmetric convex bodies. This last result is used in a smart construction to show that  $\phi(\Delta) = 0$  for arbitrary simplices  $\Delta$ . Approximating an arbitrary convex body  $K$  by polytopes  $P$  and these polytopes again by simplices finally establishes the lemma.

**Proof. (Theorem 2.1.4)** The assertion is proved inductively in dimension  $d$ . For  $d = 0$  there's nothing to prove. For  $d > 0$  let  $H$  be a hyperplane and consider the restriction  $\phi|_H$ . This functional is of course additive, continuous and invariant under motions of  $H$  into itself. Therefore  $\phi(K) = \sum_{m=0}^{d-1} c_m \cdot V_m(K)$ ,  $\forall K \in \mathcal{K}$ ,  $\dim(K) < d$  for some constants  $c_0, \dots, c_n$  independent of the hyperplane  $H$ . Consider now the new functional  $\tilde{\phi}(K) := \phi(K) - \sum_{m=0}^d c_m \cdot V_m(K)$  where  $c_n$  is chosen such that  $\tilde{\phi}$  vanishes on an arbitrary fixed unit cube. Then  $\tilde{\phi}$  fulfills the assumptions of Lemma 2.1.5 and therefore vanishes which proves the Theorem.  $\square$

## 2.1.4 Extended Definition of Minkowski Functionals

In this subsection the definition of Minkowski Functionals is extended to *the convex ring*  $\mathcal{R}$ , a set-class rich enough for mathematical modelling of porous structures and an even farther reaching extension of the definition of Minkowski functionals is mentioned. Proofs for the theorems used here can be found in the appendix.

**Definition 2.1.6. (convex ring)** *The set-system consisting of all finite unions of convex sets is called the **convex ring**  $\mathcal{R}$ . A set  $R \in \mathcal{R}$  has a representation of the following form:*

$$R = \bigcup_{i=1}^n K_i \text{ for } n \in \mathbb{N}, K_i \in \mathcal{K}.$$

Sometimes the elements  $R \in \mathcal{R}$  are called **polyconvex sets**.

Note that because of the additivity property (Theorem 2.1.3, part vi) for intrinsic volumes  $V_m(\cdot)$ , there is no freedom in choosing their functional values on  $\mathcal{R}$ . In fact they are dictated by the *inclusion-exclusion-formula*:

**Theorem 2.1.7. (inclusion-exclusion-formula)** *Every additive functional  $\phi : \mathcal{K} \rightarrow \mathbb{R}$  (or more generally with values in an abelian group) satisfies*

$$\phi(K_1 \cup \dots \cup K_n) = \sum_{m=1}^n (-1)^{m-1} \sum_{i_1 < \dots < i_m} \phi(K_{i_1} \cap \dots \cap K_{i_m}).$$

**Proof. (Theorem 2.1.7)** additivity of  $\phi$  and induction in  $n$ .

**Remark 2.1.8.** It would be tempting to use Theorem 2.1.7 to extend the definition of  $V_m(\cdot)$  to the convex ring  $\mathcal{R}$ . Unfortunately this only specifies the values of  $V_m(\cdot)$  for a particular representation of a polyconvex set  $R$ . To make this extension sensible one must make sure that the extension leads to the same value  $V_m(R)$  for *all* representations of  $R$  as a finite union of convex sets.

**Remark 2.1.9.** Because the convex ring  $\mathcal{R}$  is closed under finite unions and finite intersections the assertion of Theorem 2.1.7 makes sense also for functionals  $\phi : \mathcal{R} \rightarrow \mathbb{R}$ . In fact, the same induction argument used in the proof above also works to extend the theorem to such functionals (and therefore holds for  $V_m(\cdot)$  on  $\mathcal{R}$  once this extension is well-defined.)

The following theorem guarantees the well-definedness of the unique domain-extension for intrinsic volumes onto  $\mathcal{R}$  according to Theorem 2.1.7:

**Theorem 2.1.10. (Well-definedness of intrinsic volume extension)** *Let  $\phi$  a Hausdorff-continuous, additive mapping from  $\mathcal{K} \rightarrow X$  where  $X$  is a topological space, then  $\phi$  has a unique additive extension to the convex ring  $\mathcal{R}$ . The values  $\phi(R)$  for  $R \in \mathcal{R} \setminus \mathcal{K}$  are given by Theorem 2.1.7.*

For a proof of this theorem see the appendix, section 7.2. We summarize the results from this section in the following definition:

**Definition 2.1.11.** (*intrinsic volumes and Minkowski functionals on the convex ring*)

- Let  $R = \bigcup_{i=1}^n K_i$  a representation of a  $d$ -dimensional polyconvex set  $R \in \mathcal{R}$ . Then the *intrinsic volumes* for  $R$  are defined as:

$$V_m(R) := \sum_{k=1}^n (-1)^{k-1} \sum_{i_1 < \dots < i_k} V_m(K_{i_1} \cap \dots \cap K_{i_k}), \quad m \in 0, \dots, d,$$

where the  $V_m(K_{i_1} \cap \dots \cap K_{i_m})$  are defined according to the Steiner-formula (Definition 2.1.2).

- The functionals  $W_m : \mathcal{R} \rightarrow \mathbb{R}, m = 0, \dots, d$  which are scaled versions of the intrinsic volumes  $V_m(R)$  such that  $\kappa_m \cdot V_{d-m}(R) = \binom{d}{m} \cdot W_m(R)$  are called the **Minkowski-functionals** of a polyconvex set  $R \in \mathcal{R}$ .

The properties of Minkowski-functionals on  $\mathcal{K}$  stated in Theorem 2.1.3 carry over to  $\mathcal{R}$  (which is obvious for all of them except ii), v) and vii) from Theorem 2.1.7). Monotonicity and Hausdorff-continuity are lost while extending the functionals to the convex ring. This can be seen with the following instructive example: Consider the unit circle  $B^d(0) \in \mathcal{K}$ . Of course we have  $V_0(B^d(0)) = 1$ . Now consider a sequence of sets  $K_h := B^d(0) \cap h\mathbb{Z} \subset B^d(0)$ . Then  $K_h \in \mathcal{R}, \forall h, V_0(K_h) > 1$  for  $h$  small enough and  $V_0(K_h)$  strictly increasing as  $h \downarrow 0$ . Furthermore  $\lim_{h \rightarrow 0} K_h = B^d(0)$ . That negative values for  $V_m(\cdot)$  can show up on  $\mathcal{R}$  is clear from Definition 2.1.11. Non-negativity is therefore not preserved.

Just for completeness we would like to mention here that Minkowski functionals are special cases of so called *generalized curvature measures* which arise as coefficients in local versions of Steiner-type formulae expressing the volume of local parallel sets. It is possible to go beyond the convex ring  $\mathcal{R}$  and define curvature measures on more general (but also more abstract) classes of sets. The only such class we mention here is the class of *sets of positive reach*  $\mathcal{P}$ . This class contains nonempty, closed sets  $P$  with  $\text{reach}(P) > 0$ . The *reach* of  $P \in \mathcal{P}$  is defined as the smallest number  $r$ , such that  $\forall x \in \mathbb{R}^d$  with  $d(x, P) < r$  there exists a unique nearest point in  $P$ . For  $K \in \mathcal{K}$  we have  $\text{reach}(K) = \infty$ . An example for a set  $P$  with  $\text{reach}(P) = 1$  is the unit circle. The set that consists of two intersecting segments is not of positive reach (because in every neighborhood of the intersection point there exists a point on the bisecting line which has two closest points).

For details about generalized curvature measures consult (Schneider and Weil 1992) or (Schneider 2004) which contains an excellent bibliography for many introductory topics in integral geometry.

### 2.1.5 The Euler Characteristic

We have seen from the proof of the Steiner formula in section 2.1.1 that the functional we call *Euler characteristic*  $V_0(\cdot)$  (usually denoted by  $\chi$ ) is equal to 1 for every convex polytope  $P$ . By Hausdorff-continuity we also have  $\chi(K) = 1 \forall K \in \mathcal{K}$ . Being trivial on  $\mathcal{K}$  this functional becomes much more interesting for general polyconvex sets on  $\mathcal{R}$  where it is

no longer constant. An alternative definition of the Euler-characteristic due to Hadwiger can be given recursively for arbitrary dimension  $d$ . This recursive definition will allow us to derive the Ohser-Mücklich-estimators for  $\chi$  described in sections 2.3.1 and 2.3.2. For this purpose let  $\omega \in S^{d-1}$  denote an arbitrary direction and  $E_{y,\omega}$  a hyperplane with unit normal vector  $\omega$  and distance vector  $y$  in the orthogonal complement of  $E_{0,\omega}$ . Then the *recursive definition of the Euler-characteristic* can be stated as follows:

**Definition 2.1.12. (Hadwiger recursive definition of the Euler-characteristic)**  
Let  $R \in \mathcal{R}$ ,  $\omega \in S^{d-1}$ .

$$\chi(R) = \sum_{y \in E_{0,\omega}^c} \lim_{\Delta \rightarrow 0} [\chi(R \cap E_{y,\omega}) - \chi(R \cap E_{y+\Delta,\omega})]. \quad (2.4)$$

Please note that for  $R \in \mathcal{R}$  this sum is well defined because only finitely many summands are different from zero. Furthermore this definition is independent of the choice of the “screening direction”  $w$  also for non-isotropic sets.

For convex sets  $K \in \mathcal{K}$  and arbitrary  $\omega \in S^{d-1}$  we can always find a pair of parallel supporting hyperplanes to enclose the set and exactly one of those gives a unit contribution to equation 2.4. Hence the Hadwiger definition for convex sets coincides with our former definition  $\chi(K) = 1$ ,  $\forall K \in \mathcal{K}$ . Also note that equation 2.4 is additive and therefore the recursively defined  $\chi$  coincides with  $V_0$  also on  $\mathcal{R}$ . By means of the Hadwiger-Definition the Euler-characteristic can also be defined for the (closed) complement  $\bar{R}^c$  of a set  $R \in \mathcal{R}$ . Although the set  $\bar{R}^c$  is unbounded, there are still only finitely many points which contribute to the sum in equation 2.4. For example this leads immediately to  $\chi(\mathbb{R}^d) = 0$ ,  $d > 0$ . Furthermore in  $d = 2$  for a set  $K \in \mathcal{K}$  with inner points, application of equation 2.4 gives  $\chi(\bar{K}^c) = -1 = -\chi(K)$ . In fact, screening the set  $K^c$  from the opposite direction than when screening  $K$  gives a contribution for the same hyperplane  $E_{y,\omega}$  in both cases but with opposite sign. In  $d = 3$  screening  $K^c$  from the opposite direction again gives a contribution for the same hyperplane but this time with the *same* sign. This argument can be used inductively and proves the following important *consistency relation* for the Euler characteristic  $\chi$  which can immediately be carried over to the convex ring  $\mathcal{R}$  by additivity as long as one ensures that the set  $R \in \mathcal{R}$  has a representation  $R = \cup_{i=1}^n K_i$  for which none of the  $K_i$  have lower dimensional intersections (and thus could possibly enter the calculation for the Euler-characteristic of the complement with the wrong sign).

**Theorem 2.1.13. (Consistency of the Euler-characteristic)** Assume a set  $R \in \mathcal{R}$  in  $d$  dimensions has inner points and a representation  $R = \cup_{i=1}^n K_i$ ,  $K_i \in \mathcal{K}$  such that intersections  $K_{i_1} \cap \dots \cap K_{i_\nu}$ ,  $\nu = 2, \dots, n$  are either empty or have inner points, then:

$$\chi(\bar{R}^c) = (-1)^{d-1} \cdot \chi(R),$$

where  $\chi(\bar{R}^c)$  is defined by equation 2.4.

**Proof. (Theorem 2.1.13)** For more details see (Ohser, Nagel and Schladitz 2002), Theorem 1.

The definition of the Euler-characteristic may be extended beyond the Hadwiger-definition. For the following generalization of the concept we use (Adler 1981). For the rest of this section we denote by  $E_k$  a *regular k-plane* which is a  $k$ -dimensional affine subspace of  $\mathbb{R}^d$  which is generated by  $k$  vectors parallel to  $k$  of the unit-vectors  $\{e_1, \dots, e_d\}$ . Furthermore we call a compact set  $B \in \mathbb{R}^d$  a *basic set*, whenever its intersections with regular  $k$ -planes



$B \cap E_k$  are simply connected. The class of all basic sets is denoted by  $\mathcal{B}$ . Note for example that  $\mathcal{K} \subset \mathcal{B}$ . Next we define the class of *basic complexes*  $\mathcal{CB}$  to represent finite unions of basics such that any intersections of these basics remain in  $\mathcal{B}$ :

$$A \in \mathcal{CB} \Leftrightarrow A = \bigcup_{m=1}^n B_m \text{ and } B_{\nu_1} \cap \dots \cap B_{\nu_k} \in \mathcal{B} \text{ for } k = 1, \dots, n.$$

The sets  $p(A) = \{B_1, \dots, B_n\}$  form a *partition of A* and we call  $n$  the *order* of the partition  $p(A)$ . To give examples for sets contained in  $\mathcal{CB}$  note that  $\mathcal{R} \subset \mathcal{CB}$  and  $E_k \cap A \in \mathcal{CB}, \forall A \in \mathcal{CB}$ . We can now define:

**Definition 2.1.14.** (*Euler-Poincaré-characteristic for basic complexes*) The *Euler-Poincaré-Characteristic*  $\chi$  for a basic complex  $A \in \mathcal{CB}$  is

$$\begin{cases} \chi(A) = \mathbf{1}_{A \neq \emptyset} & \text{if } A \in \mathcal{B}, \\ \chi(A) = \sum_{m=1}^n (-1)^{m-1} \sum_{\nu_1 < \dots < \nu_m} \chi(B_{\nu_1} \cap \dots \cap B_{\nu_m}) & \text{otherwise} \end{cases}$$

where  $\{B_1, \dots, B_n\}$  is an arbitrary partition of  $A$ .

Please note that this definition is in accordance with the definition from convex geometry above on  $\mathcal{R}$ . The definition is sensible because of the following Theorem:

**Theorem 2.1.15.** (*unique extension of Euler-Poincaré-characteristic to CB*) Let  $A \in \mathcal{CB}$  and  $p(A)$  a partition of  $A$ . Then the quantity  $\chi(A, p) = \sum_{m=1}^n (-1)^{m-1} \sum_{\nu_1 < \dots < \nu_m} \chi(B_{\nu_1} \cap \dots \cap B_{\nu_m})$  is independent of the partition  $p$  and additive on  $\mathcal{CB}$ .

We give the proof of this theorem because it provides a constructive extension of the Hadwiger recursive definition 2.1.12 to basic complexes which is a set-class containing most subsets of  $\mathbb{R}^d$  which are of interest in differential topology.

**Proof. (Theorem 2.1.15)** The proof proceeds by induction in  $d$ . If  $d = 1$  basics are closed intervals, isolated points and the empty set. Then  $\chi(A, p)$  is the number of disjoint intervals and isolated points in  $A$  and therefore independent of  $p$ . Additivity is obvious. Set  $\phi^{(1)}(A) := \chi(A, p)$ . For  $d > 1$  choose one of the unit vectors  $e_j$  and  $\forall x$  denote by  $E^x$  the hyperplane  $x_j = x$ . Let  $A \in \mathcal{CB}^d$  and  $p(A) = \{B_1, \dots, B_n\}$  an arbitrary partition. Then the cross-sections  $E^x \cap A$  (projected on  $E_0$ ) are in  $\mathcal{CB}^{d-1}$  and the induction hypothesis guarantees that there exists a partition-independent functional  $\phi_x$  defined on  $\{A \cap E_x, A \in \mathcal{CB}^d\}$  and we can define the partition-independent function  $f(A, x) := \phi_x(A \cap E^x)$ . By the induction hypothesis we have:

$$f(A, x) = \sum_{m=1}^n (-1)^{m-1} \sum_{\nu_1 < \dots < \nu_k} \phi^{(d-1)}(B_{\nu_1} \cap \dots \cap B_{\nu_k} \cap E^x).$$

Note that  $\phi^{(d-1)}(B_{\nu_1} \cap \dots \cap B_{\nu_k} \cap E^x)$  is a 0-1-function and because  $B_{\nu_1} \cap \dots \cap B_{\nu_k}$  is a basic (and therefore connected) there exist  $a, b \in \mathbb{R}$  such that  $\phi^{(d-1)}(B_{\nu_1} \cap \dots \cap B_{\nu_k} \cap E^x) = \mathbf{1}_{\{a \leq x \leq b\}}$ . Hence  $f(A, x)$  is a stepfunction with only a finite number of discontinuities. Therefore the right-hand limit  $f(A, x^+) := \lim_{y \downarrow x} f(A, y)$  always exists. Now we define:

$$\phi^{(d)}(A) := \sum_{x \in \mathbb{R}} f(A, x) - f(A, x^+). \quad (2.5)$$

Note that there are only finitely many summands different from 0 on the right hand side and therefore the sum is well-defined. To show that  $\phi^{(d)} = \chi^{(d)}$  we note similarly as above

that for a basic  $B \in \mathcal{B} \setminus \emptyset$  there exist  $a, b \in \mathbb{R}$  such that  $f(B, x) = \mathbf{1}_{\{a \leq x \leq b\}}$ . The only contribution to the sum on the right hand side of equation 2.5 is therefore in  $x = b$  and equals 1. Equation 2.5 trivially holds for  $B = \emptyset$ . Now let  $A, B, A \cup B, A \cap B$  all be basic complexes in dimension  $d$ . Then the projections onto  $E^0$  of the intersections of these sets with  $E^x$  are basic complexes in dimension  $d - 1$  and therefore:

$$f(A \cup B, x) = f(A, x) + f(B, x) - f(A \cap B, x), \forall x \in \mathbb{R}.$$

The same relationship holds for  $x$  replaced by  $x^+$  and therefore also after summing over  $x$ :

$$\phi^{(d)}(A \cup B, x) = \phi^{(d)}(A, x) + \phi^{(d)}(B, x) - \phi^{(d)}(A \cap B, x), \forall x \in \mathbb{R}.$$

This proves both the additivity and the equality of  $\phi^{(d)}$  with  $\chi$  on  $\mathcal{CB}$ . Because the right hand side of equation 2.5 is partition-independent the theorem follows.  $\square$

Note that equation 2.5 is exactly the recursive Hadwiger definition of the Euler-characteristic which extends definition 2.1.12 to  $\mathcal{CB}$ . But note that on  $\mathcal{CB}$  the possible screening directions in 2.5 are limited to the directions of the coordinate axes because non-convex basics in general don't need to stay basics when rotated. 2.5 can also be seen as an alternative to definition 2.1.14. In the following we will not distinguish anymore between the Euler-characteristic from convex geometry and the Euler-Poincaré-characteristic from topology and call the functional simply Euler-characteristic.

We are now in a position to calculate the Euler-characteristic for very general sets including surfaces and submanifolds of  $\mathbb{R}^d$ . To give a few simple examples, note that  $\chi(B^d) = 1, \forall d$  (by convexity) but  $\chi(S^d) = 1 + (-1)^{d-1}$ . For  $d = 2$  partition the unit circle into four arcs delimited by the 4 intersection points with the cartesian coordinate axes. Then each of these arcs is a basic. (Partition in only two arc would not be feasible because the intersection of these arcs would *not* be a basic!) It follows by additivity that  $\chi(S^1) = 4 - 4 = 0$ . A similar construction in  $d = 3$  works by partitioning the sphere for example into 8 basics delimited by the equator and two meridional great circles intersecting each other at a right angle in north and south pole (all circles parallel to the coordinate planes). Each of the 8 faces of this partition contributes one unit to  $\chi$  whereas each of the 12 edges contributes  $-1$ . Finally the contributions of each of the 6 vertices are  $-2 + 4 - 1$ . Therefore  $\chi(S^2) = 6 - 12 + 8 = 2$ . Of course this can be more easily seen using the Hadwiger recursive definition (where one gets a contribution of  $+1$  at north and south pole). For a torus  $T$  in  $\mathbb{R}^3$  the Hadwiger-recursive definition 2.5 gives  $\chi(T) = 0$  and if we consider a three-dimensional ball with  $l$  disjoint handles attached its Euler-characteristic is  $1 - l$ . This can be seen by decomposing the set into the ball and  $l$  tori and noting that the intersections of ball and tori are topologically equivalent to a right cylinder which is convex. An interesting remark is that this last result holds independent of dimension.

That the Euler-characteristic is a topological-invariant quantity becomes clear from the following reasoning due to (Hadwiger 1957). Let  $A_1, A_2$  two sets in  $\mathcal{CB}$  and  $B_1^1, \dots, B_n^1$  a partition of  $A_1$  into basics. If we now are able to find a partition  $B_1^2, \dots, B_n^2$  for  $A_2$  of the same order  $n$  and if we can identify each basic  $B_{\nu_1}^1$  with a basic  $B_{\nu_2}^2$  such that

$$B_{\nu_1}^1 \cap \dots \cap B_{\nu_k}^1 \neq \emptyset \Leftrightarrow B_{\nu_1}^2 \cap \dots \cap B_{\nu_k}^2 \neq \emptyset, \quad (2.6)$$

then definition 2.1.14 ensures that  $\chi(A_1) = \chi(A_2)$ . So far we have developed the concept of the Euler-characteristic in convex geometry and generalized it to basic complexes. But there are also truly topological definitions of the Euler characteristic. In geometrical topology the so-called *Euler-Poincaré-Characteristic* is known to be a topologically invariant quantity of a closed surface. It is true that the topological notion of Euler-characteristic and the one from convex geometry we have used so far in this text coincide for sets where both are defined.

The topological definition of the Euler-characteristic starts from the the famous *Euler-polyhedral-formula* that provides a relationship between the number of vertices  $V$ , the number of edges  $E$  and the number of faces  $F$  of a polyhedron  $P$  in space which is homeomorphic to a sphere. The *Euler-polyhedral-formula* states:

$$V - E + F = 2. \quad (2.7)$$

Proofs for this formula can be found widespread in literature. (I have have found at least 19 different ones by research in the world wide web...). The simplest one notes that an incidence-structure of a polyhedron homeomorphic to a sphere is always equivalent to a finite, connected planar graph and uses a trivial induction in the number of graph vertices, edges or faces to show the formula for planar graphs (which are graphs that can be drawn on a sphere without intersections).

For a general polyhedron  $P$  in  $d$ -dimensions the topological *Euler-Poincaré-characteristic* is defined as

$$\chi(P) = \sum_{m=0}^d (-1)^m |\mathcal{F}^m(P)|. \quad (2.8)$$

where again  $\mathcal{F}_m$  denotes the set of  $m$ -faces of the polyhedron.

This topological definition of the Euler characteristic can be generalized in several directions. After triangulation, a polyhedron can be understood as a realization of the abstract concept of a *simplicial complex*.

**Definition 2.1.16. (Simplicial Complex)** A (finite) abstract *simplicial complex* is a finite collection  $\Sigma$  of nonempty subsets of a finite set with the property that if  $\sigma \in \Sigma$  and  $\tau \subset \sigma$  is a non-empty subset of  $\sigma$ , then also  $\tau \in \Sigma$ . Elements  $\sigma$  of cardinality  $n + 1$  are called  *$n$ -simplices* and the *dimension*  $d$  of  $\Sigma$  is the largest  $n$  for which there are  $n$ -simplices contained in  $\Sigma$ .

For finite simplicial complexes the Euler-Poincaré-characteristic is defined according to equation 2.8, where now  $\mathcal{F}^m$  denotes the number of  $m$ -simplices within  $\Sigma$ . A further generalization defines the EC-characteristic in the same way for so-called *CW-complexes*. The *Betti-numbers*  $b_k$  are the ranks of the  $k$ th homology group of the CW-complex under consideration. The Euler-characteristic is then defined as the alternating sum  $\sum_k (-1)^k b_k$ . That this last definition comprises the one for simplicial complexes follows from the *Hopf-Trace-Theorem*. The homology-definition of the Euler characteristic has the advantage that it allows to prove quite easily that the Euler-characteristic is a topologically invariant quantity. Please refer to a standard differential-topology book for further details.

Since the Euler-Poincaré-characteristic is topologically invariant, the topological definition of Euler-characteristic for simplicial complexes comprises the convex-geometry-definition in three-dimensions: Any triangulation of the surface of a  $K \in \mathcal{K}$  can be interpreted as a simplicial complex. And by topological invariance a combination of formulae 2.7 and 2.8 leads to  $\chi(K) = 1, \forall K \in \mathcal{K}$ .

In three dimensions, for compact, closed two-dimensional manifolds, the Euler-characteristic is given by the famous *Gauss-Bonnet-Theorem*. In arbitrary dimensions for compact subsets  $S \subset \mathbb{R}^d$  whose boundary is a  $d-1$ -dimensional  $\mathcal{C}^2$ -manifold the Euler-characteristic  $\chi(S)$  can be calculated using the *Morse-Cairns-Theorem*.

**Theorem 2.1.17. (*Gauss-Bonnet-Theorem*)** *Let  $Z$  be a compact, boundaryless two-dimensional Riemannian manifold, the Euler-characteristic  $\chi(Z)$  is:*

$$\chi(Z) = \frac{1}{2\pi} K(Z),$$

where  $K(Z)$  is the integral of total curvature (see definition 2.2.5 for a definition of  $K(Z)$ ).

**Proof. (Theorem 2.1.17)** The proof for  $Z = \partial K$  where  $K$  convex is given in the derivation of equation 2.19 below. For the general case consider a basic textbook in differential geometry.

**Theorem 2.1.18. (*Morse-Cairns-Theorem*)** *Let  $Z$  a compact subset of  $\mathbb{R}^d$  bounded by a regular  $d-1$ -dimensional  $\mathcal{C}^2$ -manifold. Let further  $f(t) : \mathbb{R}^d \rightarrow \mathbb{R}$  a  $\mathcal{C}^2$ -function which has no critical points on  $\partial Z$  and such that the critical points of the restrictions  $f|Z$  and  $f|\partial Z$  are all non-degenerate. Then:*

$$\chi(Z) = \sum_{k=0}^d (-1)^k m_k + \sum_{k=0}^{d-1} (-1)^k m'_k,$$

where:

$m_k :=$  the number of critical points of  $f|Z$  for which the Hessian  $\frac{\partial f}{\partial x_i \partial x_j}$  has  $k$  negative Eigenvalues ( $k = 0, \dots, d$ ).

$m'_k :=$  the number of critical points of  $f|\partial Z^-$  for which the Hessian  $\frac{\partial f}{\partial x_i \partial x_j}$  has  $k$  negative Eigenvalues ( $k = 0, \dots, d-1$ ), and where  $\partial Z^-$  is the submanifold of  $\partial Z$  of points  $x$  in which the derivative of  $f$  in direction of the outer normal vector of  $\partial Z$  in  $x$  is negative.

**Proof. (Theorem 2.1.18)** The proof requires extensive knowledge of algebraic homotopy theory and can be found in (Morse and Cairns 1969), Theorem 10.2'.

To establish the Ohser-Mücklich-estimators in sections 2.3.1, 2.3.2 we need to know how to calculate the Euler-characteristic of a porous structure from discrete approximations. The Hadwiger-Definition will not help us much if we note that in general exchange of limit and sum in equation 2.4 is not permitted and may lead to wrong results. For examples refer to (Ohser and Mücklich 2000). But this means that definition 2.1.12 is useless for implementation on a computer which can only deal with finite resolutions. However in (Ohser and Nagel 1996) conditions can be found for which this exchange is possible and

therefore a recursive procedure for calculating the Euler-characteristic can be derived which is suitable for implementation on a computer and leads directly to estimators for  $\chi$  from pixel images (see sections 2.3.1 and 2.3.2). To describe these conditions we need to have a closer look at the basic morphological operations *opening* and *closing*.

**Definition 2.1.19. (Morphological operations)** Let  $A$  and  $B$  two arbitrary sets and  $\tilde{B}$  the image of  $B$  under reflection at the origin.

- The set  $A \oplus B$ , **A dilated by B**, is the set

$$A \oplus B = \{x + y, x \in A, y \in B\}.$$

- The set  $A \ominus B$ , **A eroded by B**, is the set

$$A \ominus B = \bigcap_{y \in B} (A + y)$$

- A **morphological opening of A by B** is performed as an erosion of  $A$  by  $\tilde{B}$  followed by a dilation by  $B$ :

$$A \circ B = (A \ominus \tilde{B}) \oplus B.$$

- A **morphological closing of A by B** is performed as a dilation of  $A$  by  $\tilde{B}$  followed by an erosion by  $B$ :

$$A \bullet B = (A \oplus \tilde{B}) \ominus B.$$

- A set  $A$  is called **morphologically open** with respect to  $B$  if  $A = A \circ B$ .

- A set  $A$  is called **morphologically closed** with respect to  $B$  if  $A = A \bullet B$ .

The set  $B$  is usually called **the structuring element**. Opening is used in morphology to separate two sets which are connected by tiny capillary bridges whereas closing is used to remove and fill up tiny holes in a set.

The exchange of limit and summation in equation 2.4 is possible for sets which are both morphologically open and closed which is shown in (Ohser and Nagel 1996). Once this is shown, it is easily possible to provide a version of Hadwiger recursive definition for  $\chi$  which is suitable for computer implementation. To state the corresponding theorem we need to extend the notation used in definition 2.1.12 slightly. Let  $R \in \mathcal{R}$ . Denote by  $-(y, \omega)$  the vector of the same length as  $y$  pointing in the direction of  $-\omega$  and  $Y_{y, \omega} := (E_{y, \omega} \cap R)_{-(y, \omega)}$  the set  $E_{y, \omega} \cap R$  shifted back to the hyperplane parallel to  $E_{y, \omega}$  through the origin, then we have:

**Theorem 2.1.20. (recursive calculation of the Euler-characteristic)** Let  $\omega \in S^{d-1}$ ,  $\Delta > 0$ . Consider  $R \in \mathcal{R}$  that is morphologically open and closed with respect to the structuring element  $B := -(\Delta, \omega)$ , ie.  $R = R \circ B = R \bullet B$ , then the following integral geometric representations of the Euler-characteristic  $\chi(R)$  hold:

$$\chi(R) = \frac{1}{\Delta} \int [\chi(Y_{y, \omega} \cup Y_{y+\Delta, \omega}) - \chi(Y_{y, \omega})] dy. \quad (2.9)$$

$$\chi(R) = \frac{1}{\Delta} \int [\chi(Y_{y, \omega}) - \chi(Y_{y, \omega} \cap Y_{y+\Delta, \omega})] dy. \quad (2.10)$$

**Proof. (Theorem 2.1.20)** The proof is due to (Ohser and Nagel 1996). The derivation of equation 2.9 can be found in the appendix. Equation 2.10 then immediately follows from 2.9 by exploiting the fact that in the second integrand  $y$  may be replaced by  $y + \Delta$  and then applying additivity.

For  $w \in S^{d-1}$  and  $\Delta > 0$  let now  $E_k := E_{k\Delta, \omega}$  be a sequence of *fixed* parallel hyperplanes. The following rather surprising corollary reveals that the integration in equation 2.10 can be replaced by a sum over subsequent pairs  $E_k, E_{k+1}$  without changing the value of the Euler-characteristic  $\chi$ : This means that the usual approximation of integrals by sums in the Riemannian sense still gives the exact value in this case!

**Corollary 2.1.21.** *Let  $R \in \mathcal{R}$ . Under the conditions of Theorem 2.1.20 the following representations of the Euler-characteristic  $\chi(R)$  hold:*

$$\chi^d(R) = \sum_k (\chi(Y_{k\Delta} \cup Y_{(k+1)\Delta}) - \chi(Y_{k\Delta})) \quad (2.11)$$

$$\chi^d(R) = \sum_k (\chi(Y_{k\Delta}) - \chi(Y_{k\Delta} \cap Y_{(k+1)\Delta})) \quad (2.12)$$

This corollary confirms that the Euler-characteristic is a truly topological quantity and not dependent on the metric of the underlying space at all.

**Proof. (Corollary 2.1.21)** The proof of Theorem 2.1.20 can be literally translated by replacing  $\frac{1}{\Delta} \int$  with  $\sum_k$ . For details see (Nagel, Ohser and Pischang 2000).

As already mentioned, we will use Corollary 2.1.21 to derive the *Ohser-Mücklich-estimators* for the Euler-characteristic in two and three dimensions. We will always assume in the following that the specimens of porous-media which we examine are topologically open and closed with respect to the unit lattice vectors such that corollary 2.1.21 can be recursively applied (see sections 2.3.1 and 2.3.2). However we will see that Corollary 2.1.21 is useless if we approximate a two-dimensional set on a hexagonal graph instead of a cuboidal one. In image analysis, if only a pixel image of the set  $A \subset \mathbb{R}^2$  is available, the value  $\chi(A)$  is often estimated by the Euler-characteristic of a set  $\tilde{A}$  which is formed by intersecting  $A$  with a hexagonal graph whose vertices are the pixels of the image.  $\tilde{A}$  consists of the union of all triangles which have all three vertices in  $A$  plus all edges with both vertices in  $A$  plus all graph vertices contained in  $A$ . Note the implicit assumption that whenever two vertices of the hexagonal graph are contained in  $A$ , then the whole edge joining these vertices belongs to  $A$  and similarly if the three edges bounding a triangle belong to  $A$ , so does the whole triangle which of course doesn't need to be correct for non-convex  $A$ .

Essentially from definition 2.1.12 it follows that in 2-dimensions the Euler-characteristic of a set  $\chi(A)$  can be interpreted as  $\#(\text{connected components}) - \#(\text{holes})$ . A similar relationship also holds for the 3-dimensional case where in addition the number of completely enclosed cavities must be added.

The Euler-characteristic  $\chi(\tilde{A})$  can then be calculated because  $\tilde{A} \subset \mathcal{CB}$  and the decomposition into basics is naturally provided by the hexagonal graph. Hence in principle one could use definition 2.1.14 to calculate  $\chi(\tilde{A})$ , but here we give a graph-theoretical result from (Serra 1982) which again coincides with the definition of 2.8:

**Theorem 2.1.22.** (*Euler's relation for hexagonal graphs*) Let  $A \subset \mathbb{R}^2$  and  $\tilde{A}$  its representation on a hexagonal graph. If we denote:

$$\begin{aligned} n &:= \text{number of connected components in } \tilde{A}, \\ h &:= \text{number of holes in } \tilde{A}, \\ t &:= \text{number of triangles in } \tilde{A}, \\ e &:= \text{number of edges in } \tilde{A}, \\ v &:= \text{number of vertices in } \tilde{A}, \end{aligned}$$

then: 
$$\chi(\tilde{A}) = n - h = v + t - e. \quad (2.13)$$

**Proof.** (**Theorem 2.1.22**) The proof proceeds by induction in the number of graph vertices. For  $v = 0$  the statement reads  $0 = 0$  and for  $v = 1$  we have  $1 - 0 = 1 + 0 - 0 = 1$ . Now assume equation 2.13 holds for  $v = k$  vertices. It remains to show that adjunction of a further vertex and associated edges does not change equation 2.13. This in fact true, since for each of the (modulo hexagonal rotation) 14 possible neighbourhood configurations of the added vertex the balance is kept. Exemplarily consider the configuration where the six neighbours belong alternatingly to the graph or not. Adding a vertex means  $\Delta v = 1$ ,  $\Delta t = 0$ ,  $\Delta e = 3$ . Hence, the change in the right hand side of equation 2.13 is  $-2$ . On the left hand side, three different situations are possible. Either the three neighbouring graph vertices belong all to the same connected component, then  $\Delta n = 0$  and  $\Delta h = +2$  or two of them belong to the same connected component, but then  $\Delta n = -1$ ,  $\Delta h = +1$  or all of them belong to different connected components, then  $\Delta n = -2$  and  $\Delta h = 0$ . Therefore in either case also the left hand side changes by  $-2$  and equation 2.13 is preserved. For the other neighbourhood configurations the same argument is even easier. A detailed listing of all possible configurations and the above argument for each such configuration is provided in (Serra 1982, p. 185).  $\square$

## 2.2 The Crofton and the Kinematic-Formula

To be able to work comfortably with Minkowski functionals we need more explicit representations. A basic prerequisite for modelling porous media using Minkowski functionals is being able to compute their values for a pixel image of such a porous structure. The porous media we are investigating are always assumed to be sets from the convex ring  $\mathcal{R}$  but definitions 2.1.2 and 2.1.11 are inappropriate for computing purposes. But luckily integral geometry can provide us with integral expressions for the Minkowski functionals that can be discretized in a straightforward manner and will allow us to (at least approximately) compute values for Minkowski functionals of polyconvex sets from pixel images. A second advantage of these integral representations is that they will provide us with a more intuitive geometrical understanding of what Minkowski functionals really are. This section is devoted to the derivation and interpretation of these integral expressions. Again we follow (Schneider 2004).

### 2.2.1 The Crofton Formula

Let  $SO_d$  be the usual orthogonal group in  $\mathbb{R}^d$ . It has the structure of a compact topological group when equipped with the usual topology by representing elements  $\theta \in SO_d$  with orthogonal matrices  $R \in \mathbb{R}^{d^2}$  (after choice of an arbitrary orthonormal basis). It is well known that (locally) compact topological groups carry an up to a proportionality constant unique measure defined on the Borel- $\sigma$ -algebra which is regular and invariant with respect

to group multiplication. This measure is called the *Haar measure* and we denote the Haar-measure of  $SO_d$  by  $\nu$ , normalized such that it becomes a probability measure. Let further  $\mathcal{L}_d^q$  denote the collection of  $q$ -dimensional linear subspaces of  $\mathbb{R}^d$ . If we fix an arbitrary  $L \in \mathcal{L}_d^q$  we can define a rotation-invariant probability measure on  $\mathcal{L}_d^q$  as the image measure of  $\nu$  under the mapping  $R \rightarrow R \cdot L$ . This measure we denote by  $\nu_q$ . Because of rotation-invariance this measure is independent of the choice of  $L$ . Similarly  $\mathcal{E}_q^n$  denotes the space of  $q$ -dimensional affine subspaces of  $\mathbb{R}^d$  (space of  $q$ -flats). We can define a measure  $\mu_q$  on  $\mathcal{E}_q^d$  as the image measure of  $\lambda_{d-q} \times \nu$  under the mapping  $(x, R) \rightarrow R \cdot (L + x)$ ,  $x \in L^\perp$ , where  $\lambda_k$  is  $k$ -dimensional Lebesgue-measure. With these definitions at hand, it makes perfect sense to consider the following integral:

$$\int_{\mathcal{E}_q^d} V_j(R \cap E) \mu_q(dE) \text{ for } R \in \mathcal{R}.$$

The solution of this integral involving an intrinsic volume is again a (scaled) intrinsic volume, a fact which is the content of the *Crofton-formula*. We will first give the result for  $K \in \mathcal{K}$  and then extend it to the convex ring  $\mathcal{R}$ . Because the Crofton-formula is the basis for all the estimators for Minkowski-functionals we use in later sections, we include the proofs here.

**Theorem 2.2.1. (Crofton-formula for convex sets)** *Let  $K \in \mathcal{K}$ . For  $q \in 1, \dots, d-1$  and  $j \leq q$  we have*

$$\int_{\mathcal{E}_q^d} V_j(K \cap E) \mu_q(dE) = \frac{\Gamma(\frac{q+1}{2}) \Gamma(\frac{d-q+j+1}{2})}{\Gamma(\frac{j+1}{2}) \Gamma(\frac{d+1}{2})} V_{d-q+j}(K).$$

**Proof. (Theorem 2.2.1)** The proof is a direct consequence from the Hadwiger characterization Theorem 2.1.4. Consider the integral as a functional  $\phi$  on  $\mathcal{K}$ :

$$\phi(K) := \int_{\mathcal{E}_q^d} V_j(K \cap E) \mu_q(dE) \quad \forall K \in \mathcal{K}.$$

$\phi$  is motion-invariant, additive and Hausdorff-continuous. According to Theorem 2.1.4 it must have the following representation:

$$\phi(K) = \sum_{m=0}^d c_m V_m(K). \quad (2.14)$$

Writing  $\phi$  a little bit more complicated,

$$\phi(K) = \int_{\mathcal{L}_q^n} \int_{L^\perp} V_j(K \cap (L + x)) \lambda_{d-q}(x) \nu_q(dL),$$

one learns that  $\phi$  is homogeneous of degree  $d-q+j$  because  $\lambda_{d-q}$  is  $d-q$ -homogeneous and  $V_j$  is  $j$ -homogeneous. But then only the  $c_{d-q+j} V_{d-q+j}(K)$ -summand in (2.14) can be different from 0.

The value of the constant can be found by choosing  $K = B^d(0)$ . The Steiner-formula 2.1.1 completely determines all intrinsic volumes for the ball:

$$\sum_{m=0}^d \rho^{d-m} \kappa_{d-m} V_m(B^d(0)) = V_d(B^d(0) + \rho B^d(0)) = (1 + \rho)^d \kappa_d = \sum_{m=0}^d \rho^{d-m} \binom{d}{m} \kappa_d.$$



which implies  $V_j(B^d(0)) = \frac{\binom{d}{j}\kappa_d}{\kappa_{d-j}}$ . Finally observe that:

$$\begin{aligned} c_{d-q+j}V_{d-q+j}(B^d(0)) &= \int_{\mathcal{E}_q^d} V_j(B^d(0) \cap E)\mu_q(dE) \\ &= \int_{SO_d} \int_{L^\perp} V_j(B^d(0) \cap R(L+x))\lambda_{d-q}(dx)\nu(dR) \\ &= \int_{L^\perp \cap B^d} (1 - \|x\|^2)^{j/2} V_j(B^d(0) \cap L)\lambda_{d-q}(dx) = \frac{\binom{q}{j}\kappa_q}{\kappa_{q-j}} \int_{L^\perp \cap B^d} (1 - \|x\|^2)^{j/2} \lambda_{d-q}(dx), \end{aligned}$$

and the last integral is of the beta-type and can therefore explicitly be done.  $\square$

The extension of the Crofton-formula to the convex ring  $\mathcal{R}$  is now easy thanks to Theorem 2.1.7. To simplify notation we denote  $\mathcal{A}(n) := 2^{\{1, \dots, n\}} \setminus \emptyset$  and set  $K_A := K_{i_1} \cap \dots \cap K_{i_{|A|}}$ . Then the assertion of Theorem 2.1.7 can be written more concisely as:

$$\phi(K_1 \cup \dots \cup K_n) = \sum_{A \in \mathcal{A}(n)} (-1)^{|A|-1} \phi(K_A), \quad (2.15)$$

where  $\phi$  is an additive functional on  $\mathcal{K}$ .

For  $R \in \mathcal{R}$  we choose a representation  $R = \bigcup_{m=1}^n K_m$ ,  $K_m \in \mathcal{K}$  and then we calculate:

$$\begin{aligned} \int_{\mathcal{E}_q^d} V_j(R \cap E)\mu_q(dE) &= \int_{\mathcal{E}_q^d} V_j\left(\bigcup_{m=1}^n (K_m \cap E)\right)\mu_q(dE) \\ &= \int_{\mathcal{E}_q^d} \sum_{A \in \mathcal{A}(n)} (-1)^{|A|-1} V_j\left(\underbrace{K_A \cap E}_{\in \mathcal{K}}\right)\mu_q(dE) \\ &= \sum_{A \in \mathcal{A}(n)} (-1)^{|A|-1} \frac{\Gamma(\frac{q+1}{2})\Gamma(\frac{d-q+j+1}{2})}{\Gamma(\frac{j+1}{2})\Gamma(\frac{d+1}{2})} V_{d-q+j}(K_A \cap E) = \frac{\Gamma(\frac{q+1}{2})\Gamma(\frac{d-q+j+1}{2})}{\Gamma(\frac{j+1}{2})\Gamma(\frac{d+1}{2})} V_{d-q+j}(R). \end{aligned}$$

So that we can summarize:

**Theorem 2.2.2. (Crofton-formula for polyconvex sets)** Let  $R \in \mathcal{R}$ . For  $q \in 1, \dots, d-1$  and  $j \leq q$  we have

$$\int_{\mathcal{E}_q^d} V_j(R \cap E)\mu_q(dE) = \frac{\Gamma(\frac{q+1}{2})\Gamma(\frac{d-q+j+1}{2})}{\Gamma(\frac{j+1}{2})\Gamma(\frac{d+1}{2})} V_{d-q+j}(R).$$

For compressed and easily readable information about more general *local* versions of the Crofton-formula and other fundamental integral geometric relationships involving generalized curvature measures see (Schneider 2004).

### 2.2.2 Basic Consequences from the Crofton-Formula

Due to the Crofton-formula are now able to get a little bit more insight in the nature of the Minkowski functionals and gain some geometrical intuition about them which we have already conjectured before from their explicit representation for convex polytopes (equation 2.1). Setting  $j = 0$  in the Crofton-formula (Theorem 2.2.1) above reveals:

**Corollary 2.2.3.** (*Integral geometric representation of intrinsic volumes*) For  $m \in \{0, \dots, d\}$  and  $R \in \mathcal{R}$  the intrinsic volumes can be written as

$$V_q(R) = c_q^d \cdot \int_{\mathcal{E}_{d-q}^d} \chi(R \cap E) \mu_{d-q}(dE), \quad (2.16)$$

where  $c_q^d = \left( \frac{\Gamma(\frac{q+1}{2})\Gamma(\frac{d-q+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{d+1}{2})} \right)^{-1}$ .

For convex sets  $K \in \mathcal{K}$  the value  $\chi(K \cap E)$  is simple:  $\chi(K \cap E) = \mathbf{1}_{\{E \cap K \neq \emptyset\}}$ . Therefore:

For convex sets  $K \in \mathcal{K}$ , the intrinsic volumes  $V_q(K)$  are equal to the (normalized) measure of the set of  $d-q$ -flats intersecting  $K$ .

From corollary 2.2.3 the missing points v),vii) in the proof of Theorem 2.1.3 are now obvious.

For convex sets  $K \in \mathcal{K}$ , equation (2.16) can be further simplified to:

$$V_q(K) = c_q^d \int_{\mathcal{L}_q^d} \lambda_q(p_q(K|L)) \nu_q(dL), \quad (2.17)$$

where  $p(\cdot|L)$  denotes projection onto  $L$ .

From equation (2.17) we immediately derive a handy geometrical interpretation for the intrinsic volume  $V_1(K)$  for convex bodies when setting  $q = 1$ . Up to a constant it is equal to the **mean breadth**  $\bar{b}(K)$  of  $K$ . Another geometrical interpretation for  $V_1(K)$  on  $\mathcal{K}$  when the boundary of a convex set  $\partial K$  is a  $C^2$ -hypersurface will be given in section 2.2.4.

### 2.2.3 The Principal Kinematic Formula

To be able to give explicit expressions for the Minkowski functionals in Chapter 3 we need to state another important integral geometric theorem. Whereas the Crofton formulae describe how to calculate averages of intrinsic volumes for convex bodies over spaces of flats, the so-called *iterated kinematic formula* describes how to calculate such averages over the whole group of rigid motions  $G_d$ . Again we follow (Schneider 2004). For  $K \in \mathcal{K}$  we denote  $gK$  its image under the rigid motion  $g \in G_d$ . Similarly as for flats the Haar measure  $\mu$  on  $G_d$  can be described as the image measure of  $\lambda_d \times \nu$  under the mapping  $(x, R) \rightarrow g$ , where  $gy = Ry + x$  for  $y \in \mathbb{R}^d$ . Then the following theorem holds:

**Theorem 2.2.4.** (*principal and iterated kinematic formula*) Let  $n \geq 2$ . Furthermore  $K_1, \dots, K_n \in \mathcal{K}, g_2, \dots, g_n \in G_d, j \in \{0, \dots, d\}$ . Then:

i)

$$\int_{G_d} V_j(K_1 \cap g_2 K_2) \mu(dg) = \sum_{m=j}^d \frac{m! \kappa_m (d-m+j)! \kappa_{d-m+j}}{j! \kappa_j d! \kappa_d} V_m(K_1) V_{d-m+j}(K_2).$$

ii)

$$\begin{aligned}
 & \int_{G_d} \dots \int_{G_d} V_j(K_1 \cap g_2 K_2 \cap \dots \cap g_n K_n) \mu(dg_2) \dots \mu(dg_n) \\
 = & \sum_{\substack{m_1, \dots, m_n = j \\ m_1 + \dots + m_n = (n-1) \cdot d + j}}^d \frac{m_1! \kappa_{m_1} \dots m_n! \kappa_{m_n}}{j! \kappa_j (d! \kappa_d)^{n-1}} V_{m_1}(K_1) \dots V_{m_n}(K_n).
 \end{aligned}$$

**Proof.** For a proof see the appendix, section 7.2.

### 2.2.4 Curvature Integrals and Minkowski-Functionals

In this subsection we restrict our attention to three-dimensional Euclidean space  $\mathbb{R}^3$ . We will see that for convex sets with suitably regular boundary the intrinsic volume  $V_1(\cdot)$  coincides up to proportionality with the *integral of mean curvature* which leads to a useful second interpretation of this functional and as a side-product establishes a connection between the integral of mean curvature and mean breadth. Furthermore we will see that  $2V_2(\cdot)$  plays the role of the surface area. We will follow an easy argument from (Santaló 2004). Let us assume that  $K \in \mathcal{K}$  and  $\partial K$  has the regularity of a  $C^2$ -surface. This ensures that  $\forall x \in \partial K$  the two *principal curvatures*  $p_1(x), p_2(x)$  exist. Recall that principal curvatures denote the inverse radii  $r_1, r_2$  of the largest and smallest osculating circle in  $x$  respectively. Then we can make the following definition:

**Definition 2.2.5. (integral of mean and total curvature)** Let  $\partial K$  a  $C^2$ -surface in  $\mathbb{R}^3$  and  $\forall x \in \partial K$  denote by  $p_1(x), p_2(x)$  the principal curvatures in  $x$ . If furthermore  $d\sigma_K$  denotes the area element on  $\partial K$ , then

- the quantity

$$M(K) := \int_{\partial K} \frac{p_1(x) + p_2(x)}{2} d\sigma_K(x)$$

is called *integral of mean curvature* of the convex set  $K$ .

- the quantity

$$K(K) := \int_{\partial K} p_1(x)p_2(x) d\sigma_K(x)$$

is called *integral of total curvature* of the convex set  $K$ .

These definitions can readily be generalized to higher dimensions and larger set classes. Details follow at the end of the section. The product  $p_1(x)p_2(x)$  is the well-known *Gauss-Kronecker-curvature* which relates the area element  $d\sigma_K$  to the Hausdorff-measure  $\mathcal{H}_2$  on  $S^2$  in the following sense:

$$p_1(x)p_2(x)d\sigma = d\mathcal{H}_2.$$

This relationship can now be used to describe the area element  $d\sigma_{K_\rho}$  for the parallel body  $K_\rho$  of  $K$ . Because the curvature radii  $r_i(x)$  are expanded during the dilation process to  $r_i(x) + \rho$  we get:

$$d\sigma_{K_\rho} = (r_1(x) + \rho)(r_2(x) + \rho)d\mathcal{H}_2.$$

This allows us to calculate the surface area  $s$  of the parallel body:

$$\begin{aligned}
s(K_\rho) &= \int_{\partial K_\rho} d\sigma_{K_\rho} = \int_{S^2} r_1(x)r_2(x)d\mathcal{H}_2 + \rho \int_{S^2} (r_1(x) + r_2(x))d\mathcal{H}_2 + \rho^2 \int_{S^2} d\mathcal{H}_2 \\
&= \int_{\partial K} d\sigma_K + 2\rho \int_{\partial K} \frac{p_1(x) + p_2(x)}{2} d\sigma_K + \int_{\partial K} \rho^2 p_1(x)p_2(x) d\sigma_K \\
&= s(K) + 2\rho M(K) + \rho^2 K(K).
\end{aligned} \tag{2.18}$$

The volume  $v$  of  $K_\rho$  can now be calculated in two different ways. The first method is a direct application of the Steiner formula 2.1.1:

$$v(K_\rho) = V_3(K) + 2\rho V_2(K) + \pi\rho^2 V_1(K) + \frac{4\pi}{3}\rho^3 V_0(K),$$

but we can also use equation 2.18 above to get:

$$v(K_\rho) = v(K) + \int_0^\rho s(K_\rho) d\tilde{\rho} = v(K) + \rho s(K) + \rho^2 M(K) + \frac{1}{3}\rho^3 K(K).$$

Comparing coefficients (and using equation 2.17) finally establishes:

$$s(K) = 2V_2(K), \quad M(K) = \pi V_1(K) = 2\pi\bar{b}(K), \quad K(K) = 4\pi\chi(K), \tag{2.19}$$

which nicely connects all intrinsic volumes with geometric quantities. The same argument can be applied for other dimensions as well. The last equation proves the *Gauss-Bonnet-Theorem* (Theorem 2.1.17) from subsection 2.1.5 in the case of convex  $K$  with  $C^2$ -boundary. If the boundary of  $K$  is not  $C^2$ , then curvature integrals can be defined as:

$$M(K) = \lim_{\rho \rightarrow 0} M(K_\rho), \quad K(K) = \lim_{\rho \rightarrow 0} K(K_\rho).$$

The geometric quantities  $s(\cdot)$ ,  $M(\cdot)$  and  $K(\cdot)$  are additive which allows us to carry over the geometric interpretations of the Minkowski functionals to the convex ring  $\mathcal{R}$ .

For the rest of this thesis we restrict our attention to dimensions  $d = 2, 3$ . Hence in the following there are only four Minkowski-functionals to consider for all of which we have an intuitive geometrical interpretation:

geometric characteristic	notation	functional in $d = 2$	functional in $d = 3$
area/volume	$v$	$V_2 = W_0$	$V_3 = W_0$
boundary length/surface area	$s$	$2V_1 = 2W_1$	$2V_2 = 3W_1$
integral of mean curvature	$M$	—	$\pi V_1 = 3W_2$
Euler characteristic	$\chi$	$V_0 = \frac{1}{\pi}W_2$	$V_0 = \frac{3}{4\pi}W_3$

**Table 2.1:** *intrinsic volumes  $V_m(\cdot)$  and Minkowski-functionals  $W_m(\cdot)$  in two and three dimensions together with their geometrical interpretations.*

In the following we don't make a difference between the Minkowski-functionals or intrinsic volumes and the geometrical characteristics they represent. We will take the freedom and call all those quantities simply Minkowski-functionals. It will be clear from the context what particular quantity is addressed.

For arbitrary dimension  $d$ , definition 2.2.5 can be generalized as follows:

**Definition 2.2.6.** (*n<sup>th</sup> general integrals of mean curvature*) Let  $\partial S$  be a  $C^2$ -hypersurface in  $\mathbb{R}^d$ . Then the *n<sup>th</sup> general integral of mean curvature* is defined for  $n \in \{1, \dots, d-1\}$  as:

$$M_n(\partial S) = \binom{d-1}{n} \int_S \{p_{i_1}, \dots, p_{i_n}\} d\sigma(S),$$

where  $\{p_{i_1}, \dots, p_{i_n}\}$  denotes the *n<sup>th</sup>-elementary symmetric function* of the  $d-1$  principal curvatures.

Analogously as we did above, relationships between the Minkowski-functionals in  $d$  dimensions and the curvature integrals of definition 2.2.6 can be established. For details refer to (Santaló 2004).

## 2.3 Ohser-Mücklich-Estimators for Minkowski-Functionals

Because the specimens of porous media we examine in this thesis are available to us only as digitized pixel images on a computer, we need a method for approximate calculation of Minkowski-functionals from pixel images. Because we always assume that the porous media under consideration consist of a single homogeneous two-phase-material, their solid phase can be represented by a subset  $S \subset \mathbb{R}^d$ , where  $d = 2, 3$ . As a technical restriction, we always assume that the set  $S$  is an element of the *extended convex ring*  $\mathcal{S}$ , ie.  $S^W := S \cap W \in \mathcal{R}$ ,  $\forall W \in \mathcal{K}$ . The collection  $\mathcal{S}$  is the right class to consider because this allows us to understand the porous structures under investigation as infinitely extended which will be convenient in the following for mathematical reasons whereas in practice of course we can only observe *bounded* specimens of the porous medium, ie. we observe the compact set  $S^W$ . Usually  $W$  is assumed to be a convex body  $W \in \mathcal{K}$  which we call *observation window*. In most cases for convenience  $W$  is a rectangle ( $d = 2$ ) or a cuboid ( $d = 3$ ). For example in two dimensions  $W$  could play the role of the computer screen. The pixel image of  $S^W$  can be interpreted as the intersection  $S_{\Gamma(h)}^W = S^W \cap \Gamma(h)$  where  $\Gamma(h)$  is a cubic lattice  $\Gamma(h) = h\mathbb{Z}^d$ . We call  $h$  the *lattice spacing* which is used to control the resolution of the pixel image  $S_{\Gamma(h)}^W$ . Other lattices would also be possible (eg. cuboidal), but for the sake of simplicity of formulae we restrict ourselves to the cubic case. It will be convenient later on to interpret the sets  $S^W$  and  $S_{\Gamma(h)}^W$  as mappings  $S^W(x)$  and  $S_{\Gamma(h)}^W(x)$  which are defined as the restriction of the  $\mathbb{R}^d$ -indicator function for  $S$  to  $W$  and  $W \cap \Gamma(h)$  respectively:

$$\begin{aligned} S^W(x) : W &\rightarrow \{0, 1\} & S_{\Gamma(h)}^W(h)(x) : W \cap L(h) &\rightarrow \{0, 1\} \\ S^W(x) &= \mathbf{1}_S(x) & S_{\Gamma(h)}^W(x) &= \mathbf{1}_S(x) \end{aligned}$$

The whole information about a pixel image  $S_{\Gamma(h)}^W$  is stored in a binary string consisting of the values of the map  $b_{S_{\Gamma(h)}^W}$  (defined below) along with an indication about the dimension of the image and the number of pixels per dimension. A further convenient technical restriction we make in the following is that for our chosen observation window  $W$  the set  $S^W$  is morphologically open and closed with respect to all edges, face diagonals and spatial diagonals (only  $d = 3$ ) of a unit  $\Gamma$ -lattice cell.

The set  $W \cap L(h)$  is partitioned into quadratic (or cubic) lattice cells. Ohser and Mücklich (Ohser and Mücklich 2000) have developed formulae which give accurate estimators for the Minkowski-functionals  $V_m(\cdot)$  of continuous porous specimens  $S^W$  from discrete approximations  $S_{\Gamma(h)}^W$  on the lattice  $S^W \cap L(h)$ . These estimators depend only on the vertex-

configurations of lattice cells and reduce the workload to simple pixel counting. The estimators can therefore easily be implemented on a computer. The next two subsections are devoted to the description of these estimators and how they are derived from the Crofton-formula-representation (Corollary 2.2.3). First we discuss the Ohser-Mücklich-estimators for the two-dimensional case and then describe the formulae for the three-dimensional situation. In section 2.4 we will represent porous media specimens  $S$  as realizations of stationary random sets  $\Theta$  with values in  $\mathcal{S}$ . Such sets are of infinite extent by definition and for unbounded  $S \in \mathcal{S}$  usually extensions of Minkowski-functionals (if they exist) lead to  $V_m(S) = \infty$ . On the other hand the values  $V_m(S^W)$  are finite and rigorously defined in our framework. However, they depend heavily on the choice of the observation window  $W$ . Therefore we need expressions for the Minkowski functionals which are independent of the size of the observation window  $W$ . Note that it makes perfect sense if we formulate the Ohser-Mücklich-estimators for the *Minkowski-functionals per unit volume*,

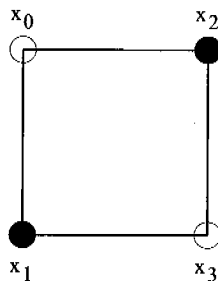
$$\bar{V}_m(S^W) := \frac{V_m(S^W)}{V_d(W)}, \quad m \in 0, \dots, d. \quad (2.20)$$

a definition which will be generalized in definition 2.4.3 for the mean values of Minkowski functionals for random sets with values in  $\mathcal{S}$ . The functional  $\bar{v} = \bar{V}_d$  is usually called *porosity* or *volume-fraction*. Please be aware that the expressions defined in equation 2.20 are indeed independent of the size *but not on the shape* of the observation window  $W$  (see also subsection 3.2.1 below).

### 2.3.1 Ohser-Mücklich-Estimators for $d = 2$ .

In two dimensions there are only three Minkowski functionals to consider, namely the porosity  $\bar{v}$ , the boundary length per unit volume  $\bar{s}$  and the Euler-characteristic  $\bar{\chi}$  per unit volume.

Each vertex  $x \in \Gamma(h)$  belongs either to  $S_\Gamma^W(h)$  or  $[S_\Gamma^W(h)]^c$  depending on the binary value of  $S^W(x)$ . Therefore to every quadratic lattice cell  $\Gamma_{x_0}(h) = \{x_0, x_1 = x_0 - h \cdot e_y, x_2 = x_0 + h \cdot e_x, x_3 = x_0 + h \cdot e_x - h \cdot e_y\}$  which we index by the upper left vertex  $x_0$  following the nomenclature in (Ohser and Mücklich 2000) and where  $e_x$  and  $e_y$  are unit vectors in the directions of the coordinate axes, we can assign a unique vertex configuration number  $b_{S_\Gamma^W}(x_0) = \sum_{i=0}^3 S_\Gamma^W(x_i) \cdot 2^i$  by binary coding. An example is illustrated in figure 2.1.



**Figure 2.1:** A lattice cell  $\Gamma_{x_0}(h)$  with  $b_{S_\Gamma^W}(x_0) = 6$ .

The only information needed to construct the Ohser-Mücklich estimators involves the *summary statistic*  $h_{S_\Gamma^W}(i) = \sum_{x_0 \in \Gamma(h)} \mathbf{1}_{\{b_{S_\Gamma^W}(x_0)=i\}}$ ,  $i \in 0 \dots 15$  which simply counts the occurrences of the 16 different vertex configurations throughout the image.

The most natural estimator for the porosity  $\bar{v}$  trivially counts a lattice cell whenever it contains a vertex in  $S_\Gamma^W$ . Only occurrences of *one* of the four vertices (here  $x_0$ ) have to be counted because the other three ( $x_1, x_2, x_3$ ) are accounted for by neighbouring lattice cells. Therefore we have simply:

$$\hat{v}(S_\Gamma^W) = \frac{1}{n_0} \sum_{i=0}^7 h_{S_\Gamma^W}(2i+1), \quad (2.21)$$

where  $n_0$  denotes the total number of lattice-cells in  $\Gamma(h)$ . Of course this estimator doesn't depend on the lattice spacing  $h$  at all. Note that this estimator can be understood as a discretized Crofton-integral. Setting  $q = d$  in the formula of Corollary 2.2.3 we obtain a representation for the volume (or area) per unit volume:

$$\bar{v}(S^W) = \bar{V}_d(S^W) = \frac{1}{v(W)} \int_W \chi(S^W \cap \{x\}) dx. \quad (2.22)$$

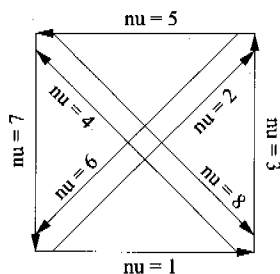
If we discretize equation 2.22 on the lattice  $\Gamma(h) \cap W$ , we immediately get the Ohser-Mücklich-estimator of equation 2.21 for the porosity  $\bar{v}$ .

For the boundary length  $\bar{s}(S^W)$  we use the Crofton-formula-representation from Corollary 2.2.3 while setting  $q = d - 1 = 1$  to get:

$$\bar{s}(S^W) = 2\bar{V}_1(S^W) = \frac{\pi}{v(W)} \int_0^{2\pi} \int_{-\infty}^{\infty} \chi(S^W \cap e_{r,\phi}) dr \frac{d\phi}{2\pi}, \quad (2.23)$$

where  $e_{r,\phi}$  denotes the straight line with angle  $\phi$  to the  $x$ -axis and distance  $r$  from the origin. The quantity  $\chi(S^W \cap e_{r,\phi})$  counts the number of *chords*, i.e. pieces on the line  $e_{r,\phi}$  which lie completely in  $S^W$ .

The simplest way to discretize the integral in equation 2.23 is to consider the 8 *prominent directions*  $\nu \in 1, \dots, 8$  dictated by the 2D-quadratic lattice. These are those directions, for which counting chords is easy, namely the 4 directions along the lattice edges and the 4 diagonal directions. For an illustration of the rose of relevant directions, see figure 2.2 below.



**Figure 2.2:** The prominent directions in a 2D quadratic-lattice.

Once a direction  $\nu$  is fixed counting chords on lines pointing in this direction is easy because instead of counting the chords themselves we can count the number of transitions from  $S^W$  to  $[S^W]^c$ . The estimator for  $\bar{s}$  can now be written as follows:

$$\hat{\bar{s}}(S_\Gamma^W) = \frac{\pi}{8 \cdot n_0} \sum_{\nu=0}^7 \frac{1}{r_\nu} \sum_{i=0}^{15} h_{S_\Gamma^W}(i) \cdot \mathbf{1}_{\{x_{g_1(\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_2(\nu)} \notin S_\Gamma^W\}}, \quad (2.24)$$

where  $n_0$  is again the total number of lattice cells in  $\Gamma(h)$ , further  $g_1(\nu), g_2(\nu)$  for every direction  $\nu$  denote the vertex index according to our specified nomenclature of the initial point and the endpoint of the corresponding arrow in figure 2.2 and  $r(\nu) = |x_{g_2(\nu)} - x_{g_1(\nu)}|$  stands for their Euclidean distance. Please note that the dependence of the estimator on the lattice spacing  $h$  is given via  $r_\nu$ , e.g.  $r_2 = \sqrt{2} \cdot h$ .

Finally we describe the Ohser-Mücklich-estimator for the Euler-characteristic  $\chi$ . For this purpose we denote the points of our quadratic lattice  $\Gamma(h)$  by  $x_{ij} := \{(ih, jh) | i, j \in \mathbb{Z}\}$ . Furthermore we denote by  $Y_j := S_\Gamma^W \cap e_{\{y=jh\}_{-(jh)}}$  the intersection of  $S_\Gamma^W$  with the line  $y = jh$  shifted back to the  $x$ -axis. A two-fold recursive application of equation 2.12 first in  $y$ -, then in  $x$ -direction reveals:

$$\begin{aligned}
\chi(S_\Gamma^W) &= \sum_j (\chi(Y_j) - \chi(Y_j \cap Y_{j+1})) \\
&= \sum_{ij} \left( \mathbf{1}_{S_\Gamma^W}(x_{ij}) - \mathbf{1}_{S_\Gamma^W}(x_{ij}) \mathbf{1}_{S_\Gamma^W}(x_{(i+1)j}) \right) - \left( \mathbf{1}_{S_\Gamma^W}(x_{ij}) \mathbf{1}_{S_\Gamma^W}(x_{i(j+1)}) \right. \\
&\quad \left. - \mathbf{1}_{S_\Gamma^W}(x_{ij}) \mathbf{1}_{S_\Gamma^W}(x_{i(j+1)}) \mathbf{1}_{S_\Gamma^W}(x_{(i+1)j}) \mathbf{1}_{S_\Gamma^W}(x_{(i+1)(j+1)}) \right), \\
&= \sum_{ij} \mathbf{1}_{S_\Gamma^W}(x_{ij}) - \mathbf{1}_{S_\Gamma^W}(x_{ij}) \mathbf{1}_{S_\Gamma^W}(x_{(i+1)j}) - \mathbf{1}_{S_\Gamma^W}(x_{ij}) \mathbf{1}_{S_\Gamma^W}(x_{i(j+1)}) \\
&\quad + \mathbf{1}_{S_\Gamma^W}(x_{ij}) \mathbf{1}_{S_\Gamma^W}(x_{i(j+1)}) \mathbf{1}_{S_\Gamma^W}(x_{(i+1)j}) \mathbf{1}_{S_\Gamma^W}(x_{(i+1)(j+1)}) \\
&= |\Gamma\text{-vertices in } S_\Gamma^W| - |\Gamma\text{-edges completely in } S_\Gamma^W| \\
&\quad + |\Gamma\text{-cells completely in } S_\Gamma^W|. \tag{2.25}
\end{aligned}$$

where summation can be restricted to the observation window  $W$ . Note that we have used here that the set  $S^W$  must be morphologically open and closed with respect to edges and diagonals of the unit cell. Note also that in this case 2.1.21 guarantees that this estimator gives the *exact* value of  $\chi(S^W)$ . Furthermore equation 2.25 confirms the relation between the Hadwiger and the topological definition of the Euler characteristic.

Instead of counting vertices, edges and cells we can count vertex configurations  $h_{S_\Gamma^W}(\cdot)$ . There are several possibilities to translate equation 2.25 into a formula for vertex configurations because for a given lattice cell we are free to choose one of the cell vertices, one of the horizontal edges and one of the vertical edges to count. If we choose for example the vertex  $x_1$  and edges  $x_0x_1$  and  $x_1x_2$  the translation of equation 2.25 reads:

$$\chi(S_\Gamma^W) = h_{S_\Gamma^W}(2) + h_{S_\Gamma^W}(6) - h_{S_\Gamma^W}(11).$$

After choosing all other possible collections of elements to count which corresponds to interpreting equation 2.25 from different directions (positive and negative  $x$ -, as well as positive and negative  $y$ -direction) and averaging over the rose of directions we get the final Ohser-Mücklich estimator for the Euler-characteristic per unit volume:

$$\hat{\chi}(S_\Gamma^W) = \frac{1}{4 \cdot n_0 \cdot h^2} \cdot \left( \sum_{i=0}^3 h_{S_\Gamma^W}(2^i) - \sum_{i=0}^3 h_{S_\Gamma^W}(15 - 2^i) + 2 \cdot (h_{S_\Gamma^W}(6) + h_{S_\Gamma^W}(9)) \right). \tag{2.26}$$

This provides us with another very natural interpretation of the Euler characteristic which might not be obvious at first glance. The first summand counts exactly the number of



convex vertices in  $S_{\Gamma}^W$  whereas the second one counts the number of concave vertices in  $S_{\Gamma}^W$ . In fact the topological Euler-Poincaré-characteristic can be defined in two dimensions as the difference between the number of convex and concave vertices. The appearance of the third part of the sum which we call *excess Euler characteristic* is rather unexpected. In fact, if the set  $S^W$  is topologically open and closed with respect to a sufficiently small structuring element the configurations of the third term will not show up. The excess-term, which we refer to as *excess Euler-characteristic* in the following, should be treated separately as we have done in our experiments with the Gibbsian Model (see section 5.9).

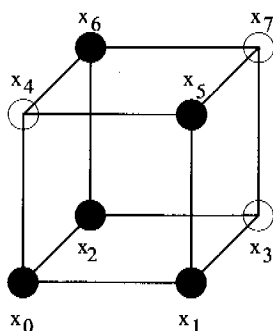
### 2.3.2 Ohser-Mücklich-Estimators for $d = 3$ .

We now describe the Ohser-Mücklich-estimators for the three-dimensional situation. In three dimensions the relevant Minkowski-functionals per unit volume are the porosity  $\bar{v}$ , the surface area per unit volume  $\bar{s}$ , the integral of mean curvature per unit volume  $\bar{M}$  and the Euler-characteristic  $\bar{\chi}$  per unit volume.

The lattice  $\Gamma(h)$  which is used to collect information about the porous structure to calculate the Ohser-Mücklich-estimators now consists of cubic lattice cells  $\Gamma_{x_0}(h)$  which are indexed by the front lower left vertex  $x_0$ . The vertex nomenclature is again chosen to be as in (Ohser and Mücklich 2000) where it is as follows:

$$\Gamma_{x_0}(h) = \{x_0, x_1 = x_0 + h \cdot e_x, x_2 = x_0 + h \cdot e_y, x_3 = x_0 + h \cdot e_x + h \cdot e_y, x_4 = x_0 + h \cdot e_z, x_5 = x_0 + h \cdot e_z + h \cdot e_x, x_6 = x_0 + h \cdot e_z + h \cdot e_y, x_7 = h \cdot e_z + h \cdot e_x + h \cdot e_y\}, \quad (2.27)$$

where  $e_x, e_y, e_z$  denote the Cartesian unit vectors. Depending on the binary values of the underlying field  $S^W(x_i)$  at the cell vertices  $x_i$  the vertex configuration numbers  $b_{S_{\Gamma}^W} = \sum_{i=0}^7 S_{\Gamma}^W(x_i) \cdot 2^i$  associated with a lattice cell  $\Gamma_{x_0}(h)$  in the binary image  $S_{\Gamma}^W$  are now between 0 and 255. An illustration is given in figure 2.3.



**Figure 2.3:** A 3D-lattice cell  $\Gamma_{x_0}(h)$  with  $b_{S_{\Gamma}^W}(x_0) = 103$ .

Again the Ohser-Mücklich-estimators are determined by the summary statistic  $h_{S_{\Gamma}^W}(i) = \sum_{x_0 \in \Gamma(h)} \mathbf{1}_{\{b_{S_{\Gamma}^W}(x_0)=i\}}, i \in 0 \dots 255$ . The exact formulae for most of these estimators are difficult to present in a concise way although all they do is counting occurrences of specific configurations  $h(i)$  and weighing them appropriately analogously as in the 2D-case above for which the formulae are precisely stated. For the 3D case we

will content ourselves with giving the global structure of the estimators which suffices to understand the proof of the 3D-convergence results in Chapter 4. For exact formulac please refer to (Ohser and Mücklich 2000).

For the porosity  $\bar{v}$  there is not much to say because equation 2.22 holds independent of dimension. The 3D-Ohser Mücklich estimator for the porosity  $\bar{p}$  is therefore a trivial generalization of the two-dimensional case. We only count lattice cells for which a specific vertex (here  $x_0$ ) belongs to the binary approximation  $S_{\Gamma(h)}^W$  of the porous phase  $S^W$ , where we recall that  $S^W = \{x \in W \mid S^W(x) = 1\}$  and  $W$  is the observation window:

$$\hat{v}(S_{\Gamma}^W) = \frac{1}{n_0} \cdot \sum_{i=0}^{127} h_{S_{\Gamma}^W}(2i+1),$$

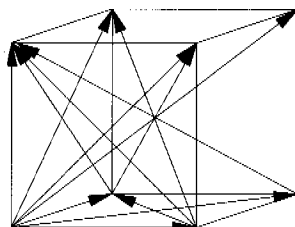
where  $n_0$  is the total number of lattice cells in  $\Gamma(h)$ .

For the surface area  $\bar{s}$  per unit volume we refer once more to the Crofton-formula-representation from Corollary 2.2.3 but this time with  $d = 3$  and  $q = 2$  which leads to:

$$\bar{s}(S^W) = \frac{4}{v(W)} \int_{S^2} \int_{\mathbb{R}^2} \chi(S^W \cap e_{y,\omega}) dy \frac{d\omega}{4\pi}, \quad (2.28)$$

where  $e_{y,\omega}$  denotes the straight line with direction  $\omega \in S^2$  and (two-dimensional) distance vector  $y$  from the origin.

For discretization we note that the number  $\chi(S^W \cap e_{y,\omega})$  still counts chords on lines  $e_{y,\omega}$ , so we can set up the estimator in principle as for  $d = 2$ . It is again a directional average over contributions to the estimator that now come from 26 prominent directions (3 cell edges, 6 face diagonals and 4 spatial diagonals each in both directions) which are shown in figure 2.4.



**Figure 2.4:** *Half the rose of relevant directions for the 3D-specific surface estimator  $\bar{s}$ .*

The estimator then becomes:

$$\hat{\bar{s}}(S_{\Gamma}^W) = \frac{4}{n_0} \sum_{\nu=0}^{25} \frac{c_{\nu}}{r_{\nu}} \sum_{i=0}^{255} h_{S_{\Gamma}^W}(i) \cdot \mathbf{1}_{\{x_{g_1(\nu)} \in S_{\Gamma}^W\}} \cdot \mathbf{1}_{\{x_{g_2(\nu)} \notin S_{\Gamma}^W\}}, \quad (2.29)$$

where again  $n_0$  is the total number of lattice cells in  $\Gamma(h)$ , and  $g_1(\nu), g_2(\nu)$  for each direction  $\nu$  denote the vertex index according to our specified nomenclature of the initial point and the endpoint of the corresponding arrow in figure 2.4 and  $r(\nu) = |x_{g_2(\nu)} - x_{g_1(\nu)}|$  stands for their Euclidean distance. The contributions of the prominent directions  $\nu$  to the estimator are now weighted by the relative area  $c_{\nu}$  of the corresponding *voronoi-cell* on the unit sphere which is associated with each direction. Let  $v_{\nu}$  denote the unit vector

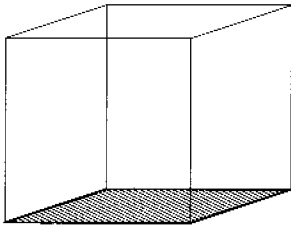
attached in the unit cubes' center of gravity  $G$  and pointing in direction  $\nu$ . Then the points  $G + v_\nu$  lie on the unit sphere  $\forall \nu$ . The voronoi-cell on the unit sphere  $S^2$  associated with a direction  $\nu_0$  is then  $V_{\nu_0} = \{x \in S^2 \mid |x - (G + v_{\nu_0})| = \min_\nu |x - (G + v_\nu)|\}$ , ie.  $V_{\nu_0}$  consists of all the points  $x \in S^2$  which are closer to the point  $G + v_{\nu_0}$  than to any other point  $G + v_\nu$ . The  $c_\nu$  can be calculated relatively easily for a regular cubic grid. The exact values can be learned from (Ohser and Mücklich 2000, pg. 116).

The next estimator we want to describe here is the one for the specific integral of mean curvature  $\bar{M}$ . One more time referring to the Crofton-formula-representation from Corollary 2.2.3, this time for  $d = 3$  and  $q = 1$ , reveals:

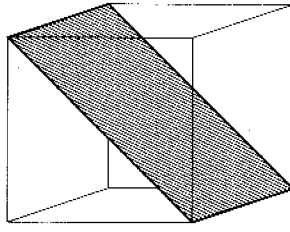
$$\bar{M}(S^W) = \frac{2\pi}{v(W)} \int_{S^2} \int_{-\infty}^{\infty} \chi(S^W \cap E_{y,\omega}) dy \frac{d\omega}{4\pi}, \quad (2.30)$$

where  $E_{y,\omega}$  is the plane with normal direction  $\omega$  and distance vector  $y$ .

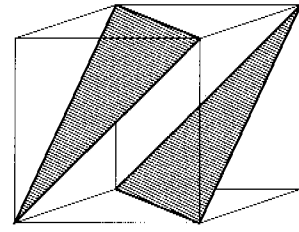
Whereas for the specific surface estimator  $\hat{s}$  the contributions came from lines  $(x_{g_1(\nu)}, x_{g_2(\nu)})$ , the contributions from  $\chi(S^W \cap E_{y,\omega})$  to  $\bar{M}$  are associated with *prominent planes*, namely cuboid faces, diagonal rectangles and diagonal triangles. An example for each of the three types of planes is given in figure 2.5.



A cuboidal face which contributes to  $\hat{M}$



A diagonal rectangle which contributes to  $\hat{M}$



Two diagonal triangles which contribute to  $\hat{M}$

**Figure 2.5:** The prominent planes involved in the calculation of the estimator for the 3D-specific mean curvature  $\bar{M}$ .

In total, the 26 planes prominent for the cubic lattice  $\Gamma$  are exactly those associated with the vectors  $G + v_\nu$  in the sense that these vectors are the normal directions to the planes under consideration. If we denote by  $g_1(\nu), \dots, g_4(\nu)$  functions that for every normal direction  $\nu$  return the vertices of the corresponding plane (in a fixed but for our purposes unimportant order) then contributions  $\mathcal{P}_\nu$  to the estimator from the quadrilateral planes (faces and diagonal rectangles) can be written as:

$$\begin{aligned} \mathcal{P}_{\nu_{\text{quad}}} = & \\ & \frac{1}{n_0 \cdot a_\nu} \cdot \left[ \sum_{i=0}^{255} h_{S_\Gamma^W}(i) \cdot \mathbf{1}_{\{x_{g_1(\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_2(\nu)} \notin S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_3(\nu)} \notin S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_4(\nu)} \notin S_\Gamma^W\}} - \right. \\ & \left. \sum_{i=0}^{255} h_{S_\Gamma^W}(i) \cdot \mathbf{1}_{\{x_{g_1(\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_2(\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_3(\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_4(\nu)} \notin S_\Gamma^W\}} \right], \quad (2.31) \end{aligned}$$

where  $a_\nu$  denotes the area of the planes, i.e.  $h^2$  or  $\sqrt{2} \cdot h^2$  respectively. In principle on rectangular grid the Euler-characteristic  $\chi(S^W \cap E_{y,\omega})$  can be calculated directly using Corollary 2.1.21 and equation 2.25. But the rectangular grid leads to a slightly more complicated expression for the rectangular contributions than the above. It is easy to see that the rectangular contributions take the simple form of equation 2.31 if one adds one of the (facial or spatial) diagonals to the edge graph on which the two-dimensional Euler-characteristic of the quadrilateral prominent planes is estimated. Please note that an application of Corollary 2.1.21 is no longer possible because the corollary doesn't give us a hint how to calculate the Euler-characteristic  $\chi(S^W \cap E_{y,\omega})$  on a triangulated graph. But fortunately this calculation can be done thanks to Theorem 2.1.22. Analogously for the contributions from the triangular planes an application of Corollary 2.1.21 is not fruitful. Using Theorem 2.1.22 to calculate  $\chi(S^W \cap E_{y,\omega})$  it is straightforward to see that the contributions from the triangular planes are as follows:

$$\mathcal{P}_{\nu_{\text{tri}}} = \frac{1}{n_0 \cdot a_\nu} \cdot \left[ \sum_{i=0}^{255} h_{S_\Gamma^W}(i) \cdot \mathbf{1}_{\{x_{g_1(\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_2(\nu)} \notin S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_3(\nu)} \notin S_\Gamma^W\}} - \sum_{i=0}^{255} h_{S_\Gamma^W}(i) \cdot \mathbf{1}_{\{x_{g_1(-\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_2(-\nu)} \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_{g_3(-\nu)} \notin S_\Gamma^W\}} \right], \quad (2.32)$$

where  $-\nu$  denotes the direction which is opposite to  $\nu$  (this direction belongs to the second of the two triangles) and  $a_\nu$  is the sum of the two involved triangles' areas. Again one recognizes the lower dimensional Euler characteristic that is used to build up the contribution. It is important to note that because we have to resort to 2.1.22 to calculate both the contributions of the triangular planes and the simple form for the contributions of the quadrilateral planes given in equation 2.31, these contributions are not exactly the lower-dimensional Euler-characteristics. On the hexagonal graph we have to approximate the Euler characteristic of the set  $A := S^W \cap E_{y,\omega}$  by  $\chi(\tilde{A})$  where  $\tilde{A}$  is the graph representation of  $A$  introduced in subsection 2.1.5 above. Hence an additional approximation step is involved here. This additional error could be avoided for the quadrilateral contributions by replacing equation 2.31 by a slightly more complicated one but there is no easy way around it for the triangular contributions.

Performing the directional weighing according to the voronoi-weights  $c_\nu$ , the final estimator  $\hat{M}$  which is a discrete analogon of equation 2.30 is then:

$$\hat{M}(S_\Gamma^W) = 2\pi \sum_{\nu=0}^{25} c_\nu \cdot \mathcal{P}_\nu. \quad (2.33)$$

It thus remains to describe the estimator for the 3D-specific Euler-characteristic. With a threefold recursive application (once in each direction of the cartesian coordinate system) of equation 2.12 the identical calculation as in equation 2.25 shows:

$$\chi(S_\Gamma^W) = |\Gamma\text{-vertices in } S_\Gamma^W| - |\Gamma\text{-edges completely in } S_\Gamma^W| + |\Gamma\text{-faces completely in } S_\Gamma^W| - |\Gamma\text{-cells completely in } S_\Gamma^W|.$$

Again we have  $\chi(S_\Gamma^W) = \chi(S^W)$  only in the case where  $S^W$  is morphologically open and closed with respect to all edges, face diagonals and spatial diagonals of the unit cell. Therefore the estimator  $\hat{\chi}$  becomes simply:

$$\begin{aligned} \hat{\chi}(S_\Gamma^W) = & \frac{1}{n_0 \cdot h^3} \cdot \sum_{i=0}^{255} h_{S_\Gamma^W}(i) \cdot \left( \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} - \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_1 \in S_\Gamma^W\}} - \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_2 \in S_\Gamma^W\}} - \right. \\ & \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_4 \in S_\Gamma^W\}} + \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_1 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_2 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_3 \in S_\Gamma^W\}} + \\ & \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_2 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_4 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_6 \in S_\Gamma^W\}} + \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_1 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_4 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_5 \in S_\Gamma^W\}} - \\ & \left. \mathbf{1}_{\{x_0 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_1 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_2 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_3 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_4 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_5 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_6 \in S_\Gamma^W\}} \cdot \mathbf{1}_{\{x_7 \in S_\Gamma^W\}} \right) \end{aligned}$$

The estimator can be made more accurate by applying this formula from different directional perspectives (which is equivalent to selective permutation of the vertex indices in the above formula) and averaging over the directional contributions.

One can find other techniques for estimating the Minkowski-functionals in the literature. An especially appealing approach where all Minkowski-functionals are estimated simultaneously is given in (Schmidt and Spodarev 2005).

## 2.4 Random Sets and Particle Processes

In this chapter we want to introduce the stochastic framework we are using to model porous media. Because a specimen of a porous medium can be seen as a loosely connected aggregation of geometrical shapes placed into space with a certain degree of randomness, the concepts and techniques of stochastic geometry prove to be the right tools for our purposes. We mainly concentrate on the notion of a *random set* in the following but also briefly mention some basic concepts from *point process* theory because the Boolean model which we introduce in Chapter 3 can be interpreted as a so-called *particle process*. This is a point process whose "points" are random sets. Among the many books about point processes I like (Daley and Vere-Jones 2002) the best. An excellent text devoted to random sets is (Molchanov 2005), as a summary also (Stoyan et al. 1985) and (Schneider and Weil 2000), the last of which we take once more as a guideline also for this section.

### 2.4.1 Random Sets

We describe the theory for Euclidean space  $\mathbb{R}^d$  although the concept of random sets can also be defined on general LCHS-spaces, ie. locally compact, second-countable Hausdorff-spaces. We restrict our attention to random closed sets but one could develop a theory for random open sets as well. The family of all closed subsets of  $\mathbb{R}^d$  is denoted by  $\mathcal{F}$ . Furthermore  $\mathcal{C}$  refers to the collection of all compact subsets of  $\mathbb{R}^d$ , and the family of open subsets of  $\mathbb{R}^d$  is called  $\mathcal{G}$ . Usually  $\mathcal{F}$  is equipped with the so-called *Fell-topology* sometimes called *topology of closed convergence*. A subbasis for the Fell-topology is provided by:

$$\mathcal{F}_G = \{F \in \mathcal{F} | F \cap G \neq \emptyset\} \quad \forall G \in \mathcal{G} \quad \text{and} \quad \mathcal{F}^C = \{F \in \mathcal{F} | F \cap C = \emptyset\} \quad \forall C \in \mathcal{C},$$

where we have introduced the two pieces of common notation  $\mathcal{F}^A$  and  $\mathcal{F}_A$  for  $A \subset \mathbb{R}^d$  respectively. It can be shown (Schneider and Weil 2000, Thm 1.1.1.) that the collection  $\mathcal{F}$  equipped with the Fell-topology is a compact space with a countable basis and hence can be metrized according to the Urysohn-Lemma. Continuity for functionals on  $\mathcal{F}$  can therefore be tested by considering sequences. The Borel- $\sigma$ -algebra on  $\mathcal{F}$  with respect to the Fell-topology is denoted by  $\mathcal{B}(\mathcal{F})$ . It is not hard to see that  $\mathcal{B}(\mathcal{F})$  is generated by each of the collections  $\{\mathcal{F}_G : G \in \mathcal{G}\}$ ,  $\{\mathcal{F}^G : G \in \mathcal{G}\}$ ,  $\{\mathcal{F}_C : C \in \mathcal{C}\}$ ,  $\{\mathcal{F}^C : C \in \mathcal{C}\}$ . Furthermore

$\mathcal{K}, \mathcal{R}$  and  $\mathcal{S}$  are all elements of  $\mathcal{B}(\mathcal{F})$ . On  $\mathcal{C}$  besides the induced Fell topology another natural topology is often used, namely the one induced by the Hausdorff-metric. The two topologies are comparable, the Hausdorff-topology being strictly finer, thus having less converging sequences but more continuous functionals. We can now define the crucial object in this section:

**Definition 2.4.1. (random closed set)** Let  $(\Omega, \mathcal{A}, P)$  a probability space. A **random closed set**  $\Theta$  is a measurable mapping  $\Theta : (\Omega, \mathcal{A}, P) \rightarrow (\mathcal{F}, \mathcal{B}(\mathcal{F}))$ .

The trivial examples of random closed sets are the singleton  $\Theta = \{\xi\}$  and  $\Theta = (-\infty, \xi]$  where  $\xi$  is an  $\mathbb{R}$ -valued random variable. Much more important for us in the following is the fact that for a real-valued stochastic process  $(\xi_x)_{x \in \mathbb{R}^d}$  with a.s. continuous sample paths, the excursion set  $\Theta = \{x \in \mathbb{R}^d | \xi_x \geq u\}$  is a random closed set  $\forall u \in \mathbb{R}$ . This can trivially be seen as follows: Because second-countable spaces are separable for  $C \in \mathcal{C}$  we can find a dense countable subset  $(k_i)_{i \in \mathbb{N}}$  of  $C$  and note that because of continuity  $\{\Theta \in \mathcal{F}^C\} = \{\sup_{x \in C} \xi_x < u\} = \{\sup_{k_i} \xi_{k_i} < u\}$ . Finally use that the supremum of sequences of measurable functions are still measurable. A broad variety of random sets can be constructed by using simple geometrical set operations. The topology and  $\sigma$ -algebra on  $\mathcal{F}$  have been chosen such that all elementary set operations are measurable. Hence if  $\Theta_1$  and  $\Theta_2$  are random closed sets, then also  $\Theta_1 \cup \Theta_2$ ,  $\Theta_1 \cap \Theta_2$ ,  $\text{cl}(\Theta_1 + \Theta_2)$ ,  $\partial\Theta_1$ ,  $\text{cl}\Theta_1^c$ ,  $\text{cl conv}(\Theta_1)$ ,  $\alpha\Theta_1, \alpha \geq 0$ ,  $g\Theta_1, g \in G_d$  are still random closed sets (see Schneider and Weil 2000, Thm 1.3.4, p.23 for proofs). If we denote  $P_\Theta := \Theta \circ P$  the *distribution* of  $\Theta$  we can use familiar concepts as independence, stationarity and isotropy for random sets as well. The invariance properties of random sets are of special importance for us. We therefore recall them briefly:

**Definition 2.4.2. (Invariance properties of random sets)**

- i) A random set  $\Theta$  is called **stationary** if  $P_\Theta = P_{\Theta+x} \forall x \in \mathbb{R}^d$ .
- ii) A random set  $\Theta$  is called **isotropic** if  $P_\Theta = P_{R\Theta} \forall R \in SO_d$ .

In the following we usually assume the random sets  $\Theta$  we use to model porous media to be both stationary and isotropic. It is intuitively obvious that a nonempty stationary random closed set must be unbounded (see Schneider and Weil 2000, Th.1.3.5 for a rigorous proof). Hence calculating Minkowski-functionals for realizations of such sets is impossible. However we will only work with random sets whose realizations belong to  $\mathcal{S}$ . For such  $\Theta$  the concept of specific Minkowski functionals can be introduced. Because we can observe realizations of such a random set only through a compact and convex observation window  $W$ , Minkowski functionals  $V_j$  can be calculated for  $\Theta^W := \Theta(\omega) \cap W \in \mathcal{R}$ . Because we are looking for a quantity which characterizes the random set  $\Theta$  as a whole and not only single realizations, a reasonable quantity to look at is  $\mathbf{E}[V_j(\Theta^W)]$ , but with the drawback that because of the monotonicity of Minkowski functionals on  $\mathcal{K}$  (Theorem 2.1.3 the quantity  $V_j(\Theta^W(\omega))$  and thus also  $\mathbf{E}[V_j(\Theta^W)]$  depends heavily on the size of the observation window  $W$ . A natural choice to overcome this problem is considering the ratio  $\mathbf{E}[V_j(\Theta^W)]/V_d(W)$  as we already did for deterministic  $\mathcal{S}$ -sets in equation 2.20. However, in general this does not solve the problem since this ratio still depends on the geometry of the chosen observation window. For example this can be seen in the case where  $\Theta$  is a stationary and isotropic Boolean Model from equation 3.2 below. The remedy that works generally is to take an increasing sequence of observation windows  $r \cdot W$  ( $r > 0$  is a scaling factor) and consider the limit of the Minkowski quantities per unit volume for  $r \rightarrow \infty$ :

**Definition 2.4.3. (Specific Minkowski Functionals)** Let  $W$  be a compact and convex observation window and  $\Theta$  a random set with values in  $\mathcal{S}$ . Then the **specific Minkowski Functionals**  $\bar{V}_j(\Theta)$  are defined as

$$\bar{V}_j(\Theta) = \lim_{r \rightarrow \infty} \frac{\mathbf{E} [V_j(\Theta \cap rW)]}{V_d(\Theta \cap rW)}.$$

where  $r$  is a scaling factor and  $\Theta^K := \Theta \cap K$  for  $K \in \mathcal{K}$ .

The quantities  $\bar{V}_j$  are in fact independent of the sequence of windows chosen as can be seen below in Theorem 3.2.1 for the stationary and isotropic Boolean model. For general random  $\mathcal{S}$ -sets we can rely on the following theorem which shows that definition 2.4.3 is sensible even in a much wider context:

**Theorem 2.4.4. (Existence of Specific Limit and Independence on the Sequence of Observation Windows)** Let  $\Theta$  be a stationary random set with values in  $\mathcal{S}$  and  $\phi : \mathcal{R} \rightarrow \mathbb{R}$  additive, translation-invariant and conditionally bounded, ie.  $\phi$  is bounded on the sets  $C_K = \{K' | K' \subset K\}$  for  $K \in \mathcal{K}$ . Then  $\forall W \in \mathcal{K}$  with  $V_d(W) > 0$  the limit

$$\bar{\phi}(\Theta) = \lim_{r \rightarrow \infty} \frac{\mathbf{E} [\phi(\Theta \cap rW)]}{V_d(\Theta \cap rW)}.$$

exists and satisfies:

$$\bar{\phi}(\Theta) = \mathbf{E} \left[ \phi(\Theta \cap I_0^d) - \phi(\Theta \cap \partial^+ I_0^d) \right],$$

where  $I_0^d$  is the  $d$ -dimensional unit cube and  $\partial^+ I_0^d$  its upper left boundary.

**Proof. (Theorem 2.4.4)** see (Schneider and Weil 2000, Thm 5.2.4).

To estimate the specific Minkowski functionals  $\bar{V}_j(\Theta)$  for a single realization  $S := \Theta(\omega)$  of a stationary random set  $\Theta$ , all we can do is to use the estimator of equation 2.20.

Similarly as in ordinary probability theory where it is well-known that the characteristic function determines the distributional properties of a random variable uniquely, there is a functional which uniquely determines the distribution of a random set. We give the definition of this functional and the relevant theorem for  $\mathbb{R}^d$ , but an analogous result holds for general LCHS spaces:

**Definition 2.4.5. (Capacity functional)** Let  $\Theta$  be a random set. Then the functional  $T_\Theta : \mathcal{C} \rightarrow \mathbb{R}$  is defined as  $T_\Theta(C) = \mathbf{P}[\Theta \cap C \neq \emptyset]$ .  $T_\Theta$  is called the **Capacity-functional** of  $\Theta$ .

We conclude this subsection with a theorem which is at the heart of the theory of random sets:

**Theorem 2.4.6. ((Choquet-Theorem))** If two random sets  $\Theta_1, \Theta_2$  have the same capacity functional  $T_{\Theta_1} = T_{\Theta_2}$  then they are equal in distribution  $\Theta_1 \stackrel{d}{=} \Theta_2$ .

**Proof. (Theorem 2.4.6)** The theorem directly follows from the standard measure-theoretic extension theorem. See (Schneider and Weil 2000), section 2.2, for details.

### 2.4.2 Point and Particle Processes

A prominent subclass of random sets is given by the simple *Point Processes*. They are especially well studied and they have a well developed theory on their own and extensive literature devoted to them, eg. (Daley and Vere-Jones 2002) and (Schneider and Weil 2000, Chapter 3). Similarly as for general random sets we will only need a minimal extract from this theory which we develop in this subsection. We denote the set of all locally finite point patterns  $\eta$  in  $\mathbb{R}^d$  by  $N$ . Locally finite means that in every compact subset  $C \in \mathcal{C}$  the point pattern has only a finite number of points. Such a point pattern  $\eta$  can also be interpreted as a locally finite *counting measure* on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ , hence the condition of local finiteness can be written as  $\eta(C) < \infty, \forall C \in \mathcal{C}$ . A convenient  $\sigma$ -algebra on  $N$  which we denote by  $\mathcal{N}$  is the one generated by the *counting functions*  $\Psi_A : \eta \rightarrow \eta(A)$  which for every  $A \in \mathcal{B}(\mathbb{R}^d)$  give the number of points, the point pattern  $\eta$  contained in  $A$ .

**Definition 2.4.7. (Point Processes)** A *point process*  $X$  is a measurable mapping from a probability space  $(\Omega, \mathcal{A}, P)$  into  $(N, \mathcal{N})$ . The point process  $X$  is called **simple** if almost surely  $X$  takes values in the set  $\{\eta \mid \eta(\{x\}) \leq 1 \forall x \in \mathbb{R}^d\}$ .

We will concentrate on simple point processes in the following. The connection between point processes and random sets is the *support-function*  $\text{supp} : N \rightarrow \mathcal{F}$  which removes the (possible) multiplicity of points from the counting measures, hence the result is an ordinary closed subset of  $\mathbb{R}^d$ :

$$\text{supp}(\eta) = \{x \in \mathbb{R}^d \mid \eta(\{x\}) > 0\}.$$

The support function is obviously measurable since  $\text{supp}^{-1}(\mathcal{F}^C) = \{\eta \mid \Psi_C(\eta) = 0\}$ . Hence simple point processes are random sets in the sense of definition 2.4.1. It is therefore not surprising that the measurable spaces of simple locally finite counting measures and locally finite closed sets are indeed isomorphic (Schneider and Weil 2000, Theorem 3.1.2.), hence the results for random sets can be used for simple point processes and vice versa locally finite random sets can be interpreted as simple point processes. Therefore it is clear how the invariance properties given in definition 2.4.2 carry over to point processes. Important quantities used to characterize a point process are given by the (factorial and non-factorial) moment measures. Their definition relies on the fact that since a point process  $X$  can be thought of as a random locally finite counting measure it is perfectly fine to consider the  $m^{\text{th}}$  *product process*  $X^m := X \otimes \dots \otimes X$ , where the product is over  $m$  instances of  $X$ .

**Definition 2.4.8. (Moment measures)**

- Let  $X$  be a point process. Then its  $m^{\text{th}}$  **-moment measure**  $\nu : \mathcal{B}(\mathbb{R}^d)^m \rightarrow \mathbb{R}$  for a set  $A$  of the product Borel  $\sigma$ -algebra is defined as the expected number of points  $\nu^{(m)}(A) := \mathbf{E}[(X \otimes \dots \otimes X)(A)]$  the product point process has in  $A$ . Especially important is the case  $m = 1$ . The measure  $\nu^{(1)}$  is called the **intensity measure** of the point process  $X$ .
- Let  $\mathbb{R}_{\neq}^d := \{(x_1, \dots, x_d) \mid x_i \neq x_j \text{ for } i \neq j\}$ . Then for  $A_1 \times \dots \times A_m \in \mathcal{B}(\mathbb{R}^d)$  the  $m^{\text{th}}$  **-factorial moment measure**  $\chi^{(m)}$  of the point process  $X$  is defined as  $\chi^{(m)}(A_1 \times \dots \times A_m) = \mathbf{E}[(X \otimes \dots \otimes X)(A_1 \times \dots \times A_m \cap \mathbb{R}_{\neq}^d)]$ .

For a concise discussion of general moment measures, see for example (Stoyan et al. 1985, Chapter 4.3).



For the discussion of the Boolean Model in Chapter 3 we need to introduce an especially important class of point processes:

**Definition 2.4.9. (Poisson-Processes)** A simple point process  $X$  is called a **Poisson Process** if  $\forall A \in \mathcal{B}(\mathbb{R}^d)$  (with  $\nu(A) < \infty$ ) the random variable  $X(A)$  has a Poisson distribution, ie.  $X(A) \sim \text{Pois}(\nu(A))$ .

**Remark 2.4.10.** In literature often one finds the additional condition that for a sequence of pairwise disjoint Borel sets  $A_1, A_2, \dots$ , the random variables  $X(A_1), X(A_2), \dots$  must be independent if  $X$  is to be a Poisson-process. However for atom-free intensity measures this can be shown to follow from definition 2.4.9 (Schneider and Weil 2000, Thm. 3.2.3). Using the fact that simple point processes which fulfill the condition of definition 2.4.9 are *always* atom-free (Schneider and Weil 2000, Lemma 3.2.1), including the independence condition into the definition is superfluous.

It is easy to make plausible that Poisson-Processes exist. For an increasing sequence of compact sets  $C_i \uparrow \mathbb{R}^d$  consider the point processes given by  $X_i = \sum_{j=1}^{n_i} \delta_{\xi_j^i}$  where the  $\xi_j^i$  are independent and uniformly distributed on the  $C_i$ . Then the  $X_i(A)$  are obviously binomial random variables with success parameter  $p_i = \lambda(C_i \cap A)/\lambda(C_i)$  where  $\lambda$  is the Lebesgue measure. Letting  $n_i$  tend to infinity in such a way that the ratio  $n_i/\lambda(C_i)$  stays finite should lead to a limit process  $X$  which is Poisson. A theorem in the appendix (Theorem 7.2.4) reveals even that for every locally finite and atom-free measure  $\nu$  on  $\mathcal{B}(\mathbb{R}^d)$  an a.s. unique Poisson process exists. (The uniqueness is obvious from Theorem 2.4.6). If a point process  $X$  is stationary or isotropic respectively then its intensity measure is invariant under translations or rotations respectively. For Poisson processes due to uniqueness the converse is also true. The only translation invariant measure on  $\mathbb{R}^d$  is the Lebesgue measure  $\lambda$  up to a multiplicative constant, hence the intensity measure of a stationary Poisson process is  $\nu = \gamma \cdot \lambda$ . The constant  $\gamma$  is referred to as *intensity* of the stationary Poisson process. From the uniqueness remark above it is obvious that for every  $\lambda \in [0, \infty)$  an a.s. unique stationary Poisson process exists. Every stationary Poisson process in  $\mathbb{R}^d$  is also isotropic (Schneider and Weil 2000, Theorem 3.3.1). Another important fact which makes Poisson processes special is that all factorial moment measures can be expressed through the intensity measure. Because we will need this fact in Chapter 3 we formulate it as a theorem:

**Theorem 2.4.11.** Let  $X$  be a Poisson process whose intensity measure is denoted by  $\nu$ . Then for the factorial moment measures  $\chi^{(m)}$  we have that

$$\chi^{(m)} = \nu^m. \quad (2.34)$$

**Proof. (Theorem 2.4.11)** The proof is pure measure theory and can be found in (Schneider and Weil 2000, Theorem 3.2.3., part c))

Please be aware here of the notational pitfall  $\nu^m \neq \nu^{(m)}$  which basically means that on the right hand side of equation 2.34 we *don't* have the  $m^{\text{th}}$ -non-factorial moment measure.

In point process theory one is often interested in computing the expected value of the sum of a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  evaluated in the (random) points of a Poisson-Process  $X$ . This can be accomplished using the intensity measure  $\nu$ :

**Theorem 2.4.12. (Campbell-Theorem)** Let  $X$  be a Poisson-Process in  $\mathbb{R}^d$  with intensity measure  $\nu$  and  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  non-negative and measurable. Then:

$$\mathbf{E} \left[ \sum_{x \in \mathbb{R}^d} X(\{x\}) f(x) \right] = \int_{\mathbb{R}^d} f \, d\nu.$$

**Proof.** (Theorem 2.4.12) follows from measure-theoretic induction since for  $f = \mathbf{1}_A$  with  $A \in \mathbf{B}(\mathbb{R}^d)$  we have

$$\mathbf{E} \left[ \sum_{x \in \mathbb{R}^d} X(\{x\}) \mathbf{1}_A \right] = \mathbf{E}[X(A)] = \nu(A) = \int_{\mathbb{R}^d} \mathbf{1}_A \, d\nu$$

□

Point processes can be defined on general LCHS-spaces. Note that the proof of Theorem 2.4.12 doesn't use the special structure of the Euclidean space and therefore holds generally. Of particular interest for us is the space  $\mathcal{F}' := \mathcal{F} \setminus \{\emptyset\}$  equipped with the Fell-topology (see subsection 2.4.1). This space is indeed a locally compact Hausdorff space with countable basis (Schneider and Weil 2000, Thm. 1.1.1.), hence we can define point processes on it. The "points" are then closed subsets of  $\mathbb{R}^d$ . The measurable space  $(N_{\mathcal{F}}, \mathcal{N}_{\mathcal{F}})$  of locally finite counting measures on  $\mathcal{F}$  can be constructed analogously as for  $\mathbb{R}^d$ . Because we concentrate on compact grains in the following, we can restrict ourselves to considering point processes on  $\mathcal{C}' := \mathcal{C} \setminus \{\emptyset\}$  with the induced topology. Such point processes are called *particle processes*:

**Definition 2.4.13. (Particle Processes)** A measurable mapping  $X$  from a probability space  $(\Omega, \mathcal{A}, P)$  into  $(N_{\mathcal{F}}, \mathcal{N}_{\mathcal{F}})$  is called a **Particle Process** if its intensity measure  $\nu$  is concentrated on  $\mathcal{C}'$ .

It is obvious that the union sets  $\Theta$  of the particles in a Particle Process  $X$  are random closed sets in the sense of definition 2.4.1 since  $\{\Theta \in \mathcal{F}^C\} = \{X(\mathcal{F}_C) = 0\} \forall C \in \mathcal{C}$ . The standard example for a Particle Process  $X$  arises from attaching iid. compact random sets  $\Theta_i$  in the points of an ordinary point process  $X = \{x_1, x_2, \dots\}$  leading to the process  $X = \{x_i + \Theta_i\}$ . Of course one must ensure local finiteness which can be done by imposing technical conditions. This simple example is of special interest for us since it leads to the so-called *germ-grain-models* for porous media.

**Definition 2.4.14. (germ-grain-models)** Let  $X = \{x_1, x_2, \dots\}$  be an (ordinary) point process in  $\mathbb{R}^d$  and  $\Theta_i$  iid. compact random sets such that  $\Phi = \{x_1 + \Theta_1, x_2 + \Theta_2, \dots\}$  is a particle process. Then the random closed union set  $\Theta$  of  $\Phi$

$$\Theta = \bigcup_{i=1}^{\infty} (x_i + \Theta_i)$$

is called a *germ-grain-model*.

We conclude this section with a decomposition result for the intensity measure of *stationary* Particle Processes which we will need later on. Since particles  $K$  are compact we can always define their center  $m(K)$  (eg. as the midpoint of the circumcircle). The geometrical information of the particle can then be separated into "location"  $m(K)$  and "shape"  $K - m(K)$  where the shapes live on the *shape space*  $\mathcal{C}_0 = \{K \in \mathcal{C}' \mid m(K) = 0\}$ . According to this separation also the intensity measure  $\nu$  of a stationary particle process can be split up:

**Theorem 2.4.15. (Location-Shape-Decomposition of the intensity measure)** Let  $X \neq 0$  be a stationary Particle Process, then

$$\nu(A) = \gamma \int_{\mathcal{C}_0} \int_{\mathbb{R}^d} \mathbf{1}_A(x + K) \lambda_d(dx) Q(dK), \quad \forall A \in \mathcal{B}(\mathcal{C}')$$

where  $\gamma > 0$  and  $Q$  is a probability measure on  $\mathcal{C}_0$ . If  $X$  is isotropic, then  $Q$  is invariant under rotations.

**Proof. (Theorem 2.4.15)** The image  $\tilde{\nu}$  of  $\nu$  under the mapping  $m : F \rightarrow (\mathbb{R}^d, \mathcal{C}_0) K \rightarrow (m(K), K - m(K))$  is a measure which is translation invariant in the first component and hence  $\tilde{\nu}(A \times B) = c(B) \cdot \lambda_d(A)$  for  $A \in \mathcal{B}(\mathbb{R}^d), B \in \mathcal{B}(\mathcal{C}_0)$ . Now, just define  $\gamma = c(\mathcal{C}_0)$  (which is finite as long as  $\nu$  is locally finite) and  $Q(B) = c(B) \cdot \gamma^{-1}$ . The isotropy statement is obvious.  $\square$

## 2.5 Other Geometrical Characteristics

Besides Minkowski-functionals there are many other characteristics in use when it comes to modelling porous media. We just mention some of the most important ones here, namely:

- the *m-point-covariance functions*,
- the *chord-length-distribution*,
- the *pore-size-distribution*,
- the *Minkowski-functions*.

A more detailed discussion of these and more microstructural descriptors can be found in (Torquato 2002, in particular Chapter 2). We start out with the *m-point-covariance functions*.

### 2.5.1 The m-Point-Covariance Functions

**Definition 2.5.1. (m-point-covariance-functions)** Let  $\Theta$  be a random set in  $\mathbb{R}^d$ . Then the *m-point-covariance-function*  $C_\Theta^{(m)}$  for  $\Theta$  is defined as the function

$$C_\Theta^{(m)} : \quad \mathbb{R}^d \oplus \dots \oplus \mathbb{R}^d \rightarrow \mathbb{R}$$

$$(x_1, \dots, x_m) \rightarrow \mathbf{E} \left[ \prod_{i=1}^m \mathbf{1}_{\{x_i \in \Theta\}} \right] = \mathbf{P} \left[ \bigcap_{i=1}^m \{x_i \in \Theta\} \right],$$

ie.  $C_\Theta^{(m)}$  is the probability that the  $m$  points  $x_1, \dots, x_m$  belong to  $\Theta$ .

Of special importance is the *two-point-covariance function*  $C_\Theta := C_\Theta^{(2)}$ . Is it closely related to the Minkowski-functionals. The following relationships are obvious for stationary  $\Theta$  without long range dependence from definition 2.5.1:

$$\lim_{\|x_1, x_2\| \rightarrow 0} C_\Theta(x_1, x_2) = \bar{v}(\Theta), \quad \lim_{\|x_1, x_2\| \rightarrow \infty} C_\Theta(x_1, x_2) = \bar{v}^2(\Theta),$$

however at first glance it is surprising that the derivative of  $C_\Theta$  at the origin completely determines the specific surface area  $\bar{s}(\Theta)$  for stationary and isotropic  $\Theta$  which almost surely coincide with the closure of their interior:

**Theorem 2.5.2. (Relationship between two-point-covariance and specific surface area)** Let  $\Theta$  be a random set in  $\mathbb{R}^d$  with realizations in  $\mathcal{S}$  which is stationary and isotropic and fulfills  $\Theta = \text{cl}(\text{int}(\Theta))$  and let  $C_\Theta(r) := C_\Theta(\|x_1 - x_2\|)$  be its two-point-covariance function. Then for the specific surface area  $\bar{s}(\Theta)$  the following is true:

$$\bar{s}(\Theta) = -\frac{d\kappa_d}{\kappa_{d-1}} \frac{\partial C_\Theta}{\partial r}(0).$$

Theorem 2.5.2 provides an alternative way to estimate the specific surface area on a pixel grid because in the stationary case estimation of the two-point correlation function is straightforward by parallel shifting of a rod across the image and counting the relative frequency of the event that both endpoints of the rod belong to  $\Theta$ . However it remains difficult to discretely approximate the (onesided) derivative in zero. Theorem 2.5.2 can also be used as a starting point if one tries to analytically express the specific surface area by the underlying parameters of a specific model. Usually for two neighboring lattice points  $x_1$  and  $x_2$ , the expression  $\mathbf{P}[x_1 \in \Theta, x_2 \in \Theta]$  can be analytically calculated and the strategy is to investigate what happens if the lattice spacing tends to 0 (see section 4.3 for such an approach in the case of the Thresholded Gaussian Field Model).

**Proof. (Theorem 2.5.2)** We choose the observation window  $W$  to be the  $d$ -dimensional ball with radius  $r$ , ie.  $W_r = r \cdot B_d(0)$ . Then an application of the Crofton-formula 2.2.3 for  $q = d - 1$  and the fact that  $s = 2 \cdot V_{d-1}$  leads to:

$$\begin{aligned} \bar{s}(\Theta) &= \lim_{r \rightarrow \infty} \frac{2 \cdot c_{d-1}^d}{\kappa_d r^d} \cdot \mathbf{E} \left[ \int_{SO_3} \int_{\mathbb{R}^{d-1}} \chi(W_r \cap \Theta \cap e_{y,\omega}) \mu_1(de) \right] \\ &= \lim_{r \rightarrow \infty} \frac{2 \cdot c_{d-1}^d}{\kappa_d r^d} \cdot \int_{\mathbb{R}^{d-1}} \mathbf{E}[\chi(W_r \cap \Theta \cap e_{y,\omega})] dy \\ &= \lim_{r \rightarrow \infty} \frac{2 \cdot c_{d-1}^d}{\kappa_d r^d} \cdot \int_{\mathbb{R}^{d-1}} \mathbf{E} \left[ \lim_{h \downarrow 0} \sum_{\substack{x_i \in e_{y,\omega} \\ |x_{i+1} - x_i| = h}} \mathbf{1}_{\{x_i \in W_r \cap \Theta, x_{i+1} \in W_r \cap \Theta^c\}} \right] dy \\ &= \lim_{r \rightarrow \infty} \frac{4 \cdot c_{d-1}^d \cdot (d-2)\kappa_{d-2}}{\kappa_d r^d} \cdot \int_0^r x^{d-2} \sqrt{r^2 - x^2} \lim_{h \downarrow 0} \frac{\mathbf{P}[x_i \in \Theta, x_{i+1} \in \Theta^c]}{h} dx \\ &= (-C'_\Theta(0)) \cdot \frac{4 \cdot c_{d-1}^d \cdot (d-2)\kappa_{d-2}}{\kappa_d} \int_0^1 x^{d-2} \sqrt{1 - x^2} dx, \end{aligned}$$

where we have used in this order isotropy, the Hadwiger-definition of the Euler-characteristic, the formula  $d \cdot \kappa_d$  for the surface of the  $d$ -dimensional unit sphere (which has to be replaced by the respective values for  $d = 0, 1$ ) and finally the substitution  $x' = x/r$  which shows that the final expression doesn't depend on the ball radius  $r$  anymore and thus the limit  $r \rightarrow 0$  becomes trivial. Solving the final Beta-Integral and using a well-known identity for Gamma-functions gives the desired result. An alternative heuristic is given in the appendix, subsection 7.2.5.  $\square$

## 2.5.2 The Chord-Length-Distribution

Assume that  $\Theta$  is both stationary and isotropic. Then a measure for its connectivity along linear paths is given by the so-called *chord-length-distribution*. Its name is misleading since it is a probability *density* function and not a distribution function. That's why some

authors prefer to call it *chord-length-density* instead. Here the notion of *chords* is used for the (finite) closed segments between two subsequent intersection points of the interface  $\partial\Theta$  with a straight line. One can consider the chords lying completely in  $\Theta$  or those lying completely in  $cl(\Theta^c)$  instead.

**Definition 2.5.3. (Chord-Length-Distribution-Function)** For  $\Theta$  stationary and isotropic and  $r > 0$  its **chord-length-distribution**  $c_\Theta(r)$  is the probability density with the property that  $c_\Theta(r) dr$  is the probability that the length of a chord of  $\Theta$  falls into  $[r, r + dr]$ .

The chord-length-distribution function is of great importance for investigating fluid-flow properties in porous media (see Lehmann et al. n.d. and references therein). It is closely related to other structural quantities such as the *lineal-path-function*  $L(r)$  which measures the probability that a segment of *given* length  $r$  thrown into space at random lies completely in  $\Theta$  (Torquato 2002, p.44ff.). For the lineal path function we obviously have  $L(0) = \bar{v}(\Theta)$  and  $L(\infty) = 0$ . A close relationship also exists to the *contact distribution function*  $H_K(r)$  which is defined as  $H_K(r) := P[0 \in \Theta \oplus r(-K) | 0 \notin \Theta]$  where  $K$  is a compact structuring element and  $0$  denotes the origin. It is easy to see that in the case where the structuring element  $K$  is a line segment of unit length, then  $H_K(r)$  is the distribution function of chord lengths of  $cl(\Theta^c)$  for chords which have one endpoint in  $0$ .

### 2.5.3 The Pore-Size-Distribution

The *pore-size-distribution* can be introduced for arbitrary random sets  $\Theta$  and their complements  $cl(\Theta^c)$ . It is a density function, too, and can be defined as follows:

**Definition 2.5.4. (pore-size-distribution)** Let  $\Theta$  be a random set and  $r > 0$ . Then its **pore size distribution**  $p_\Theta(r)$  is the probability density function such that  $p_\Theta(r) dr$  is the probability that the distance from an arbitrarily chosen point  $\in \Theta$  to the interface  $\partial\Theta$  falls into  $[r, r + dr]$ .

Basic properties of the pore size distribution are  $p_\Theta(0) = \frac{\bar{s}(\Theta)}{\bar{v}(\Theta)}$  and  $p_\Theta(\infty) = 0$ . Also the pore-size-distribution is widely used to characterize flow properties in porous media (Lehmann et al. n.d.). In real data samples the pore size distribution is usually measured by mercury porosimetry, however, the mathematical pore size distribution and the experimentally measured function in general are not the same.

### 2.5.4 The Minkowski-Functions

Instead of just considering the Minkowski-functionals  $V_j(\cdot)$  of a (deterministic) set  $R \in \mathcal{R}$  itself, one can also study its Minkowski functionals after it has undergone one of the morphological transformations given in definition 2.1.19. The most important and most frequently applied such transformation in literature is the dilation with a spherical structuring element  $B$  of radius  $r$ . This gives rise to the so-called *Minkowski-functions*:

**Definition 2.5.5.** The  $d+1$  **Minkowski-functions**  $V_j^* : \mathcal{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ ,  $j \in \{0, \dots, d\}$  for a set  $R \in \mathcal{R}$  are defined as the ordinary Minkowski-functionals  $V_j(\cdot)$  of the sets obtained by dilating  $R$  with a spherical structuring element  $B(r)$  when understood as functions of the dilation radius  $r$ , i.e.:

$$V_j^*(R, r) := V_j(R \oplus B(r)).$$

Instead of just considering the  $d+1$  single *numbers*  $V_j(R)$  to summarize an image (with solid phase  $R$ , say), one can gain a lot more information about the image when characterizing it by the  $d+1$  *functions*  $V_j^*(R, \cdot)$ . It will turn out in later chapters that Minkowski-functionals  $V_j(\cdot)$  on their own are by no means sufficient to uniquely summarize the information-content of a porous structure. If one's goal is to construct artificial images with the same optical appearance as an original natural structure, one should obtain much better results when matching the whole Minkowski-function-paths and thus exploiting the information provided by the dilated images in the analysis. For a random set  $\Theta$  one would like the expected specific Minkowski-function-path  $\mathbf{E} [V_j^*(R, r)]$  to agree with the function path of the structure to be modelled.

## Chapter 3

# The Boolean Model

The *Boolean Model* is certainly one of the simplest probability models one can think of to model irregular two-phase-structures. Thanks to its simplicity and analytical tractability it has been used to describe random patterns in nearly all branches of natural sciences but most prominently in applications of materials science. First scientific papers where it was used date back to the thirties of the last century and even beyond. The main reason why the Boolean Model is attractive for our purposes is that the theory of Minkowski functionals for Boolean structures is well understood and there are explicit formulae expressing how the specific Minkowski functionals depend on the parameters of the Boolean Model. (see section 3.2 for a derivation of these formulae.) Good surveys about the Boolean Model can be found for example in (Schneider and Weil 2000) for a profound presentation of its mathematical theory and (Stoyan et al. 1985) for the more practical aspects.

### 3.1 Definition and Relevant Properties

The *Boolean Model* is a *germ-grain-model* in the sense of definition 2.4.14. As the name already suggests, these models consist of two basic ingredients: The *germs* which basically determine the *location* in the surrounding space of the random objects making up the realizations of the Boolean Model and the *grains* which determine the *shape* of these random objects. In the case of the stationary Boolean Model the germs are distributed uniformly and independently across space. Mathematically this is realized by the stationary Poisson Process. Once the germs are determined, to each of the germs a grain of random shape is attached independently of the others. As far as these grains are concerned the general Boolean Model is very flexible allowing for arbitrary random compact sets imposing only slight regularity conditions. However for our purposes it is enough to consider *compact and convex* random closed grains which usually belong all to the same class of geometrical objects (eg. spheres, ellipsoids, triangles or lines). The union of all germ-grain pairs constructed in this manner constitute a realization of the Boolean Model. Let us now give a first formal definition and a few examples:

**Definition 3.1.1.** (*stationary Boolean Model in  $\mathbb{R}^d$  with compact grains*) Let  $X = \{x_1, x_2, \dots\}$  be a stationary Poisson Process in  $\mathbb{R}^d$  and  $\Theta_1, \Theta_2, \dots$  a sequence of iid. random compact sets in  $\mathbb{R}^d$ . Then the union  $\Theta$  of all germ-grain pairs

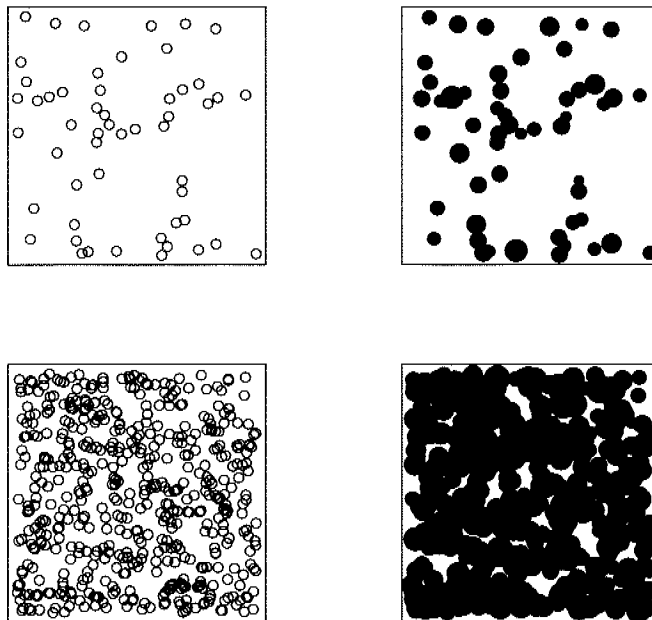
$$\Theta = \bigcup_{i=1}^{\infty} (x_i + \Theta_i)$$

is called a **Boolean Model with compact grains**.  $\Theta_0 \stackrel{d}{=} \Theta_i$  is called the **typical grain** and its distribution is referred to as the **grain distribution**  $Q$  of the Boolean Model.

Usually one needs some moderate technical conditions to ensure that the Boolean Model as defined in 3.1.1 is a random closed set in the sense of definition 2.4.1. For example one could use  $\mathbf{E}[V_d(\Theta_0 \oplus C)] < \infty$ ,  $\forall C \in \mathcal{C}$ .

This ensures that a.s. the realizations of the Boolean Model  $\Theta$  are closed subsets of  $\mathbb{R}^d$ . If one considers a sequence of compact subsets increasing to the whole space  $\mathbb{R}^d$  (eg. balls around the origin of integer radius) then the union of the (a.s. finitely many) grains hitting such a compact subset is a random closed set in the sense of definition 2.4.1. Therefore a realization of the Boolean Model can be seen (almost surely) as the set limit of an increasing sequence of random closed sets and is therefore itself a random closed set in the sense of definition 2.4.1 as long as the technical condition is met. For similar conditions see (Stoyan et al. 1985). If the grains  $\Theta_i$  are convex, then the realizations of the Boolean Model are in  $\mathcal{S}$  which is what we need in the following.

### Stationary Poisson Process and Boolean Model



**Figure 3.1:** (upper left) stationary Poisson Process in  $[0, 1]^2$  with low intensity  $\gamma = 30$ . (upper right) Boolean Model with spherical typical grain and Gamma-distributed radius. The realization of the upper left image was used as the grain process. The individual grains usually have physical meaning, eg. represent physical entities. (lower left) Stationary Poisson Process in  $[0, 1]^2$  with high intensity  $\gamma = 500$ . (lower right) Boolean Model with spherical typical grain and Gamma-distributed radius. The realization of the lower left image was used as the grain process. The individual grains have no longer physical meaning. Due to overlapping they simply contribute to the macrostructure.



Because of the high degree of flexibility for the choice of the grains, the Boolean Model can reproduce a wide variety of structures. The degree of sparseness (or density of space covered by the realizations) of the Boolean Model can easily be controlled by the intensity of the underlying stationary Poisson Process and the range of the grain distribution. The overlapping of the grains can be seen as a defect of the Boolean Model when modelling porous media. However this problem is no longer present if one does not interpret the Boolean grains as physical grains but rather does understand them as structuring elements which contribute to the macrostructure.

Stationarity of the Boolean Model is inherited from the stationarity of the underlying Poisson Process. Because every stationary Poisson Process in Euclidean space must also be isotropic, the Boolean Model is also isotropic as long as its grain distribution has that property which we will always assume in the following.

For theoretical considerations as in section 3.2 it is convenient to give an alternative (less intuitive) definition of the stationary Boolean Model as union set (or support)  $\Theta$  of a stationary Poisson particle process. It can be shown (merely as a consequence of Theorem 2.4.15) that we can always find an underlying ordinary stationary Poisson Process  $X$  (The germ process of grain centers) and a grain distribution  $Q$  such that  $\Theta$  is also a Boolean Model in the sense of definition 3.1.1. This new point of view has the advantage that we can use the whole arsenal of methods and theorems for point processes in general spaces. Together with the integral geometry developed in section 2.2 this will allow us to give explicit expressions for the specific Minkowski functionals in the Boolean Model in section 3.2.

## 3.2 Miles-Formulae

### 3.2.1 Derivation and Interpretation

Because the grains of the Boolean Model  $\Theta$  may intersect arbitrarily one might think that calculating specific Minkowski-functionals as a function of the basic parameters  $\gamma$  and  $Q$  is impossibly difficult. However this is not the case if the model is isotropic and the grains are restricted to be convex. Formulae which express  $\bar{V}_j(\Theta)$ ,  $j = 0, \dots, d$  through the intensity  $\gamma$  of the underlying stationary Poisson Process and mean values of Minkowski functionals  $\mathbf{E}[V_j(\Theta_0)]$ ,  $j = 1, \dots, d$  for the typical grain  $\Theta_0$  can be derived rather directly from the iterated kinematic formula 2.2.4. We will refer to these formulae as *Miles-formulae* but an early and complete derivation can also be found in (Mecke and Wagner 1991, Mecke 2001). The formulae date back into the seventies of the last century and are already contained in (Miles 1976), (Davy 1976) and (Davy 1978). Because the derivation of the Miles formulae is both short and elegant within the particle process framework we present it here following (Schneider and Weil 2000, Chapter 5, in particular Corollary 5.4.5.). One can even drop the isotropy property and still find appealing but more difficult formulae (Schneider and Weil 2000, Th 5.5.3.).

For the derivation of the Miles formulae let us now assume that  $\Theta$  is a stationary and isotropic Boolean Model with convex grains  $\in \mathcal{K}' = \mathcal{K} \cap \mathcal{C}'$ , intensity  $\gamma$  and grain distribution  $Q$  where  $\mathcal{C}' := \mathcal{C} \setminus \{\emptyset\}$ . The Boolean Model  $\Theta$  is the union set of a stationary Poisson Particle Process  $X$ :

$$\Theta = \bigcup_{K \in X} K.$$

Furthermore  $\Theta$  is observed through a compact and convex observation window  $W \in \mathcal{K}$ . Starting out with the inclusion-exclusion formula 2.1.7 we can write

$$\begin{aligned} V_j(\Theta \cap W) &= \sum_{m=1}^n (-1)^{m+1} \sum_{1 \leq i_1 < \dots < i_m \leq n} V_j(W \cap K_{i_1} \cap \dots \cap K_{i_m}) \\ &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \sum_{(K_1, \dots, K_m) \in X_{\neq}^m} V_j(W \cap K_1 \cap \dots \cap K_m), \end{aligned}$$

where  $n$  is of course random but can be replaced by  $\infty$  because only finitely many grains  $K$  intersect the compact observation window  $W$  (local-finiteness of the Particle Process) and where the notation  $X_{\neq}^m$  stands for the process

$$X_{\neq}^m(A_1 \times \dots \times A_m) := (X \otimes \dots \otimes X)(A_1 \times \dots \times A_m \cap \mathcal{F}_{\neq}^d), \quad A_i \in \mathcal{F}.$$

whose intensity measure is the factorial moment measure  $\chi^{(m)}$  of  $X$ . The inner summation in the last equation sums function values of a by Theorem 2.1.3, part vii) nonnegative function evaluated in the points of a general point process (namely  $X_{\neq}^m$ ). Hence for the expectation the general LCHS-space version of Theorem 2.4.12 can be applied to achieve:

$$\begin{aligned} \mathbf{E}[V_j(\Theta \cap W)] &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \mathbf{E} \left[ \sum_{(K_1, \dots, K_m) \in X_{\neq}^m} V_j(W \cap K_{i_1} \cap \dots \cap K_{i_m}) \right] \\ &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \int_{\mathcal{K}'} \dots \int_{\mathcal{K}'} V_j(W \cap K_1 \cap \dots \cap K_m) d\chi^{(m)}(K_1, \dots, K_m) \\ &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \int_{\mathcal{K}'} \dots \int_{\mathcal{K}'} V_j(W \cap K_1 \cap \dots \cap K_m) \nu(dK_1) \dots \nu(dK_m), \end{aligned}$$

where the last step follows from the representation of the factorial moment measure for Poisson Processes, Theorem 2.4.11, and the exchange of summations can be justified. Because of stationarity we can now use the decomposition of the intensity measure of the Particle Process  $\Theta$  from Theorem 2.4.15 to arrive at:

$$\mathbf{E}[V_j(\Theta \cap W)] = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \gamma^m \int_{\mathcal{K}'} \dots \int_{\mathcal{K}'} \Phi_j(K_1, \dots, K_m) Q(dK_1) \dots Q(dK_m), \quad (3.1)$$

where

$$\Phi_j(K_1, \dots, K_m) := \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} V_j(W \cap (K_1 + x_1) \cap \dots \cap (K_m + x_m)) \lambda_d(dx_1) \dots \lambda_d(dx_m)$$

is used as a shorthand-notation. Because we also assume isotropy these last Lebesgue-integrals can be replaced by integrals with respect to the Haar measure  $\mu$  on the group of rigid motions  $g \in G_d$  because averaging with respect to the Haar measure on  $SO_d$  is

redundant by isotropy. Hence the iterated kinematic formula 2.2.4 can be applied to the  $\Phi_j(K_1, \dots, K_m)$ :

$$\begin{aligned}\Phi_j(K_1, \dots, K_m) &= \int_{G^d} \dots \int_{G^d} V_j(W \cap g_1 K_1 \cap \dots \cap g_m K_m) \mu(dg_1) \dots \mu(dg_m) \\ &= \sum_{\substack{k_0, \dots, k_m=j \\ k_0 + \dots + k_m = md+j}}^d \frac{k_0! \kappa_{k_0} \dots k_m! \kappa_{k_m}}{j! \kappa_j (d! \kappa_d)^m} V_{k_0}(W) V_{k_1}(K_1) \dots V_{k_m}(K_m).\end{aligned}$$

This can now be plugged back into equation 3.1 and the integrations with respect to the grain distribution can be performed:

$$\mathbf{E}[V_j(\Theta \cap W)] = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \sum_{\substack{k_0, \dots, k_m=j \\ k_0 + \dots + k_m = md+j}}^d \frac{k_0! \kappa_{k_0} \dots k_m! \kappa_{k_m}}{j! \kappa_j (d! \kappa_d)^m} V_{k_0}(W) \bar{V}_{k_1}(\Theta) \dots \bar{V}_{k_m}(\Theta),$$

where we used the shorthand  $\bar{V}_i(\Theta) := \gamma \cdot \mathbf{E}[V_j(\Theta_0)]$ . This already establishes the important relationship between mean Minkowski quantities for the whole model  $\Theta$  and the basic "grain" quantities for  $\Theta_0$ . What now follows is pure algebra such that the defining limit for the specific Minkowski functionals  $\bar{V}_j(\Theta)$  (see definition 2.4.3) can be read off:

$$\begin{aligned}\mathbf{E}[V_j(\Theta \cap W)] &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \sum_{k_0=j}^d \frac{k_0! \kappa_{k_0}}{j! \kappa_j} V_{k_0}(W) \sum_{\substack{k_1, \dots, k_m=j \\ k_1 + \dots + k_m = md+j-k_0}}^d \prod_{i=1}^m \frac{k_i! \kappa_{k_i}}{d! \kappa_d} \bar{V}_{k_i}(\Theta) \\ &= V_j(W) \left(1 - e^{-\bar{V}_d(\Theta)}\right) + A,\end{aligned}\tag{3.2}$$

where

$$A = \sum_{k_0=j+1}^d \frac{k_0! \kappa_{k_0}}{j! \kappa_j} V_{k_0}(W) \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m!} \sum_{\substack{k_1, \dots, k_m=j \\ k_1 + \dots + k_m = md+j-k_0}}^d \prod_{i=1}^m \frac{k_i! \kappa_{k_i}}{d! \kappa_d} \bar{V}_{k_i}(\Theta)$$

The conditions in the subarray of the sum in  $A$  reveal that for given  $k_0$  at most  $k_0-j$  of the summands  $k_i$  can be different from  $d$  and thus  $A$  is equal to:

$$\begin{aligned}&\sum_{k_0=j+1}^d \frac{k_0! \kappa_{k_0}}{j! \kappa_j} V_{k_0}(W) \sum_{s=1}^{k_0-j} \sum_{r=0}^{\infty} \binom{r+s}{r} \frac{(-1)^{r+s+1}}{(r+s)!} \bar{V}_d(\Theta)^r \sum_{\substack{k_1, \dots, k_s=j \\ k_1 + \dots + k_s = sd+j-k_0}}^{d-1} \prod_{i=1}^s \frac{k_i! \kappa_{k_i}}{d! \kappa_d} \bar{V}_{k_i}(\Theta) \\ &= -e^{-\bar{V}_d(\Theta)} \sum_{k_0=j+1}^d \frac{k_0! \kappa_{k_0}}{j! \kappa_j} V_{k_0}(W) \sum_{s=1}^{k_0-j} \frac{(-1)^s}{s!} \sum_{\substack{k_1, \dots, k_s=j \\ k_1 + \dots + k_s = sd+j-k_0}}^{d-1} \prod_{i=1}^s \frac{k_i! \kappa_{k_i}}{d! \kappa_d} \bar{V}_{k_i}(\Theta).\end{aligned}$$

If we now divide the resulting expression for  $\mathbf{E}[V_j(\Theta \cap W)]$  by  $V_d(W)$  we recognize that this quotient still depends on the geometrical characteristics of the observation window  $W$  in which the Minkowski characteristics are calculated. This illustrates the need for defining the specific Minkowski-functionals as we did in definition 2.4.3. This dependence

on the observation window vanishes if we replace  $W$  by  $r \cdot W$  for a scaling factor  $r > 0$  and consider the defining limits. Note that for the specific volume  $\bar{v}$  the quantity  $A$  in the above derivation is not present and thus we simply get:

$$\bar{v}(\Theta) = \lim_{r \rightarrow \infty} \frac{\mathbf{E}[V_d(\Theta \cap W)]}{V_d(W)} = 1 - e^{-\bar{V}_d(\Theta_0)}.$$

For  $j < d$  only the term for  $k_0 = d$  leads to a non zero limit, if we additionally separate the summand for  $s = 1$  in this term, we conclude:

$$\bar{V}_j(\Theta) = \lim_{r \rightarrow \infty} \frac{\mathbf{E}[V_j(\Theta \cap W)]}{V_d(W)} = e^{-\bar{V}_d(\Theta_0)} \left( \bar{V}_j(\Theta_0) - \frac{d! \kappa_d}{j! \kappa_j} \sum_{s=2}^{d-j} \frac{(-1)^s}{s!} \sum_{\substack{k_1, \dots, k_s = j+1 \\ k_1 + \dots + k_s = (s-1)d+j}}^{d-1} \prod_{i=1}^s \frac{k_i! \kappa_{k_i}}{d! \kappa_d} \bar{V}_{k_i}(\Theta_0) \right),$$

where we may start indexing the  $k_i$  at  $j+1$  instead of  $j$ . We collect these results in the following theorem:

**Theorem 3.2.1. (General Miles-Formulae)** *Let  $\Theta$  be a stationary and isotropic Boolean Model with convex grains and denote by  $\gamma$  the intensity of the underlying stationary Poisson Process and by  $\Theta_0$  the typical grain. Then the following equations hold:*

$$\begin{aligned} \bar{V}_d(\Theta) &= 1 - e^{-\bar{V}_d(\Theta_0)} \\ \bar{V}_j(\Theta) &= e^{-\bar{V}_d(\Theta_0)} \left( \bar{V}_j(\Theta_0) - \frac{d! \kappa_d}{j! \kappa_j} \sum_{s=2}^{d-j} \frac{(-1)^s}{s!} \sum_{\substack{k_1, \dots, k_s = j+1 \\ k_1 + \dots + k_s = (s-1)d+j}}^{d-1} \prod_{i=1}^s \frac{k_i! \kappa_{k_i}}{d! \kappa_d} \bar{V}_{k_i}(\Theta_0) \right) \quad (j < d), \end{aligned}$$

where  $\bar{V}_i(\Theta_0) := \gamma \cdot \mathbf{E}[V_i(\Theta_0)]$  for  $i \in \{0, \dots, d\}$ .

**Proof. (Theorem 3.2.1)** See the derivation above.

For modelling purposes it is more convenient to work with the geometrical quantities the Minkowski functionals represent instead of the Minkowski functionals themselves. The following two corollaries give the Miles-equations in two and three dimensions translated for the geometrical quantities using the scaling factors from table 2.2.4. These corollaries are the key for all modelling within the Boolean Model that follows. Note that  $\mathbf{E}[\chi(\Theta_0)] = 1$  for convex grains.

**Corollary 3.2.2. (Miles-Equations in two dimensions)** *Let  $d = 2$  and  $\Theta$  as in Theorem 3.2.1, then:*

$$\begin{aligned} \bar{v}(\Theta) &= 1 - e^{-\gamma \cdot \mathbf{E}[v(\Theta_0)]} \\ \bar{s}(\Theta) &= e^{-\gamma \cdot \mathbf{E}[v(\Theta_0)]} \cdot \gamma \mathbf{E}[s(\Theta_0)] \\ \bar{\chi}(\Theta) &= e^{-\gamma \cdot \mathbf{E}[v(\Theta_0)]} \cdot \left( \gamma - \frac{1}{4\pi} \gamma^2 \mathbf{E}[s(\Theta_0)]^2 \right). \end{aligned}$$

**Proof. (Corollary 3.2.2)** Direct consequence of Theorem 3.2.1 and Table 2.2.4.

**Corollary 3.2.3.** (*Miles-Equations in three dimensions*) Let  $d = 3$  and  $\Theta$  as in Theorem 3.2.1, then:

$$\begin{aligned}\bar{v}(\Theta) &= 1 - e^{-\gamma \mathbf{E}[v(\Theta_0)]} \\ \bar{s}(\Theta) &= e^{-\gamma \mathbf{E}[v(\Theta_0)]} \cdot \gamma \mathbf{E}[s(\Theta_0)] \\ \bar{M}(\Theta) &= e^{-\gamma \mathbf{E}[v(\Theta_0)]} \cdot \left( \gamma \cdot \mathbf{E}[M(\Theta_0)] - \frac{\pi^2}{32} \cdot \gamma^2 \mathbf{E}[s(\Theta_0)]^2 \right) \\ \bar{\chi}(\Theta) &= e^{-\gamma \mathbf{E}[v(\Theta_0)]} \cdot \left( \gamma - \frac{1}{4\pi} \gamma^2 \cdot \mathbf{E}[M(\Theta_0)] \cdot \mathbf{E}[s(\Theta_0)] + \frac{\pi}{384} \cdot \gamma^3 \cdot \mathbf{E}[s(\Theta_0)]^3 \right).\end{aligned}$$

**Proof.** (Corollary 3.2.3) Direct consequence of Theorem 3.2.1 and Table 2.2.4.

It is straightforward to see that these systems of equations can be solved for the parameters of the Boolean Model by forward insertion. An especially interesting fact is that the intensity  $\gamma$  of the underlying stationary Poisson Process is determined if the whole set of specific Minkowski functionals for the stationary and isotropic Boolean structure is known and doesn't depend on the grain distribution  $Q$  at all. For  $d = 2$  inverting the equation system in 3.2.2 one gets:

**Corollary 3.2.4.** (*Finding intensity and grain quantities in 2D*)

$$\begin{aligned}\gamma &= \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \frac{\bar{s}^2(\Theta)}{(1 - \bar{v}(\Theta))^2} \\ \mathbf{E}[v(\Theta_0)] &= -\frac{1}{\gamma} \cdot \log(1 - \bar{v}(\Theta)) = -\frac{4\pi \cdot \log(1 - \bar{v}(\Theta)) \cdot (1 - \bar{v}(\Theta))^2}{4\pi(1 - \bar{v}(\Theta))\bar{\chi}(\Theta) + \bar{s}^2(\Theta)} \\ \mathbf{E}[s(\Theta_0)] &= \frac{\bar{s}(\Theta)}{\gamma \cdot (1 - \bar{v}(\Theta))} = \frac{4\pi \cdot \bar{s}(\Theta) \cdot (1 - \bar{v}(\Theta))}{4\pi(1 - \bar{v}(\Theta))\bar{\chi}(\Theta) + \bar{s}^2(\Theta)}.\end{aligned}$$

**Proof.** (Corollary 3.2.4) Simple arithmetics.

And similarly for  $d = 3$  inverting the system 3.2.3 one obtains:

**Corollary 3.2.5.** (*Finding intensity and grain quantities in 3D*)

$$\begin{aligned}\gamma &= \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \frac{\bar{M}(\Theta)\bar{s}(\Theta)}{(1 - \bar{v}(\Theta))^2} + \frac{\pi}{192} \frac{\bar{s}(\Theta)^3}{(1 - \bar{v}(\Theta))^3} \\ \mathbf{E}[v(\Theta_0)] &= -\frac{1}{\gamma} \cdot \log(1 - \bar{v}(\Theta)) \\ &= -\frac{192\pi \log(1 - \bar{v}(\Theta))(1 - \bar{v}(\Theta))^3}{192\pi\bar{\chi}(\Theta)(1 - \bar{v}(\Theta))^2 + 48\bar{M}(\Theta)\bar{s}(\Theta)(1 - \bar{v}(\Theta)) + \pi^2\bar{s}^3(\Theta)} \\ \mathbf{E}[s(\Theta_0)] &= \frac{\bar{s}(\Theta)}{\gamma \cdot (1 - \bar{v}(\Theta))} = \frac{192\pi\bar{s}(\Theta)(1 - \bar{v}(\Theta))^2}{192\pi\bar{\chi}(\Theta)(1 - \bar{v}(\Theta))^2 + 48\bar{M}(\Theta)\bar{s}(\Theta)(1 - \bar{v}(\Theta)) + \pi^2\bar{s}^3(\Theta)} \\ \mathbf{E}[M(\Theta_0)] &= \frac{1}{\gamma} \left( \frac{\bar{M}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{\pi^2}{32} \frac{\bar{s}^2(\Theta)}{(1 - \bar{v}(\Theta))^2} \right) \\ &= \frac{192\pi\bar{M}(\Theta)(1 - \bar{v}(\Theta))^2 + 6\pi^3\bar{s}^2(\Theta)(1 - \bar{v}(\Theta))}{192\pi\bar{\chi}(\Theta)(1 - \bar{v}(\Theta))^2 + 48\bar{M}(\Theta)\bar{s}(\Theta)(1 - \bar{v}(\Theta)) + \pi^2\bar{s}^3(\Theta)}.\end{aligned}$$

**Proof.** (Corollary 3.2.5) Simple arithmetics.

According to Theorem 3.2.1 the set of specific Minkowski functionals  $\bar{V}_j(\Theta)$  thus uniquely determines the intensity of the underlying Poisson Process and the moments  $\mathbf{E}[V_j(\Theta_0)]$  of the typical grain. Because only the *moments* of the expected grain quantities are uniquely defined, there is still a large freedom for choosing the grain distribution  $Q$  and one might conjecture already from here that there is a wide range of visually and structurally different Boolean structures leading to the same specific Minkowski functionals. This phenomenon is referred to as *indeterminacy* in the following.

### 3.2.2 Attainability of Specific Minkowski-Values Within the Homothetic Boolean Model

From the Miles equations it can be seen that not every combination of preset Minkowski values  $v_j := \bar{V}_j(\Theta)$  can be realized within the Boolean Model. However there are only four types of necessary restrictions that could possibly be violated. Three of them are trivial: First the intensity of the underlying Poisson Process must of course be positive, second for convex grains  $\Theta_0$  the mean Minkowski-functionals  $\mathbf{E}[V_j(\Theta_0)]$  must be positive (this follows from Theorem 2.1.3, part vii)) and third the Jensen inequality for the moments of the  $V_j(\Theta_0)$  and their parameters must be fulfilled:

$$\begin{array}{l} \gamma > 0, \\ \mathbf{E}[V_j^p(\Theta_0)] \geq \mathbf{E}[V_j^q(\Theta_0)]^{\frac{p}{q}} \quad \text{for } p > q. \end{array}$$

In addition to that, the Minkowski functionals of the typical grain must fulfill the set of *isoperimetric inequalities* which implies that also a set of corresponding inequalities must hold true for the moments of their distributions.

**Theorem 3.2.6. (isoperimetric inequalities)** Let  $K \in \mathcal{K}$  and  $j, k \in \mathbb{N}$  with  $0 < j < k \leq d$ . Then:

$$\left( \frac{\kappa_{d-j}}{\binom{d}{j}} \cdot V_j(K) \right)^k \geq \kappa_d^{k-j} \cdot \left( \frac{\kappa_{d-k}}{\binom{d}{k}} \cdot V_k(K) \right)^j.$$

**Proof. (Theorem 3.2.6)** The result follows from the theory of mixed volumes. It is a direct consequence of the *Aleksandrov-Fenchel-inequality* proved in convexity theory. Details can be found in (Schneider 1993, Chapter 6, especially equation 6.4.6).

In dimensions  $d = 2$  and  $d = 3$  the moment inequalities derived from Theorem 3.2.6 read:

$$\begin{array}{ll} d = 2 : & d = 3 : \\ \mathbf{E}[s^2(\Theta_0)] \geq 4\pi \cdot \mathbf{E}[v(\Theta_0)] & \mathbf{E}[s^3(\Theta_0)] \geq 36\pi \cdot \mathbf{E}[v^2(\Theta_0)] \\ & \mathbf{E}[M^3(\Theta_0)] \geq 48\pi^2 \cdot \mathbf{E}[v(\Theta_0)] \\ & \mathbf{E}[M^2(\Theta_0)] \geq 4\pi \cdot \mathbf{E}[s(\Theta_0)] \end{array}$$

In this section we want to restrict ourselves to the *homothetic* Boolean model for which the typical grain  $\Theta_0$  is formed by expanding an arbitrary convex *prototype grain*  $\Theta_{00} \in \mathcal{K}$  by some random expansion factor  $c$ , ic.

$$\Theta_0 = c \cdot \Theta_{00}.$$

For the general homothetic setting it is not possible to calculate the attainable region exactly as long as the prototype grain  $\Theta_{00}$  is not specified. However for some selected special cases this can in fact be done. In the following we want to investigate the attainable region of specific Minkowski-functionals for the following special cases of the homothetic Boolean Model:

- 2D-circular Boolean Model,
- 3D-spherical Boolean Model,
- 2D-elliptical Boolean Model with fixed axis ratio  $R = \frac{b}{a}$ ,
- 3D-spheroidal Boolean Model (ie.  $b = c$ ) with fixed axis ratio  $R = \frac{b}{a}$ .

In the case of the spherical Boolean Model the isoperimetric inequalities are of course exactly fulfilled for every grain and hence we also have equality in the moment equations above. This leaves us with two necessary restrictions in the 2D-case and three necessary restrictions in the 3D-case:

2D-restrictions	3D-restrictions
$\gamma > 0,$ $\mathbf{E} [r^2] \geq \mathbf{E} [r]^2,$	$\gamma > 0,$ $\mathbf{E} [r^2] \geq \mathbf{E} [r]^2,$ $\mathbf{E} [r^3] \geq \mathbf{E} [r^2]^{\frac{3}{2}},$

where  $r$  is the radius of the circle or sphere respectively. Note that in the 3D-case the inequality  $\mathbf{E} [r^3] \geq \mathbf{E} [r]^3$  follows from the above and is therefore not required. Because the expected grain quantities in the spherical case are directly linked to the moments of the radius these restrictions translate directly into inequalities for the specific model quantities  $\bar{V}_j(\Theta)$  via the Miles equations in 3.2.2 and 3.2.3. Exemplarily for the 2D-case we have:

$$\begin{aligned} \gamma > 0 &\Leftrightarrow 4\pi \cdot \bar{\chi}(\Theta)(1 - \bar{v}(\Theta)) + \bar{s}^2(\Theta) > 0 \\ 4\pi \mathbf{E} [v(\Theta_0)] = \mathbf{E} [s^2(\Theta_0)] &\geq \mathbf{E} [s(\Theta_0)]^2 \geq 0 \Leftrightarrow -4\pi \frac{Z_v}{N} \geq \frac{Z_s^2}{N^2}, \end{aligned}$$

where the quantities  $Z_v := 4\pi \log(1 - \bar{v}(\Theta))(1 - \bar{v}(\Theta))^2 \leq 0$ ,  $Z_s := 4\pi \bar{s}(\Theta) \cdot (1 - \bar{v}(\Theta)) \geq 0$  and  $N := 4\pi(1 - \bar{v}(\Theta))\bar{\chi}(\Theta) + \bar{s}^2(\Theta)$  are shortcuts for the nominators and denominators of the corresponding ratios on the right hand side of the equation system in corollary 3.2.4. Note that  $\gamma > 0 \Leftrightarrow N > 0$  follows from the first condition. Hence the only restriction remaining is  $N \geq -\frac{Z_s^2}{4\pi Z_v}$  which finally translates to:

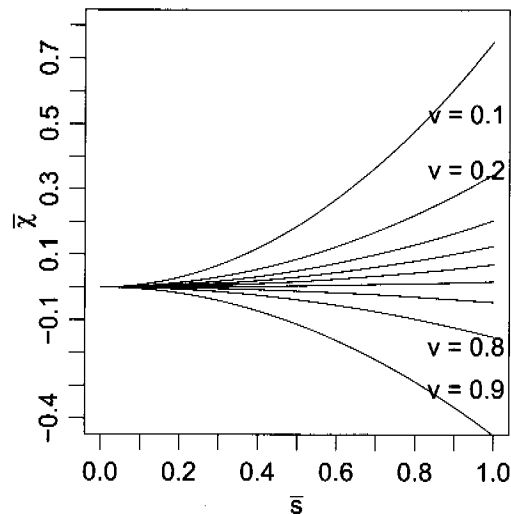
$$\bar{\chi}(\Theta) \geq -\bar{s}^2(\Theta) \frac{1 + \log(1 - \bar{v}(\Theta))}{4\pi \log(1 - \bar{v}(\Theta))(1 - \bar{v}(\Theta))}. \quad (3.3)$$

This confirms the intuition that the smaller the specific volume is, the harder it becomes to attain low (especially negative) values for the specific Euler characteristic. In fact the lower bound on the right hand side converges to infinity as  $\bar{v}(\Theta) \downarrow 0$ . On the other hand with  $\bar{v}(\Theta) \uparrow 1$  every value for the specific Euler characteristic should be attainable.

Of course when fitting a Boolean Model to real data, one usually imposes a radius distribution from a certain model class. Usually within such a class not every combination of moments which is feasible according to the above restrictions can be achieved, in fact there are additional class-specific restrictions.

Based on inequality 3.3, figure 3.2 shows for some selected fixed values out of the range of  $\bar{v}(\Theta)$  the region of attainable pairs  $(s(\Theta), \chi(\Theta))$  within the 2D-spherical model. The attainable region lies *above* the drawn boundary line. Figure 3.2 exhibits the symmetry which can be expected from the relationships between the Minkowski functionals of the solid and pore space although one must be aware that the (closed) complement of a Boolean Model is *not* again a Boolean Model in general.

### Attainable region in 2D Spherical Boolean Model



**Figure 3.2:** The figure shows for selected values of  $\bar{v}(\Theta)$  the region of attainable pairs  $(s(\Theta), \chi(\Theta))$  within  $[0, 1] \times [-1, 1]$  for the 2D-spherical Boolean Model. For each  $\bar{v}(\Theta) \in \{0.1, 0.2, \dots, 0.9\}$  the boundary lines between the region of attainable pairs and non-attainable-ones is drawn. The attainable region is above the drawn line.

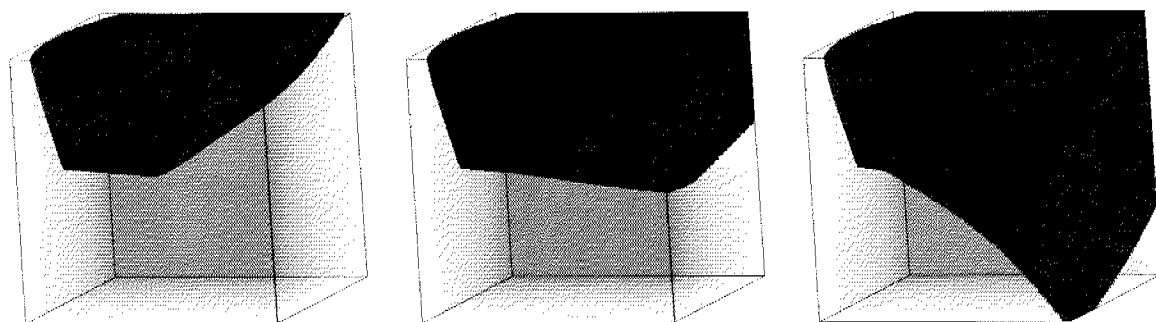
A typical cross-section from the coarse-sand specimen in figure 1.1 has specific Minkowski functionals around:

$$\bar{v}(\Theta) = 0.6, \bar{s}(\Theta) = 0.05, \bar{\chi}(\Theta) = 1.6 \cdot 10^{-4},$$

hence according to inequality 3.3 this set of specific Minkowski-values does not violate the Jensen and intensity constraints and should thus be attainable within the 2D spherical Boolean Model. However it turns out that to obtain these values in practice may become difficult even for reasonable parametric classes for the grain distribution.



A similar analysis for the attainability of sets of specific Minkowski functionals  $(\bar{v}(\Theta), \bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  can be done for 3D by numerical determination of the attainable region. Again the isoperimetric restrictions must not be taken into account, and we are left with the Jensen restrictions for the radius and  $\gamma > 0$ , but because the radius distribution must be concentrated on positive values, we must additionally require that the third 3D-Miles equation for the mean integral of mean curvature  $\mathbf{E}[M(\Theta_0)]$  from corollary 3.2.5 leads to  $\mathbf{E}[M(\Theta_0)] > 0$  or equivalently  $\mathbf{E}[r] > 0$ , otherwise we wouldn't be able to find a *nonnegative* random variable  $r$  with the required first three moments. Note that from the Miles equations we have  $\gamma > 0 \Rightarrow \mathbf{E}[r^3] > 0, \mathbf{E}[r^2] > 0$ , but opposed to the 2D-case  $\gamma > 0$  *not* necessarily ensures  $\mathbf{E}[r] > 0$ , hence this is an additional condition. Figure 3.3 shows the attainable region  $(\bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  for three selected values of specific volume  $\bar{v}(\Theta) \in \{0.1, 0.5, 0.9\}$ .



**Figure 3.3:** The black region contains all the attainable triples  $(\bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  for a fixed value of  $\bar{v}(\Theta)$  in the 3D-spherical Boolean Model. This value is  $\bar{v}(\Theta) = 0.1$  (left),  $\bar{v}(\Theta) = 0.5$  (middle),  $\bar{v}(\Theta) = 0.9$  (right). The  $\bar{s}$ -values extend from left to right in  $[0, 1]$  ( $x$ -direction), the  $\bar{M}$ -values extend from front to back in  $[-1, 1]$  ( $y$ -direction) and the  $\bar{\chi}$ -values extend from bottom to top in  $[-1, 1]$  ( $z$ -direction). Hence the lower left front vertex of the cubes is the point  $(\bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta)) = (0, -1, -1)$ .

Figure 3.3 displays a similar behaviour as in 2 dimensions, for small  $v(\Theta)$  it is hard to attain small  $\chi$ -values and as  $\bar{v}(\Theta)$  increases, the attainable region might grow substantially. To attain small values for  $\bar{M}(\Theta)$  seems to be difficult for small  $\bar{s}(\Theta)$ .

The attainable region can be further enlarged substantially by relaxing the assumption of spherical grains and allow for elliptical or ellipsoidal ones instead. In the 2D case let us denote  $a$  and  $b$  the semi-major and minor axis of the typical ellipse  $\Theta_0$  with  $a > b$  and assume for simplicity that the axis-ratio  $R = b/a$  is deterministically fixed. The expected volume and surface for  $\Theta_0$  are given by elementary geometry:

$$\begin{aligned} \mathbf{E}[v(\Theta_0)] &= \mathbf{E}[\pi \cdot a \cdot b] = \pi \mathbf{E}[a^2] \cdot R, \\ \mathbf{E}[s(\Theta_0)] &= \mathbf{E}\left[a \int_0^{2\pi} \sqrt{\sin^2 \phi + R^2 \cdot \cos^2 \phi} d\phi\right] := \mathbf{E}[a] \cdot I(R), \end{aligned}$$

where  $I(R) = 4 \cdot E(\sqrt{1 - R^2})$  and  $E(\cdot)$  is the complete elliptic integral of the second kind (see Abramowitz and Stegun 1972, p 589 ff.). The necessary restrictions for this setting are  $\gamma > 0$  and  $\mathbf{E}[a^2] \geq \mathbf{E}[a]^2$  and the latter translates to:

$$\frac{\mathbf{E}[v(\Theta_0)]}{\pi R} \geq \frac{\mathbf{E}[s(\Theta_0)]^2}{I(R)^2}$$

Note that although the (expected) isoperimetric inequality is strict in the elliptical case it does not add an additional necessary constraint if the axis ratio  $R$  is non-random because it reads

$$4\pi^2 R \mathbf{E}[a^2] = 4\pi \mathbf{E}[V(\Theta_0)] \leq \mathbf{E}[s^2(\Theta_0)] = \mathbf{E}[a^2] I(R)^2,$$

and  $4\pi^2 R \leq I(R)^2$  holds  $\forall R \in [0, 1]$  because this is the deterministic isoperimetric inequality for ellipses with  $a = 1$ . Analogously as in the spherical case we find a lower bound for the attainable Euler characteristic by plugging in the expressions for  $\mathbf{E}[v(\Theta_0)]$  and  $\mathbf{E}[s(\Theta_0)]$  from the 2D-Miles-equations (Corollary 3.2.2):

$$\bar{\chi}(\Theta) \geq -s(\Theta)^2 \frac{4\pi^2 R + \log(1 - \bar{v}(\Theta)) I(R)^2}{4\pi \log(1 - \bar{v}(\Theta)) (1 - \bar{v}(\Theta)) I(R)^2}. \quad (3.4)$$

For  $R = 1$  we have  $I(R) = 2\pi$  and inequality 3.3 is rediscovered. Note that  $I(R)$  is strictly increasing in  $R$  and that  $\lim_{R \downarrow 0} I(R) = 4$ . From equation 3.4 it is obvious that the lowest attainable value for  $\bar{\chi}(\Theta)$  in the elliptical Boolean Model with fixed deterministic axis ratio for given specific volume  $\bar{v}(\Theta)$  and surface  $\bar{s}(\Theta)$  is:

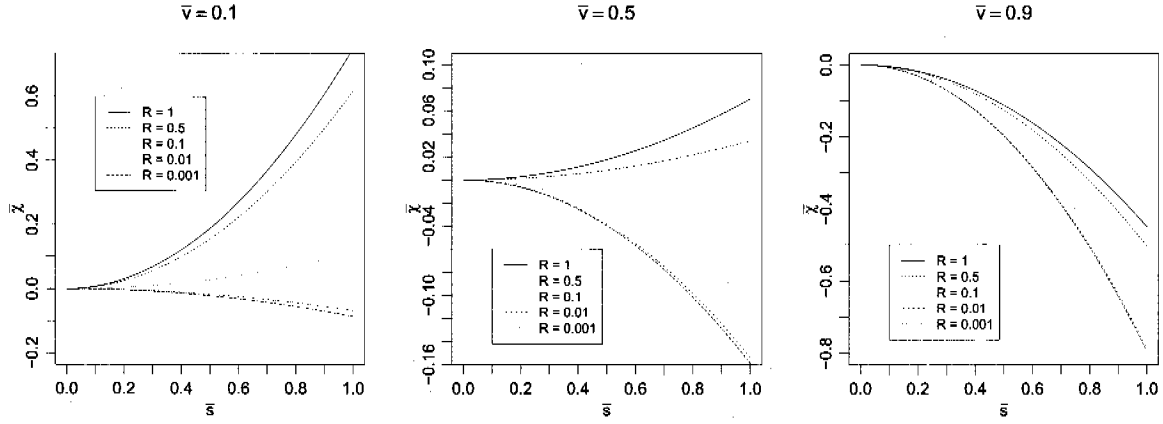
$$\bar{\chi}_{\min}(\Theta) = -\frac{\bar{s}^2(\Theta)}{4\pi(1 - \bar{v}(\Theta))}, \quad (3.5)$$

but again one must be aware that the choice of the grain-distribution class for  $a$  might impose further restrictions. The lower bound from equation 3.5 in fact holds for *all* convex 2D-Boolean Models as it immediately follows from Corollary 3.2.4 and the requirement  $\gamma > 0$ . Thus a simple example which is not realizable within the (general) Boolean Model is a finite circular disc with sufficiently many identical tiny holes. These holes do not influence volume and surface substantially but reduce the Euler-characteristic to an arbitrarily low value.

Figure 3.4 shows how substantially the attainable region of pairs  $(\bar{s}(\Theta), \bar{\chi}(\Theta))$  can be enlarged by allowing eccentric elliptical grains ( $R \ll 1$ ) for a selection of values for the specific volume  $\bar{v}(\Theta) \in \{0.1, 0.5, 0.9\}$ .

Figure 3.4 shows that for high eccentricity values  $R \ll 1$  the attainable region grows considerably and even for small porosity  $\bar{v}$ , very low  $\bar{\chi}$  can be obtained. However also for elliptical grains there exists a  $\bar{\chi}$ -value which can not be underrun (given by equation 3.5). One must also bear in mind that the shape of the Boolean grains heavily influences the optical appearance of the artificial random structures generated by the Boolean Model. If one tries to model "physical" grains of approximately spherical shape in a real structure (as when modelling sand data), very eccentric elliptical grains have a good chance to replicate the Minkowski functionals of the original structure, however the optical appearance of the Boolean structure will probably be completely different from the original image.

For the three dimensional Boolean Model with ellipsoidal grains the same analysis of the attainable region will be done in the following. For simplicity, we assume that the typical



**Figure 3.4:** The boundary lines between attainable and non-attainable region of pairs  $(\bar{s}(\Theta), \bar{\chi}(\Theta))$  in the 2D-elliptical Boolean Model with elliptical grains. For selected values of the specific volume  $\bar{v}(\Theta) = 0.1$  (left),  $\bar{v}(\Theta) = 0.5$  (middle) and  $\bar{v}(\Theta) = 0.9$  (right) boundary lines are drawn for a selection of eccentricity values  $R$  for the typical elliptical grain, varying from  $R = 1$  (spherical case) down to  $R = 0.001$  (extreme eccentricity). The attainable region lies above the drawn lines.

ellipsoid with semi-axes  $a \geq b \geq c$  is of spheroidal shape, ie.  $b = c$  and its axis ratio  $0 \leq R = b/a = c/a \leq 1$  is deterministic. For a general ellipsoid the elliptic integrals for specific surface  $\bar{s}$  and specific mean curvature  $\bar{M}$  are unfortunately not analytically tractable and one has to resort to tables or numerical approximation. However, thanks to the spheroidal symmetry assumed here, all Minkowski functionals of the typical ellipsoid can explicitly be given:

$$\begin{aligned} \mathbf{E}[v(\Theta_0)] &= \mathbf{E}\left[\frac{4}{3}\pi abc\right] = \frac{4}{3}\pi \mathbf{E}[a^3] R^2, \\ \mathbf{E}[s(\Theta_0)] &= \mathbf{E}\left[2\pi a^2 R \cdot \left(R + \frac{1}{\sqrt{1-R^2}} \arcsin \sqrt{1-R^2}\right)\right] =: \mathbf{E}[a^2] \cdot I_s(R), \\ \mathbf{E}[M(\Theta_0)] &= \mathbf{E}\left[2\pi a \cdot \left(R^2 + \frac{1}{\sqrt{1-R^2}} \arcsin \sqrt{1-R^2}\right)\right] =: \mathbf{E}[a] \cdot I_M(R), \end{aligned}$$

where  $I_s(R) \in [0, 4\pi]$ ,  $I_M(R) \in [\pi^2, 4\pi]$ . As in the spherical case, the relevant restrictions are  $\gamma > 0$  and the two moment inequalities implied by Jensen. The latter can be written as:

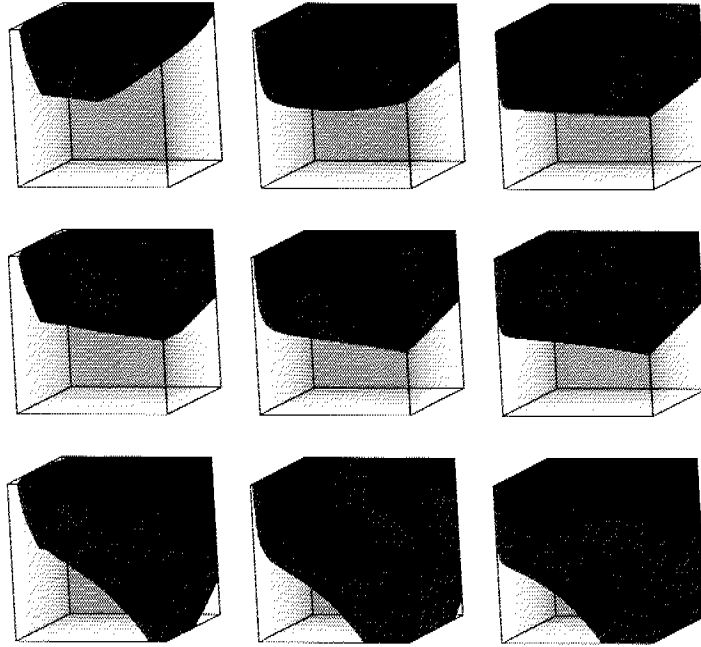
$$\frac{\mathbf{E}[s(\Theta_0)]}{I_s(R)} \geq \frac{\mathbf{E}[M(\Theta_0)]^2}{I_M(R)^2}, \quad \frac{3 \mathbf{E}[v(\Theta_0)]}{4\pi R^2} \geq \left(\frac{\mathbf{E}[s(\Theta_0)]}{I_s(R)}\right)^{\frac{3}{2}}.$$

There is no further restriction for the moments coming from the isoperimetric inequalities because of the spheroidal structure with deterministic axis ratio: Exemplarily for the isoperimetric relationship between volume and surface of the typical grain, we have:

$$\begin{aligned} \mathbf{E}[a^6] I_s(R)^3 = \mathbf{E}[s^3(\Theta_0)] &\geq 36\pi \mathbf{E}[v(\Theta_0)^2] = 36\pi \left(\frac{4}{3}\pi R^2\right)^2 \mathbf{E}[a^6] \\ \iff I_s(R)^3 &\geq 36\pi \left(\frac{4}{3}\pi R^2\right)^2. \end{aligned} \quad (3.6)$$

The latter inequality holds trivially because it is the deterministic isoperimetric inequality for a spheroid with semi-major axis  $a = 1$ . But again similarly as in the 3D-spherical case because the semi-major axis  $a$  is non-negative we must additionally require that the Miles equation for the mean integral of mean curvature  $\mathbf{E}[M(\Theta_0)]$  (third equation in Corollary 3.2.2) leads to  $\mathbf{E}[a] > 0$ , such that a *nonnegative* random variable  $a$  with the required first three moments can be found.

Figure 3.5 displays the attainable region  $(\bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  for selected values of the porosity  $v(\Theta)$  and different eccentricity-levels parametrized by the axis-ratio  $R$ . By relaxing the spheroid assumptions this region can be further enlarged although one must be aware that there are additional restrictions because the axis ratio  $R$  is then an additional source of randomness.



**Figure 3.5:** The black region is the set of attainable triples  $(\bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  within the 3D-Boolean Model with spheroidal grains for selected values of the porosity  $v(\Theta)$  and selected axis-ratios  $R$ . The porosity values chosen are  $\bar{v}(\Theta) = 0.1$  (top row),  $\bar{v}(\Theta) = 0.5$  (middle row) and  $\bar{v}(\Theta) = 0.9$  (bottom row). The axis-ratios chosen are  $R = 0.5$  (left column),  $R = 0.1$  (middle column),  $R = 0.01$  (right column). In all displayed cubes the  $\bar{s}$ -values extend from left to right within  $[0, 1]$ , the  $\bar{M}$ -values extend from front to back within  $[-1, 1]$  and the  $\bar{\chi}$ -values extend from bottom to top within  $[-1, 1]$ . Hence in all cubes the front lower left vertex is the point  $(\bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta)) = (0, -1, -1)$ .

Now we turn back to the general homothetic Boolean Model. For the attainability of a given quadruple of specific Minkowski-values  $(\bar{v}(\Theta), \bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  in three dimensions, the following can be said:

**Theorem 3.2.7. (Attainability in the 3D-homothetic Boolean Model)** *Let  $\Theta$  be a 3D-Boolean Model whose typical grain  $\Theta_0$  is a random expansion of a fix convex prototype-grain  $\Theta_{00} \in \mathcal{K}$  with random expansion factor  $c$ , ie.  $\Theta_0 = c \cdot \Theta_{00}$ . Denote further the right hand side of the Miles-equations for the Poisson intensity  $\gamma$  and the mean Minkowski functionals  $\mathbf{E}[V_j(\Theta_0)]$  of the typical grain in Corollary 3.2.5 by  $f_\gamma, f_v, f_s$  and  $f_M$ . Now let a quadruple of values for the specific Minkowski-functionals  $(\bar{v}(\Theta), \bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  be given.*

*i). If the inequalities  $f_\gamma > 0, f_M > 0$  are simultaneously fulfilled for the given quadruple, one can always find a convex prototype-grain  $\Theta_{00}$  such that the quadruple can be attained within the homothetic Boolean Model. The prototype-grain can be chosen to be an ellipsoid.*

*ii). If the inequalities of i). hold and additionally*

$$\frac{f_s}{f_M^2} \geq \frac{1}{4\pi}, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \geq \frac{1}{\sqrt{36\pi}}$$

*are simultaneously fulfilled, the quadruple can be attained with an arbitrary convex prototype-grain  $\Theta_{00}$ . The coefficient of variation of the expansion factor  $c$  is smallest for  $\Theta_{00}$  being the unit sphere.*

*iii). If the inequalities of i). hold and additionally we have both*

$$\frac{f_s}{f_M^2} \leq \frac{1}{4\pi}, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \leq \frac{1}{\sqrt{36\pi}},$$

*the expansion factor  $c$  may be chosen to be a deterministic constant and  $\Theta_{00}$  may be chosen to be an ellipsoid.*

**Proof. (Theorem 3.2.7)** We can always find a distribution for the expansion factor  $c$  with the first three moments  $\mathbf{E}[c], \mathbf{E}[c^2]$  and  $\mathbf{E}[c^3]$  arbitrarily fixed as long as the Jensen inequalities  $\mathbf{E}[c^2] \geq \mathbf{E}[c]^2$  and  $\mathbf{E}[c^3] \geq \mathbf{E}[c^2]^{3/2}$  are fulfilled. (The inequality  $\mathbf{E}[c^3] \geq \mathbf{E}[c]^3$  follows from the latter two). Because only a nonnegative expansion factor makes sense, we must additionally ensure that  $\mathbf{E}[c] > 0$ . Furthermore we have:

$$\begin{aligned} f_v &= \mathbf{E}[v(\Theta_0)] = \mathbf{E}[c^3] \cdot v(\Theta_{00}), \\ f_s &= \mathbf{E}[s(\Theta_0)] = \mathbf{E}[c^2] \cdot s(\Theta_{00}), \\ f_M &= \mathbf{E}[M(\Theta_0)] = \mathbf{E}[c] \cdot M(\Theta_{00}). \end{aligned}$$

Hence from the third equation we have that  $f_M > 0 \Rightarrow \mathbf{E}[c] > 0$ . Furthermore we must require:

$$\frac{\mathbf{E}[c^2]}{\mathbf{E}[c]^2} = \frac{f_s}{f_M^2} \cdot \frac{M^2(\Theta_{00})}{s(\Theta_{00})} \geq 1, \quad (3.7)$$

$$\frac{\mathbf{E}[c^3]}{\mathbf{E}[c^2]^{\frac{3}{2}}} = \frac{f_v}{f_s^{\frac{3}{2}}} \cdot \frac{s^{\frac{3}{2}}(\Theta_{00})}{v(\Theta_{00})} \geq 1. \quad (3.8)$$

ad i):

From the deterministic isoperimetric inequalities (Theorem 3.2.6) we immediately get a lower bound on the following two quotients:

$$q_1(\Theta_{00}) := \frac{s^{\frac{3}{2}}(\Theta_{00})}{v(\Theta_{00})} \geq (36\pi)^{1/2}, \quad q_2(\Theta_{00}) := \frac{M^3(\Theta_{00})}{s^{\frac{3}{2}}(\Theta_{00})} \geq (4\pi)^{3/2}, \quad (3.9)$$

but there exists no upper bound on these quotients. We can find a sequence of prototype-ellipsoids  $\Theta_{00}^i$  such that

$$\lim_{i \rightarrow \infty} \frac{s^{\frac{3}{2}}(\Theta_{00}^i)}{v(\Theta_{00}^i)} = \infty, \quad \lim_{i \rightarrow \infty} \frac{M^3(\Theta_{00}^i)}{s^{\frac{3}{2}}(\Theta_{00}^i)} = \infty. \quad (3.10)$$

Such a sequence  $\Theta_{00}^i$  can already be found within the set of ellipsoids by letting one of the semi-major axes grow towards infinity. This can be seen from equations 3.15 below and is also illustrated in figure 3.6. But then because  $f_\gamma > 0$  ensures that also  $f_v > 0$  and  $f_s > 0$ , we can choose  $i_0$  high enough such that simultaneously equations 3.7, 3.8 are fulfilled:

$$\frac{f_s}{f_M^2} \cdot \frac{M^2(\Theta_{00}^{i_0})}{s(\Theta_{00}^{i_0})} \geq 1, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \cdot \frac{s^{\frac{3}{2}}(\Theta_{00}^{i_0})}{v(\Theta_{00}^{i_0})} \geq 1,$$

and thus a distribution for  $c$  concentrated on  $\mathbb{R}^+$  with the required moments  $\mathbf{E}[c^3]$ ,  $\mathbf{E}[c^2]$ ,  $\mathbf{E}[c]$  can be found.

ad ii):

Let  $\Theta_{00}$  be an arbitrary convex prototype-grain. If the inequalities

$$\frac{f_s}{f_M^2} \geq \frac{1}{4\pi}, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \geq \frac{1}{\sqrt{36\pi}}$$

hold, then together with the isoperimetric bounds from equation 3.9 we conclude:

$$\frac{f_s}{f_M^2} \cdot \frac{M^2(\Theta_{00})}{s(\Theta_{00})} \geq \frac{1}{4\pi} \cdot 4\pi = 1, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \cdot \frac{s^{\frac{3}{2}}(\Theta_{00})}{v(\Theta_{00})} \geq \frac{1}{\sqrt{36\pi}} \cdot \sqrt{36\pi} = 1,$$

hence equations 3.7, 3.8 are fulfilled.

In the light of equation 3.7, we see that the quotient  $\mathbf{E}[c^2] / \mathbf{E}[c]^2$  becomes minimal if we let  $q_1(\Theta_{00})$  attain its minimum  $4\pi$  for the unit sphere. This is only feasible for

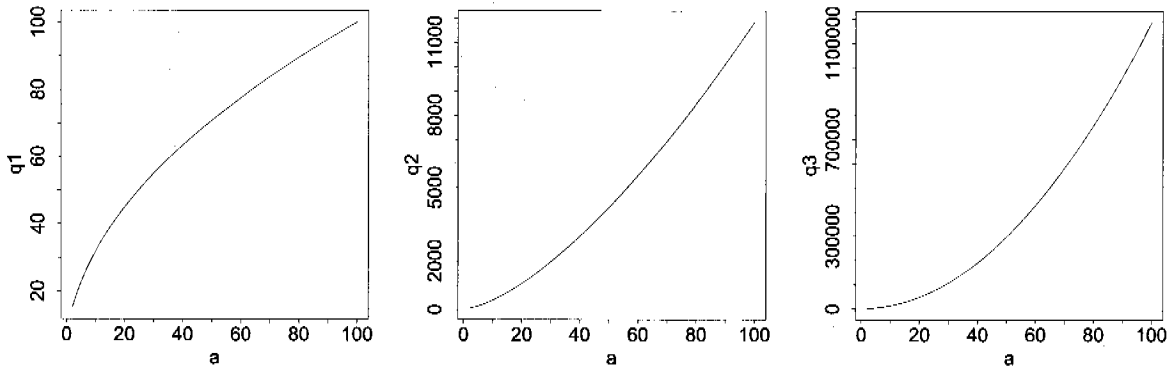
$$\frac{f_s}{f_M^2} \geq \frac{1}{4\pi}, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \geq \frac{1}{\sqrt{36\pi}}$$

because we must ensure that inequalities 3.7, 3.8 hold true.

ad iii):

If both

$$\frac{f_s}{f_M^2} \leq \frac{1}{4\pi}, \quad \frac{f_v}{f_s^{\frac{3}{2}}} \leq \frac{1}{\sqrt{36\pi}},$$



**Figure 3.6:** The three quotients  $q_1 := s^{\frac{3}{2}}(\Theta_{00})/v(\Theta_{00})$ ,  $q_2 := M^3(\Theta_{00})/s^{\frac{3}{2}}(\Theta_{00})$  and  $q_3 := M^3(\Theta_{00})/v(\Theta_{00})$ , which are lower bounded by the 3D-isoperimetric inequalities (Theorem 3.2.6) for  $\Theta_{00}$  being a general ellipsoid with the two smaller semi-axes being  $b = 1$ ,  $c = \frac{1}{2}$  as a function of the largest semi-axis  $a$ .

because of equation 3.10 there exists an ellipsoid  $\Theta_{00}$  with

$$\frac{M^2(\Theta_{00})}{s(\Theta_{00})} = \frac{f_M^2}{f_s} \geq 4\pi, \quad \frac{s^{\frac{3}{2}}(\Theta_{00})}{v(\Theta_{00})} = \frac{f_s^{\frac{3}{2}}}{f_v} \geq \sqrt{36\pi},$$

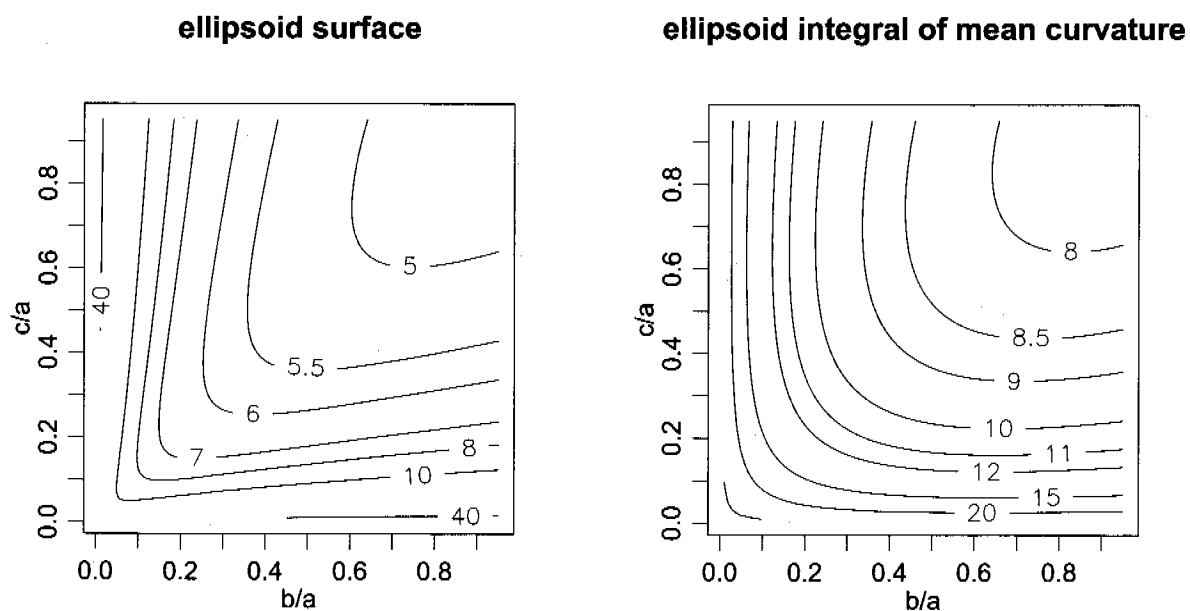
such that we have equality in equations 3.7, 3.8. □

Figure 3.6 illustrates that already within the class of ellipsoids the quotients  $q_1, q_2, q_3$  can be made arbitrarily large. According to the figure the quotient  $q_1$  might be bounded from above, but it is obvious that by "stretching" an ellipsoid of given volume, its surface can be made arbitrarily large. Rigorous arguments can be derived from equation system 3.15 below and the properties of the elliptic integrals  $F$  and  $E$ . Additionally figure 3.7 gives a general illustration of how the surface area  $s$  and the integral of mean curvature depend on the ellipsoid axes.

Of course an analogous result to Theorem 3.2.7 exists also in two dimensions. We state it here for the sake of completeness:

**Theorem 3.2.8. (Attainability in the 2D-homothetic Boolean Model)** Let  $\Theta$  be a 2D-Boolean Model whose typical grain  $\Theta_0$  is a random expansion of a fix convex prototype grain  $\Theta_{00} \in \mathcal{K}$  with random expansion factor  $c$ , ie.  $\Theta_0 = c \cdot \Theta_{00}$ . Denote further the right hand side of the Miles-equations for the Poisson intensity  $\gamma$  and the mean Minkowski functionals  $\mathbf{E}[V_j(\Theta_0)]$  of the typical grain in Corollary 3.2.4 by  $f_\gamma, f_v$  and  $f_s$ . Now let a triple of values for the specific Minkowski-functionals  $(\bar{v}(\Theta), \bar{s}(\Theta), \bar{\chi}(\Theta))$  be given.

- i). If  $f_\gamma > 0$  for the given triple, one can always find a convex prototype-grain  $\Theta_{00}$  such that the triple can be attained within the homothetic Boolean Model. The prototype-grain can be chosen to be an ellipse.
- ii). If  $f_\gamma > 0$  and additionally  $\frac{f_v}{f_s^2} \geq \frac{1}{4\pi}$ , the triple can be attained with an arbitrary convex prototype-grain  $\Theta_{00}$  and the variability of the expansion factor  $c$  is smallest for  $\Theta_{00}$  being the unit disk.
- iii). If  $f_\gamma > 0$  and additionally  $\frac{f_v}{f_s^2} \leq \frac{1}{4\pi}$ , the expansion factor  $c$  may be chosen to be a deterministic constant and  $\Theta_{00}$  may be chosen to be an ellipse.



**Figure 3.7:** Contours of the surface area  $s$  (left) and the integral of mean curvature  $M$  (right) as a function of the axis quotients  $\frac{b}{a}$  and  $\frac{c}{a}$  for a general ellipsoid with unit volume. It is assumed that  $a$  is the largest semi-axis, but  $b$  and  $c$  are in arbitrary order which makes the plots symmetrical to the main diagonal.



**Proof.** (Theorem 3.2.8) analogous to the proof of Theorem 3.2.7.

We conclude this section with a final remark on the 3D-convex homothetic Boolean Model. Please note that similarly as in two dimensions from Corollary 3.2.5 and the requirement  $\gamma > 0$  a lower bound for the Euler characteristic in the *general* 3D-convex Boolean Model is:

$$\bar{\chi}(\Theta) \geq -\frac{\bar{s}(\Theta)}{1 - \bar{v}(\Theta)} \cdot \left( \frac{\bar{M}(\Theta)}{4\pi} + \frac{\pi}{192} \cdot \frac{\bar{s}^2(\Theta)}{1 - \bar{v}(\Theta)} \right),$$

However, in the convex *homothetic* Boolean Model we have the requirement that the random expansion factor  $c$  is positive and thus also  $\mathbf{E}[c] > 0$  which implies  $\mathbf{E}[M(\Theta_0)] \geq 0$ . Hence from Corollary 3.2.5 we learn that in the convex homothetic Boolean Model we also have a lower bound on  $M$ :

$$\bar{M}(\Theta) \geq -\frac{\pi^2}{32} \cdot \frac{\bar{s}^2(\Theta)}{1 - \bar{v}(\Theta)}.$$

### 3.3 Simulation Procedures

Due to the Miles formulae, simulation of a stationary and isotropic Boolean Model is quite straightforward although there are some technical difficulties one may encounter which will be mentioned in the following. We start out with the set of *predetermined* specific Minkowski functionals  $(v_0, \dots, v_d)$  and the goal is to generate a realization of a Boolean Model whose specific Minkowski functionals match the predetermined values:

$$\bar{V}_0(\Theta), \dots, \bar{V}_d(\Theta) = (v_0, \dots, v_d), \quad d \in \{2, 3\}. \quad (3.11)$$

In this section we restrict ourselves to considering the stationary Boolean Models with circular (spherical) and elliptical (ellipsoidal) grains in two and three dimensions. In principle one could generate Boolean Models with much more complicated (even non-convex) grains but the practical realization of such models is usually difficult and often the benefit over the simple spherical models is moderate. In addition to that, many porous structures which turn up in practice can be reasonably modelled by spherical or ellipsoidal Boolean grains because their solid phase is granular and these real material grains can be roughly approximated by spheres or ellipsoids. This is the case for the sand structures we examine in this thesis but one must also be aware that there exist many structures for which a spherical or ellipsoidal Boolean Model is not appropriate, for example if the structure consists of thin fibres (then a fibre-process or tessellation model might help).

Of course there are many methods to simulate Boolean Models and we have tried several of them. In the following we present the techniques that worked best for us with the least effort. In general one can say that the (continuous) realization of a Boolean Model is easy, for example in the circular case, a realization simply consists of two vectors of disk midpoint coordinates and a vector of radii. The hard and time-consuming part is the accurate discretization of a realization on a grid which is crucial if one intends to measure the Minkowski-functionals and evaluate the simulation-process in the end.

### 3.3.1 Simulation of the 2D-Circular Boolean Model

Fixing the specific Minkowski functionals for the stationary Boolean Model  $\Theta$  according to equation 3.11 determines the expected Minkowski functionals of the typical grain  $\mathbf{E}[V_j(\Theta_0)]$ ,  $j \in 1, 2$  and the intensity  $\gamma$  of the underlying Poisson Process. If we now consider circular grains with random radius  $r$ , the  $\mathbf{E}[V_j(\Theta_0)]$ ,  $j \in 1, 2$  are proportional to the first two moments  $\mathbf{E}[r^2(\Theta_0)]$  and  $\mathbf{E}[r(\Theta_0)]$  of the typical disc  $\Theta_0$ . Because of indeterminacy - as long as one respects the restrictions discussed in subsection 3.2.2 - there is still a large freedom to choose the grain distribution  $Q$  which is simply the distribution of the random radius  $r$  in this case. The choice first suggesting itself is the Gamma-distribution, hence we assume

$$r \sim \text{Gamma}(\gamma_r, \lambda_r), \quad (3.12)$$

where  $\gamma_r, \lambda_r > 0$  and the circular grains are independent.

Of course there are plenty of other appropriate choices. Especially the choice of a distribution with finite support might be considered if one knows for example that the real grains of the structure to be modelled are of some characteristic size. With the assumption 3.12 made, our final model consists of only three parameters, namely  $\gamma, \gamma_r$  and  $\lambda_r$  which all must be positive. Thus the positivity of  $\gamma_r$  and  $\lambda_r$  might impose additional restrictions on the attainability of specific Minkowski functionals  $\bar{V}_j(\Theta)$  coming from the Gamma-distribution but this is not the case as soon as the Jensen inequality for the radius is fulfilled and the intensity of the underlying Poisson Process is positive. Therefore the choice of the Gamma-distribution in fact does not impose any further restrictions. This makes the choice of the Gamma-distribution a very flexible one. Using the inverted Miles-equations 3.2.4 and the expressions for the Gamma-moments we have:

$$\begin{aligned} \gamma &= \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \cdot \frac{\bar{s}^2(\Theta)}{(1 - \bar{v}(\Theta))^2}, \\ \mathbf{E}[r^2] &= \frac{-4 \log(1 - \bar{v}(\Theta)) \cdot (1 - \bar{v}(\Theta))^2}{4\pi \cdot (1 - \bar{v}(\Theta)) \cdot \bar{\chi}(\Theta) + \bar{s}^2(\Theta)} \stackrel{!}{=} \frac{\gamma_r \cdot (\gamma_r + 1)}{\lambda_r^2}, \\ \mathbf{E}[r] &= \frac{2\bar{s}(\Theta) \cdot (1 - \bar{v}(\Theta))}{4\pi \cdot (1 - \bar{v}(\Theta)) \cdot \bar{\chi}(\Theta) + \bar{s}^2(\Theta)} \stackrel{!}{=} \frac{\gamma_r}{\lambda_r}, \end{aligned}$$

and hence the algorithm can be summarized as follows:

**Algorithm 3.3.1. (stationary circular 2D-Boolean Model)** To simulate a stationary 2D-Boolean Model with circular grains,  $\Gamma$ -distributed radius and predetermined specific Minkowski functionals  $\bar{V}_j(\Theta)$  in an observation window  $W$

- i). generate inside  $W$  a 2D-stationary Poisson Process  $X = \{x_1, x_2, \dots\}$  with intensity  $\gamma$ , where:

$$\gamma = \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \cdot \frac{\bar{s}^2(\Theta)}{(1 - \bar{v}(\Theta))^2},$$

- ii). generate  $n = \gamma \cdot |W|$  independent circular grains  $\Theta_i = B(x_i, r_i)$ , where  $r_i \sim \Gamma(\gamma_r, \lambda_r)$  with:

$$\gamma_r = \frac{\mathbf{E}[r(\Theta)]^2}{\mathbf{E}[r^2(\Theta)] - \mathbf{E}[r(\Theta)]^2} = \frac{4 \cdot \bar{s}^2(\Theta) \cdot (1 - \bar{v}(\Theta))^2}{(-4 \log[1 - \bar{v}(\Theta)](1 - \bar{v}(\Theta))^2) \cdot (4\pi \cdot (1 - \bar{v}(\Theta)) \cdot \bar{\chi}(\Theta) + \bar{s}^2(\Theta)) - 4s^2(\Theta)(1 - \bar{v}(\Theta))^2},$$

$$\lambda_r = \frac{\mathbf{E}[r(\Theta)]}{\mathbf{E}[r^2(\Theta)] - \mathbf{E}[r(\Theta)]^2} = \frac{(2 \cdot \bar{s}(\Theta) \cdot (1 - \bar{v}(\Theta))) \cdot (4\pi \cdot (1 - \bar{v}(\Theta)) \cdot \bar{\chi}(\Theta) + \bar{s}^2(\Theta))}{(-4 \log[1 - \bar{v}(\Theta)](1 - \bar{v}(\Theta))^2) \cdot (4\pi \cdot (1 - \bar{v}(\Theta)) \cdot \bar{\chi}(\Theta) + \bar{s}^2(\Theta)) - 4s^2(\Theta)(1 - \bar{v}(\Theta))^2}.$$

**Remark 3.3.2.** In step i) of course we have the problem of *edge effects*. If we simulate the Poisson grain only in the observation window  $W$  we neglect those Boolean spheres  $\Theta_i$  intersecting  $W$  whose center lies outside  $W$  but is still close to its boundary. This would obviously result in severe errors when calculating the (estimated) specific Minkowski-functionals  $\bar{V}(\Theta)$  using such a realization in the end. Thus in practice we *must* simulate the Poisson-germs in an enlarged window  $\tilde{W}$  which in the circular/spherical case adds at least a margin of a sphere radius to the dimensions of  $W$  on each side. In the elliptical/ellipsoidal we add at least the size of the semi-major axis. The same problem occurs for all the simulation algorithms for the Boolean Model in the following and can be accounted for in the same way. We will not mention the problem again for each algorithm in the following.

### 3.3.2 Simulation of the 2D-Elliptical Boolean Model

Relaxing the circular grain-assumption and allowing for elliptical grains makes the model more powerful in the sense that the attainable region of specific Minkowski-values is extended (see section 3.2.2). If the eccentricity of the elliptical grains is chosen to be small the results of the elliptical model are of course similar to the ones generated by the circular model but by choosing the eccentricity large one can generate quite different looking realizations consisting of long, thin "needles" or fibres. The main drawback of the elliptical model is the fact that the circumference of an ellipse cannot be calculated analytically but only through numerical approximation of the *complete elliptic integral of the second kind*. The approximation errors typically become larger for high eccentricities ( $\epsilon \approx 1$ ). With a typical intensity of  $\gamma = 10^{-3}$  in a  $1000 \times 1000$  pixel observation window  $W$  about  $10^3$  ellipses have to be placed and the approximation errors for the circumference in the calculation of  $\bar{s}$  may quickly become relevant. This accumulating error-effect of course gets much worse in 3-dimensions where in a  $1000^3$ -pixel-sized observation window with intensity  $\gamma = 10^{-3}$  the surface area of one million ellipsoids is involved. In 3D the same problem occurs also for the mean-curvature because  $M$  for the ellipsoid as well cannot be calculated analytically but only by approximation of an elliptic integral. An other consequence of this effect shows up in discretization of elliptical/ellipsoidal Boolean Models: To explicitly draw the ellipse/ellipsoid its semi-major axes have to be computed from its Minkowski-functionals which also has to be done numerically and which again is subject to approximation error. In general to find the axes for each ellipse/ellipsoid a small numerical approximation problem has to be solved which prolongates computation time considerably compared to the circular/spherical models. This can be avoided by choosing a homothetic Boolean Model, where all the grains are dilated or shrunk version of the same prototype-ellipsoid. In this case there is a only one single axes calculation to perform whose computational cost is negligible. The restriction to the homothetic

model can also be justified by the fact that grains in real granular porous structures often can be thought of as a product of minor random modifications of the same basic shape. Therefore in 3D the homothetic model will be the model of our choice in the following.

However in the 2D-situation we implement the general (non-homothetic) stationary Boolean Model with elliptical grains. In our algorithm the typical grain  $\Theta_0$  is built by random expansion of a prototype ellipse:

$$\Theta_0 = c \cdot \Theta_{00},$$

where  $c$  is a random expansion factor which we assume to follow a Gamma-distribution:

$$c \sim \text{Gamma}(\gamma_c, \lambda_c), \quad \gamma_c, \lambda_c > 0.$$

But opposed to the homothetic situation in Theorem 3.2.8 the prototype ellipse is random itself. We assume  $\Theta_{00}$  to have unit volume, ie.  $v(\Theta_{00}) = 1$  and random circumference  $s(\Theta_{00})$  independent of  $c$ . Of course the isoperimetric inequality (Theorem 3.2.6) requires that  $s \geq 2\sqrt{\pi} := s_{\text{circ}}$ , where  $s_{\text{circ}}$  is the circumference of a circle with unit volume. We assume  $s(\Theta_{00})$  to follow a translated Gamma-distribution starting at  $s_{\text{circ}}$ :

$$s(\Theta_{00}) := s_{\text{circ}} + \Delta, \quad \Delta \sim \text{Gamma}(\gamma_\Delta, \lambda_\Delta), \quad \gamma_\Delta, \lambda_\Delta \geq 0.$$

Our final model then consists of 5 parameters  $\gamma, \gamma_c, \lambda_c, \gamma_\Delta$  and  $\lambda_\Delta$  but from the inverted Miles-equations 3.2.4 we only get three conditions. Hence the model is underdetermined and we are free to choose 2 parameters ourselves. A meaningful parameter to tune the model is  $\mathbf{E}[\Delta]$  which allows us to control the mean eccentricity of the grains. It turns out that the second parameter which should be controlled is  $\lambda_\Delta$ . This parameter has influence on the variability of the ellipsoidal shapes which are present in the model.  $\lambda_\Delta$  should be chosen big enough for two reasons: Firstly in real structures often all of the grains are of rather similar shape and it is very unlikely that a model with too much shape variability will fit the data well and secondly one should prevent the model to choose ellipses which are too eccentric because the numerical axes calculation considerably loses precision for very eccentric ellipses.

With the above assumptions one now finds for the mean Minkowski functionals of the typical grain:

$$\begin{aligned} \mathbf{E}[v(\Theta_0)] &= \mathbf{E}[c^2] = \frac{\gamma_c + \gamma_c^2}{\lambda_c^2} \\ \mathbf{E}[s(\Theta_0)] &= \mathbf{E}[c] \cdot \mathbf{E}[s(\Theta_{00})] = \frac{\gamma_c}{\lambda_c} \cdot (s_{\text{circ}} + \mathbf{E}[\Delta]). \end{aligned}$$

Hence once  $\mathbf{E}[\Delta]$  is determined,  $\gamma_c$  and  $\lambda_c$  can be calculated from the above equations. Furthermore  $\gamma_\Delta = \lambda_\Delta \cdot \mathbf{E}[\Delta]$  and the expression for the Poisson intensity  $\gamma$  is the same as in the circular case and follows directly from the inverted Miles-equations 3.2.4. Finally to draw the ellipses for each ellipse *one* of the semi-axes  $a$  and  $b$  has to be computed and the other one follows from the condition  $1 \stackrel{!}{=} v(\Theta_{00}) = \pi \cdot a \cdot b$ . We now summarize all these steps in the following algorithm:

**Algorithm 3.3.3. (stationary elliptical 2D-Boolean Model)** To simulate a stationary 2D-Boolean Model with predetermined specific Minkowski functionals  $\bar{V}_j(\Theta)$  and

elliptical grains in an observation window  $W$  with grains of the form  $\Theta_0 = c \cdot \Theta_{00}$ , where  $\Theta_{00}$  has unit volume and  $c$  and  $s(\Theta_{00}) - 2\sqrt{\pi}$  are Gamma-distributed,

- i). choose a value for the mean excess circumference  $\mathbf{E}[\Delta]$  and  $\lambda_\Delta$  to regulate the mean eccentricity of the elliptical grains and its variability and calculate mean volume  $\mathbf{E}[v(\Theta_0)]$  and mean circumference  $\mathbf{E}[s(\Theta_0)]$  of the typical grain by applying the inverted Miles-equations 3.2.4.
- ii). generate inside  $W$  a 2D-stationary Poisson Process  $X = \{x_1, x_2, \dots\}$  with intensity  $\gamma$ , where:

$$\gamma = \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \cdot \frac{\bar{s}^2(\Theta)}{(1 - \bar{v}(\Theta))^2},$$

- ii). for  $n = \gamma \cdot |W|$  generate  $n$  independent realizations  $c_i$  of the expansion factor  $c$  and  $n$  independent realizations  $s_i$  of the circumference  $s(\Theta_{00})$  of the prototype ellipse:

$$\begin{aligned} c_i &\sim \Gamma(\gamma_c, \lambda_c) \text{ iid.}, i \in \{1, \dots, n\} \\ s_i &\sim 2 \cdot \sqrt{\pi} + \Gamma(\gamma_\Delta, \lambda_\Delta) \text{ iid.}, i \in \{1, \dots, n\} \end{aligned}$$

where the parameters are determined in this order:

$$\begin{aligned} \gamma_c &= \frac{\mathbf{E}[s(\Theta_0)]^2}{\mathbf{E}[v(\Theta_0)] \cdot (2\sqrt{\pi} + \mathbf{E}[\Delta])^2 - \mathbf{E}[s(\Theta_0)]^2}, \\ \lambda_c &= \sqrt{\frac{\gamma_c + \gamma_c^2}{\mathbf{E}[v(\Theta_0)]}}, \\ \gamma_\Delta &= \lambda_\Delta \cdot \mathbf{E}[\Delta]. \end{aligned}$$

- iii). calculate the semi-major axes  $a_i$  of the  $n$  elliptical grains  $\Theta_i(x_i, a_i, b_i)$  by solving numerically for each  $i \in \{1, \dots, n\}$  the following equation for  $a_i$ :

$$s_i = 4 \cdot a_i \int_0^{\frac{\pi}{2}} \sqrt{1 - \left(1 - \frac{1}{\pi^2 \cdot a_i^4}\right) \sin^2 \phi} d\phi = 4 \cdot E \left[ 1 - \frac{1}{\pi^2 \cdot a_i^4} \right], \quad (3.13)$$

where  $E[\cdot]$  is the complete elliptic integral of the second kind (see Abramowitz and Stegun 1972, p. 590, eq. 17.3.7). Set  $b_i = 1/(\pi \cdot a_i)$  and draw the grains  $\Theta(x_i, a_i, b_i)$  centered in the points  $x_i$  of the stationary Poisson Process  $X$  in random orientation.

The random orientation in step iii). can trivially be achieved by simulating the polar angle  $\phi$  of the large semi-major axis  $a$  of the ellipses, i.e.  $\phi \sim \text{Unif}[0, 2\pi]$ .

Please note that opposed to the circular model from section 3.3.1 in the elliptical case the Gamma-assumption *does* impose additional restrictions on the attainability of specific Minkowski-functionals for a given mean eccentricity. In step ii) the denominator of the expression for  $\gamma_c$  can be negative. But in practice this problem can always be overcome

by increasing  $\mathbf{E}[\Delta]$ . Step iii). of the above algorithm is the most subtle one with respect to approximation errors and time, however, in principle the approximation problem is trivial because for fixed volume the circumference  $s(\Theta_{00})$  is a convex and monotone function of the ellipse eccentricity  $\epsilon = \sqrt{1 - b^2/a^2}$ . Equation 3.13 follows trivially from the parametrization  $x = a \cdot \cos \phi$ ,  $y = b \cdot \sin \phi$ ,  $0 \leq \phi \leq 2\pi$  and:

$$\begin{aligned} s(\Theta_{00}) &= \int_0^{2\pi} ds = \int_0^{2\pi} \sqrt{(dx)^2 + (dy)^2} = \int_0^{2\pi} \sqrt{\left(\frac{dx}{d\phi}\right)^2 + \left(\frac{dy}{d\phi}\right)^2} d\phi \\ &= \int_0^{2\pi} \sqrt{a^2 \sin^2 \phi + b^2 \cos^2 \phi} d\phi = 4a \cdot \int_0^{\frac{\pi}{2}} \sqrt{1 - \left(1 - \frac{b^2}{a^2}\right) \cos^2 \phi} d\phi \\ &= 4a \int_0^{\frac{\pi}{2}} \sqrt{1 - \left(1 - \frac{b^2}{a^2}\right) \sin^2 \phi} d\phi \stackrel{1=\pi ab}{=} 4a \int_0^{\frac{\pi}{2}} \sqrt{1 - \left(1 - \frac{1}{\pi^2 \cdot a^4}\right) \sin^2 \phi} d\phi. \end{aligned}$$

### 3.3.3 Simulation of the 3D-Spherical-Boolean Model

Also in three dimensions our goal is to simulate Boolean Models with predetermined values for the specific Minkowski-characteristics. In 3D there is an additional Minkowski-functional, namely the integral of mean-curvature  $M$  (see subsection 2.2.4 for a discussion). Starting from the preset values  $(\bar{v}(\Theta), \bar{s}(\Theta), \bar{M}(\Theta), \bar{\chi}(\Theta))$  we can directly compute the intensity  $\gamma$  of the underlying stationary Poisson Process and the mean values  $\mathbf{E}[V_j(\Theta_0)]$ ,  $j \in \{1, 2, 3\}$  for volume, surface and integral of mean curvature of the typical grain using the inverted 3D-Miles-equations from Corollary 3.2.5. Because the grains are spherical, the grain distribution is fixed as soon as the distribution of the random radius  $r$  is fixed. Because the  $\mathbf{E}[V_j(\Theta_0)]$ ,  $j \in \{1, 2, 3\}$  correspond to the first three moments of the radius distribution, we therefore need to find a distribution with  $\mathbf{E}[r]$ ,  $\mathbf{E}[r^2]$  and  $\mathbf{E}[r^3]$  fixed. A natural choice of a three-parametric distribution is the translated Gamma-distribution, ie.

$$r = \delta + R, \quad R \sim \text{Gamma}(\alpha, \beta), \quad \alpha, \beta, \delta > 0.$$

It should be mentioned that opposed to the 2D-setting, in 3D the choice of the (translated) Gamma-distribution *does* impose additional restrictions on the attainability of specific Minkowski-functionals within the 3D-Spherical-Boolean Model. It has turned out that these additional constraints are rather severe and in many situations other (often less comfortable) radial distributions must be considered. Again a distribution with constant support may be considered if the real structural grains are of a characteristic size.

However, with our assumptions we can express:

$$\begin{aligned} \mathbf{E}[v(\Theta_0)] &= \frac{4\pi}{3} \mathbf{E}[r^3] = \frac{4\pi}{3} \mathbf{E}[(\delta + R)^3] \\ &= \frac{4\pi}{3} \left( \frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} + \frac{3\delta\alpha(\alpha+1)}{\beta^2} + 3\frac{\delta^2\alpha}{\beta} + \delta^3 \right). \\ \mathbf{E}[s(\Theta_0)] &= 4\pi \mathbf{E}[r^2] = 4\pi \mathbf{E}[(\delta + R)^2] = 4\pi \left( \frac{\alpha(\alpha+1)}{\beta^2} + \frac{2\delta\alpha}{\beta} + \delta^2 \right). \\ \mathbf{E}[M(\Theta_0)] &= 4\pi \mathbf{E}[r] = 4\pi \mathbf{E}[\delta + R] = 4\pi \left( \delta + \frac{\alpha}{\beta} \right). \end{aligned} \tag{3.14}$$

This is an equation system for  $\alpha, \beta, \delta$  which can be solved analytically by removing the constants:

$$\tilde{v} := \frac{3 \cdot \mathbf{E}[v(\Theta_0)]}{4\pi}, \quad \tilde{s} := \frac{\mathbf{E}[s(\Theta_0)]}{4\pi}, \quad \tilde{m} := \frac{\mathbf{E}[M(\Theta_0)]}{4\pi},$$

and introducing the new variables:

$$x := \frac{1}{\beta}, \quad y := \frac{\alpha}{\beta}.$$

Thus the whole simulation procedure can be summarized in the following algorithm:

**Algorithm 3.3.4. (stationary spherical 3D-Boolean Model)** To simulate a stationary 3D-Boolean Model with spherical grains, a radius which follows a translated Gamma-distribution and predetermined specific Minkowski functionals  $\bar{V}_j(\Theta)$ ,  $j \in \{0, \dots, 3\}$  in an observation window  $W$

- i). use the 3D-inverted-Miles-equations from Corollary 3.2.5 to calculate the mean volume  $\mathbf{E}[v(\Theta_0)]$ , the mean surface  $\mathbf{E}[s(\Theta_0)]$  and the mean integral curvature  $\mathbf{E}[M(\Theta_0)]$  for the typical grain  $\Theta_0$ .
- ii). generate inside  $W$  a 3D-stationary Poisson Process  $X = \{x_1, x_2, \dots\}$  with intensity  $\gamma$ , where:

$$\gamma = \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \frac{\bar{M}(\Theta)\bar{s}(\Theta)}{(1 - \bar{v}(\Theta))^2} + \frac{\pi}{192} \frac{\bar{s}(\Theta)^3}{(1 - \bar{v}(\Theta))^3}.$$

- iii.) generate  $n = \gamma \cdot |W|$  independent spherical grains  $\Theta_i = B(x_i, r_i)$ , where the radii  $r_i \sim \delta + \Gamma(\alpha, \beta)$ . The parameters  $\alpha, \beta, \delta$  can be determined by performing the following calculations in this order:

1. Calculate

$$\tilde{v} = \frac{3 \cdot \mathbf{E}[v(\Theta_0)]}{4\pi}, \quad \tilde{s} = \frac{\mathbf{E}[s(\Theta_0)]}{4\pi}, \quad \tilde{m} = \frac{\mathbf{E}[M(\Theta_0)]}{4\pi}.$$

2. Calculate

$$y = \frac{2 \cdot (\tilde{s} - \tilde{m}^2)^2}{\tilde{v} - 3 \cdot \tilde{m} \cdot (\tilde{s} - \tilde{m}^2) - \tilde{m}^3}, \quad x = \frac{\tilde{s} - 2 \cdot (\tilde{m} - y) \cdot y - (\tilde{m} - y)^2 - y^2}{y}.$$

3. Calculate

$$\beta = \frac{1}{x}, \quad \alpha = y \cdot \beta, \quad \delta = \tilde{m} - y.$$

The calculations in step iii.) of the above algorithm stem from a step by step-solution of equation-system 3.14. Of course the algorithm only works if the parameters  $\alpha, \beta, \gamma, \delta$  are all non-negative. If not, as long as  $\gamma > 0$ , a change of the radius distribution may be considered.

### 3.3.4 Simulation of the 3D-Homothetic Ellipsoidal-Boolean Model

As far as the 3D-ellipsoidal model is concerned we restrict ourselves to the homothetic case, ie. all Boolean grains  $\Theta_0$  are scaled versions of a single *fixed* prototype ellipsoid  $\Theta_{00}$ . The main reason for this is that the ellipsoid surface  $s$  and its integral of mean curvature  $M$  cannot be calculated analytically but only expressed as elliptic integrals. The Minkowski-functionals for a general triaxial ellipsoid  $E$  with semi-major axes  $a \geq b > c > 0$  are:

$$\begin{aligned}
 v(E) &= \frac{4}{3} \cdot \pi \cdot a \cdot b \cdot c \\
 s(E) &= 2\pi c^2 + \frac{2\pi b}{\sqrt{a^2 - c^2}} \cdot \left[ c^2 \cdot F \left( \arcsin \left( \frac{a^2 - c^2}{a} \right), \frac{a^2 \cdot (b^2 - c^2)}{b^2 \cdot (a^2 - c^2)} \right) + \right. \\
 &\quad \left. (a^2 - c^2) \cdot E \left( \arcsin \left( \frac{a^2 - c^2}{a} \right), \frac{a^2 \cdot (b^2 - c^2)}{b^2 \cdot (a^2 - c^2)} \right) \right], \\
 M(E) &= 2\pi \cdot \left[ \frac{bc}{a} + \frac{c^2}{\sqrt{a^2 - c^2}} \cdot F \left( \arcsin \left( \frac{a^2 - c^2}{a} \right), \frac{a^2 - b^2}{a^2 - c^2} \right) \right. \\
 &\quad \left. + \sqrt{a^2 - c^2} \cdot E \left( \arcsin \left( \frac{a^2 - c^2}{a} \right), \frac{a^2 - b^2}{a^2 - c^2} \right) \right], \tag{3.15}
 \end{aligned}$$

where  $F$  and  $E$  are the *Elliptic integrals of the first and second kind* respectively which are defined as follows (see also (Abramowitz and Stegun 1972), Ch. 17 for a complete survey of elliptic integrals. Especially formulae 17.2.6, 17.2.8):

$$F(\phi, m) = \int_0^\phi \frac{1}{\sqrt{1 - m \cdot \sin^2 \theta}} d\theta, \quad E(\phi, m) = \int_0^\phi \sqrt{1 - m \cdot \sin^2 \theta} d\theta.$$

Because numerical approximation is needed for each ellipsoid when calculating the Minkowski-functionals  $S$  and  $M$  from the semi-major axes  $a, b, c$  or vice versa and the number of ellipsoids in reasonably sized 3D-samples may be immense, accumulating approximation error is a serious problem for the 3D-ellipsoidal Boolean Model. Also the myriad of small approximation steps needed to calculate the ellipsoid-axes when drawing the ellipsoids slows down the algorithms considerably. The homothetic model has the advantage that only a *single* approximation step is needed for the prototype ellipsoid  $\Theta_{00}$ . Of course this single approximation must be performed as accurate as possible because the approximation error is transferred to the other grains by scaling. From the modelling point of view, restriction to the homothetic case is often no serious issue because the grains in many real material structures have a typical basic grain shape with relatively small variation compared to the variation in grain size. Hence we assume in the following that the typical ellipsoidal grain  $\Theta_0$  is generated by expanding a *non-random* prototype-ellipsoid  $\Theta_{00}$  with a random expansion factor  $c$ :

$$\Theta_0 = c \cdot \Theta_{00},$$

where  $v(\Theta_{00}) = 1$ ,  $s(\Theta_{00}) = s_{00}$  and  $m(\Theta_{00}) = m_{00}$  and where  $s_{00}$  and  $m_{00}$  are fixed constants. For the expansion factor  $c$  we assume a Gamma-distribution:

$$c \sim \text{Gamma}(\gamma_c, \lambda_c), \quad \gamma_c, \lambda_c > 0.$$

It is a fact that both the restriction to the homothetic model as well as the choice of the Gamma-distribution for  $c$  impose serious restrictions on the attainable region



of Minkowski-values. These can be partially overcome by dropping the homotheticity assumption and choosing  $s_{00}$  and  $m_{00}$  random obeying the isoperimetric inequalities from 3.2.6. However, because of the problems mentioned above we will not do that in the following. We would also like to mention that instead of fixing distributions for  $s_{00}$  and  $m_{00}$  one could also fix distributions for the semi-major axes  $a$ ,  $b$  and  $c$ . This would bypass the time-consuming numerical axes-calculation in the non-homothetic model. But then relating the distributional parameters to the the mean Minkowski-functionals of the typical grain becomes more difficult and furthermore we could not reach the same level of accuracy in reproducing preset values for the specific Minkowski-characteristics with this simpler procedure as with the algorithm described below.

From the inverted Miles-equations given in Corollary 3.2.5 we can calculate the intensity  $\gamma$  of the underlying stationary Poisson Process and the mean Minkowski-characteristics  $\mathbf{E}[v(\Theta_0)]$ ,  $\mathbf{E}[s(\Theta_0)]$  and  $\mathbf{E}[M(\Theta_0)]$ . The latter can be expressed by the additional model parameters  $\gamma_c, \lambda_c, s_{00}, m_{00}$  as follows:

$$\begin{aligned}\mathbf{E}[v(\Theta_0)] &= \mathbf{E}[c^3] = \frac{\gamma_c \cdot (\gamma_c + 1) \cdot (\gamma_c + 2)}{\lambda_c^3}, \\ \mathbf{E}[s(\Theta_0)] &= s_{00} \cdot \mathbf{E}[c^2] = s_{00} \cdot \frac{\gamma_c \cdot (\gamma_c + 1)}{\lambda_c^2}, \\ \mathbf{E}[M(\Theta_0)] &= m_{00} \cdot \mathbf{E}[c] = m_{00} \cdot \frac{\gamma_c}{\lambda_c}.\end{aligned}\tag{3.16}$$

These are three equations for four model parameters ( $s_{00}, m_{00}, \gamma_c, \lambda_c$ ) hence we may choose one of these parameters free. The most meaningful parameter is  $s_{00}$  which we can use to tune the (deterministic!) eccentricity of the ellipsoidal grains. However by the isoperimetric inequality (Theorem 3.2.6) we must have  $s_{00} \geq \sqrt[3]{36\pi} \approx 4.84$ . Equation system 3.16 is readily solved and we can summarize the whole simulation procedure in the following algorithm:

**Algorithm 3.3.5. (stationary homothetic ellipsoidal 3D-Boolean Model)** To simulate a stationary homothetic 3D-Boolean Model with predetermined specific Minkowski functionals  $\bar{V}_j(\Theta)$  and ellipsoidal grains in an observation window  $W$  with grains of the form  $\Theta_0 = c \cdot \Theta_{00}$ , where  $\Theta_{00}$  is a *fixed deterministic* prototype-ellipsoid of unit volume, fixed surface area  $s_{00}$  and fixed integral mean curvature  $m_{00}$  and  $c \sim \text{Gamma}(\gamma_c, \lambda_c)$  is a Gamma-distributed scaling factor,

- i). use the 3D-inverted-Miles-equations from Corollary 3.2.5 to calculate the mean volume  $\mathbf{E}[v(\Theta_0)]$ , the mean surface  $\mathbf{E}[s(\Theta_0)]$  and the mean integral curvature  $\mathbf{E}[M(\Theta_0)]$  for the typical grain  $\Theta_0$  and choose a value for the surface area  $s_{00}$  of the prototype-ellipsoid.
- ii). generate inside  $W$  a 3D-stationary Poisson Process  $X = \{x_1, x_2, \dots\}$  with intensity  $\gamma$ , where:

$$\gamma = \frac{\bar{\chi}(\Theta)}{1 - \bar{v}(\Theta)} + \frac{1}{4\pi} \frac{\bar{M}(\Theta)\bar{s}(\Theta)}{(1 - \bar{v}(\Theta))^2} + \frac{\pi}{192} \frac{\bar{s}(\Theta)^3}{(1 - \bar{v}(\Theta))^3}.$$

- iii). For  $n = \gamma \cdot |W|$  generate  $n$  independent realizations  $c_i$  of the expansion factor  $c \sim \Gamma(\gamma_c, \lambda_c)$  where  $\gamma_c$  is a (positive) solution of the quadratic equation

$$\left(s_{00}^3 \cdot \mathbf{E}[v(\Theta_0)]^2 - \mathbf{E}[s(\Theta_0)]^3\right) \gamma_c^2 + \left(s_{00}^3 \cdot \mathbf{E}[v(\Theta_0)]^2 - 4 \cdot \mathbf{E}[s(\Theta_0)]^3\right) \cdot \gamma_c - 4 \cdot \mathbf{E}[s(\Theta_0)]^3 = 0,$$

and

$$\lambda_c = \sqrt{\frac{(\gamma_c + 1) \cdot \gamma_c \cdot s_{00}}{\mathbf{E}[s(\Theta_0)]}}.$$

- iv.) Set  $m_{00} = \mathbf{E}[M(\Theta_0)] \frac{\lambda_c}{\gamma_c}$  and calculate the axes  $a, b, c$  of the prototype ellipsoid  $\Theta_{00}(a, b, c)$  by numerically solving the equation system 3.15. Then draw the homothetic ellipsoidal grains  $\Theta_i(x_i, a, b, c) = c_i \cdot \Theta_{00}(a, b, c)$  centered at the points  $x_i$  of the stationary Poisson Process  $X$  in random orientation.

In step iv). of the above algorithm to place the ellipsoids in random orientation three angles  $\phi, \theta, \nu$  need to be simulated to uniquely determine the position of the ellipsoids  $\Theta_i$  in space. An easy way to do this is to take the pair  $(\phi, \theta)$  as the polar angles of the largest semi-major axis  $a$  and  $\nu$  the rotation angle of the ellipsoid around  $a$  with respect to a certain reference point in the normal plane  $\Pi_a$  to  $a$  through the ellipsoid center. For example one of the intersection points of  $\Pi_a$  with the plane parallel to  $x = 0$  through the ellipsoid center could serve as a reference point (one can take the one with the lower  $z$ -coordinate to define it uniquely). Then  $\nu$  is trivial and  $(\phi, \theta) \sim \text{Unif}(S^2)$ , where  $S^2$  is the unit sphere in three dimensions, ie.

$$\phi \sim \text{Unif}[-\pi, \pi], \quad \theta \sim \arcsin(2 \cdot \text{Unif}[0, 1] - 1), \quad \nu \sim \text{Unif}[0, 2\pi].$$

Again only positive values for  $\gamma, \gamma_c, \lambda_c$  make sense, of course.

### 3.3.5 The Discretization-Process

We conclude this section with a few remarks about the discretization process. Because realizations of a Boolean Model are continuous sets in  $\mathbb{R}^d$  and the Ohser-Mücklich-estimators (subsections 2.3.1, 2.3.2) for the Minkowski-functionals only can be used on a grid we must discretize the Boolean realizations generated with the above algorithms to evaluate the simulation process. At least in three dimensions this is definitely the most-time consuming step of the simulation procedure. Discretization must be done at a reasonably high resolution such that the curved grain boundaries of the circles/ellipses can be approximated with an edge path on the grid with sufficiently small error. Please note however that discretization is *not* necessary just to define the artificial images for further calculations (eg. midpoints and radii of spheres) without plotting them and without checking for their specific Minkowski-functionals. Once a target resolution is specified (around  $1000^d$  pixels in our case), the 2D-discretization is still fast. We start with a white image, and process each circle/ellipse sequentially. For pixels  $p$  lying close enough to the grain center  $x_0$  (only pixels in the square centered at  $x_0$  and side-length equal to the perimeter  $2r$  of the grain disk or the length of the major axis  $2a$  need to be considered), we check whether they lie inside the grain and thus need to be blackened. In the circular case for example, we simply check whether the pixel's grid coordinates fulfill  $|p - x_0|^2 \leq r^2$ . Once a pixel is blackened, it stays black until discretization finishes. In three dimensions where discretization time is crucial, there are basically two possible generalizations of this 2D-procedure:

- Process the spherical/ellipsoidal Boolean grains  $\Theta_i$  sequentially and for each grain blacken the points which lie inside the Boolean grain. For a spherical grain  $B(x_0, r)$  for example the pixel with coordinates  $p$  is blackened if  $|p - x_0|^2 \leq r^2$ .

- Process the horizontal layers  $z = z_0$  sequentially. Calculate the intersection spheres/ellipses of the 3D-Boolean grains with the layers  $z = z_0$  and apply the 2D-procedure for each layer.

Although the second procedure initially involves a few simple spatial geometry calculations, it turned out in our experiments that the second method performed much faster on large 3D-images and is well worth the slightly higher implementational effort. It is even possible to carry this idea one step further and reduce the needed calculations to one dimension. The interior of an ellipsoid in general position is given by the usual inequality for a quadratic form  $(x - x_0)^T A(x - x_0) \leq 1$ , where  $A$  is a positive definite  $3 \times 3$ -matrix and  $x = (x_1, x_2, x_3)$  and  $x_0 = (x_{01}, x_{02}, x_{03})$  are 3-vectors, where  $x_0$  is the midpoint of the ellipsoid. By completing the square this can be transformed into:

$$\begin{aligned} (x - x_0)^T A(x - x_0) &= a_{11} \left( (x_1 - x_{01}) + \frac{a_{12}}{a_{11}}(x_2 - x_{02}) + \frac{a_{13}}{a_{11}}(x_3 - x_{03}) \right)^2 \\ &+ \left( a_{22} - \frac{a_{12}^2}{a_{11}} \right) \left( (x_2 - x_{02}) + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2}(x_3 - x_{03}) \right)^2 \\ &+ \left( a_{33} - \frac{a_{13}^2}{a_{11}} - \frac{(a_{11}a_{23} - a_{12}a_{13})^2}{a_{11}^2 a_{22} - a_{11}a_{12}^2} \right) (x_3 - x_{03})^2 \leq 1, \end{aligned} \quad (3.17)$$

where all the coefficients of the squares on the left hand side of the inequality are strictly positive by positive-definiteness of the matrix  $A$ . From equation 3.17 one can directly devise the following discretization-algorithm for the 3D-ellipsoidal (and spherical) Boolean Model:

**Algorithm 3.3.6. (Discretization of 3D-Ellipsoidal Boolean Model)** Let  $E$  be an ellipsoid given by  $(x - x_0)^T A(x - x_0) \leq 1$ . The points  $x = (x_1, x_2, x_3)$  falling into the interior of  $E$  can be found by successively solving one-dimensional inequalities as follows:

- i) Solve for  $x_3$ :

$$\left( a_{33} - \frac{a_{13}^2}{a_{11}} - \frac{(a_{11}a_{23} - a_{12}a_{13})^2}{a_{11}^2 a_{22} - a_{11}a_{12}^2} \right) (x_3 - x_{03})^2 \leq 1.$$

- ii) For each  $x_3$  that fulfills i). solve for  $x_2$ :

$$\begin{aligned} &\left( a_{22} - \frac{a_{12}^2}{a_{11}} \right) \left( (x_2 - x_{02}) + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2}(x_3 - x_{03}) \right)^2 \\ &\leq 1 - \left( a_{33} - \frac{a_{13}^2}{a_{11}} - \frac{(a_{11}a_{23} - a_{12}a_{13})^2}{a_{11}^2 a_{22} - a_{11}a_{12}^2} \right) (x_3 - x_{03})^2. \end{aligned}$$

- iii) For each solution  $(x_2, x_3)$  to i). and ii). solve for  $x_1$ :

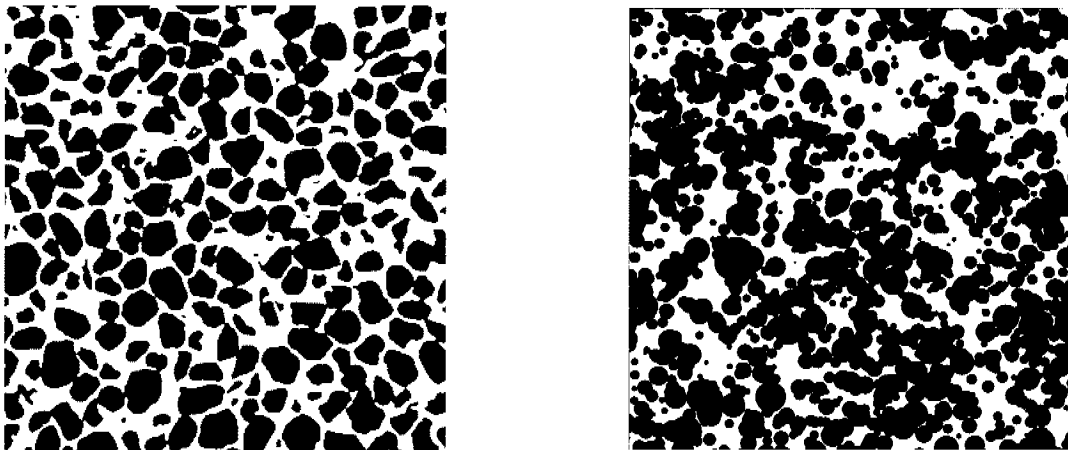
$$\begin{aligned} &a_{11} \left( (x_1 - x_{01}) + \frac{a_{12}}{a_{11}}(x_2 - x_{02}) + \frac{a_{13}}{a_{11}}(x_3 - x_{03}) \right)^2 \leq \\ &1 - \left( a_{22} - \frac{a_{12}^2}{a_{11}} \right) \left( (x_2 - x_{02}) + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2}(x_3 - x_{03}) \right)^2 \\ &- \left( a_{33} - \frac{a_{13}^2}{a_{11}} - \frac{(a_{11}a_{23} - a_{12}a_{13})^2}{a_{11}^2 a_{22} - a_{11}a_{12}^2} \right) (x_3 - x_{03})^2. \end{aligned}$$

### 3.4 2D-Results

In this section we want to apply the algorithms 3.3.1 and 3.3.3 to model 2-dimensional cross-sections as displayed in figure 1.3 of the synchrotronized coarse-sand cube given in figure 1.1. We also generate images where one of the 3 Minkowski-characteristics  $(v(\Theta), s(\Theta), \bar{\chi}(\Theta))$  is changed and the others kept fixed to illustrate the change in optical appearance of the artificial Boolean structure with respect to a single Minkowski characteristic. This also helps us to illustrate the obvious defects of the Boolean Model in modelling porous media. One should bear in mind that main reason why the Boolean Model has been intensively studied in literature is not because it generates good model fitting results but because it is analytically tractable to quite a far extent.

#### 3.4.1 Typical Results for the 2D-Spherical Boolean Model

Figure 3.8 shows a typical cross-section extracted from our data coarse-sand-cube in figure 1.1 and a realization of the stationary 2D circular Boolean Model generated with algorithm 3.3.1 which has the same specific Minkowski characteristics as the original sand cross-section up to a negligible relative error of 3% for the Euler-characteristic and an absolute error of  $< 10^{-3}$  for the other characteristics.

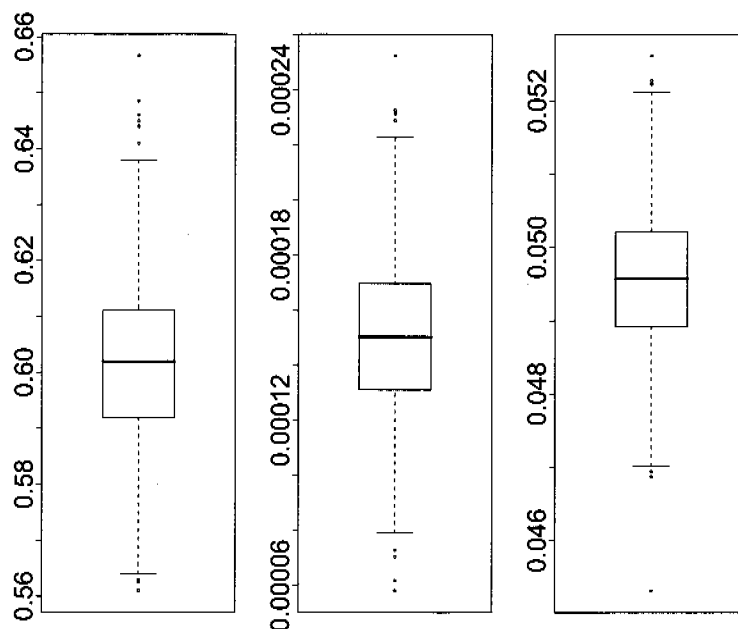


**Figure 3.8:** An  $800 \times 800$ -pixel-sized cross-section of the coarse-sand-cube from figure 1.1 with  $\bar{v} = 0.600$ ,  $s = 0.051$  and  $\bar{\chi} = 1.63 \cdot 10^{-4}$  (left) and a realization of the circular Boolean model with the same Minkowski characteristics generated according to the algorithm 3.3.1. (right)

The Boolean Model doesn't perform bad, it passably captures the basic structure of the true image. Because of indeterminacy we cannot expect a perfect result, because there are a lot of potential Boolean Models leading to the same specific Minkowski functionals, hence the Minkowski functionals on their own are simply not enough to capture the whole information stored in the true image. If one understands the Boolean Model in the sense that every model grain in the Boolean realization should stand for a true grain in the real image, then the Boolean Model has a major defect: its grains are allowed to overlap ad libitum whereas the sand grains in the real image do not overlap at all. More realistic models such as hardcore and cherry-pit don't show this defect. In the hardcore model the

circular grains don't overlap at all whereas in the cherry-pit model they are allowed to overlap only up to a certain depth. (This is the origin of the name-analogy. A cherry has a soft and penetrable pulp and a hard impenetrable pit.) The drawback of these models is that they are no longer analytically tractable and more difficult to implement. For details concerning the cherry-pit model see for example (Stoyan and Kadashevich 2005) and for the basic so-called force-biased hardcore algorithm (Moscinski, Bargiel, Rycerz and Jacobs 1989). But as figure 3.8 and subsequent figures show, if the natural grains are in average much larger than the Boolean grains, it may be that a natural grain corresponds to a whole non-circular compound of many small overlapping circular Boolean grains and hence the Boolean grains themselves may have no interpretation, they are just the constituents of the whole structure. With this viewpoint, the overlapping of the Boolean grains is no disadvantage at all, but a flexible means of building up non-circular grains.

Because in the Boolean Model only the *mean* Minkowski-characteristics  $\bar{V}_j(\Theta)$  coincide with the specific Minkowski-characteristics estimated in the true image, the Minkowski-characteristics of a particular realization only approximately correspond to the ones of the true image. Sometimes one may have to simulate several times to get a satisfactory result. It is hard to get analytic expressions for the variability of the  $V_j$  but one can get quite a good impression of it by simulation. Figure 3.9 shows boxplots for the distribution of the three Minkowski-characteristics for 1000 realizations of the Boolean Model used to generate the right image in figure 3.8:



**Figure 3.9:** The distributions of  $\hat{v}(\Theta)$ ,  $\hat{\chi}(\Theta)$  and  $\hat{s}(\Theta)$  for 1000 realizations of the stationary circular 2D-Boolean Model with  $\bar{v}(\Theta) = 0.600$ ,  $\bar{\chi}(\Theta) = 1.63 \cdot 10^{-4}$  and  $\bar{s}(\Theta) = 0.051$ .

Figure 3.9 shows that there is considerable variation of the estimated specific Minkowski functionals and the distribution seems to be centered at the right means but it is a

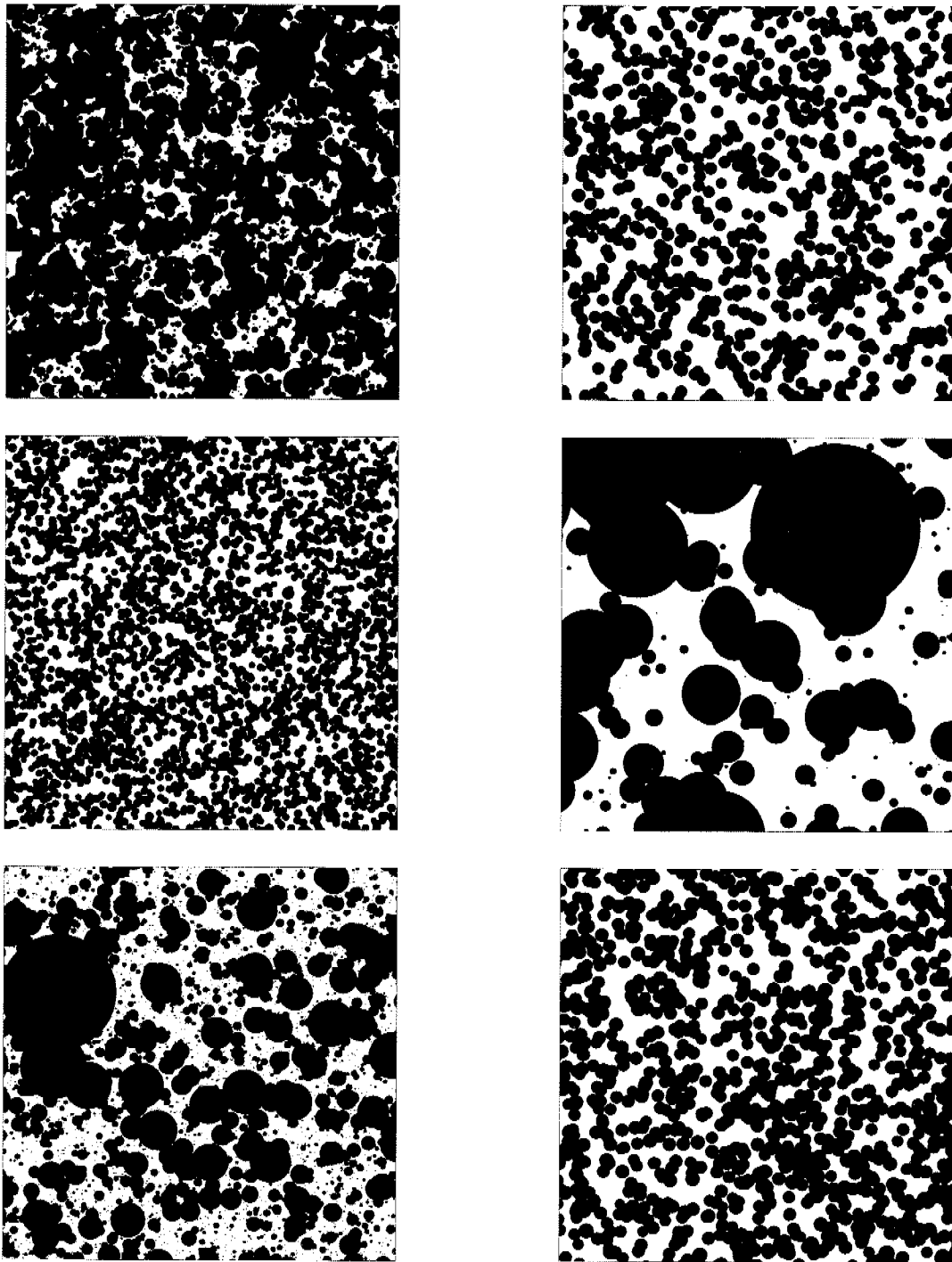
fact that the estimators  $\hat{s}$  and  $\hat{\chi}$  defined in equations 2.24 and 2.24 are biased which is not very surprising because a smooth boundary is not digitizable. The bias depends on the adjacency defined on the underlying grid. In this thesis we are always using (4,4)-adjacency which means that for both the black and white phase we use the natural adjacency given by the lattice squares such that each pixel has 4 direct neighbours. One could use other adjacencies instead, for example 8-adjacency, where also the lattice-cell diagonals are part of the edge-graph and hence each pixel has 8 direct neighbours. The surprising result found in (Ohser, Nagel and Schladitz 2003) is that for the spherical Boolean Model this bias for  $\hat{\chi}$  doesn't vanish even when the lattice spacing  $h$  tends to zero. For 3D-Spherical Boolean Models and for some reasonable adjacencies it is even asymptotically infinitely large! Please note that the realizations of the Boolean Model are not morphologically regular, and in particular not morphologically closed with respect to the segments which form the edges of the cubic lattice cells on which the Euler-characteristic is estimated. This holds for arbitrarily small such structuring segments. This phenomenon can easily be seen already for the 2D-Boolean Model with spherical grains and fixed deterministic radius  $r$ . Whenever two such spherical grains slightly overlap they form a compound which has two boundary points at which the curvature is not continuous (These are exactly the extremal points of the intersection sickle of the two spheres.) After morphological closing the compound will have grown a little near these two boundary points regardless of the compound's orientation relative to the structuring segment and hence corollary 2.1.21 cannot be applied which means that the specific Euler characteristic estimation of the Boolean realization may be erroneous, even for an arbitrarily small lattice constant  $h$ . Because such overlapping of the spherical grains occurs with positive probability it is well plausible that the OM-estimator for the Euler characteristic is asymptotically biased for  $h \downarrow 0$ .

At last in figure 3.10 we show realizations of the stationary circular Boolean Model where one of the specific Minkowski functionals is enlarged or reduced and the other two specific Minkowski-functionals agree with the ones of the original sand cross-section in figure 3.8 (left image). The Minkowski-functionals are modified towards the boundary of the attainable region as indicated in table 3.1.

Image	modified functional	modification
top left	specific volume	increased to 0.8
top right	specific volume	decreased to 0.515
middle left	specific surface	increased by factor 1.7
middle right	specific surface	decreased by factor 0.3
bottom left	specific Euler	increased by factor 100
bottom right	specific Euler	decreased by factor 0.3

**Table 3.1:** *Modifications of specific Minkowski functionals with respect to the ones of the original cross-section in figure 3.8. The specific Minkowski functionals of the reference image are  $\bar{v} = 0.600$ ,  $\bar{s} = 0.051$ ,  $\bar{e} = 1.63 \cdot 10^{-4}$ . The corresponding images are shown in figure 3.10.*

Figure 3.10 shows that for the increased volume (top left) the overlapping of the circular grains become so intensive that the singular grains can hardly be discerned any longer. The structural grains which are a result of the clustering can still be identified by eye but



**Figure 3.10:** *The influence of changing a single specific Minkowski-characteristic in the 2D-circular Boolean Model. As a reference the original sand cross-section in figure 3.8 was taken. Top row: enlarged/reduced specific volume, middle row: enlarged/reduced specific surface, bottom row: enlarged/reduced specific Euler characteristic.*

their shape is no longer circular. On the other hand the image with the reduced volume uses the Boolean grains directly as structural grains. The grains are connected via large

chains and thus form large holes to account for the low Euler characteristic. In the high surface image (middle left) the Boolean grains form thin connections to stretch out the given amount of black volume into thin fibres over the image to make the surface as large as possible including smaller holes as in the top right image because of the higher volume. In the middle right image the volume is concentrated in large spheres to reduce the surface to the required small value. The Euler characteristic is corrected upwards by introducing a lot of additional tiny spheres each of which increases the Euler characteristic by one at a very low cost of increasing the surface. The bottom left image uses the same technique to increase the Euler characteristic. The whole interspace between the few large spheres which make up volume and surface is blotched with spheres of tiniest size to screw up the Euler characteristic. Finally in the bottom right image the Boolean spheres seem to be arranged "the least random" but rather quite regularly in a grid-like fashion. It's easy to see that a grid-like arrangement of pixels achieves the least possible Euler characteristic for a given volume and surface. Reducing the Euler-Characteristic further (which is not possible for the given values of volume and surface) would reveal the grid-effect more clearly.

### 3.4.2 Typical Results for the 2D-Elliptical Boolean Model

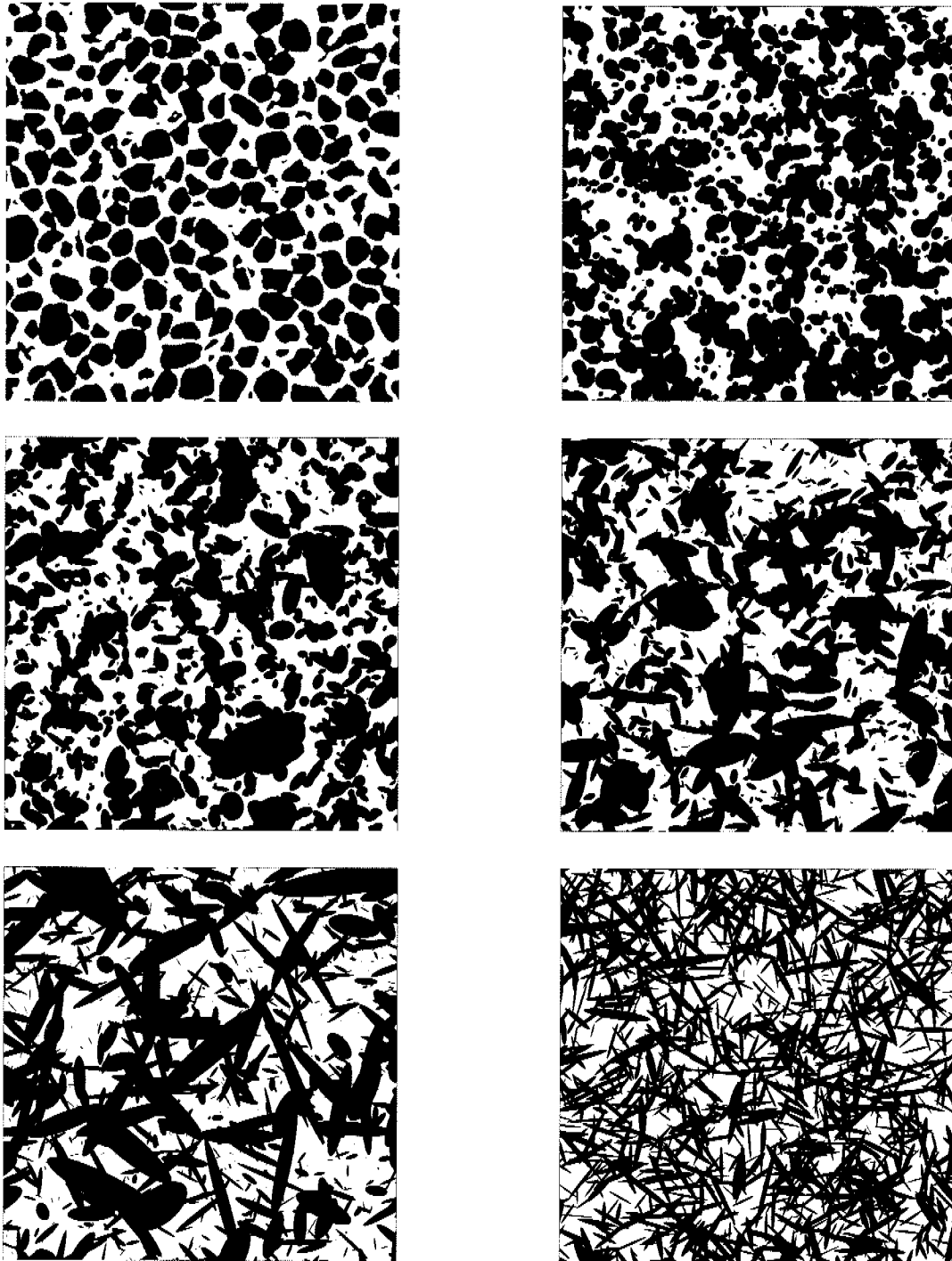
In this subsection we show a few realizations of the 2D-stationary elliptical Boolean Model. Again we use a cross-section from figure 1.1 as a benchmark. Figure 3.11 shows realizations of the stationary 2D-elliptical Boolean Model which all share their specific Minkowski-functionals with the cross-section taken from the real sand-cube in the upper left image of figure 3.8. The realizations differ in their value for the mean eccentricity parameter  $\mathbf{E}[\Delta]$  that was chosen in algorithm 3.3.3 which is varying from moderate to high eccentricity. Of course the real sand grains exhibit only moderate eccentricity and thus the large eccentricity images should of course not be interpreted as grain-to-grain reconstructions.

The images in figure 3.11 illustrate the flexibility of the elliptical Boolean Model. The granular structures in the top images are optically quite different from the fibre-like structures in the bottom images. This fact also confirms the indeterminacy phenomenon: Already within the elliptical Boolean Model a large variety of different structures with the same specific Minkowski functionals can be generated and hence a lot of important structural information is lost by summarizing an image into its three specific Minkowski-values.

Figure 3.12 illustrates the sensitivity of the artificial elliptical Boolean structure with respect to changes of a single specific Minkowski characteristic while the other two are kept fixed and coincide with the values of the top left image in figure 3.11. The changes in the model parameters that were made to produce the different images of figure 3.12 are summarized in table 3.2.

The images of figure 3.12 show similar effects as the ones in the corresponding figure 3.10 for the circular model. For the bottom right image the sign of the Euler characteristic was flipped which results in a densely interconnected network leaving small holes which give negative contributions to the Euler number. Such a low Euler characteristic can only be achieved for high eccentricities, hence the eccentricity parameter had to be chosen much higher than for the other images. We now turn our attention to the presentation of some three dimensional results.





**Figure 3.11:** An  $800 \times 800$ -pixel-sized cross-section of the sandcube from figure 1.1 with  $\bar{v} = 0.600$ ,  $\bar{s} = 0.051$  and  $\bar{\chi} = 1.63 \cdot 10^{-4}$  (top left) and several realizations of the elliptical Boolean Model with the same Minkowski characteristics generated according to algorithm 3.3.3. The eccentricity parameter was chosen to be  $\mathbf{E}[\Delta] \in \{0.1, 0.3, 0.8, 2.0, 3.7\}$  in this order from left to right and top to bottom. The variability of ellipse shapes was kept small by choosing the parameter  $\lambda_{\Delta}$  to be as large as 4.

Image	modified functional	modification
top left	specific volume	increased to 0.8
top right	specific volume	decreased to 0.3
middle left	specific surface	increased by factor 2
middle right	specific surface	decreased by factor 0.5
bottom left	specific Euler	increased by factor 100
bottom right	specific Euler	decreased by factor $-2$

**Table 3.2:** *Modifications of specific Minkowski functionals with respect to the ones of the original cross-section in figure 3.11. The specific Minkowski functionals of the reference image are  $\bar{v} = 0.600$ ,  $\bar{s} = 0.051$ ,  $\bar{\epsilon} = 1.63 \cdot 10^{-4}$ . The corresponding images are shown in figure 3.12.*

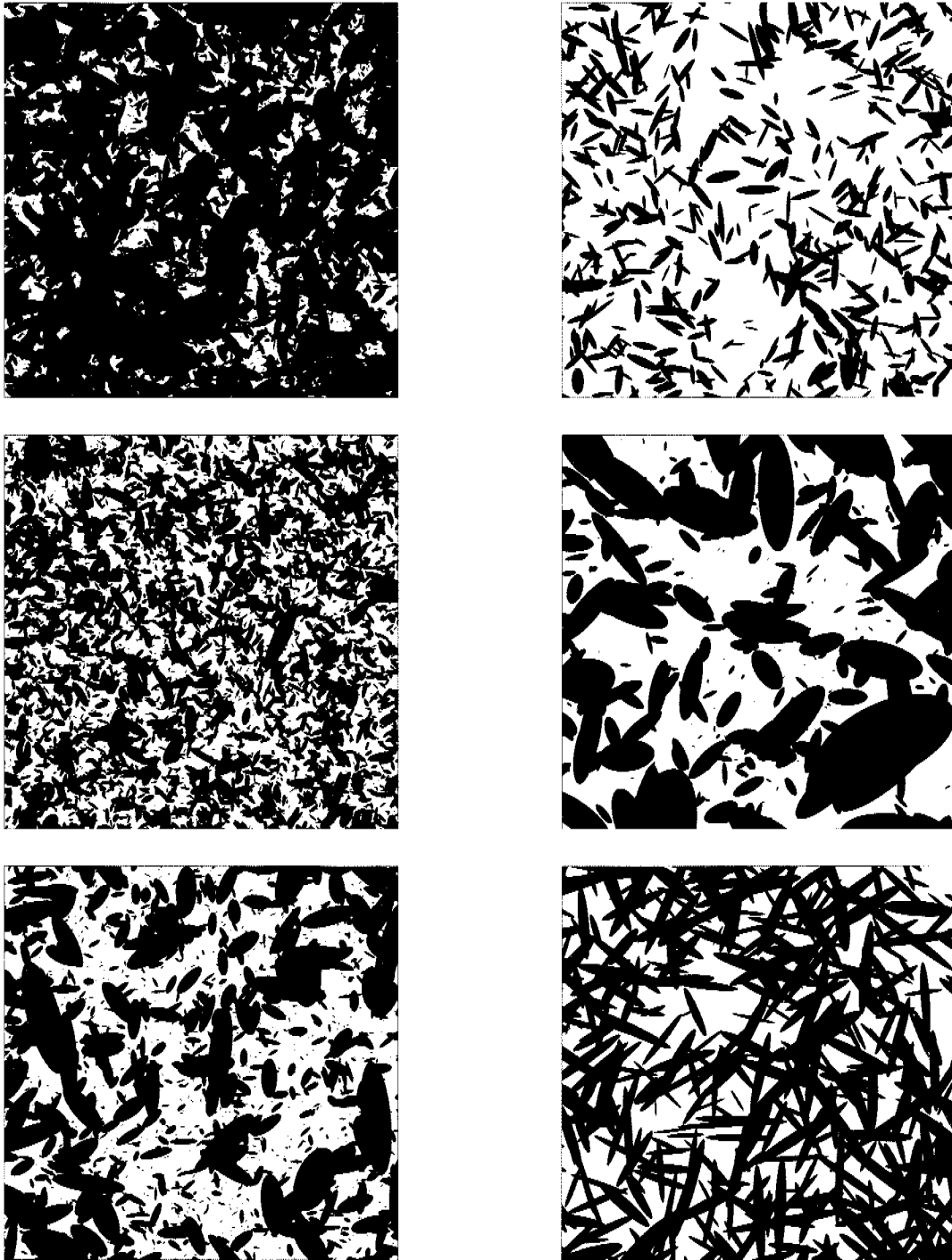
### 3.5 3D-Results

In this section we present a few realizations of three-dimensional Boolean models. The images were generated using the module-based 3D-visualization-tool AVS (Advanced Visual Systems 2006). Firstly we present some modelling results of our experiments with the spherical Boolean Model and then summarize a rather extensive study where artificial structures generated with the ellipsoidal Boolean Model and real sand-specimens with the same specific Minkowski functionals are compared with respect to their geophysical properties such as fluid-flow and permeability. A more detailed summary of this work was published in (Lehmann et al. n.d.). The study was part of the FIMOTUM (First principle based MOdelling of Transport in Unsaturated Media) project which was jointly pursued during the last few years by 4 groups of geophysicists from the Technical University of Braunschweig, the University of Stuttgart and the Swiss Federal Institute of Technology (2 groups) (visit the FIMOTUM-website for more detailed information [www.fimotum.stroemungsmechanik.org](http://www.fimotum.stroemungsmechanik.org)).

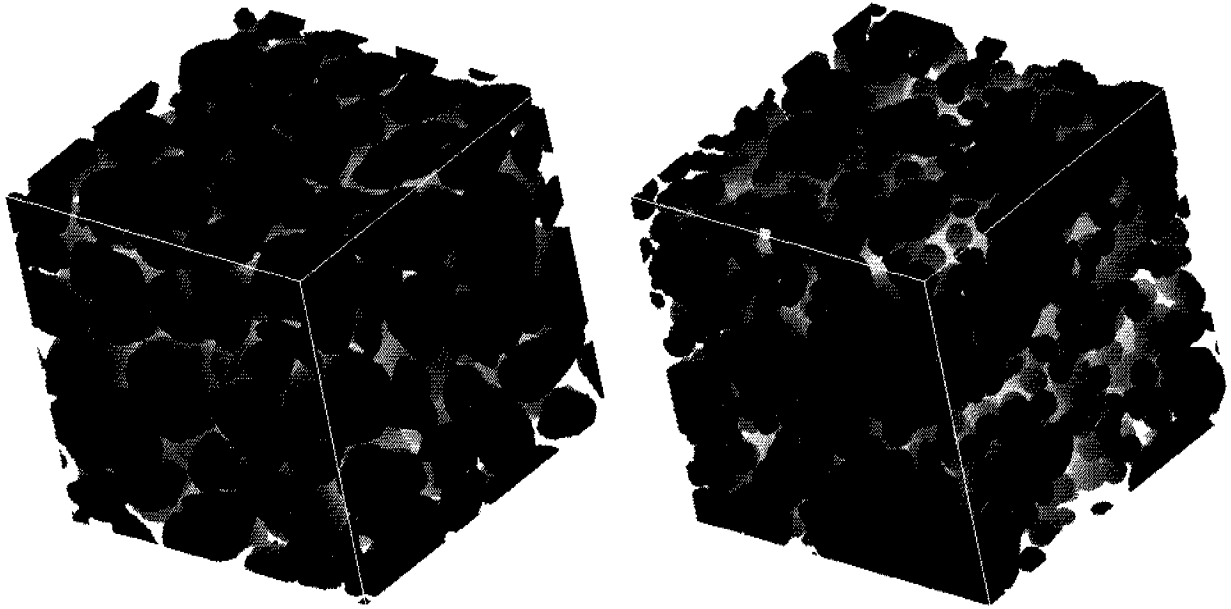
#### 3.5.1 Typical Results for the 3D-Spherical Boolean Model

Figure 3.13 displays the same  $256^3$ -pixel-sized cubic extract from an  $800^3$ -pixel synchrotronized HASYLAB coarse-sand-specimen that was already displayed as the left hand-side image of figure 1.1 along with a realization from the 3D-spherical Boolean Model generated according to algorithm 3.3.4. The two structures exhibit the same specific Minkowski-functionals with high accuracy. The fine sand-sample on the right hand-side of figure 1.1 cannot be modelled by the spherical Boolean Model if the grain distribution is chosen as in algorithm 3.3.4 although its specific Minkowski functionals are in the attainable region.

Figure 3.13 shows that the variability of grain sizes is much higher in the artificial structure than in the real sand-image. A possible explanation for this effect is the overlapping of grains in the Boolean structure. The volume and surface loss due to overlapping must be compensated by additional touching small spheres. If the goal is to artificially reproduce the sand-structure on the left hand-side of figure 3.13, a possible solution to this problem would be to chose a radial distribution which concentrates its mass on the radii observed in the real coarse-sand-sample.



**Figure 3.12:** *The influence of changing a single specific Minkowski-characteristic in the 2D-elliptical Boolean Model. As a reference the original sand cross-section in figure 3.8 was taken. Top row: enlarged/reduced specific volume, middle row: enlarged/reduced specific surface, bottom row: enlarged/reduced specific Euler characteristic.  $\mathbf{E}[\Delta]$  was chosen to be 0.8 in all but the last image, were it was chosen to be 3.7,  $\lambda_{\Delta} = 4$  for all images.*

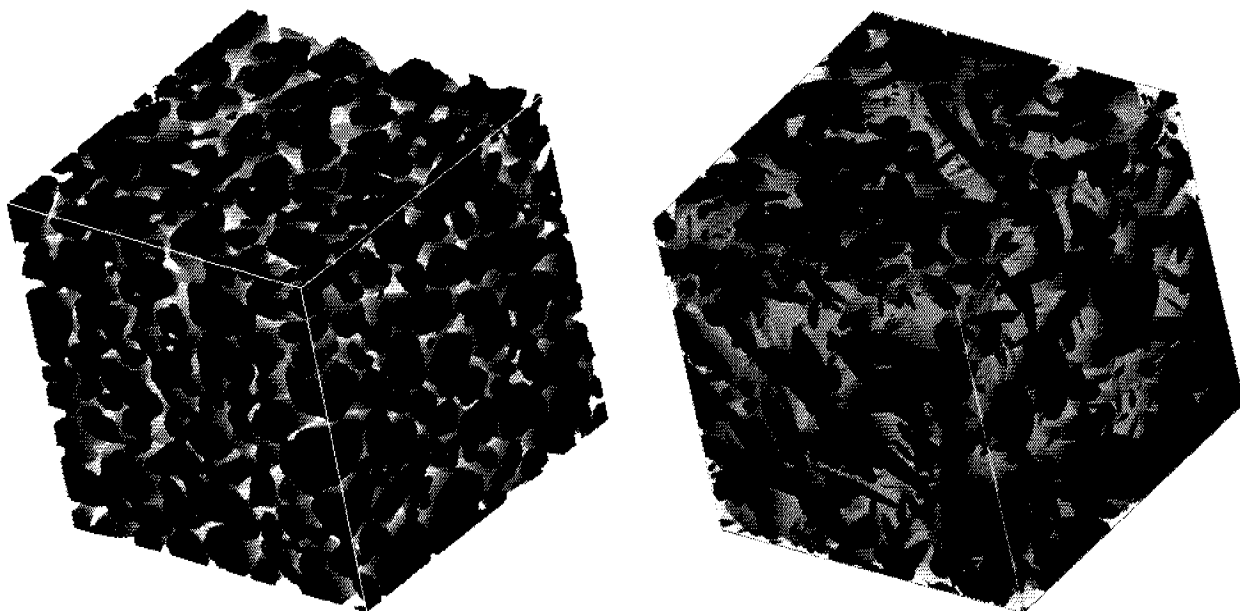


**Figure 3.13:** *An original  $256^3$ -pixel-sized cube taken from the  $800^3$ -pixel samples of the HASYLAB synchrotron sand data with coarse granularity and side length =  $11\mu\text{m}$  displayed on the left hand-side of figure 1.1. The right image shows a realization of the 3D-spherical Boolean Model generated according to algorithm 3.3.4. The two structures share the same specific Minkowski functionals.*

### 3.5.2 Typical Results for the 3D-Ellipsoidal Boolean Model

Figure 3.14 again displays a  $256^3$ -pixel-sized cubic extract from an  $800^3$ -pixel synchrotronized HASYLAB sand-specimen but this time the fine granularity-specimen that was already displayed as the right hand-side image of figure 1.1. The image comes along with a realization of a 3D-ellipsoidal-Boolean Model with rather eccentric grains which was generated according to algorithm 3.3.5 and has the same specific Minkowski functionals as the fine-sand-sample. The reason why we chose the fine-sand specimen this time is because its Minkowski-functionals are a little bit farther away from the boundary of the attainable region within the Ellipsoidal Boolean Model and thus exhibit more flexibility for a sensitivity analysis as it is done further below. Because of the high eccentricity of the ellipsoidal grains we cannot expect a high optical similarity of the two structures in figure 3.14, but the figure is an instructive example for the indeterminacy principle and once more shows that optically very different-looking structures may still have the same specific Minkowski-functionals.

If the goal is to reproduce the real fine sand-structure artificially, less eccentric ellipsoids must be used. Unfortunately the convenient choice of the Gamma-distribution for the extension factor  $c$  in algorithm 3.3.5 is not compatible with low eccentricities. The minimal surface area  $s_{00}$  of the prototype ellipsoid in algorithm 3.3.5 with which a structure with the same specific Minkowski functionals as the fine-sand-sample in figure 3.14 can be achieved is  $s_{00} = 6.74$  which is far above the theoretical minimum 4.84 given by the isoperimetric inequality. For less eccentric grains different grain distributions must be used and the simulation-algorithm must be adapted to the target sample. But this is not our goal here.



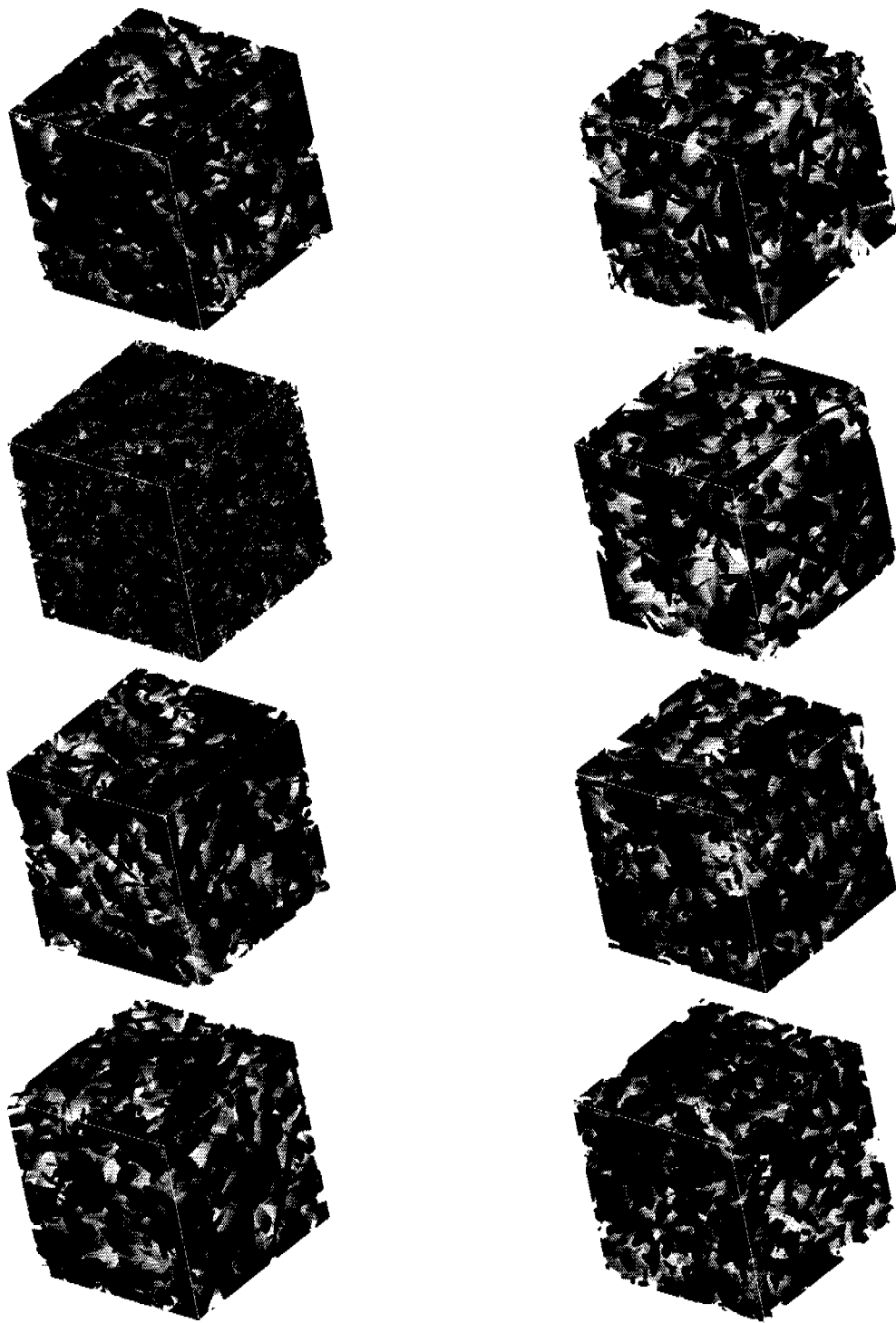
**Figure 3.14:** An original  $256^3$  pixel-sized cube taken from the  $800^3$ -pixel samples of the IASLAB synchrotron sand data with fine granularity and side length =  $11\mu\text{m}$  on the left hand-side. The right image shows a realization of the 3D-ellipsoidal Boolean Model generated according to algorithm 3.3.5. The two structures share the same specific Minkowski functionals.

Because of the greater flexibility of the ellipsoidal compared to the spherical Boolean Model similarly as in the discussion of the 2D models above it is possible to do a sensitivity analysis with respect to the change in a single specific Minkowski characteristic while keeping the other three specific Minkowski values fixed. In this spirit, figure 3.15 shows 6 realizations of the 3D-Ellipsoidal Boolean Model for each of which one of the specific Minkowski-functionals was changed according to table 3.3 whereas the others were chosen equal to the ones of the real fine-sand reference-structure shown in figures 1.1 and 3.14.

Image	modified functional	modification
top left	specific volume	increased to 0.65
top right	specific volume	decreased to 0.55
middle left	specific surface	decreased by factor 0.75
middle right	specific surface	decreased by factor 2
middle left	specific integral mean curvature	decreased by factor 0.5
middle right	specific integral mean curvature	increased by factor 2
bottom left	specific Euler	increased by factor 1.5
bottom right	specific Euler	decreased by factor 0.5

**Table 3.3:** Modifications of specific Minkowski functionals with respect to the ones of the original fine sand-sample shown on the left hand side of figure 3.14. The specific Minkowski-functionals of the reference-sample are  $v = 0.598$ ,  $s = 0.106$ ,  $\bar{M} = 1.201 \cdot 10^{-3}$ ,  $\bar{e} = -1.165 \cdot 10^{-4}$ . The corresponding images are shown in figure 3.15.

Figure 3.15 is not as instructive as the corresponding two dimensional figure. Apart from the image with increased specific surface, where the same thin fibre-effect as in 2D is clearly visible, the other images all look very similar. This gives us a hint that that Minkowski-functionals are even less powerful in three than in two dimensions to characterize a porous structure and there is a strong need to consider other aspects of the underlying geometry as well.



**Figure 3.15:** *The influence of changing a single specific Minkowski-characteristic in the 3D-ellipsoidal Boolean Model. As a reference the original fine-sand cross-section from figure 3.14 was taken. Top row: enlarged/reduced specific volume, second row: enlarged/reduced specific surface, third row: enlarged/reduced integral of mean curvature, bottom row: enlarged/reduced specific Euler characteristic. The parameter  $s_{00}$  in algorithm 3.3.5 which governs the mean grain-eccentricity was chosen to be 9 for all images which is high above its theoretical minimum 4.84.*





## Chapter 4

# The Thresholded Gaussian Model

In this chapter we describe how Thresholded Gaussian Random Fields can be used to model porous media with Minkowski functionals. In order to ensure that the geometrical quantities related to the Minkowski functionals exist and can be calculated for realizations of Thresholded Gaussian Fields, the underlying Gaussian Field must be smooth enough. Hence the first section below - besides giving necessary definitions - presents the relevant aspects of continuity and differentiability of (Gaussian) random fields.

### 4.1 Definition and Relevant Properties

The *Gaussian Field Model* is widely used to model spatio-temporal data with a certain dependence structure in various scientific branches, the most prominent of which being geology and geostatistics. The special appeal of the Gaussian Field Model lies in its simplicity and analytical tractability because the model is widely based on the Gaussian distribution which is both omnipresent in statistics and well-understood. Because the Gaussian Field Model can be combined with any feasible covariance structure in  $\mathbb{R}^d$ , it is also extremely flexible. A Gaussian field is usually defined as a continuous model where in every point of Euclidean space a standard Gaussian variable is placed and a covariance structure describing the dependence of those Gaussian variables is imposed. For our purposes we can restrict ourselves to the stationary (and isotropic) case.

**Definition 4.1.1.** (*Gaussian Field in  $\mathbb{R}^d$* )

- A **Gaussian Field**  $Z$  in  $\mathbb{R}^d$  is a collection of random variables  $(Z(x))_{x \in \mathbb{R}^d}$  such that the finite dimensional distributions of the random vectors  $Z(x_1), \dots, Z(x_n)$  are jointly Gaussian  $\forall x_1, \dots, x_n \in \mathbb{R}^d, n \in \mathbb{N}$ .
- A Gaussian field is **stationary** if and only if its expectation  $\mathbf{E}[Z(x)]$  is constant for all  $x \in \mathbb{R}^d$  and its **covariance**  $\rho(x, y) := \text{Cov}(Z(x), Z(y))$  depends only on  $h := y - x \in \mathbb{R}^d$ . We then write  $\rho(h) := \text{Cov}(Z(x), Z(x + h))$ .
- A stationary Gaussian field  $Z$  is **isotropic** if  $\rho(h) = \rho(\|h\|)$  is a function on  $\mathbb{R}$ .

Without loss of generality we may assume  $\mathbf{E}[Z(x)] = 0$  and  $\text{Var}(Z(x)) = 1$ . A function  $\rho$  is called the *covariance function* of a stationary Gaussian Random Field if and only if it is positive definite. There is a plethora of standard covariance models used in literature. A large collection along with techniques to construct one's own positive definite functions can be found in (Schlather 1999, p.12 ff.).

We now turn to continuity and differentiability of general random fields. Noting that  $\mathbf{E} [(Z(x) - Z(y))^2] = \rho(x, x) + \rho(y, y) - 2\rho(x, y)$  we recognize that continuity in the mean square sense for a random field  $Z(x)$  at  $x$  follows directly from continuity of the covariance function at  $(x, x)$ . For mean square differentiability we have the following important theorem:

**Theorem 4.1.2. (mean-square differentiability of random fields)** *Let  $Z(x), x \in \mathbb{R}^d$  be a random field with covariance function  $\rho(x, y)$  and differentiable expectation. If the derivative  $\partial^2 \rho(x, y) / \partial x_i \partial y_i$  exists and is finite  $\forall i \in 1, \dots, d$  at the point  $(x, x)$  then  $Z(x)$  is mean square differentiable in  $x$ . The covariance function of  $\frac{\partial Z}{\partial x_i}$  is given by  $\partial^2 \rho(x, y) / \partial x_i \partial y_i$ .*

**Proof.** (Theorem 4.1.2) see for example (Cramér and Leadbetter 1968, p. 84)

Of course the theorem can be extended to higher order derivatives. From the above statements it becomes evident that in the stationary case we have mean-square continuity at  $x$  if and only if  $\rho$  is continuous in 0 and mean-square-differentiability follows from existence and finiteness of  $\partial^2 \rho / \partial x_i^2$  in zero. Furthermore from the Cauchy-Schwarz inequality it immediately follows that a stationary covariance model attains its maximum value at the origin, ie.:

$$|\rho(x)| \leq \rho(0) \forall x \in \mathbb{R}^d.$$

Therefore in the isotropic case a necessary condition for the second-order derivatives of the covariance function  $\partial^2 \rho(x) / \partial x_i^2$  to exist is that the covariance model  $\rho(h), h \in \mathbb{R}$  has a vanishing derivative at zero, ie.  $\rho'(0) = 0$ .

For modelling porous media we restrict ourselves in the following to stationary and isotropic Gaussian Fields. For such fields we saw above that continuity of the covariance model in 0 is already sufficient for continuity in mean square but not for almost sure continuity of the sample paths which is the property we will need in the following. However, sufficient conditions for the latter can be given if the covariance model  $\rho(h)$  doesn't move away from  $\rho(0)$  too quickly:

**Theorem 4.1.3. (Sample path continuity of Gaussian Fields)** *Let  $Z(x), x \in \mathbb{R}^d$  be a real-valued Gaussian field with  $\mathbf{E} [Z(x)] = 0$  and a continuous covariance function  $\rho$ . Then  $Z$  has almost surely continuous sample paths in the unit cube if one of the following conditions is fulfilled:*

i)

$$\int_0^1 \sqrt{-\log u} dp(u) < \infty.$$

$$\text{where } p(u) = \max_{\|x_1 - x_2\| \leq |u| \sqrt{d}} \sqrt{\mathbf{E} [|Z(x_1) - Z(x_2)|^2]}.$$

ii)  $\exists C$  with  $0 < C < \infty$  and  $\exists \epsilon > 0$ ,

$$\mathbf{E} [|Z(x_1) - Z(x_2)|^2] = 2(\rho(0) - \rho(\|x_1 - x_2\|)) \leq \frac{C}{|\log \|x_1 - x_2\||^{1+\epsilon}},$$

$\forall x_1, x_2$  in the unit cube.

iii)  $Z$  is stationary and  $\exists \epsilon > 0$  such that

$$\int_{\mathbb{R}^d} |\log(1 + \|\lambda\|)|^{1+\epsilon} dF(\lambda) < \infty,$$

where  $F(\cdot)$  is the spectral distribution function of  $Z$ , ie  $\rho(h) = \int e^{iht} dF(t), \forall h \in \mathbb{R}^d$ .

**Proof. (Theorem 4.1.3)** see (Adler 1981), Thms 3.3.2, 3.4.1, 3.4.3.

**Definition 4.1.4. (Thresholded Gaussian Field)**

$$\Theta(\omega) = \{x \in \mathbb{R}^d | Z(x)(\omega) \geq b\}.$$

It is trivial to verify that this set  $\Theta$  is a random closed set in the sense of definition 2.4.1 if we assume that the Gaussian field  $Z$  has almost surely continuous sample paths (see the examples following definition 2.4.1 for the argument.) Specific realizations of  $\Theta$  we will call  $\theta$  in the following and we will also use the notation  $\theta(x)(\omega) = \mathbf{1}_{Z(x)(\omega) \geq b}$  to denote the binary state of a pixel within a realization  $\theta$ .

While Theorem 4.1.3 answers the question, when the *excursion set*  $\Theta(\omega)$  of a Gaussian field is a random set, it is not at all clear whether we can compute its specific Minkowski functionals  $\bar{V}_j(\Theta)$ . However, if we can make sure that the *level set over the observation window*  $\theta_b^W := \{x \in W | \theta(x) = b\}$  is a smooth  $d-1$  dimensional  $C^2$ -manifold then all the geometrical quantities related to the Minkowski functionals such as surface, curvatures and Euler-characteristic are well-defined in differential geometry and thus Minkowski functionals for  $\theta^W := \theta \cap W$  can be computed. It turns out that with moderate regularity assumptions this is indeed the case for stationary Gaussian fields which are our focus in this chapter. The following rather involved result is developed in (Adler 1981):

**Theorem 4.1.5. (boundary smoothness of excursion sets)** *Let  $Z(x), x \in \mathbb{R}^d$  be a stationary Gaussian field with almost surely pathwise continuous partial derivatives up to second order such that the joint distribution of  $Z(x)$  and all its first and distinct second order derivatives are non-degenerate. Let further  $W$  be a compact observation window whose boundary has zero Lebesgue measure. Then the level set  $\theta_b^W$  is almost surely a  $d-1$ -dimensional  $C^2$ -manifold and also a basic complex and therefore the Minkowski-functionals of  $\theta^W$  are well-defined.*

**Proof. (Theorem 4.1.5)** The theorem follows from (Adler 1981, Thm. 3.2.4) because all the conditions required for suitable regularity are fulfilled if the Gaussian field is as specified above. The definition of suitable regularity (Adler 1981, Definition 3.1.1) then immediately implies that  $\theta_b^W$  is a  $C^2$ -manifold via the implicit-function-theorem. The fact that the excursion set is a basic complex follows from (Adler 1981, Thm. 4.3.1).

**Remark 4.1.6.** The existence of first and second order partial derivatives of a stationary and isotropic follows according to Theorem 4.1.2 from the existence and finiteness of the second and fourth derivative of the covariance function in the origin. From Theorem 4.1.2 and the discussion thereafter it follows further that necessary and in the case of smooth enough covariance models also sufficient conditions for the existence of the second and fourth derivative of the covariance function in the origin is that both the first and the third derivatives of the covariance model  $\rho$  must vanish (otherwise the covariance function or its second derivative respectively would not be differentiable at the origin):

$$\frac{\partial \rho}{\partial h}(0) = 0 \quad \text{and} \quad \frac{\partial^3 \rho}{\partial h^3}(0) = 0. \quad (4.1)$$

Because (mean square sense) derivatives of Gaussian fields are obviously still Gaussian we can then use Theorem 4.1.3 to check whether their sample paths are almost surely continuous which gives us a method to verify the assumptions of Theorem 4.1.5.

Hence for our purposes we can restrict our attention to smooth covariance models with the property 4.1. In literature such covariance models are sometimes called *non-erratic*, so we adopt this definition here:

**Definition 4.1.7. (Non-erraticity)** A stationary and isotropic Gaussian Field  $Z(x)$  is called *non-erratic*, if its covariance model  $\rho$  can be expanded as follows:

$$\rho(h) = 1 - \frac{|\rho''(0)|}{2} \cdot h^2 + \frac{\rho^{(iv)}(0)}{24} \cdot h^4 + \mathcal{O}(h^5).$$

Once the existence of first and second order partial derivatives is established, note that the conditions for almost sure continuity of sample paths stated in Theorem 4.1.3 are rather weak. For example is hard to find a covariance model with a finite first derivative which does *not* fulfill condition ii). However, a non-erratic covariance model will always meet this condition because its derivative in the origin is zero, whereas the derivative of the dominating function  $1/|\log h|^{1+\epsilon}$  at the origin is infinity.

**Assumption 4.1.8.** For the rest of the chapter we restrict ourselves (if not otherwise stated) to considering stationary, isotropic and non-erratic Gaussian random fields. Our standard example is the widely used two-parametric Whittle-Matérn-Covariance-Model for  $\kappa \geq 3$  which we introduce now.

**Definition 4.1.9. (Whittle-Matérn-Covariance Model)** The *Whittle-Matérn Covariance Model* is the class of positive definite covariance functions

$$\rho_{\kappa,\nu}(h) = 2^{1-\kappa} \cdot \frac{1}{\Gamma(\kappa)} \cdot \left(\frac{h}{\nu}\right)^\kappa K_\kappa\left(\frac{h}{\nu}\right),$$

where  $\Gamma(\cdot)$  is the Gamma-function,  $K_\kappa(\cdot)$  is the Bessel-function of the second kind,  $\kappa > 0$  is a shape parameter and  $\nu > 0$  is a scale parameter.

The popularity of the Whittle-Matérn Class stems from the fact that it makes it possible to control the degree of differentiability of the covariance function and hence of the underlying Gaussian field: If  $\kappa = (2n + 1)/2$  then  $\rho_{\kappa,\nu}$  is  $2n$  times differentiable and thus  $Z$  is  $n$  times mean square differentiable. Another property which is crucial for us is that odd derivatives  $\partial^u \rho_{\kappa,\nu}(h^u)/\partial h$  of the Whittle-Matérn covariance function vanish in  $0 \forall u < \kappa$ . This can easily be seen using two properties of the Bessel function (see Abramowitz and Stegun 1972):

$$K_\kappa(h) \sim \frac{1}{2} \Gamma(\kappa) \left(\frac{1}{2}h\right)^{-\kappa} \quad \text{for } (h \rightarrow 0, \kappa > 0) \quad (4.2)$$

$$\frac{\partial^u}{\partial h^u} (h^\kappa K_\kappa(h)) = \sum_{i=0}^{\lfloor \frac{u}{2} \rfloor} c_i \cdot h^{\kappa-i} K_{\kappa-u+i}(h) \quad \text{for } u < \kappa, \quad (4.3)$$

where the  $c_i$  are constants.

Combining these properties for  $\nu > 1$  it follows that in the expansion (4.3) for  $h \downarrow 0$  all summands vanish if  $u < \kappa$  and  $u$  is odd because of:

$$\frac{u}{2} > i \iff h^{\kappa-i} \cdot K_{\kappa-u+i}(h) \xrightarrow{h \downarrow 0} 0.$$

## 4.2 Adler-Tomita-Analytical-Formulae

Once one has made sure that the level set for  $\theta_b^W$  is smooth enough, general expressions for the specific Minkowski functionals in the Thresholded Gaussian Model can be obtained using the relationships between the general integrals of mean curvature and Minkowski functionals from section 2.2.4. Instead of calculating Minkowski-functionals directly one can use the arsenal of methods from differential geometry to get explicit expressions for the general mean curvatures (definition 2.2.6) which are easier to obtain than expressions for the Minkowski functionals directly.

An explicit expression for the specific Euler characteristic  $\bar{\chi}(\Theta)$  for the excursion set  $\Theta$  of suitably regular Gaussian fields  $\Theta$  in arbitrary dimension was first obtained by R.J. Adler and a considerable part of the book (Adler 1981) is devoted to its derivation. However it was eventually the paper of H. Tomita (Tomita 1990) which contained expressions for *all* general integrals of mean curvature in arbitrary dimension. The following result by Tomita was a milestone and lies at the very heart of modelling porous media with Minkowski functionals:

**Theorem 4.2.1. (*explicit expressions for mean curvatures*)** *Let  $Z(x), x \in \mathbb{R}^d$  be a stationary, isotropic and non-erratic Gaussian field. Then the specific mean curvatures  $M_r$  of the pore space  $cl(\Theta^c) = \{x \in \mathbb{R}^d | Z(x) \leq b\}$  can be calculated as follows:*

$$M_r(\partial cl(\Theta^c)) = \left( \frac{-\rho''(0)}{2\pi} \right)^{\frac{r+1}{2}} \left[ 2\pi^{(r+1)/2} \frac{\Gamma(\frac{d-r+1}{2})}{\Gamma(\frac{d}{2})} \right] (-1)^r \left( \frac{\partial^r}{\partial x^r} \phi \right) (-b),$$

where  $\rho$  is the covariance model associated with  $Z(x)$  and  $\phi$  is the standard Gaussian density.

**Proof. (Theorem 4.2.1)** The involved proof of this result is covered in (Tomita 1990) and references therein.

For our purposes we are interested in these expressions for dimensions  $d = 2$  and  $d = 3$ . Note that the volume is not covered by the set of mean curvatures, however it is of course trivial to get an explicit expression for the specific volume  $\bar{v}(\Theta)$ . For the other Minkowski functionals recall that the connection between mean curvatures of a compact  $C^2$ -hypersurface  $S \in \mathbb{R}^d$  and Minkowski functionals is given by

$$M_r(\partial S) = d \cdot W_{r+1}(S)$$

(see equation 2.19 or (Santaló 2004, p.224) for a proof in the convex case.) To switch between expressions for  $S$  and the closure of its complement recall Theorem 2.1.13 and note  $M(S) = -M(cl(S^c))$ . Now Theorem 4.2.1 together with these remarks permits us to formulate the following corollaries summarizing Tomita's results in two and three dimensions. For the sake of completeness also the one-dimensional functionals are given:

**Corollary 4.2.2. (*1D-Minkowski functionals for Thresholded Gaussian Fields*)** *Let  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  a one-dimensional stationary, isotropic, non-erratic thresholded Gaussian field. Then its specific Minkowski functionals can be expressed as follows:*

$$\bar{v}(\Theta) = 1 - \Phi(b). \quad (4.4)$$

$$\bar{s}(\Theta) = \sqrt{\frac{2}{\pi}} \cdot \phi(b) \cdot \sqrt{-\rho''(0)}. \quad (4.5)$$

**Corollary 4.2.3. (2D-Minkowski functionals for Thresholded Gaussian Fields)**

Let  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  a two-dimensional stationary, isotropic, non-erratic thresholded Gaussian field. Then its specific Minkowski functionals can be expressed as follows:

$$\bar{v}(\Theta) = 1 - \Phi(b). \quad (4.6)$$

$$\bar{s}(\Theta) = \sqrt{\frac{\pi}{2}} \cdot \phi(b) \cdot \sqrt{-\rho''(0)}. \quad (4.7)$$

$$\bar{\chi}(\Theta) = -\frac{1}{2\pi} \cdot b \cdot \phi(b) \cdot \rho''(0), \quad (4.8)$$

**Corollary 4.2.4. (3D-Minkowski functionals for Thresholded Gaussian Fields)**

Let  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  a three-dimensional stationary, isotropic, non-erratic thresholded Gaussian field. Then its specific Minkowski functionals can be expressed as follows:

$$\bar{v}(\Theta) = 1 - \Phi(b). \quad (4.9)$$

$$\bar{s}(\Theta) = \frac{4}{\sqrt{2 \cdot \pi}} \cdot \phi(b) \cdot \sqrt{-\rho''(0)}. \quad (4.10)$$

$$\bar{M}(\Theta) = -b \cdot \phi(b) \cdot \rho''(0), \quad (4.11)$$

$$\bar{\chi}(\Theta) = \frac{1}{(2 \cdot \pi)^{3/2}} \cdot (b^2 - 1) \cdot \phi(b) \cdot (-\rho''(0))^{\frac{3}{2}}, \quad (4.12)$$

From these corollaries it becomes evident, that all the specific Minkowski-functionals are determined if the threshold  $b$  and  $\rho''(0)$  are given. Therefore only very little information about the covariance function  $\rho$  of the underlying field is incorporated in the Minkowski-functionals, however, quite surprisingly it is the very *same* information for all of them.

### 4.3 Asymptotical Unbiasedness of O-M-Type-Estimators

As far as the modelling of a Thresholded Gaussian Field  $\Theta$  is concerned, for a given random closed set  $\Theta$  we can only generate compact extracts  $\theta^W := \theta \cap W$  of its realizations  $\theta$  in an observation window  $W$  and hence being able to compute Minkowski functionals for  $\theta^W$  is enough. But as an additional restriction we can only simulate Thresholded Gaussian Fields  $\Theta_\Gamma^W = \Theta^W \cap \Gamma$  on a grid  $\Gamma$ . It is then obvious that  $\Theta_\Gamma^W$  is a random set in  $\mathbb{R}^d \cap W$  (because  $W \cap \Gamma$  is compact) and that its realizations  $\theta_\Gamma^W := \Theta_\Gamma^W(\omega)$  are elements of  $\mathcal{S}$  (more precisely of  $\mathcal{R}$ ). Therefore the Thresholded Gaussian Field on a grid  $\Theta_\Gamma$  fits in our framework and specific Minkowski-functionals for its realizations exist and can be estimated by Ohser-Mücklich estimators from subsections 2.3.1, 2.3.1. It is then interesting to ask whether the specific Minkowski functionals for the estimated  $\bar{V}_j(\Theta_\Gamma)$  converge to the  $\bar{V}_j(\Theta)$  for the continuous Thresholded Gaussian field. This question will be addressed in this section. More precisely we will show that both the  $d = 2$  and  $d = 3$  means for the Ohser-Mücklich estimators converge to the true theoretical values for the specific Minkowski functionals of the continuous Thresholded Field Model as the lattice spacing  $h$  tends to 0. We assume that estimation takes place on an equidistant finite cubic lattice  $\Gamma(h) = h \cdot \mathbb{Z}^d \cap W$  with lattice spacing  $h$ . Recall that we denote the continuous model by  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  where  $b$  is an arbitrary real threshold and  $Z(x)$  stands for a stationary and isotropic non-erratic Gaussian Random Field with  $\mathbf{E}[Z(x)] = 0$ ,  $\text{Var}(Z(x)) = 1$ ,  $\forall x \in \mathbb{R}^d$  and covariance function  $\rho(h)$ .

### 4.3.1 Integrals and Expansions

Throughout this and also the next section we will need to calculate a plethora of elementary integrals involving the standard Gaussian probability density and cumulative distribution function. Furthermore many of our theoretical results involve the explicit calculation of multivariate Gaussian probabilities for correlations tending to one. We thus decided to devote this subsection to the collection of all the necessary integrals and expansions we will need later to establish our main results. We start out by a lemma giving explicit expressions for some elementary integrals:

**Lemma 4.3.1. (Some elementary integrals)** Let  $\phi(x)$  and  $\Phi(x)$  denote the density and cumulative distribution functions of  $\mathcal{N}(0, 1)$ . Then:

- i).  $\int \Phi(x) dx = x \cdot \Phi(x) + \phi(x)$ .
- ii).  $\int x \cdot \phi(x) dx = -\phi(x)$ .
- iii).  $\int x \cdot \Phi(-x) dx = \frac{1}{2} \cdot x^2 \cdot \Phi(-x) - \frac{1}{2} \cdot x \cdot \phi(x) + \frac{1}{2} \cdot \Phi(x)$ .
- iv).  $\int x \cdot \Phi^2(-x) dx = \frac{1}{2} \cdot x^2 \cdot \Phi^2(-x) - x \cdot \phi(x) \cdot \Phi(-x) + \frac{1}{2} \cdot \Phi(x) \cdot \Phi(-x) + \frac{1}{2} \cdot \phi^2(x) + \frac{1}{2} \cdot \Phi(x)$ .
- v).  $\forall a \neq 0, b \in \mathbb{R} : \int_0^\infty \phi(a \cdot x) \cdot \Phi(b \cdot x) dx = \frac{\pi + 2 \cdot \arctan \frac{b}{|a|}}{4 \cdot \pi \cdot |a|}$ .
- vi).  $\forall a \neq 0, b \in \mathbb{R} : \int_{-\infty}^0 \phi(a \cdot x) \cdot \Phi(b \cdot x) dx = \frac{\pi - 2 \cdot \arctan \frac{b}{|a|}}{4 \cdot \pi \cdot |a|}$ .
- vii).  $\int_0^\infty x \cdot \Phi(-x) \cdot \Phi\left(-\frac{x}{\sqrt{2}}\right) dx = \frac{1}{4} - \frac{1}{4 \cdot \pi} \cdot \arctan \sqrt{2} - \frac{1}{2 \cdot \sqrt{2} \cdot \pi}$ .
- viii).  $\int_{-\infty}^0 x \cdot \Phi(x) \cdot \Phi\left(\frac{x}{\sqrt{3}}\right) dx = \frac{\sqrt{3}}{4 \cdot \pi} - \frac{5}{24}$ .
- ix).  $\int_{-\infty}^0 x^2 \cdot \Phi(x) dx = \frac{2}{3} \cdot \frac{1}{\sqrt{2\pi}}$ .
- x).  $\int_0^\infty x^2 \cdot \phi(x) dx = \frac{1}{2}$ .
- xi).  $\int_{-\infty}^0 x^2 \cdot \phi(x) \cdot \Phi(\sqrt{3} \cdot x) dx = \frac{1}{12} - \frac{3}{4\sqrt{3}} \frac{1}{\sqrt{2\pi}}$ .
- xii).  $\int_{-\infty}^0 x^2 \cdot \phi(x) \cdot \Phi(x) dx = \frac{1}{8} - \frac{1}{4\pi}$ .
- xiii).  $\int_{-\infty}^0 x^3 \cdot \Phi^2(x) dx = \frac{3}{16} - \frac{1}{2\pi}$ .
- xiv).  $\int_0^\infty x \cdot \phi^2(x) dx = \frac{1}{4\pi}$ .

**Proof. (Lemma 4.3.1)** Apart from parts v) and vi) respectively, the proofs are trivial applications of integration by parts and simple substitutions. Some of them, such as x), are even obvious. The proof of v) follows for example from the exposition in (Kendall and Stuart 1969), pg 351 and vi), of course, follows from v).

The main problem we face when it comes to find explicit expressions for Gaussian integrals is the fact that for  $a, b, c, d$  pairwise different real constants, the following integral *cannot* be explicitly calculated:

$$\int \phi(ax + b) \cdot \phi(cx + d) dx$$

This problem will cause us much difficulty in the following and to bypass it some of our proofs get considerably lengthier and less elegant. It is also one of the main reasons which prevents a more direct derivation of Theorem 4.2.1 without the use of the machinery of differential-geometry.

The following lemmata are statements about pairs, triples and quadruples of general standard Gaussian random variables  $Z_i$  in Euclidean spaces of arbitrary dimension  $d$  with correlation function  $\rho(\cdot)$ . However, as our focus lies on stationary and isotropic Gaussian fields we will adapt the formulation of these lemmata to this context.

Here is a first lemma about the probability that two Gaussian variables at small distance take different values:

**Lemma 4.3.2. (Gaussian Pairs)** *If  $Z(x), x \in \mathbb{R}^d$  is a stationary and isotropic Gaussian Field with covariance Function  $\rho(h)$  and  $x_h \in \mathbb{R}^d$  is a point at distance  $h$  from the origin, then for arbitrary threshold  $b \in \mathbb{R}$*

$$\begin{aligned} & \mathbb{P}[Z(0) \geq b, Z(x_h) \leq b] \\ &= \sqrt{2}\phi(0)\phi(b)\sqrt{1-\rho(h)} - \frac{\sqrt{2}}{12}(b^2-1)\phi(0)\phi(b)(1-\rho(h))^{\frac{3}{2}} + \mathcal{O}((1-\rho(h))^2). \end{aligned}$$

If  $\rho \in C^4$  then:

$$\begin{aligned} & \mathbb{P}[Z(0) \geq b, Z(x_h) \leq b] \\ &= \frac{1}{\sqrt{2\pi}}\phi(b)\sqrt{-\rho''(0)} \cdot h - \frac{\sqrt{-\rho''(0)}}{24} \left[ (b^2-1)\rho''(0)^2 + \frac{\rho^{(iv)}(0)}{\rho''(0)} \right] \cdot h^3 + \mathcal{O}(h^4), \end{aligned}$$

**Proof. (Lemma 4.3.2)** We have that:

$$\begin{pmatrix} Z(0) \\ Z(x_h) \end{pmatrix} \sim \mathcal{N}_2 \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho(h) \\ \rho(h) & 1 \end{pmatrix} \right),$$

for the conditional distribution we get  $Z(x_h)|Z(0) \sim \mathcal{N}(\mathbf{E}[Z(x_h)|Z(0)], \text{Var}(Z(x_h)|Z(0)))$  with:

$$\begin{aligned} \mathbf{E}[Z(x_h)|Z(0) = z_0] &= \rho(h) \cdot z_0, \\ \text{Var}(Z(x_h)|Z(0)) &= 1 - \rho^2(h). \end{aligned}$$

Hence introducing the new variable  $\tau = \sqrt{1-\rho(h)}$  and performing the substitution  $z_0 = b + u \cdot \tau$  we can write:

$$\begin{aligned} \mathbb{P}[Z(0) \geq b, Z(x_h) \leq b] &= \int_b^\infty \phi(z_0) \cdot \Phi \left( \frac{b - \rho(h) \cdot z_0}{\sqrt{1 - \rho^2(h)}} \right) dz_0 \\ &= \tau \cdot \underbrace{\int_0^\infty \phi(b + u \cdot \tau) \Phi \left( \frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}} \right) du}_{:=f(u,\tau)}, \end{aligned}$$



where we use  $f(u, \tau)$  as a shortcut for the underbraced expression. The next step is to expand  $f(u, \tau)$  around zero up to second order. The first and second derivative of  $f(u, \tau)$  are found to be:

$$\begin{aligned}
f'(u, \tau) &= -(b + u\tau) \cdot u \cdot \phi(b + u\tau) \cdot \Phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \\
&+ \phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \cdot \phi(b + u\tau) \cdot \frac{-u\tau^3 + 3u\tau + 2b}{(2 - \tau^2)\sqrt{2 - \tau^2}}. \\
f''(u, \tau) &= -u^2 \cdot \phi(b + u\tau) \Phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \\
&+ (b + u\tau)^2 \cdot u^2 \cdot \phi(b + u\tau) \Phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \\
&- 2 \cdot (b + u\tau) \cdot u \cdot \phi(b + u\tau) \cdot \phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \frac{-u\tau^3 + 3u\tau + 2b}{(2 - \tau^2)\sqrt{2 - \tau^2}} \\
&- \phi(b + u\tau) \cdot \phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \cdot \left(\frac{-u\tau^3 + 3u\tau + 2b}{(2 - \tau^2)\sqrt{2 - \tau^2}}\right)^2 \frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}} \\
&+ \phi(b + u\tau) \cdot \phi\left(\frac{\tau(b + u\tau) - u}{\sqrt{2 - \tau^2}}\right) \\
&\times \frac{(2 - \tau^2)\sqrt{2 - \tau^2} \cdot (-3u\tau^2 + 3u) - (-u\tau^3 + 3u\tau + 2b) \cdot (-2\tau) \cdot \frac{3}{2}\sqrt{2 - \tau^2}}{(2 - \tau^2)^3}.
\end{aligned}$$

and thus:

$$\begin{aligned}
f(u, 0) &= \phi(b) \cdot \Phi\left(\frac{-u}{\sqrt{2}}\right), \\
f'(u, 0) &= -b \cdot u \cdot \phi(b) \cdot \Phi\left(\frac{-u}{\sqrt{2}}\right) + \phi\left(\frac{u}{\sqrt{2}}\right) \cdot \phi(b) \cdot \frac{b}{\sqrt{2}}, \\
f''(u, 0) &= -u^2 \cdot \phi(b) \cdot \Phi\left(\frac{-u}{\sqrt{2}}\right) + b^2 \cdot u^2 \cdot \phi(b) \cdot \Phi\left(\frac{-u}{\sqrt{2}}\right) - 2 \cdot b \cdot u \cdot \phi(b) \cdot \phi\left(\frac{u}{\sqrt{2}}\right) \frac{b}{\sqrt{2}} \\
&+ \frac{u}{\sqrt{2}} \cdot \phi\left(\frac{u}{\sqrt{2}}\right) \frac{b^2}{2} \cdot \phi(b) + \phi\left(\frac{u}{\sqrt{2}}\right) \phi(b) \frac{3 \cdot \sqrt{2} \cdot u}{4}.
\end{aligned}$$

Therefore we can express  $P[Z(0) \geq b, Z(x_h) \leq b]$  as:

$$\begin{aligned}
P[Z(0) \geq b, Z(x_h) \leq b] &= \tau\phi(b) \int_0^\infty \Phi\left(\frac{-u}{\sqrt{2}}\right) du \\
&+ \tau^2 \left[ -b\phi(b) \int_0^\infty u \cdot \Phi\left(\frac{-u}{\sqrt{2}}\right) du + \frac{b}{\sqrt{2}} \cdot \phi(b) \int_0^\infty \phi\left(\frac{u}{\sqrt{2}}\right) du \right] \\
&+ \tau^3 \left[ \frac{b^2 - 1}{2} \cdot \phi(b) \int_0^\infty u^2 \cdot \Phi\left(\frac{-u}{\sqrt{2}}\right) du - \frac{3\sqrt{2}}{8}(b^2 - 1) \cdot \phi(b) \int_0^\infty u \cdot \phi\left(\frac{u}{\sqrt{2}}\right) du \right] \\
&+ \mathcal{O}(\tau^4).
\end{aligned}$$

Working out the integrals using Lemma 4.3.1, parts i), ii), iii) and ix) together with some elementary substitutions we conclude:

$$\begin{aligned}
\mathbb{P}[Z(0) \geq b, Z(x_h) \leq b] &= \tau \phi(b) \phi(0) \sqrt{2} \\
&+ \tau^2 \left[ b \phi(b) \left( \frac{1}{2} - \frac{1}{2} \right) \right] + \tau^3 \left[ (b^2 - 1) \phi(b) \phi(0) \sqrt{2} \left( \frac{2}{3} - \frac{3}{4} \right) \right] + \mathcal{O}(\tau^4) \\
&= \sqrt{2} \phi(0) \phi(b) \tau - \frac{\sqrt{2}}{12} (b^2 - 1) \phi(0) \phi(b) \tau^3 + \mathcal{O}(\tau^4),
\end{aligned}$$

which was to be established. The second equation immediately follows from expanding  $\sqrt{1 - \rho(h)}$  and  $(1 - \rho(h))^{3/2}$ .  $\square$

The next lemma is about the probability that standard Gaussian triples which span small similar triangles take different values:

**Lemma 4.3.3. (Gaussian triples)** *If  $Z(x), x \in \mathbb{R}^d$  is a (non-erratic) stationary and isotropic Gaussian Field with covariance function  $\rho(h)$  and the points  $x_1, x_2, x_3$  form a triangle  $\in \mathbb{R}^d$  such that:*

$$|x_1 - x_2| = h, \quad |x_3 - x_1| = c_2 \cdot h, \quad |x_3 - x_2| = c_3 \cdot h,$$

where  $c_2, c_3 \in \mathbb{R}^+$  are positive constants respecting the triangle inequality  $c_3 < 1 + c_2$ , then for an arbitrary threshold  $b \in \mathbb{R}$  we have:

$$\begin{aligned}
\mathbb{P}[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b] &= \\
&\sqrt{1 - \rho(h)} \cdot \phi(b) \cdot \int_0^\infty \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{c_2 \sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1+c_2^2-c_3^2}{2c_2} \\ \frac{1+c_2^2-c_3^2}{2c_2} & 1 \end{array} \right) \right] du \\
&+ (1 - \rho(h)) \cdot b \cdot \phi(b) \left\{ \frac{1}{\sqrt{2}} \int_0^\infty \phi \left( \frac{u}{\sqrt{2}} \right) \Phi \left( \frac{(-1 + c_2^2 - c_3^2) \cdot u}{\sqrt{2} \cdot (2c_2^2(1 + c_3^2) - (c_3^2 - 1)^2 - c_2^4)} \right) du \right. \\
&+ \frac{c_2}{\sqrt{2}} \int_0^\infty \phi \left( \frac{u}{c_2 \sqrt{2}} \right) \Phi \left( \frac{(1 - c_2^2 - c_3^2) \cdot u}{c_2 \sqrt{2} \cdot (2c_2^2(1 + c_3^2) - (c_3^2 - 1)^2 - c_2^4)} \right) du \\
&\left. - \int_0^\infty u \cdot \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{c_2 \sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1+c_2^2-c_3^2}{2c_2} \\ \frac{1+c_2^2-c_3^2}{2c_2} & 1 \end{array} \right) \right] du \right\} + \mathcal{O}((1 - \rho(h))^{\frac{3}{2}}).
\end{aligned}$$

**Remark 4.3.4.** Note that the correlation term asymptotically vanishes exactly if the three points  $x_1, x_2, x_3$  span a rectangular triangle. Because for the OM-estimators (defined in section 2.3) considered later only triples forming rectangular triangles must be taken into account, the involved calculations are easy enough to obtain explicit analytical results, because the two-dimensional cumulative distribution function  $\Phi_2(\cdot)$  decouples. The only exception being the triangular contributions to the estimator of the specific integral of mean curvature given in equation 2.32. However it turns out later, that also this case is still analytically tractable. We collect these two especially important cases in the following corollary:

**Corollary 4.3.5. (Special Triangles)** *Consider the situation of Lemma 4.3.3. Set  $\tau := \sqrt{1 - \rho(h)}$ . Then:*

i). For  $c_3^2 = 1 + c_2^2$  we have:

$$\begin{aligned} \mathbb{P}[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b] &= \tau \cdot \phi(b) \cdot \int_0^\infty \Phi\left(-\frac{u}{\sqrt{2}}\right) \cdot \Phi\left(-\frac{u}{c_2\sqrt{2}}\right) du \\ &+ \tau^2 \cdot b \cdot \phi(b) \left\{ \frac{1}{\sqrt{2}} \int_0^\infty \phi\left(\frac{u}{\sqrt{2}}\right) \Phi\left(\frac{-2u}{\sqrt{2} \cdot (2c_2^2(1+c_3^2) - 2c_2^4)}\right) du \right. \\ &+ \frac{c_2}{\sqrt{2}} \int_0^\infty \phi\left(\frac{u}{c_2\sqrt{2}}\right) \Phi\left(\frac{-2c_2 \cdot u}{\sqrt{2} \cdot (2c_2^2(1+c_3^2) - 2c_2^4)}\right) du \\ &\left. - \int_0^\infty u \cdot \Phi\left(-\frac{u}{\sqrt{2}}\right) \cdot \Phi\left(-\frac{u}{c_2\sqrt{2}}\right) du \right\} + \mathcal{O}(\tau^3). \end{aligned}$$

ii) For  $c_2^2 = c_3^2 = 1$  we have:

$$\begin{aligned} \mathbb{P}[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b] &= \tau \cdot \phi(b) \cdot \int_0^\infty \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{\sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{array} \right) \right] du \\ &+ b \cdot \phi(b) \cdot \left[ \sqrt{2} \cdot \int_0^\infty \phi\left(\frac{u}{\sqrt{2}}\right) \cdot \Phi\left(-\sqrt{\frac{3}{2}} \cdot u\right) du \right. \\ &\left. - \int_0^\infty u \cdot \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{\sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{array} \right) \right] du \right] \cdot \tau^2 + \mathcal{O}(\tau^3). \end{aligned}$$

**Proof.** (Corollary 4.3.5) obvious.

**Proof.** (Lemma 4.3.3)

We define the variables  $\rho_1 := \rho(h)$ ,  $\rho_2 := \rho(c_2 \cdot h)$ ,  $\rho_3 := \rho(c_3 \cdot h)$  and  $\tau = \sqrt{1 - \rho(h)}$ . With this notation, we obviously have  $\rho_i \rightarrow 1$ ,  $i \in \{1, 2, 3\}$  as  $h \downarrow 0$ .

$$\begin{pmatrix} Z(x_1) \\ Z(x_2) \\ Z(x_3) \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix} \right).$$

In the following we are going to expand the triple probability on the left hand side in the statement of Lemma 4.3.3 as a function of  $\tau$ . The involved calculations are lengthy and tedious but straightforward applications of Taylor's theorem. We will therefore suppress the details of these calculations to a large degree.

Expanding  $\rho_2$  and  $\rho_3$  as functions of  $\tau$  we find:

$$\begin{aligned} \rho_2 &= 1 - c_2^2 \cdot \tau^2 + \frac{\rho^{(iv)}(0)c_2^2(c_2^2 - 1)}{6\rho''(0)^2} \cdot \tau^4 + \mathcal{O}(\tau^6). \\ \rho_3 &= 1 - c_3^2 \cdot \tau^2 + \frac{\rho^{(iv)}(0)c_3^2(c_3^2 - 1)}{6\rho''(0)^2} \cdot \tau^4 + \mathcal{O}(\tau^6). \end{aligned}$$

Furthermore by conditioning onto  $Z(x_1)$  we can write:

$$\mathbb{P}[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b] = \int_b^\infty \phi(z_1) \cdot \mathbb{P}[Z(x_2) \leq b, Z(x_3) \leq b | z_1] dz_1.$$

Next we note for the conditional distribution  $Z(x_2), Z(x_3)|Z(x_1)$ :

$$Z(x_2), Z(x_3)|Z(x_1) \sim \mathcal{N} \left( \begin{pmatrix} \rho_1 z_1 \\ \rho_2 z_1 \end{pmatrix}, \begin{pmatrix} 1 - \rho_1^2 & \rho_3 - \rho_1 \rho_2 \\ \rho_3 - \rho_1 \rho_2 & 1 - \rho_2^2 \end{pmatrix} \right).$$

If we now perform the change of variables  $z_1 = b + u \cdot \tau$ , we are left with the following expression for  $P[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b]$ :

$$\begin{aligned} & P[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b] \\ &= \tau \cdot \int_0^\infty \phi(b + u \cdot \tau) \cdot \Phi \left( \begin{pmatrix} b \\ b \end{pmatrix}, \begin{pmatrix} \rho_1(b + u\tau) \\ \rho_2(b + u\tau) \end{pmatrix}, \begin{pmatrix} 1 - \rho_1^2 & \rho_3 - \rho_1 \rho_2 \\ \rho_3 - \rho_1 \rho_2 & 1 - \rho_2^2 \end{pmatrix} \right) du \\ &= \tau \cdot \int_0^\infty \phi(b + u \cdot \tau) \cdot \underbrace{\Phi \left( \begin{pmatrix} \frac{b - \rho_1(b + u\tau)}{\sqrt{1 - \rho_1^2}} \\ \frac{b - \rho_2(b + u\tau)}{\sqrt{1 - \rho_2^2}} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \frac{\rho_3 - \rho_1 \rho_2}{\sqrt{1 - \rho_1^2} \sqrt{1 - \rho_2^2}} \\ \frac{\rho_3 - \rho_1 \rho_2}{\sqrt{1 - \rho_1^2} \sqrt{1 - \rho_2^2}} & 1 \end{pmatrix} \right)}_{:=f(u, \tau)} du. \end{aligned}$$

It remains to perform a Taylor expansion of the integrand  $f(u, \tau)$  up to at least first order. To do this we need the following expansions in  $\tau$ :

$$\begin{aligned} b_1(\tau) &:= \frac{b - \rho_1(b + u\tau)}{\sqrt{1 - \rho_1^2}} = -\frac{u}{\sqrt{2}} + \frac{b}{\sqrt{2}}\tau + \frac{3u}{4\sqrt{2}}\tau^2 + \mathcal{O}(\tau^3), \\ b_2(\tau) &:= \frac{b - \rho_2(b + u\tau)}{\sqrt{1 - \rho_2^2}} = -\frac{u}{c_2\sqrt{2}} + \frac{bc_2}{\sqrt{2}}\tau + \frac{((1 - c_2^2)\rho^{(iv)}(0) + 9\rho''(0)^2c_2^2)u}{12\sqrt{2}\rho''(0)^2c_2}\tau^2 + \mathcal{O}(\tau^3), \\ \rho(\tau) &:= \frac{\rho_3 - \rho_1\rho_2}{\sqrt{1 - \rho_1^2}\sqrt{1 - \rho_2^2}} = \frac{1 + c_2^2 - c_3^2}{2c_2} + \mathcal{O}(\tau^2). \end{aligned} \quad (4.13)$$

Next we calculate the derivative of  $f(u, \tau)$ . By calculating the partial derivatives with respect to  $b_1(\tau), b_2(\tau), \rho(\tau)$  and using the expansions 4.25 the chain rule gives (after a couple of simple algebraic manipulations with the Gaussian density):

$$\begin{aligned} \frac{\partial f}{\partial \tau}(u, \tau) &= -(b + u\tau) \cdot u \cdot \phi(b + u\tau) \Phi_2 \left[ \begin{pmatrix} b_1(\tau) \\ b_2(\tau) \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho(\tau) \\ \rho(\tau) & 1 \end{pmatrix} \right] \\ &+ \phi(b + u\tau) \left\{ \phi(b_1(\tau)) \cdot \Phi \left( \frac{b_2(\tau) - \rho(\tau)b_1(\tau)}{\sqrt{1 - \rho^2(\tau)}} \right) \cdot \left( \frac{b}{\sqrt{2}} + \mathcal{O}(\tau) \right) \right. \\ &\left. + \phi(b_2(\tau)) \cdot \Phi \left( \frac{b_1(\tau) - \rho(\tau)b_2(\tau)}{\sqrt{1 - \rho^2(\tau)}} \right) \cdot \left( \frac{bc_2}{\sqrt{2}} + \mathcal{O}(\tau) \right) + \mathcal{O}(\tau) \right\} \end{aligned}$$

This shows that we will be in need for two further expansions in the following:

$$\begin{aligned} \frac{b_2(\tau) - \rho(\tau)b_1(\tau)}{\sqrt{1 - \rho^2(\tau)}} &= \frac{(c_2^2 - c_3^2 - 1) \cdot u}{\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} \\ &+ \frac{b \cdot (c_2^2 + c_3^2 - 1)}{\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} \cdot \tau + \mathcal{O}(\tau^2), \\ \frac{b_1(\tau) - \rho(\tau)b_2(\tau)}{\sqrt{1 - \rho^2(\tau)}} &= \frac{(1 - c_2^2 - c_3^2) \cdot u}{c_2\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} \\ &+ \frac{b \cdot c_2 \cdot (1 - c_2^2 + c_3^2)}{\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} \cdot \tau + \mathcal{O}(\tau^2). \end{aligned} \quad (4.14)$$

Hence we conclude:

$$\begin{aligned}
f(u, 0) &= \phi(b) \cdot \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{c_2\sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1+c_2^2-c_3^2}{2c_2} \\ \frac{1+c_2^2-c_3^2}{2c_2} & 1 \end{array} \right) \right], \\
\int_0^\infty \frac{\partial f}{\partial \tau}(u, 0) du &= b\phi(b) \left\{ \frac{1}{\sqrt{2}} \int_0^\infty \phi\left(\frac{u}{\sqrt{2}}\right) \Phi\left(\frac{(-1+c_2^2-c_3^2) \cdot u}{\sqrt{2} \cdot (2c_2^2(1+c_3^2) - (c_3^2-1)^2 - c_2^4)}\right) du \right. \\
&\quad + \frac{c_2}{\sqrt{2}} \int_0^\infty \phi\left(\frac{u}{c_2\sqrt{2}}\right) \Phi\left(\frac{(1-c_2^2-c_3^2) \cdot u}{c_2\sqrt{2} \cdot (2c_2^2(1+c_3^2) - (c_3^2-1)^2 - c_2^4)}\right) du \\
&\quad \left. - \int_0^\infty u \cdot \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{c_2\sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1+c_2^2-c_3^2}{2c_2} \\ \frac{1+c_2^2-c_3^2}{2c_2} & 1 \end{array} \right) \right] du \right\}.
\end{aligned}$$

Recalling that  $\tau = \sqrt{1 - \rho(h)}$  establishes the Lemma.  $\square$

The next lemma is about a probability for standard Gaussian quadruples. This lemma will simplify the tasks we face in later sections considerably because it will allow us to identify a class of lattice cells with a special vertex configuration as negligible:

**Lemma 4.3.6. (Negligible configurations)** Let  $Z_0 := Z(x_0), Z_1 := Z(x_1), Z_2 := Z(x_2), Z_3 := Z(x_3)$  be standard Gaussian variables of a stationary and isotropic Gaussian Field with covariance function  $\rho(\cdot)$  such that  $x_0, x_1 := x_0 + he_1, x_2 := x_0 + he_2, x_3 := x_0 + he_1 + he_2$  span a square with side length  $h$ . Then it holds:

$$P[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] = O(h^4), \quad \forall b \in \mathbb{R}.$$

**Proof. (Lemma 4.3.6)** To start out we introduce the following new set of random variables  $Y_1, Y_2, Y_3$ :

$$Y_1 := \frac{Z_1 - Z_0}{h}, \quad Y_2 := \frac{Z_2 - Z_0}{h}, \quad Y_3 := \frac{Z_3 - Z_1 - Z_2 + Z_0}{h^2},$$

which are discrete approximations of the first and one of the second directional derivatives. We then have  $(Z_0, Y_1, Y_2, Y_3) \sim \mathcal{N}(0, \Sigma)$  with and

$$\Sigma = \begin{pmatrix} 1 & \frac{\rho(h)-1}{h} & \frac{\rho(h)-1}{h} & \frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^2} \\ \frac{\rho(h)-1}{h} & \frac{2(1-\rho(h))}{h^2} & \frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^2} & -2\frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^3} \\ \frac{\rho(h)-1}{h} & \frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^2} & \frac{2(1-\rho(h))}{h^2} & -2\frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^3} \\ \frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^2} & -2\frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^3} & -2\frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^3} & 4\frac{1+\rho(\sqrt{2}h)-2\rho(h)}{h^4} \end{pmatrix}.$$

Then we can write:

$$\begin{aligned}
& \mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] \\
&= \mathbb{P}\left[Z_0 \geq b, Y_1 \leq \frac{b - Z_0}{h}, Y_2 \leq \frac{b - Z_0}{h}, Y_3 \geq \frac{b - Z_0 - h(Y_1 + Y_2)}{h^2}\right] \\
&= \int_b^\infty \int_{-\infty}^{\frac{b-Z_0}{h}} \int_{-\infty}^{\frac{b-Z_0}{h}} \mathbb{P}\left[Y_3 \geq \frac{b - Z_0 - h(Y_1 + Y_2)}{h^2} \middle| Z_0 = z_0, Y_1 = y_1, Y_2 = y_2\right] \\
&\times \mathbb{P}[Z_0 \in dz_0, Y_1 \in dy_1, Y_2 \in dy_2] \\
&= h^4 \cdot \int_0^\infty \int_u^\infty \int_u^\infty \mathbb{P}\left[Y_3 \geq v + w - u \middle| Z_0 = b + uh^2, Y_1 = -vh, Y_2 = -wh\right] \\
&\times \phi\left[\begin{pmatrix} b + uh^2 \\ -vh \\ -wh \end{pmatrix}, 0, \Sigma_{-44}\right] dw dv du, \tag{4.15}
\end{aligned}$$

where we have substituted  $Z_0 := b + uh^2$ ,  $Y_1 = -vh$ ,  $Y_2 = -wh$  and  $\Sigma_{-44}$  emerges from  $\mu$  and  $\Sigma$  by deleting the fourth row and column respectively.

The conditional distribution  $Y_3|Z_0, Y_1, Y_2$  is  $\mathcal{N}(\tilde{\mu}, \tilde{\Sigma})$ , where with some tedious calculations  $\tilde{\mu}$  and  $\tilde{\Sigma}$  can be found to be:

$$\begin{aligned}
\tilde{\mu} &= \frac{1}{h^2} \cdot \frac{1 - 2\rho(h) + \rho(\sqrt{2}h)}{1 - 2\rho^2(h) + \rho(\sqrt{2}h)} \cdot \left[ b(\rho(\sqrt{2}h) - 1) + h^2 \left( (v + w)(1 + \rho(h)) - u(1 - \rho(\sqrt{2}h)) \right) \right] \\
&= \frac{-\rho^{(iv)}(0)(\rho''(0)b + 2(v + w))}{12\rho''(0)} \cdot h^2 + \mathcal{O}(h^4). \\
\tilde{\Sigma} &= \frac{1}{h^4} \cdot \frac{(1 - \rho(\sqrt{2}h)) \cdot ((1 + \rho(\sqrt{2}h))^2 - 4\rho^2(h))}{1 - 2\rho^2(h) + \rho(\sqrt{2}h)} = \frac{\rho^{(iv)}(0)}{3} + \mathcal{O}(h^2).
\end{aligned}$$

Thus in first order the expression of equation 4.15 boils down to:

$$\begin{aligned}
& \mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] \sim \\
& h^4 \cdot \phi(b) \cdot \phi(0)^2 \int_0^\infty \int_u^\infty \int_u^\infty \left[ 1 - \Phi\left(\frac{\sqrt{3}(v + w - u)}{\sqrt{\rho^{(iv)}(0)}}\right) \right] dw dv du + \mathcal{O}(h^5).
\end{aligned}$$

The integrals in the last expression are finite. They can be related to the third absolute moment of the standard normal. If  $Z \sim \mathcal{N}(0, 1)$  by substituting  $w' := v + w - u$  we obtain:

$$\begin{aligned}
& \int_0^\infty \int_u^\infty \int_u^\infty \left[ 1 - \Phi\left(\frac{\sqrt{3}(v + w - u)}{\sqrt{\rho^{(iv)}(0)}}\right) \right] dw dv du \\
&= \int_0^\infty \int_u^\infty \int_v^\infty \left[ 1 - \Phi\left(\frac{\sqrt{3}w'}{\sqrt{\rho^{(iv)}(0)}}\right) \right] dw' dv du = \int_0^\infty \int_0^{w'} \int_0^v \left[ 1 - \Phi\left(\frac{\sqrt{3}w'}{\sqrt{\rho^{(iv)}(0)}}\right) \right] du dv dw' \\
&= \int_0^\infty \frac{1}{2}w'^2 \cdot \left[ 1 - \Phi\left(\frac{\sqrt{3}w'}{\sqrt{\rho^{(iv)}(0)}}\right) \right] dw' = \frac{1}{6} \left( \frac{\rho^{(iv)}(0)}{3} \right)^{\frac{3}{2}} \int_0^\infty y^3 \phi(y) dy \\
&= \frac{1}{12} \left( \frac{\rho^{(iv)}(0)}{3} \right)^{\frac{3}{2}} \mathbf{E}[|Z|^3] < \infty,
\end{aligned}$$

and this concludes the proof.  $\square$

The last lemma we prove here is about a probability concerning a quadruple  $Z_1, Z_2, Z_3, Z_4$  of standard Gaussian variables on the real line consisting of two collapsing pairs  $(Z_1, Z_2)$  and  $(Z_3, Z_4)$  at a fixed distance  $x$ :

**Lemma 4.3.7. (Collapsing Gaussian pairs at fixed distance)** Let  $(Z_1, Z_2, Z_3, Z_4)$  be a quadruple of standard Gaussian random variables with the following covariance structure:

$$\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{pmatrix} \sim \mathcal{N} \left[ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_3 & \rho_4 \\ \rho_1 & 1 & \rho_5 & \rho_6 \\ \rho_3 & \rho_5 & 1 & \rho_2 \\ \rho_4 & \rho_6 & \rho_2 & 1 \end{pmatrix} \right],$$

such that  $\rho_1, \rho_2 \rightarrow 1, \rho_3$  fixed and  $\rho_4, \rho_5, \rho_6 \rightarrow \rho_3$  for  $\tau = \sqrt{1 - \rho_1} \rightarrow 0$  in the following way:

$$\begin{aligned} 1 - \rho_2 &= c_2 \tau^2 + \mathcal{O}(\tau^4) \\ \rho_4 &= \rho_3 + c_4 \tau + \mathcal{O}(\tau^2) \\ \rho_5 &= \rho_3 + c_5 \tau + \mathcal{O}(\tau^2) \\ \rho_6 &= \rho_3 + (c_4 + c_5) \tau + \mathcal{O}(\tau^2) \\ \rho_6 - \rho_5 - \rho_4 + \rho_3 &= c_6 \tau^2 + o(\tau^2). \end{aligned} \tag{4.16}$$

Then:

$$\begin{aligned} &P[Z_1 \geq b, Z_2 \leq b, Z_3 \geq b, Z_4 \leq b] \sim \\ &\sqrt{2} \phi(b) \tau^2 \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{2}}\right) \phi\left(\frac{2b(1 - \rho_3) - c_5 w_2}{\sqrt{\Sigma_{11}}}\right) \frac{1}{\sqrt{\Sigma_{11}}} \left[ m \Phi\left(-\frac{m}{s}\right) - s \phi\left(\frac{m}{s}\right) \right] dw_2, \end{aligned}$$

where:

$$\begin{aligned} m &= 2b \left[ \frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}} (1 - \rho_3) \right] + w_2 \left[ \frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}} c_5 \right], \\ s &= \sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}, \\ \Sigma &= \begin{pmatrix} 4(1 - \rho_3^2) - 2c_5^2 & -2c_4\rho_3 - c_5c_6 \\ -2c_4\rho_3 - c_5c_6 & 2c_2 - c_4^2 - \frac{c_5^2}{2} \end{pmatrix} \end{aligned}$$

**Proof. (Lemma 4.3.7)** We introduce the following set of new variables  $(W_1, \dots, W_4)$ :

$$\begin{aligned} W_1 &= Z_1 + Z_2, & W_2 &= \frac{Z_2 - Z_1}{\tau}, \\ W_3 &= Z_3 + Z_4, & W_4 &= \frac{Z_4 - Z_3}{\tau} \end{aligned}$$

Then with the assumptions of equation-system 4.16 we find for the distribution of  $(W_1, \dots, W_4)$ :

$$\begin{pmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \end{pmatrix} \sim \mathcal{N} \left[ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 + \mathcal{O}(\tau^2) & 0 & 4\rho_3 + \mathcal{O}(\tau) & 2c_4 + \mathcal{O}(\tau) \\ 0 & 2 + \mathcal{O}(\tau^2) & 2c_5 + \mathcal{O}(\tau) & c_6 + o(1) \\ 4\rho_3 + \mathcal{O}(\tau) & 2c_5 + \mathcal{O}(\tau) & 4 + \mathcal{O}(\tau^2) & 0 \\ 2c_4 + \mathcal{O}(\tau) & c_6 + o(1) & 0 & 2c_2 + \mathcal{O}(\tau^2) \end{pmatrix} \right],$$

Furthermore we have the following set equivalence:

$$\begin{aligned} & \{Z_1 \geq b, Z_2 \leq b, Z_3 \geq b, Z_4 \leq b\} = \\ & \{W_2 \leq 0, 2b + \tau W_2 \leq W_1 \leq 2b - \tau W_2, W_4 \leq 0, 2b + \tau W_4 \leq W_3 \leq 2b - \tau W_4\}. \end{aligned}$$

By conditioning on the independent pair  $W_1, W_2$  we calculate:

$$\begin{aligned} & \mathbb{P}[Z_1 \geq b, Z_2 \leq b, Z_3 \geq b, Z_4 \leq b] = \tag{4.17} \\ & \frac{1}{2\sqrt{2}} \int_{-\infty}^0 \phi\left(\frac{w_2}{\sqrt{2}}\right) \int_{2b+\tau w_2}^{2b-\tau w_2} \phi\left(\frac{w_1}{2}\right) \mathbb{P}[2b - \tau W_4 \leq W_3 \leq 2b + \tau W_4, W_4 \leq 0 | w_1, w_2] dw_1 dw_2 \end{aligned}$$

If we now denote the conditional distribution  $(W_3, W_4 | W_1, W_2)$  as:

$$(W_3, W_4 | w_1, w_2)^T \sim \mathcal{N} \left[ \begin{pmatrix} \beta_{31}(\tau)w_1 + \beta_{32}(\tau)w_2 \\ \beta_{41}(\tau)w_1 + \beta_{42}(\tau)w_2 \end{pmatrix}, \Sigma(\tau) \right]$$

and consider the limit  $\tau \rightarrow 0$  in 4.17 we find that both the limits  $\Sigma := \Sigma(0)$  and  $\beta := \beta(0)$  exist and thus:

$$\begin{aligned} & \mathbb{P}[Z_1 \geq b, Z_2 \leq b, Z_3 \geq b, Z_4 \leq b] \sim \tag{4.18} \\ & \sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 \phi\left(\frac{w_2}{\sqrt{2}}\right) w_2 \int_{-\infty}^0 w_4 \phi_2 \left[ \begin{pmatrix} 2b(1 - \beta_{31}) - \beta_{32}w_2 \\ w_4 - 2b\beta_{41} - \beta_{42}w_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right] dw_4 dw_2, \end{aligned}$$

where with some standard calculations one finds  $\beta_{31} = \rho_3$ ,  $\beta_{32} = c_5$ ,  $\beta_{41} = \frac{c_4}{2}$ ,  $\beta_{42} = \frac{c_6}{2}$  and

$$\Sigma = \begin{pmatrix} 4(1 - \rho_3^2) - 2c_5^2 & -2c_4\rho_3 - c_5c_6 \\ -2c_4\rho_3 - c_5c_6 & 2c_2 - c_4^2 - \frac{c_6^2}{2} \end{pmatrix}.$$

Because the first entry of  $\phi_2(\cdot)$  is independent of  $w_4$ , the inner integral in equation 4.18 can be worked out explicitly. By decomposing  $\phi_2(\cdot)$  as:

$$\begin{aligned} & \phi_2 \left[ \begin{pmatrix} 2b(1 - \beta_{31}) - \beta_{32}w_2 \\ w_4 - 2b\beta_{41} - \beta_{42}w_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right] = \\ & \phi(2b(1 - \beta_{31}) - \beta_{32}w_2, 0, \Sigma_{11}) \cdot \phi \left( w_4 - 2b\beta_{41} - \beta_{42}w_2, \frac{\Sigma_{12}}{\Sigma_{11}} (2b(1 - \beta_{31}) - \beta_{32}w_2), \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}} \right) \end{aligned}$$

and thus by abbreviating  $m = 2b \left( \beta_{41} + \frac{\Sigma_{12}}{\Sigma_{11}} (1 - \beta_{31}) \right) + w_2 \left( \beta_{42} - \frac{\Sigma_{12}}{\Sigma_{11}} \beta_{32} \right)$  and furthermore  $s = \sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}$  equation 4.18 can be transformed into:



$$\begin{aligned} & \mathbb{P}[Z_1 \geq b, Z_2 \leq b, Z_3 \geq b, Z_4 \leq b] \sim \\ & \sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 \phi\left(\frac{w_2}{\sqrt{2}}\right) w_2 \phi\left(\frac{2b(1-\beta_{31})-\beta_{32}w_2}{\sqrt{\Sigma_{11}}}\right) \frac{1}{\sqrt{\Sigma_{11}}} \int_{-\infty}^0 w_4 \phi\left(\frac{w_4-m}{s}\right) \frac{1}{s} dw_4 dw_2 \\ & = \sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 \phi\left(\frac{w_2}{\sqrt{2}}\right) w_2 \phi\left(\frac{2b(1-\beta_{31})-\beta_{32}w_2}{\sqrt{\Sigma_{11}}}\right) \frac{1}{\sqrt{\Sigma_{11}}} \left[m\Phi\left(-\frac{m}{s}\right) - s\phi\left(\frac{m}{s}\right)\right] dw_2, \end{aligned}$$

which is what we wanted to show. □

### 4.3.2 Two Dimensional Convergence Results

First of all we state and prove the 2D-version of the convergence results. The specific Minkowski functionals arising for  $d = 2$  are the porosity or volume fraction  $\bar{v}$ , the specific surface area  $\bar{s}$  and the specific Euler-characteristic  $\bar{\chi}$ .

**Theorem 4.3.8. (Asymptotical Unbiasedness of OM-Estimators (2D-Version))**

Let  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  a two-dimensional stationary, isotropic, non-erratic Thresholded Gaussian Field. Then the mean Ohser-Mücklich estimates for the specific Minkowski functionals on a finite grid  $\Gamma(h)$  converge to the continuous limit, ie. the rates of the specific Minkowski-functionals in the continuous setting, if we let the grid-spacing  $h$  tend to zero:

$$\begin{aligned} i) \quad & \lim_{h \downarrow 0} \mathbf{E} [\hat{v}(\theta_\Gamma(h))] = \bar{v}(\Theta) = 1 - \Phi(b). \\ ii) \quad & \lim_{h \downarrow 0} \mathbf{E} [\hat{s}(\theta_\Gamma(h))] = \bar{s}(\Theta) = \sqrt{\frac{\pi}{2}} \cdot \phi(b) \cdot \sqrt{-\rho''(0)}. \\ iii) \quad & \lim_{h \downarrow 0} \mathbf{E} [\hat{\chi}(\theta_\Gamma(h))] = \bar{\chi}(\Theta) = -\frac{1}{2\pi} \cdot b \cdot \phi(b) \cdot \rho''(0), \end{aligned}$$

where  $\theta_\Gamma(h) = \{x \in \Gamma(h) | \Theta(x) = 1\}$  denotes the binary approximation of the continuous observed pore phase  $\theta \cap W, \theta = \{x \in \mathbb{R}^2 | \Theta(x) = 1\}$  and  $\phi(x)$  and  $\Phi(x)$  denote the density and cumulative distribution function of the standard-Normal-distribution.

The main step towards establishing Theorem 4.3.8, namely the proof of part iii), is provided by Lemma 4.3.6 and Corollary 4.3.5.

**Proof. (Theorem 4.3.8)**

**ad i).** Because counting grid points and counting lattice squares is not quite the same we need some more notation. Let's denote  $\Gamma^*(h) = \Gamma(h) \setminus \{\text{lower and right boundary of } W\}$ . Part i) is then a trivial observation because of:

$$\begin{aligned} \mathbf{E} [\hat{v}(\theta_\Gamma(h))] &= \mathbf{E} \left[ \frac{1}{n_0} \cdot \sum_{i=0}^7 h_{\theta_\Gamma(2i+1)} \right] = \mathbf{E} \left[ \frac{1}{n_0} \cdot \sum_{x_0 \in \Gamma^*(h)} \mathbf{1}_{\{x_0 \in \theta_\Gamma(h)\}} \right] \\ &= \mathbf{P}[x_0 \in \theta_\Gamma(h)] = \mathbf{P}[Z(x_0) \geq b] = 1 - \Phi(b). \end{aligned}$$

**ad ii).** The proof of asymptotical unbiasedness for the OM-specific-surface estimator is postponed to section 4.4 where we will prove asymptotical unbiasedness for a more general class of surface estimators. It turns out that the OM-surface-estimator considered here is a member of this class and hence the proof directly follows from applying Theorem 4.4.4 below.

**ad iii).** The proof of this part is the most involved one because four-dimensional Gaussian probabilities must be calculated. In fact we have because of the isotropy property:

$$\begin{aligned}
\mathbf{E} [\hat{\chi}] &= \mathbf{E} \left[ \frac{1}{4 \cdot n_0 \cdot h^2} \cdot \left( \sum_{i=1}^4 h_{\theta_\Gamma}(2^i) - \sum_{i=1}^4 h_{\theta_\Gamma}(15 - 2^i) + 2 \cdot (h_{\theta_\Gamma}(6) + h_{\theta_\Gamma}(9)) \right) \right] \\
&= \frac{1}{4 \cdot n_0 \cdot h^2} \sum_{\Gamma_{x_0}(h)} \left( \sum_{i=1}^4 \left[ \mathbf{P} [x_i \in \theta_\Gamma, x_i^c \subset \theta_\Gamma^c] - \mathbf{P} [x_i^c \subset \theta_\Gamma, x_i \in \theta_\Gamma^c] \right] \right. \\
&\quad \left. + 2 \cdot \mathbf{P} [\{x_0, x_3\} \subset \theta_\Gamma, \{x_1, x_2\} \subset \theta_\Gamma^c] + 2 \cdot \mathbf{P} [\{x_1, x_2\} \subset \theta_\Gamma, \{x_0, x_3\} \subset \theta_\Gamma^c] \right) \\
&= \frac{1}{h^2} \left( \mathbf{P} [x_0 \in \theta_\Gamma, x_0^c \subset \theta_\Gamma^c] - \mathbf{P} [x_0^c \subset \theta_\Gamma, x_0 \in \theta_\Gamma^c] + \mathbf{P} [\{x_0, x_3\} \subset \theta_\Gamma, \{x_1, x_2\} \subset \theta_\Gamma^c] \right) \\
&= \frac{1}{h^2} \left( \mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 < b] - \mathbf{P} [Z_0 < b, Z_1 \geq b, Z_2 \geq b, Z_3 \geq b] + \right. \\
&\quad \left. \mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] \right),
\end{aligned}$$

where we have slightly abused the notation by setting  $x_i^c = \{x_0, x_1, x_2, x_3\} \setminus \{x_i\}$ .

Please note that there are only two different types of four-dimensional Gaussian probabilities involved here. In fact, due to  $X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow -X \stackrel{d}{=} X$  and  $\text{Cov}(X, Y) = \text{Cov}(-X, -Y)$  which of course holds generally, we have:

$$\mathbf{P} [Z_0 < b, Z_1 \geq b, Z_2 \geq b, Z_3 \geq b] = \mathbf{P} [Z_0 \geq -b, Z_1 < -b, Z_2 < -b, Z_3 < -b].$$

Because of high positive correlation of  $Z_0, \dots, Z_3$  when the lattice spacing  $h$  is small, we expect the term  $\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b]$  to be of smaller magnitude than  $\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 < b]$ . From Lemma 4.3.6 we see that this intuition is true. In fact, the term  $\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b]$  is even negligible for our purposes. This is somehow reassuring that the ambiguous vertex configurations won't have a considerable influence on the characteristic as long as the lattice spacing is small enough to discern between constrictions of a single and the presence of two different connected components. A further benefit is of course that the lemma considerably simplifies our computational task. We note that what Lemma 4.3.6 provides is even more than we need. In fact a decay of order  $o(h^2)$  would already suffice for our purposes.

$$\begin{aligned}
\lim_{h \downarrow 0} \mathbf{E} [\hat{\chi}] &= \frac{1}{h^2} \left( \mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b] - \mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] - \right. \\
&\quad \left. \mathbf{P} [Z_0 \geq -b, Z_1 < -b, Z_2 < -b] + \mathbf{P} [Z_0 \geq -b, Z_1 < -b, Z_2 < -b, Z_3 \geq -b] + \right. \\
&\quad \left. \mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] \right)
\end{aligned}$$

$$\begin{aligned}
 &= \lim_{h \downarrow 0} \frac{1}{h^2} \left( \mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b] - \mathbb{P}[Z_0 \geq -b, Z_1 < -b, Z_2 < -b] + o(h^3) \right) \\
 &= \lim_{h \downarrow 0} \frac{1}{h^2} \left( \mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b] - \mathbb{P}[Z_0 \geq -b, Z_1 < -b, Z_2 < -b] \right), \quad (4.19)
 \end{aligned}$$

Hence it remains to evaluate  $\mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b]$  for arbitrary  $b \in \mathbb{R}$ . To calculate this probability we may directly apply the corollary to Lemma 4.3.3 with  $c_2 = 1$  and  $c_3 = \sqrt{2}$ . Recalling  $\tau = \sqrt{1 - \rho(h)}$  we learn from corollary 4.3.5, part i):

$$\begin{aligned}
 &\mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b] \\
 &= \tau \cdot \phi(b) \int_0^\infty \Phi^2\left(\frac{-u}{\sqrt{2}}\right) du + \tau^2 \cdot \left[ 2 \frac{\phi(b) \cdot b}{\sqrt{2}} \int_0^\infty \phi\left(\frac{u}{\sqrt{2}}\right) \cdot \Phi\left(-\frac{u}{\sqrt{2}}\right) du \right. \\
 &\quad \left. - b \cdot \phi(b) \int_0^\infty u \cdot \Phi^2\left(-\frac{u}{\sqrt{2}}\right) du \right] + \mathcal{O}(\tau^3).
 \end{aligned}$$

From equation 4.19 it becomes obvious that we only need to consider the  $\tau^2$ -term in the above expression because the terms which are even functions of  $b$  drop by taking the difference in equation 4.19. The remaining integrals in the  $\tau^2$ -term can be calculated using Lemma 4.3.1, parts iv) and v), to arrive at:

$$\begin{aligned}
 \lim_{h \downarrow 0} \mathbf{E}[\hat{\chi}] &= \lim_{h \downarrow 0} \frac{1}{h^2} \left[ 4 \frac{\phi(b) \cdot b}{\sqrt{2}} \cdot \frac{1}{4\sqrt{2}} - 2 \cdot b \cdot \phi(b) \cdot \left( \frac{1}{4} - \frac{1}{2\pi} \right) \right] \cdot \tau^2 + \mathcal{O}(\tau^3) \\
 &= -\frac{(\rho''(0))}{2\pi} \cdot b \cdot \phi(b), \quad (4.20)
 \end{aligned}$$

where we have used the definition of  $\tau$  in the last step. This concludes the proof.  $\square$

### 4.3.3 Three Dimensional Convergence Results

We now turn our attention to the three-dimensional situation. The main goal here is to prove the 3D-analogue to Theorem 4.3.8 for the 3D-Ohser-Mücklich estimators from subsection 2.3.2. The specific Minkowski-functionals in three dimensions are the porosity (or specific volume)  $\bar{v}$ , the specific surface area  $\bar{s}$ , the specific the integral of mean curvature  $\bar{M}$  and again the Euler-characteristic  $\bar{\chi}$ .

**Theorem 4.3.9. (Asymptotical Unbiasedness of OM-Estimators (3D-Version))**

Let  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  a three-dimensional stationary and isotropic Thresholded Gaussian Field. Then the mean Ohser-Mücklich estimates for the specific Minkowski functionals on a finite grid  $\Gamma(h)$  converge to the continuous limit, ie. the rates of the specific Minkowski-functionals in the continuous setting, if we let the grid-spacing  $h$  tend to zero.

- i)  $\lim_{h \downarrow 0} \mathbf{E}[\hat{v}(\theta_\Gamma(h))] = \bar{v}(\Theta) = 1 - \Phi(b).$
- ii)  $\lim_{h \downarrow 0} \mathbf{E}[\hat{s}(\theta_\Gamma(h))] = \bar{s}(\Theta) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \phi(b) \cdot \sqrt{-\rho''(0)}.$
- iii)  $\lim_{h \downarrow 0} \mathbf{E}[\hat{M}(\theta_\Gamma(h))] = \bar{M}(\Theta) = -b \cdot \phi(b) \cdot \rho''(0),$

where  $\theta_\Gamma(h) = \{x \in \Gamma(h) | \Theta(x) = 1\}$  denotes the binary approximation of the continuous observed pore phase  $\theta \cap W$ ,  $\theta = \{x \in \mathbb{R}^3 | \Theta(x) = 1\}$  and  $\phi(x)$  and  $\Phi(x)$  denote the density and cumulative distribution function of the Standard-Normal-distribution.

Also in three dimensions the covariance function  $\rho(h)$  of the underlying Gaussian field enters the formulac only through the second derivative at zero. A further interesting observation is that the 3D-Euler-characteristic changes sign twice at thresholds  $b = \pm 1$  (see below).

**Remark 4.3.10.** Of course there is also a formula for the 3D-Euler-characteristic for Thresholded Gaussian Fields  $\Theta$ . In the book of Adler (Adler 1981) the following expression for the continuous limit from corollary 4.2.4 is proved:

$$\bar{\chi}(\Theta) = \frac{1}{(2 \cdot \pi)^{3/2}} \cdot (b^2 - 1) \cdot \phi(b) \cdot (-\rho''(0))^{\frac{3}{2}}.$$

Because under the conditions of Theorem 4.1.5 the realizations of a Thresholded Gaussian Field are  $C^2$ -manifolds, the curvature at the boundary points within  $W$  is always defined and finite and thus the realizations should be morphologically open and closed almost surely. But then summation and limit in the Hadwiger definition of the Euler-characteristic (definition 2.1.12) may be exchanged. Thus we conjecture that in analogy to Theorems 4.3.8 and 4.3.9 we also have asymptotical unbiasedness for the 3D-Euler-characteristic:

$$\lim_{h \downarrow 0} \mathbf{E} [\hat{\chi}] (\theta_\Gamma(h)) = \bar{\chi}(\Theta) = \frac{1}{(2 \cdot \pi)^{3/2}} \cdot (b^2 - 1) \cdot \phi(b) \cdot (-\rho''(0))^{\frac{3}{2}}, \quad (4.21)$$

but although having invested considerable effort and time, we have not been able to prove equation 4.21 by direct calculation so far. If working with the usual 6-adjacency, the main difficulty is that in order to calculate  $\mathbf{E} [\hat{\chi}]$  explicitly, we need to determine Gaussian probabilities involving all the eight dependent jointly-Gaussian variates corresponding to the eight vertices of the typical cubic lattice cell. Because the Euler-characteristic is dimensionless, the involved probabilities have to be expanded at least up to order  $\mathcal{O}(h^3)$ . However, taking into account Lemma 4.3.6 of the 256 binary vertex-configurations we may neglect those that contain a either a cell-face or a rectangle bounded by two cell-face-edges and two cell-face-diagonals whose 4 vertices are of alternating colour (eg. clockwise black-white-black-white). Further exploiting symmetries and complementarities only 3 of the 256 vertex-configurations need to be considered to calculate the Euler-characteristic, namely:

$$\begin{aligned} & \mathbf{P} [Z_0 \geq b, Z_1 \leq b, \dots, Z_8 \leq b], \quad \mathbf{P} [Z_0 \geq b, \dots, Z_2 \geq b, Z_3 \leq b, \dots, Z_7 \leq b], \\ & \mathbf{P} [Z_0 \geq b, \dots, Z_2 \geq b, Z_3 \leq b, Z_4 \geq b, Z_5 \leq b, \dots, Z_7 \leq b]. \end{aligned}$$

Unfortunately, this does not remove the difficulty of calculating the probabilities for eight-tuples. Not even with the assumption of linearity for the underlying Gaussian field, ie.

$$Z(x) = Z(0) + \left[ \frac{\partial Z(x)}{\partial x} (0) \right]^T \cdot x \text{ for } x \in \mathbb{R}^d,$$

we were able to prove equation 4.21. It turned out that this considerable simplification is not even feasible in the two-dimensional situation in the sense that with the linear

approximation we cannot reproduce the result of the 2D-Theorem 4.3.8, part iii), and there is no reason to believe that the linear approximation is accurate enough for  $d = 3$  if it already fails for  $d = 2$ .

Another approach we performed used the fact that the Euler-characteristic of a set does not depend on the edge-graph on which it is calculated. However as already mentioned in the discussion of the Boolean model its estimation on a discrete lattice may be biased. Instead of using the 6-adjacency induced by an ordinary cubic lattice one might consider other more complicated graphs defined by different adjacencies and it might well be that the configurations needed to calculate the Euler characteristic for other adjacencies might be better analytically tractable. Of course one must be aware that different adjacencies might well lead to different biases. However we were able to set up a small computer algorithm to show that the expectations of the OM-estimator for the specific Euler-characteristic with respect to all the adjacencies considered in (Ohser et al. 2003) only differ in contributions that stem from negligible configurations in the sense of Lemma 4.3.6. This means that the differences of the biases for the different adjacencies are of order  $\mathcal{O}(h^4)$ . Therefore the approach to consider different adjacencies to prove equation 4.21 is well-justified. Using the 14.1-adjacency defined in (Ohser et al. 2003) which decomposes three-dimensional space into tetrahedrons we were able to show that:

$$\begin{aligned} \mathbf{E} [\hat{\chi}] &= \frac{1}{h^3} \cdot (\mathbf{P} [Z_0 \geq b, Z_1 \leq b, Z_3 \leq b, Z_7 \leq b] + \mathbf{P} [Z_0 \geq b, Z_4 \geq b, Z_6 \geq b, Z_7 \leq b] \\ &\quad - \mathbf{P} [Z_0 \geq b, Z_1 \leq b, Z_5 \geq b, Z_7 \leq b] - \mathbf{P} [Z_0 \geq b, Z_2 \geq b, Z_3 \leq b, Z_7 \leq b] \\ &\quad - \mathbf{P} [Z_0 \geq b, Z_2 \leq b, Z_6 \geq b, Z_7 \leq b] - \mathbf{P} [Z_0 \geq b, Z_4 \geq b, Z_5 \leq b, Z_7 \leq b]), \end{aligned}$$

where  $Z_0, \dots, Z_7$  are the standard Gaussian variates associated with the vertices of a lattice cube numbered according to the nomenclature of equation 2.27. Hence we only need to consider *quadruple*-probabilities. Explicit calculation of the involved integrals, however, remains unfeasible.

The main part of the proof of Theorem 4.3.9 to which we now proceed can again be derived from the fundamental Lemma 4.3.3.

**Proof. (Theorem 4.3.9)**

**ad i).** Again the proof for the specific-volume estimator  $\hat{v}$  is trivial and can be copied from the 2D-version:

$$\begin{aligned} \mathbf{E} [\hat{v}(\theta_\Gamma(h))] &= \mathbf{E} \left[ \frac{1}{n_0} \cdot \sum_{i=0}^{127} h_{\theta_\Gamma}(2i+1) \right] = \mathbf{E} \left[ \frac{1}{n_0} \cdot \sum_{x_0 \in \Gamma^*(h)} \mathbf{1}_{\{x_0 \in \theta_\Gamma(h)\}} \right] \\ &= \mathbf{P} [x_0 \in \theta_\Gamma(h)] = \mathbf{P} [Z(x_0) \geq b] = 1 - \Phi(b), \end{aligned}$$

where  $\Gamma^*(h) = \Gamma(h) \setminus \{\text{lower and right boundary of } W\}$  is the boundary-reduced lattice.

**ad ii).** The proof for the specific surface estimator can be postponed to section 4.4 where an analogous result will be proved (Theorem 4.4.4) for a more general class of surface estimators of which also the 3D-OM-surface estimator defined in equation 2.29 is a member.

**ad iii).** The main part of the proof is thus devoted to the OM-estimator of the specific integral of mean curvature  $\bar{M}$ . For its expectation we have:

$$\lim_{h \downarrow 0} \mathbf{E} \left[ \hat{M}(\theta_\Gamma(h)) \right] = 2\pi \cdot \lim_{h \downarrow 0} \left( \mathbf{E} [\mathcal{P}_{\nu_{\text{faces}}}] + \mathbf{E} [\mathcal{P}_{\nu_{\text{diagonals}}}] + \mathbf{E} [\mathcal{P}_{\nu_{\text{trig}}}] \right). \quad (4.22)$$

We can now have a look at the three types of contributions separately. We can use the familiar isotropy argument. The expected directional contributions  $\mathbf{E} [\mathcal{P}_\nu]$  of all  $2 \cdot 3$  cell faces are thus the same. With lemma 4.3.6 they can be calculated as:

$$\begin{aligned} \mathbf{E} [\mathcal{P}_{\nu_{\text{faces}}}] &= \frac{6 \cdot c_0}{h^2} \cdot (\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 < b] - \mathbf{P} [Z_0 \geq b, Z_1 \geq b, Z_2 \geq b, Z_3 < b]) \\ &= \frac{6 \cdot c_0}{h^2} \cdot (\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_2 < b] - \mathbf{P} [Z_0 \geq -b, Z_1 < -b, Z_2 < -b] + o(h^3)). \end{aligned}$$

Similarly for the  $2 \cdot 6$  mean diagonal rectangle contributions we get:

$$\begin{aligned} \mathbf{E} [\mathcal{P}_{\nu_{\text{diagonals}}}] &= \\ &= \frac{12 \cdot c_3}{\sqrt{2} \cdot h^2} \cdot (\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_6 < b, Z_7 < b] - \mathbf{P} [Z_0 \geq b, Z_1 \geq b, Z_6 \geq b, Z_7 < b]) = \\ &= \frac{12 \cdot c_3}{\sqrt{2} \cdot h^2} \cdot (\mathbf{P} [Z_0 \geq b, Z_1 < b, Z_6 < b] - \mathbf{P} [Z_0 \geq -b, Z_1 < -b, Z_6 < -b] + o(h^3)). \end{aligned}$$

And finally the mean contributions of the  $2 \cdot 4$  triangle planes are:

$$\mathbf{E} [\mathcal{P}_{\nu_{\text{trig}}}] = \frac{8 \cdot c_9}{\sqrt{3} \cdot h^2} \cdot (\mathbf{P} [Z_0 \geq b, Z_5 < b, Z_6 < b] - \mathbf{P} [Z_0 \geq b, Z_5 \geq b, Z_6 < b]).$$

The contributions  $\mathbf{E} [\mathcal{P}_{\nu_{\text{faces}}}]$  of the cell faces have already been calculated in the proof of Theorem 4.3.8, part iii). From equations 4.19 and 4.20 we learn:

$$\mathbf{E} [\mathcal{P}_{\nu_{\text{faces}}}] = 6 \cdot c_0 \cdot \left( -\frac{1}{2\pi} \cdot b \cdot \phi(b) \cdot \rho''(0) \right) + \mathcal{O}(h).$$

For the rectangular contributions  $\mathbf{E} [\mathcal{P}_{\nu_{\text{diagonals}}}]$  we can apply Corollary 4.3.5, part i), with  $c_2 = \sqrt{2}$  and  $c_3 = \sqrt{3}$  (Note that the  $c_i$  of Lemma 4.3.3 and the Voronoi-coefficients  $c_i$  in the contributions above are different!) to find:

$$\begin{aligned} \mathbf{P} [Z_0 \geq b, Z_1 \leq b, Z_6 \leq b] &= \\ &= \tau \cdot \phi(b) \int_0^\infty \Phi \left( -\frac{u}{\sqrt{2}} \right) \cdot \Phi \left( -\frac{u}{2} \right) du + \phi(b) \cdot b \cdot \left[ \frac{1}{\sqrt{2}} \left( \int_0^\infty \Phi \left( -\frac{u}{2} \right) \cdot \phi \left( \frac{u}{\sqrt{2}} \right) du \right. \right. \\ &\quad \left. \left. + \sqrt{2} \int_0^\infty \Phi \left( -\frac{u}{\sqrt{2}} \right) \cdot \phi \left( \frac{u}{2} \right) du \right) - \int_0^\infty u \cdot \Phi \left( -\frac{u}{\sqrt{2}} \right) \cdot \Phi \left( -\frac{u}{2} \right) du \right] \cdot \tau^2 + \mathcal{O}(\tau^3). \end{aligned}$$

Again by taking differences only the terms which are odd-functions in  $b$  survive, hence we need only consider the  $\mathcal{O}(\tau^2)$ -term. The integrals can be solved using Lemma 4.3.1, parts v). and vii.). This leads to:

$$\begin{aligned}
\mathbf{E}[\mathcal{P}_{\nu_{\text{diagonals}}}] &= \frac{12 \cdot c_3}{\sqrt{2} \cdot h^2} \cdot b \cdot \phi(b) \left[ \frac{2}{\sqrt{2}} \cdot \left( \frac{-\sqrt{2}}{4} + \frac{\sqrt{2}}{2} \cdot \frac{\pi + 2 \cdot \arctan \sqrt{2}}{2\pi} \right. \right. \\
&\quad \left. \left. + \sqrt{2} \cdot \frac{\pi - 2 \cdot \arctan \sqrt{2}}{2\pi} \right) - 2 \cdot \left( \frac{1}{2} - \frac{1}{2\pi} \arctan \sqrt{2} - \frac{1}{\sqrt{2} \cdot \pi} \right) \right] \cdot \tau^2 + \mathcal{O}(\tau^3). \\
&= \frac{12 \cdot c_3}{\sqrt{2}} \left( -\frac{1}{\sqrt{2} \cdot \pi} \cdot b \cdot \phi(b) \cdot \rho''(0) \right) + \mathcal{O}(h),
\end{aligned}$$

where we have used  $\tau = \sqrt{1 - \rho(h)}$  in the last step.

It remains to consider the triangular contributions  $\mathbf{E}[\mathcal{P}_{\nu_{\text{tri}}}]$ . The triangle spanned by  $Z_0, Z_1$  and  $Z_6$  is equilateral with side-length  $\sqrt{2}$ . Therefore we may apply corollary 4.3.5, part ii), and the rescaled lattice spacing  $\tilde{h} = \sqrt{2} \cdot h$ . Consequently we use 4.3.5, part ii), for  $\tau(\tilde{h}) := \sqrt{1 - \rho(\tilde{h})}$  to arrive at:

$$\begin{aligned}
\mathbf{P}[Z_0 \geq b, Z_5 \leq b, Z_6 \leq b] &= \tilde{\tau} \cdot \phi(b) \cdot \int_0^\infty \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{\sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{array} \right) \right] du + b \cdot \phi(b) \cdot \\
&\quad \left[ \sqrt{2} \cdot \int_0^\infty \phi \left( \frac{u}{\sqrt{2}} \right) \cdot \Phi \left( -\sqrt{\frac{3}{2}} \cdot u \right) du \right. \\
&\quad \left. - \int_0^\infty u \cdot \Phi_2 \left[ \left( \begin{array}{c} -\frac{u}{\sqrt{2}} \\ -\frac{u}{\sqrt{2}} \end{array} \right), \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{array} \right) \right] du \right] \cdot \tilde{\tau}^2 + \mathcal{O}(\tilde{\tau}^3). \tag{4.23}
\end{aligned}$$

Because the triangle  $(Z_0, Z_5, Z_6)$  is not rectangular, the two-dimensional Gaussian distribution-functions do not decouple which makes the integration a bit more tedious. By the usual symmetry argument when taking differences we only need to consider the  $\tilde{\tau}^2$ -term.

As a next step we solve the last integral involving the non-factorizing two-dimensional Gaussian distribution function  $\Phi_2(\cdot)$  by a transformation to independent variables. To do this consider the Gaussian random variables  $X_1, X_2$  with  $\text{Var}(X_1) = \text{Var}(X_2) = 1$  and  $\text{Cov}(X_1, X_2) = \frac{1}{2}$  and let us introduce the pair of transformed Gaussian variables  $W_1, W_2$  as:

$$W_1 := X_1 + X_2, \quad W_2 := X_1 - X_2.$$

Then we obviously have:

$$\text{Var}(W_1) = 3, \quad \text{Var}(W_2) = 1, \quad \text{Cov}(W_1, W_2) = 0,$$

hence  $W_1$  and  $W_2$  are independent. Thus:

$$\begin{aligned}
& \int_0^\infty u \cdot \mathbf{P} \left[ X_1 \leq -\frac{u}{\sqrt{2}}, X_2 \leq -\frac{u}{\sqrt{2}} \right] du \\
&= \frac{1}{\sqrt{3}} \int_0^\infty \int_{-\infty}^{-\sqrt{2}u} \int_{\sqrt{2}u+w_1}^{\sqrt{2}u-w_1} u \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) \cdot \phi(w_2) dw_2 dw_1 du \\
&= \frac{1}{\sqrt{3}} \int_0^\infty \int_{-\infty}^{-\sqrt{2}u} u \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) \cdot [1 - 2 \cdot \Phi(\sqrt{2} \cdot u + w_1)] dw_1 du \\
&= \frac{1}{\sqrt{3}} \int_{-\infty}^0 \int_0^{-\frac{w_1}{\sqrt{2}}} u \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) \cdot [1 - 2 \cdot \Phi(\sqrt{2} \cdot u + w_1)] du dw_1 \\
&= \frac{1}{4\sqrt{3}} \int_{-\infty}^0 w_1^2 \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) dw_1 - \frac{1}{2\sqrt{3}} \int_{-\infty}^0 w_1^2 \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) \cdot \Phi(w_1) dw_1 \\
&\quad - \int_{-\infty}^0 w_1 \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) \cdot \phi(w_1) dw_1 + \frac{1}{4\sqrt{3}} \int_{-\infty}^0 \phi\left(\frac{w_1}{\sqrt{3}}\right) dw_1 \\
&\quad - \frac{1}{2\sqrt{3}} \int_{-\infty}^0 \phi\left(\frac{w_1}{\sqrt{3}}\right) \cdot \Phi(w_1) dw_1 + \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{3}} \int_{-\infty}^0 w_1 \cdot \phi\left(\frac{w_1}{\sqrt{3}}\right) dw_1 \\
&= \frac{1}{3} - \frac{\sqrt{3}}{2} \cdot \frac{1}{2\pi}, \tag{4.24}
\end{aligned}$$

where we have used Lemma 4.3.1, parts i),ii),v),xi) and x) and some simple substitution and integration by parts-arguments.

Combining equations 4.23 and 4.24 and a further application of Lemma 4.3.1, part v), takes us to:

$$\begin{aligned}
\mathbf{E} [\mathcal{P}_{\nu_{\text{orig}}}] &= \frac{8 \cdot c_9}{\sqrt{3} \cdot h^2} \cdot \left( b \cdot \phi(b) \cdot \left[ 2 \cdot \sqrt{2} \cdot \frac{\sqrt{2}}{6} - 2 \cdot \left( \frac{1}{3} - \frac{\sqrt{3}}{2} \cdot \frac{1}{2\pi} \right) \right] \cdot \tilde{\tau}^2 + \mathcal{O}(\tilde{\tau}^3) \right) \\
&= \frac{8 \cdot c_9}{\sqrt{3} \cdot h^2} \cdot \left( b \cdot \phi(b) \cdot \frac{\sqrt{3}}{2\pi} \cdot \tilde{\tau}^2 + \mathcal{O}(\tilde{\tau}^3) \right) = \frac{8 \cdot c_9}{\sqrt{3} \cdot h^2} \cdot \left( -b \cdot \phi(b) \cdot \frac{\sqrt{3}}{2\pi} \cdot \frac{\rho''(0)}{2} \cdot \tilde{h}^2 + \mathcal{O}(\tilde{h}^3) \right) \\
&= \frac{8 \cdot c_9}{\sqrt{3}} \cdot \left( -\frac{\sqrt{3}}{2\pi} \cdot b \cdot \phi(b) \cdot \rho''(0) + \mathcal{O}(h) \right).
\end{aligned}$$

Finally plugging the three contributions into equation 4.22 gives:

$$\begin{aligned}
\lim_{h \downarrow 0} \mathbf{E} [\hat{M}] &= 2\pi \cdot \left[ 6 \cdot c_0 \cdot \left( -\frac{1}{2\pi} \cdot b \cdot \phi(b) \cdot \rho''(0) \right) + \frac{12 \cdot c_3}{\sqrt{2}} \cdot \left( -\frac{1}{\sqrt{2\pi}} \cdot b \cdot \phi(b) \cdot \rho''(0) \right) \right. \\
&\quad \left. + \frac{8 \cdot c_9}{\sqrt{3}} \cdot \left( -\frac{\sqrt{3}}{2\pi} \cdot b \cdot \phi(b) \cdot \rho''(0) \right) \right] \\
&= -b \cdot \phi(b) \cdot \rho''(0) \cdot (6 \cdot c_0 + 12 \cdot c_3 + 8 \cdot c_9) = -b \cdot \phi(b) \cdot \rho''(0),
\end{aligned}$$

where we have used the fact that the Voronoi-coefficients add up to one. This proves the theorem.  $\square$



## 4.3.4 Bias of the 2D-Euler-Characteristic Estimator

By expanding the probability  $P[Z(x_1) \geq b, Z(x_2) \leq b, Z(x_3) \leq b]$  further than only to first order in  $\tau = \sqrt{1 - \rho(h)}$  as given in Lemma 4.3.3, one can find expressions for the biases of the OM-type estimators. Because by the symmetry arguments applied above the expressions for the expected values of these estimators only consist of terms which are odd functions of the threshold  $b$ , for the Euler-characteristic an expansion to at least fourth order in  $\tau$  is needed. This involves the calculation of the third derivative of the function  $f(u, \tau)$  in the proof of Lemma 4.3.3. The resulting calculations are extremely tedious and the Gaussian integrals can no longer be solved in general. However, for the 2D-Euler characteristic considerable simplifications can be made. This is because the decoupling effect leads to integrals involving only products of one-dimensional Gaussian distribution and density functions and polynomials which can still be solved analytically. Hence with some effort an explicit expression for the bias can be obtained. This is the content of the following theorem.

**Theorem 4.3.11.** (*Bias of the 2D-OM-estimator for Euler-characteristic*) Let  $\Theta(x) = \mathbf{1}_{Z(x) \geq b}$  a two-dimensional stationary, isotropic, non-erratic Thresholded Gaussian Field. Then the bias of the two-dimensional Ohser-Mücklich estimator for the specific Euler-characteristic given in equation 2.26 when used on a finite grid  $\Gamma(h)$  with lattice spacing  $h$  is as follows:

$$\mathbf{E}[\hat{\chi}(\Theta)] - \bar{\chi} = -\phi(b) \cdot b \cdot \frac{\rho''(0)^2 \cdot (4b^2 - 18) + \rho^{(iv)}(0) \cdot (6 + \pi)}{96\pi} \cdot h^2 + \mathcal{O}(h^4).$$

**Proof. (Theorem 4.3.11)** Because the calculations involved here could easily fill a dozen pages without offering the reader some deeper insight, we content ourselves with giving a few intermediate results. We use the notation introduced in the proof of Lemma 4.3.3. There one can easily see that the second derivative of  $f(u, \tau)$  evaluated in  $\tau = 0$  is an even function in  $b$  and thus does not contribute to the bias. Hence contributions to the bias come from terms up to order  $\mathcal{O}(\tau^4)$ .

Expanding the expressions defined in equation-systems 4.25 and 4.14 a bit further we find:

$$\begin{aligned} b_1(\tau) &= -\frac{u}{\sqrt{2}} + \frac{b}{\sqrt{2}}\tau + \frac{3u}{4\sqrt{2}}\tau^2 + \frac{b}{4\sqrt{2}} \cdot \tau^3 + \mathcal{O}(\tau^4), \\ b_2(\tau) &= -\frac{u}{c_2\sqrt{2}} + \frac{bc_2}{\sqrt{2}} \cdot \tau + \frac{((1 - c_2^2)\rho^{(iv)}(0) + 9\rho''(0)^2 c_2^2) u}{12\sqrt{2}\rho''(0)^2 c_2} \cdot \tau^2 \\ &\quad + \frac{bc_2 \cdot (\rho^{(iv)}(0) + 3\rho''(0)^2 c_2^2 - \rho^{(iv)}(0)c_2^2)}{12\sqrt{2}\rho''(0)^2} \cdot \tau^3 + \mathcal{O}(\tau^4), \\ \rho(\tau) &= \frac{1 + c_2^2 - c_3^2}{2c_2} + \frac{1}{24\rho''(0)^2 c_2} \left[ -\rho^{(iv)}(0) \cdot (1 + c_2^4 + c_3^2 - 2c_3^4 + c_2^2(c_3^2 - 2)) \right. \\ &\quad \left. + 3\rho''(0)^2 (1 + c_2^4 - c_3^2 - c_2^2(2 + c_3^2)) \right] \cdot \tau^2 + \mathcal{O}(\tau^4), \end{aligned}$$

and:

$$\begin{aligned}
\frac{b_2(\tau) - \rho(\tau)b_1(\tau)}{\sqrt{1 - \rho^2(\tau)}} &= \frac{(c_2^2 - c_3^2 - 1) \cdot u}{\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} + \frac{b \cdot (c_2^2 + c_3^2 - 1)}{\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} \cdot \tau \\
&+ \left\{ \frac{-4\rho^{(iv)}(0)c_3^2[-2c_2^4 + (c_3^2 - 1)^2 + c_2^2(1 + c_3^2)]}{12\sqrt{2}\rho''(0)^2[2c_2^2(1 + c_3^2) - (c_3^2 - 1)^2 - c_2^4]^{\frac{3}{2}}} \right. \\
&+ \left. \frac{3\rho''(0)^2[c_2^6 - 3c_2^4(1 + c_3^2) + (c_3^2 - 1)^2(3c_3^2 - 1) - c_2^2(-3 + 2c_3^2 + c_3^4)]}{12\sqrt{2}\rho''(0)^2[2c_2^2(1 + c_3^2) - (c_3^2 - 1)^2 - c_2^4]^{\frac{3}{2}}} \right\} \cdot u \cdot \tau^2 + \mathcal{O}(\tau^3), \\
\frac{b_1(\tau) - \rho(\tau)b_2(\tau)}{\sqrt{1 - \rho^2(\tau)}} &= \frac{(1 - c_2^2 - c_3^2) \cdot u}{c_2\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} + \frac{b \cdot c_2 \cdot (1 - c_2^2 + c_3^2)}{\sqrt{2(2c_2^2(1 + c_3^2) - (c_3 - 1)^2 - c_2^4)}} \cdot \tau \\
&+ \left\{ \frac{\rho^{(iv)}(0)[c_2^8 - (c_3^2 - 1)^3 + c_2^6(3c_3^2 - 4) + c_2^4(6 + 3c_3^2 - 9c_3^4) + c_2^2(-4 - 3c_3^2 + 2c_3^4 + 5c_3^6)]}{12\sqrt{2}\rho''(0)^2c_2[2c_2^2(1 + c_3^2) - (c_3^2 - 1)^2 - c_2^4]^{\frac{3}{2}}} \right. \\
&+ \left. \frac{3\rho''(0)^2c_2^2[-1 + c_2^6 + 3c_3^2 + c_3^4 - 3c_3^6 - c_2^4(3 + 5c_3^2) + c_2^2(3 + 2c_3^2 + 7c_3^4)]}{12\sqrt{2}\rho''(0)^2c_2[2c_2^2(1 + c_3^2) - (c_3^2 - 1)^2 - c_2^4]^{\frac{3}{2}}} \right\} \cdot u \cdot \tau^2 + \mathcal{O}(\tau^3)
\end{aligned}$$

For the third derivative we also need to take into account the derivatives of

$$\Phi(\rho) := \Phi \left( \begin{pmatrix} b_1(\tau) \\ b_2(\tau) \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho(\tau) \\ \rho(\tau) & 1 \end{pmatrix} \right)$$

with respect to  $\rho$ . Because for the Euler characteristic we have  $x_1, x_2, x_3$  forming a rectangular triangle with  $c_2 = 1, c_3 = \sqrt{2}$  and thus  $\rho(\tau) \rightarrow 0$ , these derivatives become especially simple:

$$\begin{aligned}
\frac{\partial \Phi}{\partial \rho}(0) &= \phi^2 \left( \frac{u}{\sqrt{2}} \right), \\
\frac{\partial^2 \Phi}{\partial \rho^2}(0) &= \frac{u^2}{2} \phi^2 \left( \frac{u}{\sqrt{2}} \right).
\end{aligned}$$

Finally for the third derivative of  $f$  in zero for  $c_2 = 1, c_3 = \sqrt{2}$  we end up with:

$$\begin{aligned}
\frac{\partial^3 f}{\tau^3}(u, 0) &= \phi(b) \cdot \frac{7b \cdot (b^2 - 3)}{12\sqrt{2}} \cdot \int_0^\infty u^2 \cdot \phi \left( \frac{u}{\sqrt{2}} \right) \cdot \Phi \left( -\frac{u}{\sqrt{2}} \right) du \\
&+ \phi(b) \cdot b \cdot \left( \frac{1}{2} - \frac{b^2}{6} \right) \cdot \int_0^\infty u^3 \cdot \Phi^2 \left( -\frac{u}{\sqrt{2}} \right) du \\
&+ \phi(b) \cdot \left( -\frac{b^3}{4} + b - \frac{b}{6} \cdot \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right) \cdot \int_0^\infty u \cdot \phi^2 \left( \frac{u}{\sqrt{2}} \right) du \\
&+ \phi(b) \cdot \left( \frac{1}{18} \frac{\rho^{(iv)}(0)}{\rho''(0)^2} - \frac{1}{6} \right) \cdot \int_0^\infty u^2 \cdot \phi^2 \left( \frac{u}{\sqrt{2}} \right) du \\
&+ \phi(b) \cdot \frac{b}{\sqrt{2}} \cdot \left( -\frac{b^2}{6} + \frac{1}{2} \right) \cdot \int_0^\infty \phi \left( \frac{u}{\sqrt{2}} \right) \cdot \Phi \left( -\frac{u}{\sqrt{2}} \right) du.
\end{aligned}$$

The integrals can be worked out using Lemma 4.3.1, parts v),xii),xiii),xiv) to obtain:

$$\frac{\partial^3 f}{\tau^3}(u, 0) = \frac{\rho^{(iv)}(0) \cdot (\sqrt{2\pi} - 6b) - 3\rho''(0)^2 \cdot (2b^3 - 9b + \sqrt{2\pi})}{72\rho''(0)^2\pi} \cdot \phi(b)$$

Recalling once again that  $\tau = \sqrt{1 - \rho(h)}$  and even terms in  $b$  drop when taking the difference in equation 4.19, we see that the contribution to the bias of the  $\tau^4$ -term is:

$$\phi(b) \cdot \frac{\rho''(0)^2 \cdot (9b - 2b^3) - 2b \cdot \rho^{(iv)}(0)}{48\pi} \cdot h^2 + \mathcal{O}(h^4).$$

The other contribution that stems from the  $\tau^2$ -term can be derived from equation 4.20 and is:

$$\frac{-\phi(b) \cdot b \cdot \rho^{(iv)}(0) \cdot (2 + \pi)}{96\pi} \cdot h^2 + \mathcal{O}(h^4).$$

Hence we can conclude for the bias:

$$\mathbf{E} [\hat{\chi}(\Theta)] = -\phi(b) \cdot b \cdot \frac{\rho''(0)^2 \cdot (4b^2 - 18) + \rho^{(iv)}(0) \cdot (6 + \pi)}{96\pi} \cdot h^2 + \mathcal{O}(h^4),$$

which was to establish. □

We would like to mention that in principle also for the bias of the integral of mean curvature-estimator one could provide a (less explicit) expression using the same expansions above and work out the integrals numerically. We will not give this expression here. For the specific-surface estimators in both two and three dimensions to obtain explicit expressions for the bias is much easier because the calculation of their expectations only involves two-point-probabilities. The biases of the surface-estimators are given in Theorem 4.4.4 in section 4.4.

### 4.4 General Surface Estimators Based on the Covariance Function

In this section we still consider the Thresholded Gaussian Field  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  and generalize some results for the 2D- and 3D-Ohser-Mücklich estimators defined in equations 2.24 and 2.29 of the previous section 4.3 to a more general class. Besides proving asymptotical unbiasedness we will gain explicit information about the bias of these estimators when applied on a finite regular grid  $\Gamma$  with non-zero lattice spacing  $h$ . The class of estimators we consider in this section can be defined for general stationary and isotropic random fields  $\Theta$  and arises naturally from the fundamental relationship for stationary and isotropic random sets which is the content of Theorem 2.5.2. The relationship connects the two-point Covariance function  $C^{(2)}(h)$  given in definition 2.5.1 (we will suppress the superscript (2) in the following) with the specific surface area  $\bar{s}$ : The latter is - up to a dimension-dependent constant - equal to the first derivative of the covariance function in 0. Approximating the derivative by its difference quotient, we have that:

$$\lim_{h \downarrow 0} \sum_{j=1}^m \beta_j \frac{C(h \cdot |a_j|) - C(0)}{h} = -C'(0) \sum_{j=1}^m \beta_j |a_j|. \tag{4.25}$$

Hence because we have for arbitrary coefficients  $\beta_j \in \mathbb{R}$  and vectors  $a_j \in \mathbb{R}^d$ :

$$\begin{aligned} C(h \cdot |a_j|) - C(0) &= \\ \mathbf{P}[0 \in \Theta, h \cdot a_j \in \Theta] - \mathbf{P}[0 \in \Theta] &= -\mathbf{P}[0 \in \Theta, h \cdot a_j \in \Theta^c], \end{aligned} \tag{4.26}$$

on a finite grid  $\Gamma$  it is natural to consider estimators of the following class to estimate the specific surface area  $\bar{s}(\Theta)$ :

**Definition 4.4.1.** Let  $\Theta$  be a stationary and isotropic random field observed on a finite regular grid  $\Gamma$  with lattice spacing  $h$  consisting of  $n$  lattice points  $x_i$ . An estimator of the type

$$\hat{s} = -\frac{1}{nh_n} \sum_{i=1}^n \sum_{j=1}^m \beta_j \cdot \mathbf{1}_{\{x_n \in \Theta, x_n + h_n \cdot a_j \in \Theta^c\}}, \quad (4.27)$$

is called a **general specific surface estimator**, if  $a_j \in \mathbb{Z}^d, j \in 1, \dots, m$  are vectors such that  $x_i + h \cdot a_j$  are lattice points  $\forall i, j$  and  $\beta_j \in \mathbb{R}, j \in 1, \dots, m$  are suitable weights such that:

$$\sum_{j=1}^m \beta_j \cdot |a_j| = -\frac{d \cdot b_d}{b_{d-1}}, \quad (4.28)$$

for  $b_d$  the volume of the unit sphere  $S^{d-1}$ .

Hence in the case of the Thresholded Gaussian Field a general specific surface estimator is of the form:

$$\hat{s} = -\frac{1}{nh} \cdot \sum_{i=1}^n \sum_{j=1}^m \beta_j \cdot \mathbf{1}_{\{Z(x_i) \geq b, Z(x_i + h \cdot a_j) \leq b\}}.$$

**Remark 4.4.2.** The reason for the normalizing condition 4.28 in definition 4.4.1 is that it ensures the asymptotic unbiasedness property 4.25 for general specific surface estimators as will be proved in Theorem 4.4.4 below. Also note that definition 4.4.1 is completely dimension-free. Note also that we do not take into account any edge effects to keep the notation uncluttered, but of course for real applications the  $x_i$  must be the points of a reduced subgrid  $\Gamma_{\text{red}}$  such that still all  $x_i + h \cdot a_j$  are contained in  $\Gamma$  and thus can be observed.

**Example 4.4.3.** Both Ohser-Mücklich surface-estimators for 2D (equation 2.24) and 3D (equation 2.29) are examples of general surface-estimators in the sense of definition 4.4.1. In fact for the 2D-estimator we have:

$$\begin{aligned} m &= 8, \\ |a_1| &= |a_3| = |a_5| = |a_7| = 1, \\ |a_2| &= |a_4| = |a_6| = |a_8| = \sqrt{2}, \\ \beta_j &= -\frac{1}{|a_j|} \cdot \frac{\pi}{8}, \quad j \in \{1, \dots, 8\}. \end{aligned}$$

and thus:

$$\sum_{j=1}^m \beta_j \cdot |a_j| = -\pi = -\frac{2 \cdot \pi}{2} = -\frac{2 \cdot b_2}{b_1}.$$

And for the 3D-estimator we see:

$$\begin{aligned}
m &= 26, \\
|a_1| &= |a_2| = |a_3| = 1, |a_4| = \dots = |a_9| = \sqrt{2}, |a_{10}| = \dots = |a_{13}| = \sqrt{3}, \\
|a_{13+k}| &= |a_k| \text{ for } k \in \{1, \dots, 13\}, \\
\beta_j &= -\frac{4}{|a_j|} c_j, \quad j \in \{1, \dots, 26\}
\end{aligned}$$

where the  $c_j$  are the *Voronoi-coefficients* from equation 2.29 and hence:

$$\sum_{j=1}^m \beta_j \cdot |a_j| = -4 \sum_{i=1}^{26} c_i = -4 = -\frac{3 \cdot \frac{4\pi}{3}}{\pi} = -\frac{3 \cdot b_3}{b_2}.$$

In section 4.5 we will also consider the 1D-Ohser-Mücklich surface-estimator. First we have to clarify what we mean by specific surface area in one dimension and how its Ohser-Mücklich estimator can be defined. Note that in 1D there are only two Minkowski-functionals, namely the segment length  $V_1^1(\cdot)$  and the Euler characteristic  $V_0^1(\cdot)$ . From the discussion in Chapter 2 we know that the surface area  $s$  equals  $2V_{d-1}$  and thus in  $d = 1$  is just the doubled Euler-characteristic. This is in accordance with a general definition of specific surface area which is provided by stochastic geometry. Let  $\Theta$  be a random closed set, then the specific surface area  $\bar{s}$  for arbitrary dimension  $d \geq 1$  is the density of the *surface measure*  $S$ :

$$S_{\Theta}(B) = h_{d-1}(B \cap \partial\Theta),$$

where  $B \in \mathcal{B}(R^d)$  is a Borel-set and  $h_d$  is the  $d$ -dimensional Hausdorff-measure. Thus for  $d = 1$  the surface measure simply counts all the phase transitions from solid into pore space and vice versa. Note that this definition is in agreement with the surface area as it is defined for (finite) one-dimensional objects in two-dimensional space. The surface area of a line segment  $l$  in 2D is proportional to  $V_1^2(l)$  and equals *twice* the length of  $l$ . Hence the one-dimensional Ohser-Mücklich surface-estimator is simply:

$$\hat{s}_n(Z) = \frac{1}{n \cdot h_n} \sum_{i=1}^n (\mathbf{1}_{Z(ih_n) \in \Theta} \cdot \mathbf{1}_{Z((i+1)h_n) \in \Theta^c} + \mathbf{1}_{Z(ih_n) \in \Theta^c} \cdot \mathbf{1}_{Z((i+1)h_n) \in \Theta}). \quad (4.29)$$

Also this estimator is a general surface estimator according to definition 4.4.1 because:

$$\begin{aligned}
m &= 1, \quad |a_1| = 1, \quad |a_2| = |-1| = 1, \quad \beta_1 = \beta_2 = -\frac{1}{|a_i|}, \quad i \in \{1, 2\}, \\
\beta_1 \cdot |a_1| + \beta_2 \cdot |a_2| &= -2 = -\frac{1 \cdot 2}{1} = -\frac{1 \cdot b_2}{b_1}.
\end{aligned}$$

The following result collects some important properties about general surface estimators in arbitrary dimension  $d$ :

**Theorem 4.4.4. (Properties of General Surface Estimators)** *Consider a stationary and isotropic Thresholded Gaussian Field  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  in  $\mathbb{R}^d$  with a covariance function  $\rho(h) \in C^4$  that is observed on a finite regular grid  $\Gamma$ . Let further  $\hat{s}$  be a general surface estimator according to 4.4.1 with vectors  $a_j \in \mathbb{R}^d$  and coefficients  $\beta_j \in \mathbb{R}$  such that equation 4.28 holds. Then the following statements about  $\hat{s}$  hold true:*

- i).  $\hat{s}$  is an asymptotically unbiased estimator for  $h \downarrow 0$ .
- ii). The bias of  $\hat{s}$  is of order  $\mathcal{O}(h^2)$  and can explicitly be given as:

$$\mathbf{E}[\hat{s}(h)] - s = \frac{1}{24} \sum_{j=1}^m \beta_j |a_j|^3 \phi(0) \phi(b) (-\rho''(0))^{\frac{3}{2}} \left( (b^2 - 1) + \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right) \cdot h^2 + \mathcal{O}(h^3).$$

Thus for  $\sum_{j=1}^m \beta_j |a_j|^3 = 0$  we can even obtain an  $\mathcal{O}(h^3)$ -biased general surface estimator.

- iii). If  $m = 2$  and  $\rho^{(iv)}(0) \neq 0$ , the  $\mathcal{O}(h^3)$ -biased estimators are exactly those which fulfill:

$$\beta_1 = -\frac{|a_2|^2 \cdot c_d}{|a_1|(|a_1|^2 - |a_2|^2)}, \quad \beta_2 = \frac{|a_1|^2 \cdot c_d}{|a_2|(|a_1|^2 - |a_2|^2)}, \quad |a_1| \neq |a_2| \neq 0.$$

$$\text{where } c_d = -\frac{d \cdot b_d}{b_d - 1}.$$

**Remark 4.4.5.**

- i) The fact that the bias for the surface-estimator in the Thresholded Gaussian Field model is of order  $\mathcal{O}(h^2)$  seems quite surprising at first glance. This result heavily relies on the assumption that  $\rho \in C^4$ . Relaxing this assumption means less differentiability of the underlying Gaussian field resulting in less smooth boundaries of the excursion set. In fact if  $\rho \in C^4$  doesn't hold, the bias is of order  $\mathcal{O}(h)$ . This is in accordance with a (quite heuristic) argument in (Frisch and Stillinger 1963) where it is shown by differential-geometric arguments that for random sets whose boundaries have no edges, corners, multiple points or generally no singular points of infinite curvature an  $\mathcal{O}(h^2)$ -bias can be achieved. The paper also provides an example of a random system consisting of impenetrable identical 3D-spheres which may exhibit inter-particle-contacts and thus have singular contact points. It is shown that the bias for this system is  $\mathcal{O}(h)$ . The same problem also arises in the Boolean Model where the inter-particle contacts also lead to singular boundary points. Such realizations are obviously not morphologically closed.
- ii) Choosing  $\sum_{j=1}^m \beta_j |a_j|^3 = 0$  to achieve an  $\mathcal{O}(h^3)$ -bias is only possible if there are both positive and negative coefficients among the  $\beta_j$ . From definition 4.4.1 it becomes clear that this might lead to *negative* estimates of the specific surface  $\bar{s}$  which is an obvious disadvantage of this bias-optimized estimator.

**Corollary 4.4.6.** Let  $\Theta(x) = \mathbf{1}_{\{Z(x) \geq b\}}$  be a stationary and isotropic Thresholded Gaussian Field in  $\mathbb{R}^d$  with a covariance function  $\rho(h) \in C^4$  that is observed on a finite regular grid  $\Gamma$ . Let further  $\hat{s}_1, \hat{s}_2, \hat{s}_3$  be the Ohser-Mücklich surface estimators defined in equations 4.29, 2.24, 2.29 respectively. Then the  $\hat{s}_i, i \in \{1, 2, 3\}$  are asymptotically unbiased for  $h \downarrow 0$  and on the grid for  $h > 0$  their biases are:

- i).  $\mathbf{E}[\hat{s}_1] - \bar{s}(\Theta) = -\frac{1}{12} \cdot (-\rho''(0))^{\frac{3}{2}} \cdot \phi(0) \cdot \phi(b) \cdot \left( (b^2 - 1) + \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right) \cdot h^2 + \mathcal{O}(h^3),$
- ii).  $\mathbf{E}[\hat{s}_2] - \bar{s}(\Theta) = -\frac{\pi}{16} \cdot (-\rho''(0))^{\frac{3}{2}} \cdot \phi(0) \cdot \phi(b) \cdot \left( (b^2 - 1) + \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right) \cdot h^2 + \mathcal{O}(h^3),$
- iii).  $\mathbf{E}[\hat{s}_3] - \bar{s}(\Theta) = -0.334486 \cdot (-\rho''(0))^{\frac{3}{2}} \cdot \phi(0) \cdot \phi(b) \cdot \left( (b^2 - 1) + \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right) \cdot h^2 + \mathcal{O}(h^3).$

**Proof. (Corollary 4.4.6)** Combining Theorem 4.4.4 part i). with Example 4.4.3. For iii). the factor

$$-0.334486 \approx -\frac{4}{24} (6 \cdot c_0 + 24 \cdot c_3 + 24 \cdot c_9)$$

stems from the *Voronoi*-coefficients defined in the discussion following equation 2.29.  $\square$

The proof of Theorem 4.4.4 is based on Lemma 4.3.2.

**Proof. (Theorem 4.4.4)**

ad i).

We start with the definition of a general surface estimator given in definition 4.4.1. For its expectation we calculate using Lemma 4.3.2:

$$\begin{aligned} \mathbf{E}[\hat{s}] &= \mathbf{E} \left[ -\frac{1}{n \cdot h} \cdot \sum_{i=1}^n \sum_{j=1}^m \beta_j \cdot \mathbf{1}_{\{Z(x_i) \geq b, Z(x_i + h \cdot a_j) \leq b\}} \right] = \\ &= -\frac{1}{h} \sum_{j=1}^m \beta_j \mathbf{P}[Z(0) \geq b, Z(h \cdot a_j) \leq b] = -\frac{1}{h} \sum_{j=1}^m \left[ \beta_j \sqrt{2} \phi(0) \phi(b) \sqrt{1 - \rho(h \cdot |a_j|)} + \mathcal{O}(1 - \rho(h)) \right] \\ &= -\frac{1}{h} \sum_{j=1}^m \left[ \beta_j \sqrt{2} \phi(0) \phi(b) \sqrt{-\rho''(0)} \frac{|a_j|}{\sqrt{2}} \cdot h + \mathcal{O}(h^2) \right] = \\ &= -\phi(0) \phi(b) \sqrt{-\rho''(0)} \sum_{j=1}^m \beta_j |a_j| + \mathcal{O}(h), \end{aligned} \tag{4.30}$$

where we have expanded  $\sqrt{1 - \rho(h \cdot |a_j|)}$ .

Similarly we learn from Theorem 2.5.2 and equation 4.26:

$$\begin{aligned} \bar{s} &= -\frac{d \cdot b_d}{b_{d-1}} C'(0) = -\frac{d \cdot b_d}{b_{d-1}} \lim_{h \downarrow 0} \frac{C(h) - C(0)}{h} \\ &= \frac{d \cdot b_d}{b_{d-1}} \lim_{h \downarrow 0} \frac{1}{h} \mathbf{P}[Z(0) \geq b, Z(h \cdot a) \leq b] = \frac{d \cdot b_d}{b_{d-1}} \sqrt{2} \phi(0) \phi(b) \lim_{h \downarrow 0} \frac{1}{h} \sqrt{1 - \rho(h)} \\ &= \frac{d \cdot b_d}{b_{d-1}} \phi(0) \phi(b) \sqrt{-\rho''(0)}, \end{aligned} \tag{4.31}$$

where  $C(\cdot)$  is the two-point-covariance function (definition 2.5.1) and  $a$  is a unit-vector. Comparing this with equation 4.30 we see that  $\hat{s}$  is asymptotically unbiased for  $h \downarrow 0$  if and only if equation 4.28 holds which is true by definition of  $\hat{s}$ .

ad ii).

To calculate the bias explicitly we need to expand  $\sqrt{1 - \rho^2(h)}$  a bit further:

$$\sqrt{1 - \rho(h)} = \sqrt{\frac{-\rho''(0)}{2}} \cdot h - \frac{1}{24} \frac{1}{\sqrt{2}} \frac{\rho^{(iv)}(0)}{\sqrt{-\rho''(0)}} \cdot h^3 + \mathcal{O}(h^4).$$

If we plug this into the expression for  $\mathbf{P}[Z(0) \geq b, Z(h|a_j|) \leq b]$  calculated in Lemma 4.3.2 we find that up to order  $h^3$  we have:

$$\begin{aligned} \mathbb{P}[Z(0) \geq b, Z(h|a_j|) \leq b] &= h \cdot |a_j| \cdot \phi(0)\phi(b)\sqrt{-\rho''(0)} \\ &- h^3 \cdot |a_j|^3 \cdot \frac{1}{24} \left[ \phi(0)\phi(b) \frac{\rho^{(iv)}(0)}{\sqrt{-\rho''(0)}} + (-\rho''(0))^{\frac{3}{2}} \phi(b)\phi(0)(b^2 - 1) \right] + \mathcal{O}(h^4), \end{aligned}$$

where we have noted that there is an additional term of order  $h^3$  coming from the expansion of  $(1 - \rho)^{3/2}$ . Hence for the bias we arrive at:

$$\mathbb{E}[\hat{s}] - \bar{s} = \frac{1}{24} \sum_{j=1}^m \beta_j |a_j|^3 \left( \phi(0)\phi(b) \frac{\rho^{(iv)}(0)}{\sqrt{-\rho''(0)}} + (-\rho''(0))^{\frac{3}{2}} \phi(b)\phi(0)(b^2 - 1) \right) h^2 + \mathcal{O}(h^4),$$

which was to be shown.

ad iii.)

If  $|a_1|, |a_2|$  are given,  $\beta_1, \beta_2$  must solve the following equation system:

$$\begin{aligned} \beta_1 \cdot |a_1| + \beta_2 \cdot |a_2| &= c_d \\ \beta_1 \cdot |a_1|^3 + \beta_2 \cdot |a_2|^3 &= 0. \end{aligned}$$

It's trivial to verify that this equation system is fulfilled for the  $\beta_i, i \in \{1, 2\}$  given in statement iii).  $\square$

## 4.5 Asymptotic Normality of Surface Estimators

The main goal of this section is to prove asymptotic normality of the general surface-estimators of definition 4.4.1 for Thresholded Gaussian Random Fields when the field size  $n$  tends to infinity and the grid spacing  $h_n$  tends to zero accordingly. We start out with some general model-independent remarks on the variance of general-surface estimators.

### 4.5.1 Variance of General Surface Estimators Based on the Covariance Function

In this subsection we compute the asymptotic order of the variance of general surface estimators as the lattice-spacing  $h$  tends to zero. For this purpose let  $\Theta$  be an arbitrary stationary and isotropic random set in  $\mathbb{R}^d$  and consider a sequence of refining grids  $\Gamma_n = \prod_{s=1}^d [0, h_n, 2h_n, \dots, (n_s - 1)h_n]$  with lattice-spacings  $h_n$  and set  $n = \prod_{s=1}^d n_s$  the total number of lattice points. We assume that as  $n \rightarrow \infty$  (and all  $n_s \rightarrow \infty$ ) we have  $h_n \downarrow 0$  in such a way that  $nh_n^d \rightarrow \infty$ . Furthermore we denote by  $(x_{ni})_{0 \leq i \leq n}$  a numbering of the gridpoints of  $\Gamma_n$  and note that for every  $x = (x^1, \dots, x^d) \in \mathbb{R}^{d^+}$  there is a sequence  $x_{ni} \rightarrow x$  for  $n \rightarrow \infty$ . In this setting we recall definition 4.4.1 of general surface estimators:

$$\hat{s}_n = -\frac{1}{nh_n} \sum_{i=1}^n \sum_{j=1}^m \beta_j \cdot \mathbf{1}_{\{x_{ni} \in \Theta, x_{ni} + h_n \cdot a_j \in \Theta^c\}}, \quad \sum_{j=1}^m \beta_j \cdot |a_j| = -\frac{db_d}{b_{d-1}}. \quad (4.32)$$

In practice of course, one can compute  $\hat{s}_n$  only on a reduced grid and one must assure that  $x_{ni} + h_n \cdot a_j$  are gridpoints of  $\Gamma_n$ , too, and thus observable. We ignore this fact here to



keep the notation and the formulae simple. It's then straightforward to see that we can write the variance of  $\hat{s}$  as:

$$\begin{aligned} \text{Var}(\hat{s}) = & \\ & \frac{1}{n^2 h_n^2} \sum_{j=1}^m \sum_{l=1}^m \beta_j \cdot \beta_l \sum_{i=1}^n \sum_{k=1}^n \left[ \text{P}[x_{ni} \in \Theta, x_{ni} + h_n \cdot a_j \in \Theta^c, x_{nk} \in \Theta, x_{nk} + h_n \cdot a_l \in \Theta^c] \right. \\ & \left. - \text{P}[x_{ni} \in \Theta, x_{ni} + h_n \cdot a_j \in \Theta^c] \text{P}[x_{nk} \in \Theta, x_{nk} + h_n \cdot a_l \in \Theta^c] \right]. \end{aligned}$$

If we now use stationarity this can be simplified to:

$$\begin{aligned} \text{Var}(\hat{s}_n) \sim & \\ & \frac{1}{n^2 h_n^2} \sum_{j=1}^m \sum_{l=1}^m \beta_j \cdot \beta_l \sum_{x_{ni} \in \Gamma_n^*} \left[ \text{P}[0 \in \Theta, h_n \cdot a_j \in \Theta^c, x_{ni} \in \Theta, x_{ni} + h_n \cdot a_l \in \Theta^c] \right. \\ & \left. - \text{P}[0 \in \Theta, h_n \cdot a_j \in \Theta^c] \text{P}[0 \in \Theta, h_n \cdot a_l \in \Theta^c] \right] \cdot n \cdot \prod_{s=1}^d \left( 1 - \frac{|x_{ni}^s|}{n_s} \right), \quad (4.33) \end{aligned}$$

where  $\Gamma_n^* = \prod_{s=1}^d \{-(n_s - 1)h_n, \dots, (n_s - 1)h_n\}$  and  $x_{ni}^s$  is the  $s$ -th component of the gridpoint  $x_{ni}$ .

We now need to discuss the covariance terms in square brackets of equation 4.33 a little more in detail and study their order of magnitude for  $h_n \downarrow 0$ . This discussion applies for many stochastic models  $\Theta$  and examples will be given in the following. It turns out that the covariance terms exhibit a different behaviour for dimension  $d = 1$  than for higher dimensions. Let us first study the one-dimensional case. Then the covariance-terms in equation 4.33 simplify to:

$$\begin{aligned} & \text{P}[Z(0) \in \Theta, Z(jh_n) \in \Theta^c, Z(kh_n) \in \Theta, Z((k+l)h_n) \in \Theta^c] \\ & - \text{P}[Z(0) \in \Theta, Z(jh_n) \in \Theta^c] \text{P}[Z(0) \in \Theta, Z(lh_n) \in \Theta^c], \end{aligned}$$

where  $1 \leq j, l \leq m$  and  $-n \leq k \leq n$ , hence for the discussion without loss of generality we restrict ourselves to the case where both  $j$  and  $l$  are positive. Now for fixed  $j$  and  $l$  the covariance term depends on the ordering of the involved quadruple of variables. For  $k \geq j$  or  $k \leq -l$ , and thus for almost all values of  $k$ , we have that the ordered variables of the quadruple take the values 1 – 0 – 1 – 0 alternatingly. Therefore the covariance term is then of order  $\mathcal{O}(h_n^2)$ , but as we will see below, the sum of all such covariance terms can be written as a Riemannian integral and is thus of order  $\mathcal{O}(h_n)$ . On the other hand for  $-l < k < j$  the variables of the ordered quadruple take the values 1 – 1 – 0 – 0 and thus the covariance term is of order  $\mathcal{O}(h_n)$ . Because the sum over  $k$  is then only over a fixed finite number of terms (always  $j + l - 1$  terms not depending on  $n$ ), we are left with an  $\mathcal{O}(h_n)$ -term after summing over  $k$ . This reasoning shows that the finite part of the sum over  $k$  where the intervals  $[0, j]$  and  $[k, (k+l)]$  intersect is not negligible. Thus for the one-dimensional case we can summarize:

**Assumption 4.5.1. (Covariance-terms in 1-dimension)**

$$\begin{aligned}
& \mathbb{P}[Z(0) \in \Theta, Z(jh_n) \in \Theta^c, Z(kh_n) \in \Theta, Z((k+l)h_n) \in \Theta^c] \\
& - \mathbb{P}[Z(0) \in \Theta, Z(jh_n) \in \Theta^c] \mathbb{P}[Z(0) \in \Theta, Z(lh_n) \in \Theta^c] \\
& \sim \begin{cases} g_{jl}(x_{ni}) \cdot h_n^2 & \text{for } k \geq j \text{ or } k \leq -l, \\ c_{jl}(k) \cdot h_n & \text{for } -l < k < j. \end{cases} \quad (4.34)
\end{aligned}$$

where  $c_{jl}(k)$  are constants and  $g_{jl}(x_{ni})$  are model-dependent functions.

Lemmata 4.3.2 and 4.3.7 show that this is true for the Thresholded Gaussian Model. A further example where the assumption 4.5.1 is obviously met is the Markovian Model treated in section 7.1 of the appendix. The important observation is now that the sum of the  $g_{jl}(\cdot)h_n^2$ -terms is of Riemannian nature. Thus from equation 4.33 we see for the variance variance of  $\hat{s}_n$  for  $n \rightarrow \infty$  in the one-dimensional case:

$$\begin{aligned}
\text{Var}(\hat{s}_n) & \sim \frac{1}{nh_n} \sum_{j=1}^m \sum_{l=1}^m \beta_j \cdot \beta_l \left[ \sum_{k \geq j, k \leq -l} g_{jl}(kh_n) h_n + \sum_{-l < k < j} c_{jl}(k) \right] \\
& \sim \frac{1}{nh_n} \sum_{j=1}^m \sum_{l=1}^m \beta_j \cdot \beta_l \left( \int_{\mathbb{R}^d} g_{jl}(x) dx + \sum_{-l < k < j} c_{jl}(k) \right). \quad (4.35)
\end{aligned}$$

Hence the asymptotic order is  $\mathcal{O}\left(\frac{1}{nh_n}\right)$ .

For the multi-dimensional case  $d > 1$  and without loss of generality for  $k_{ni} \in \mathbb{Z}^d \cap [-n, n]^d$  the covariance-terms are of order

$$\begin{aligned}
& \mathbb{P}[0 \in \Theta, h_n \cdot a_j \in \Theta^c, h_n k_{ni} \in \Theta, h_n(k_{ni} + a_l) \in \Theta^c] \\
& - \mathbb{P}[0 \in \Theta, h_n \cdot a_j \in \Theta^c] \mathbb{P}[0 \in \Theta, h_n \cdot a_l \in \Theta^c] \sim r_{jl}(k_{ni}) h_n
\end{aligned}$$

or even smaller depending on the relative position of  $a_j, a_l$  and  $k_{ni}h_n$ . The  $r_{jl}$  stand for model-dependent constants. If we now assume that the  $r_{jl}(\cdot)$  are summable, ie.:

$$\sum_{k \in \mathbb{Z}^d} |r_{jl}(k)| < \infty,$$

there exists a number  $k_0$  such that for  $\|k_{ni}\| > k_0$  the covariance terms are of order  $\mathcal{O}(h_n^2)$  independent of the relative position of  $a_j, a_l$  and  $k_{ni}h_n$ . Hence because of  $k_{ni}h_n \rightarrow x$  for  $\|k_{ni}\| > k_0$  the covariance terms can be written as:

$$\begin{aligned}
& \mathbb{P}[0 \in \Theta, h_n \cdot a_j \in \Theta^c, h_n k_{ni} \in \Theta, h_n(k_{ni} + a_l) \in \Theta^c] \\
& - \mathbb{P}[0 \in \Theta, h_n \cdot a_j \in \Theta^c] \mathbb{P}[0 \in \Theta, h_n \cdot a_l \in \Theta^c] \sim g_{jl}(k_{ni}h_n) h_n^2,
\end{aligned}$$

where  $g_{jl}(\cdot)$  is a Riemann-integrable function. Thus we conclude from this reasoning and equation 4.33 for the variance of  $\hat{s}_n$ :

$$\text{Var}(\hat{s}_n) \sim \sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l \left[ \frac{1}{nh_n^d} \int_{\mathbb{R}^d} g_{jl}(x) dx + \frac{1}{nh_n} \sum_{k \in \mathbb{Z}^d, \|k\| \leq k_0} r_{jl}(k) \right] \sim \mathcal{O}\left(\frac{1}{nh_n^d}\right),$$

Therefore opposed to the 1D-case, in the multi-dimensional case the Riemannian sum dominates the finite-sum.

### 4.5.2 Asymptotic Normality for the Surface Estimator in the Gaussian Thresholded Field Model

In this section we investigate the properties of the one-dimensional Ohser-Mücklich surface estimator  $\hat{s}$  defined in equation 4.29 for the Thresholded Gaussian Random Field Model. The main goal is to compute its variance and to assert its asymptotic normality as the grid on which the estimator is computed becomes finer, ie. the number of lattice points  $n \rightarrow \infty$  and the lattice spacing  $h_n \xrightarrow{n \rightarrow \infty} 0$  in such a way that  $n \cdot h_n \rightarrow \infty$ .

The same strategies which will be used to calculate the variance of  $\hat{s}$  and to prove asymptotic normality in the Thresholded Gaussian Model in the following are applied - as an introductory example - to the *Continuous On-Off-Markov-System* in section 7.1 of the appendix. We have included this example because before reading on, the reader might want to familiarize himself with these techniques and see them working in a more easily and explicitly tractable setting first before tackling the Thresholded Gaussian Model.

Throughout this section we will assume that the random set under consideration is  $\Theta(x) = \mathbf{1}_{X(x) \geq b}$ , where  $Z(x), x \in \mathbb{R}$  is a strictly stationary Gaussian Random Field in one dimension. We start out with the formula for the estimator  $\hat{s}$  in this setting. As mentioned in section 4.4 the surface of a one-dimensional object is equal to the total number of its endpoints. Hence the estimator  $\hat{s}$  simply counts the number of phase transitions from black to white and vice versa:

$$\hat{s}_{n,h_n} = \frac{1}{n \cdot h_n} \sum_{i=1}^n \mathbf{1}_{\{Z(i \cdot h_n) \geq b\}} \mathbf{1}_{\{Z((i+1) \cdot h_n) \leq b\}} + \mathbf{1}_{\{Z(i \cdot h_n) \leq b\}} \mathbf{1}_{\{Z((i+1) \cdot h_n) \geq b\}}. \quad (4.36)$$

Of course a transition from black to white must be followed by a transition from white to black until the next black-white transition follows and thus the two summands that contribute to the estimator are identical up to edge effects. The analogous remark also applies for the same estimator in the continuous Markov model considered in subsection 7.1.1 of the appendix.

The next theorem collects the crucial properties of  $\hat{s}$  in the 1D-Gaussian Thresholded Field Model:

**Theorem 4.5.2. (Surface-estimator properties in the 1D-Gaussian-Model)** *Consider the one-dimensional Thresholded Gaussian Random Field Model  $\Theta(x) = \mathbf{1}_{X(x) \geq b}$ , where  $b \in \mathbb{R}$  and  $Z(x), x \in \mathbb{R}$  is a strictly stationary Gaussian Random Field with covariance function  $\rho(h)$  observed on a sequence of grids  $\Gamma_n = h_n \cdot \mathbb{Z}$  with  $n \rightarrow \infty, h_n \downarrow 0$  such that  $n \cdot h_n \rightarrow \infty$ . Let further  $\hat{s}$  be the Ohser-Mücklich surface estimator defined as in equation 4.36. Then the following statements are true:*

- i). *The specific surface area  $\bar{s}(\Theta)$  in the 1D-Thresholded Gaussian Field Model  $\Theta$  is:*

$$\bar{s}(\Theta) = \sqrt{\frac{2}{\pi}} \phi(b) \sqrt{-\rho''(0)},$$

where  $\phi(\cdot)$  is the density function of the standard normal.

- ii). *For  $h_n \downarrow 0$  the estimator  $\hat{s}$  is asymptotically unbiased for  $\bar{s}$  and the bias has the form:*

$$\mathbf{E} [\hat{s}_{n,h_n}] - \bar{s}(\Theta) = -\frac{1}{12} \cdot \frac{(-\rho''(0))^{\frac{3}{2}}}{\sqrt{2\pi}} \cdot \phi(b) \cdot \left( (b^2 - 1) + \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right) \cdot h_n^2 + \mathcal{O}(h_n^3).$$

iii). If  $\rho(x), \rho'(x), \rho''(x)$  are integrable and

$$\int_0^\infty |x| \cdot |\rho(x)| dx < \infty, \quad \int_0^\infty |x| \cdot |\rho'(x)| dx < \infty, \quad \int_0^\infty |x| \cdot |\rho''(x)| dx < \infty,$$

there exists an integrable function  $g(x) : \mathbb{R} \rightarrow \mathbb{R}$  such that  $x \cdot g(x)$  is still integrable and

$$\text{Var}(\hat{s}) \sim \frac{1}{n \cdot h_n} \left( k_1 + \int_{-\infty}^\infty g(x) dx \right),$$

where  $k_1$  is a constant. Hence the variance  $\text{Var}(\hat{s})$  is of order  $\mathcal{O}(\frac{1}{nh_n})$ .

**Proof:(Theorem 4.5.2)**

**ad i) and ii):** Since the specific surface estimator as given in equation 4.36 is a special case of the general specific surface estimator form definition 4.4.1 for  $m = 2, \beta_1 = \beta_2 = -1$  and  $a_1 = 1, a_2 = -1$ , these claims follow directly from equation 4.31 and Theorem 4.4.4, part ii).

**ad iii):**

Since the estimator  $\hat{s}$  is of the form given by equation 4.36 for  $m = 2, \beta_1 = -1, \beta_2 = -1, a_1 = 1, a_2 = -1$ , we get for the variance:

$$\begin{aligned} \text{Var}(\hat{s}_n) &= \frac{1}{nh_n^2} \sum_{k=-n}^n \left(1 - \frac{|k|}{n}\right) \left\{ \text{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k+1)h_n) \leq b] \right. \\ &+ 2 \cdot \text{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \leq b, Z((k+1)h_n) \geq b] \\ &+ \text{P}[Z(0) \leq b, Z(h_n) \geq b, Z(kh_n) \leq b, Z((k+1)h_n) \geq b] \\ &- \text{P}[Z(0) \geq b, Z(h_n) \leq b]^2 - 2 \text{P}[Z(0) \geq b, Z(h_n) \leq b] \text{P}[Z(0) \leq b, Z(h_n) \geq b] \\ &\left. - \text{P}[Z(0) \leq b, Z(h_n) \geq b]^2 \right\}. \end{aligned} \quad (4.37)$$

We will see that for  $n, k \rightarrow \infty, nh_n \rightarrow \infty, h_n \rightarrow 0$  and  $k \cdot h_n \rightarrow x$  most summands of the expression in curly braces of equation 4.37 are of order  $\mathcal{O}(h_n^2)$ . However by Lemmata 4.3.2 and 4.3.3 the terms for which the intervals  $[0, h_n], [kh_n, (k+1)h_n]$  intersect are of order  $\mathcal{O}(h_n)$ , namely for  $k = 0$  and  $k = 1$  respectively.

$$\begin{aligned} &\text{P}[Z(0) \geq b, Z(h_n) \leq b, Z(0h_n) \geq b, Z((0+1)h_n) \leq b] \\ &+ \text{P}[Z(0) \leq b, Z(h_n) \geq b, Z(0h_n) \leq b, Z((0+1)h_n) \geq b] \\ &+ 2 \left(1 - \frac{1}{n}\right) \text{P}[Z(0) \geq b, Z(h_n) \leq b, Z(1h_n) \leq b, Z((1+1)h_n) \geq b] \\ &= \text{P}[Z(0) \geq b, Z(h_n) \leq b] + \text{P}[Z(0) \leq b, Z(h_n) \geq b] \\ &\quad + \left(2 - \frac{1}{n}\right) \text{P}[Z(0) \geq b, Z(h_n) \leq b, Z(2h_n) \geq b] \sim \left(k_1 + \frac{k_2}{n}\right) \cdot h_n, \end{aligned}$$

for some constants  $k_1, k_2$ . The subtracted terms in equation 4.37 are of order  $\mathcal{O}(h_n^2)$ .

We will show in the following that all other summands of the expression in curly braces in equation 4.37 where the intervals  $[0, h_n]$  and  $[kh_n, (k + 1)h_n]$  do not intersect can be expanded as  $g(kh_n)h_n^2 + \mathcal{O}(h_n^3)$ , for a function  $g$  which fulfills the integrability conditions stated in the theorem. Thus assumption 4.34 is fulfilled. Once this is shown the claim follows from expressing the variance through Riemannian sums and treating the  $\mathcal{O}(h_n)$ -terms separately:

$$\begin{aligned} \text{Var}(\hat{s}_n) &= \\ & \frac{1}{nh_n^2} \left[ (k_1 + \frac{k_2}{n}) \cdot h_n + \sum_{k=-n}^n \left( 1 - \frac{|k|}{n} \right) (g(kh_n)h_n^2 + \mathcal{O}(h_n^3)) \right] \\ & \sim \frac{1}{nh_n} \left( k_1 + \sum_{k=-n}^n g(kh_n)h_n \right) - \frac{1}{n^2h_n^2} \sum_{k=-n}^n |k|h_n g(kh_n)h_n + \frac{k_2}{n^2h_n} \\ & \stackrel{kh_n \rightarrow x}{\sim} \frac{1}{nh_n} \left( k_1 + \int_{-\infty}^{\infty} g(x) dx \right) - \frac{1}{n^2h_n^2} \int_{-\infty}^{\infty} |x|g(x) dx + \frac{k_2}{n^2h_n} \\ & \sim \frac{1}{nh_n} \left( k_1 + \int_{-\infty}^{\infty} g(x) dx \right), \end{aligned} \tag{4.38}$$

where all sums only run over such  $k$  for which the intervals  $[0, h_n]$  and  $[kh_n, (k + 1)h_n]$  do not intersect.

It then remains to show the integrability of the function  $g(\cdot)$ . In a first approach, we only consider the single probability-term  $\mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k + 1)h_n) \leq b]$ . to which we want to apply Lemma 4.3.7 with  $Z_1 := Z(0), Z_2 := Z(h_n), Z_3 := Z(kh_n), Z_4 := Z((k + 1)h_n)$ . To do this we have to determine the constants  $c_2, \dots, c_6$  of the Lemma in the one-dimensional Thresholded Gaussian Field setting. Because of  $kh_n \rightarrow x$  we have:

$$\rho_1 = \rho_2 = \rho(h), \quad \rho_3 = \rho_6 = \rho(x), \quad \rho_4 = \rho(x + h), \quad \rho_5 = \rho(x - h),$$

and thus with  $\tau = \sqrt{1 - \rho(h)}$ :

$$\rho_2 = \rho(h) = 1 - \tau^2 \Rightarrow 1 - \rho_2 = \tau^2 \Rightarrow c_2 = 1, \tag{4.39}$$

$$c_4 = \lim_{\tau \downarrow 0} \frac{\rho_4 - \rho_3}{\tau} = \lim_{\tau \downarrow 0} \frac{\rho(x + h) - \rho(x)}{\sqrt{1 - \rho(h)}} = \frac{\sqrt{2}\rho'(x)}{\sqrt{-\rho''(0)}}. \tag{4.40}$$

$$c_5 = \lim_{\tau \downarrow 0} \frac{\rho_5 - \rho_3}{\tau} = \lim_{\tau \downarrow 0} \frac{\rho(x - h) - \rho(x)}{\sqrt{1 - \rho(h)}} = -\frac{\sqrt{2}\rho'(x)}{\sqrt{-\rho''(0)}}. \tag{4.41}$$

$$c_6 = \lim_{\tau \downarrow 0} \frac{\rho_6 - \rho_5 - \rho_4 + \rho_3}{\tau^2} = \lim_{\tau \downarrow 0} \frac{2\rho(x) - \rho(x + h) - \rho(x - h)}{1 - \rho(h)} = \frac{2\rho''(x)}{\rho''(0)}. \tag{4.42}$$

Hence a direct application of Lemma 4.3.7 leads us to:

$$\begin{aligned}
& \mathbb{P}[Z_1 \geq b, Z_2 \leq b, Z_3 \geq b, Z_4 \leq b] \sim \tag{4.43} \\
& \sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{2}}\right) \phi\left(\frac{2b(1-\rho_3) - c_5 w_2}{\sqrt{\Sigma_{11}}}\right) \frac{1}{\sqrt{\Sigma_{11}}} \left[ m\Phi\left(-\frac{m}{s}\right) - s\phi\left(\frac{m}{s}\right) \right] dw_2 = \\
& \sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{2}}\right) \phi\left(\frac{2b(1-\rho(x)) + \frac{\sqrt{2}\rho'(x)}{\sqrt{-\rho''(0)}}w_2}{\sqrt{\Sigma_{11}}}\right) \frac{1}{\sqrt{\Sigma_{11}}} \left[ m\Phi\left(-\frac{m}{s}\right) - s\phi\left(\frac{m}{s}\right) \right] dw_2,
\end{aligned}$$

where:

$$\begin{aligned}
m &= 2b \left[ \frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1-\rho_3) \right] + w_2 \left[ \frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}}c_5 \right] \\
&= 2b \left[ \frac{\rho'(x)}{\sqrt{-2\rho''(0)}} + \frac{\Sigma_{12}}{\Sigma_{11}}(1-\rho(x)) \right] + w_2 \left[ \frac{\rho''(x)}{\rho''(0)} + \frac{\Sigma_{12}}{\Sigma_{11}} \frac{\sqrt{2}\rho'(x)}{\sqrt{-\rho''(0)}} \right], \\
s &= \sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}, \\
\Sigma &= \begin{pmatrix} 4(1-\rho(x)^2) + 4\frac{\rho'(x)^2}{\rho''(0)} & -2\sqrt{2}\rho'(x) \left[ \frac{\rho(x)}{\sqrt{-\rho''(0)}} + \frac{\rho''(x)}{(-\rho''(0))^{\frac{3}{2}}} \right] \\ -2\sqrt{2}\rho'(x) \left[ \frac{\rho(x)}{\sqrt{-\rho''(0)}} + \frac{\rho''(x)}{(-\rho''(0))^{\frac{3}{2}}} \right] & 2 + \frac{2\rho'(x)^2}{\rho''(0)} - \frac{2\rho''(x)^2}{\rho''(0)^2} \end{pmatrix}.
\end{aligned}$$

From equation 4.37 we see that the function  $g(x) = g_1(x) + g_2(x) + g_3(x)$  is defined through:

$$\begin{aligned}
g_1(x) &:= \tag{4.44} \\
& \lim_{n \rightarrow \infty} \frac{1}{h_n^2} \left[ \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(x) \geq b, Z(x+h_n) \leq b] - \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b]^2 \right]
\end{aligned}$$

and two similar terms  $g_2(x)$  and  $g_3(x)$  which can also be read off from equation 4.37. All the calculations that follow are identical for  $g_1(\cdot), g_2(\cdot), g_3(\cdot)$ , so without loss of generality we may only consider  $g_1(\cdot)$  in the following and drop the index to simply denote it  $g(\cdot)$  in the following. Please note further that only the term coming from the quadruple probability depends on  $x$  and that obviously

$$\mathbb{P}[Z(0) \geq b, Z(h_n) \leq b]^2 = \lim_{x \rightarrow \infty} \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(x) \geq b, Z(x+h) \leq b].$$

To show integrability of  $g$  we certainly must show that  $g(x)$  decays fast enough for  $x \rightarrow \infty$ . The integrability for  $x \rightarrow -\infty$  then follows from symmetry considerations. But another problem arises for  $g(\cdot)$  near zero. One can easily see that  $\Sigma_{11}, \Sigma_{22}, s \rightarrow 0$  for  $x \rightarrow 0$  while the correlation  $\Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}} \rightarrow 1$  for  $x \rightarrow 0$ . Hence in view of equation 4.43 we must also ensure that  $g(x)$  does not explode for  $x$  near zero. This is the problem we will tackle first. Because only the quadruple probability in the above expression for  $g$  depends on  $x$  it is enough to show that the integral in equation 4.43 stays bounded for  $x \rightarrow 0$ . This we will show now.

Estimating from above  $\phi\left(\frac{w_2}{\sqrt{2}}\right) \leq \frac{1}{\sqrt{2\pi}}$  and  $\Phi\left(-\frac{m}{s}\right) \leq 1$  and dropping any constants where possible the integral in equation 4.43 can be decomposed into the following three terms:

$$\begin{aligned}
 \text{i)} \quad & \frac{1}{\sqrt{\Sigma_{11}}} \left[ \frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1 - \rho_3) \right] \int_{-\infty}^0 w_2 \phi \left( \frac{2b(1 - \rho_3) - c_5 w_2}{\sqrt{\Sigma_{11}}} \right) dw_2 \\
 \text{ii)} \quad & \frac{1}{\sqrt{\Sigma_{11}}} \left[ \frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}} c_5 \right] \int_{-\infty}^0 w_2^2 \phi \left( \frac{2b(1 - \rho_3) - c_5 w_2}{\sqrt{\Sigma_{11}}} \right) dw_2 \\
 \text{iii)} \quad & \frac{1}{\sqrt{\Sigma_{11}}} \sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}} \int_{-\infty}^0 w_2 \phi \left( \frac{2b(1 - \rho_3) - c_5 w_2}{\sqrt{\Sigma_{11}}} \right) \\
 & \times \phi \left( \frac{2b \left[ \frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1 - \rho_3) \right] + w_2 \left[ \frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}} c_5 \right]}{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}} \right) dw_2,
 \end{aligned}$$

We now show that in fact all these three integrals vanish for  $x \rightarrow 0$ :

ad i):

$$\begin{aligned}
 & \left| \frac{\frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1 - \rho_3)}{c_5} \int_{-\infty}^0 w_2 \frac{c_5}{\sqrt{\Sigma_{11}}} \phi \left( \frac{w_2 + \frac{2b(1 - \rho_3)}{c_5}}{\frac{\sqrt{\Sigma_{11}}}{c_5}} \right) dw_2 \right| \\
 & \leq \underbrace{\left| \frac{\frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1 - \rho_3)}{c_5} \right|}_{\sim x^2 \text{ for } x \rightarrow 0} \underbrace{\left( \left| \frac{2b(1 - \rho_3)}{c_5} \right| + \left| \frac{2}{\sqrt{2\pi}} \frac{\sqrt{\Sigma_{11}}}{c_5} \right| \right)}_{\sim x \text{ for } x \rightarrow 0} \sim x^3 \xrightarrow{x \rightarrow 0} 0.
 \end{aligned}$$

where we have standardized such that the Gaussian integral can be estimated from above by the first moment  $\mathbf{E}[|X|]$  of a Gaussian random variable  $X \sim \mathcal{N}(\mu, \sigma^2)$  which can be estimated from above by  $\mathbf{E}[|X|] \leq |\mu| + \frac{2}{\sqrt{2\pi}} \cdot \sigma$ . With some tedious but straightforward calculations we have expanded the two remaining factors into Taylor-Series for  $x \rightarrow 0$  to find out their order of magnitude.

ad ii): This can be treated similarly:

$$\begin{aligned}
 & \left| \frac{\frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}} c_5}{c_5} \int_{-\infty}^0 w_2^2 \frac{c_5}{\sqrt{\Sigma_{11}}} \phi \left( \frac{w_2 - \frac{2b(1 - \rho_3)}{c_5}}{\frac{\sqrt{\Sigma_{11}}}{c_5}} \right) dw_2 \right| \\
 & \leq \underbrace{\left| \frac{\frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}} c_5}{c_5} \right|}_{\sim \frac{1}{x} \text{ for } x \rightarrow 0} \underbrace{\left[ \frac{\Sigma_{11}}{c_5^2} + \frac{4b^2(1 - \rho_3)^2}{c_5^2} \right]}_{\sim x^2 \text{ for } x \rightarrow 0} \sim x \xrightarrow{x \rightarrow 0} 0,
 \end{aligned}$$

where we have standardized such that the Gaussian integral can be estimated from above by the second moment  $\mathbf{E}[X^2] = \text{Var}(X) + \mathbf{E}[X]^2$  of some Gaussian random variable  $X$  this time and again expanded the remaining factors into Taylor-series for  $x \rightarrow 0$ .

ad iii):

Here we first combine the two standard Gaussian factors into a single one. For arbitrary real numbers  $a, b, c, d$  completing the square reveals:

$$\begin{aligned} & \phi\left(\frac{w_2 - a}{b}\right) \cdot \phi\left(\frac{w_2 - c}{d}\right) = \tag{4.45} \\ & \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{1}{b^2} + \frac{1}{d^2}}} \text{Exp}\left(-\frac{1}{2}\left[\frac{a^2}{b^2} + \frac{c^2}{d^2} - \frac{\left(\frac{a}{b^2} + \frac{c}{d^2}\right)^2}{\left(\frac{1}{b^2} + \frac{1}{d^2}\right)}\right]\right) \cdot \phi\left(w_2, \frac{\frac{a}{b^2} + \frac{c}{d^2}}{\frac{1}{b^2} + \frac{1}{d^2}}, \frac{1}{\frac{1}{b^2} + \frac{1}{d^2}}\right). \end{aligned}$$

And hence by applying equation 4.45 to iii) and comparing to the first moment as in i) we find:

$$\begin{aligned} & \left| \frac{1}{\sqrt{\Sigma_{11}}} \sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}} \int_{-\infty}^0 w_2 \phi\left(\frac{2b(1-\rho_3) - c_5 w_2}{\sqrt{\Sigma_{11}}}\right) \right. \\ & \times \phi\left(\frac{2b\left[\frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1-\rho_3)\right] + w_2\left[\frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}}c_5\right]}{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}}\right) dw_2 \left. \right| \\ & \leq \frac{1}{\sqrt{2\pi}} \underbrace{\frac{1}{\sqrt{\Sigma_{11}}}}_{\sim \frac{1}{x^2} \text{ for } x \rightarrow 0} \underbrace{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}}_{\sim x^2 \text{ for } x \rightarrow 0} \underbrace{\frac{1}{\sqrt{\frac{1}{b^2} + \frac{1}{d^2}}}}_{\sim x^2 \text{ for } x \rightarrow 0} \underbrace{\text{Exp}\left(-\frac{1}{2}\left(\frac{\overbrace{a^2}^{\sim x^2}}{b^2} + \frac{\overbrace{c^2}^{\sim \text{cst.}}}{d^2} - \frac{\overbrace{\left(\frac{a}{b^2} + \frac{c}{d^2}\right)^2}^{\sim x^2}}{\left(\frac{1}{b^2} + \frac{1}{d^2}\right)}\right)\right)}_{\sim \text{cst. for } x \rightarrow 0} \\ & \times \left( \underbrace{\left|\frac{\frac{a}{b^2} + \frac{c}{d^2}}{\frac{1}{b^2} + \frac{1}{d^2}}\right|}_{\sim x^3 \text{ for } x \rightarrow 0} + \underbrace{\frac{2}{\sqrt{2\pi}} \left|\frac{1}{\sqrt{\frac{1}{b^2} + \frac{1}{d^2}}}\right|}_{\sim x^2 \text{ for } x \rightarrow 0} \right) \xrightarrow{x \rightarrow 0} 0, \end{aligned}$$

where:

$$\begin{aligned} a &= -\frac{2b\left[\frac{c_4}{2} + \frac{\Sigma_{12}}{\Sigma_{11}}(1-\rho_3)\right]}{\frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}}c_5} & b &= \frac{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}}{\frac{c_6}{2} - \frac{\Sigma_{12}}{\Sigma_{11}}c_5} \\ c &= \frac{2b(1-\rho_3)}{c_5} & d &= \frac{\sqrt{\Sigma_{11}}}{c_5} \end{aligned}$$

and hence  $g(0)$  exists and is even zero. It remains to show that the decay of  $g(x)$  for  $x \rightarrow \infty$  is fast enough to ensure integrability of  $g$ . We start again from equation 4.44 and only consider the  $g_1(\cdot)$ -part of  $g(\cdot)$  given in equation 4.44. It is enough to show that  $|\int_{x_0}^{\infty} g(x) dx|$  is finite for a well chosen  $x_0$ . For a fixed  $1/2 > \epsilon > 0$ , we can easily choose an  $x_0$  such that the following conditions are simultaneously fulfilled  $\forall x \geq x_0$ :

$$|\rho(x)| \leq 1, \quad |\rho'(x)| \leq 1, \quad |\rho''(x)| \leq 1,$$

$$\tilde{C}_1 := \frac{1}{2} - \epsilon \leq \frac{1}{\sqrt{\Sigma_{11}(x)}} \leq \frac{1}{2} + \epsilon := C_1,$$

$$\tilde{C}_2 := \sqrt{2} - \epsilon \leq s(x) = \sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}} \leq \sqrt{2} + \epsilon := C_2,$$

$$\left| \frac{2b(1-\rho(x))}{\sqrt{\Sigma_{11}(x)}} \right| \leq |b| + \epsilon := C_3, \quad \left| 1 - \frac{\sqrt{\Sigma_{11}(x)}}{2} \right| \leq |\rho(x)|,$$



The last of these inequalities can be made true for  $x$  large enough because de L'Hôpital's rule gives:

$$\begin{aligned} & \left| \lim_{x \rightarrow \infty} \frac{\rho(x)}{1 - \frac{\sqrt{\Sigma_{11}(x)}}{2}} \right| = \left| \lim_{x \rightarrow \infty} \frac{\rho'(x)}{-\frac{1}{\sqrt{\Sigma_{11}(x)}}(-2\rho(x)\rho'(x) + \frac{2}{\rho''(0)}\rho'(x)\rho''(x))} \right| \\ & = \left| \lim_{x \rightarrow \infty} \frac{1}{-\frac{1}{\sqrt{\Sigma_{11}(x)}}(-2\rho(x) + \frac{2}{\rho''(0)}\rho''(x))} \right| = \infty. \end{aligned}$$

Now we define the following quantities for  $x \geq x_0$ :

$$\begin{aligned} \Delta_1(x) & := \frac{1}{\sqrt{\Sigma_{11}(x)}} \left[ 2b(1 - \rho(x)) + \frac{\sqrt{2}\rho'(x)}{\sqrt{-\rho''(0)}}w_2 \right] - b \\ |\Delta_1(x)| & \leq \left| \frac{2b(1 - \rho(x))}{\sqrt{\Sigma_{11}(x)}} - b \right| + \left| \frac{\sqrt{2}\rho'(x)}{\sqrt{\Sigma_{11}(x)}\sqrt{-\rho''(0)}} \right| |w_2| \leq 4|b|C_1|\rho(x)| + \sqrt{2}C_1 \frac{|\rho'(x)|}{\sqrt{-\rho''(0)}} |w_2|. \\ \Delta_2(x) & := \frac{m(x)}{s(x)} = \frac{1}{s(x)} \left( 2b \left[ \frac{\rho'(x)}{\sqrt{-2\rho''(0)}} + \frac{\Sigma_{12}}{\Sigma_{11}}(1 - \rho(x)) \right] + w_2 \left[ \frac{\rho''(x)}{\rho''(0)} + \frac{\Sigma_{12}}{\Sigma_{11}} \frac{\sqrt{2}\rho'(x)}{\sqrt{-\rho''(0)}} \right] \right) \\ |\Delta_2(x)| & \leq \frac{1}{\tilde{C}_2} \left\{ 2b \left[ \frac{|\rho'(x)|}{\sqrt{-2\rho''(0)}} + 4C_1^2\sqrt{2}|\rho'(x)| \left( \frac{1}{\sqrt{-\rho''(0)}} + \frac{1}{(-\rho''(0))^{\frac{3}{2}}} \right) \right] \right. \\ & \quad \left. + |w_2| \left( \frac{|\rho''(x)|}{|\rho''(0)|} + \frac{4C_1^2}{\sqrt{-\rho''(0)}} |\rho'(x)| \left( \frac{1}{\sqrt{-\rho''(0)}} + \frac{1}{(-\rho''(0))^{\frac{3}{2}}} \right) \right) \right\} \\ |m(x)| & \leq \left\{ 2b \left[ \frac{|\rho'(x)|}{\sqrt{-2\rho''(0)}} + 4C_1^2\sqrt{2}|\rho'(x)| \left( \frac{1}{\sqrt{-\rho''(0)}} + \frac{1}{(-\rho''(0))^{\frac{3}{2}}} \right) \right] \right. \\ & \quad \left. + |w_2| \left( \frac{|\rho''(x)|}{|\rho''(0)|} + \frac{4C_1^2}{\sqrt{-\rho''(0)}} |\rho'(x)| \left( \frac{1}{\sqrt{-\rho''(0)}} + \frac{1}{(-\rho''(0))^{\frac{3}{2}}} \right) \right) \right\}. \end{aligned}$$

Then we also have the following inequalities:

$$\begin{aligned} \xi(x) & \in \left[ \min \left\{ b, b + \Delta_1(x) \right\}, \max \left\{ b, b + \Delta_1(x) \right\} \right] \\ |\xi(x)| & \leq |b| + \left| \frac{2b(1 - \rho(x))}{\sqrt{\Sigma_{11}(x)}} \right| + \left| \frac{\sqrt{2}\rho'(x)}{\sqrt{\Sigma_{11}(x)}\sqrt{-\rho''(0)}} \right| |w_2| \leq |b| + C_3 + \sqrt{2}C_1 \frac{|\rho'(x)|}{\sqrt{-\rho''(0)}} |w_2|. \\ \nu(x) & \in \left[ \min \left\{ 0, \Delta_2(x) \right\}, \max \left\{ 0, \Delta_2(x) \right\} \right] \\ |\nu(x)| & \leq |\Delta_2(x)| \leq \frac{1}{\tilde{C}_2} \left\{ 2b \left[ \frac{|\rho'(x)|}{\sqrt{-2\rho''(0)}} + 4C_1^2\sqrt{2}|\rho'(x)| \left( \frac{1}{\sqrt{-\rho''(0)}} + \frac{1}{(-\rho''(0))^{\frac{3}{2}}} \right) \right] \right. \\ & \quad \left. + |w_2| \left( \frac{|\rho''(x)|}{|\rho''(0)|} + \frac{4C_1^2}{\sqrt{-\rho''(0)}} |\rho'(x)| \left( \frac{1}{\sqrt{-\rho''(0)}} + \frac{1}{(-\rho''(0))^{\frac{3}{2}}} \right) \right) \right\} \end{aligned}$$

Further we can deduce from the mean value-theorem the existence of  $\theta \in (\frac{1}{\Sigma_{11}(x)}, \frac{1}{4})$  and  $\eta \in (2, s^2)$  with:

$$\begin{aligned} \left| \frac{1}{\sqrt{\Sigma_{11}(x)}} - \frac{1}{2} \right| &= \left| \sqrt{\frac{1}{\Sigma_{11}(x)}} - \sqrt{\frac{1}{4}} \right| = \frac{1}{2\sqrt{\theta}} \left| \frac{1}{\Sigma_{11}(x)} - \frac{1}{4} \right| \leq \frac{C_1^2}{2\tilde{C}_1} \left| 1 - (1 - \rho(x)^2) - \frac{\rho'(x)^2}{\rho''(0)} \right| \\ &\leq \frac{C_1^2}{2\tilde{C}_1} \left( |\rho(x)| + \frac{|\rho'(x)|}{|\rho''(0)|} \right), \\ |s - \sqrt{2}| &\leq |\sqrt{s^2} - \sqrt{2}| \leq \frac{1}{2\sqrt{\eta}} |s^2 - 2| \\ &\leq \frac{1}{2\tilde{C}_2} \left[ \frac{2|\rho'(x)|}{|\rho''(0)|} + 2\frac{|\rho''(x)|}{\rho''(0)^2} + 8C_1^2|\rho'(x)| \left( \frac{1}{-\rho''(0)} + \frac{2}{\rho''(0)^2} + \frac{1}{(-\rho''(0))^3} \right) \right]. \end{aligned}$$

As a last preparation step due to Lemma 4.3.2 we see that the constant term not dependent on  $x$  in the definition of  $g(x)$  in equation 4.44 can be written as:

$$\begin{aligned} \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b]^2 &\sim \\ \frac{2}{2\pi} \phi(b)^2 \tau^2 &= -\sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 w_2 dw_2 \phi\left(\frac{w_2}{\sqrt{2}}\right) \phi(b) \frac{1}{2} \sqrt{2} \frac{1}{\sqrt{2\pi}}. \end{aligned} \quad (4.46)$$

Repeated use of the mean-value theorem together with the preparations above allows us now to rewrite equation 4.43 as:

$$\begin{aligned} \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(x) \geq b, Z(x+h_n) \leq b] &\sim \\ \sqrt{2}\phi(b)\tau^2 \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{2}}\right) &[\phi(b) - \xi\phi(\xi)\Delta_1(x)] \left[ \frac{1}{2} + \left( \frac{1}{\sqrt{\Sigma_{11}(x)}} - \frac{1}{2} \right) \right] \\ \times \left\{ m\Phi\left(-\frac{m}{s}\right) - \left[ \sqrt{2} + (s - \sqrt{2}) \right] \right. &[\phi(0) - \nu\phi(\nu)\Delta_2(x)] \left. \right\} dw_2, \end{aligned} \quad (4.47)$$

Hence we recognize that (besides the  $m\Phi(\cdot)$ -term) the integrand in equation 4.47 is up to a constant of the form  $(a+b)(c+d)(e+f)(g+h)$  and comparing with equation 4.46 we see that with these abbreviations the integrand in equation 4.46 equals up to the same constant the term  $aceg$ . If we now use the decomposition:

$(a+b)(c+d)(e+f)(g+h) - aceg = b(c+d)(e+f)(g+h) + ad(e+f)(g+h) + acf(g+h) + aceh$ , consider the  $m\Phi(\cdot)$ -term separately and estimate  $\phi(\cdot) \leq \frac{1}{\sqrt{2\pi}}$  and  $\Phi(\cdot) \leq 1$ , we can write:

$$\begin{aligned} \left| \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(x) \geq b, Z(x+h_n) \leq b] - \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b]^2 \right| &\sim \sqrt{2}\phi(b)\tau^2 \times \\ \left\{ \frac{C_1}{\sqrt{2\pi}} \int_{-\infty}^0 |w_2| \phi\left(\frac{w_2}{\sqrt{2}}\right) |m(x)| dw_2 + \frac{C_1 C_2}{2\pi} \int_{-\infty}^0 |w_2| \phi\left(\frac{w_2}{\sqrt{2}}\right) |\xi(x)| |\Delta_1(x)| dw_2 \right. \\ + \frac{C_2}{2\pi} \int_{-\infty}^0 |w_2| \phi\left(\frac{w_2}{\sqrt{2}}\right) \left| \frac{1}{\sqrt{\Sigma_{11}(x)}} - \frac{1}{2} \right| dw_2 + \frac{1}{4\pi} \int_{-\infty}^0 |w_2| \phi\left(\frac{w_2}{\sqrt{2}}\right) |s(x) - \sqrt{2}| dw_2 \\ \left. + \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^0 |w_2| \phi\left(\frac{w_2}{\sqrt{2}}\right) |\nu(x)| |\Delta_2(x)| dw_2 \right\}. \end{aligned}$$

As we have seen further above, the quantities  $|m(x)|$ ,  $|\xi(x)\Delta_1(x)|$ ,  $|1/\Sigma_{11}(x) - 1/2|$ ,  $|s - \sqrt{2}|$  and  $|\nu(x)\Delta_2(x)|$  can be estimated from above by summands each of which containing one of the terms  $|\rho(x)|$ ,  $|\rho'(x)|$  or  $|\rho''(x)|$ . Hence the integrability of  $\rho(\cdot)$ ,  $\rho'(\cdot)$  and  $\rho''(\cdot)$  carries over to  $g(\cdot)$  in the sense that the leading term in the probability difference considered in equation 4.48 can be bounded by summands of the form:

$$\text{const} \cdot |\rho(x)^{(i)}| \int_0^\infty |w_2|^k \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) dw_2 = \text{const} \cdot |\rho(x)^{(i)}|, \quad i \in \{0, 1, 2\},$$

because the involved Gaussian integrals exist and are finite. Hence in total we see that:

$$\int_{x_0}^\infty |g(x)| dx \leq K_0 \cdot \int_{x_0}^\infty |\rho(x)| dx + K_1 \cdot \int_{x_0}^\infty |\rho'(x)| dx + K_2 \cdot \int_{x_0}^\infty |\rho''(x)| dx \leq \infty,$$

where  $K_i, i \in \{0, 1, 2\}$  are some constants, which finally shows the  $x$ -integrability of  $g$ .

Please note that in the proof we have only used the integrability required for  $\rho^{(i)}(x), i \in \{0, 1, 2\}$ . However from equation 4.38 it becomes clear that to ensure that  $\text{Var}(\hat{s}_n)$  has the order of magnitude stated in the theorem, also  $x \cdot g(x)$  must be integrable. Obviously the same proof can be repeated to show the integrability of  $x \cdot g(x)$  under the condition that  $x \cdot \rho^{(i)}(x), i \in \{0, 1, 2\}$  are all integrable.  $\square$

For the proof of asymptotic normality because both the number of grid points  $n$  and the grid-spacing  $h_n$  are variable during the limit process we must resort to a central limit theorem for dependent random variables in a triangular array which usually involves proving some type of classical *Lindeberg*-condition. For both the On-Off-System in the appendix and also the Thresholded Gaussian Field Model below a fundamental result from (Peligrad 1996) about strong mixing arrays proves to be useful. Before we present this theorem, we need to introduce the concept of strong mixing:

**Definition 4.5.3. (Mixing)**

*i). Let  $(\Omega, \mathcal{A}, P)$  be a probability space and  $\mathcal{A}_1, \mathcal{A}_2$  be sub- $\sigma$ -algebras of  $\mathcal{A}$ . Then we define by*

$$\alpha(\mathcal{A}_1, \mathcal{A}_2) := \sup_{(A_1, A_2) \in \mathcal{A}_1 \times \mathcal{A}_2} |P[A_1 \cap A_2] - P[A_1]P[A_2]|$$

*the strong mixing coefficient of the  $\sigma$ -algebras  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . Similarly we denote:*

$$\phi(\mathcal{A}_1, \mathcal{A}_2) := \sup_{(A_1, A_2) \in \mathcal{A}_1 \times \mathcal{A}_2, P[A_1] \neq 0} \left| \frac{P[A_1 \cap A_2] - P[A_1]P[A_2]}{P[A_1]} \right|,$$

*the  $\phi$ -mixing coefficient of  $\mathcal{A}_1$  and  $\mathcal{A}_2$  and:*

$$\rho(\mathcal{A}_1, \mathcal{A}_2) := \sup_{f \in L_2(\mathcal{A}_1), g \in L_2(\mathcal{A}_2)} |\text{corr}(f, g)|,$$

*the maximal coefficient of correlation of the  $\sigma$ -algebras  $\mathcal{A}_1$  and  $\mathcal{A}_2$ .*

ii). Let  $(X_i)_{i \in \mathbb{Z}}$  be a stochastic process. Then the **strong mixing coefficients** of  $(X_i)_{i \in \mathbb{Z}}$  are

$$\alpha_\tau := \sup_{t \in \mathbb{Z}} \alpha(\mathcal{F}_{t-\tau}, \mathcal{G}_t),$$

where  $\mathcal{F}_l = \sigma(X_i, i \leq l)$  is the past and  $\mathcal{G}_l = \sigma(X_i, i \geq l)$  is the future of the process at time  $l$ . Similarly:

$$\phi_\tau := \sup_{t \in \mathbb{Z}} \phi(\mathcal{F}_{t-\tau}, \mathcal{G}_t), \quad \rho_\tau := \sup_{t \in \mathbb{Z}} \rho(\mathcal{F}_{t-\tau}, \mathcal{G}_t),$$

are the  $\phi$ -mixing coefficients and maximal correlation coefficients of  $(X_i)_{i \in \mathbb{Z}}$  respectively. If  $\{X_{n,i}, 1 \leq i \leq k_n\}$  is a triangular array we analogously define:

$$\begin{aligned} \alpha_{n,k} &:= \sup_{t \in \mathbb{Z}} \alpha(\sigma(X_{n,i}), i \leq t, \sigma(X_{n,j}), j \geq t+k), \quad \text{and} \quad \alpha_k := \sup_n \alpha_{n,k}, \\ \phi_{n,k} &:= \sup_{t \in \mathbb{Z}} \phi(\sigma(X_{n,i}), i \leq t, \sigma(X_{n,j}), j \geq t+k), \quad \text{and} \quad \phi_k := \sup_n \phi_{n,k}, \\ \rho_{n,k} &:= \sup_{t \in \mathbb{Z}} \rho(\sigma(X_{n,i}), i \leq t, \sigma(X_{n,j}), j \geq t+k), \quad \text{and} \quad \rho_k := \sup_n \rho_{n,k}. \end{aligned}$$

iii). We say that a stochastic process  $(X_i)_{i \in \mathbb{Z}}$  satisfies the **strong mixing condition** (or is **strongly mixing**) if  $\alpha_\tau \xrightarrow{\tau \rightarrow \infty} 0$ , whereas the triangular array  $\{X_{n,i}, 1 \leq i \leq k_n\}$  is called to satisfy the **strong mixing condition** if  $\lim_{k \rightarrow \infty} \alpha_k = 0$ .

Similarly the process  $(X_i)_{i \in \mathbb{Z}}$  is called  $\phi$ -mixing (or **uniformly strong mixing**) if  $\phi_\tau \xrightarrow{\tau \rightarrow \infty} 0$ , whereas the triangular array  $\{X_{n,i}, 1 \leq i \leq k_n\}$  is called to satisfy the  $\phi$ -mixing condition if  $\lim_{k \rightarrow \infty} \phi_k = 0$ .

Analogously the mixing coefficients can be defined for continuous processes  $(X_t)_{t \in \mathbb{R}}$ . It can easily be seen that  $\phi$ -mixing implies strong mixing. For a summary and discussion of other mixing-conditions and their relationships, see for example (Lahiri 2003 pg. 45,46). We now state the fundamental central limit theorem for triangular arrays:

**Theorem 4.5.4. (Central Limit Theorem for Strongly Mixing Triangular Arrays)** Let  $\{X_{n,i}, 1 \leq i \leq k_n\}$  be a triangular array of centered random variables, which is strongly mixing and has finite second moments. Assume further that  $\lim_{k \rightarrow \infty} \rho_k < 1$ . and denote  $S_n := \sum_{i=1}^{k_n} X_{n,i}$  and  $\sigma_n^2 := \text{Var}(S_n)$ . If the following two conditions are fulfilled:

i)

$$\sup_n \frac{1}{\sigma_n^2} \sum_{i=1}^{k_n} \mathbf{E} [X_{n,i}^2] < \infty,$$

ii)

$$\forall \epsilon > 0: \quad \frac{1}{\sigma_n^2} \sum_{i=1}^{k_n} \mathbf{E} [X_{n,i}^2 \mathbf{1}_{\{|X_{n,i}| > \epsilon \sigma_n\}}] \xrightarrow{n \rightarrow \infty} 0,$$

then the Central Limit Theorem holds, ie.  $\frac{S_n}{\sigma_n} \xrightarrow{d} \mathcal{N}(0, 1)$ .

**Proof.** (Theorem 4.5.4) see (Peligrad 1996).

We now turn to the main result in this section, the proof of asymptotic normality of  $\hat{s}_n$  under moderate assumptions about the decay of the strong mixing coefficients of the underlying Gaussian Field. Similarly as in the Markov case of the appendix we can rely on the fundamental central limit theorem for strong mixing sequences, Theorem 4.5.4.

In the following we consider again the classical standardized linear combination  $a_n \cdot (\hat{s}_n - \bar{s})$ . From Theorem 4.5.2 we learn that to stabilize expectation and variance in the case of the thresholded Gaussian field, we need:

$$\begin{aligned} \text{Var}[\hat{s}] \xrightarrow{n \rightarrow \infty} 0 &\iff n \cdot h_n \xrightarrow{n \rightarrow \infty} \infty \\ \mathbf{E} [a_n(\hat{s} - s)] \xrightarrow{n \rightarrow \infty} c_1 &\iff a_n \cdot h_n^2 \xrightarrow{n \rightarrow \infty} c_1, \\ \text{Var} (a_n(\hat{s} - s)) \xrightarrow{n \rightarrow \infty} c_2 &\iff \frac{a_n^2}{n \cdot h_n} \xrightarrow{n \rightarrow \infty} c_2, \end{aligned}$$

where  $c_1$  and  $c_2$  are some constants. But then:

$$c_2 \sim \frac{a_n^2}{\sqrt{\frac{c_1}{a_n}} n} \iff \frac{a_n^{\frac{5}{2}}}{n} = \sqrt{c_1} c_2 \iff a_n \sim n^{\frac{2}{5}},$$

and thus  $h_n \sim \frac{1}{\sqrt{a_n}} \sim n^{-\frac{1}{5}}$ .

Note that this relationship between  $n, a_n$  and  $h_n$  is different from the Markovian situation in the appendix because in the thresholded Gaussian model the bias is of smaller order.

As already mentioned for the asymptotic normality of  $\hat{s}_n$  we want to rely on the strong mixing property (definition 4.5.3) of the underlying Gaussian Field. It can be seen that stationary Gaussian Fields are strongly mixing under mild conditions concerning their spectral density. For example if the spectral density of a Gaussian sequence is continuous and bounded away from zero, it satisfies the strong mixing condition (Ibragimov 1971, Theorem 17.3.3). But even the convergence rate of the strong mixing coefficients can be controlled. The conditions in the following discrete-time theorem not only ensure that a Gaussian process is strongly mixing but even provide information about the rate of convergence of the mixing coefficients which will be useful for us later.

**Theorem 4.5.5. (Convergence rate for strong mixing coefficients)** *Let  $(X_i)_{i \in \mathbb{Z}}$  be a Gaussian process. A necessary and sufficient condition for*

$$\alpha(\tau) = \mathcal{O}(\tau^{-\tau-\beta}), \text{ where } 0 < \beta < 1,$$

*is that the spectral density  $f(\lambda)$  of the process permits a representation of the form  $f(\lambda) = |P(e^{i\lambda})|^2 w(\lambda)$ , where  $P(z)$  is a polynomial with zeroes on  $\{|z| = 1\}$  and where the function  $w(\lambda)$  is strictly positive,  $w(\lambda) \geq m > 0$ , and is  $r$  times differentiable with the  $r$ -th derivative satisfying a Hölder-condition of order  $\beta$ .*

**Proof. (Theorem 4.5.5)** The theorem is mainly Theorem 8 in Chapter V.6 of (Ibragimov and Rozanov 1978). There the statement is proved for the maximal correlation coefficient but it is well known that in the Gaussian case the strong mixing coefficients are bounded by the maximal correlation coefficients.

Similar theorems with slightly more complicated conditions can be found for the continuous case as well in (Ibragimov and Rozanov 1978). We can now state the main result of this subsection:

**Theorem 4.5.6. (Asymptotic normality of the surface-estimator in Thresholded Gaussian Field Model)** Let  $\Theta(x)_{x \in \mathbb{R}} = \mathbf{1}_{\{Z(x) \geq b\}}$  be a stationary continuous Thresholded Gaussian Field with threshold  $b$  and assume the underlying Gaussian process  $Z(x)_{x \in \mathbb{R}}$  is strongly mixing with a covariance function  $\rho(h)$  which fulfills the integrability conditions of Theorem 4.5.2. Assume further that  $\Theta(x)_{x \in \mathbb{R}}$  is observed on a refining sequence of lattices  $\Gamma_n$  containing  $n$  points and lattice spacing  $h_n$  such that  $nh_n^3 \rightarrow \infty$  and  $\lim_{n \rightarrow \infty} nh_n^5 < \infty$ . Let the estimator  $\hat{s}_n$  be defined as in equation 4.36. Then for  $a_n := \sqrt{nh_n}$ , the linear combination  $a_n(\hat{s}_n - s)$  converges in distribution to a Gaussian variable with constant finite expectation and variance. Especially for the (with respect to bias-variance-tradeoff) optimal lattice spacing  $h_n = n^{-1/5}$  we get:

$$n^{\frac{2}{5}} \cdot (\hat{s}_n - s) \xrightarrow{d} \mathcal{N} \left( -\frac{1}{12} \cdot \frac{(-\rho''(0))^{\frac{3}{2}}}{\sqrt{2\pi}} \cdot \phi(b) \cdot \left( (b^2 - 1) + \frac{\rho^{(iv)}(0)}{\rho''(0)^2} \right), k_1 + \int_{-\infty}^{\infty} g(x) dx \right).$$

**Remark 4.5.7.** Before we prove the theorem we have to mention that the results of Theorems 4.5.2 and 4.5.6 unfortunately are not new but independently rediscovered versions of some known results. Because calculating the one-dimensional specific-surface-estimator of equation 4.36 basically means calculating the number of crossings of a continuous one-dimensional Gaussian process, some results from the vast theory on (up-)crossings of Gaussian processes apply. We address here only the most important results in this context. A complete survey on the actual state of the research in this area can be found in (Kratz 2006). We emphasize that throughout this remark we consider continuous stationary one-dimensional Gaussian processes. We denote the number of crossings of a certain level  $x$  in the interval  $[0, t]$  by  $N_t(x)$  and the number of upcrossings by  $N_t^+(x)$ . Cramér and Leadbetter in (Cramér and Leadbetter 1968) gave a condition for the finiteness of the variance of the specific surface estimator if the threshold is zero:

$$\exists \delta > 0, \text{ such that } \frac{\rho''(t) - \rho''(0)}{t} \in L^1([0, \delta], dx) \Rightarrow \mathbf{E} [N_t^2(0)] < \infty \quad (4.48)$$

The condition on the covariance function on the left hand side of equation 4.48 is called the *Geman-condition*. Cramér and Leadbetter also were able to provide quite an explicit integral expression for this variance even for general threshold  $x$  involving joint densities of values and derivatives:

$$\mathbf{E} [N_t^2(x)] = \frac{t}{\pi} e^{-\frac{x^2}{2}} \sqrt{-\rho''(0)} + 2 \int_0^t (t-u) \int_{\mathbb{R}^2} |\dot{x}| |\dot{y}| \phi_u(x, \dot{x}, x, \dot{y}) dx dy du.$$

Recently Kratz and León in (Kratz and León 2006) generalized the result of equation 4.48 for a general threshold  $x$  and also showed that in fact the Geman condition is *both* sufficient *and* necessary.

But the most important result for our purposes which directly implies the central limit theorem of the 1D-specific surface estimator under the same integrability condition as in Theorem 4.5.2 is the following result by Peterburg:

**Theorem 4.5.8.** (CLT for the Number of Upcrossings, Piterbarg, 1978) Let  $Z_t, t \in \mathbb{R}$  be a stationary Gaussian Process, mean zero, unit variance, satisfying the Geman-condition from equation 4.48 and the integrability condition on the covariance function and its first two derivatives from Theorem 4.5.2. Then:

$$\text{Var} (N_t^+(x)) = \sigma^2 t (1 + o(1)) \text{ as } t \rightarrow \infty,$$

where:

$$\sigma^2 := \int_0^\infty \int_0^\infty \int_0^\infty yz \left( \phi_s(x, y, x, z) - \frac{e^{-x^2} \sqrt{-\rho''(0)}}{2\pi} \text{Exp} \left( \frac{y^2 + z^2}{2\rho''(0)} \right) \right) dy dz ds + \frac{\sqrt{-\rho''(0)} e^{-x^2}}{2\pi} > 0.$$

In addition to that, the central limit theorem for  $N_x^+(t)$  holds.

**Proof.** (Theorem 4.5.8) see (Piterbarg 1978) for details.

We now turn to the proof of Theorem 4.5.6:

**Proof.** (Theorem 4.5.6) The proof of Theorem 4.5.6 resorts to the same fundamental central limit theorem (Theorem 4.5.4) as the proof of the asymptotic normality in the on-off-Markov-system (Theorem 7.1.1, part iv)). Therefore for some arguments of this proof we ask the reader to consult the proof of the aforementioned theorem in the appendix. Instead of the estimator of equation 4.36 we may consider without loss of generality the modified estimator of equation 7.12 with  $E_{i,n} = \{\Theta(ih_n) \neq \Theta((i+1)h_n)\}$ . Again we consider the standardized sum  $S_n$  and introduce the block-variables  $X_{i,n}$  as in equation 7.13. Now we can proceed as in the proof of Theorem 7.1.1, part iv). The strong mixing property of the triangular array  $\{X_{n,i}, 1 \leq i \leq nh_n\}$  and  $\lim_{k \rightarrow \infty} \phi_k < 1$  both follow from equation 7.14 and the relationship between  $\rho_k$  and  $\alpha_k$ . The verification of condition i) in Theorem 4.5.4 can be literally copied from the proof of Theorem 7.1.1 and can be done analogously as in equation 7.15. For the verification of condition ii) in Theorem 4.5.4 we see that:

$$|X_{n,i}| \leq \frac{j_n}{\sqrt{nh_n}} \sim \frac{1}{\sqrt{nh_n^3}} \xrightarrow{n \rightarrow \infty} 0,$$

and thus  $\forall \epsilon > 0$  and  $\forall i$  we have  $\mathbf{1}_{\{|X_{n,i}| > \epsilon \sigma_n\}} = 0$  for  $n$  large enough. Therefore condition ii) in 4.5.4 becomes trivial. Hence the asymptotic normality follows from Theorem 4.5.4. The moments of the limiting distribution are derived from Theorem 4.5.2.  $\square$

## 4.6 The Generalized Thresholded Field Model

We now come to the more practical aspects of our survey of the Thresholded Gaussian Model. Again as for the Boolean Model the goal is to simulate Thresholded Gaussian realizations with preset values  $v_i$  for the specific Minkowski-functionals in two and three dimensions:

$$\bar{V}_0(\Theta), \dots, \bar{V}_d(\Theta) = (v_0, \dots, v_d), \quad d \in \{2, 3\}. \tag{4.49}$$

Note that opposed to the Boolean Model the Thresholded Gaussian Field Model can be simulated directly on a grid and hence no cumbersome discretization process is needed.

The Gaussian Model is not of germ-grain type and hence its realizations do not consist of physical entities which can be compared to corresponding entities in the real structure. In this sense the Gaussian Model is more flexible than the Boolean one and doesn't exhibit the obvious defects of the latter such as overlapping grains or the restriction of the grain shapes to a certain geometric family. Moreover because Gaussian Random Fields have been studied extensively in literature, the Gaussian Model also has the advantage of being analytically easily tractable. However the Thresholded Gaussian Field Model as is has been defined in definition 4.1.4 unfortunately cannot be used directly to produce artificial binary structures with given specific Minkowski functionals. This can be seen from the Adler-Tomita-formulae (Corollaries 4.2.3 and 4.2.3). In both the two-dimensional and the three-dimensional situation the same problem arises. Once the specific volume is preset, the threshold  $b$  is determined and besides this threshold  $b$ , all other specific Minkowski functionals only depend on the second derivative of the covariance function in zero. Once the specific volume is given, it is therefore not possible to specify more than *one* further specific Minkowski functional in advance. It is somehow surprising that to determine the specific Minkowski functionals only very little information about the covariance function of the underlying Gaussian field is needed and also that it is the very same information for all of them. This fact also shows that at least in the Gaussian case the covariance function contains much more information than the specific Minkowski functionals and varying the covariance function outside a small region around zero might completely change the appearance of the corresponding realized artificial structures whereas the specific Minkowski still stay the same. This is again the *indeterminacy*-phenomenon already discussed for the Boolean Model before.

However already a slight generalization of the defining model equation 4.49 removes the problem of dependent specific Minkowski functionals. Instead of just truncating the underlying stationary Gaussian field  $Z(x)$  when it exceeds a certain threshold  $b$ , we introduce a second threshold  $a$  and also truncate  $Z(x)$  when it drops below the level  $a$ . This idea can be generalized to introduce as much flexibility into the model as needed by considering truncation of more than one disjoint intervals. This leads us to the *Generalized Thresholded Gaussian Field Model*:

**Definition 4.6.1. (*Generalized Thresholded Gaussian Field Model*)** Let  $Z(x), x \in \mathbb{R}^d$  be a (real-valued) stationary, isotropic and non-erratic Gaussian Field. Let further  $n \in \{1, 2, \dots\}$  and  $[a_i, b_i], i \in \{1, \dots, n\}$  be pairwise disjoint real intervals. The **Generalized Thresholded Field Model**  $\Theta(x), x \in \mathbb{R}^d$  is then defined as:

$$\Theta(x) = \sum_{i=1}^n \mathbf{1}_{\{a_i \leq Z(x) \leq b_i\}}.$$

The Minkowski functionals in the Generalized Thresholded Field Model are still computable due to the additivity property (Theorem 2.1.3, part vi.):

**Corollary 4.6.2. (*2D-Minkowski-functionals in the Generalized Thresholded Gaussian Model*)** Let  $\Theta(x) = \sum_{i=1}^n \mathbf{1}_{\{a_i \leq Z(x) \leq b_i\}}$  be a two-dimensional stationary, isotropic and non-erratic generalized Thresholded Gaussian Field. Then its specific Minkowski functionals can be expressed as follows:



$$\begin{aligned}\bar{v}(\Theta) &= \sum_{i=1}^n (\Phi(b_i) - \Phi(a_i)), \\ \bar{s}(\Theta) &= \sqrt{\frac{\pi}{2}} \cdot \sqrt{-\rho''(0)} \sum_{i=1}^n (\phi(a_i) + \phi(b_i)) \\ \bar{\chi}(\Theta) &= \frac{1}{2\pi} \cdot \rho''(0) \cdot \sum_{i=1}^n (b_i \cdot \phi(b_i) - a_i \cdot \phi(a_i)).\end{aligned}$$

**Proof. (Corollary 4.6.2)**

Because the sets  $\{a_i \leq Z(x) \leq b_i\}$  are disjoint  $\forall i$  we can use additivity and sum up all contributions of the different intervals and restrict ourselves to the single-interval case  $n = 1$ :

ad  $\bar{v}$ :

$$\bar{v}(\mathbf{1}_{\{a \leq Z(x) \leq b\}}) = P[a \leq Z(x) \leq b] = \Phi(b) - \Phi(a).$$

ad  $\bar{s}$ :

$$\begin{aligned}\bar{s}(\mathbf{1}_{\{a \leq Z(x) \leq b\}}) &= \bar{s}(\mathbf{1}_{\{Z(x) \leq a\}}) + \bar{s}(\mathbf{1}_{\{Z(x) \geq b\}}) \\ &= \bar{s}(\mathbf{1}_{\{Z(x) \geq -a\}}) + \bar{s}(\mathbf{1}_{\{Z(x) \geq b\}}),\end{aligned}$$

where we have used  $Z(x) \sim -Z(x)$  and  $\bar{s}(\Theta) = \bar{s}(\Theta^c)$  and now use corollary 4.2.3 for the two last summands.

ad  $\bar{\chi}$ :

$$\begin{aligned}\bar{\chi}(\mathbf{1}_{\{a \leq Z(x) \leq b\}}) &= -\bar{\chi}(\mathbf{1}_{\{Z(x) \leq a\}}) - \bar{\chi}(\mathbf{1}_{\{Z(x) \geq b\}}) \\ &= -\bar{\chi}(\mathbf{1}_{\{Z(x) \geq -a\}}) - \bar{\chi}(\mathbf{1}_{\{Z(x) \geq b\}}),\end{aligned}$$

where we have used  $Z(x) \sim -Z(x)$  and  $\bar{\chi}(\Theta) = -\bar{\chi}(\Theta^c)$  in two dimensions according to Theorem 2.1.13 and now use corollary 4.2.3 for the two last summands.  $\square$

**Corollary 4.6.3. (3D-Minkowski-functionals in the Generalized Thresholded Gaussian Model)** Let  $\Theta(x) = \sum_{i=1}^n \mathbf{1}_{\{a_i \leq Z(x) \leq b_i\}}$  be a three-dimensional stationary, non-erratic generalized thresholded Gaussian field. Then its specific Minkowski functionals can be expressed as follows:

$$\begin{aligned}\bar{v}(\Theta) &= \sum_{i=1}^n (\Phi(b_i) - \Phi(a_i)), \\ \bar{s}(\Theta) &= \frac{4}{\sqrt{2 \cdot \pi}} \cdot \sqrt{-\rho''(0)} \cdot \sum_{i=1}^n (\phi(a_i) + \phi(b_i)), \\ \bar{M}(\Theta) &= \rho''(0) \cdot \sum_{i=1}^n (b_i \cdot \phi(b_i) - a_i \cdot \phi(a_i)), \\ \bar{\chi}(\Theta) &= \frac{1}{(2 \cdot \pi)^{3/2}} \cdot (-\rho''(0))^{\frac{3}{2}} \cdot \sum_{i=1}^n [(a_i^2 - 1) \cdot \phi(a_i) + (b_i^2 - 1) \cdot \phi(b_i)].\end{aligned}$$

**Proof. (Corollary 4.6.3)** Completely analogous to the proof of corollary 4.6.2 but here we note that  $\bar{M}(\Theta) = -\bar{M}(\Theta^c)$  which is clear from definition 2.2.5 and in three dimensions  $\bar{\chi}(\Theta) = \bar{\chi}(\Theta^c)$  which stems from Theorem 2.1.13.  $\square$

## 4.7 Simulation for the Whittle-Matérn-Covariance Model

The crucial step in the simulation of the Generalized Thresholded Gaussian Field Model  $\Theta(x)$  is of course the simulation of the underlying stationary Gaussian Random Field  $Z(x)$  on a grid  $\Gamma$ . This is usually easy in one-dimension and still well tractable in two, but in three dimensions the curse of dimensionality leads to problems which are difficult to overcome. The trivial solution  $Z(x) = \sqrt{\Sigma} \cdot Y$ , where  $Y$  is a vector of iid. univariate  $\mathcal{N}(0, 1)$ -variables and  $\sqrt{\Sigma}$  is such that  $\sqrt{\Sigma} \cdot \sqrt{\Sigma}^T$  equals the covariance matrix  $\Sigma$  becomes unfeasible for large grids  $\Gamma$  because the numerical computation of  $\sqrt{\Sigma}$  quickly gets too difficult for general covariance functions  $\rho(h)$ . However for covariance functions of finite support direct simulation is still an option also for large fields because of the sparse band-structure of the covariance matrix  $\Sigma$ . For simulation of general Gaussian Fields on a grid various techniques are known in literature eg. the circulant embedding, turning bands or random coin methods among which the circulant embedding technique is the fastest for general Gaussian Fields. Gaussian random fields can also be simulated using marked point processes. For an overview of simulation methods see (Schlather 1999). The same author provides algorithms for simulating Random Fields (not necessarily Gaussian) implemented as a user-friendly software package called *RandomFields* (Schlather 2001) written in the R language for statistical computing (R Development Core Team 2006) and available at <http://cran.r-project.org>. For our experiments, we did always resort to the *RandomFields*-package to simulate the underlying Gaussian Random Field. However, it turns out that the simulation of three-dimensional Gaussian fields of considerable size such as used in this thesis ( $800^3 \approx 10^9$  pixels) may well exhaust the limits of these algorithms for certain covariance models at least when a simulation result is expected in reasonable time on an ordinary machine. Highly efficient algorithms and an exploration of how far we can reach using contemporary computer power with respect to image size are given in (Gneiting, Sevcikova, Percival, Schlather and Jiang 2006) for the two-dimensional situation.

We should mention in this context also the *R*-package *geoR* (Ribeiro Jr. and Diggle 2001) which was originally written for statistical analysis in geostatistics and is very helpful if one attempts to fit not only specific Minkowski-functionals but whole covariance functions or variograms to data.

In the following we restrict ourselves to considering the Whittle-Matérn-Covariance-Model introduced in definition 4.1.9 for both our two and three-dimensional simulation algorithms. This choice is to some extent arbitrary. We chose it because it is very popular in geostatistical modelling and offers a certain degree of flexibility because it is two-parametric and the degree of differentiability of the underlying Gaussian Random Field can be controlled ( $\kappa \geq (2m + 1)/2$  implies that the field is at least  $m$  times differentiable). But the plethora of other standard covariance models could be used as well because only the second derivative at zero of the covariance function  $\rho(h)$  enters the equations for the specific Minkowski functionals in corollaries 4.6.2 and 4.6.3.

### 4.7.1 2D-Simulation of Generalized Thresholded Gaussian Fields with the Whittle-Matérn Covariance Model

To simulate artificial images with predetermined specific Minkowski-functionals with the 2D-Generalized Thresholded Field Model  $\Theta$  and the Whittle-Matérn covariance function,

we use the expressions from Corollary 4.6.2 for  $(\bar{v}(\Theta), \bar{s}(\Theta), \bar{\chi}(\Theta))$ . Because in two dimensions we have only three Minkowski functionals, a model with only one single interval  $[a, b]$  provides already enough flexibility. The four model parameters are then the two thresholds  $a$  and  $b$  and the two Whittle-Matérn parameters  $\kappa$  and  $\nu$  (see definition 4.1.9) which must be chosen such that the three equations for the specific Minkowski-functionals are fulfilled. The model is under-determined and thus we can choose the Matérn shape-parameter  $\kappa$  to control the covariance-structure of the model. The equation system from Corollary 4.6.2 can easily be solved numerically for the thresholds  $a$  and  $b$  and the second derivative of the Whittle-Matérn covariance function in zero,  $\rho''_{\kappa,\nu}(0)$ . In general it is hard to express  $\rho''_{\kappa,\nu}(0)$  as a function of the two parameters  $\kappa$  and  $\nu$ . However once an expression for  $\rho''_{\kappa,1}(0)$  is available (which could be estimated for example by a simple polynomial approximation for arbitrary  $\kappa$ ) we can use the relationship

$$\rho''_{\kappa,\nu}(0) = \frac{\rho''_{\kappa,1}(0)}{\nu^2}, \quad (4.50)$$

to determine the shape-parameter  $\nu$ . This follows directly from  $\rho_{\kappa,1}(h) = \rho_{\kappa,\nu}(\nu h)$ . The shape parameter  $\kappa$  can be an arbitrary value in  $\mathbb{R}^+$ . However, if one contents oneself with considering integer values  $\geq 3$  for  $\kappa$ , then no estimation is needed at all. From (Abramowitz and Stegun 1972, equation 9.6.28) we learn from the properties of the K-Bessel-function:

$$\rho''_{\kappa,1}(0) = -\frac{(\kappa - 2)!}{2 \cdot (\kappa - 1)!}, \quad \kappa \in \mathbb{N}, \quad \kappa \geq 3. \quad (4.51)$$

Hence once  $\rho''_{\kappa,\nu}(0)$  is determined using corollary 4.6.2 and the shape parameter  $\kappa$  is chosen we can use equation 4.51 in conjunction with equation 4.50 to determine the remaining scale parameter  $\nu$ . This procedure is summarized in the following algorithm:

**Algorithm 4.7.1. (2D-Generalized Thresholded Field with Whittle-Matérn-Covariance)** To simulate on a finite grid  $\Gamma$  a 2D-stationary and isotropic Generalized Thresholded Gaussian Field-Model with a single interval  $[a, b]$ , predetermined values for the specific Minkowski functionals  $\bar{V}_j(\Theta)$  and the Whittle-Matérn Covariance-Model  $\rho_{\kappa,\nu}(h)$ , proceed as follows:

- i). Choose the Matérn shape-parameter  $\kappa \in \mathbb{N}, \kappa \geq 3$  to determine the covariance structure of the underlying stationary Gaussian Field.
- ii). Solve the equation system from corollary 4.6.2 numerically for  $a, b$  and  $\rho''_{\kappa,\nu}(0)$ .
- iii). Determine the remaining Matérn scale parameter  $\nu$  according to:

$$\nu = \sqrt{\frac{(\kappa - 2)!}{2 \cdot (-\rho''_{\kappa,\nu}(0)) \cdot (\kappa - 1)!}} \quad (4.52)$$

- iv). Simulate a Gaussian Random Field  $Z(x)$  on  $\Gamma$  with covariance model  $\rho_{\kappa,\nu}(h)$  and generate  $\Theta(x) = \mathbf{1}_{\{a \leq Z(x) \leq b\}}$ .

**Remark 4.7.2.**

- i) In step i) we choose the shape parameter  $\kappa$ . Of course  $\kappa$  should be chosen to match the covariance-function of the given structure as closely as possible. To achieve this, we always estimated the covariance-function of the sand-structure first (see figure 1.2 for an example) and chose  $\kappa$  to provide at least a reasonable optical fit. The same remark applies for algorithm 4.7.3 below.

- ii) For step iv). for example the R-function *GaussRF* from the Random Fields-package can be used. This fully-automated function has the advantage that it searches for the fastest valid simulation method for the given model specifications itself.

#### 4.7.2 3D-Simulation of Generalized Thresholded Gaussian Fields with the Whittle-Matérn Covariance Model

The simulation of the Generalized Thresholded Field Model in three dimensions is completely analogous to the simulation in the 2D-case described in subsection 4.7.1. Because in 3D we have 4 specific Minkowski functionals we need to consider a model with two disjoint truncation intervals  $[a, b]$  and  $[c, d]$ . The model is then under-determined and our experiments have shown that it is best to choose a target scale  $\nu \in \mathbb{R}^+$  and a target shape  $\kappa \in \mathbb{N}$ ,  $\kappa \geq 3$  in advance which then determine  $\rho''_{\kappa, \nu}(0)$  via equations 4.50 and 4.51. We then solve the equation system in corollary 4.6.3 for the four interval boundary thresholds  $a, b, c$  and  $d$ . Note that it is crucial that the two intervals  $[a, b]$  and  $[c, d]$  are disjoint and in the derivation of Corollary 4.6.3 it was implicitly assumed that  $a < b < c < d$  or  $c < d < a < b$ . However it is possible that the solution  $(a, b, c, d)$  of the equation system in corollary 4.6.3 violates both these restrictions. From a certain symmetry inherent in the equations of the system it is obvious that if  $(a, b, c, d)$  is a solution, also  $(c, b, a, d)$ ,  $(a, d, c, b)$  and  $(c, d, a, b)$  are. Hence by permutation we have also found a feasible solution if  $c < b < a < d$  or  $a < d < c < b$ . If the solution doesn't exhibit either of the four possible orderings one has to change the values for  $\nu$  and  $\kappa$ . Because of the curse of dimensionality it is not possible to generate very large realizations in reasonable time. Images larger than the ones shown in section 4.9 ( $\approx 10^9$  pixels) are already cumbersome to work with. But fortunately the Matérn scale parameter  $\nu$  can be used to simulate low resolution images of large extracts of the porous structure. For example an  $800^3$ -extract can be simulated as a  $256^3$ -pixel image by simply choosing  $\nu = 256/800$ . Again we summarize the simulation procedure in an algorithm:

**Algorithm 4.7.3. (3D-Generalized Thresholded Field with Whittle-Matérn-Covariance)** To simulate on a finite grid  $\Gamma$  a 3D-stationary and isotropic Generalized Thresholded Gaussian Field-Model with two disjoint intervals  $[a, b]$  and  $[c, d]$ , predetermined values for the specific Minkowski functionals  $\bar{V}_j(\Theta)$  and the Whittle-Matérn Covariance-Model  $\rho_{\kappa, \nu}(h)$ ,

- i). Choose the Matérn scale-parameter  $\nu$  according to the size of the image and the desired resolution  $r$ . To generate a cubic image of size  $n^3$  pixels with a resolution of  $r^3$  pixels, choose  $\nu = \frac{r}{n}$ . Also choose the Matérn shape-parameter  $\kappa \in \mathbb{N}$ ,  $\kappa \geq 3$  to determine the covariance structure of the model.
- ii). Calculate  $\rho''_{\kappa, \nu}(0)$  as:

$$\rho''_{\kappa, \nu}(0) = -\frac{1}{\nu^2} \cdot \frac{(\kappa - 2)!}{2 \cdot (\kappa - 1)!}$$

and solve the equation system from Corollary 4.6.3 numerically for  $a, b, c$  and  $d$ . The solution is feasible if one of the following orderings holds:

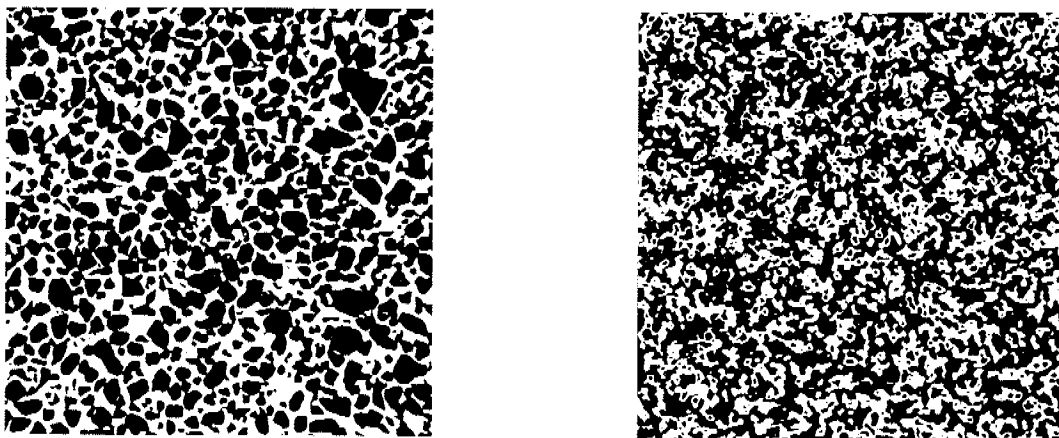
$$\begin{aligned} a < b < c < d, & \quad c < b < a < d, \\ a < d < c < b, & \quad c < d < a < b. \end{aligned}$$

If the solution is feasible, permute  $a \leftrightarrow c, b \leftrightarrow d$  or both such that  $a < b < c < d$  holds.

- iii). Simulate a Gaussian Random Field  $Z(x)$  on  $\Gamma$  with covariance model  $\rho_{\kappa, \nu}(h)$  and generate  $\Theta(x) = \mathbf{1}_{\{a \leq Z(x) \leq b\}} + \mathbf{1}_{\{c < Z(x) \leq d\}}$ .

## 4.8 2D-Results

In this section we show a few images generated with the Thresholded Gaussian Field Model and the Whittle-Matérn covariance function according to algorithm 4.7.1. The goal was to produce artificial structures which have the same specific Minkowski-functionals as a cross-section from the fine-sand specimen shown in the right image of figure 1.1. We have tried out various choices for the shape parameter  $\kappa$  and we have also simulated images of different size by varying the scale parameter  $\nu$  while always keeping the resolution fixed at 800 pixels. Figure 4.1 shows an 800<sup>2</sup>-pixel-sized cross-section taken from the center of the fine sand-specimen shown in figure 1.1 along with a realization of the Generalized Thresholded Gaussian Field Model. Both have the same specific Minkowski functionals up to a negligible error.

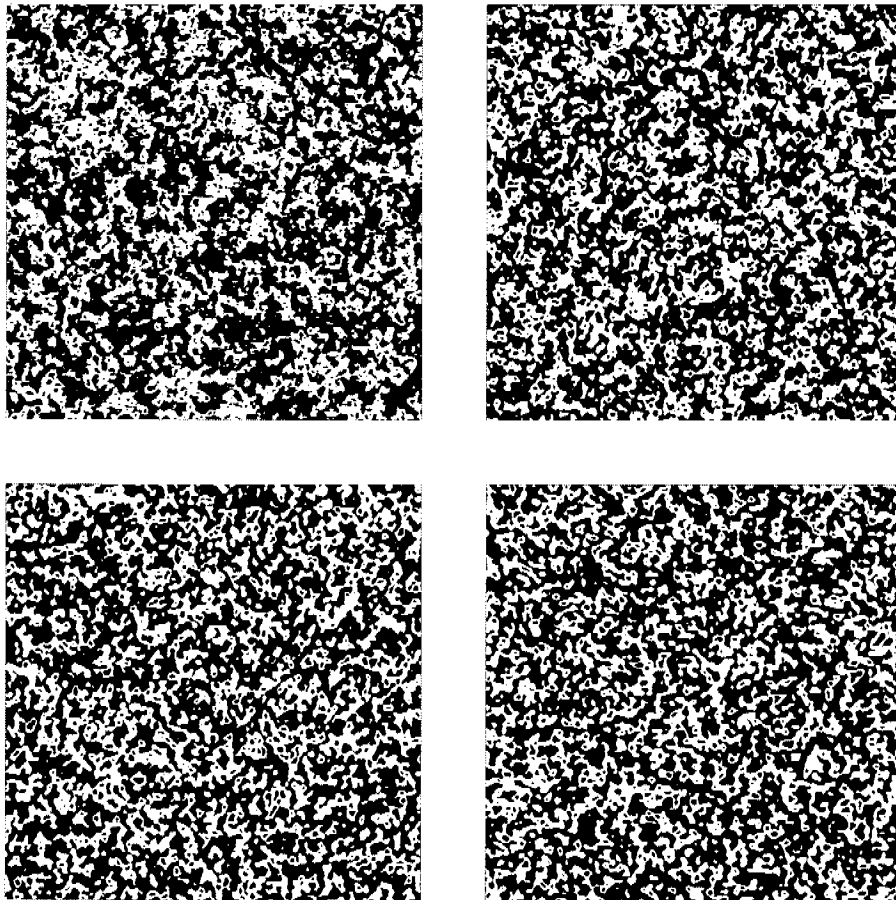


**Figure 4.1:** *An original 800<sup>2</sup>-pixel-cross section taken from the 800<sup>3</sup>-pixel samples of the HASYLAB-synchrotron sand data with fine granularity and side length = 11 $\mu$ m on the left hand-side. The right image shows a realization of the 2D-Generalized Thresholded Gaussian Field Model according to algorithm 4.7.1. The two structures share the same specific Minkowski functionals.  $\kappa$  was chosen to be 7.*

Figure 4.1 shows that the Generalized Thresholded Field Model is able to generate a granular structure but compared to the real data cross-sections the "grains" are too small. An other obvious defect is that it produces too many small black artifacts which increases the Euler characteristic and thus the system of black grains is too densely connected to compensate for this excess in Euler characteristic coming from the small artifacts. This effect can be alleviated by choosing a larger  $\kappa$  and hence imposing stronger positive dependence. This inhibits too tiny black (or white) regions. But figure 4.2 shows that the effect cannot completely be removed. This is because the shift towards a stronger

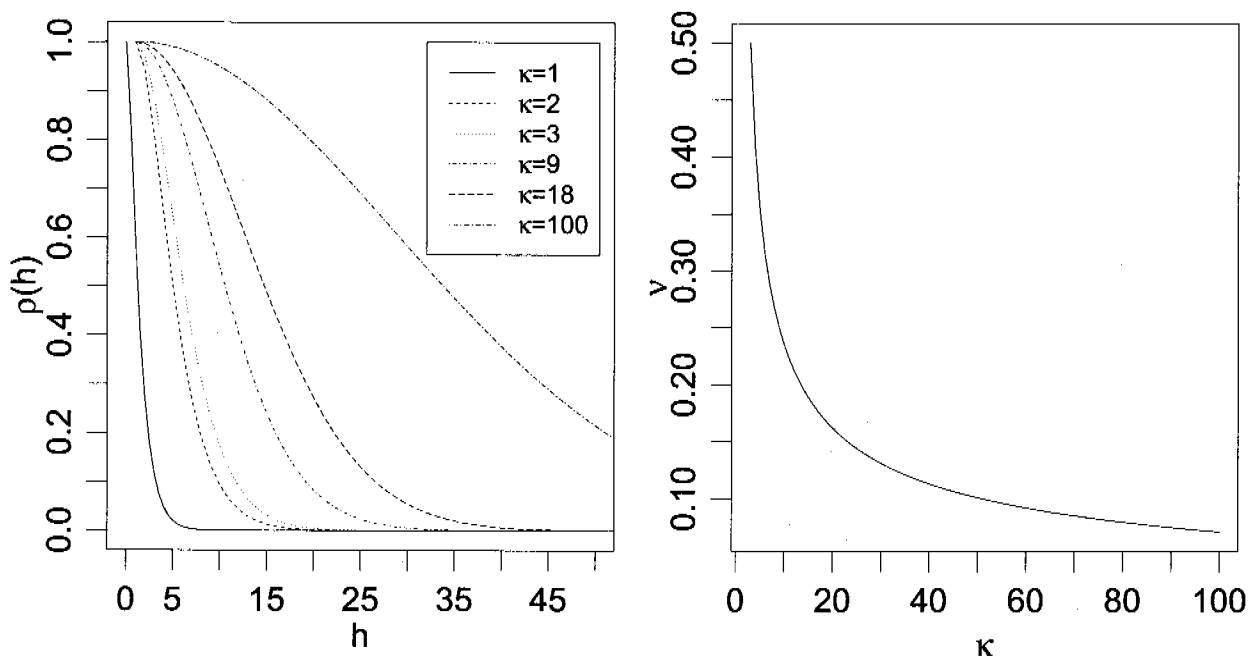
dependence due to a large  $\kappa$  is compensated by a smaller scale parameter  $\nu$  because according to equation 4.52 increasing  $\kappa$  rather quickly decreases  $\nu$ . The influence of  $\kappa$  on the Whittle-Matérn-covariance function and the relationship between  $\nu$  and  $\kappa$  is illustrated in figure 4.3 further below.

In figure 4.2 we display four more realizations of the Generalized Thresholded Gaussian Field Model with Matérn covariance, each of which has the same specific Minkowski functionals as the original sand cross-section in the left image of figure 4.1. The four realizations differ in the shape-parameter  $\kappa$ . The top left image corresponding to the smallest  $\kappa$ -value 3 has the most artifacts whereas the bottom right image with  $\kappa = 100$  has much less and also larger ones. But apparently apart from increasing boundary smoothness choosing  $\kappa$  larger does not change the optical appearance of the realizations substantially.



**Figure 4.2:** *Four realizations of the Generalized Thresholded Gaussian Model with Whittle-Matérn covariance structure and the same specific Minkowski functionals as the original  $800^2$  pixel-fine sand cross-section in the left image of figure 4.1. The shape-parameter  $\kappa$  which controls boundary smoothness was chosen to be  $\kappa = \{3, 9, 18, 100\}$  from left to right and top to bottom.*

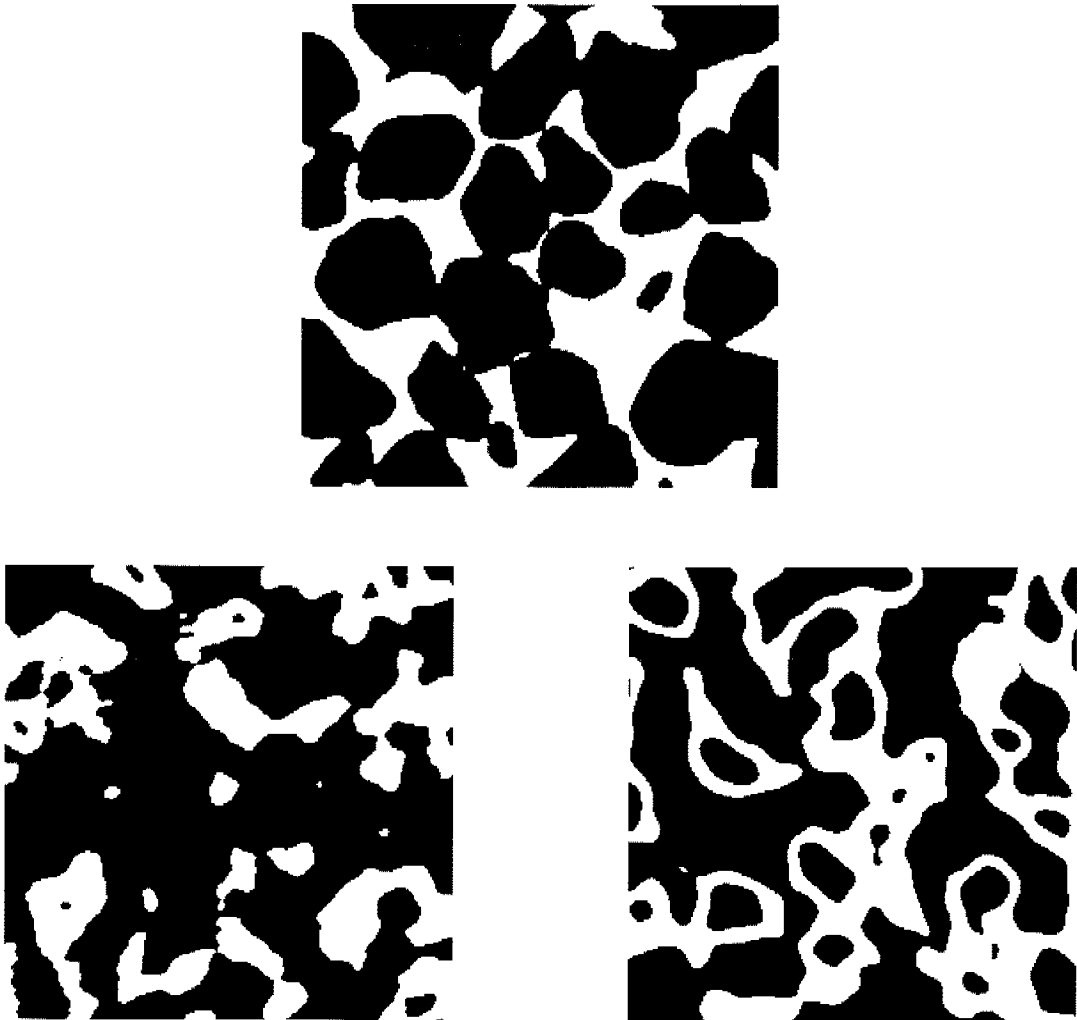
Figure 4.3 gives an impression how the Whittle-Matérn covariance function changes with increasing  $\kappa$ . For  $\kappa$  larger than 100 the precision in matching the Minkowski functionals of the original quickly decreases because the simulation of the covariance structure of the underlying Gaussian Random becomes imprecise. Figure 4.3 also shows how fast the scale  $\nu$  drops when  $\kappa$  is increased.



**Figure 4.3:** left image: The Whittle-Matérn-Covariance function for selected values of the shape-parameter  $\kappa$ . The scale-parameter  $\nu$  was held constant at 1. Increasing  $\kappa$  means increasing dependence. right image: In algorithm 4.7.1 the two parameters  $\nu$  and  $\kappa$  are dependent according to equation 4.52. Increasing  $\kappa$  implies quickly decreasing  $\nu$ . These two effects on the dependence structure partly compensate each other.

Finally figure 4.4 compares the original sand cross-section and the Generalized Thresholded Gaussian Random Field on the micro-scale. The leftmost image shows a  $256^2$ -pixel-sized extract from the  $800^2$ -pixel sized cross-section in the left image of figure 4.1 and two  $256^2$ -pixel realizations from the Generalized Thresholded Field with the same specific Minkowski functionals but different boundary smoothness.

Figure 4.4 illustrates that the Generalized Thresholded Gaussian Field Model, at least if only the Minkowski functionals are matched, does not produce grains of the right size and shape. The Gaussian grains are too lengthy or too clumped which is mainly because the thin connections between the real sand grains are not accurately modelled. They often are too broad and thus clusters of separate grains appear to be a single one. The figure also illustrates that on the micro-scale the effect of  $\kappa$  on the boundary smoothness is perceptible more clearly. The phase-boundaries of the left hand side realization ( $\kappa = 2$ ) seem to be much rougher than the rather smooth ones of the right hand side realization ( $\kappa = 7$ ).



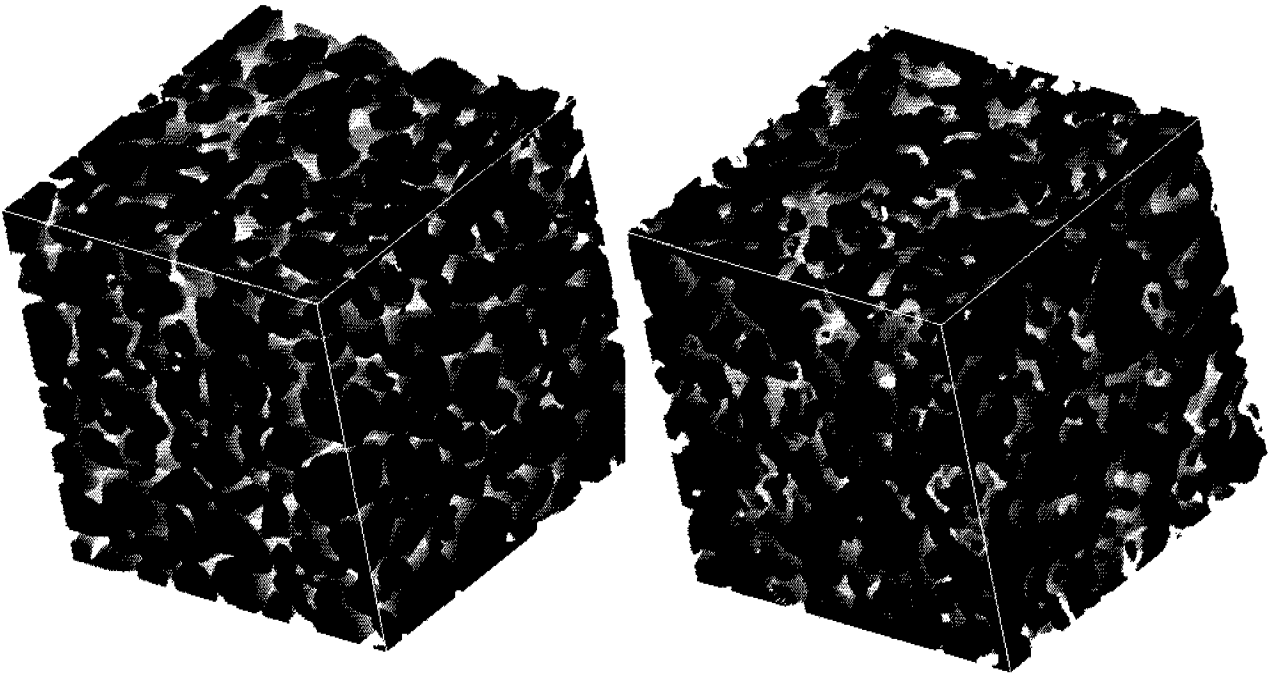
**Figure 4.4:** A  $256^2$ -pixel-sized extract from the  $800^2$  pixel fine-sand-cross section shown in the left image of figure 4.1 (top) and two realizations of the Generalized Thresholded Gaussian Model with Matérn covariance function and  $\kappa$  chosen to be 2 (left) and 7 (right).

## 4.9 3D-Results

In this short section we show a three-dimensional realization generated with the Generalized Thresholded Gaussian Field Model with Whittle-Matérn covariance function according to algorithm 4.7.3 which shares its specific Minkowski-functionals with the  $800^3$ -pixel sized fine-sand cube displayed in the right hand side image of figure 1.1. The realization is of size  $256^3$  pixels and the scale was chosen to be  $\nu = 256/800$  in order to achieve a low resolution image of a  $800^3$ -pixel sized Thresholded Gaussian-structure which can directly be compared to the original fine-sand specimen. Both the original and the artificial image are shown in figure 4.5.

The two images in figure 4.5 look quite similar although as already in the two-dimensional situation the Generalized Thresholded Gaussian Field Model produces small artifacts which are not present in the original structure and it has too broad inter-grain-connections.





**Figure 4.5:** An  $800^3$  pixel-sized extract from the HASYLAB fine-sand specimen which was already displayed as the right hand side image in figure 1.1 together with a  $256^3$  pixel-sized realization of the 3D-Generalized Thresholded Field Model with Whittle-Matérn covariance function generated according to algorithm 4.7.3. The two structures exhibit the same specific Minkowski-functionals.  $\kappa$  was chosen to be 25.

Please note that in the way algorithm 4.7.3 is formulated the two Matérn-parameters  $\nu$  and  $\kappa$  are independent. Hence  $\kappa$  can be raised to reduce the number of artifacts without transforming the model to a wrong scale. However our experiments have shown that the matching of specific Minkowski characteristics becomes imprecise if  $\kappa$  is chosen too high. We can conclude this section with the remark that for reconstruction purposes also in the Gaussian Model one should not only consider the specific Minkowski functionals but additionally include other geometrical characteristics (for example the ones of section 2.5) in the analysis. However these have been less intensively studied so far than for the Boolean Model and their distributional properties are often not explicitly known in the Thresholded Gaussian case.



## Chapter 5

# The Gibbsian Model

The last model we consider in this thesis is the so-called *Gibbsian Model*. Opposed to the Boolean and Gaussian Models considered in previous chapters (3 and 4) it is a *discrete* model that only makes sense on finite (or at least countable) lattices. Hence if the Gibbs Model is to be applied, it is a necessity that discretized binary images such as the one in figure 1.1 of the structure to be modelled are available. The idea of Gibbsian Fields and most of the corresponding terminology originates from Statistical Mechanics. There, Gibbsian Fields became popular descriptions for the equilibrium states of large finite particle systems. This is justified by the *variational principle* (see section 5.4 for details) which tells us that among all possible distributions of the states of a finite particle system exactly the Gibbs-distributions are the ones with the lowest free energy for a given potential. In the case of finite lattices, the Gibbsian assumption for the distribution of the lattice configurations is no restriction at all since every strictly positive probability distribution on the configurations of a finite lattice can be written in Gibbsian form. The main advantage of the Gibbsian viewpoint on a distribution is the close relationship between Gibbsian and Markov random fields provided by the *Hammersley-and-Clifford-Theorem* (see section 5.2) which will make the algorithms involving calculations of conditional distributions very easy and fast for Gibbsian fields. Among these, the *Gibbs-sampler* is the most prominent example of which we will present the adapted version to our problem in section 5.6. The restriction of the lattice  $\Gamma$  being finite can be relaxed and Gibbsian-distribution on countable sets can be defined. In the case of  $\Gamma = \mathbb{Z}^d$ , a Gibbsian distribution can be defined by a so-called *thermodynamic limit-procedure* which basically means to consider the Gibbsian-distributions on an arbitrary series of finite lattices  $\Gamma_n$  increasing to  $\mathbb{Z}^d$  and taking the limit of these distributions. It can be shown that this procedure is well-defined and that the limit does not depend on the series  $\Gamma_n$  that was chosen. An other approach to define Gibbsian Random Fields on countable sets uses the idea of *coherent conditional kernels*. For both these methods and properties of Gibbs fields on countable sets an excellent reference is the book of Guyon (Guyon 1995). The most famous example for a Gibbsian Model on  $\mathbb{Z}^d$  is the so-called *Ising-Model* which has been studied extensively in literature. It is one of the very few models so far for which both necessary and sufficient conditions for the uniqueness of the Gibbs-distribution (and thus the absence of *phase transitions*) are known. In general only sufficient conditions for the uniqueness can be given. Another possibility to extend the finite Gibbsian Model would be to let the lattice spacing  $h$  shrink to zero. This type of limit procedure is not very well known so far. We will mainly restrict ourselves to the case of finite Gibbsian fields with binary state space in the following. For Gibbsian fields on finite lattices with finite state space the theory is relatively simple and a quick overview can be gained within the book by

(Winkler 1995) which we use as a guideline to develop the theory in the following. To simulate a Gibbsian Field with predetermined values for the Minkowski functionals many different techniques are available, we have tried several of them for the simulation of random porous media with quite dissimilar success. Of these techniques we will describe the *Pseudo-likelihood-Method* (subsection 5.7.1), the *Sequential Newton Method* (subsection 5.7.2) and the *Simulated Annealing-Method* (subsection 5.8) which has turned out to be the most promising. Because the algorithms involved to simulate from Gibbsian Fields are *equilibrium algorithms* such as Markov-Chain-Monte Carlo methods, they are relatively slow. This and the fact that the 2D-simulation results obtained with Gibbsian Fields did not quite match our expectations prevented us from applying the Gibbsian theory to 3D-porous media so far and thus the results presented at the end of this chapter in section 5.9 unfortunately are only two-dimensional.

## 5.1 Definition and Properties

We start out with the definition of the  $d$ -dimensional binary Gibbsian Model on a finite lattice. For this purpose let  $\Gamma$  be a finite (cubic) lattice,  $\Gamma \subset h \cdot \mathbb{Z}^d$ . The space of **binary image configurations**  $\mathcal{S}$  can then be defined as  $\mathcal{S} = \prod_{x \in \Gamma} \{0, 1\}$ . The Gibbsian Model is now given by a probability distribution on  $\mathcal{S}$ . The **Gibbsian Field** is a random set  $\Theta$  taking values in  $\mathcal{S}$ . Note that  $\mathcal{S} \subset \mathcal{F}$  (and even  $\mathcal{R}$ ). We define  $\Theta$  by giving its probability distribution  $\pi_G$  on  $\mathcal{S}$ . It is the so-called *Gibbsian distribution*:

**Definition 5.1.1.** (*Gibbsian distribution*) *The Gibbsian distribution on  $\mathcal{S}$  is:*

$$\pi_G(s) = \frac{\text{Exp}(-H(s))}{\sum_{t \in \mathcal{S}} \text{Exp}(-H(t))}, \quad (5.1)$$

where  $H(\cdot) : \mathcal{S} \mapsto \mathbb{R}$  is the *energy function*.

Furthermore we denote by  $\Theta(x), x \in \Gamma$  the random variables induced by  $\Theta$  giving the random pixel value at a lattice position  $x$ . Similarly for  $s \in \mathcal{S}$  we denote  $s(x)$  the pixel value of  $s$  in  $x \in \Gamma$ . Please note that for the denominator  $Z := \sum_{t \in \mathcal{S}} \text{Exp}(-H(t))$  on the right hand side of definition 5.1.1 to make sense, the state space of the  $\Theta(x)$  must be finite otherwise integrability conditions for  $\text{Exp}(-H(t))$  with respect to a reference measure would be needed. Also note that every strictly positive probability distribution  $\pi$  on  $\mathcal{S}$  can be written in Gibbsian form by using the energy function  $H_\pi(s) := -\log \pi(s)$  leading to  $Z = 1$ . Hence a specific distribution is specified only if the potential  $H(x)$  is provided. On the other hand, as is obvious from equation 5.1, a Gibbsian distribution  $\pi_G$  defines the corresponding potential function  $H$  only up to an additive constant.

Of special interest are those energy functions  $H$  which are driven by a potential  $U$ :

**Definition 5.1.2.** (*Potential and its Energy*)

*i). A **potential**  $U$  is a collection of (real-valued) functions  $\{U_A : A \subset \Gamma\}$  defined on  $\mathcal{S}$  such that*

*a).  $U_\emptyset \equiv 0$ ,*

*b).  $U_A(s) = U_A(t)$  if  $s$  and  $t$  agree on  $A$ .*

ii). The energy function  $H$  of a potential  $U$  is defined as:

$$H_U(s) := \sum_{A \subset \Gamma} U_A(s).$$

The potential  $U$  and its corresponding energy function  $H_U$  are equivalent in the sense that the potential can be recalculated from the energy-function by Moebius-inversion (see equations 7.28, 7.30). Among the potentials we are especially interested in those which are constantly equal to zero for all subsets  $A$  but the cliques of a neighbourhood system on  $\Gamma$ :

**Definition 5.1.3. (Neighbourhood-Systems and Cliques on  $\Gamma$ )**

- i). A collection  $\delta = \{\delta(x), x \in \Gamma\}$  of subsets of  $\Gamma$  is called a **neighbourhood-system** on  $\Gamma$  if:
- a.)  $x \notin \delta(x)$ ,
  - b.)  $y \in \delta(x) \Rightarrow x \in \delta(y)$ .
- ii). The **Cliques of a neighbourhood-system** on  $\Gamma$  are the subsets  $C \subset \Gamma$  such that every two distinct points  $x$  and  $y$  ( $x \neq y$ ) are neighbours of each other, ie.  $y \in \delta(x) \forall x \neq y \in C$ .

It is evident that each neighbourhood system can be represented by a graph with the lattice points  $x \in \Gamma$  as vertices which are joined by edges if and only if they are neighbours. The cliques  $C$  correspond to complete subgraphs. Conversely each graph and hence each of the adjacency systems discussed in this thesis so far induces a neighbourhood system in a natural way. As an example the ordinary cubic lattice graph in 3D induces the 6-neighbourhood in which each lattice point  $x$  is surrounded by the 6 neighbours to which it is joined by the six edges that meet in  $x$  and are parallel to the coordinate axes. The non-trivial cliques consist of vertex pairs which delimit a lattice edge. For the three dimensional 26-neighbourhood on a cubic lattice (which means including face and spatial diagonals as edges) that we will need in the following, the largest cliques are the vertices of the single lattice-cell-cubes. This neighbourhood will be interesting for us because all the vertices, edges, faces and volumes of the adjacencies considered in this thesis and all the subsets needed to calculate the different OM-estimators for the specific Minkowski-functionals are cliques with respect to the 26-neighbourhood.

**Definition 5.1.4. (Neighbourhood potentials)** Let a neighbourhood-system  $\delta$  on  $\Gamma$  be given. A **neighbourhood potential** for  $\delta$  on  $\mathcal{S}$  is a potential  $U$  such that  $U_A \equiv 0, \forall A$  which are not cliques.

We will call Gibbs-Fields with a Neighbourhood-Potential *Neighbourhood-Gibbs-Fields* in the following. It will turn out that the potentials we will need in the following are in fact neighbourhood-potentials with respect to the 26-neighbourhood on the regular finite cubic lattice  $\Gamma$ . As will be explored in the next section, the main advantage of Gibbs-models  $\Theta$  with neighbourhood potentials is that the *conditional distributions*  $P[\Theta(x) = s_x | \cap_{\{y \in \Gamma, y \neq x\}} \Theta(y) = s_y]$  for  $s_x, s_y \in \{0, 1\}$  can be easily determined without having to calculate the sum in the denominator  $Z$  of equation 5.1. This is crucial for the algorithms in later sections to attain the desired results in reasonable time. Note that already in the binary case the number of summands in  $Z$  grows exponentially, ie. if  $|\Gamma| = n$ , we have  $|\mathcal{S}| = 2^n$ .

The next section is devoted to the *Hammersley and Clifford-Theorem* which establishes the connection between Gibbs Models with neighbourhood potentials and Markovian Random Fields and thus justifies why such Gibbs Models can quite easily be simulated.

## 5.2 The Hammersley and Clifford-Theorem

In this section the connection between neighbourhood-Gibbs-Fields and Markovian Random Fields is established. For this purpose, we quickly recall the notion of Markov Random Fields in our setting:

**Definition 5.2.1. (Markovian Random Fields)** *A random field  $\Theta$  is called a **Markov Random Field** on  $\Gamma$  with respect to a neighbourhood-system  $\delta$  if  $\forall s \in \mathcal{S}$  and  $\forall x \in \Gamma$  the following holds:*

$$P[\Theta(x) = s_x | \Theta(y) = s_y, y \neq x] = P[\Theta(x) = s_x | \Theta(y) = s_y, y \in \delta(x)]. \quad (5.2)$$

For simulation algorithms involving the successive update of single pixels (or regions) of  $\Gamma$  conditioned on the rest of  $\Gamma$ , the Markov property of equation 5.2 is most welcome. The following well-known theorem by Hammersley and Clifford shows that the Neighbourhood-Gibbs-Fields and Markovian-Fields are in fact the same class.

**Theorem 5.2.2. (Hammersley and Clifford)** *Let  $\delta$  be a neighbourhood-system.*

- i). A random field  $\Theta$  is a Markov Field if and only if it is a Neighbourhood-Gibbs-Field.*
- ii). For a Neighbourhood-Gibbs-Field  $\Theta$  and an arbitrary subset  $A \subset \Gamma$ :*

$$P[\Theta(x) = s_x, x \in A | \Theta(y) = s_y, y \in A^c] = P[\Theta(x) = s_x, x \in A | \Theta(y) = s_y, y \in \delta A],$$

where  $\delta A := \cup_{x \in A} \delta(x) \setminus A$ .

**Proof. (Theorem 5.2.2)**

We show part ii). first.

**ad ii):** We can write:

$$\pi_G(s) = \frac{\text{Exp}(-\sum_{C \subset \Gamma} U_C(s))}{\sum_{t \in \mathcal{S}} \text{Exp}(-\sum_{C \subset \Gamma} U_C(t))},$$

where the sum in the exponential runs through all the cliques  $C \subset \Gamma$  with respect to  $\delta$ . For  $A \subset \Gamma$  we define the following partition of the set of cliques  $\mathcal{C}$ :

$$\mathcal{C}_A := \{C \text{ Clique of } \Gamma : C \cap A \neq \emptyset\}, \quad \mathcal{C}_{A^c} := \{C \text{ Clique of } \Gamma : C \cap A = \emptyset\},$$

then we have for  $s$  with  $s(x) = s_x, x \in A$  and  $s(y) = s_y, y \in A^c$ :

$$\begin{aligned}
\mathbb{P}[\Theta(x) = s_x, x \in A | \Theta(y) = s_y, y \in A^c] &= \frac{\text{Exp}\left(-\sum_{C \in \mathcal{C}} U_C(s)\right)}{\sum_{t \in \mathcal{S}, t(y)=s_y, y \in A^c} \text{Exp}\left(-\sum_{C \in \mathcal{C}} U_C(t)\right)} \\
&= \frac{\text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(s)\right) \cdot \text{Exp}\left(-\sum_{C \in \mathcal{C}_{A^c}} U_C(s)\right)}{\sum_{t \in \mathcal{S}, t(y)=s_y, y \in A^c} \text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(t)\right) \cdot \text{Exp}\left(-\sum_{C \in \mathcal{C}_{A^c}} U_C(t)\right)} \\
&= \frac{\text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(s)\right)}{\sum_{t \in \mathcal{S}, t(y)=s_y, y \in A^c} \text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(t)\right)},
\end{aligned}$$

where we have used that for  $C \in \mathcal{C}_{A^c}$  and  $t \in \mathcal{S} : t(y) = s_y, y \in A^c$  we have  $U_C(t) = U_C(s)$  by definition 5.1.2, part i)b).

If we now exploit that for  $C \in \mathcal{C}_A$  we have  $C \subset A \cup \delta A$ , we can conclude:

$$\begin{aligned}
\mathbb{P}[\Theta(x) = s_x, x \in A | \Theta(y) = s_y, y \in A^c] &= \\
&= \frac{\text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(s)\right) \cdot \sum_{t \in \mathcal{S}: t(x)=s_x, x \in A \cup \delta A} \text{Exp}\left(-\sum_{C \in \mathcal{C}_{A^c}} U_C(t)\right)}{\sum_{t \in \mathcal{S}, t(y)=s_y, y \in A^c} \text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(t)\right) \cdot \sum_{t \in \mathcal{S}: t(x)=s_x, x \in A \cup \delta A} \text{Exp}\left(-\sum_{C \in \mathcal{C}_{A^c}} U_C(t)\right)} \\
&= \frac{\sum_{t \in \mathcal{S}: t(x)=s_x, x \in A \cup \delta A} \text{Exp}\left(-\sum_{C \in \mathcal{C}} U_C(t)\right)}{\sum_{t \in \mathcal{S}, t(y)=s_y, y \in A^c} \cdot \sum_{l \in \mathcal{S}: l(x)=s_x, x \in A \cup \delta A} \text{Exp}\left(-\sum_{C \in \mathcal{C}_A} U_C(t)\right) \cdot \text{Exp}\left(-\sum_{C \in \mathcal{C}_{A^c}} U_C(l)\right)} \\
&= \frac{\sum_{t \in \mathcal{S}: t(x)=s_x, x \in A \cup \delta A} \text{Exp}\left(-\sum_{C \in \mathcal{C}} U_C(t)\right)}{\sum_{l \in \mathcal{S}: t(x)=s_x, x \in \delta A} \text{Exp}\left(-\sum_{C \in \mathcal{C}} U_C(t)\right)} = \mathbb{P}[\Theta(x) = s_x, x \in A | \Theta(y) = s_y, y \in \delta A],
\end{aligned}$$

which was to show.

**ad i):** The fact that Neighbourhood-Gibbs-fields are Markov fields follows immediately from ii). by specializing  $A = \{x\}$  for arbitrary  $x \in \Gamma$ . It remains to show that Markov Random Fields must be Gibbs-Fields with a neighbourhood-potential.

The converse direction of i). is not directly relevant for us in the following and can thus be found in the appendix (subsection 7.4.1).  $\square$

From the proof it becomes clear that the theorem holds for arbitrary finite state spaces and that a restriction to binary fields is not necessary. It should also be mentioned here that the Hammersley-and-Clifford-Theorem holds in a more general setting where the state space is allowed to be discrete and infinite. For details please refer to (Guyon 1995, Theorem 2.2.1).

With the Hammersley-Clifford-Theorem at hand it becomes clear why Gibbsian Fields can easily be simulated by MCMC-type algorithms such as the *Gibbs-Sampler* (see section 5.6). For example in 2 dimensions with the usual 4-neighbourhood induced by the edges of a square lattice  $\Gamma$ , the conditional distribution needed to update a single pixel of a Gibbs field at a particular lattice point  $x \in \Gamma$  conditioned on the pixel values of all the other

lattice points thanks to the Markovian structure only depends on the 4 neighbouring pixel values regardless of the size of the finite image lattice. With image size the number of possible (binary) configuration grows exponentially and to implement a Gibbs-sampler on large images ( $\approx 10^9$  pixels) where each pixel (or block of pixels) has to be updated a myriad of times ( $\approx 10^5$  runs) until the equilibrium has been reached, the Markov property is absolutely crucial.

### 5.3 Choice of the Gibbs-Potential

We will now start adapting the Gibbsian Model developed in the former sections to our problem of simulating porous binary structures with predetermined specific Minkowski-values. The important question to answer here is how we should choose the energy function  $H$  for our Gibbs-Model. The model has to distinguish configurations in  $\mathcal{S}$  by means of their Minkowski-functionals  $V_j(s), j \in \{0, \dots, d\}$  and configurations with the same Minkowski-functionals should be assigned the same probability under the Gibbs-Model. Hence the natural choice for the energy function  $H$  for our purposes is a linear combination of the Ohser-Mücklich estimators  $\hat{V}_j(s)$  as defined in subsections 2.3.1, 2.3.2:

**Definition 5.3.1. (OM-energy)** Let  $\hat{V}_j(s), j \in \{0, \dots, d\}$  be the OM-estimators for the Minkowski-functionals  $V_j(s)$  of the binary configurations  $s \in \mathcal{S}$ . Then the energy function

$$H_{OM}(s) := \sum_{j=0}^d \beta_j \cdot \hat{V}_j(s),$$

where the  $\beta_j$  are real coefficients, is referred to as the **OM-energy** in the following.

Please note that we work here with the OM-estimators  $\hat{V}_j$  for the *unspecific Minkowski functionals*  $V_j$ . The field volume  $h^d \cdot |\Gamma|$  enters the formula for  $H_{OM}$  only implicitly as a contribution to the coefficients  $\beta_j$ . Assume that the grid  $\Gamma$  consists of  $n$  voxels with grid constant  $h$ . The parameters  $\tilde{\beta}_j$  for a differently sized extract of size  $\tilde{n}$  voxels with grid resolution  $\tilde{h}$  can then be calculated by the following rescaling formula:

$$\frac{\beta_j}{\tilde{\beta}_j} = \frac{\tilde{n} \cdot \tilde{h}^d}{n \cdot h^d}.$$

For the Gibbs Model with energy function  $H_{OM}$  every choice of a parameter set  $\beta_j, j \in \{0, \dots, d\}$  defines a random set  $\Theta$  on  $\mathcal{S}$  with different Minkowski functionals per unit volume,  $\bar{V}_j(\Theta)$ . One would like to choose the parameters  $\beta_j$  such that the basic modelling equation 1.1 is fulfilled, ie.,

$$v_j(s_0) \stackrel{!}{=} \bar{V}_j(\Theta), \quad j \in \{0, \dots, d\}, \quad (5.3)$$

where the  $v_j(s_0)$  are the predetermined values for the Minkowski functionals per unit volume which possibly stem from some digitized reference image  $s_0 \in \mathcal{S}$ . Note that it doesn't make sense here to consider the specific Minkowski-functionals as defined in 2.4.3 since the realizations of the Gibbsian Model  $\Theta$  are restricted to a finite lattice and therefore not stationary in  $\mathbb{R}^d$ . Therefore  $\bar{V}_j(\Theta)$  simply means here the expected Minkowski-functional  $V_j$  per unit volume. Unfortunately the specific Minkowski functionals for the Gibbsian Model are unknown analytically and thus instead of solving equation 5.3 one has to resort to considering:



$$v_j(s_0) \stackrel{!}{=} \frac{\mathbf{E} [\hat{V}_j(\Theta)]}{|W|}, \quad j \in \{0, \dots, d\}. \quad (5.4)$$

Solving equation 5.4 for a parameter set  $\beta$  corresponds to finding the *Maximum-Likelihood-estimator*  $\hat{\beta}^{MLE}$ . In fact the solution of the optimization problem

$$\hat{\beta}^{MLE} = \max_{\beta} \frac{\text{Exp} \left( - \sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_0) \right)}{\sum_{s \in \mathcal{S}} \text{Exp} \left( - \sum_{j=0}^d \beta_j \cdot \hat{V}_j(s) \right)} \quad (5.5)$$

is given by the following equation-system for  $j_0 \in \{0, \dots, d\}$ :

$$0 \stackrel{!}{=} \sum_{s \in \mathcal{S}} \text{Exp} \left( - \sum_{j=0}^d \beta_j^{MLE} \cdot \hat{V}_j(s) \right) \cdot (-\hat{V}_{j_0}(s_0)) + \sum_{s \in \mathcal{S}} \hat{V}_{j_0}(s) \cdot \text{Exp} \left( - \sum_{j=0}^d \beta_j^{MLE} \cdot \hat{V}_j(s) \right)$$

which is equivalent to equation 5.4 after dividing by the field's volume.

As can easily be seen from the definitions of the OM-estimators in subsections 2.3.1, 2.3.2 the energy-function  $H_{OM}$  corresponds to a neighbourhood-potential for the 8-neighbourhood in 2D and the 26-neighbourhood in 3D on the squared (cubic) lattice respectively and hence the Hammersley-and-Clifford-Theorem (Theorem 5.2.2) may be applied and the Markovian structure with respect to these neighbourhoods may be exploited. In fact the estimators boil down to counting vertices, edges (lattice cell edges, face diagonals and spatial diagonals, and planes (lattice cell faces, diagonal faces, triangular faces) within certain selected pixel configurations. All these are cliques with respect to the 8-(26-)neighbourhood and hence all the information needed to calculate the estimators can be calculated by considering the cliques which are subsets of lattice squares (or lattice cubes). It will turn out in section 5.7 that not only simulation of the Gibbsian Field but also determination of the parameters  $\beta_j$  becomes much easier because of the Markov property.

## 5.4 The Variational Principle

In this section and the next one we can give two justifications why the Gibbs Model is a natural choice when it comes to modelling porous media. The *variational principle* discussed in this section is in fact a mathematical formulation of the *maximum entropy principle* which is one of the most fundamental principles at the very heart of physics in general and especially statistical mechanics where it is usually referred to as the *second law of thermodynamics*. In simplified terms, the maximum entropy-principle means the fact that any self-supporting process in any isolated system of free particles increases the physical entropy  $S$  of the system and thus the system spontaneously organizes itself in such a way that the physical entropy is maximized and equivalently the amount of free energy available in the system is minimized. In stochastic terms the maximum-entropy-principle can be formulated as the so-called *variational principle*. This is the core statement of this section and means the following: For a random physical particle system with a certain constant amount of intrinsic energy  $U$ , the distributional law which maximizes entropy and thus according to which matter organizes itself is exactly the Gibbsian distribution for this energy function. This crucial result explains the importance of the Gibbsian distribution

for statistical physics and why it originated from there. If we understand the natural formation process of porous media as a random system and if we agree upon the fact that nature respects the maximum entropy principle during this formation process, then the distribution of the porous structures that can be observed in nature should follow a Gibbsian law and thus the Gibbsian Model should perform well when it comes to modelling such structures. It is important to know that the variational principle as a mathematical statement only holds for systems on the *infinite* lattice  $h \cdot \mathbb{Z}^d$  and thus only *after* passing to the *thermodynamic limit* with the necessarily finite lattices  $\Gamma$  we use in practice. In order to give a more precise formulation of the variational principle, we need to introduce some basic definitions of entropy in probability theory in the following. We will not recall here the basic notions of energy and physical entropy of thermodynamics since they are discussed in every basic textbook about statistical physics, but we will make use of the following fundamental relationship:

**Definition 5.4.1. (Free energy)** *The Free energy  $F$  of a physical system is*

$$F = U - T \cdot S,$$

where  $U$  is the *intrinsic energy* of the system,  $S$  is its (thermodynamic) **entropy** and  $T$  is the temperature.

An entropy concept also exists in probability. For two arbitrary finite measures  $\mu$  and  $\nu$  we can define their *relative entropy* (or equivalently Kullback-Leibler-divergence) in terms of the following non-symmetrical "distance":

**Definition 5.4.2. (Relative entropy)** *Let  $\mu$  and  $\nu$  be two finite measures on  $\mathbb{Z}^d$  and denote by  $\mu_\Gamma$  and  $\nu_\Gamma$  their marginals on a finite subset  $\Gamma \subset \mathbb{Z}^d$ . Then the **relative entropy**  $\mathcal{E}_\Gamma(\mu|\nu)$  of  $\mu$  with respect to  $\nu$  on  $\Gamma$  is defined as:*

$$\mathcal{E}_\Gamma(\mu|\nu) = \begin{cases} -E_{\mu_\Gamma} \left[ \log \frac{\partial \mu_\Gamma}{\partial \nu_\Gamma} \right] & \text{if } \mu_\Gamma \ll \nu_\Gamma, \\ -\infty & \text{otherwise.} \end{cases}$$

The concept of Kullback-Leibler divergence is of course by no means restricted to  $\mathbb{Z}^d$  and can be introduced for general probability spaces. To an arbitrary stationary probability measure  $\mu$  on  $\mathbb{Z}^d$ , we can now assign its *specific entropy* by a thermodynamic limit procedure:

**Definition 5.4.3. (Specific entropy)** *Let  $\Gamma_n$  be an increasing sequence of lattices with  $\lim_{n \rightarrow \infty} \Gamma_n = \mathbb{Z}^d$  and let  $\mu$  be a stationary probability measure on  $\mathbb{Z}^d$  with marginals  $\mu_{\Gamma_n}$ . Then the **specific entropy**  $\mathcal{E}(\mu)$  of  $\mu$  is defined as the thermodynamic limit of the relative entropy of  $\mu$  with respect to the counting measure  $\lambda$  on the sequence  $\Gamma_n$ :*

$$\mathcal{E}(\mu) = \lim_{n \rightarrow \infty} \mathcal{E}_{\Gamma_n}(\mu|\lambda) = - \lim_{n \rightarrow \infty} E_{\mu_n}(\log \mu_{\Gamma_n}).$$

With these definitions at hand, we can now formulate the variational principle:

**Theorem 5.4.4. (Variational Principle)** *Let  $\mathcal{P}(\mathbb{Z}^d)$  be the set of all stationary probability measures on the set of binary configurations over  $\mathbb{Z}^d$  and denote by  $H : \mathcal{S} \rightarrow \mathbb{R}$  an energy function and  $\mu_G$  the corresponding Gibbs measure according to definition 5.1.1 after passing to the thermodynamic limit. Then:*

$$\mu_G = \arg \min_{\mu \in \mathcal{P}(\mathbb{Z}^d)} [E_\mu[H] - \mathcal{E}(\mu)]. \quad (5.6)$$

**Proof. (Theorem 5.4.4)** The variational principle is an immediate consequence of (Guyon 1995, Theorem 2.1.4, p.54).

In thermodynamics the energy function  $H(s)$  corresponds to the physical energy of the state  $s$  scaled by inverse temperature  $T$  (and eventually the Boltzmann-factor  $k_B$ ). The quantity  $E_\mu[H]$  then plays the role of  $U/T$  and if we treat the temperature as constant equation 5.6 becomes:

$$\mu_G = \arg \min_{\mu \in \mathcal{P}(\mathbb{Z}^d)} F(\mu),$$

hence the Gibbs measure is in fact the distribution on  $\mathbb{Z}^d$  which minimizes the amount of free energy. In our case the energy function  $H$  is the OM-energy  $H_{OM}$  according to definition 5.3.1 with the  $\beta_j$  chosen such that the basic modelling equation 5.4 is fulfilled. Therefore we conclude that among all distributions in  $\mathcal{P}(\mathbb{Z}^d)$  which lead to the same predetermined values for the specific Minkowski-functionals  $v_j(s_0)$ , the Gibbsian distribution  $\pi_G$  with energy function  $H_{OM}$  is the one which maximizes entropy (or equivalently minimizes free energy). One should bear in mind that the variational principle gives a justification in favour of the Gibbs Model on  $\mathbb{Z}^d$ , and thus only *after* the thermodynamic limit procedure, however it is almost trivial to see that on a finite grid  $\Gamma$ , the Gibbsian distribution maximizes entropy as well. For later reference we provide the argument here. On a finite grid  $\Gamma$  the maximum entropy problem reduces to the following optimization problem with side-constraints:

$$\begin{aligned} \max_{p \in \mathcal{P}(\mathcal{S})} \left\{ - \sum_{s \in \mathcal{S}} p(s) \cdot \log p(s) \right\}, \quad (5.7) \\ \text{subject to:} \\ \sum_{s \in \mathcal{S}} V_j(s) \cdot p(s) = V_j(s_0) \text{ for } j \in \{0, \dots, d\} \\ \sum_{s \in \mathcal{S}} p(s) = 1. \end{aligned}$$

where  $V_j(s_0)$  are the predetermined unspecific Minkowski values. This can easily be solved by introducing the Lagrange-multiplicators  $\lambda_0, \dots, \lambda_d$  for the first set of constraints and  $\lambda$  for the last. We then must have:

$$-\log p(s) - 1 \stackrel{!}{=} \sum_{j=0}^d \lambda_j \cdot V_j(s) + \lambda,$$

and thus:

$$p(s) = \text{Exp} \left( - \sum_{j=0}^d \lambda_j \cdot V_j(s) - \lambda - 1 \right),$$

where  $\lambda$  can be determined by summing up the last equation over  $s \in \mathcal{S}$  and using the constraint that  $p$  must be a probability distribution on  $\Gamma$ . We find:

$$p(s) = \frac{\text{Exp} \left( - \sum_{j=0}^d \lambda_j \cdot V_j(s) \right)}{\sum_{t \in \mathcal{S}} \text{Exp} \left( - \sum_{j=0}^d \lambda_j \cdot V_j(t) \right)},$$

which shows that in fact  $p = \pi_G$  for the OM-energy function and with  $\beta_j = \lambda_j, j \in \{0, \dots, d\}$ . The only drawback here is that the Minkowski-functionals  $V_j(s)$  cannot be determined exactly and thus the OM-energy function uses their estimators  $\hat{V}_j(s)$  instead. In the next section we will see that on a finite grid  $\Gamma$  among all probability distributions concentrated on binary images with the same Minkowski functionals, entropy is maximized for the uniform distribution. The maximum entropy principle will enable us to establish a connection between the Gibbsian distribution and the uniform distribution on the set of images  $s \in \mathcal{S}$  which share the same Minkowski-functionals.

## 5.5 Julész-Ensembles

In this section we discuss the fact that the Gibbsian distribution with the OM-energy function converges in the thermodynamic limit to the uniform distribution on the set of images with the same specific Minkowski functionals. Such a set will be called a *Julész-ensemble*, named after the famous Hungarian experimental psychologist who revolutionized the theory of visual perception with his studies about ensembles of neurons. The convergence of the neighbourhood-Gibbs distribution to the uniform distribution on the Julész-ensemble still holds if we replace the OM-energy with a finite linear combination of arbitrary image statistics as long as they arise from a neighbourhood potential and thus can be calculated as spatial averages of local image features. These image features  $\phi_j(\cdot) : \mathcal{S} \rightarrow \mathbb{R}, j \in \{1, \dots, n\}$  in our case must be computable from the  $2^{2^d}$  specific binary vertex combinations  $i$  of the  $2^d$  vertices of the unit cell of the cubic lattice  $\Gamma$ . Similarly as for the OM-estimators for the Minkowski-functionals in Chapter 2 such an image statistic is computable from the *summary statistics* (or *configuration vector*)  $h$  whose components  $h(i), i \in \{0, \dots, 2^{2^d} - 1\}$  indicate how often each of the vertex configurations can be found among the voxels of the lattice  $\Gamma$ . For simplicity we assume for the rest of this section that all considered lattices  $\Gamma$  have the same unit lattice spacing. It is often more convenient to work with the *specific configuration vector*  $\bar{h}(s) := h(s)/|\Gamma|$  which indicates the relative frequencies of the  $2^{2^d}$  vertex configurations in an image  $\mathcal{S}$ . The summary statistics  $\phi(\cdot)$  we consider in this section can then be obtained by multiplying the configuration vector  $h$  with a  $2^{2^d}$ -*contribution vector*  $c$ , whose components  $c_i$  give the contributions of a specific vertex configuration  $i$  to the statistic  $\phi(\cdot)$ . Analogously the specific summary statistic  $\bar{\phi}$  is obtained by multiplying  $c$  with the specific configuration vector  $\bar{h}$ :

$$\phi(s) := \sum_{i=0}^{2^{2^d}-1} c_i \cdot h(i), \quad \bar{\phi}(s) := \sum_{i=0}^{2^{2^d}-1} c_i \cdot \bar{h}(i) \quad (5.8)$$

Furthermore the Gibbsian energy functions  $H$  we consider here are of the following form:

$$H^\phi(s) := \sum_{j=1}^n \beta_j \cdot \phi_j(s), \quad (5.9)$$

where the  $\beta_j$  are chosen to fulfill the basic modelling-equation  $E_{\pi_{G,n}}[\bar{\phi}_j(\Theta)] = v_j$  and the  $v_j$  are preset values for the specific feature statistics. It is easy to see that the OM-estimators  $\hat{V}_j(\cdot)$  of subsections 2.3.1 and 2.3.2 constituting the OM-energy defined in 5.3.1 are in fact of this type.

It is even possible to establish an *equivalence* between the thermodynamic limits of the Gibbs-distribution and the uniform distribution on the Julész-ensemble. An introductory

discussion of this equivalence and its derivation can be found in a series of computer science papers by Wu, Zhu and Liu: See (Wu, Zhu and Liu 1999, Wu, Zhu and Liu 2000) for the theoretical aspects and (Zhu, Liu and Wu 2000) concerning simulation.

Let us now describe the *ensemble-equivalence* in mathematical terms. We consider an increasing sequence of finite lattices  $\Gamma_n$  with  $\lim_{n \rightarrow \infty} \Gamma_n = \mathbb{Z}^d$  and denote the corresponding configuration spaces with  $\mathcal{S}_n := \{0, 1\}^{|\Gamma_n|}$ . The uniform distribution on  $\mathcal{S}_n$  we denote by  $U_{\Gamma_n}$ . We are interested in those configurations  $s \in \mathcal{S}_n$  which share the same specific Minkowski-functionals  $\bar{V}_j(\cdot)$  or more generally the same values  $v_j$  for the specific feature statistics  $\bar{\phi}_j(\cdot)$ . We write  $v := (v_j)_{j \in \{1, \dots, n\}}$  for the vector of target feature statistics and introduce the notation

$$\Omega_{\Gamma_n}(v) = \{s \in \mathcal{S}_n : \bar{\phi}_j(s) = v_j, \forall j \in \{1, \dots, n\}\}$$

to denote the set of all configurations over  $\Gamma_n$  which share the same values for the specific feature statistics. Similarly  $\Omega_A(v)$  for a sub-lattice  $A \subset \Gamma_n$  denotes all configurations over the sub-lattice  $A$  whose values for the specific images statistics  $\bar{\phi}$  are  $v_j$ . Note that because the specific configuration vector  $\bar{h}$  can only attain discretely many values, it is well possible that the set  $\Omega_{\Gamma_n}(v)$  is empty for a specific value of  $v$ . This minor complication can be handled by considering a small set  $\mathcal{V}$  containing  $v$  and its nearest attainable discrete approximation and arguing for  $\mathcal{V}$  instead of  $v$ . However for the sake of simplicity we will ignore this problem in the following and directly work with  $v$  instead. We write  $U_{n,v}$  for the uniform distribution on  $\Omega_{\Gamma_n}(v)$ , ie:

$$U_{n,v}(s) := \begin{cases} 1/|\Omega_{\Gamma_n}(v)| & \text{if } s \in \Omega_{\Gamma_n}(v) \\ 0 & \text{if } s \in \mathcal{S} \setminus \Omega_{\Gamma_n}(v). \end{cases}$$

The *Julész-ensemble* can then be defined as the thermodynamic limit of the uniform distributions  $U_{n,v}(s)$  for  $n \rightarrow \infty$ :

**Definition 5.5.1. (Julész-Ensemble)** Given a set of specific feature-statistics-values  $v = (v_j)_{j \in \{1, \dots, n\}}$  for the  $\bar{\phi}_j$ , the corresponding *Julész-ensemble* is the limiting distribution of  $U_{n,v}(\cdot)$  as  $\Gamma_n \rightarrow \mathbb{Z}^d$ .

We are well aware that definition 5.5.1 is a little vague because we do not describe the limit process in detail neither take into account any boundary conditions. However to establish the ensemble-equivalence to an extent which suffices for our purposes, this is not necessary. Often besides to the limiting distribution the term *Julész-ensemble* also refers to the *set* of all configurations over  $\mathbb{Z}^d$  which have the values  $v_j$  for their specific feature characteristics. In thermodynamics the Julész-ensemble corresponds to the so called *micro-canonical ensemble* which describes an isolated particle system with a certain fixed energy. On the other hand the *Gibbs-ensemble* corresponds to the so called *canonical ensemble* which describes a system which is in equilibrium with a heat reservoir of constant temperature  $T$ . Likewise the Gibbs-ensemble can be defined as a thermodynamic limit of the *Gibbs-distribution*:

**Definition 5.5.2. (Gibbs-ensemble)** Given a set of specific feature-statistics values  $v = (v_j)_{j \in \{1, \dots, n\}}$  for the  $\bar{\phi}_j$ , the corresponding *Gibbs-ensemble* is the limiting distribution of the *Gibbsian distribution*  $\pi_{G,n}$  on  $\mathcal{S}_n$  with the energy function  $H^\phi$  as  $\Gamma_n \rightarrow \mathbb{Z}^d$ , where:

$$\pi_{G,n}(s) = \frac{\text{Exp}\left(-\sum_{j=1}^n \beta_j^n \cdot \phi_j(s)\right)}{\sum_{t \in \mathcal{S}_n} \text{Exp}\left(-\sum_{j=1}^n \beta_j^n \cdot \phi_j(t)\right)}, \quad \beta^n \text{ such that } E_{\beta^n}[\bar{\phi}_j] \stackrel{!}{=} v_j.$$

Hence the main difference between the ensembles of definitions 5.5.1 and 5.5.2 is that the Julész-ensemble selects configurations  $s$  over  $\mathbb{Z}^d$  which satisfy the hard constraint  $\bar{\phi}_j(s) = v_j$  whereas the Gibbs ensemble only implies the soft constraint  $E_{\pi_G}[\bar{\phi}_j] = v_j$ , thus in the Gibbsian-ensemble the predetermined values  $v_j$  must only be met in average. Note that the Gibbs distribution  $\pi_{G,n}$  assigns positive mass to *all* configurations  $s \in \mathcal{S}_n$  whereas  $U_{n,v}$  is concentrated on the Julész-ensemble only. Using the above terminology the ensemble-equivalence can now be described as follows: Under certain regularity conditions the distribution  $\pi_{G,n}(s)$  finally concentrates its mass uniformly on the Julész-ensemble and therefore the two ensembles (over  $\mathbb{Z}^d$ ) can be regarded as equivalent.

A heuristic argument for the ensemble-equivalence relies on the *variational principle* for the Gibbs-distribution. On finite lattices  $\Gamma$  we have seen in section 5.4 that the Gibbs distribution with  $E_{\pi_G}[\bar{\phi}_j] = v_j$  is the solution to the maximal entropy problem 5.7. It is equally easy to see that on finite lattices  $\Gamma$  it is the the uniform distribution  $U$  on  $\mathcal{S}$  that maximizes entropy, ie.  $U$  is the solution to:

$$\max_{p \in \mathcal{P}(\mathcal{S})} \left\{ - \sum_{s \in \mathcal{S}} p(s) \cdot \log p(s) \right\}, \quad (5.10)$$

subject to:

$$\sum_{s \in \mathcal{S}} p(s) = 1.$$

This can be seen as follows: The entropy  $S_U$  of the uniform distribution  $U$  is:

$$S_U = - \sum_{s \in \mathcal{S}} \frac{1}{|\mathcal{S}|} \cdot \log \frac{1}{|\mathcal{S}|} = \log |\mathcal{S}|.$$

If  $p = (p_1, \dots, p_{|\mathcal{S}|})$  is an arbitrary other distribution on  $\mathcal{S}$  with entropy  $S_P$  we have:

$$\begin{aligned} & \ln x \leq x - 1 \quad \forall x \in \mathbb{R}^+, \\ \Rightarrow & \log \left( \frac{1}{|\mathcal{S}| \cdot p_i} \right) \leq \frac{1}{|\mathcal{S}| \cdot p_i} - 1 \quad \forall i \in \{1, \dots, |\mathcal{S}|\}, \\ \Leftrightarrow & -p_i \cdot (\log |\mathcal{S}| + \log p_i) \leq \frac{1}{|\mathcal{S}|} - p_i \quad \forall i \in \{1, \dots, |\mathcal{S}|\}, \\ \xrightarrow{\sum_{s \in \mathcal{S}}} & -\log |\mathcal{S}| + S_P \leq 0 \\ \Rightarrow & S_P \leq \log |\mathcal{S}| = S_U. \end{aligned}$$

Obviously this argument still works if we look for a distribution with maximal entropy on a subset of  $\mathcal{S}$ . Thus  $U_{n,v}$  is the entropy maximizing distribution on  $\Omega_{\Gamma_n}(v)$ . If we summarize so far, we have that the Gibbsian distribution  $\pi_{G,n}$  maximizes entropy under the soft constraint  $E_{\pi_G}[\bar{\phi}_j] = v_j$  and the uniform distribution  $U_{n,v}$  maximizes entropy under the hard constraint  $\bar{\phi}_j = v_j$ . The equivalence of ensembles then follows if we can resort to a law of large numbers in the following sense:

$$v_j = \bar{\phi}_j \xrightarrow{\Gamma_n \rightarrow \mathbb{Z}^d} E_{\pi_G}[\bar{\phi}_j] = v_j, \quad \forall j \in \{1, \dots, n\}.$$

Of course this can also be proved more rigorously. The following theorem is proved for example in (Wu et al. 1999) and a sketch of its proof is provided in the appendix:

**Theorem 5.5.3. (Equivalence of Ensembles)** *If there is a unique  $v_0$ , where  $s_\beta(v)$  (defined as in equation 7.32) attains its maximum, then the Gibbs-distribution  $\pi_{G,n}$  for  $\Gamma_n \rightarrow \mathbb{Z}^d$  eventually concentrates its mass on the Julész-ensemble corresponding to  $v_0$ , and conversely the uniform distribution  $U_{n,v_0}$  for  $\Gamma_n \rightarrow \mathbb{Z}^d$  converges to the limit of the Gibbsian distribution  $\pi_{G,n}$ .*

**Proof. (Theorem 5.5.3)** see Appendix, subsection 7.4.2.

The equivalence of ensembles will give us a powerful tool to simulate from from the Gibbsian ensemble or more precisely from an approximation of it on a sufficiently large lattice  $\Gamma$ . In subsection 5.8 we will simulate from an approximation to the Julész ensemble instead which is a parameter-free model and hence the Gibbs-Model can be applied via the equivalence of ensembles without having to estimate any parameters at all!

## 5.6 Gibbs-Sampling

### 5.6.1 General Definition

We have now accumulated enough theoretical facts to be ready to apply the Gibbsian Model to the simulation of porous media. The basic question we have to answer when attempting to do so is how to simulate from the Gibbsian distribution (definition 5.1.1). Especially in the case of neighbourhood-potentials (definition 5.1.4) the well-known *Gibbs-Sampler* is perfectly suited to accomplish this task. In this section we will briefly recall the Gibbs-Sampling algorithm and exemplarily describe its most important application for us, namely the simulation from the Gibbsian distribution with the OM-energy (definition 5.3.1). For all techniques to simulate artificial porous structures with predefined values for the specific Minkowski-functionals that will be presented in section 5.7 it is always the Gibbs-sampler which we now describe that is used to simulate from the Gibbsian distribution and the methods only differ in choosing an appropriate parameter set  $\beta = (\beta_0, \dots, \beta_d)$ .

The Gibbs-Sampling-Algorithm can be used (in principle) to simulate from an arbitrary *multidimensional* distribution whose density we denote by  $f$ . Note that Gibbs-Sampling is a truly multi-dimensional procedure which makes no sense in the one dimensional setting. Thus let  $d \geq 2$ . It is crucial for the Gibbs-Sampler that the density  $f$  from which we intend to simulate is known analytically and all its one-dimensional conditional densities  $f_i(z_i|z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_d)$  are known and can be simulated. The Gibbs-Sampler is a Metropolis-Hastings-type MCMC-algorithm and therefore iteratively generates a Markov chain which has  $f$  as its limiting distribution regardless of the starting value  $z_0$  which must be provided. The main advantage of the Gibbs-Sampler over the usual Metropolis-Hastings procedure is that no simulated values are wasted, ie. each simulated value gets accepted with probability 1. If the above conditions are met, the algorithm works as follows:

#### Algorithm 5.6.1. (Gibbs-Sampler)

- i). Choose an arbitrary starting value  $z^{(0)} = (z_1^0, \dots, z_d^0)$
- ii). For  $t = 0, 1, 2, 3, \dots$  update the vector  $z^{(t)}$  as follows:
  - First draw  $z_1^{(t+1)}$  at random from  $f_1(z_1|z_2^{(t)}, \dots, z_d^{(t)})$ .
  - Then for  $i = 2, \dots, d-1$  draw  $z_i^{(t+1)}$  from  $f_i(z_i|z_1^{(t+1)}, \dots, z_{i-1}^{(t+1)}, z_{i+1}^{(t)}, \dots, z_d^{(t)})$ .
  - Finally draw  $z_d^{(t+1)}$  from  $f_d(z_d|z_1^{(t+1)}, \dots, z_{d-1}^{(t+1)})$ .

- iii). Repeat step ii). sufficiently often until the chain has reached it stationary distribution. Then  $(z_1, \dots, z_d) \sim f$ .

The correctness of the Gibbs-Sampler can be proved under mild regularity conditions. A detailed discussion with many application examples including proofs can be found in (Robert and Casella 1999, Chapter 7).

Although simulating from an (in principle) arbitrarily high-dimensional density each step of algorithm 5.6.1 only involves simulations from one-dimensional distributions. But often these distributions are not even analytically calculable or are at least difficult to simulate. However, for neighbourhood-Gibbs fields because of the Markov property (definition 5.2.1) it turns out that applying the Gibbs Sampler becomes very easy.

### 5.6.2 The Gibbs Sampler for Neighbourhood-Potentials

We will now specialize algorithm 5.6.1 to the case where we want to simulate from the Gibbsian distribution with the OM-energy (definition 5.3.1) on the space of binary configurations  $\mathcal{S}$  over a finite grid  $\Gamma$ . The neighbourhood used in our experiments was the classical 8-neighbourhood in two dimensions on the square grid but other neighbourhoods on other grid types (eg. a hexagonal grid) would as well be possible. In our case the density from which we want to simulate has the dimensionality of the grid size  $|\Gamma|$  and the components  $s(x_i)$  of the Gibbsian target vector  $\Theta = (\Theta(x_i)_{x_i \in \Gamma})$  we intend to simulate are the black/white-pixel values corresponding to the lattice point in an arbitrary but fixed order. In each iteration of the Gibbs-Sampler only a single pixel  $\Theta(x_i)$  is updated according to the conditional distribution  $f_i(\Theta(x_i)|\Theta(x_j), j \neq i)$  given all the other pixels of the image. Because of the Markov property (definition 5.2.1) only the pixels in the neighbourhood  $\delta(x_i)$  of  $x_i$  have to taken into account which in case of the 8-neighbourhood are only 8 pixels instead of  $|\Gamma| - 1$  in total or equivalently the 4 vertex configurations in a square of  $2 \times 2$  lattice cells centered at  $x_i$ . This makes the Gibbsian approach computationally feasible also for large three-dimensional images. If we order the configurations  $s \in \mathcal{S}$  in an arbitrary way, ie.  $s = \{s_1, \dots, s_{|\Gamma|}\}$  and if we denote by a slight abuse of notation  $s_i^c = \{s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_{|\Gamma|}\}$  the sub-image of  $s$  image when leaving out the  $i$ .th-pixel value and by  $\delta s_i = \{s_j \in s : x_j \in \delta(x_i)\}$  the sub-image of  $s$  consisting of the pixel values at neighbouring lattice points of  $x_i$ , we can write:

$$\begin{aligned}
 & P[\Theta(x_i) = s_i | \Theta(x_k) = s_k, x_k \in \Gamma, k \neq i] \\
 &= \frac{\text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_i \cup s_i^c)\right)}{\text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_i \cup s_i^c)\right) + \text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j((1-s_i) \cup s_i^c)\right)} \\
 &= \frac{\text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_i \cup \delta s_i)\right)}{\text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_i \cup \delta s_i)\right) + \text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j((1-s_i) \cup \delta s_i)\right)}, \quad (5.11)
 \end{aligned}$$

because all the pixel values of  $s$  which correspond to lattice points outside the neighbourhood  $\delta(x_i)$  show up in the exponential of the numerator and both summands of the denominator as well and thus cancel out.

How conveniently the Gibbs-sampler-algorithm adapts to our problem can be demonstrated best if we write the unspecific OM-estimators in the form of equation 5.8 and denote the 4 lattice cells meeting in  $x_i$  by  $\Gamma_{x_i}^{ur}$  (upper right),  $\Gamma_{x_i}^{ul}$  (upper left),  $\Gamma_{x_i}^{ll}$  (lower left),



$\Gamma_{x_i}^{lr}$  (lower right). We further denote the set of the corresponding vertex-configurations  $\in [0, 15]$  of an image  $s$  after setting  $s_i := 1$ .

$$\mathcal{I}_{\delta x_i} := \{i_{\Gamma_{x_i}^{ur}}, i_{\Gamma_{x_i}^{ul}}, i_{\Gamma_{x_i}^{ll}}, i_{\Gamma_{x_i}^{lr}}\} \quad (5.12)$$

accordingly. Then with the help of the *proposed configuration vector difference*  $\Delta h(i), i \in [0, 15]$  whose components we define according to the vertex nomenclature introduced in Chapter 2 (basically figure 2.1) to be 0 except for:

$$\begin{aligned} \Delta h(i_{\Gamma_{x_i}^{lr}}) &:= 1, & \Delta h(i_{\Gamma_{x_i}^{lr}} - 1) &:= -1, & \Delta h(i_{\Gamma_{x_i}^{ur}}) &:= 1, & \Delta h(i_{\Gamma_{x_i}^{ur}} - 2) &:= -1, \\ \Delta h(i_{\Gamma_{x_i}^{ul}}) &:= 1, & \Delta h(i_{\Gamma_{x_i}^{ul}} - 4) &:= -1, & \Delta h(i_{\Gamma_{x_i}^{ul}}) &:= 1, & \Delta h(i_{\Gamma_{x_i}^{ul}} - 8) &:= -1, \end{aligned} \quad (5.13)$$

it is easy to see that we can rewrite equation 5.11 as:

$$\begin{aligned} P[\Theta(x_i) = 0 | \Theta(x_k) = s_j, x_k \in \Gamma, k \neq i] &= \frac{1}{1 + \text{Exp}\left(-\sum_{j=0}^d \sum_{i=0}^{15} \beta_j \cdot c_{ji} \cdot \Delta h(i)\right)}, \\ P[\Theta(x_i) = 1 | \Theta(x_k) = s_j, x_k \in \Gamma, k \neq i] &= 1 - P[\Theta(x_i) = 0 | \Theta(x_k) = s_j, x_k \in \Gamma, k \neq i], \end{aligned} \quad (5.14)$$

where the  $c_{ij}$  are the contributions of configuration  $i$  to the OM-estimator of  $V_j$ . This demonstrates how easily the pixel updates can be performed by simply keeping track of the change  $\Delta h$  in the configuration vector  $h$ . The proposed configuration difference vector  $\Delta h$  describes the change in the configuration vector  $h$  if the pixel  $x_i$  flips from 0 to 1. However if the *actual* flip when updating the pixel is the other way around, we must of course also keep track of the *real* change in the configuration vector which can be similarly done. If the pixel doesn't flip its value at all, there is evidently no real change of the configuration vector at all. After a pixel has been updated the *updating probabilities* for the next pixel in equation 5.14 can again be computed using the new proposed configuration vector change  $\Delta h$  given above.

As a starting point for the Gibbs-sampler an arbitrary image  $s \in \mathcal{S}$  will do. We mainly used one of the following three natural choices for our experiments:

- Completely white image  $s_i \equiv 0, \forall i$  or completely black image  $s_i \equiv 1, \forall i$ .
- White noise image as a realization of  $\Theta$  with  $P[\Theta(x_i) = 0] = P[\Theta(x_i) = 1] = 0.5$  independently for all  $i$ .
- The true image for which an artificial Gibbsian realization with the same Minkowski-functionals should be created.

Of course the last choice should only be used to test the Gibbs-Sampling-Algorithm in the case of convergence problems. In general of course, the algorithm should be able to find the equilibrium distribution on its own.

Evidently, the Gibbs-sampler algorithm works not only for the OM-energy but analogously for general image feature statistics  $\phi(\cdot)$  defined according to equation 5.8. We can summarize the above arguments about the Gibbs-sampler with potentials of the form given in equation 5.9 in the following algorithm:

**Algorithm 5.6.2. (Gibbs-Sampler for  $H^\phi$ -energies)** Let a finite 2-dimensional square lattice  $\Gamma = \{x_1, \dots, x_{|\Gamma|}\}$  with a specified ordering of its lattice points be given and denote by  $\mathcal{S}$  the space of binary configurations over  $\Gamma$ . By  $h^{(t,i)}$  we denote the configuration vector of the actual image after  $i$  pixels have been updated in the  $t$ -th image-sweep and  $s^{(t,i)}$  denotes the actual image after  $i$  pixels have been updated in the  $t$ -th sweep. Further let the  $c_{ij}$  be the contribution of the  $i$ th configuration to the OM-estimator of the  $j$ th Minkowski functional. To simulate a random image  $\Theta = (\Theta(x_i))_{x_i \in \Gamma}$  as a realization from the Gibbsian distribution driven by an energy function  $H^\phi$  of the form given in equation 5.9 with known coefficients  $\beta_j, j \in \{1, \dots, n\}$ :

- i). Choose an arbitrary starting image  $s^{(0,0)}$  from  $\mathcal{S}$  and determine its configuration vector  $h^{(0,0)}$ .
- ii). For  $t \in \{0, 1, 2, 3, \dots\}$  sweep the image by updating for  $i \in \{1, \dots, |\Gamma|\}$  the pixel values corresponding to the lattice points  $x_i$  according to the pre-specified order. The update of a single pixel at  $x_i$  is done with the following procedure:

1. Determine the configurations the neighbouring lattice cells to  $x_i$  would have if  $s_i$  were 1 (see equation 5.12 and preceding paragraph):

$$\mathcal{I}_{\delta x_i} := \{i_{\Gamma_{x_i}^{ur}}, i_{\Gamma_{x_i}^{ul}}, i_{\Gamma_{x_i}^{lr}}, i_{\Gamma_{x_i}^{ll}}\} \in [0, 15]^4.$$

2. Calculate the proposed configuration vector difference  $\Delta h(i)$  for the configurations  $\mathcal{I}_{\delta x_i}$  given in equation 5.13.
3. Generate a realization  $U(\omega)$  of the uniform random variable  $U \sim Unif[0, 1]$ .
4. If

$$U(\omega) \leq \frac{1}{1 + \text{Exp}\left(-\sum_{j=0}^d \sum_{i=0}^{15} \beta_j \cdot c_{ji} \cdot \Delta h(i)\right)},$$

set  $s_i^{(t,i)} := 0$ , otherwise set  $s_i^{(t,i)} := 1$ . Set further  $s_j^{(t,i)} := s_j^{(t,i-1)} \forall j \neq i$ .

5. Because due to the pixel-update the vertex-configurations of the four lattice cells  $\Gamma_{x_i}^{ur}, \Gamma_{x_i}^{ul}, \Gamma_{x_i}^{lr}, \Gamma_{x_i}^{ll}$  meeting in  $x_i$  might have been altered, update the configuration vector  $h^{(t,i-1)}$  to  $h^{(t,i)}$  by changing the 8 relevant entries if a pixel-flip has occurred (ie. if  $s_i^{(t,i)} \neq s_i^{(t,i-1)}$ ), otherwise set  $h^{(t,i)} := h^{(t,i-1)}$ . If  $i = |\Gamma|$ , set  $s^{(t+1,0)} := s^{(t,|\Gamma|)}$  and  $h^{(t+1,0)} := h^{(t,|\Gamma|)}$

- iii.) Repeat step ii). until the equilibrium has been reached. The actual image  $s^{(t_0,0)}$  (for  $t_0$  large enough) is then a realization from  $\Theta$ .

**Remark 5.6.3.**

- i). The update of the configuration vector in step ii), part 5 in case of a pixel flip is done according to equation 5.13 if the pixel flips from 0 to 1:  $h^{(t,i)} := h^{(t,i-1)} + \Delta(h)$ . If the flip is from 1 to 0 the configuration vector change is also given by equation 5.13 after flipping signs:  $h^{(t,i)} := h^{(t,i-1)} - \Delta(h)$ .
- ii). Obviously the pixels at the boundary of  $\Gamma$  cannot be updated according to the above algorithm because they have no complete neighbourhood. Either one uses a slightly modified update-procedure by just considering the alterations in the vertex-configurations of the one or two lattice cells meeting in such a boundary point or

one simply keeps the boundary fixed. For this second method a good choice for the fixed boundary would be to take the boundary of an image of the real structure one intends to model. Of course one can also create a boundary at random.

- iii). Instead of updating the pixel according to a fixed order one can also draw the index number of the pixel to be updated at random according to the uniform distribution on the lattice  $\Gamma$ .
- iv). The same algorithm works also for 3-dimensions. For every pixel update the 256-dimensional configuration vector then might change in 16 entries because on the square lattice there are 8 cubic lattice cells which meet in an inner lattice point.
- v). The algorithm can easily be generalized to work for general neighbourhood potentials, the form of  $H^\phi$  energies given in equation 5.9 is by no means mandatory, and the Markov property can still be used to keep the simulation simple. However for such potentials the update probabilities no longer depend on the configuration vector only and equation 5.11 must be used with the  $\hat{V}_j(\cdot)$  replaced by general feature statistics  $\phi(\cdot)$  which are no longer calculable locally. To calculate the probabilities in equation 5.11 efficiently one must keep track of the influence of a single-pixel update on these feature statistics. For example in our experiments we used potentials including the two-point-covariance function (definition 2.5.1), the chord-length-distribution (definition 2.5.3) and also the Minkowski-*functions* (see definition 2.5.5). The influence of a pixel-flip on these quantities is more subtle than for the configuration vector, for example in case of the chord-length-distribution flipping a pixel from 0 to 1 may or may not give rise to an increase or decrease of the total number of chords in the image depending on the state of its neighbours. To determine the influence of a pixel swap on the Minkowski-*functions* as functions of the dilation radius is much more difficult and time-consuming, because the change of Minkowski-*functions* can obviously no longer be determined by simply considering the 8-neighbourhood of a pixel. This means that a potential including the Minkowski-*functions* is no longer a neighbourhood-potential with respect to the 8-neighbourhood and the Markov-property can no longer be exploited. Of course one can make any potential a neighbourhood-potential by just enlarging the neighbourhood, but evidently this does not help reducing the computational workload because the Markov-property quickly loses its value when the neighbourhood becomes large. Nevertheless we actually did also include the Minkowski-*functions* in our experiments (see section 5.9).

As an alternative to updating only one single pixel in every step one could employ a *Block-Gibbs-sampler* strategy instead where a whole block (eg. a square sub-lattice) of the image is updated in a single step. The advantage of this procedure is the by a factor 1/block-size decreased number of update steps but the drawback of this block-wise updating is that the multi-dimensional conditional distributions for the blocks are far more difficult to handle.

What remains to answer for the simulation of random porous structures with the Gibbsian model once we know how to simulate from the Gibbsian distribution is the problem how to determine the parameters  $\beta$ . This crucial question shall be addressed in the next section.

## 5.7 Parameter Determination

Our goal is to simulate from the Gibbsian distribution which corresponds to a random set  $\Theta$  that fulfills the basic modelling equation 5.3 and whose realizations thus in average exhibit the predetermined values for the specific Minkowski-functionals. Equation 5.3 can be satisfied by choosing the parameter values  $\beta$  of the Gibbsian distribution accordingly. In section 5.3 the problem of choosing the right parameter set  $\beta = (\beta_0, \dots, \beta_d)$  has basically been solved: Equation 5.5 tells us that  $\beta$  has to be chosen to be the maximum likelihood estimator  $\hat{\beta}_{MLE}$  under the Gibbs model. Unfortunately this is not feasible for numerical optimization in practice -at least for reasonably sized images - because the denominator of the likelihood  $Z(\beta)$  involves  $2^{|\Gamma|}$  exponential terms which would have to be evaluated in each optimization step. A natural way out of this problem is to consider a *pseudo-likelihood* to be minimized which is more amenable for numerical optimization. This is going to be our first parameter-estimation method which we describe in the following:

### 5.7.1 Pseudo-Likelihood-Logistic-Regression-Method

As mentioned above it is hard to calculate probabilities of the form  $P[\Theta(x_i) = s_i, i \in \Gamma]$  in the Gibbs model but conditional probabilities of the form  $P[\Theta(x_i) = s_i | \Theta(x_k) = s_k, k \neq i]$  are easy to deal with because of the Markov property. Therefore a natural substitute for the likelihood to consider is the following quantity:

**Definition 5.7.1. (Pseudo-Likelihood)** Let  $s \in \mathcal{S}$  be a given pixel configuration over  $\Gamma$ . We define the *Pseudo-Likelihood* for  $s$  in the Neighbourhood-Gibbsian Model  $\Theta$  to be the quantity

$$PSL_s(\beta) = \prod_{x_i \in \Gamma} P[\Theta(x_i) = s_i | \Theta(x_k) = s_k, k \neq i] = \prod_{x_i \in \Gamma} P[\Theta(x_i) = s_i | \Theta(x_k) = s_k, x_k \in \delta(x_i)].$$

For a given image  $s \in \mathcal{S}$  (usually the image whose Minkowski functionals our artificial structure  $\Theta$  should reproduce in average), we then estimate the Gibbsian parameters  $\beta$  as:

$$\hat{\beta}_{PSL} = \arg \max_{\beta} PSL_s(\beta) \quad (5.15)$$

In case of the OM-energy (definition 5.3.1) this optimization problem becomes:

$$\hat{\beta}_{PSL} = \arg \max_{\beta} \prod_{x_i \in \Gamma} \frac{\text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_i \cup \delta s_i)\right)}{\text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j(s_i \cup \delta s_i)\right) + \text{Exp}\left(-\sum_{j=0}^d \beta_j \cdot \hat{V}_j((1-s_i) \cup \delta s_i)\right)},$$

which is a convex optimization problem whose target function for a given  $\beta_0$  can easily be calculated in a single sweep of the image  $s$ . However for computational reasons it is better to maximize  $\log PSL_s(\beta)$  which is a sum of negative terms instead of a product of small ones.

Of course the estimator  $\hat{\beta}_{PSL}$  in general will *not* fulfill the basic modelling equation 5.3 but it can be understood as an estimate of the maximum-likelihood-estimator  $\hat{\beta}_{MLE}$  which does. The pseudo-likelihood-approach is well justified because it is a well-know fact, that the pseudo-likelihood estimator is consistent in the thermodynamic limit, ie. it

converges to the  $\hat{\beta}_{MLE}$  in the thermodynamic limit with probability one. (for a proof in a very general context, see (Comets 1992)).

The above optimization problem (equation 5.15) can be written as a *logistic regression* model and thus the optimization can be done using the optimized functions for logistic regression of standard statistical software packages. To see this exemplarily in the case of the OM-energy we note that from equation 5.11 we can write:

$$\begin{aligned} & \text{P} [\Theta(x_i) = s_i | \Theta(x_k) = s_k, x_k \in \Gamma, k \neq i] \\ &= \frac{\text{Exp} \left( - \sum_{j=0}^d \beta_j \cdot \left[ \hat{V}_j(s_i \cup \delta s_i) - \hat{V}_j((1-s_i) \cup \delta s_i) \right] \right)}{1 + \text{Exp} \left( - \sum_{j=0}^d \beta_j \cdot \left[ \hat{V}_j(s_i \cup \delta s_i) - \hat{V}_j((1-s_i) \cup \delta s_i) \right] \right)} = \frac{\text{Exp}(\beta \cdot \eta)}{1 + \text{Exp}(\beta \cdot \eta)}, \end{aligned}$$

where we have introduced the vector  $\eta$  with components  $\eta_j = -\hat{V}_j(s_i \cup \delta s_i) + \hat{V}_j((1-s_i) \cup \delta s_i)$  and we can recognize the cumulative distribution function of the *logistic distribution* on the last line. Hence we have the equivalence to the following logistic regression model:

$$\log \left( \frac{\text{P} [\Theta(x_i) = s_i | \Theta(x_k) = s_k, x_k \in \Gamma, k \neq i]}{\text{P} [\Theta(x_i) = 1 - s_i | \Theta(x_k) = s_k, x_k \in \Gamma, k \neq i]} \right) = \beta \cdot \eta,$$

where  $\eta$  plays the role of the vector of the  $d+1$  explanatory variables. If we now treat the pixel values  $s_i, i \in \{1, \dots, |\Gamma|\}$  as *independent* realizations of a binary response variable  $Y$  with

$$\text{P} [Y_i = 1] := \text{P} [\Theta(x_i) = 1 | \Theta(x_k) = s_k, x_k \in \Gamma, k \neq i],$$

in the above logistic regression problem, then the likelihood of the logistic regression problem evidently agrees with the pseudo-likelihood in definition 5.7.1. Because the standard estimation procedure for the logistic regression coefficients is maximum likelihood, the logistic regression coefficients  $\hat{\beta}$  estimated by standard logistic regression programs are exactly the solution to the optimization problem of equation 5.15.

Our experiments have shown that the pseudo-likelihood-approach may perform poorly to find a good parameter set  $\hat{\beta}$  which at least approximately fulfills the basic modelling equation 5.3. It can be refined by a grid search to invert the functions

$$g_j : \beta \rightarrow E_\beta[\hat{V}_j(\Theta)] \tag{5.16}$$

in the neighbourhood of the parameter  $\hat{\beta}_{PSL}$  proposed by the pseudo-likelihood-method and then choose as a final parameter set the  $\beta_{\text{final}}$  which minimizes  $\sum_{j=0}^d |g_j(\beta_{\text{final}}) - V_j(s_0)|$  where the  $V_j(s_0)$  are the predetermined values for the unspecific Minkowski functionals. To perform such a grid search is difficult because to approximate  $E_{\beta_0}[\hat{V}_j(\Theta)]$  for a specific parameter vector  $\beta_0$  a sample of images  $s_i$  must be generated by Gibbs-Sampling, then the  $\hat{V}_j(s_i)$ -values must be computed and averaged. It is obvious that it would be extremely inefficient and time-consuming to run a different Gibbs-sampler for *each*  $\beta$  of the grid. However the techniques of *reweighting mixtures* and *reverse logistic regression* introduced by (Geyer 1991) can be used which allow us to perform a Gibbs-sampling only for a few well-chosen evenly spread  $\beta$  of the grid and estimate  $E_\beta[\hat{V}_j(\Theta)]$  for the other parameter values  $\beta$  in between *without* further Gibbs-Sampling. We will briefly summarize these techniques in the following.

Let us assume that  $g$  is a measurable real-valued function and that  $X \sim \pi$  for an arbitrary distribution  $\pi$  (absolutely continuous with respect to Lebesgue measure). We further assume that about  $\pi$  with density  $f_\pi$  we only know its unnormalized density  $h_\pi$ . To approximate the expectation  $E_\pi[g(X)]$  by means of a sample  $X_i, i \in \{1, \dots, n\}$  following *another* distribution  $\nu$ , with (unnormalized) density  $h_\nu$ , the well-known *importance sampling* technique can be used: Because of

$$E_\pi[g(X)] = \int g(x) \cdot f_\pi(x) dx = \int g(x) \cdot \frac{f_\pi(x)}{f_\nu(x)} \cdot f_\nu(x) dx = E_\nu \left[ g(x) \cdot \frac{f_\pi(x)}{f_\nu(x)} \right] = \frac{E_\nu \left[ g(x) \cdot \frac{f_\pi(x)}{f_\nu(x)} \right]}{E_\nu \left[ \frac{f_\pi(x)}{f_\nu(x)} \right]},$$

where we have used in the last equation that  $E_\nu \left[ \frac{f_\pi(x)}{f_\nu(x)} \right] = 1$ , we can estimate:

$$E_\pi[g(X)] \approx \sum_{i=1}^n \omega(X_i) \cdot g(X_i), \quad \text{where} \quad \omega(x) := \frac{h_\pi(x)/h_\nu(x)}{\sum_{i=1}^n h_\pi(x)/h_\nu(x)}. \quad (5.17)$$

Note that by using the special form of the importance sampling weights  $\omega(\cdot)$  introduced in equation 5.17 we can use the unnormalized densities because the normalizing constants of  $\pi$  and  $\nu$  cancel in the expression for  $\omega(\cdot)$ . It is obvious that importance sampling makes sense only if the two distributions  $\pi$  and  $\nu$  have similar support and mass allocation, hence in our case it is certainly not enough to use a single Gibbsian distribution  $\pi_{G, \beta_0}$  to calculate the expectations given in equation 5.16 for all the other  $\beta$ -values of the grid. But assume that we have used Gibbs-sampling to generate samples  $s_{ik}, k \in \{1, \dots, m\}, i \in \{1, \dots, n_k\}$  for a selection of evenly spread  $\beta$ -vectors  $\{\beta_{01}, \dots, \beta_{0m}\}$  of our parameter grid and suppose we want to calculate the expectation of equation 5.16 for another  $\beta$  of our parameter grid lying for example in the convex hull of the  $\beta_{0k}$ . We could use each of the samples  $s_k$  to propose an estimate for the desired expectation with  $g(\cdot) = \hat{V}_j(\cdot)$  calculated according to the ordinary importance sampling equation 5.17 leading to  $m$  different estimates. If we denote by

$$h_{G, \beta_{0k}}(\cdot) := c_k \cdot \pi_{G, \beta_{0k}}(\cdot)$$

the unnormalized versions of the Gibbsian densities  $\pi_{G, \beta_{0k}}$  for the parameters  $\beta_{0k}$  with normalization constants  $c_k$  a much better *single* estimate can be produced as follows: Consider the *mixture-distribution* with density

$$\pi_{mix}(s) := \sum_{k=1}^m \frac{n_k}{n} \cdot \pi_{G, \beta_{0k}}(s) = \sum_{k=1}^m \frac{n_k}{n} \cdot \frac{h_{G, \beta_{0k}}(s)}{c_k}, \quad (5.18)$$

where  $n = \sum_{k=1}^m n_k$  and assume that the joint sample  $s..$  of the  $s_{ik}$  was generated by this distribution. We would then like to use this mixture distribution as a sampling distribution for importance sampling. Unfortunately the formula for the mixture-distribution involves the normalization constants  $c_k$  which cannot be calculated for the Gibbs-distribution with reasonable effort as we have mentioned several times now. However, it turns out that the  $c_k$  can be estimated up to a common proportionality factor  $c$  by reverse logistic regression which we describe later. Hence if we set  $C_k := c \cdot c_k$  and use the the following unnormalized version of the mixture density,

$$h_{mix}(s) := \sum_{k=1}^m \frac{n_k}{n} \cdot \frac{h_{G, \beta_{0k}}(s)}{\hat{C}_k},$$

with  $\hat{C}_k$  being the estimate for  $C_k$ , we can use the importance sampling scheme of equation 5.17 with  $g := \hat{V}_j$ ,  $h_\pi := Z(\beta) \cdot \pi_{G,\beta}$  and  $h_\nu := h_{mix}$  to produce a single estimate of  $E_\beta[\hat{V}_j(\Theta)]$  exploiting the information of all the generated samples  $s_{ik}$  simultaneously. Note that the  $\hat{C}_j$  need only to be estimated once and hence the mean Minkowski functionals per unit volume can be estimated for many different parameter sets  $\beta$  always using the *same* mixture distribution. Furthermore because the normalizing constants  $Z(\beta)$  and  $c$  don't have to be calculated this estimation method is very efficient. Please note also that this procedure only works because for importance sampling according to the scheme 5.17 both the target and the sampling distribution only need to be known up to a proportionality factor which may be different for the two distributions. Using importance sampling in this way when assuming the pooled sample stems from the above mixture distribution is referred to as the method of *reweighting mixtures*.

It remains to describe the method of *reverse logistic regression* used to estimate the  $C_k$ . Note that because of the positivity of  $n$ ,  $n_k$  and  $c_k$  the mixture density of equation 5.18 can be written as:

$$\pi_{mix}(s) := \sum_{k=1}^m h_{G,\beta_{0k}}(s) \cdot e^{\eta_k},$$

with  $\eta_k := -\log c_k + \log \frac{n_k}{n}$ . Furthermore we denote

$$p_k(s, \eta) := \frac{h_{G,\beta_{0k}}(s) \cdot e^{\eta_k}}{\sum_{l=1}^m h_{G,\beta_{0l}}(s) \cdot e^{\eta_l}} \quad (5.19)$$

which is the probability that if an image  $s \in \mathcal{S}$  was observed in the joint sample  $s_{..}$ , it belongs to the  $k$ -th sample  $s_{.k}$ . We now estimate the  $\eta_j$  by maximizing the *log-quasi-likelihood*  $l_n(\eta)$  defined as

$$l_n(\eta) := \sum_{k=1}^m \sum_{i=1}^{n_k} \log p_k(s_{ik}, \eta) \quad (5.20)$$

with respect to  $\eta$ . This is a concave optimization problem which can be solved computationally. Finally we find the estimates  $\hat{C}_k$  to be:

$$\hat{C}_k := e^{-\eta_k} \cdot \frac{n_k}{n}.$$

Note that from equation 5.19 it is evident, that the  $\eta_j$  are only determined up to addition of a common constant and hence the  $\hat{C}_k$  only estimate the normalization constants up to a common factor of proportionality. To make the solution identifiable we use the convention  $\eta_m = 0$  in our experiments.

The optimization of  $l_n(\eta)$  in equation is more complicated that it might seem to be at first glance. To demonstrate this let us assume for simplicity that all the  $m$  samples are equal in size, i.e.  $n_k = N \forall k$  and hence  $n = m \cdot N$ . With the above convention the log-quasi-likelihood then reads:

$$l_\eta(\eta) = N \cdot \sum_{k=1}^{m-1} \eta_k - \sum_{k=1}^m \sum_{i=1}^N \log \left[ \sum_{l=1}^m \text{Exp} \left( \eta_l - \beta_{0l} \cdot \hat{V}(s_{ik}) \right) \right] + A, \quad (5.21)$$

where  $A$  collects the remaining terms not containing  $\eta$  and  $\hat{V}$  denotes the vector of the joint Minkowski-functionals for  $j = 0, \dots, d$ . It is obvious that it is easily possible that

computational overflow may occur in equation 5.21 when calculating the argument of the logarithm. It is therefore wise to transform equation into the following more computationally stable form before optimizing:

$$l_\eta(\eta) - A = N \cdot \sum_{k=1}^{m-1} \eta_k - \sum_{k=1}^m \sum_{i=1}^N \left[ \eta_{l_0^{ik}} - \beta_{0l_0^{ik}} \cdot \hat{V}(s_{ik}) + \log \left( 1 + \sum_{l \neq l_0^{ik}} \text{Exp} \left( \underbrace{\eta_l - \eta_{l_0^{ik}} - (\beta_{0l} - \beta_{0l_0^{ik}}) \hat{V}(s_{ik})}_{\leq 0} \right) \right) \right],$$

where  $l_0^{ik} := \arg \max_l (\eta_l - \beta_{0l} \cdot \hat{V}(s_{ik}))$ .

This makes all the  $n$  arguments of the exponentials negative and thus prevents overflow.

Our experiments have shown that even with the pseudo-likelihood-method refined by the grid-search involving reweighting mixtures and reverse logistic regression, determining a parameter set  $\beta$  solving the basic modelling equation satisfactorily close for all  $j \in \{1, \dots, d\}$  is sometimes difficult. It seems promising to split up the task of finding all  $d$  Gibbsian parameters  $\beta_j$  *simultaneously* into smaller and easier problems and find them *sequentially* instead. This approach is pursued by the *Sequential-Newton-Method* discussed in the next subsection.

### 5.7.2 Sequential Newton-Parameter Determination

In this subsection we present a trial-and-error method for the determination of a Gibbsian parameter vector  $\beta$  which solves the basic modelling equations 5.3 for the Gibbsian model  $\Theta$  with the OM-energy (definition 5.3.1). The method iteratively proposes new parameter values according to a Newton-scheme. This iterative procedure allows us to determine the parameters *sequentially*. The method consists of two nested iteration cycles. For the *outer iteration* we start out with the uniform distribution over  $\mathcal{S}$  corresponding to the Gibbs-parameter vector  $\beta = (0, \dots, 0)$ . In the first iteration step we find a vector  $\beta^0 = (\beta_0^0, 0, \dots, 0)$  which fulfills only the modelling equation for the volume, ie.  $E_{\beta^0}[\hat{V}_0(\Theta)] \approx v_0$ . This solution  $\beta_0$  is found by means of the *inner iteration* cycle which uses the ordinary Newton-method to approximate the solution of the modelling equation for the volume. The parameter vector  $\beta_0$  is then used as a starting vector for the second outer iteration step in which we find a solution  $\beta^1 = (\beta_0^1, \beta_1^1, 0, \dots, 0)$  such that the first two modelling equations  $E_{\beta^1}[\hat{V}_j(\Theta)] \approx v_j, j \in \{0, 1\}$  hold simultaneously. This is achieved again by the same inner iteration Newton-scheme. We proceed in this way finding in each step a parameter vector with one less zero entry that fulfills one more modelling equation until a solution  $\beta = \beta^d = (\beta_0^d, \dots, \beta_d^d)$  is found which (at least approximately) makes all the  $d+1$  modelling equations hold true. Please note that the Newton method changes also components that have been found in an earlier outer iteration step and thus  $\beta_s^i \neq \beta_s^j$  for  $i \neq j$ . The obvious advantage of this sequential procedure over a method that fits all the  $d+1$  parameters simultaneously is that the original problems is divided up into smaller subproblems which may be easier to solve. The main drawback as with every Newton-type algorithm is that one may end up cycling between wrong parameter sets and never reach a satisfactorily precise solution.



We now come to the precise mathematical description of the Sequential-Newton-Method. For this purpose we need to introduce some more pieces of notation. In the following the first superscript  $k \in \{0, \dots, d\}$  in the notation  $\beta^{(k,n)}$  indicates that the vector  $\beta_k^n$  was produced in the  $k$ -th outer iteration step (where counting iteration steps starts with 0) whereas the second superscript  $n$  refers to the  $n$ -th inner iteration step. All considered vectors are of the following form:

$$\beta^{(k,n)} = (\beta_0^{(k,n)}, \dots, \beta_k^{(k,n)}, 0, \dots, 0) \in \mathbb{R}^{d+1}, \quad (5.22)$$

hence of the vectors considered in the  $k$ -th outer iteration step only the first  $k+1$  components are possibly different from zero and  $\beta_s^{(k,n)} = 0, s \in \{k+1, \dots, d\}$ . Furthermore by  $V = (V_0, \dots, V_d)$  we denote the vector of preset (unspecific) Minkowski-values and for an image  $s \in \mathcal{S}$  we use  $\hat{V}(s) = (\hat{V}_0(s), \dots, \hat{V}_d(s))$  for the vector of OM-estimators of its (unspecific) Minkowski functionals. Similarly as above the vectors

$$V^k := (V_0, \dots, V_k, 0, \dots, 0), \quad \hat{V}^k(s) := (\hat{V}_0(s), \dots, \hat{V}_k(s), 0, \dots, 0) \quad \text{for } k \in \{0, \dots, d\}$$

are built from these Minkowski-vectors  $V$  and  $\hat{V}(s)$  by setting all but the first  $k+1$  entries to zero. The notation  $\beta^{(k,\text{opt})}$  is used to denote a (not necessarily unique) optimal solution which is the goal for the  $k$ -th iteration step of the outer iteration, ie.:

$$V^k \approx E_{\beta^{(k,\text{opt})}}[\hat{V}^k(\Theta)].$$

Unfortunately there is no way to analytically evaluate expectations of the form  $E_{\beta}[\hat{V}^k(\Theta)]$  in the Gibbs model. We only can approximate them by averaging the Minkowski-values  $\hat{V}^k(s_t)$  of a sample of images  $s_t \in \mathcal{S}$  where  $s_t \sim \pi_{G,\beta}, t \in \{1, \dots, T\}$  are produced by Gibbs-sampling and  $T$  denotes the size of the sample. In the following we use the notations  $s_t^{(k,n)}$  for sample images with  $s_t^{(k,n)} \sim \pi_{G,\beta^{(k,n)}}$  and  $s_t^{(k,\text{opt})}$  for sample images with  $s_t^{(k,\text{opt})} \sim \pi_{G,\beta^{(k,\text{opt})}}$ . The definition of  $\beta^{(k,\text{opt})}$  must therefore be relaxed in the following way:  $\beta^{(k,\text{opt})}$  denotes a (not necessarily unique) solution to the equation:

$$V^k \approx \frac{1}{T} \sum_{t=1}^T \hat{V}^k(s_t^{(k,\text{opt})}) \approx E_{\beta^{(k,\text{opt})}}[\hat{V}^k(\Theta)], \quad (5.23)$$

To be a little more precise, in the  $k$ -th outer iteration step,  $\beta^{(k,\text{opt})}$  is found by an inner iteration producing a Newton sequence of parameter vectors  $\beta^{(k,n)}$  until equation 5.23 is satisfied up to some predefined tolerance, ie:

$$\beta^{(k,\text{opt})} := \beta^{(k,n_k)}$$

with

$$n_k = \min_n \left\{ \beta^{(k,n)} \left\| \left\| V^k - \frac{1}{T} \sum_{t=1}^T \hat{V}^k(s_t^{(k,n)}) \right\|_{\mathcal{L}^1} < \epsilon \right\} \right\}. \quad (5.24)$$

For simplicity we use the same sample size  $T$  for all  $k$  (and also for all pairs  $(k,n)$  respectively) although this is not necessary as long as the sample sizes for the averages approximating the expectations are chosen large enough. We must now describe how the Newton sequences  $\beta^{(k,n)}$  for a given  $k$  are generated. For this purpose we consider the functions  $f^k : \mathbb{R}^{d+1} \rightarrow \mathbb{R}^{d+1}$  on the  $k+1$ -dimensional subspace of  $\mathbb{R}^{d+1}$  which is spanned by the first  $k+1$ -unit vectors.

$$f^k(\beta) := E_\beta[\hat{V}^k(\Theta)] - V^k = \frac{\sum_{s \in \mathcal{S}} \hat{V}^k(s) \cdot \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)}{\sum_{s \in \mathcal{S}} \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)} - V^k.$$

Obviously  $f^k(\beta)$  is the objective function of which we search a root by the Newton-algorithm in the  $k$ -th outer iteration step. Note that because we only consider  $f_k$  on the subspace spanned by the first  $k+1$  coordinates, the sums in the exponentials only involve the first  $k+1$  terms. For the Newton method we need to consider the derivatives

$$\begin{aligned} \frac{\partial f_i^k}{\partial \beta_{j_0}} &= - \frac{\sum_{s \in \mathcal{S}} \hat{V}_i(s) \hat{V}_{j_0}(s) \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)}{\sum_{s \in \mathcal{S}} \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)} \\ &+ \frac{\left(\sum_{s \in \mathcal{S}} \hat{V}_i(s) \cdot \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)\right) \cdot \left(\sum_{s \in \mathcal{S}} \hat{V}_{j_0}(s) \cdot \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)\right)}{\left(\sum_{s \in \mathcal{S}} \text{Exp}\left(-\sum_{j=0}^k \beta_j \cdot \hat{V}_j(s)\right)\right)^2}, \end{aligned}$$

where  $0 \leq j_0 \leq k$  and thus returning to matrix notation we conclude for the derivative:

$$\frac{\partial f^k}{\partial \beta} = \left[ E_\beta[\hat{V}^k(\Theta)] \right] \cdot \left[ E_\beta[\hat{V}^k(\Theta)] \right]^T - E_\beta[\hat{V}^k(\Theta) \cdot (\hat{V}^k(\Theta))^T] = -\text{Cov}_\beta[\hat{V}^k(\Theta)], \quad (5.25)$$

where only the upper left  $k+1 \times k+1$  square sub-matrix is non-zero. Applying the Newton-algorithm gives us the following update-step for the inner iteration:

$$\begin{aligned} \beta^{(k,n+1)} &= \beta^{(k,n)} - \left[ \frac{\partial f^k}{\partial \beta}(\beta^{(k,n)}) \right]^{-1} \cdot f^k(\beta^{(k,n)}) \\ &= \beta^{(k,n)} + \left[ \text{Cov}_{\beta^{(k,n)}}[\hat{V}^k(\Theta)] \right]^{-1} \cdot \left( E_{\beta^{(k,n)}}[\hat{V}^k(\Theta)] - V^k \right), \end{aligned}$$

where again the expectations must be replaced by sample averages. The sample version of the update-step of the inner iteration thus reads:

$$\begin{aligned} \beta^{(k,n+1)} - \beta^{(k,n)} &= \\ &\left[ \frac{1}{T} \sum_{t=1}^T \hat{V}^k(s_t^{(k,n)}) \cdot \left( \hat{V}^k(s_t^{(k,n)}) \right)^T - \left( \frac{1}{T} \sum_{t=1}^T \hat{V}^k(s_t^{(k,n)}) \right) \cdot \left( \frac{1}{T} \sum_{t=1}^T \hat{V}^k(s_t^{(k,n)}) \right)^T \right]^{-1} \times \\ &\left( \frac{1}{T} \sum_{t=1}^T \hat{V}^k(s_t^{(k,n)}) - V^k \right), \end{aligned} \quad (5.26)$$

where we recall that the  $s_t^{(k,n)} \sim \pi_{G,\beta^{(k,n)}}$  are independently generated by Gibbs-sampling.

Please note that in the light of equation 5.25 alternatively the method of reweighting mixtures presented in the last subsection can be used to estimate the covariance matrix in 5.25 to speed up the algorithm and reduce the number of samples.

Computing the inverse in the typical cases  $d = 2, 3$  is neither difficult nor time-consuming and in the unlikely event that the sample covariance is singular one can for example resort to averaging over a sufficiently large *subsample* of the  $s_i^{(k,n)}$ .

As a starting value for the inner iteration used in the  $k$ -th step of the outer iteration we use the optimal solution found in the  $k-1$ -th step, ie.:

$$\beta^{(k,0)} := \beta^{(k-1,\text{opt})},$$

whereas at the beginning for the outer iteration step  $k = 0$  we simply use  $\beta^{(0,0)} := (0, \dots, 0)$  as a starting value. The final solution  $\beta^{\text{opt}}$  is of course the parameter  $\beta^{\text{opt}} := \beta^{(d,\text{opt})}$ . We summarize the procedure in the following algorithm:

**Algorithm 5.7.2. (Sequential Newton Parameter Determination)** To determine a parameter set  $\beta^{\text{opt}} \in \mathbb{R}^{d+1}$  which approximately satisfies the basic modelling equations  $E_{\beta^{\text{opt}}}[\hat{V}_j(\Theta)] = V_j$  for the Gibbsian Model  $\Theta$  with the OM-energy, one can proceed as follows using the notation introduced further above in this subsection:

- i). Set  $\beta^{(0,0)} = (0, \dots, 0) \in \mathbb{R}^d$  as a starting value.
- ii). (*inner iteration*) For  $k = 0$ , iterate  $n = 0, \dots, n_k - 1$  computing  $\beta^{(k,n+1)}$  using the Newton scheme from equation 5.26 until a  $\beta^{(k,n_k)}$  has been found such that the tolerance of equation 5.24 is met. Set  $\beta^{(k,\text{opt})} := \beta^{(k,n_k)}$ .
- iii). (*outer iteration*) Repeat step ii). for  $k = 1, \dots, d + 1$  instead of  $k = 0$  using the initialization  $\beta^{(k,0)} = \beta^{(k-1,\text{opt})}$ .
- iv.) Set  $\beta^{\text{opt}} := \beta^{(d,\text{opt})}$ .

We conclude this subsection with the remark that the choice of the Newton method in algorithm 5.7.2 is by no means a crucial one. One could devise similar sequential algorithms involving other iterative root-finding techniques for multidimensional equation-systems.

Our experiments have shown that in the Gibbsian Model the (specific) Minkowski-functionals  $\hat{V}_j(\Theta)$  are rather sensitive to slight changes in the parameter vector  $\beta$  (see the results in section 5.9) which makes the estimation of a useful parameter set a difficult task. However, in the next section we present the so-called *Simulated Annealing*-method which can simulate from the Gibbsian distribution  $\pi_{G,\beta}$  without any parameter-estimation at all! This seems quite surprising at first glance but can be justified by the *equivalence of ensembles*-principle discussed in section 5.5. The advantages of such a non-parametric method are evident and sensitivity of the Minkowski-functionals with respect to the Gibbs-parameters becomes a negligible issue.

## 5.8 Simulated Annealing-Method

In this subsection we discuss a non-parametric method to simulate from the Gibbs-ensemble without knowing the Gibbsian-parameter-vector  $\beta$ . This can be justified by the *equivalence of ensembles* discussed in section 5.8. Hence for this method the problems one faces when applying the parameter estimation techniques from section 5.7 are completely irrelevant. We present the method as it is described in (Zhu et al. 2000) and call it *Simulating-Annealing-Method* in the following because it involves the

well-known *Simulated-Annealing* optimization algorithm adapted to our setting. The simulated annealing method for stochastic reconstruction has been extremely popular in recent years especially with the geo-physical and geo-statistical communities. Exemplarily we refer to (Talukdar, Torsaeter, Ioannidis and Howard 2002) for an example of three-dimensional chalk reconstruction from two-dimensional sections. Among all the simulation methods we have applied in our experiments and described in this thesis, the Simulated-Annealing-method has performed best (see results-section 5.9 for details). Another reason why we prefer the Simulated-Anncaling-Method to others is its flexibility. It can be used (at least in principle) to match an arbitrarily large number of image statistics and to include additional image features to be matched by the algorithm can be done very easily. Hence the method not only allows one to generate artificial Gibbsian structures with the same specific Minkowski-functionals as an original image but also with the same chord-length-distribution-function (definition 2.5.3), the same covariance-function (definition 2.5.1), the same pore-size-distribution (definition 2.5.4), the same Minkowski-functions (definition 2.5.5) and even combinations of these. However, in practice unfortunately, as the number of images features to be matched grows, the performance of the method quickly decreases. Before we describe how *Simulated Annealing* is used in our setting, we are to briefly recall how Simulated Annealing works in general.

Simulated Annealing owes its name and terminology to the physical annealing process used in metallurgy to grow crystals which exhibit as little defects as possible. Simulated Annealing starts with an arbitrary state of the system and sequentially performs small random alterations of the present state while forcing the sequence of visited states to converge to the optimum. To measure the distance of an arbitrary state from the (not necessarily unique) optimal state a distance-measure  $d$  is introduced. If the random alteration of a state reduces the distance from the optimum, it is accepted with high probability (or even for sure in some versions) and if the random alteration increases this distance it is accepted only with a small probability. At first glance it may seem unlogical to accept alterations increasing the distance from the optimum but this has the purpose to prevent the algorithm from getting trapped in a local minimum of the distance  $d$ . However, as the algorithm proceeds over time, the acceptance probability of such "bad" alterations successively gets smaller. The rate at which this happens is controlled by the so-called *cooling schedule*. In the end, the algorithm finishes hopefully with a state close enough to the optimal one. To be a little bit more precise, in a system with the optimal state  $s^{\text{opt}}$  (and thus  $d(s^{\text{opt}}) = 0$ ) the *acceptance probabilities* in the  $n$ -th step  $p_n(s, s', T)$  of a state  $s'$  when the system is currently in the state  $s$  are of the form:

$$p_n(s, s', T(n)) = \text{const} \cdot \text{Exp} \left( \frac{d(s) - d(s')}{T(n)} \right).$$

where the strictly decreasing sequence  $T(n)$  which we call the *temperature* - according to its significance in the physical model process - determines the predefined cooling schedule. For small  $n$  and thus large  $T(n)$  an increase of the distance to the optimum is accepted roughly with the same probability as a decrease of the same size. But as  $n$  increases and  $T(n)$  decreases, the probability of accepting a state switch with a decrease of the distance to the optimum becomes overwhelmingly more probable than an equally sized increase. It can be proved that for every finite optimization problem the probability that the Simulated Annealing algorithms ends up in the optimum approaches 1 if the cooling is performed infinitely slow. However, this theoretical result is of not much practical use,

since we have to find a sufficiently close solution in a reasonable time. Unfortunately in practice it is well possible that the convergence of the distance to the optimum approaches zero too slow to be of practical use or even that the algorithm gets trapped in local minima. The selection of an appropriate cooling schedule is crucial for the success of the method and although guidelines for the selection of a cooling schedule exist, a good cooling schedule can often not be found without some heuristic experimenting.

We now describe how the Simulated Annealing approach can be used to simulate from the Gibbsian Model with predetermined Minkowski-functionals. The key idea is that because of *ensemble-equivalence* (see section 5.5) we can simulate from the Julész-ensemble (definition 5.5.1) instead of the Gibbs-ensemble (definition 5.5.2). But the Julész-ensemble is simply the uniform distribution on the set  $\Omega_{\mathbb{Z}^d}(v)$  of all images on  $\mathbb{Z}^d$  whose specific Minkowski-functionals agree with the preset values  $v$ . Although this uniform distribution is completely parameter-free, it is still quite difficult to simulate from the Julész-ensemble because among all binary images over  $\mathbb{Z}^d$  the Julész-ensemble only has negligible volume. Please note further that ensemble-equivalence in the sense of section 5.5 only holds on the *infinite* lattice  $\mathbb{Z}^d$  but for practice we are bound to deal with finite lattices instead and the equivalence of ensembles does not hold. However, for large enough lattices  $\Gamma$  the use of the equivalence of ensembles can be justified.

To apply Simulated Annealing we need to define a distance  $d$  which measures the distance of a given image  $s \in \mathcal{S}$  from the Julész-ensemble. Any image in the Julész-ensemble can be considered as an optimum state. The most natural choice for  $d$  is the following:

$$d(s) = \sum_{j=0}^d \left| \hat{V}_j(s) - v_j \right|, \quad (5.27)$$

where again  $\hat{V}$  denotes the OM-estimators for the specific Minkowski-functionals from subsections 2.3.1, 2.3.2 and  $v$  is the vector of predetermined specific Minkowski-functionals. Of course, because we work on a finite lattice  $\Gamma$ , one could also work with the unspecific quantities  $\hat{V}$  and  $V$  instead or replace the  $\mathcal{L}^1$ -norm by some other reasonable distance (eg. the usual  $\mathcal{L}^2$ -norm). Our experiments have shown that different choices do not affect the performance of the method severely. From the theory of simulated annealing it is then clear that the sequence of distributions on  $\mathcal{S}$

$$q(s, T) = \frac{1}{Z(T)} \cdot \text{Exp} \left( -\frac{d(s)}{T} \right) = \frac{1}{Z(T)} \cdot \text{Exp} \left( -\frac{\sum_{j=0}^d |\hat{V}_j(s) - v_j|}{T} \right), \quad (5.28)$$

converges to (the finite version of) the Julész-ensemble if  $T \rightarrow 0$  sufficiently slow. As for the Gibbs-distribution,  $Z(T)$  is just a temperature-dependent normalizing factor. The distribution  $q(s, T)$  is perfectly amenable to Gibbs-sampling and hence the combination of Simulated Annealing and Gibbs-Sampling leads us to the following algorithm to simulate from the Julész-ensemble:

**Algorithm 5.8.1. (Simulated Annealing Method)** Consider a finite grid  $\Gamma = \{x_1, \dots, x_{|\Gamma|}\}$  whose gridpoints  $x_i$  are enumerated according to some predefined ordering, eg. row- or column-wise. To simulate a realization  $s \in \mathcal{S}$  of the Gibbsian Model  $\Theta$  with the OM-energy function and predetermined specific Minkowski-functionals  $v$ , one can use the following parameter-free procedure:

- i). Define and initialize the cooling schedule, eg. for  $n \in \mathbb{N}$  choose a decreasing real-valued function  $T(n)$  with the property  $T(n) \downarrow 0$  as  $n \rightarrow \infty$ . Further determine an approximation tolerance  $\epsilon > 0$ .
- ii). Initialize an image  $s^{(0,0)} \in \mathcal{S}$  arbitrarily, eg. as completely black  $s_i^{(0,0)} = 1, \forall x_i \in \Gamma$ , completely white  $s_i^{(0,0)} = 0, \forall x_i \in \Gamma$  or as white noise in which case  $s^{(0,0)}$  is a realization of the random set  $\Theta_0$  with  $P[\Theta_0(x_i) = 1] = P[\Theta_0(x_i) = 0] = 0.5$  independently,  $\forall x_i \in \Gamma$ .
- iii) For  $n = 0$  do an annealing step by generating for  $T := T(n)$  a realization  $s^{(n,|\Gamma|)} \in \mathcal{S} \sim q(\cdot, T)$  given in equation 5.28 by repeating the following steps for  $k \in \{0, \dots, |\Gamma| - 1\}$ :

1. Select at random uniformly a number  $i_k \in \{1, \dots, |\Gamma|\}$ .
2. Calculate the conditional probability

$$P \left[ \Theta(x_{i_k}) = s_{i_k}^{(n,k)} \mid \Theta(x_j) = s_j^{(n,k)}, j \neq i_k \right] = \frac{1}{1 + \text{Exp} \left( \frac{1}{T} \left[ \sum_{j=0}^d |\hat{V}_j(s^{(n,k)}) - v_j| - |\hat{V}_j(\tilde{s}^{(n,k)}) - v_j| \right] \right)},$$

where  $\tilde{s}^{(n,k)}$  is the image with  $\tilde{s}_{i_k}^{(n,k)} = 1 - s_{i_k}^{(n,k)}$  and  $\tilde{s}_j^{(n,k)} = s_j^{(n,k)}$  for  $j \neq i_k$ .

3. Generate  $U_k \sim \text{Unif}[0, 1]$  and if

$$U_k < P \left[ \Theta(x_{i_k}) = s_{i_k}^{(n,k)} \mid \Theta(x_j) = s_j^{(n,k)}, j \neq i_k \right]$$

set  $s_{i_k}^{(n,k+1)} := s_{i_k}^{(n,k)}$  and  $s_{i_k}^{(n,k+1)} := 1 - s_{i_k}^{(n,k)}$  otherwise.

4. Set  $s_j^{(n,k+1)} := s_j^{(n,k)}$  for all  $j \neq i_k$ .

- iv). Set  $s^{(n+1,0)} := s^{(n,|\Gamma|)}$ .

- v). Repeat steps iii). and iv). for  $n = 1, 2, 3, \dots, n_0 - 1$ , where

$$n_0 = \min \left\{ n \mid \sum_{j=0}^d |\hat{V}_j(s^{(n,0)}) - v_j| < \epsilon \right\}.$$

The output of the algorithm is the image  $s := s^{(n_0,0)}$ .

### Remark 5.8.2.

- 1.) When sweeping the image in step iii.) to simulate from  $q(\cdot, T)$  only the actual image  $s$  and its configuration vector  $h$  must be stored. There is no need to actually regenerate the whole image for each  $k$  as only a single pixel is possibly swapped when passing from  $s^{(n,k)}$  to  $s^{(n,k+1)}$ . During the sweep of the image if a pixel is swapped we only record the pixel swap in the actual image instead of copying the whole image and further record only *the changes* in the configuration vector instead of recalculating it for the new image. But this is easy and fast since per pixel swap at most  $2^{d+1}$  entries of the configuration vector are affected (either incremented or

decremented). This procedure was already discussed in detail for algorithm 5.6.2. Hence the recalculation of the Minkowski-functionals of an image with a single pixel changed is easy and quick. For the same reason there is no need to actually generate the image  $\tilde{s}$  in step iii.) to calculate the conditional probability.

- 2.) Instead of picking a pixel at random in step iii.), part 1, one could also sweep the image according to a predefined ordering scheme, eg. update pixels row by row or column by column.
- 3.) It is easy to generalize the procedure to match predefined values  $v$  of general image feature statistics  $\phi : \mathcal{S} \rightarrow \mathbb{R}$ . Assume that  $m$  general image feature statistics  $\phi_j, j \in \{1, \dots, m\}$  are given and we want to produce a Gibbsian realization  $s$  such that each of the  $m$  feature statistics takes for  $s$  a predetermined value  $v_j$ . Then we can apply algorithm 5.8.1 by simply replacing the distance  $d$  from equation 5.27 by:

$$d(s) = \sum_{j=0}^m \left| \phi_j(s) - v_j \right|.$$

It is therefore very easy to enlarge a given set of feature statistics to be matched since in the whole programming code only one element, namely the distance  $d$ , must be slightly altered. However, if one uses features  $\phi$  that are not locally computable and are not generated by a neighbourhood-potential for a reasonably small neighbourhood, we lose the Markov property and the calculation of the conditional probabilities in step iii), part 2, might be involved. In some cases one might even have to recalculate the statistic for the *whole* image after each pixel swap.

- 4.) The more statistics that are to be matched the higher the risk that the Simulating-Annealing-Algorithm gets trapped in some local optimum and thus the more crucial is a careful choice of the cooling schedule. It is important that the initial temperature  $T(0)$  is high enough to permit that the algorithm moves away from the starting configuration. At least for low temperatures it is also extremely important that the cooling is performed slowly enough. In some of our experiments we had to use cooling schedules with  $\Delta(T) = T(n+1) - T(n)$  as small as  $10^{-6}$  for high  $n$  in order to get close to the optimum.
- 5.) Algorithm 5.8.1 only does a *single* sweep of the image to simulate from the distribution  $q(\cdot, T)$  for fixed temperature. However as in algorithm 5.6.2 it can be wise to do *several* image sweeps to ensure that the equilibrium has been reached, ic. repeat step iii). not only  $|\Gamma|$  times for  $k \in \{0, \dots, |\Gamma| - 1\}$  but rather  $N \cdot |\Gamma|$  times for  $k \in \{0, \dots, N \cdot |\Gamma| - 1\}$ , where  $N \geq 1$  is a natural number.
- 6.) For the Gibbs-Sampling part of algorithm 5.8.1 the statements of remark 5.6.3 apply as well.

We would like to stress once more that among all the methods presented for the Gibbsian Model in this thesis, the Simulated-Annealing-Method 5.8.1 performed best once an appropriate cooling schedule had been chosen. The main disadvantage of the method is that it can be rather time-consuming if the number of sweeps  $N$  is large and the cooling is slow. It is therefore advisable to use some sort of *adaptive cooling schedule*: Start with a rather rough schedule and successively save some intermediate images. If the algorithm

gets trapped in a local minimum, step back to a higher temperature for which an intermediate image has been saved, refine the cooling schedule and restart the cooling starting from the saved image and the corresponding temperature. We conclude this chapter by showing some 2D-results achieved within the Gibbsian Model.

## 5.9 2D-Results

We now describe some two-dimensional results achieved with the Gibbsian Model and the Simulated-Annealing Method (algorithm 5.8.1). We have restricted ourselves to considering images of size  $256 \times 256$  pixels for reasons of computer time but conceptually the method applies also for 3D-images of arbitrary size. As a reference image we extracted from the coarse sand specimen (left image in figure 1.1) a  $256 \times 256$ -pixel-sized extract from the xy-cross-section at position  $z = 400$  shown as the top left image in figure 1.3. The reference image is shown as the upper left image of figure 5.1. The image on the right hand side in figure 5.1 shows a  $256 \times 256$ -pixel-sized artificial image with the same specific Minkowski functionals as the reference image generated according to algorithm 5.8.1 when starting from a white noise image. The algorithm was run with as many as 800'000 image sweeps which implies more than  $52 \cdot 10^9$  pixel updates! The two lower images show realizations of algorithm 5.8.1 when starting on an all black or all white image respectively. Beside the specific volume, surface and Euler-characteristic also the so-called specific *excess Euler characteristic*  $\bar{\chi}_{exc}$  has been matched (see the last paragraph preceding subsection 2.3.2). We recall that the excess Euler characteristic is the difference between the OM-Euler-characteristic estimator (equation 2.26) and the topological definition "number of convex vertices minus number of concave ones". The degree of precision with which a matching of the Minkowski-functionals can be achieved with the Simulated-Annealing-algorithm is stunning as can be seen in table 5.1.

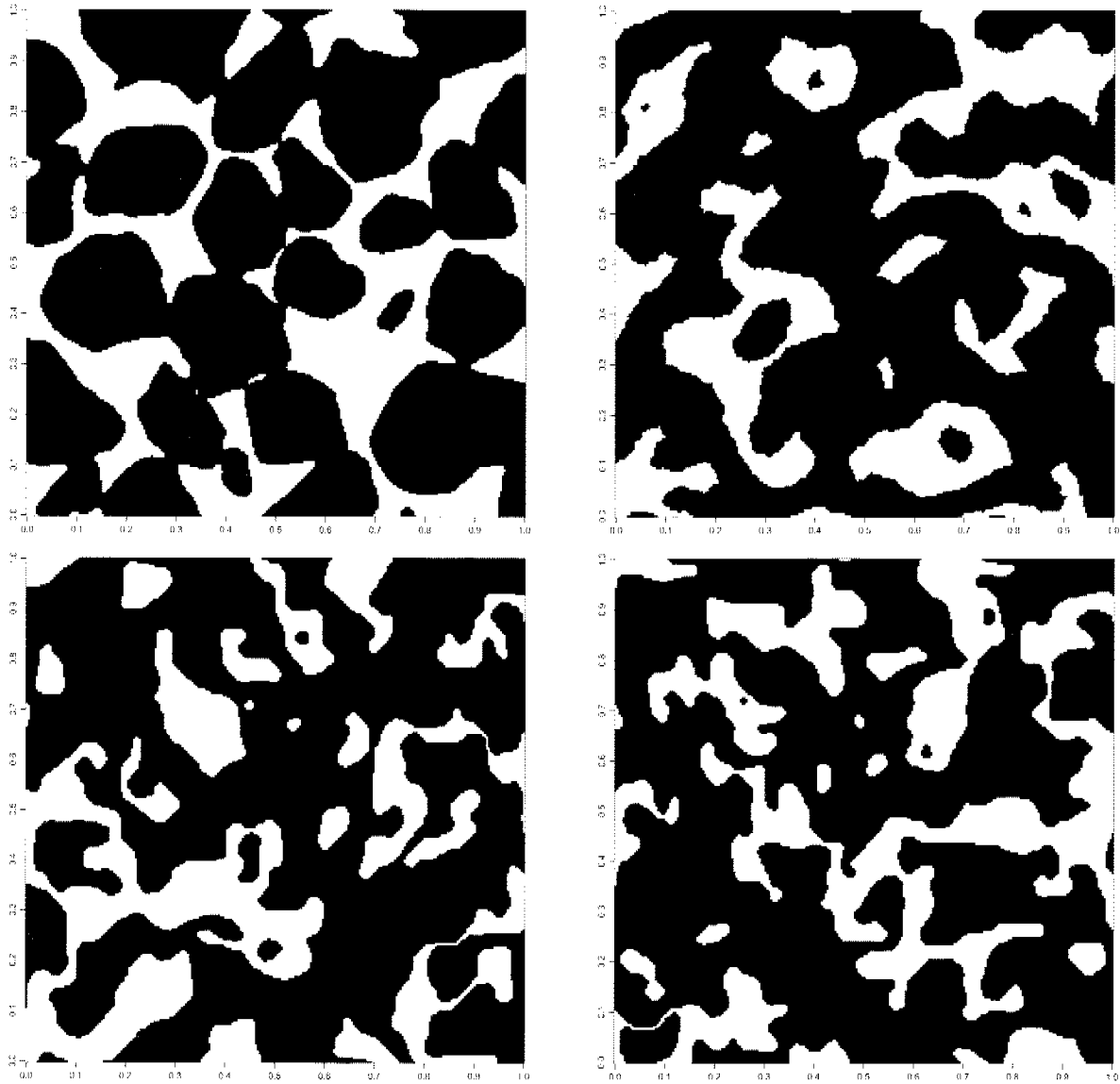
	$\bar{v}$	$\bar{s}$	$\bar{\chi}$	$\bar{\chi}_{exc}$
reference image	0.6978854287	0.0000461361	0.0468725915	0.0000000000
Gibbs realization	0.6978854287	0.0000461361	0.0468726525	0.0000000000

**Table 5.1:** *specific Minkowski functionals of the images in figure 5.1. The reference image is the left hand side original extract and the Gibbsian realization is on the right.*

It is quite obvious that the *indeterminacy*-phenomenon is also present within the Gibbsian Model. Although the (estimated) specific Minkowski Functionals are equal to as many as 6 significant digits, the images are look quite different. The Gibbsian realization appears to be too densely connected and the Euler characteristic is artificially increased with small black artifacts and the surface by frazzling the boundary. To reduce the number of artifacts one can add an additional additive term which contributes the number of such artifacts to the potential. Then the equilibrium states don't show any artifacts anymore. We have achieved such images in our experiments but the frazzling boundary remains. At first glance at least it seems difficult to enforce a smooth boundary since adding terms to control the boundary-smoothness to the potential is of little help because the realizations are generated on a finite lattice. Figure 5.2 shows how the 4 matched characteristics have evolved while algorithm 5.8.1 has been run to produce the Gibbsian realization on the right hand side of figure 5.1. The starting image was chosen to be a white noise image.

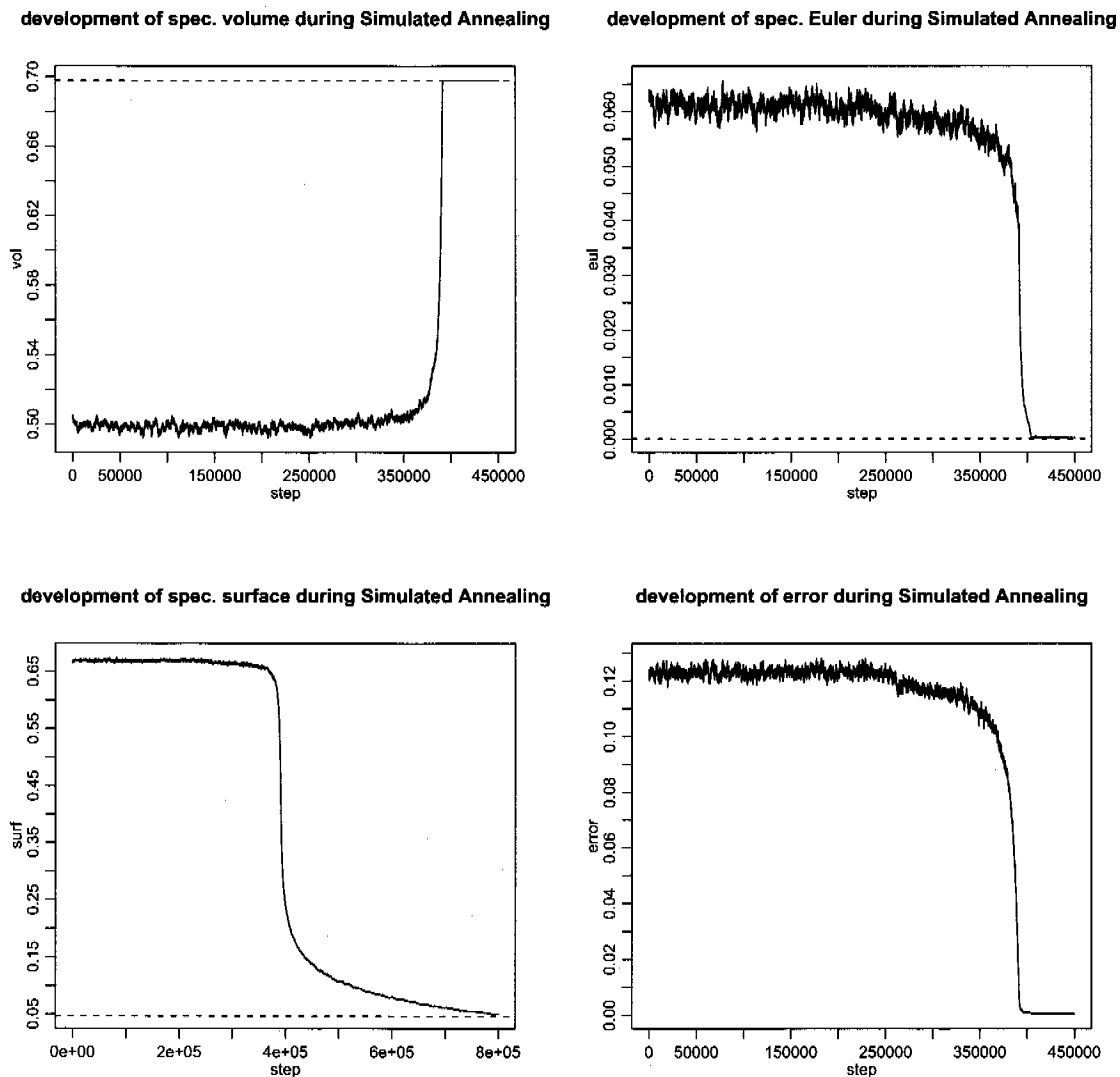
It is quite surprising how long the the algorithm "oscillates" around the starting values (for about 350'000 steps!). One is tempted to believe that the initial temperature of the





**Figure 5.1:** (upper left) 256×256-pixel-sized extract from a cross-section through the coarse sand specimen given in figure 1.1. (Others) 256×256-pixel-sized Gibbsian realizations generated with the Simulated-Annealing-Method which has the same specific Minkowski-functionals as the reference image on the upper left side with very high precision. Algorithm 5.8.1 was used with an initial temperature  $T_0 = 100$  with an adaptive cooling schedule whose smallest temperature difference near the equilibrium was chosen to be  $\Delta T = 10^{-6}$ . The starting images were different for the three realizations: (upper right) white noise, (lower left) all black, (lower right) all white image.

Simulated-Annealing-procedure was chosen to be much too high (about  $T_0 = 100$  in this case). However, our experiments have shown that this is not the case. When choosing significantly smaller starting temperature, the algorithm gets trapped in a local optimum even if the temperature is reduced in steps as tiny as  $\Delta T = 10^{-6}$ . After this long oscillating phase, the algorithm adjusts the specific volume away from the white noise value 0.5 to the predetermined one ( $\approx 0.7$ ) in only a handful steps and from this timepoint



**Figure 5.2:** Evolution of the Minkowski-functionals from a white noise image to the image with predefined Minkowski functionals on the right hand side of figure 5.1 during Simulated Annealing. Besides volume, surface and Euler-characteristic also the excess Euler-characteristic was matched (not shown). The last image shows the evolution of the  $\mathcal{L}_1$ -norm of the total error.

on only produces images with exactly the correct predetermined value for the specific volume! After some 400'000 steps it adjusts to the correct Euler characteristic and needs additional 400'000 until the specific surface is matched. This procedure of successively matching the characteristics without straying away from them anymore is quite characteristic for the Simulated-Annealing-Method. By considering only the first 250'000 steps say, one would be tempted to believe that the sequence of specific volume estimates stays stationary around 0.5. However, figure 5.2 impressively proves that this is wrong and emphasizes that much patience may be needed before the Simulated Annealing unfolds its power.

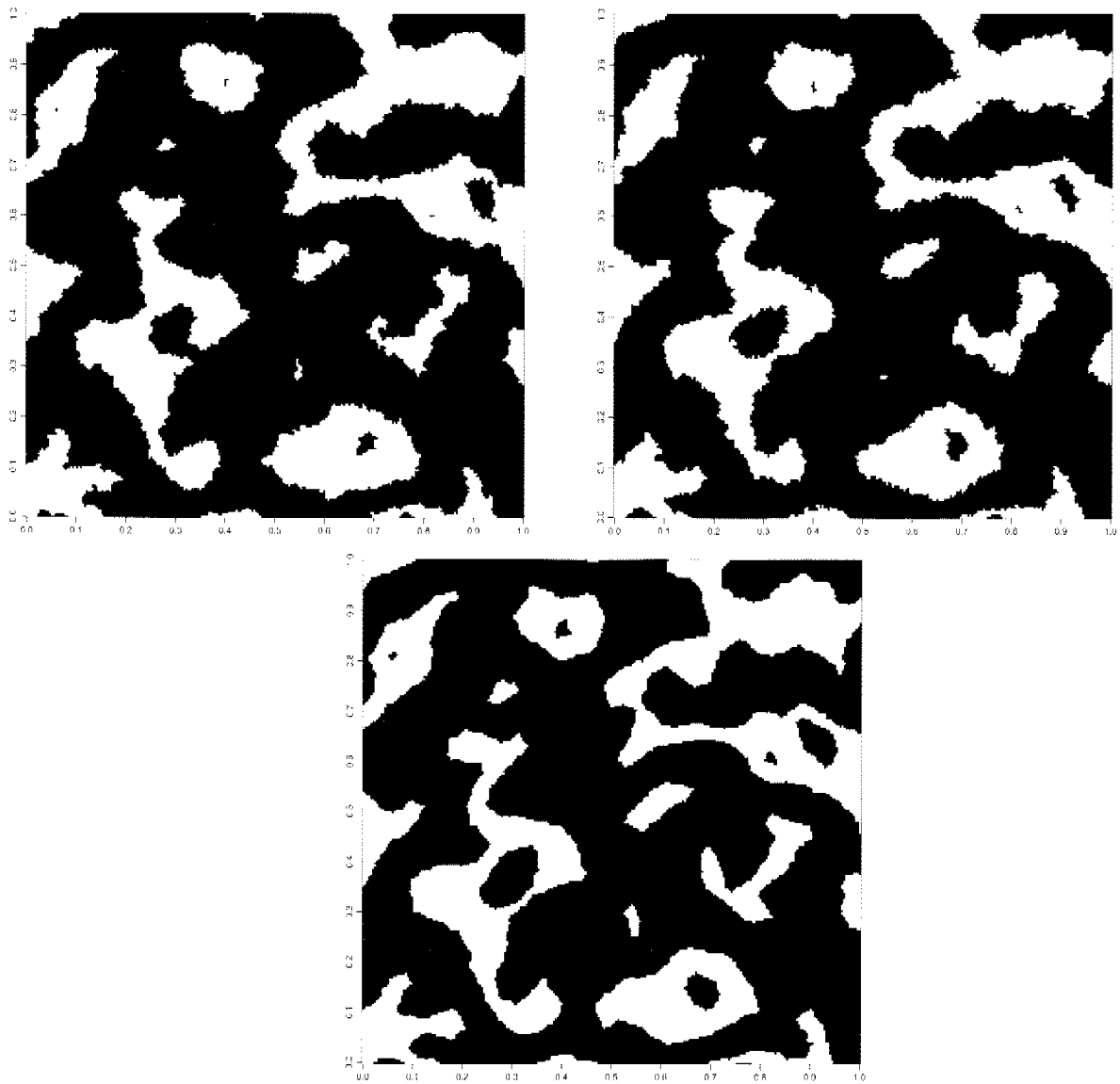
To illustrate the path that the Simulated-Annealing-Algorithm samples within the Julész-ensemble figure 5.3 shows the generated image on the right hand side of 5.1 together with two intermediate images of the sample path. The final image was produced after 850'000 steps of algorithm 5.8.1. The intermediate images were drawn after the 600'000 and 800'000 step of the same run of the algorithm respectively. We learn from figure 5.2 shows that at these stages of the algorithm volume and Euler-characteristic have already converged to their final values, whereas the surface is still incorrect for the 600'000-step image, but has already converged for the 800'000-step image. Figure 5.3 that the surface is finally adjusted (decreased from its too high initial value) by smoothing the boundary between the two phases of the image. The global structure of the image however, doesn't change at all between steps 600'000 and 850'000.

Figure 5.4 shows some Gibbsian realizations each of which besides the specific Minkowski functionals matches one additional image feature: The top right image matches the specific Minkowski functionals together with the two-point-covariance function (definition 2.5.1), the bottom left image matches the specific Minkowski functionals together with the chordlength-distribution (definition 2.5.3) and the bottom right image matches the whole Minkowski-functions (definition 2.5.5).

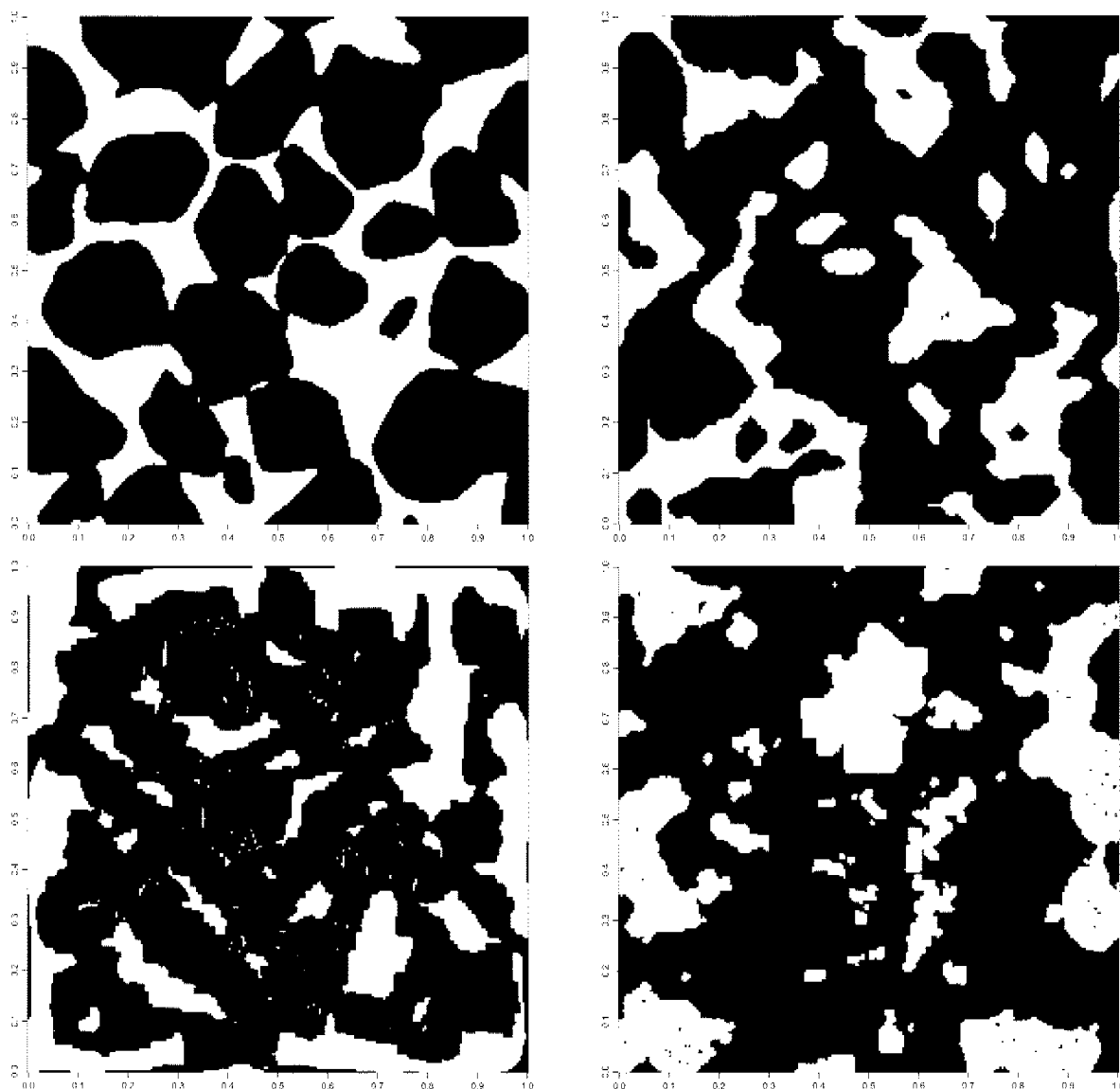
It becomes evident from figure 5.4 that also the matching of additional image features is by no means suffices to make the porous structure unique. An obvious defect of the model is that opposed to the Boolean Model for example the Gibbsian model doesn't produce any artificial entities that play the role of the real sand grains but only generates the global pattern. Thus it is very difficult to reproduce the granular shapes within the Gibbsian Model. Figure 5.5 shows how close the additional characteristics of the artificial images in figure 5.4 come to those of the reference image.

Figure 5.5 shows that it is hardest to match the Minkowski-function for the Euler characteristic since a single pixel flip may increase or decrease the unspecific Euler characteristic by as much as 1 unit. The same applies for the somewhat less important excess-Euler-characteristic. However, we are confident, that by investing even more time in algorithm 5.8.1 and using and even more cautious cooling schedules one would be able to remove these minor differences for large dilation radii as well. It is not necessary to consider large dilation radii (here: larger than 25) because evidently such images would be completely black.

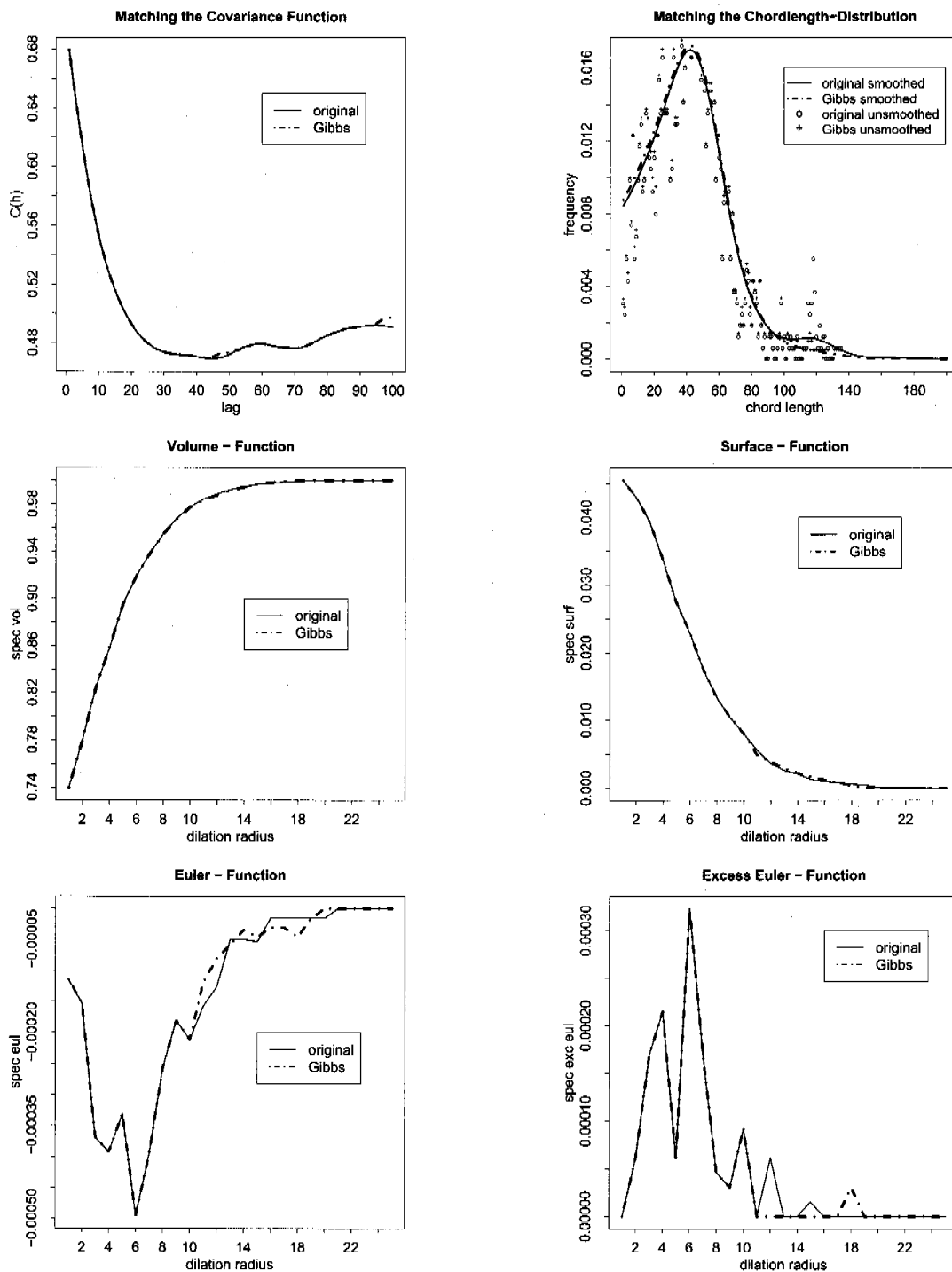
We conclude this chapter with a somewhat surprising image. The image in figure 5.6 also shares the same Minkowski functionals as our reference image. This image was generated by the pseudo-likelihood method (subsection 5.7.1). Its mass is concentrated almost completely in a centered black square with frayed boundary and some additional separated frayed black spots near the image boundary. This image illustrates a simple recipe how to generate images with predetermined Minkowski functionals and thus somehow heuristically explains the indeterminacy phenomenon: Place (almost) the whole black mass in a circle (to guarantee the lowest possible surface) centered in the middle of the image and add some additional small black spots with low volume to account for the Euler characteristic if it is positive. (Add white spots accordingly within the black circle if the Euler characteristic is negative.) Finally add boundary fringes to increase the surface to the desired value.



**Figure 5.3:** (upper left) Intermediate image after 600'000 steps of the Simulated-Annealing-Algorithm run which was used to produce the right hand side image in figure 5.1. (upper right) Intermediate image after 800'000 steps. (lower) The final image is given as the right hand side image in figure 5.1. The surface is decreased by smoothing out the boundary between the phases.



**Figure 5.4:** (top left)  $256 \times 256$ -pixel-sized extract from a cross-section through the coarse sand specimen given in figure 1.1 as a reference image. (top right)  $256 \times 256$ -pixel-sized Gibbsian realization with the same specific Minkowski-functionals and the same two-point-covariance function up to lag 100 as the reference image. (bottom left)  $256 \times 256$ -pixel-sized Gibbsian realization with the same specific Minkowski-functionals and the same chord-length-distribution up to chord-length 200 as the reference image. (bottom right)  $256 \times 256$ -pixel-sized Gibbsian realization with the same Minkowski-functions up to dilation radius 25 as the reference image. For the top right and the bottom left image additionally the appearance of black artifacts has been suppressed.



**Figure 5.5:** Additional characteristics matched for the images in figure 5.4. Shown is the characteristic for the reference image (top left in figure 5.4) together with the characteristic for the corresponding artificial image of figure 5.4. Shown are the two-point-covariance-function (top left), the chordlength-distribution (top right) and the Minkowski functions for selected dilation radii: volume (middle left), surface (middle right), Euler characteristic (bottom left) and excess Euler characteristic (bottom right).



**Figure 5.6:** *A surprising image which shares the same specific Minkowski-functionals as the reference image used throughout this section. This image was produced using the pseudo-likelihood-method (subsection 5.7.1)*





## Chapter 6

# Conclusion

Generating random porous media with predefined values for the Minkowski-functionals is possible with a high accuracy in both two and three dimensions as has been demonstrated in this thesis at least for the three simple models which have been investigated. Although theory is often straightforward, implementation in practice is far from trivial. This is mainly because of two reasons: First, estimating specific Minkowski functionals of a continuous porous structure from a discrete pixel image may be afflicted with considerable approximation error (see for example (Ohser et al. 2003) for the case of the Boolean Model). And second, the massive data sizes of high-resolution three-dimensional images easily overstrain many of present-day computers. This asks for sophisticated fast algorithms and makes memory management a challenging task.

We have demonstrated in this thesis that the Minkowski functionals on their own are by no means sufficient to summarize the geometrical information content of a porous image. There is still a large variety of porous structures which share the same specific Minkowski functionals but not the same optical appearance. This indeterminacy problem is still present even in two-dimensions, if one considers the Minkowski-functions instead of the Minkowski-functionals or if one fits additional characteristics such as the  $m$ -point-covariance-function and/or the chord-length-distribution (at least in the Gibbsian case where we studied it).

However, the artificial structures we generated exhibit quite similar flow and transport properties as their real world counterparts. Especially the three-dimensional Gaussian structures with matched specific Minkowski-characteristics are able to mimic the liquid flow through real sand soil although still being easily optically distinguishable from the real structure. For a discussion of the geophysical properties of our artificial Boolean structures see (Lehmann et al. n.d.). The even better flow properties of our Gaussian structures will be analyzed in a further paper by the same authors.

From a theoretical point of view we were able to clarify the asymptotical behaviour of the Ohser-Mücklich estimators if the lattice spacing tends to zero in the Thresholded Gaussian case and have also been able to quantify their bias. The main problem for analytical calculations within the Gaussian Model is that most integrals involving the multivariate Gaussian density cannot be analytically solved. This makes it extremely difficult to find explicit solutions in the Thresholded Gaussian model for all but the very simplest probability statements.



# Chapter 7

## Appendix

### 7.1 The On-Off-Markov-System

#### 7.1.1 Theory of the On-Off-Markov-System

An ordinary discrete *Markov chain* with finite state space  $\mathcal{S}$  consists of a sequence of random variables  $Z_0, Z_1, \dots$ , which is characterized by the property that  $\forall n \in \mathbb{N}, i, j, x_i \in \mathcal{S}$  we have:

$$\mathbb{P}[Z_n = j | Z_0 = x_0, \dots, Z_{n-2} = x_{n-2}, Z_{n-1} = i] = \mathbb{P}[Z_n = j | Z_{n-1} = i]. \quad (7.1)$$

A discrete stationary Markov chain is determined by the initial distribution  $q_i := \mathbb{P}[Z_1 = i], i \in \mathcal{S}$  and the transition probabilities  $q_{ij} := \mathbb{P}[Z_n = j | Z_{n-1} = i]$ .

A *continuous-time Markov chain* is a stochastic process  $\{Z(t), t \geq 0\}$  indexed by the non-negative real line. The defining property is the continuous analogon of equation 7.1.  $\forall n \in \mathbb{N}, 0 \leq t_1 < \dots < t_n, x_i \in \mathcal{S}$  we have:

$$\mathbb{P}[Z(t_n) = j | Z(t_1) = x_1, \dots, Z(t_{n-2}) = x_{n-2}, Z(t_{n-1}) = i] = \mathbb{P}[Z(t_n) = j | Z(t_{n-1}) = i]. \quad (7.2)$$

The probabilities  $P_{ij}(t_n - t_{n-1}) := \mathbb{P}[Z(t_n) = j | Z(t_{n-1}) = i]$  involved in the above statement are again called *transition probabilities*. Because of the stationarity of the process  $Z(t)$  we have that for arbitrary timepoints  $t, s \geq 0$ :

$$\mathbb{P}[Z(t+s) = j | Z(t) = i] = P_{ij}(s) \quad (7.3)$$

is independent of  $t$ . For such a continuous Markov chain  $\{Z(t), t \geq 0\}$  defined by equations 7.2 and 7.3 a basic construction can be given. Start out with a *discrete* Markov chain  $\{Z_0, Z_1, \dots\}$  with finite state space  $\mathcal{S}$ , initial distribution  $\{q_i\}$  and transition matrix  $(Q_{ij})_{ij} := q_{ij}$ . Furthermore  $E_0, E_1, \dots$  is a sequence of independent standard exponentially distributed random variables and we are given a finite set of non-negative values  $\{\lambda(i), i \in \mathcal{S}\}$ . Now the continuous Markov chain  $\{Z(t), t \geq 0\}$  is constructed as follows:

- Set  $T_0 := 0$ .
- For  $n \geq 0$ : Define the sojourn-time in state  $Z_n$  as  $(\Delta T)_n \sim \text{Exp}(\lambda(Z_n))$ , set the jump time  $T_{n+1} := T_n + (\Delta T)_n$  and let  $Z(t) := Z_n$  for  $T_n \leq t < T_{n+1}$ .

To make the chain identifiable we always assume  $q_{ii} = 0 \forall i \in \mathcal{S}$ . in the following. It can be shown (see Resnick 1992, Chapter 5.2.1) that the so-constructed continuous process indeed has the properties of equations 7.2 and 7.3.

The local transitional behaviour of the continuous Markov chain  $\{Z(t), t \geq 0\}$  is given by the so-called *generator matrix*  $A$  which is defined as:

$$A_{ij} := \begin{cases} -\lambda(i) & \text{if } i = j, \\ \lambda(i) \cdot q_{ij} & \text{if } i \neq j. \end{cases} \quad (7.4)$$

It's easy to see (Resnick 1992, Prop. 5.4.3) that

$$\frac{\partial P}{\partial t}(0) = A \quad \text{and} \quad \frac{\partial P}{\partial t} = AP(t), \quad (7.5)$$

where  $P(t)$  is the matrix  $(P_{ij}(t))$ . Hence:

$$P_{ij}(h) = A_{ij} \cdot h + o(h) \quad \forall i, j \in \mathcal{S}.$$

Of course the probabilities  $P[Z(t) = i]$  depend on the initial distribution  $\{q_i\}$  and the continuous Markov chain is not stationary in general. However, if we manage to find a distribution  $\eta$  on  $\mathcal{S}$  which satisfies  $\eta^T P(t) = \eta^T \forall t > 0$  then obviously we have  $P[Z(t) = i] = \eta(i), \forall t \geq 0$  and thus stationarity of  $\{Z(t), t \geq 0\}$ . Such an  $\eta$  is called a *stationary distribution* of the continuous Markov chain. The existence of a unique stationary distribution is granted if the underlying discrete Markov chain is both *irreducible* (no matter from which state the chain starts with positive probability it reaches any other state in finite time) and *recurrent* (no matter from which state the chain starts with probability one it comes back to the same state in finite time). In this case  $\eta$  is a solution of:

$$\eta^T A = 0. \quad (7.6)$$

This statement is proved for example in (Resnick 1992, Theorem 5.5.3).

If we want to model a one-dimensional binary image with a stationary continuous Markov chain, our state space  $\mathcal{S}$  only consists of two elements: 1 (image or solid phase) and 0 (complement or pore phase). From the construction of the stationary continuous Markov chain it is clear that once chain enters one of the two phases  $i \in \{0, 1\}$  it stays there for a random time which is exponentially distributed with mean  $1/\lambda(i)$ , then jumps to the other state  $1 - i$  and so forth. This very simple Markovian dynamics is widely used in literature and known as the *on-off-system*. Obviously the discrete transition matrix  $Q$  and the generator  $A$  for the on-off-system are as follows:

$$Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{pmatrix}.$$

It is then simple to calculate the continuous transition matrix  $P(t)$  for the on-off-system using the relationships from equation 7.5. The continuous transition probabilities must fulfill ( $' = \partial/\partial t$ ):

$$\begin{aligned} P'_{00}(t) &= \lambda_0 \cdot (P_{10}(t) - P_{00}(t)) \\ P'_{10}(t) &= \lambda_1 \cdot (P_{00}(t) - P_{10}(t)), \end{aligned}$$

which can be recombined to:

$$\lambda_1 P'_{00}(t) + \lambda_0 P'_{10}(t) = 0.$$

Integrating leads to:

$$\lambda_1 P_{00}(t) + \lambda_0 P_{10}(t) = \lambda_1,$$

and hence we obtain an ordinary linear initial value problem of the first order for  $P_{00}$ :

$$P'_{00}(t) = \lambda_1(1 - P_{00}(t)) - \lambda_0 P_{00} \text{ with } P_{00}(0) = 1,$$

which is trivial to solve and results in the following transition matrix:

$$P(t) = \frac{1}{\lambda_0 + \lambda_1} \begin{pmatrix} \lambda_1 + \lambda_0 \cdot e^{-(\lambda_0 + \lambda_1)t} & \lambda_0 - \lambda_0 \cdot e^{-(\lambda_0 + \lambda_1)t} \\ \lambda_1 - \lambda_1 \cdot e^{-(\lambda_0 + \lambda_1)t} & \lambda_0 + \lambda_1 \cdot e^{-(\lambda_0 + \lambda_1)t} \end{pmatrix}. \quad (7.7)$$

Since the on-off-system is of course both irreducible and recurrent we also have a stationary distribution  $\eta$  which (according to equation 7.6) is determined by:

$$\begin{aligned} \eta_0 + \eta_1 &= 1 \\ \lambda_1 \cdot \eta_1 - \lambda_0 \cdot \eta_0 &= 0, \end{aligned}$$

And thus the stationary distribution is:

$$\eta_0 = \frac{\lambda_1}{\lambda_0 + \lambda_1}, \quad \eta_1 = \frac{\lambda_0}{\lambda_0 + \lambda_1}. \quad (7.8)$$

Because the stationary distribution is also the limiting distribution  $\eta_j = \lim_{t \rightarrow \infty} P_{ij}(t)$  in this case, this can also be seen directly from equation (7.7).

Since according to the discussion preceding equation 4.29 we only need to count phase transitions, the Ohser-Mücklich estimator for the specific surface area  $\bar{s}$  in the on-off-system is easily described. Let us assume that the one-dimensional-grid consists of  $n$  points  $x_0 = 0, \dots, x_n = n \cdot h_n$  in the interval  $[0, n \cdot h_n]$  where we can observe the stationary continuous on-off-process  $Z$  and the distance between two consecutive points is  $h_n$ . Then:

$$\hat{\hat{s}}_n(Z) = \frac{1}{n \cdot h_n} \sum_{i=1}^n (\mathbf{1}_{Z(ih_n)=1} \cdot \mathbf{1}_{Z((i+1)h_n)=0} + \mathbf{1}_{Z(ih_n)=0} \cdot \mathbf{1}_{Z((i+1)h_n)=1}). \quad (7.9)$$

Instead of investigating the properties of  $\hat{\hat{s}}_n(Z)$  and prove its asymptotic normality we will do this for the one-dimensional more general class of surface estimators from definition 4.4.1. However for the sake of simplicity in the following we only consider those estimators of the type given in equation 4.32 for which the lags  $j$  and  $l$  are assumed to be positive instead of  $-m \leq j, l \leq m$ . This restriction can be made without losing anything important but has the disadvantage that the estimator  $\hat{\hat{s}}_n(Z)$  from equation 7.9 above is not a member of this restricted class of general surface estimators we consider in the following (because  $j = 1$  and  $l = -1$ ). However, it is obvious that all the statements we make in the next subsection can be proved analogously. The exact results for  $\hat{\hat{s}}_n(Z)$  are provided for completeness in section 7.3.4.

### 7.1.2 Asymptotic Normality for General Surface Estimators in the On-Off-Markov-System

In this subsection we collect some important properties and prove asymptotic normality for the general surface estimators of the type 4.32 for the one-dimensional on-off-Markov-Model. For the variance we use the procedure outlined in subsection 4.5.1. The one-dimensional version of the general surface estimator is according to equation 4.32:

$$\hat{s}_n = -\frac{1}{nh_n} \sum_{j=1}^m \beta_j \sum_{i=1}^n \mathbf{1}_{\{Z(ih_n)=1, Z((i+j)h_n)=0\}}, \quad \sum_{j=1}^m j\beta_j = -2. \quad (7.10)$$

We are now ready to collect the results about the estimator 7.10 in the following theorem which is the main result of this subsection:

**Theorem 7.1.1. (Asymptotic normality of the general surface estimator)** Consider the estimator defined in 7.10 in the on-off-Markov-Model introduced in subsection 7.1.1. Then the following statements hold true:

i). The specific surface area  $\bar{s}$  for  $Z$  is:

$$\bar{s}(Z) = 2 \cdot \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1}.$$

ii). The bias of  $\hat{s}_n$  is:

$$\mathbf{E}[\hat{s}_n] - s = h_n \cdot \frac{\lambda_0 \lambda_1}{2} \sum_{j=1}^m \beta_j \cdot j^2 + \mathcal{O}(h_n^2).$$

iii). The variance of  $\hat{s}$  is:

$$\text{Var} \hat{s}_n \sim \frac{1}{nh_n} \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l \left[ 2 \sum_{k=0}^{j-1} \min(j-k, l) - \min(l, j) \right] - \frac{8\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \right).$$

iv) For  $n \rightarrow \infty, h \downarrow 0$  such that  $n \cdot h_n \rightarrow \infty$  the distribution of  $\hat{s}$  is asymptotically normal. Especially for the rate  $h_n \sim n^{-1/3}$  expectation and variance are stabilized and we have that:

$$n^{\frac{1}{3}} \cdot (\hat{s}_n - s) \xrightarrow{d} \mathcal{N} \left( \frac{\lambda_0 \lambda_1}{2} \sum_{j=1}^m \beta_j \cdot j^2, \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l \left[ 2 \sum_{k=0}^{j-1} \min(j-k, l) - \min(l, j) \right] - \frac{8\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \right).$$

**Proof. (Theorem 7.1.1)**

ad i):

The two-point covariance function  $C(h)$  for the on-off-system is:

$$C(h) = \mathbf{P}[Z(x) = 1, Z(x+h) = 1] = \mathbf{P}[Z(x+h) = 1 | Z(x) = 1] \cdot \mathbf{P}[Z(x) = 1] = P_{11}(h) \cdot \eta_1.$$

Taking the derivative with respect to  $h$  yields:

$$\frac{\partial C}{\partial h}(0) = \frac{\partial P_{11}}{\partial h}(0) \cdot \eta_1 = -\lambda_1 \cdot \eta_1,$$

where we have used  $P'(0) = A$  from equation 7.5. Applying Theorem 2.5.2 gives the desired result:

**ad ii):**

For the expectation of  $\hat{s}$  we have because of stationarity:

$$\begin{aligned} \mathbf{E}[\hat{s}] &= -\frac{1}{h_n} \sum_{j=1}^m \beta_j \mathbf{P}[Z(0) = 1, Z(jh_n) = 0] = -\frac{1}{h_n} \sum_{j=1}^m \beta_j \eta_1 P_{10}(jh_n) \\ &= -\frac{1}{h_n} \frac{\lambda_0 \lambda_1}{(\lambda_0 + \lambda_1)^2} \sum_{j=1}^m \beta_j \left(1 - e^{-(\lambda_0 + \lambda_1)jh_n}\right) = -\frac{1}{h_n} \frac{\lambda_0 \lambda_1}{(\lambda_0 + \lambda_1)^2} \times \\ &\quad \sum_{j=1}^m \beta_j \left[ (\lambda_0 + \lambda_1)jh_n - \frac{1}{2}(\lambda_0 + \lambda_1)^2 j^2 h_n^2 + \mathcal{O}(h_n^3) \right] \\ &= \frac{2\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} + h_n \cdot \frac{\lambda_0 \lambda_1}{2} \sum_{j=1}^m \beta_j j^2 + \mathcal{O}(h_n^2), \end{aligned}$$

Using i) above we see that  $s = \frac{2\lambda_0 \lambda_1}{\lambda_0 + \lambda_1}$  and thus the claim follows.

**ad iii):**

From equation 4.33 we deduce for the variance:

$$\begin{aligned} \text{Var}(\hat{s}) &= \\ &\frac{1}{nh_n^2} \sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l \sum_{k=-n}^n \left(1 - \frac{|k|}{n}\right) \left[ \mathbf{P}[Z(0) = 1, Z(jh_n) = 0, Z(kh_n) = 1, Z((k+l)h_n) = 0] \right. \\ &\quad \left. - \mathbf{P}[Z(0) = 1, Z(jh_n) = 0] \mathbf{P}[Z(0) = 1, Z(lh_n) = 0] \right] \end{aligned} \quad (7.11)$$

Let us consider first the infinite parts of this sum. Making use of the expressions for the transition probabilities in the on-off-system given in equation 7.7 we can see:

$$\begin{aligned}
& \sum_{k=j}^n \left[ \mathbb{P}[Z(0) = 1, Z(jh_n) = 0, Z(kh_n) = 1, Z((k+l)h_n) = 0] \right. \\
& \quad \left. - \mathbb{P}[Z(0) = 1, Z(jh_n) = 0] \mathbb{P}[Z(0) = 1, Z(lh_n) = 0] \right] \\
&= \sum_{k=j}^n [\eta_1 P_{10}(jh_n) P_{01}((k-j)h_n) P_{10}(lh_n) - \eta_1^2 P_{10}(jh_n) P_{10}(lh_n)] \\
&= \sum_{k=j}^n \eta_1 P_{10}(jh_n) P_{10}(lh_n) (P_{01}((k-j)h_n) - \eta_1) \\
&\sim -h_n^2 \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^2} \cdot jl \sum_{k=j}^n e^{-(k-j) \cdot (\lambda_0 + \lambda_1) \cdot h_n} \\
&\sim -h_n \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^2} \cdot jl \int_0^\infty e^{-(\lambda_0 + \lambda_1) \cdot t} dt = -h_n \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \cdot jl.
\end{aligned}$$

And similarly by symmetry:

$$\begin{aligned}
& \sum_{k=-n}^{-l} \left[ \mathbb{P}[Z(0) = 1, Z(jh_n) = 0, Z(kh_n) = 1, Z((k+l)h_n) = 0] \right. \\
& \quad \left. - \mathbb{P}[Z(0) = 1, Z(jh_n) = 0] \mathbb{P}[Z(0) = 1, Z(lh_n) = 0] \right] \\
&= \sum_{k=-n}^{-l} [\eta_1 P_{10}(lh_n) P_{01}((-k-l)h_n) P_{10}(jh_n) - \eta_1^2 P_{10}(jh_n) P_{10}(lh_n)] \\
&\sim -h_n \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^2} \cdot jl \int_{-\infty}^0 e^{(\lambda_0 + \lambda_1) \cdot t} dt = -h_n \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \cdot jl,
\end{aligned}$$

The finite parts of the sum can be treated as follows:

$$\begin{aligned}
& \sum_{k=0}^{j-1} \left[ \mathbb{P}[Z(0) = 1, Z(jh_n) = 1, Z(kh_n) = 0, Z((k+l)h_n) = 0] \right. \\
& \quad \left. - \mathbb{P}[Z(0) = 1, Z(jh_n) = 0] \mathbb{P}[Z(0) = 1, Z(lh_n) = 0] \right] \\
&= \sum_{k=0}^{j-1} [\eta_1 P_{11}(kh_n) P_{10}(\min(j-k, l)h_n) P_{00}(|k+l-j|h_n) - \eta_1^2 P_{10}(jh_n) P_{10}(lh_n)] \\
&\sim h_n \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \sum_{k=0}^{j-1} \min(j-k, l) + \mathcal{O}(h_n^2).
\end{aligned}$$

And finally:



$$\begin{aligned}
& \sum_{k=-l+1}^0 \left[ \mathbb{P}[Z(0) = 1, Z(jh_n) = 0, Z(kh_n) = 1, Z((k+l)h_n) = 0] \right. \\
& \quad \left. - \mathbb{P}[Z(0) = 1, Z(jh_n) = 0] \mathbb{P}[Z(0) = 1, Z(lh_n) = 0] \right] \\
& = \sum_{k=-l+1}^0 [\eta_1 P_{11}(-kh_n) P_{10}(\min(l+k, j)h_n) P_{00}(|k+l-j|h_n) - \eta_1^2 P_{10}(jh_n) P_{10}(lh_n)] \\
& \sim h_n \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \sum_{k=-l+1}^0 \min(l+k, j) + \mathcal{O}(h_n^2).
\end{aligned}$$

Hence we can conclude for the variance:

$$\begin{aligned}
& \text{Var } \hat{s} \sim \\
& \frac{1}{nh_n} \sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \left[ \sum_{k=0}^{j-1} \min(j-k, l) + \sum_{k=-l+1}^0 \min(l+k, j) - \min(l, j) \right] - \frac{2\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \cdot jl \right) \\
& = \frac{1}{nh_n} \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l \left[ 2 \sum_{k=0}^{j-1} \min(j-k, l) - \min(l, j) \right] - \frac{8\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \right),
\end{aligned}$$

which is what we wanted to show, where we have used in the last step that  $\sum_{j=1}^m \sum_{l=1}^m \beta_j \beta_l jl = \left( \sum_{j=1}^m j \beta_j \right)^2 = \frac{d^2 b_d^2}{b_{d-1}^2}$ .

**ad iv)**

We now turn to the proof of asymptotic normality for  $\hat{s}$  in the on-off system. To stabilize expectation and variance of  $a_n \cdot (\hat{s} - s)$  the following conditions must hold:

$$\begin{aligned}
& \text{Var}[\hat{s}] \xrightarrow{n \rightarrow \infty} 0 \iff n \cdot h_n \xrightarrow{n \rightarrow \infty} \infty \\
& \mathbf{E} [a_n(\hat{s} - s)] \xrightarrow{n \rightarrow \infty} c_1 \iff a_n \cdot h_n \xrightarrow{n \rightarrow \infty} c_1, \\
& \text{Var} (a_n(\hat{s} - s)) \xrightarrow{n \rightarrow \infty} c_2 \iff \frac{a_n^2}{n \cdot h_n} \xrightarrow{n \rightarrow \infty} c_2,
\end{aligned}$$

where  $c_1$  and  $c_2$  are some nonzero constants. But then:

$$c_2 \sim \frac{a_n^2}{\frac{c_1}{a_n} n} \iff \frac{a_n^3}{n} \sim c_1 c_2 \iff a_n \sim n^{\frac{1}{3}},$$

and thus  $h_n \sim n^{-\frac{1}{3}}$ .

Of course one can also consider the cases where  $c_1$  or  $c_2$  or both are zero. For  $c_1 = 0$  we end up with a *centered* limiting normal distribution and for  $c_2 = 0$  the linear combination  $a_n(\hat{s} - s)$  converges to a deterministic constant.

For the rest of the proof we note that instead of considering the estimator of equation 7.10 without loss of generality we may consider the following simplified estimator:

$$\hat{s}_n = \frac{1}{nh_n} \sum_{i=1}^n \mathbf{1}_{E_{i,n}}, \quad (7.12)$$

where  $E_{i,n} = \{Z(ih_n) \neq Z((i+1)h_n)\}$  which basically means that we neglect the summation over different lags  $j$  for simplicity. In the following we consider the quantity  $S_n$  which according to iii). is a centered version of the estimator in equation 7.12 with constant variance  $\sigma_n^2$ :

$$S_n := \frac{1}{\sqrt{nh_n}} \sum_{i=1}^n (\mathbf{1}_{E_{i,n}} - P[E_{i,n}]) = \frac{1}{\sqrt{nh_n}} \sum_{i=1}^n \xi_{i,n}, \quad (7.13)$$

where we introduce the abbreviation  $\xi_{i,n} := \mathbf{1}_{E_{i,n}} - P[E_{i,n}]$ . This last sum is now divided into blocks  $X_{n,i}$  to which we want to apply Theorem 4.5.4. To do that we define the series of block limits  $j_n := \lfloor \frac{1}{h_n} \rfloor$  and have up to negligible boundary effects the following equation:

$$S_n = \frac{1}{\sqrt{nh_n}} \sum_{i=1}^n \xi_{i,n} = \sum_{i=1}^{\lfloor nh_n \rfloor} \underbrace{\sum_{j=1}^{j_n} \xi_{(i-1)j_n+j,n}}_{:=X_{n,i}} \frac{1}{\sqrt{nh_n}}.$$

We first show that the triangular array  $\{X_{n,i}, 1 \leq i \leq nh_n\}$  is  $\phi$ -mixing and thus also strongly mixing. For this we consider events  $A_1 \in \sigma(X_{n,i}, i \leq t)$  and  $A_2 \in \sigma(X_{n,i}, i \geq t+k)$  of the form:

$$\begin{aligned} A_1 &= \{Z(t_0 j_n h_n) = i_0, \dots, Z(t_l j_n h_n) = i_l, t_s \leq t+1 \text{ for } s \in 1, \dots, l\} \\ A_2 &= \{Z(t_{l+1} j_n h_n) = i_{l+1}, \dots, Z(t_{l+m} j_n h_n) = i_{l+m}, t_s \geq t+k \text{ for } s \in l+1, \dots, l+m\}. \end{aligned}$$

Now we can follow (Ibragimov 1971, Chapter 19). For  $k$  large enough:

$$\begin{aligned} P[A_1 \cap A_2] &= \\ P[A_1] \cdot P_{i_l i_{l+1}}((t_{l+1} - t_l)j_n h_n) P_{i_{l+1} i_{l+2}}((t_{l+2} - t_{l+1})j_n h_n) \dots P_{i_{l+m-1} i_{l+m}}((t_{l+m} - t_{l+m-1})j_n h_n) \\ P[A_1] P[A_2] &= P[A_1] \eta_{i_l i_{l+1}} P_{i_{l+1} i_{l+2}}((t_{l+2} - t_{l+1})j_n h_n) \dots P_{i_{l+m-1} i_{l+m}}((t_{l+m} - t_{l+m-1})j_n h_n) \end{aligned}$$

Hence:

$$\frac{|P[A_1 \cap A_2] - P[A_1] P[A_2]|}{P[A_1]} \leq |P_{i_l i_{l+1}}((t_{l+1} - t_l)j_n h_n) - \eta_{i_l i_{l+1}}| \leq C \rho^{k-1},$$

independent of  $n$  where  $C$  and  $0 < \rho < 1$  are constants and in the last inequality we have used  $|t_{l+1} - t_l| \geq k-1$  and applied *Markov's Theorem* (see for example (Resnick 1992, Chapter 2.13)). However, the last inequality can also be seen directly from equation 7.7. Hence we can conclude  $\phi_{n,k} \leq C \rho^k$  and because the right hand side is independent of  $n$  also:

$$\phi_k = \sup_n \phi_{n,k} \leq C \rho^k \xrightarrow{k \rightarrow \infty} 0,$$

which establishes the  $\phi$ -mixing-condition.

Next we must show  $\lim_{k \rightarrow \infty} \rho_k < 1$ . Because of the fundamental relationship  $\alpha_k \leq \rho_k \leq 2\pi\alpha_k$  for strictly stationary random fields developed in (Bradley 1993) this follows from:

$$\alpha_k = \sup_n \alpha_{n,k} \leq \sup_n \alpha((k-1)j_n h_n) \leq \alpha(k-1) \xrightarrow{k \rightarrow \infty} 0, \quad (7.14)$$

hence we even have  $\lim_{k \rightarrow \infty} \rho_k = 0$ . For the second moments of the  $X_{n,i}$ , we see using the calculations in iii):

$$\begin{aligned} \mathbf{E} [X_{n,i}^2] &= \frac{1}{nh_n} \sum_{j,l=1}^{j_n} \text{Cov} (\xi_{(i-1)j_n+j,n}, \xi_{(i-1)j_n+l,n}) \\ &= \frac{1}{nh_n} \sum_{j,l=1}^{j_n} (\mathbf{P} [E_{(i-1)j_n+j,n} \cap E_{(i-1)j_n+l,n}] - \mathbf{P} [E_{(i-1)j_n+j,n}] \mathbf{P} [E_{(i-1)j_n+l,n}]) \\ &\sim \frac{1}{nh_n} \cdot j_n h_n = \mathcal{O} \left( \frac{1}{nh_n} \right). \end{aligned}$$

And thus also:

$$\sup_n \frac{1}{\sigma_n^2} \sum_{i=1}^{\lfloor nh_n \rfloor} \mathbf{E} [X_{n,i}^2] = \sup_n \text{const.} \cdot \lfloor nh_n \rfloor \frac{1}{nh_n} \sim \mathcal{O}(1) < \infty, \quad (7.15)$$

which establishes condition i) of Theorem 4.5.4.

For condition ii) of theorem 4.5.4 it turns out that we need to control the fourth moment  $\mathbf{E} [X_{n,i}^4]$ . We have:

$$\begin{aligned} \mathbf{E} [X_{n,i}^4] &= \frac{1}{n^2 h_n^2} \sum_{j,k,l,m=1}^{\lfloor j_n \rfloor} \mathbf{E} [\xi_{(i-1)j_n+j,n} \xi_{(i-1)j_n+k,n} \xi_{(i-1)j_n+l,n} \xi_{(i-1)j_n+m,n}] \\ &= \frac{1}{n^2 h_n^2} \left( \sum_{j,k,l,m=1}^{\lfloor j_n \rfloor} \mathbf{P} [E_{(i-1)j_n+j,n} \cap E_{(i-1)j_n+k,n} \cap E_{(i-1)j_n+l,n} \cap E_{(i-1)j_n+m,n}] \right. \\ &\quad + 4 \lfloor j_n \rfloor \cdot \mathbf{P} [E_{1,n}] \sum_{j,k,l=1}^{\lfloor j_n \rfloor} \mathbf{P} [E_{(i-1)j_n+j,n} \cap E_{(i-1)j_n+k,n} \cap E_{(i-1)j_n+l,n}] \\ &\quad + 6 \lfloor j_n \rfloor^2 \mathbf{P} [E_{1,n}]^2 \sum_{j,k=1}^{\lfloor j_n \rfloor} \mathbf{P} [E_{(i-1)j_n+j,n} \cap E_{(i-1)j_n+k,n}] \\ &\quad \left. + 4 \lfloor j_n \rfloor^3 \mathbf{P} [E_{1,n}]^3 \sum_{j=1}^{\lfloor j_n \rfloor} \mathbf{P} [E_{(i-1)j_n+j,n}] + \lfloor j_n \rfloor^4 \mathbf{P} [E_{1,n}]^4 \right) \\ &\sim \mathcal{O} \left( \frac{1}{n^2 h_n^2} \cdot (\lfloor j_n \rfloor h_n + \lfloor j_n \rfloor^2 h_n^2 + \lfloor j_n \rfloor^3 h_n^3 + \lfloor j_n \rfloor^4 h_n^4) \right) = \mathcal{O} \left( \frac{1}{n^2 h_n^2} \right). \end{aligned}$$

Therefore we can now conclude for each  $\epsilon > 0$ :

$$\frac{1}{\sigma_n^2} \sum_{i=1}^{\lfloor nh_n \rfloor} \mathbf{E} [X_{n,i}^2 \mathbf{1}_{\{|X_{n,i}| > \epsilon \sigma_n\}}] \leq \frac{1}{\epsilon^2 \sigma_n^4} \sum_{i=1}^{\lfloor nh_n \rfloor} \mathbf{E} [X_{n,i}^4] \sim \mathcal{O} \left( \frac{1}{nh_n} \right) \xrightarrow{n \rightarrow \infty} 0,$$

which establishes condition ii) of Theorem 4.5.4 and thus Theorem 4.5.4 can be applied to prove the assertion.  $\square$

## 7.2 Proofs of Chapter 2

### 7.2.1 Proof of Theorem 2.1.10

**Lemma 7.2.1.** If  $\phi : \mathcal{K} \rightarrow X$  is Hausdorff-continuous and additive,  $X$  a topological space, then the following implication is true:

$$\sum_{m=1}^n c_m \mathbf{1}_{K_m} \equiv 0 \Rightarrow \sum_{m=1}^n c_m \phi(K_m) = 0.$$

**Proof. (Lemma 7.2.1)** Proof by contradiction. If the implication is false, there is a smallest number  $n \geq 2$  such that there exist numbers  $c_1, \dots, c_n$ ,  $K_1, \dots, K_n \in \mathcal{K}$  such that  $\sum_{m=1}^n c_m \mathbf{1}_{K_m} \equiv 0$  but  $\sum_{m=1}^n c_m \phi(K_m) =: k \neq 0$ . Let  $H \subset \mathbb{R}^n$  a hyperplane with  $K_1 \subset \text{int}H^+$ . We have

$$\sum_{m=1}^n c_m \mathbf{1}_{K_m \cap H^-} \equiv \sum_{m=1}^n c_m \mathbf{1}_{K_m \cap H} \equiv 0.$$

and each of these two sums has at most  $n-1$  non-zero summands. Therefore by minimality of  $n$ :

$$\sum_{m=1}^n c_m \phi(K_m \cap H^-) = \sum_{m=1}^n c_m \phi(K_m \cap H) = 0.$$

But then we must have:

$$\sum_{m=1}^n c_m \mathbf{1}_{K_m \cap H^+} = 0 \text{ and } \sum_{m=1}^n c_m \phi(K_m \cap H^+) = k.$$

This is the situation from which we started with  $K_m$  replaced by  $K_m \cap H^+$ . Convexity theory assures that a sequence of hyperplanes  $(H_i)_{i \in \mathbb{N}}$  can be found such that  $K_1 = \bigcap_{i=1}^{\infty} H_i^+$ . Repeated use of the above argument reveals:

$$\sum_{m=1}^n c_m \phi(K_m \cap \bigcap_{i=1}^s H_i) = k.$$

Because  $(K_m \cap \bigcap_{i=1}^s H_i^+)$  converges to  $K_m \cap K_1$  in the Hausdorff-sense, we can write:

$$\sum_{m=1}^n c_m \phi(K_m \cap K_1) = k \text{ and } \sum_{m=1}^n c_m \mathbf{1}_{K_m \cap K_1} \equiv 0.$$

Again we reached a similar statement as the one from which we started. Repeated use of the whole argument (for  $K_2, \dots, K_n$ ) leads to:

$$\sum_{m=1}^n c_m \phi(K_1 \cap \dots \cap K_n) = k \text{ and } \sum_{m=1}^n c_m \mathbf{1}_{K_1 \cap \dots \cap K_n} \equiv 0.$$

But because  $k \neq 0$  also  $\sum_{m=1}^n c_m \neq 0$  and therefore  $\mathbf{1}_{K_1 \cap \dots \cap K_n} \equiv 0$  which finally yields  $\phi(K_1 \cap \dots \cap K_n) = 0$  in contradiction to the first equation in the last pair above.  $\square$

**Proof. (Theorem 2.1.10)** The finite linear combinations of indicator functions of sets in  $\mathcal{K}$  form a linear vector space  $V$  with  $\mathbf{1}_R \in V$  according to Theorem 2.1.7. For  $f = \sum_{m=1}^n c_m \mathbf{1}_{K_m} \in V$  define:

$$\tilde{\phi}(f) := \sum_{m=1}^n c_m \phi(K_m).$$

Lemma 7.2.1 guarantees that this definition does not depend on the representation of  $f$ .  $\tilde{\phi}(f)$  is linear with  $\tilde{\phi}(\mathbf{1}_K) = \phi(K)$  for  $K \in \mathcal{K}$ . It is therefore natural to define:

$$\phi(R) := \tilde{\phi}(\mathbf{1}_R).$$

Additivity of  $\phi$  on  $\mathcal{R}$  can trivially be verified using linearity of  $\tilde{\phi}$  and additivity of  $K \rightarrow \mathbf{1}_K$  which completes the proof.  $\square$

### 7.2.2 Proof of Theorem 2.1.20

**Lemma 7.2.2.** Let  $R = \cup_{i=1}^m K_i \in \mathcal{R}$  and  $B := -(\Delta, \omega)$  a segment such that  $R = R \circ B = R \bullet B$  and  $E_{s,\omega}$  the hyperplane with normal vector  $\omega$  at (oriented) distance  $s$  from the origin. Then  $\forall s \in \mathbb{R}$ :

$$\chi((R \oplus B) \cap E_{s,\omega}) = \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \chi \left[ \left( \left( \bigcap_{i \in I} K_i \right) \oplus B \right) \cap E_{s,\omega} \right].$$

**Proof. (Lemma 7.2.2)** Let  $s \in \mathbb{R}$  and  $x \in E_{s,\omega}$  be fixed. Obviously  $\forall I \subset \{1, \dots, m\}$  we have that  $x \notin (R \oplus B) \Rightarrow x \notin \left( \bigcap_{i \in I} K_i \right) \oplus B$ . And hence:

$$\mathbf{1}_{(R \oplus B)}(x) = 0 \Rightarrow \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \mathbf{1}_{\left( \bigcap_{i \in I} K_i \right) \oplus B}(x) = 0.$$

Furthermore note that because of  $R = R \bullet B$  for  $\tilde{B}_x := -B + x$  we have  $\chi(R \cap \tilde{B}_x) \leq 1$ . Whereas additivity yields:

$$\chi(R \cap \tilde{B}_x) = \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \chi \left( \bigcap_{i \in I} K_i \cap \tilde{B}_x \right) = \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \mathbf{1}_{\left( \bigcap_{i \in I} K_i \right) \oplus B}(x).$$

If we note that  $\chi(R \cap \tilde{B}_x) = 1 \Rightarrow x \in (R \oplus B)$  we can conclude so far  $\forall x \in E_{s,\omega}$ :

$$\mathbf{1}_{R \oplus B}(x) = \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \mathbf{1}_{\left( \bigcap_{i \in I} K_i \right) \oplus B}(x).$$

And thus trivially also:

$$\mathbf{1}_{(R \oplus B) \cap E_{s,\omega}}(x) = \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \mathbf{1}_{\left( \left( \bigcap_{i \in I} K_i \right) \oplus B \right) \cap E_{s,\omega}}(x).$$

Again we use additivity to expand the left hand side of the last equation:

$$\sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \mathbf{1}_{\bigcap_{i \in I} (K_i \oplus B) \cap E_{s,\omega}}(x) = \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \mathbf{1}_{\left( \bigcap_{i \in I} K_i \right) \oplus B \cap E_{s,\omega}}(x).$$

Now we can apply Lemma 7.2.1 further above to replace the indicators by the Euler-characteristic  $\chi$  in the last equation and use additivity to transform the left hand side back to  $\chi((R \oplus B) \cap E_{s,\omega})$ .  $\square$

We still need another small lemma:

**Lemma 7.2.3.** Let  $R \in \mathcal{R}$  such that  $R = R \circ B = R \bullet B$  where  $B = -(\Delta, \omega)$ . For  $s \in \mathbb{R}$  let further  $E_{s,\omega}$  be the hyperplane with normal vector  $\omega$  at (oriented) distance  $s$  from the origin and  $Y_{s,\omega} := (E_{s,\omega} \cap R)_{(-s,\omega)}$  the intersection of  $R$  with  $E_{s,\omega}$  shifted back to the origin along  $-\omega$ . Then

$$\chi((R \oplus B) \cap E_{s,\omega}) = \chi(Y_{s,\omega} \cup Y_{s+\Delta,\omega}).$$

**Proof. (Lemma 7.2.3)** Because of  $R = R \circ B = R \bullet B$  we simply have  $R \oplus B = R \cup R_{-\Delta}$ . And thus the lemma follows directly from additivity.  $\square$

We are now ready to prove the theorem:

**Proof. (Theorem 2.1.20)** We apply in this order first Lemma 7.2.3, Lemma 7.2.2 and additivity to arrive at:

$$\begin{aligned} & \int [\chi(Y_{s,\omega} \cup Y_{s+\Delta,\omega}) - \chi(Y_{s,\omega})] ds = \int [\chi((R \oplus B) \cap E_{s,\omega}) - \chi(Y_{s,\omega})] ds \\ &= \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \int \left[ \chi \left( \left( \left( \bigcap_{i \in I} K_i \right) \oplus B \right) \cap E_{s,\omega} \right) - \chi \left( \bigcap_{i \in I} K_i \cap E_{s,\omega} \right) \right] ds \\ &= \sum_{I \subset \{1, \dots, m\}} (-1)^{|I|-1} \cdot \Delta \cdot \chi \left( \bigcap_{i \in I} K_i \right) = \Delta \cdot \chi(R), \end{aligned}$$

where the penultimate equation holds because for each  $I \in \{1, \dots, m\}$  we have either  $\bigcap_{i \in I} K_i = \emptyset$  or  $\bigcap_{i \in I} K_i$  convex.  $\square$

### 7.2.3 Proof of Theorem 2.2.4

**Proof. (Theorem 2.2.4)**

ad i): Define

$$\phi(K_2) := \int_{G_d} V_j(K_1 \cap g_2 K_2) \mu(dg) \quad \forall K_2 \in \mathcal{K}.$$

The function  $\phi(\cdot)$  is trivially integrable, additive and invariant under rigid motions. It can be shown that  $\phi(\cdot)$  is also continuous (for the technical argument see Schneider and Weil 1992). But then according to Theorem 2.1.4 there exist functions  $c_{d-i}(\cdot) : \mathcal{K} \rightarrow \mathbb{R}$  such that:

$$\phi(K_2) = \sum_{i=0}^d c_{d-i}(K_1) V_i(K_2).$$

For  $m$  in  $\{0, \dots, d\}$  choose  $L_m \in \mathcal{L}_m^d$  and set  $B_m^d := B_d(0) \cap L_m$  and  $K_r := rB_m^d$ . Then:

$$\phi(K_r) = \sum_{i=0}^m c_{d-i}(K_1) r^i V_i(B_m^d).$$

But also:

$$\begin{aligned} \phi(K_r) &= \int_{G_d} V_j(K_1 \cap g_2 r B_m^d) \mu(dg) = \int_{SO_d} \int_{\mathbb{R}^d} V_j(K_1 \cap (R \cdot r B_m^d + x)) \lambda_d(dx) \nu(dR) \\ &= \int_{SO_d} \int_{L_m^\perp} \int_{L_m} V_j(K_1 \cap (Rr(B_m^d + x_1) + Rx_2)) r^m \lambda_m(dx_1) \lambda_{d-m}(dx_2) \nu(dR) \end{aligned}$$

Comparison of coefficients gives:

$$\kappa_m c_{d-i}(K_1) = \lim_{r \rightarrow \infty} \int_{SO_d} \int_{L_m^\perp} \int_{L_m} V_j(K_1 \cap (Rr(B_m^d + x_1) + Rx_2)) \lambda_m(dx_1) \lambda_{d-m}(dx_2) \nu(dR).$$

Since  $V_j(K_1 \cap (Rr(B_m^d + x_1) + Rx_2)) \xrightarrow{r \rightarrow \infty} \mathbf{1}_{\{0 \in \text{relint} R(B_m^d + x_1)\}} V_j(K_1 \cap R(L_m + x_2))$ , applying the bounded convergence theorem leads to:

$$\begin{aligned} \kappa_m c_{d-m}(K_1) &= \int_{SO_d} \int_{L_m^\perp} V_j(K_1 \cap R(L_m + x_2)) \lambda_m(B_m^d) \lambda_{d-m}(dx_2) \nu(dR) \\ &= \kappa_m \int_{\mathcal{E}_m^d} V_j(K_1 \cap E) \mu_m(dE) = \kappa_m \frac{m! \kappa_m (d-m+j)! \kappa_{d-m+j}}{j! \kappa_j d! \kappa_d} V_{d-m+j}(K_1), \end{aligned}$$

where we have used the Crofton-formula (Theorem 2.2.1) and a well-know identity for Gamma functions at the end.

ad ii): This follows directly from an induction in  $n$ . The case  $n = 2$  is provided by *i*). The induction step follows directly from first applying *i*) and then the induction hypothesis.  $\square$

#### 7.2.4 Proof of the Existence of a Poisson Process in $\mathbb{R}^d$

**Theorem 7.2.4.** *Let  $\nu$  be an arbitrary locally-finite and atom-free measure on  $\mathcal{B}(\mathbb{R}^d)$ . Then a Poisson-Process  $X$  with intensity  $\nu$  exists.*

**Proof. (Theorem 7.2.4)** The proof is constructive. Let a countable decomposition of  $\mathbb{R}^d$  into disjoint bounded sets  $A_i$  be given. We construct the process  $X \cap A_i \forall i$  separately according to the following recipe:

i) Draw the number of points  $X(A_i)$  that  $X$  has within  $A_i$  according to:

$$X(A_i) \sim \text{Pois}(\nu(A_i)).$$

ii) Given  $X(A_i) = n$ , choose  $n$  i.i.d. points from  $A_i$ , i.e:

$$P[x_i \in B] = \frac{\nu(A_i \cap B)}{\nu(A_i)}, \quad B \subset \mathbb{R}^d.$$

We now show that the so-constructed process  $X$  is indeed a Poisson process. We must show that the number of points  $X(B)$  of  $X$  falling into a set  $B$  has a  $\text{Pois}(\nu(B))$ -distribution. We first assume  $B$  to be a subset of an  $A_i$ , ie.  $B \subset A_i$ . Then:

$$\begin{aligned}
\mathbb{P}[X(B) = k] &= \sum_{n=k}^{\infty} \mathbb{P}[X(B) = k | X(A_i) = n] \cdot \mathbb{P}[X(A_i) = n] \\
&= \sum_{n=k}^{\infty} \binom{n}{k} \left( \frac{\nu(B)}{\nu(A_i)} \right)^k \left( 1 - \frac{\nu(B)}{\nu(A_i)} \right)^{n-k} e^{-\nu(A_i)} \cdot \frac{(\nu(A_i))^n}{n!} \\
&= e^{-\nu(A_i)} \frac{(\nu(B))^k}{k!} \sum_{n=k}^{\infty} \frac{1}{(n-k)!} (\nu(A_i) - \nu(B))^{n-k} \\
&= e^{-\nu(A_i)} \frac{(\nu(B))^k}{k!} e^{\nu(A_i) - \nu(B)} = e^{-\nu(B)} \frac{(\nu(B))^k}{k!}.
\end{aligned}$$

For a general  $B \subset \mathbb{R}^d$  we can write:

$$X(B) = \sum_i X(B \cap A_i),$$

where for a specific  $B$  this is in fact a finite sum of Poisson variables which is thus Poisson itself with intensity  $\nu(B)$ .

For the independence of point-numbers  $X(B_i)$  falling into the sets  $B_i$  of a finite collection  $B_i, i \in 1, \dots, n$ , we consider for simplicity only the case where  $n = 2$  and  $B_1 \subset A_i, B_2 \subset A_i, B_1 \cap B_2 = \emptyset$ . The general case is completely analogous. We further abbreviate  $B_3 := A_i \setminus (B_1 \cup B_2)$ . Now we can calculate as follows:

$$\begin{aligned}
\mathbb{P}[X(B_1) = k_1, X(B_2) = k_2] &= \sum_{n=k_1+k_2}^{\infty} \mathbb{P}[X(B_1) = k_1, X(B_2) = k_2 | X(A_i) = n] \mathbb{P}[X(A_i) = n] \\
&= \sum_{n=k_1+k_2}^{\infty} \frac{n!}{k_1! k_2! (n - k_1 - k_2)!} \left( \frac{\nu(B_1)}{\nu(A_i)} \right)^{k_1} \left( \frac{\nu(B_2)}{\nu(A_i)} \right)^{k_2} \left( \frac{\nu(B_3)}{\nu(A_i)} \right)^{n-k_1-k_2} \cdot e^{-\nu(A_i)} \frac{(\nu(A_i))^n}{n!} \\
&= e^{-\nu(A_i)} \frac{\nu(B_1)^{k_1}}{k_1!} \frac{\nu(B_2)^{k_2}}{k_2!} \sum_{n=k_1+k_2}^{\infty} \frac{1}{(n - k_1 - k_2)!} \nu(B_3)^{n-k_1-k_2} \\
&= e^{-\nu(B_1) - \nu(B_2)} \frac{\nu(B_1)^{k_1}}{k_1!} \frac{\nu(B_2)^{k_2}}{k_2!} = \mathbb{P}[X(B_1) = k_1] \mathbb{P}[X(B_2) = k_2].
\end{aligned}$$

□

### 7.2.5 Alternative Proof of Theorem 2.5.2

**Proof. (Theorem 2.5.2)** Let us denote by  $\partial\Theta$  the boundary of  $\Theta$  and let  $\partial\Theta(\epsilon)$  be the set  $\partial\Theta(\epsilon) := \{x \in \mathbb{R}^d | d(x, \partial\Theta) < \epsilon\}$ . Now consider an arbitrary point  $x_0 \in \partial\Theta(\epsilon)$ . Then a certain region  $A(\epsilon)$  of the ball  $B_d(x_0, \epsilon)$  consists of points  $x \in \mathbb{R}^d$  with  $\mathbf{1}_{x \in \Theta} \neq \mathbf{1}_{x_0 \in \Theta}$ . If  $\epsilon$  is chosen small enough,  $A(\epsilon)$  is approximately a spherical segment. Because of stationarity and isotropy the distance  $d$  from  $x_0$  to the spherical segment  $A(\epsilon)$  has a  $U(0, \epsilon)$ -distribution. The expected volume of  $A(\epsilon)$  can easily be calculated:



We start out with the deterministic volume of a spherical segment  $A(\epsilon, t)$  of  $B_d(x_0, \epsilon)$  with distance  $t$  from the origin:

$$v(A(\epsilon, t)) = \kappa_{d-1} \cdot \int_t^\epsilon (\epsilon^2 - r^2)^{\frac{d-1}{2}} dr = \kappa_{d-1} \cdot \epsilon^d \int_{\frac{t}{\epsilon}}^1 (1 - r^2)^{\frac{d-1}{2}} dr,$$

and hence:

$$\begin{aligned} \mathbf{E}[v(A(\epsilon)) | x_0 \in \partial\Theta(\epsilon)] &= \frac{1}{\epsilon} \int_0^\epsilon v(A(\epsilon, t)) dt = \kappa_{d-1} \cdot \epsilon^{d-1} \int_0^\epsilon \int_{\frac{t}{\epsilon}}^1 (1 - r^2)^{\frac{d-1}{2}} dr dt \\ &= \kappa_{d-1} \cdot \epsilon^{d-1} \int_0^1 \int_0^{\epsilon r} (1 - r^2)^{\frac{d-1}{2}} dt dr = \frac{\kappa_{d-1} \cdot \epsilon^d}{d+1}. \end{aligned} \quad (7.16)$$

Noting further that for two arbitrary points  $x, x_0 \in \mathbb{R}^d$  we have:

$$\begin{aligned} \mathbf{P}[\mathbf{1}_{\{x \in \Theta\}} \neq \mathbf{1}_{\{x_0 \in \Theta\}}] &= 2 \cdot (\mathbf{P}[x \in \Theta] - \mathbf{P}[x \in \Theta \cap x_0 \in \Theta]) = 2(C_\Theta(0) - C_\Theta(\|x - x_0\|)) \\ &\approx -2\|x - x_0\| \cdot \frac{\partial C_\Theta}{\partial r}(0). \end{aligned}$$

For an arbitrary point  $x_0 \in \mathbb{R}^d$  we thus get:

$$\begin{aligned} \mathbf{P}[x_0 \in \partial\Theta(\epsilon)] \cdot \mathbf{E}[v(A(\epsilon)) | x_0 \in \partial\Theta(\epsilon)] &= \mathbf{E}[\mathbf{E}[v(A(\epsilon)) | x_0 \in \partial\Theta(\epsilon)]] = \mathbf{E}[v(A(\epsilon))] \\ &= \int_\Omega \int_{B_d(x_0, \epsilon)} \mathbf{1}_{\{\Theta(x) \neq \Theta(x_0)\}} dx dP = \int_{B_d(x_0, \epsilon)} \mathbf{P}[\mathbf{1}_{\{x \in \Theta\}} \neq \mathbf{1}_{\{x_0 \in \Theta\}}] dx \\ &= -2 \frac{\partial C_\Theta}{\partial r}(0) \int_{B_d(x_0, \epsilon)} \|x - x_0\| dx = -2 \frac{\partial C_\Theta}{\partial r}(0) \cdot d \cdot \kappa_d \int_0^\epsilon r \cdot r^{d-1} dr \\ &= -2 \frac{\partial C_\Theta}{\partial r}(0) \cdot d \cdot \kappa_d \frac{\epsilon^{d+1}}{d+1}. \end{aligned} \quad (7.17)$$

Now we note that for sufficiently small  $\epsilon$  and observation window  $W$ :

$$\begin{aligned} \mathbf{P}[x_0 \in \partial\Theta(\epsilon)] &= \frac{1}{v(W)} \int_W \mathbf{P}[x \in \partial\Theta(\epsilon)] dx = \frac{1}{v(W)} \int_\Omega \int_W \mathbf{1}_{\{x \in \partial\Theta(\epsilon)\}} dx dP \\ &= 2\epsilon \frac{\mathbf{E}[s(\Theta \cap W)]}{v(W)} \approx 2\epsilon \bar{s}(\Theta). \end{aligned} \quad (7.18)$$

It now remains to combine equations 7.16, 7.17 and 7.18 to achieve:

$$2\epsilon \bar{s}(\Theta) \cdot \frac{\kappa_{d-1} \cdot \epsilon^d}{d+1} \approx -2 \frac{\partial C_\Theta}{\partial r}(0) \cdot d \cdot \kappa_d \frac{\epsilon^{d+1}}{d+1},$$

which is independent on the  $\epsilon$  we chose. Solving for  $\bar{s}(\Theta)$  finally gives the desired result:

$$\bar{s}(\Theta) = -\frac{d\kappa_d}{\kappa_{d-1}} \frac{\partial C_\Theta}{\partial r}(0).$$

□

### 7.3 Proofs of Chapter 4

#### 7.3.1 Alternative Proof of Lemma 4.3.2

**Proof. (Lemma 4.3.2)**

$$\begin{aligned}
& \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b] = \int_b^\infty \mathbb{P}[Z(h_n) \leq b | Z(0) = z] \mathbb{P}[Z(0) \in dz] \\
&= \int_b^\infty \Phi\left(\frac{b - \rho(h_n)z}{\sqrt{1 - \rho^2(h_n)}}\right) \cdot \phi(z) dz \\
&= h \cdot \int_0^\infty \Phi\left(-\sqrt{\frac{u^2}{-\rho''(0)}} - b \cdot u \cdot h + \frac{1}{12} \left(-3\rho''(0)b^2 - 9u^2 + \frac{\rho^{IV}(0)}{\rho''(0)^2}u^2\right) \cdot h^2 + \mathcal{O}(h^3)\right) \\
&\quad \times \phi(b + uh) du,
\end{aligned}$$

where we have used the the substitution  $z = b + u \cdot h$  and the following expansions of the  $\Phi(\cdot)$  argument into a Taylor series around  $h = 0$ :

$$\begin{aligned}
b - \rho(h)z &= b - \left(1 + \frac{\rho''(0)}{2} \cdot h^2 + \mathcal{O}(h^4)\right) \cdot (b + u \cdot h) \\
&= -u \cdot h - \frac{b \cdot \rho''(0)}{2} \cdot h^2 + \mathcal{O}(h^3). \tag{7.19}
\end{aligned}$$

$$\begin{aligned}
1 - \rho^2(h) &= 1 - (1 + \rho''(0)h^2 + \mathcal{O}(h^4)) \\
&= -\rho''(0) \cdot h^2 + \mathcal{O}(h^4). \tag{7.20}
\end{aligned}$$

This can be expanded further using the following expansions for small  $h$ :

$$\begin{aligned}
\Phi(x + h) &= \Phi(x) + \phi(x) \cdot h - x \cdot \phi(x) \cdot \frac{h^2}{2} + \mathcal{O}(h^3), \\
\phi(b + u \cdot h) &= \phi(b) - b \cdot \phi(b) \cdot u \cdot h + \phi(b) \cdot \frac{b^2 - 1}{2} \cdot u^2 \cdot h^2 + \mathcal{O}(h^3), \\
\sqrt{x + h} &= \sqrt{x} + \frac{1}{2\sqrt{x}} \cdot h - \frac{1}{8x^{\frac{3}{2}}} \cdot h^2 + \mathcal{O}(h^3).
\end{aligned}$$

Thus we get:

$$\begin{aligned}
& \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b] \\
&= h_n \cdot \int_0^\infty \Phi\left(-\frac{u}{\sqrt{-\rho''(0)}} + \frac{\sqrt{-\rho''(0)}}{2} \cdot bh_n - \frac{(\rho^{IV}(0) - 9\rho''(0)^2)u}{24(-\rho''(0))^{\frac{3}{2}}} \cdot h_n^2 + \mathcal{O}(h_n^3)\right) \\
&\quad \times \phi(b + u \cdot h_n) du \\
&= h_n \cdot \int_0^\infty \left[ \Phi\left(-\frac{u}{\sqrt{-\rho''(0)}}\right) + \phi\left(\frac{u}{\sqrt{-\rho''(0)}}\right) \cdot \frac{\sqrt{-\rho''(0)}}{2} \cdot b \cdot h_n \right. \\
&\quad \left. - \phi\left(\frac{u}{\sqrt{-\rho''(0)}}\right) \cdot \frac{(\rho^{IV}(0) - 3\rho''(0)^2(3 + b^2))u}{24(-\rho''(0))^{\frac{3}{2}}} \cdot h_n^2 + \mathcal{O}(h_n^3) \right] \\
&\quad \cdot \left( \phi(b) - b\phi(b)u \cdot h_n + \frac{b^2 - 1}{2} \phi(b)u^2 h_n^2 + \mathcal{O}(h_n^3) \right) du
\end{aligned}$$

$$\begin{aligned}
&= h_n \cdot \phi(b) \cdot \int_0^\infty \Phi\left(-\frac{u}{\sqrt{-\rho''(0)}}\right) du \\
&+ h_n^2 \cdot \left[ \frac{1}{2} b \phi(b) \sqrt{-\rho''(0)} \int_0^\infty \phi\left(\frac{u}{\sqrt{-\rho''(0)}}\right) du - b \phi(b) \cdot \int_0^\infty u \cdot \Phi\left(-\frac{u}{\sqrt{-\rho''(0)}}\right) du \right] \\
&+ h_n^3 \left[ \frac{b^2-1}{2} \phi(b) \int_0^\infty u^2 \Phi\left(-\frac{u}{\sqrt{-\rho''(0)}}\right) du - \phi(b) \frac{(\rho^{IV}(0) - 3\rho''(0)^2(3+b^2))}{24(-\rho''(0))^{\frac{3}{2}}} \right. \\
&\times \int_0^\infty u \phi\left(\frac{u}{\sqrt{-\rho''(0)}}\right) du - \frac{1}{2} b^2 \phi(b) \sqrt{-\rho''(0)} \int_0^\infty u \phi\left(\frac{u}{\sqrt{-\rho''(0)}}\right) du \left. \right] + \mathcal{O}(h_n^4) \\
&\sim \frac{1}{\sqrt{2\pi}} \phi(b) \sqrt{-\rho''(0)} \cdot h_n + h_n^2 \cdot \left( \frac{1}{4} b \phi(b) (-\rho''(0)) - \frac{1}{4} b \phi(b) (-\rho''(0)) \right) \\
&+ h_n^3 \left[ \frac{b^2-1}{2} \phi(b) \frac{2}{3} \sqrt{\frac{1}{2\pi}} (-\rho''(0))^{\frac{3}{2}} + \frac{3}{8} \phi(b) \left( \sqrt{-\rho''(0)} \cdot (1-b^2) - \frac{\rho^{IV}(0)}{9(-\rho''(0))^{\frac{3}{2}}} \right) \right. \\
&\times \left. \frac{1}{\sqrt{2\pi}} \cdot (-\rho''(0)) \right] \\
&= \frac{1}{\sqrt{2\pi}} \phi(b) \sqrt{-\rho''(0)} \cdot h_n - \frac{\sqrt{-\rho''(0)}}{24} \left[ (b^2-1)\rho''(0)^2 + \frac{\rho^IV(0)}{\rho''(0)} \right] \cdot h_n^3 + \mathcal{O}(h_n^4), \quad (7.21)
\end{aligned}$$

where we have used equation lemma 4.3.1, parts ii), iii) and ix).  $\square$

### 7.3.2 Alternative Proof of Lemma 4.3.6

**Proof. (Lemma 4.3.6)** Following the strategy approved above, we start out expressing  $P[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b]$  by conditioning on  $Z_0$  and using the variables  $u$  (positive first-order deviation from  $b$ ), i.e.  $Z_0 = b + u \cdot h$ , the approximate first-order differences  $Y_1 = \frac{Z_1 - b}{h}$ ,  $Y_2 = \frac{Z_2 - b}{h}$  and the approximate second-order mixed derivative  $Y_3 = \frac{Z_3 - Z_1 - Z_2 + b}{h^2}$ . This leads to:

$$\begin{aligned}
&P[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] = \int_b^\infty P[Z_0 = z_0, Z_1 < b, Z_2 < b, Z_3 \geq b] dz_0 \\
&= P\left[Z_1 < b, Z_2 < b, Z_3 \geq b \mid Z_0 = z_0\right] \cdot P[Z_0 \in dz_0] \\
&= \int_b^\infty P\left[Z_1 < b, Z_2 < b, Z_3 \geq b \mid Z_0 = z_0\right] \cdot \phi(z_0) dz_0 \\
&= h \int_0^\infty \phi(b + h \cdot u) \cdot P\left[Z_1 < b, Z_2 < b, Z_3 \geq b \mid Z_0 = b + h \cdot u\right] du \\
&= h \cdot \int_0^\infty \phi(b + h \cdot u) \cdot P\left[Y_1 < 0, Y_2 < 0, Y_3 \geq -\frac{Y_1 + Y_2}{h} \mid Z_0 = b + h \cdot u\right] du,
\end{aligned}$$

where the last equality holds because of:

$$Z_3 \geq b \iff Y_3 = \frac{(Z_3 - (Z_1 - b) - (Z_2 - b) + b)}{h^2} \geq -\frac{Y_1 + Y_2}{h}.$$

The random vector  $Z = (Z_0, Z_1, Z_2, Z_3)$  is jointly Gaussian distributed with  $\mathbf{E}[Z] = 0$  and  $\text{Cov}(Z) = \Sigma(h)$ , where

$$\Sigma(h) = \begin{pmatrix} 1 & \rho(h) & \rho(h) & \rho(\sqrt{2} \cdot h) \\ \rho(h) & 1 & \rho(\sqrt{2} \cdot h) & \rho(h) \\ \rho(h) & \rho(\sqrt{2} \cdot h) & 1 & \rho(h) \\ \rho(\sqrt{2} \cdot h) & \rho(h) & \rho(h) & 1 \end{pmatrix}.$$

The conditional distribution  $(Z_1, Z_2, Z_3)^T | Z_0$  is still multivariate Gaussian with parameters:

$$\begin{aligned} \mathbf{E}[(Z_1, Z_2, Z_3)^T | Z_0] &= \begin{pmatrix} \rho(h) \\ \rho(h) \\ \rho(\sqrt{2} \cdot h) \end{pmatrix} \cdot (b + h \cdot u) \\ \text{Cov}((Z_1, Z_2, Z_3)^T | Z_0) &= \\ &\begin{pmatrix} 1 - \rho^2(h) & \rho(\sqrt{2} \cdot h) - \rho^2(h) & \rho(h) - \rho(h) \cdot \rho(\sqrt{2} \cdot h) \\ \rho(\sqrt{2} \cdot h) - \rho^2(h) & 1 - \rho^2(h) & \rho(h) - \rho(h) \cdot \rho(\sqrt{2} \cdot h) \\ \rho(h) - \rho(h) \cdot \rho(\sqrt{2} \cdot h) & \rho(h) - \rho(h) \cdot \rho(\sqrt{2} \cdot h) & 1 - \rho^2(\sqrt{2} \cdot h) \end{pmatrix} \end{aligned}$$

From this we can determine the moments  $\tilde{\mu}(u, h) = \mathbf{E}[(Y_1, Y_2, Y_3 | Z_0]$  and  $\tilde{\Sigma}(h) = \text{Cov}((Y_1, Y_2, Y_3 | Z_0)$  of the conditional distribution  $(Y_1, Y_2, Y_3 | Z_0$ :

$$\begin{aligned} \tilde{\mu}_1(u, h) &:= \mathbf{E}[Y_1 | Z_0] = \mathbf{E}\left[\frac{Z_1 - b}{h} \middle| Z_0\right] = \frac{1}{h} \cdot \left(\rho(h) \cdot (b + u \cdot h) - b\right) \\ &= u + \frac{\rho''(0)}{2} \cdot b \cdot h + \frac{\rho''(0)}{2} \cdot u \cdot h^2 + \mathcal{O}(h^3). \\ \tilde{\mu}_2(u, h) &:= \mathbf{E}[Y_2 | Z_0] = \tilde{\mu}_1(u, h). \\ \tilde{\mu}_3(u, h) &:= \mathbf{E}[Y_3 | Z_0] = \mathbf{E}\left[\frac{Z_3 - Z_1 - Z_2 + b}{h^2} \middle| Z_0\right] = \\ &\frac{1}{h^2} \left(\rho(\sqrt{2} \cdot h) \cdot (b + u \cdot h) - 2 \cdot \rho(h) \cdot (b + u \cdot h) + b\right) = -\frac{u}{h} + \frac{1}{12} \cdot b \cdot \rho^{(IV)}(0) \cdot h^2 + \mathcal{O}(h^3). \\ \tilde{\Sigma}_{11}(h) &:= \text{Var}\left(\frac{Z_1 - b}{h} \middle| Z_0\right) = \frac{1 - \rho^2(h)}{h^2} = -\rho''(0) + \left(-\frac{1}{4}\rho''(0) - \frac{1}{12}\rho^{(IV)}(0)\right) \cdot h^2 + \mathcal{O}(h^4). \\ \tilde{\Sigma}_{22}(h) &:= \text{Var}(Y_2 | Z_0) = \tilde{\Sigma}_{11}(h). \\ \tilde{\Sigma}_{33}(h) &:= \text{Var}(Y_3 | Z_0) = \text{Var}\left(\frac{Z_3 - Z_1 - Z_2 + b}{h^2} \middle| Z_0\right) = \frac{1}{3} \cdot \rho^{(IV)}(0) + \mathcal{O}(h^2). \\ \tilde{\Sigma}_{12}(h) &:= \text{Cov}(Y_1, Y_2 | Z_0) = \frac{1}{h^2} \cdot \text{Cov}(Z_1, Z_2 | Z_0) = \left(\frac{1}{12} \cdot \rho^{(IV)}(0) - \frac{1}{4} \cdot \rho''(0)\right) \cdot h^2 + \mathcal{O}(h^4). \\ \tilde{\Sigma}_{13}(h) &:= \text{Cov}(Y_1, Y_3 | Z_0) = \frac{1}{h^3} \cdot (\text{Cov}(X_1, X_3 | Z_0) - \text{Var}(X_1 | Z_0) - \text{Cov}(X_1, X_2 | Z_0)) \\ &= -\frac{1}{6} \cdot \rho^{(IV)}(0) \cdot h + \mathcal{O}(h^3). \\ \tilde{\Sigma}_{23}(h) &:= \tilde{\Sigma}_{13}(h). \end{aligned}$$

Thus we have:

$$\begin{aligned}
& \mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] = \\
& h \int_0^\infty \phi(b + h \cdot u) \int_{-\infty}^0 \int_{-\infty}^0 \int_{-\frac{y_1+y_2}{h}}^\infty \phi\left((y_1, y_2, y_3)^T, \tilde{\mu}(u, h), \tilde{\Sigma}(h)\right) dy_3 dy_2 dy_1 du \\
& \stackrel{z=y_3-\tilde{\mu}_3(u,h)}{=} h \int_0^\infty \phi(b + h \cdot u) \int_{-\infty}^0 \int_{-\infty}^0 \int_{-\frac{y_1+y_2}{h}-\tilde{\mu}_3(u,h)}^\infty \\
& \quad \phi\left((y_1, y_2, z)^T, (u + \mathcal{O}(h), u + \mathcal{O}(h), 0)^T, \tilde{\Sigma}(h)\right) dy_3 dy_2 dy_1 du \\
& = h \int_0^\infty \phi(b + h \cdot u) \int_{-\infty}^0 \int_{-\infty}^0 \left[ \phi(y_1, u, -\rho''(0)) \cdot \phi(y_2, u, -\rho''(0)) \cdot \right. \\
& \quad \left. \left(1 - \Phi\left(\frac{-\frac{y_1+y_2}{h} + \frac{u}{h}}{\sqrt{\tilde{\Sigma}_{33}(0)}}\right)\right) + \mathcal{O}(h) \right] dy_2 dy_1 du
\end{aligned}$$

To estimate the leading term in the above expression we use the following coordinate transform:

$$\begin{aligned}
x_1 &= \frac{u - y_1 - y_2}{h}, & u - y_1 &= h \cdot x_1 + y_2 \\
x_2 &= \frac{y_1}{h}, & &= h \cdot (x_1 + x_3), \\
x_3 &= \frac{y_2}{h}, & u - y_2 &= h \cdot h \cdot x_1 + y_1 \\
& & &= h \cdot (x_1 + x_2),
\end{aligned}$$

which implies:

Applying this transform and using  $\phi(b + h \cdot u) \leq (2 \cdot \pi)^{-1}$  three times allows us to write:

$$\begin{aligned}
\mathbb{P}[Z_0 \geq b, Z_1 < b, Z_2 < b, Z_3 \geq b] &\leq h^4 \cdot \frac{1}{\sqrt{2 \cdot \pi}} \int_0^\infty \left(1 - \Phi\left(\frac{x_1}{\sqrt{\tilde{\Sigma}_{33}(0)}}\right)\right) \cdot \\
& \quad \int_0^{x_1} \phi\left(\frac{h \cdot (x_1 + x_2)}{\sqrt{-\rho''(0)}}\right) dx_2 \int_0^{x_1} \phi\left(\frac{h \cdot (x_1 + x_3)}{\sqrt{-\rho''(0)}}\right) dx_3 dx_1 \\
& = h^4 \cdot \frac{1}{\sqrt{2 \cdot \pi}} \int_0^\infty \left(1 - \Phi\left(\frac{x_1}{\sqrt{\tilde{\Sigma}_{33}(0)}}\right)\right) \cdot \left(\int_0^{x_1} \phi\left(\frac{h \cdot (x_1 + x_2)}{\sqrt{-\rho''(0)}}\right) dx_2\right)^2 dx_1 \\
& \leq h^4 \cdot \frac{1}{(2 \cdot \pi)^{\frac{3}{2}}} \cdot \underbrace{\int_0^\infty x_1^2 \cdot \left(1 - \Phi\left(\frac{x_1}{\sqrt{\tilde{\Sigma}_{33}(0)}}\right)\right) dx_1}_{< \infty},
\end{aligned}$$

where on the last line we have used that because of  $0 \leq -(y_1 + y_2) \leq h \cdot x_1 - u \leq h \cdot x_1$  the integration domain of  $x_2, x_3$  is contained in the square  $[0, x_1]^2$ .  $\square$

### 7.3.3 Alternative Version of Lemma 4.3.7

**Lemma 7.3.1. (Collapsing Gaussian pairs at fixed distance II)** Let a one-dimensional stationary Gaussian Field with standard Gaussian variables  $Z(t), t \geq 0$  and

covariance function  $\rho(\cdot)$  be given. Consider the quadruple  $Z_1 := Z(0), Z_2 := Z(h_n), Z_3 := Z(kh_n), Z_4 := Z((k+1)h_n)$  for the limit process  $h_n \downarrow 0, n \rightarrow \infty$  such that  $kh_n \rightarrow x \in \mathbb{R}^+$ . Then:

$$\begin{aligned} & \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k+1)h_n) \leq b] \\ & \sim h_n^2 \cdot \left[ 2 \cdot \frac{\phi(b)}{\sqrt{-\rho''(0)}} \frac{1}{\sqrt{c_1(x)}} \rho'(x)b \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \right. \\ & \times \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \Phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) dw_2 \\ & + 2 \cdot \frac{\phi(b)}{\sqrt{-\rho''(0)}} \frac{1}{\sqrt{c_1(x)}} \frac{\rho''(x)}{\rho''(0)} \int_{-\infty}^0 w_2^2 \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \\ & \times \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \Phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) dw_2 \\ & - 2 \cdot \frac{\phi(b)}{\sqrt{-\rho''(0)}} \sqrt{\frac{c_2(x)}{c_1(x)}} \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \\ & \left. \times \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) dw_2 \right], \end{aligned}$$

where we have abbreviated:

$$c_1(x) = 4(1-\rho(x))^2 + 4\frac{\rho'(x)^2}{\rho''(0)}, \quad (7.22)$$

$$c_2(x) = -\rho''(0) - \rho'^2(x) + \frac{\rho''(x)^2}{\rho''(0)}. \quad (7.23)$$

**Proof. (Lemma 7.3.1)** We can write:

$$\begin{aligned} & \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k+1)h_n) \leq b] \quad (7.24) \\ & = \int_b^\infty \int_{-\infty}^b \mathbb{P}[Z(kh_n) \geq b, Z((k+1)h_n) \leq b \mid Z(0) = z_1, Z(h_n) = z_2] \phi_2(z_1, z_2) dz_1 dz_2, \end{aligned}$$

where  $\phi_2(\cdot)$  is the density function of the two-dimensional (standard)-normal-distribution. The covariance-matrix of the random vector  $(Z(0), Z(h_n), Z(kh_n), Z((k+1)h_n))$  is:

$$\Sigma = \begin{pmatrix} 1 & \rho(h_n) & \rho(kh_n) & \rho((k+1)h_n) \\ \rho(h_n) & 1 & \rho((k-1)h_n) & \rho(kh_n) \\ \rho(kh_n) & \rho((k-1)h_n) & 1 & \rho(h_n) \\ \rho((k+1)h_n) & \rho(kh_n) & \rho(h_n) & 1 \end{pmatrix}$$

and thus we can calculate the moments of the conditional distribution  $(Z(kh_n), Z((k+1)h_n) \mid z_1, z_2) \sim \mathcal{N}(\tilde{\mu}, \tilde{\Sigma})$  and expand the terms  $\rho((k+1)h_n)$  and  $\rho((k-1)h_n)$  around  $kh_n$ . Finally we introduce the new orthogonal variables  $W_1 := Z(0) + Z(h_n)$  and  $W_2 = (Z(h_n) - Z(0))/h_n$ :

$$\begin{aligned}
\tilde{\mu} &= \begin{pmatrix} \rho(kh_n) & \rho((k-1)h_n) \\ \rho((k+1)h_n) & \rho(kh_n) \end{pmatrix} \frac{1}{1-\rho^2(h_n)} \begin{pmatrix} 1 & -\rho(h_n) \\ -\rho(h_n) & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\
&= \frac{1}{1-\rho^2(h_n)} \begin{pmatrix} [\rho(kh_n) - \rho((k-1)h_n)\rho(h_n)]z_1 + [\rho((k-1)h_n) - \rho(kh_n)\rho(h_n)]z_2 \\ [\rho((k+1)h_n) - \rho(h_n)\rho(kh_n)]z_1 + [\rho(kh_n) - \rho(h_n)\rho((k+1)h_n)]z_2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{\rho'(kh_n)(z_2-z_1)}{\rho''(0)h} + \frac{\rho''(kh_n)(z_1-z_2) + \rho''(0)\rho(kh_n)(z_1+z_2)}{2\rho''(0)} + \mathcal{O}(h_n) \\ \frac{\rho'(kh_n)(z_2-z_1)}{\rho''(0)h} + \frac{\rho''(kh_n)(z_2-z_1) + \rho''(0)\rho(kh_n)(z_1+z_2)}{2\rho''(0)} + \mathcal{O}(h_n) \end{pmatrix} \\
&= \begin{pmatrix} \frac{\rho(kh_n)w_1}{2} + \frac{\rho'(kh_n)w_2}{\rho''(0)} + \mathcal{O}(h_n) \\ \frac{\rho(kh_n)w_1}{2} + \frac{\rho'(kh_n)w_2}{\rho''(0)} + \mathcal{O}(h_n) \end{pmatrix}. \\
\tilde{\Sigma} &= \begin{pmatrix} 1 & \rho(h_n) \\ \rho(h_n) & 1 \end{pmatrix} - \begin{pmatrix} \rho(kh_n) & \rho((k-1)h_n) \\ \rho((k+1)h_n) & \rho(kh_n) \end{pmatrix} \\
&\times \frac{1}{1-\rho^2(h_n)} \begin{pmatrix} 1 & -\rho(h_n) \\ -\rho(h_n) & 1 \end{pmatrix} \begin{pmatrix} \rho(kh_n) & \rho((k+1)h_n) \\ \rho((k-1)h_n) & \rho(kh_n) \end{pmatrix} \\
&= \begin{pmatrix} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21} & \tilde{\Sigma}_{22} \end{pmatrix},
\end{aligned}$$

where:

$$\begin{aligned}
\tilde{\Sigma}_{11} &= 1 - \frac{1}{1-\rho^2(h_n)} [\rho^2(kh_n) + \rho^2((k-1)h_n) - 2\rho(h_n)\rho(kh_n)\rho((k-1)h_n)] \\
&= 1 - \rho^2(kh_n) + \frac{\rho'^2(kh_n)}{\rho''(0)} + \rho'(kh_n) \left( \rho(kh_n) - \frac{\rho''(kh_n)}{\rho''(0)} \right) h_n + \mathcal{O}(h_n^2). \\
\tilde{\Sigma}_{12} &= \rho(h_n) - \frac{1}{1-\rho^2(h_n)} [\rho(kh_n)\rho((k+1)h_n) + \rho(kh_n)\rho((k-1)h_n) \\
&\quad - \rho(h_n)\rho((k-1)h_n)\rho((k+1)h_n) - \rho(h_n)\rho^2(kh_n)] \\
&= 1 - \rho^2(kh_n) + \frac{\rho'^2(kh_n)}{\rho''(0)} + \mathcal{O}(h_n^2). \\
\tilde{\Sigma}_{22} &= 1 - \frac{1}{1-\rho^2(h_n)} [\rho^2((k+1)h_n) + \rho^2(kh_n) - 2\rho(h_n)\rho(kh_n)\rho((k+1)h_n)] \\
&= 1 - \rho^2(kh_n) + \frac{\rho'^2(kh_n)}{\rho''(0)} - \rho'(kh_n) \left( \rho(kh_n) - \frac{\rho''(kh_n)}{\rho''(0)} \right) h_n + \mathcal{O}(h_n^2).
\end{aligned}$$

The change of variables  $(Z_1, Z_2) \mapsto (W_1, W_2)$  changes the integration region in equation 7.24:

from  $\{Z_1 \geq b, Z_2 \leq b\}$  to  $\{W_1 \leq 2b, W_2 \leq (W_1 - 2b)/h\} \cup \{W_1 \geq 2b, W_2 \leq (2b - W_1)/h_n\}$ .

Hence equation 7.24 becomes:

$$\begin{aligned}
&\left( \int_{2b}^{\infty} \int_{-\infty}^{\frac{2b-W_1}{h}} + \int_{-\infty}^{2b} \int_{-\infty}^{\frac{W_1-2b}{h}} \right) \mathbb{P}[Z(kh_n) \geq b, Z((k+1)h_n) \leq b | w_1, w_2] \times \\
&\phi_2 \left[ \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 + \mathcal{O}(h^2) & 0 \\ 0 & -\rho''(0) + \mathcal{O}(h^2) \end{pmatrix} \right] dw_1 dw_2, \quad (7.25)
\end{aligned}$$

where we have used:

$$\begin{aligned}\mathbf{E}[W_1] &= \mathbf{E}[W_2] = 0, \\ \text{Var}(W_1) &= 2 + 2 \cdot \rho(h_n) = 4 + \rho''(0)h_n^2 + \mathcal{O}(h_n^4), \\ \text{Var}(W_2) &= \frac{1}{h_n^2} (2 - 2 \cdot \rho(h_n)) = -\rho''(0)h_n^2 + \mathcal{O}(h_n^4), \\ \text{Cov}(W_1, W_2) &= 0.\end{aligned}$$

To work out the term  $\mathbb{P}[Z(kh_n) \geq b, Z((k+1)h_n) \leq b | w_1, w_2]$  in equation 7.26 we introduce the orthogonal variables  $Y_3 := Z(kh_n) + Z((k+1)h_n)$ ,  $Y_4 := (Z(kh_n) - Z((k+1)h_n))/h$  and thus have:

$$\{Z(kh_n) \geq b, Z((k+1)h_n) \leq b\} = \{Y_3 \leq 2b, Y_4 \geq \frac{2b - Y_3}{h}\} \cup \{Y_3 \geq 2b, Y_4 \geq \frac{Y_3 - 2b}{h}\}.$$

Next we need to calculate the moments of the conditional distribution  $(Y_3, Y_4 | w_1, w_2) \sim \mathcal{N}(\bar{\mu}, \bar{\Sigma})$ . We find:

$$\begin{aligned}\bar{\mu} &= \begin{pmatrix} \rho(kh_n)w_1 + \frac{2\rho'(kh_n)w_2}{\rho''(0)} + \mathcal{O}(h_n) \\ -\frac{\rho'(kh_n)w_1}{2} - \frac{\rho''(kh_n)w_2}{\rho''(0)} + \mathcal{O}(h_n) \end{pmatrix}, \\ \bar{\Sigma} &= \begin{pmatrix} 4(1 - \rho^2(kh_n)) + \frac{4\rho'^2(kh_n)}{\rho''(0)} + \mathcal{O}(h_n^2) & 0 \\ 0 & -\rho''(0) - \rho'^2(kh_n) + \frac{\rho''^2(kh_n)}{\rho''(0)} + \mathcal{O}(h_n^2) \end{pmatrix}.\end{aligned}$$

Performing the substitution  $W_1 = 2b + h_n \cdot u$  the integral in equation 7.26 becomes:

$$\begin{aligned}h_n \cdot \left( \int_0^\infty \int_{-\infty}^{-u} + \int_{-\infty}^0 \int_{-\infty}^u \right) \mathbb{P}[Z(kh_n) \geq b, Z((k+1)h_n) \leq b | 2b + h_n u, w_2] \times \\ \phi_2 \left[ \begin{pmatrix} 2b + h_n u \\ w_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 + \mathcal{O}(h^2) & 0 \\ 0 & -\rho''(0) + \mathcal{O}(h^2) \end{pmatrix} \right] dw_1 dw_2, \quad (7.26)\end{aligned}$$

Finally we substitute  $Y_3 = 2b + h_n \cdot v$  to arrive at:

$$\begin{aligned}\mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k+1)h_n) \leq b] \\ = h_n^2 \left( \int_0^\infty \int_{-\infty}^{-u} + \int_{-\infty}^0 \int_{-\infty}^u \right) \phi_2 \left[ \begin{pmatrix} 2b + h_n u \\ w_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 + \mathcal{O}(h^2) & 0 \\ 0 & -\rho''(0) + \mathcal{O}(h^2) \end{pmatrix} \right] \times \\ \left( \int_0^\infty \int_v^\infty + \int_{-\infty}^0 \int_{-v}^\infty \right) \phi_2 \left[ \begin{pmatrix} 2b + h_n v \\ y_4 \end{pmatrix}, \begin{pmatrix} \rho(kh_n)(2b + h_n u) + \frac{2\rho'(kh_n)w_2}{\rho''(0)} + \mathcal{O}(h_n) \\ -\frac{\rho'(kh_n)(2b + h_n u)}{2} - \frac{\rho''(kh_n)w_2}{\rho''(0)} + \mathcal{O}(h_n) \end{pmatrix}, (7.27) \\ \left( \begin{matrix} 4(1 - \rho^2(kh_n)) + \frac{4\rho'^2(kh_n)}{\rho''(0)} + \mathcal{O}(h_n^2) & 0 \\ 0 & -\rho''(0) - \rho'^2(kh_n) + \frac{\rho''^2(kh_n)}{\rho''(0)} + \mathcal{O}(h_n^2) \end{matrix} \right) \right] dy_4 dv dw_2 du.\end{aligned}$$

Hence in first order it holds:



$$\begin{aligned}
& \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k+1)h_n) \leq b] \\
&= h_n^2 \phi(b) \cdot \frac{1}{2} \frac{1}{\sqrt{-\rho''(0)}} \frac{1}{\sqrt{c_1(x)c_2(x)}} \times \\
&\left( \int_0^\infty \int_{-\infty}^{-u} + \int_{-\infty}^0 \int_{-\infty}^u \right) \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \times \\
&\left( \int_0^\infty \int_v^\infty + \int_{-\infty}^0 \int_{-v}^\infty \right) \phi\left(\frac{1}{\sqrt{c_2(x)}} \left[y_4 + \rho'(x)b + \frac{\rho''(x)}{\rho''(0)}w_2\right]\right) dy_4 dv dw_2 du.
\end{aligned}$$

The inner integrals with respect to  $y_4$  and  $v$  and can be worked out as follows:

$$\begin{aligned}
&\left( \int_0^\infty \int_v^\infty + \int_{-\infty}^0 \int_{-v}^\infty \right) \phi\left(\frac{1}{\sqrt{c_2(x)}} \left[y_4 + \rho'(x)b + \frac{\rho''(x)}{\rho''(0)}w_2\right]\right) dy_4 dv \\
&= 2 \cdot \int_{-\infty}^0 \int_{-v}^\infty \phi\left(\frac{1}{\sqrt{c_2(x)}} \left[y_4 + \rho'(x)b + \frac{\rho''(x)}{\rho''(0)}w_2\right]\right) dy_4 dv \\
&= 2 \cdot \sqrt{c_2(x)} \cdot \int_{-\infty}^0 \Phi\left(\frac{1}{\sqrt{c_2(x)}} \left[v - \rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right]\right) dv \\
&= 2 \cdot c_2(x) \left[ \frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right) \Phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) \right. \\
&\quad \left. + \phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) \right].
\end{aligned}$$

The outer integrals with respect to  $u$  and  $w_2$  are equal by symmetry. Because the integrand does not depend on  $u$  it is favorable to perform the  $u$ -integration first. The integration region transforms as follows:

$$\{u \geq 0, w_2 \leq -u\} \cup \{u \leq 0, w_2 \leq u\} = \{w_2 \leq 0, w_2 \leq u \leq -w_2\}.$$

Hence:

$$\begin{aligned}
& \mathbb{P}[Z(0) \geq b, Z(h_n) \leq b, Z(kh_n) \geq b, Z((k+1)h_n) \leq b] \\
&\sim h_n^2 \cdot \left[ 2 \cdot \frac{\phi(b)}{\sqrt{-\rho''(0)}} \frac{1}{\sqrt{c_1(x)}} \rho'(x)b \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \right. \\
&\times \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \Phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) dw_2 \\
&+ 2 \cdot \frac{\phi(b)}{\sqrt{-\rho''(0)}} \frac{1}{\sqrt{c_1(x)}} \frac{\rho''(x)}{\rho''(0)} \int_{-\infty}^0 w_2^2 \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \\
&\times \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \Phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) dw_2
\end{aligned}$$

$$-2 \cdot \frac{\phi(b)}{\sqrt{-\rho''(0)}} \sqrt{\frac{c_2(x)}{c_1(x)}} \int_{-\infty}^0 w_2 \phi\left(\frac{w_2}{\sqrt{-\rho''(0)}}\right) \\ \times \phi\left(\frac{1}{\sqrt{c_1(x)}} \cdot \left[2b(1-\rho(x)) - 2\frac{\rho'(x)}{\rho''(0)}w_2\right]\right) \phi\left(\frac{1}{\sqrt{c_2(x)}} \left(-\rho'(x)b - \frac{\rho''(x)}{\rho''(0)}w_2\right)\right) dw_2 \Big],$$

which is the desired form.  $\square$

### 7.3.4 Asymptotic Normality of the OM-Surface-Estimator

**Theorem 7.3.2. (surface-estimator properties in on-off-system)** Let  $Z$  be the stationary continuous on-off Markov chain with the generator  $A$  from equation 7.4 and initial distribution  $\eta$  in equation 7.8. Let further the estimator  $\hat{s}_n(Z)$  be defined as in equation 7.9. Then the following statements hold true:

i).

$$\mathbf{E} [\hat{s}_n(Z)] = \frac{1}{h_n} \cdot (P_{10}(h_n) \cdot \eta_1 + P_{01}(h_n) \cdot \eta_0) = \frac{1}{h_n} \cdot \frac{2\lambda_0\lambda_1}{(\lambda_0 + \lambda_1)^2} \left(1 - e^{-(\lambda_0 + \lambda_1) \cdot h_n}\right)$$

and the bias of  $\hat{s}_n(Z)$  is:

$$\mathbf{E} [\hat{s}_n(Z)] - \bar{s} = -\lambda_0 \cdot \lambda_1 \cdot h_n + \mathcal{O}(h_n^2).$$

Hence  $\hat{s}_n(Z)$  is asymptotically unbiased as  $h_n \downarrow 0$ .

ii). If  $n \rightarrow \infty$  and  $h_n \downarrow 0$  such that  $n \cdot h_n \rightarrow \infty$  the order of magnitude of the variance is:

$$\text{Var}(\hat{s}_n(Z)) \sim \frac{4\lambda_0\lambda_1}{nh_n} \frac{\lambda_0^2 + \lambda_1^2}{(\lambda_0 + \lambda_1)^3}.$$

#### Proof. (Theorem 7.3.2)

ad i): Using stationarity and linearity we conclude:

$$\mathbf{E} [\hat{s}_n(Z)] = \frac{1}{n \cdot h_n} \sum_{i=1}^n (\mathbf{P}[Z(0) = 1, Z(h_n) = 0] + \mathbf{P}[Z(0) = 0, Z(h_n) = 1]) \\ = \frac{1}{h_n} (P_{10}(h_n) \cdot \eta_1 + P_{01}(h_n) \cdot \eta_0).$$

Hence by using equation 7.7 and expanding exponentials into Taylor series:

$$\mathbf{E} [(\hat{s} - s)] = \left(\frac{1}{h_n} (\eta_1 P_{10}(h_n) + \eta_0 \cdot P_{01}(h_n)) - \frac{2\lambda_0\lambda_1}{\lambda_0 + \lambda_1}\right) \\ = \left(\eta_1(\lambda_1 - \lambda_1 \cdot \frac{\lambda_0 + \lambda_1}{2} h_n + \mathcal{O}(h_n^2)) + \eta_0(\lambda_0 - \lambda_0 \cdot \frac{\lambda_0 + \lambda_1}{2} h_n + \mathcal{O}(h_n^2)) - \frac{2\lambda_0\lambda_1}{\lambda_0 + \lambda_1}\right) \\ = -h_n \cdot \lambda_0 \cdot \lambda_1 \xrightarrow{h_n \downarrow 0} 0.$$

ad ii): For the variance we define the shortcuts  $X_i := \mathbf{1}_{Z(ih_n)=1} \cdot \mathbf{1}_{Z((i+1)h_n)=0}$  and  $Y_i := \mathbf{1}_{Z(ih_n)=0} \cdot \mathbf{1}_{Z((i+1)h_n)=1}$ . Then:

$$\begin{aligned}
\text{Var}(\hat{s}_n(Z)) &= \frac{1}{n^2 \cdot h_n^2} \text{Var}\left(\sum_{i=1}^n (X_i + Y_i)\right) = \frac{1}{n^2 \cdot h_n^2} \sum_{i,j=1}^n \text{Cov}(X_i + Y_i, X_j + Y_j) \\
&= \frac{1}{n^2 \cdot h_n^2} \sum_{i,j=1}^n (\mathbf{E}[(X_i + Y_i)(X_j + Y_j)] - \mathbf{E}[X_i + Y_i] \mathbf{E}[X_j + Y_j]) \\
&= \frac{1}{n^2 \cdot h_n^2} \sum_{i,j=1}^n (\mathbf{E}[X_i X_j] + 2 \mathbf{E}[X_i Y_j] + \mathbf{E}[Y_i Y_j] - \mathbf{E}[X_i] \mathbf{E}[X_j] \\
&\quad - 2 \mathbf{E}[X_i] \mathbf{E}[Y_j] - \mathbf{E}[Y_i] \mathbf{E}[Y_j])
\end{aligned}$$

Now we have a closer look at the ingredients of this last sum:

$$\begin{aligned}
\mathbf{E}[X_i] &= \mathbf{P}[Z(ih_n) = 1, Z((i+1)h_n) = 0] = \mathbf{P}[Z(0) = 1, Z(h_n) = 0] = P_{10}(h_n) \cdot \eta_1. \\
\mathbf{E}[Y_i] &= P_{01}(h_n) \cdot \eta_0 \\
\mathbf{E}[X_i X_{i+j}] &= \mathbf{P}[Z(0) = 1 \cap Z(h_n) = 0 \cap Z(jh_n) = 1 \cap Z((j+1)h_n) = 0] \\
&= \begin{cases} \mathbf{P}[Z(0) = 1, Z(h_n) = 0] = P_{10}(h_n) \cdot \eta_1 & (j = 0) \\ \mathbf{P}[Z(0) = 1, Z(h_n) = 0, Z(h_n) = 1, Z(2h_n) = 0] = 0. & (j = 1) \end{cases}
\end{aligned}$$

For  $\mathbf{E}[X_i X_{i+j}]$  when  $j \geq 2$  we can use conditional independence of  $Z(h_n)$  and  $Z((j+1)h_n)$  given  $Z(jh_n)$ :

$$\begin{aligned}
\mathbf{E}[X_i X_{i+j}] &= \mathbf{P}[Z(0) = 1 \cap Z(h_n) = 0 \cap Z(jh_n) = 1 \cap Z((j+1)h_n) = 0] \\
&= \mathbf{P}[Z(h_n) = 0 \cap Z((j+1)h_n) = 0 | Z(0) = 1, Z(jh_n) = 1] \mathbf{P}[Z(0) = 1, Z(jh_n) = 1] \\
&= \mathbf{P}[Z(h_n) = 0 | Z(0) = 1, Z(jh_n) = 1] \mathbf{P}[Z((j+1)h_n) = 0 | Z(jh_n) = 1] \mathbf{P}[Z(0) = 1, Z(jh_n) = 1] \\
&= \mathbf{P}[Z(h_n) = 0, Z(0) = 1, Z(jh_n) = 1] \mathbf{P}[Z((j+1)h_n) = 0 | Z(jh_n) = 1] \\
&= \mathbf{P}[Z((j+1)h_n) = 0 | Z(jh_n) = 1] \mathbf{P}[Z(jh_n) = 1 | Z(h_n) = 0] \mathbf{P}[Z(h_n) = 0 | Z(0) = 1] \mathbf{P}[Z(0) = 1] \\
&= P_{10}(h_n) P_{01}((j-1)h_n) P_{10}(h_n) \eta_1.
\end{aligned}$$

And similarly:

$$\begin{aligned}
\mathbf{E}[X_i Y_{i+j}] &= \mathbf{P}[Z(0) = 1 \cap Z(h_n) = 0 \cap Z(jh_n) = 0 \cap Z((j+1)h_n) = 1] \\
&= \begin{cases} 0 & (j = 0) \\ \mathbf{P}[Z(0) = 1, Z(h_n) = 0, Z(2h_n) = 1] = P_{01}(h_n) P_{10}(h_n) \cdot \eta_1 & (j = 1) \\ P_{01}(h_n) P_{00}((j-1)h_n) P_{10}(h_n) \cdot \eta_1. & (j \geq 2) \end{cases}
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}[Y_i Y_{i+j}] &= \mathbf{P}[Z(0) = 0 \cap Z(h_n) = 1 \cap Z(jh_n) = 0 \cap Z((j+1)h_n) = 1] \\
&= \begin{cases} \mathbf{P}[Z(0) = 0, Z(h_n) = 1] = P_{01}(h_n) \cdot \eta_0 & (j = 0) \\ 0 & (j = 1) \\ P_{01}(h_n) P_{10}((j-1)h_n) P_{01}(h_n) \cdot \eta_0. & (j \geq 2) \end{cases}
\end{aligned}$$

Substituting these terms back into the expression for the variance yields:

$$\begin{aligned}
\text{Var}(\hat{s}_n(Z)) &= \frac{1}{n^2 \cdot h_n^2} \underbrace{\left[ n \cdot P_{10}(h_n) \cdot \eta_1 + 2 \cdot \sum_{j=2}^{n-2} (n-j-1) P_{10}^2(h_n) P_{01}((j-1)h_n) \eta_1 \right]}_{\mathbf{E}[X_i X_{i+j}], j=0} \\
&\quad + \underbrace{2 \cdot \sum_{j=2}^{n-2} (n-j-1) P_{10}(h_n) P_{01}((j-1)h_n) \eta_1}_{\mathbf{E}[X_i X_{i+j}], |j| \geq 2} \\
&\quad + 2 \cdot \left\{ \underbrace{\left[ (n-2) P_{01}(h_n) P_{10}(h_n) \eta_1 + (n-2) P_{10}(h_n) P_{01}(h_n) \eta_0 \right]}_{\mathbf{E}[X_i Y_{i+j}], j=1} \right. \\
&\quad + \underbrace{\sum_{j=2}^{n-2} (n-j-1) P_{01}(h_n) P_{00}((j-1)h_n) P_{10}(h_n) \eta_1}_{\mathbf{E}[X_i Y_{i+j}], j \geq 2} \\
&\quad \left. + \underbrace{\sum_{j=2}^{n-2} (n-j-1) P_{10}(h_n) P_{11}((j-1)h_n) P_{01}(h_n) \eta_0}_{\mathbf{E}[X_i Y_{i+j}], j \leq -2} \right\} \\
&\quad + \underbrace{\frac{n \cdot P_{01}(h_n) \cdot \eta_0 + 2 \cdot \sum_{j=2}^{n-2} (n-j-1) P_{01}^2(h_n) P_{10}((j-1)h_n) \eta_0}_{\mathbf{E}[Y_i Y_{i+j}], j=0}}_{\mathbf{E}[Y_i Y_{i+j}], |j| \geq 2}} \\
&\quad - \left[ n^2 P_{10}^2(h_n) \eta_1^2 - 2n^2 P_{10}(h_n) P_{01}(h_n) \eta_0 \eta_1 - n^2 P_{01}^2(h_n) \eta_0^2 \right] \\
&= \frac{1}{n^2 \cdot h_n^2} \underbrace{\left[ n \cdot \eta_1 P_{10}(h_n) + n \cdot \eta_0 P_{01}(h_n) \right]}_{=n \cdot h_n \cdot \mathbf{E}[\hat{s}_n(Z)]} - n^2 \underbrace{\left[ P_{10}(h_n) \eta_1 + P_{01}(h_n) \eta_0 \right]}_{h_n \cdot \mathbf{E}[\hat{s}_n(Z)]} \\
&\quad + 2(n-2) P_{01}(h_n) P_{10}(h_n) \\
&\quad + 2 \sum_{j=2}^{n-2} (n-j-1) (P_{10}^2(h_n) P_{01}((j-1)h_n) \eta_1 + P_{01}(h_n) P_{00}((j-1)h_n) P_{10}(h_n) \eta_1 \\
&\quad + P_{10}(h_n) P_{11}((j-1)h_n) P_{01}(h_n) \eta_0 + P_{01}^2(h_n) P_{10}((j-1)h_n) \eta_0)].
\end{aligned}$$

These last sums can be seen to be of Riemannian nature and converge to simple integrals as  $n \rightarrow \infty, h_n \rightarrow 0$ . Exemplarily:

$$\begin{aligned}
&\sum_{j=2}^{n-2} (n-j-1) P_{10}^2(h_n) P_{01}((j-1)h_n) \eta_1 \\
&= P_{01}(h_n)^2 \eta_1 n \frac{\lambda_0}{\lambda_0 + \lambda_1} \sum_{j=2}^{n-2} \left( 1 - \frac{j+1}{n} \right) \left( 1 - e^{-(\lambda_0 + \lambda_1)(j-1)h_n} \right) \\
&\sim P_{01}(h_n)^2 \eta_1 n \frac{\lambda_0}{\lambda_0 + \lambda_1} \left( (n-3) - \left( \frac{n-1}{2} - \frac{3}{n} \right) - \frac{1}{h_n} \left( 1 - \frac{2}{n} \right) \int_0^\infty e^{-(\lambda_0 + \lambda_1)t} dt \right. \\
&\quad \left. + \frac{1}{nh_n^2} \int_0^\infty t e^{-(\lambda_0 + \lambda_1)t} dt \right) \\
&\sim h_n^2 n \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \right)^2 \left( \frac{n}{2} - \frac{5}{2} + \frac{3}{n} \right) + h_n n \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \left( \frac{2}{n} - 1 \right) + \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^4},
\end{aligned}$$

where we have used:

$$\int_0^{\infty} e^{-(\lambda_0+\lambda_1)t} dt = \frac{1}{\lambda_0 + \lambda_1}, \quad \int_0^{\infty} t \cdot e^{-(\lambda_0+\lambda_1)t} dt = \frac{1}{(\lambda_0 + \lambda_1)^2}.$$

The other Riemannian sums can be treated analogously:

$$\begin{aligned} & \sum_{j=2}^{n-2} (n-j-1) P_{01}(h_n) P_{10}(h_n) P_{00}((j-1)h_n) \eta_1 \\ \sim & h_n^2 n \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \right)^2 \left( \frac{n}{2} - \frac{5}{2} + \frac{3}{n} \right) - h_n n \frac{\lambda_0^3 \lambda_1}{(\lambda_0 + \lambda_1)^3} \left( \frac{2}{n} - 1 \right) - \frac{\lambda_0^3 \lambda_1}{(\lambda_0 + \lambda_1)^4} \\ & \sum_{j=2}^{n-2} (n-j-1) P_{10}(h_n) P_{01}(h_n) P_{11}((j-1)h_n) \eta_0 \\ \sim & h_n^2 n \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \right)^2 \left( \frac{n}{2} - \frac{5}{2} + \frac{3}{n} \right) - h_n n \frac{\lambda_0 \lambda_1^3}{(\lambda_0 + \lambda_1)^3} \left( \frac{2}{n} - 1 \right) - \frac{\lambda_0 \lambda_1^3}{(\lambda_0 + \lambda_1)^4} \\ & \sum_{j=2}^{n-2} (n-j-1) P_{01}^2(h_n) P_{10}((j-1)h_n) \eta_0 \\ \sim & h_n^2 n \left( \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \right)^2 \left( \frac{n}{2} - \frac{5}{2} + \frac{3}{n} \right) + h_n n \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \left( \frac{2}{n} - 1 \right) + \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^4}, \end{aligned}$$

Collecting terms finally leads us to the following expression for the variance:

$$\begin{aligned} \text{Var}(\hat{s}_n(Z)) \sim & \frac{2\lambda_0 \lambda_1}{nh_n} \left( \frac{(\lambda_0 - \lambda_1)^2}{(\lambda_0 + \lambda_1)^3} + \frac{1}{\lambda_0 + \lambda_1} \right) - \frac{4\lambda_0 \lambda_1 (\lambda_0 - \lambda_1)^2}{n^2 h_n (\lambda_0 + \lambda_1)^3} - \frac{2\lambda_0 \lambda_1 (\lambda_0 - \lambda_1)^2}{n^2 h_n^2 (\lambda_0 + \lambda_1)^4} \\ & + 2\lambda_0 \lambda_1 \frac{n-2}{n^2} + \frac{8}{n} \frac{\lambda_0^2 \lambda_1^2}{(\lambda_0 + \lambda_1)^2} \left( \frac{3}{n} - \frac{5}{2} \right). \end{aligned}$$

If  $h_n \downarrow 0$  slower than  $n^{-1}$  in the sense that  $h_n \cdot n \rightarrow \infty$  then the order of magnitude of the variance is given by the leading term in the above expression:

$$\text{Var}(\hat{s}_n(Z)) \sim \frac{2\lambda_0 \lambda_1}{nh_n} \left( \frac{(\lambda_0 - \lambda_1)^2}{(\lambda_0 + \lambda_1)^3} + \frac{1}{\lambda_0 + \lambda_1} \right) = 4\lambda_0 \lambda_1 \frac{\lambda_0^2 + \lambda_1^2}{(\lambda_0 + \lambda_1)^3}.$$

which had to be shown. □

**Theorem 7.3.3. (Asymptotic normality of the surface-estimator in the on-off-system)** Let  $Z$  be the stationary continuous on-off Markov chain with the generator given in equation 7.4  $A$  and initial distribution  $\eta$  in equation 7.8 observed on a refining sequence of lattices  $\Gamma_n$  containing  $n$  points and having lattice spacing  $h_n = n^{-1/3}$ . Let further the estimator  $\hat{s}_n(Z)$  be defined as in equation 7.9. Then:

$$n^{\frac{1}{3}} \cdot (\hat{s}_n(Z) - s) \xrightarrow{d} \mathcal{N} \left( -\lambda_0 \cdot \lambda_1, 4\lambda_0 \lambda_1 \frac{\lambda_0^2 + \lambda_1^2}{(\lambda_0 + \lambda_1)^3} \right).$$

**Proof. (Theorem 7.3.3)** Analogous to the proof of Theorem 7.1.1, part iv).

## 7.4 Proofs of Chapter 5

### 7.4.1 Converse Direction of Hammersley and Clifford (Theorem 5.2.2)

Let  $\Theta$  be a Markovian Random field on  $\Gamma$  and let  $\pi$  be its distribution on  $\mathcal{S}$ . We fix an arbitrary configuration  $v \in \mathcal{S}$  as a reference configuration and for  $s \in \mathcal{S}$  and  $B \subset \Gamma$  denote by  $s^B$  the configuration  $\in \mathcal{S}$  with  $s^B(x) = s(x)$  for  $x \in B$  and  $s^B(x) = v(x)$  for  $x \in B^c$ . Then we claim that  $U$  with  $U_\emptyset \equiv 0$  and for  $A \neq \emptyset, A \subset \Gamma$

$$U_A(s) := - \sum_{B \subset A} (-1)^{|A-B|} \log \pi(s^B), \quad (7.28)$$

is a potential for  $\Theta$ . This can be proved by Moebius-inversion. For  $B \subset \Gamma$  we set  $\Phi(B) := -U_B(s)$  and  $\Psi(B) = \log \left( \frac{\pi(s^B)}{\pi(v)} \right)$ . Then we have for  $A \neq \emptyset, A \subset \Gamma$ :

$$\Phi(A) = \sum_{B \subset A} (-1)^{|A-B|} \log \pi(s^B) - \log \pi(v) \sum_{B \subset A} (-1)^{|A-B|} = \sum_{B \subset A} (-1)^{|A-B|} \Psi(B), \quad (7.29)$$

where we have used  $\sum_{B \subset A} (-1)^{|A-B|} = \sum_{k=0}^{|A|} \binom{|A|}{k} (-1)^k = (1-1)^{|A|} = 0$ .

For  $A = \emptyset$  equation 7.29 holds too because of:

$$\Phi(\emptyset) = -U_\emptyset(s) = 0 = \log \left( \frac{\pi(s^\emptyset)}{\pi(v)} \right) = \Psi(\emptyset).$$

Hence Moebius-inversion results in:

$$\log \left( \frac{\pi(s)}{\pi(v)} \right) = \Psi(\Gamma) = \sum_{B \subset \Gamma} \Phi(B) = - \sum_{B \subset \Gamma} U_B(s) = -H_U(s),$$

where  $H_U(s)$  is the energy function associated with  $U$ . Thus  $\pi(s) = \pi(v) \cdot \text{Exp}(-H_U(s))$ , and because  $\pi$  is a distribution we conclude  $\pi(v) = Z^{-1}$ , which establishes the claim that  $U$  is a potential.

Let further  $A \neq \emptyset, A \subset \Gamma$ . If we now choose a lattice point  $a \in A$ , we can rewrite equation 7.28 as:

$$\begin{aligned} U_A(x) &= - \sum_{B \subset A-a} (-1)^{|A-B|} (\log(\pi(s^B)) - \log(\pi(s^{B+a}))) \\ &= - \sum_{B \subset A-a} (-1)^{|A-B|} \log \left( \frac{\text{P}[\Theta(a) = s^B(a) | \Theta(x) = s^B(x), x \neq a]}{\text{P}[\Theta(a) = s^{B+a}(a) | \Theta(x) = s^{B+a}(x), x \neq a]} \right). \end{aligned} \quad (7.30)$$

It remains to show that  $U$  is in fact a neighbourhood-potential. To do this, let  $A \subset \Gamma$  be a set which is not a clique. Then there exist  $a, b \in A$  such that  $b \in A \setminus \delta(a)$ . Now, equation 7.30 can further be transformed into:

$$\begin{aligned} U_A(x) &= - \sum_{B \subset A \setminus \{a, b\}} (-1)^{|A-B|} \log \left( \frac{\text{P}[\Theta(a) = s^B(a) | \Theta(x) = s^B(x), x \neq a]}{\text{P}[\Theta(a) = s^{B+b}(a) | \Theta(x) = s^{B+b}(x), x \neq a]} \right) \times \\ &\quad \frac{\text{P}[\Theta(a) = s^{B+a+b}(a) | \Theta(x) = s^{B+a+b}(x), x \neq a]}{\text{P}[\Theta(a) = s^{B+a}(a) | \Theta(x) = s^{B+a}(x), x \neq a]} \end{aligned}$$

Now note that we have the following equations:

$$\begin{aligned} \{\Theta(a) = s^B(a)\} &= \{\Theta(a) = s^{B+b}(a)\} \\ \mathbb{P} \left[ \Theta(a) = s^B(a) \mid \Theta(x) = s^{B+b}(x), x \neq a \right] &= \mathbb{P} \left[ \Theta(a) = s^B(a) \mid \Theta(x) = s^B(x), x \neq a \right], \end{aligned}$$

the second of which being true, because  $\Theta$  is Markovian and  $b \notin \delta(a)$ . This shows that the first quotient in the logarithm equals one. The same argument applies for the second quotient, which finally gives us  $U_A(s) \equiv 0 \forall A \subset \Gamma, A$  not a clique. This concludes the proof.  $\square$

### 7.4.2 Proof of the Equivalence of Ensembles (Theorem 5.5.3)

**Proof. (Theorem 5.5.3)** In a first step we show that the Julész-ensemble in fact has the Markov-property and thus it is a Gibbs-ensemble by a general version of the Hammersley-Clifford-Theorem 5.2.2. To do this we need the following lemma:

**Lemma 7.4.1.**

$$\lim_{\Gamma_n \rightarrow \mathbb{Z}^d} \frac{1}{|\Gamma_n|} \log |\Omega_{\Gamma_n}(v)| = s(v),$$

where  $s : \mathbb{R}^n \rightarrow \mathbb{R}$  is strictly concave (i.e. its Hessian is negative-definite everywhere).

**Proof. (Lemma 7.4.1)** see (Lanford 1973).

We now consider the set  $\Omega_{\Gamma_n}(v)$  and its uniform distribution  $U_{n,v}$  over a large  $\Gamma_n$  and choose arbitrary but fixed sub-lattices  $\Gamma^0 \subset \Gamma^1 \subset \Gamma_n$  such that the neighbourhood  $\delta\Gamma^0$  is completely contained in  $\Gamma^1$ . We further denote  $\Gamma_n^{-1} := \Gamma_n \setminus \Gamma^1$  and  $\Gamma^{01} := \Gamma^1 \setminus \Gamma^0$ . According to equations 5.8 the specific feature statistics  $\phi_j$  are additive in the sense that

$$\phi_j^{\Gamma_n^{-1}} = v_j \cdot |\Gamma_n| - \phi_j^{\Gamma^{01}} - \phi_j^{\Gamma^0},$$

where  $\phi_j^A$  means the contribution to the feature statistic  $\phi_j$  if only contributions on a sub-lattice  $A \subset \Gamma_n$  are counted. The number of configurations in  $\Omega_{\Gamma_n}(v)$  which exhibit a given sub-configuration  $s_{\Gamma^0}$  on  $\Gamma^0$  with joint local nonspecific feature statistic values  $V_{\Gamma^0}$  and  $s_{\Gamma^{01}}$  on  $\Gamma^{01}$  with joint local nonspecific feature statistic values  $V_{\Gamma^{01}}$  is thus equal to:

$$\left| \Omega_{\Gamma_n^{-1}} \left( \frac{v \cdot |\Gamma_n| - V_{\Gamma^{01}} - V_{\Gamma^0}}{|\Gamma_n^{-1}|} \right) \right|.$$

Hence when sampling at random from  $\Omega_{\Gamma_n}(v)$  for the conditional distribution  $\mathbb{P}[s_{\Gamma^0} | s_{\Gamma^{01}}]$  we obtain:

$$\mathbb{P}[s_{\Gamma^0} | s_{\Gamma^{01}}] \propto \left| \Omega_{\Gamma_n^{-1}} \left( \frac{v \cdot |\Gamma_n| - V_{\Gamma^{01}} - V_{\Gamma^0}}{|\Gamma_n^{-1}|} \right) \right| = \left| \Omega_{\Gamma_n^{-1}} \left( v' - \frac{V_{\Gamma^0}}{|\Gamma_n^{-1}|} \right) \right|, \quad (7.31)$$

where we have abbreviated  $v' = \frac{v \cdot |\Gamma_n| - V_{\Gamma^{01}}}{|\Gamma_n^{-1}|}$  which can be regarded as constant for the above conditional probability. Equation 7.31 shows that up to a constant,  $\mathbb{P}[s_{\Gamma^0} | s_{\Gamma^{01}}]$  is uniform and thus Lemma 7.4.1 can be applied to get for large  $\Gamma_n$ :

$$\log \mathbb{P}[s_{\Gamma^0} | s_{\Gamma^{01}}] = \text{const} + |\Gamma_n^{-1}| \cdot s \left( v' - \frac{V_{\Gamma^0}}{|\Gamma_n^{-1}|} \right),$$

which allows a Taylor-expansion of  $s$  around  $v'$  to arrive at:

$$\begin{aligned} \log P [s_{\Gamma^0} | s_{\Gamma^0}] &= \text{const} + |\Gamma_n^{-1}| \cdot s(v') - s'(v') \cdot V_{\Gamma^0} + \mathcal{O}\left(\frac{1}{|\Gamma_n|}\right) \\ &= \text{const} - s'(v') \cdot V_{\Gamma^0} + \mathcal{O}\left(\frac{1}{|\Gamma_n|}\right). \end{aligned}$$

If we now note that  $v' \rightarrow v$  for  $\Gamma_n \rightarrow \mathbb{Z}^d$  we conclude that for  $\Gamma_n \rightarrow \mathbb{Z}^d$  we have:

$$P [s_{\Gamma^0} | s_{\Gamma^0}] \xrightarrow{\Gamma_n \rightarrow \mathbb{Z}^d} \text{Exp} (\text{const} - s'(v) \cdot V_{\Gamma^0}) = P [s_{\Gamma^0} | s_{\delta\Gamma^0}],$$

because only  $V_{\Gamma^0}$  plays a role which can be computed from  $\Gamma^0 \cup \delta\Gamma^0$ . This is seen to be the Markovian property for the Gibbsian conditional distribution if we choose  $\beta = s'(v)$  and use that the mass of the distribution must sum up to one.

For the converse direction we start with the Gibbsian distribution  $\pi_{G,n}$  as given in definition 5.5.2 on  $\Gamma_n$  and argue that it concentrates its mass uniformly on the Julész-ensemble. The probability  $P [s \in \Omega_{\Gamma_n}(v)]$  under the Gibbsian Model is:

$$P [s \in \Omega_{\Gamma_n}(v)] = \frac{1}{Z_n(\beta)} \text{Exp} (-|\Gamma_n| \beta \cdot v) \cdot 2^{|\Gamma_n|} \cdot U_{\Gamma_n}(\Omega_{\Gamma_n}(v)),$$

where  $\beta \cdot v$  denotes a scalar product between vectors and  $Z_n(\beta)$  is the normalising constant of the Gibbs distribution on  $\Gamma_n$ . For the probability rate function one then gets:

$$\begin{aligned} \lim_{\Gamma_n \rightarrow \mathbb{Z}^d} \frac{1}{|\Gamma_n|} \log P [s \in \Omega_{\Gamma_n}(v)] &= -\beta \cdot v + \lim_{\Gamma_n \rightarrow \mathbb{Z}^d} \frac{1}{|\Gamma_n|} \log U_{\Gamma_n}(\Omega_{\Gamma_n}(v)) \\ &\quad - \left( \lim_{\Gamma_n \rightarrow \mathbb{Z}^d} \frac{1}{|\Gamma_n|} \log Z_n(\beta) - \log 2 \right) \\ &= s(v) - \beta \cdot v - \rho(\beta), \end{aligned}$$

where we have again used Lemma 7.4.1 for the second summand and for the third we used:

**Lemma 7.4.2.**

$$\rho(\beta) := \lim_{\Gamma_n \rightarrow \mathbb{Z}^d} \frac{1}{|\Gamma_n|} \log Z_n(\beta) - \log 2$$

*exists independent of any boundary conditions and is a strictly convex function.*

**Proof.** (Lemma 7.4.2) see (Griffiths and Ruelle 1971).

This means that the probability mass the Gibbsian distribution assigns to  $\Omega_{\Gamma_n}(v)$  is of the following order:

$$P [s \in \Omega_{\Gamma_n}(v)] \sim \text{Exp} (|\Gamma_n| \cdot (s(v) - \beta \cdot v - \rho(\beta))).$$

But this implies that

$$s_\beta(v) := s(v) - \beta \cdot v - \rho(\beta) \leq 0 \tag{7.32}$$

must hold and there must be a value  $v_0$  for which  $s_\beta(v_0) = 0$  otherwise the probability mass would become completely annihilated in the thermodynamic limit. But then the



function  $s_\beta(\cdot)$  obviously attains its (global) maximum at  $v_0$  from which we immediately rediscover the relationship we used further above:

$$\beta = s'(v_0).$$

If the value  $v_0$  is unique it characterizes the Júlész-ensemble to which the Gibbs ensemble is equivalent. The uniqueness of  $v_0$  is given whenever the Gibbs model shows no phase-transitions, i.e. there is only one Gibbs-measure possible as a thermodynamic limit for a given potential. Sufficient conditions for the absence of phase transitions are the *Dobrushin-condition* and the *Simon-condition* which are discussed in detail in (Guyon 1995, section 2.1.3.1).  $\square$



# Bibliography

- Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions*, Dover Books on Advanced Mathematics, 9th edn, Dover, New York.
- Adler, R. J. (1981). *The Geometry of Random Fields*, Wiley Series in Probability and Mathematical Statistics.
- Advanced Visual Systems (2006). *AVS5 Software*, Advanced Visual Systems, Waltham MA 02451, USA.
- Bradley, R. C. (1993). Equivalent mixing conditions for random fields, *Annals of Probability* **21**(4): 1921–1926.
- Comets, F. (1992). On consistency of a class of estimators for exponential families of markov random fields on a lattice, *Annals of Statistics* **20**: 455–468.
- Cramér, H. and Leadbetter, M. (1968). *Stationary and Related Stochastic Processes*, John Wiley and Sons, New York.
- Daley, D. and Vere-Jones, D. (2002). *An Introduction to the Theory of Point Processes*, Probability and Its Applications, 2nd edn, Springer, Heidelberg.
- Davy, P. (1976). Projected thick sections through multi-dimensional particle aggregates, *Journal of Applied Probability* **13**: 714–722.
- Davy, P. (1978). *Stereology - A Statistical Viewpoint*, PhD thesis, Australian National University, Canberra.
- Frisch, H. and Stillinger, F. (1963). Contribution to the statistical geometric basis of radiation scattering, *Journal of Chemical Physics* **38**: 2200–2207.
- Geyer, C. J. (1991). Estimating normalizing constants and reweighting mixtures in markov chain monte carlo, *Technical Report 568*, University of Minnesota.
- Gneiting, T., Sevcikova, H., Percival, D. B., Schlather, M. and Jiang, Y. (2006). Fast and exact simulation of large gaussian lattice systems in  $\mathbb{R}^2$ : Exploring the limits, *Journal of Computational and Graphical Statistics* **15**(3): 483–501.
- Griffiths, R. and Ruelle, D. (1971). Strict convexity of the pressure in a lattice system, *Comm. Math. Phys.*
- Guyon, X. (1995). *Random Fields on a Network: Modeling, Statistics, and Applications*, Probability and Its Applications, Springer-Verlag, New York, Heidelberg, Berlin.
- Hadwiger, H. (1957). *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag.

- Ibragimov, I. (1971). *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing Groningen, Netherlands.
- Ibragimov, I. and Rozanov, Y. (1978). *Gaussian Random Processes*, Springer-Verlag, New York, Heidelberg, Berlin.
- Kendall, M. G. and Stuart, A. (1969). *The Advanced Theory of Statistics*, Griffin, London.
- Klain, D. (1995). A short proof of hadwiger's characterization theorem, *Mathematika* **42**: 329–339.
- Kratz, M. F. (2006). Level crossings and other level functionals of stationary gaussian processes, *Probability Surveys* **3**: 230–288.
- Kratz, M. F. and León, J. (2006). On the second moment of the number of crossings by a stationary gaussian process, *Annals of Probability*.
- Lahiri, S. (2003). *Resampling Methods for Dependent Data*, Springer Series in Statistics, Springer-Verlag.
- Lanford, O. (1973). *Entropy and equilibrium states in classical mechanics*, Statistical Mechanics and Mathematical Problems, Springer-Verlag.
- Lehmann, P., Berchtold, M., Ahrenholz, B., Tölke, J., Krafczyk, M., Flühler, H. and Künsch, H.-R. (n.d.). Impact of geometrical properties on permeability and fluid phase distribution in porous media, *To appear in Advances in Water Resources*.
- Mecke, K. and Wagner, H. (1991). Euler characteristic and related measures for random geometric sets, *Journal of Statistical Physics* **64**(3/4): 843–850.
- Mecke, K. R. (2001). Exact moments of curvature measures in the boolean model, *Journal of Statistical Physics* **102**(5/6): 1343–1381.
- Miles, R. (1976). Estimating aggregate and overall characteristics from thick sections by transmission microscopy, *Journal of Microscopy* **107**: 227–233.
- Molchanov, I. (2005). *Theory of Random Sets*, Probability and Its Applications, 2nd edn, Springer, Heidelberg.
- Morse, M. and Cairns, S. (1969). *Critical Point Theory in Global Analysis and Differential Topology*, Academic Press, New York.
- Moscinski, J., Bargiel, M., Rycerz, Z. and Jacobs, P. (1989). The force-biased algorithm for the irregular close packing of equal hard spheres, *Mol. Simulat.* **3**: 201–212.
- Nagel, W., Ohser, J. and Pischang, K. (2000). An integral-geometric approach for the euler-poincaré-characteristic of spatial images, *Journal of Microscopy* **198**: 54–62.
- Ohser, J. and Mücklich, F. (2000). *Statistical Analysis of Microstructures in Materials Science*, John Wiley and Sons, LTD.
- Ohser, J. and Nagel, W. (1996). The estimation of the euler-poincaré-characteristic from observations on parallel sections, *Journal of Microscopy* **184**: 117–126.
- Ohser, J., Nagel, W. and Schladitz, K. (2002). The euler number of discretized sets- on choice of adjacency in homogeneous lattices, *Morphology of Condensed Matter*.

- Ohser, J., Nagel, W. and Schladitz, K. (2003). The euler number of discretised sets-surprising results in three dimensions, *Image Anal. Stereol.* **22**: 11–19.
- Peligrad, M. (1996). On the asymptotic normality of sequences of weak dependent random variables, *Journal of Theoretical Probability* **9**: 703–715.
- Piterbarg, V. (1978). The central limit theorem for the number of level crossings of a stationary gaussian process, *Theory of Probab. Applications* **23**: 178–182.
- R Development Core Team (2006). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0
- Resnick, S. I. (1992). *Adventures in Stochastic Processes*, Birkhäuser Boston, Basel, Berlin.
- Ribeiro Jr., P. and Diggle, P. (2001). *geoR: A package for geostatistical analysis*. R package version 1.6.14
- Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods*, Springer Texts in Statistics, Springer-Verlag, New York, Heidelberg, Berlin.
- Santaló, L. (2004). *Integral Geometry and Geometric Probability*, 2nd edn, Cambridge Mathematical Library.
- Schlather, M. (1999). Introduction to positive definite functions and to unconditional simulation of random fields, *Technical report*, Department of Mathematics and Statistics, Lancaster University, UK.
- Schlather, M. (2001). *RandomFields: Simulation and Analysis of Random Fields*. R package version 1.3.29
- Schmidt, V. and Spodarev, E. (2005). Joint estimators for the specific intrinsic volumes of stationary random sets, *Stochastic Processes and their Applications* **115**: 959–981.
- Schneider, R. (1993). *Convex Bodies: The Brunn-Minkowski-Theory*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press.
- Schneider, R. (2004). Integral geometric tools for stochastic geometry, *Lecture Notes from CIME-Conference on Stochastic Geometry in Martina Franca (Taranto, Italy)*.
- Schneider, R. and Weil, W. (1992). *Integralgeometrie*, Teubner, Stuttgart.
- Schneider, R. and Weil, W. (2000). *Stochastische Geometrie*, Teubner, Stuttgart.
- Serra, J. (1982). *Image Analysis and Mathematical Morphology*, Academic Press, New York.
- Stoyan, D. and Kadashevich, I. (2005). Aerated autoclaved concrete: Stochastic structure model and elastic properties, *Proc. Appl. Math. Mech.* **5**(1): 419–420.
- Stoyan, D., Kendall, W. S. and Mecke, J. (1985). *Stochastic Geometry and its Applications*, John Wiley and Sons, Chichester.
- Talukdar, M., Torsaeter, M., Ioannidis, M. and Howard, J. (2002). Stochastic reconstruction of chalk from 2d images, *Transport in Porous Media* **48**: 101–123.

- Tomita, H. (1990). Statistics and geometry of random interface systems, *Formation, Dynamics and Statistics of Patterns* **1**: 113–157.
- Torquato, S. (2002). *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*, Vol. 16 of *Interdisciplinary applied mathematics*, Springer, New York.
- Winkler, G. (1995). *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A mathematical Introduction*, Applications of Mathematics, Springer-Verlag, New York, Heidelberg, Berlin.
- Wu, Y. N., Zhu, S. C. and Liu, X. (1999). Equivalence of julész and gibbs texture ensembles, *Proceedings of the International Conference on Computer Vision* **2**: 1025–1032.
- Wu, Y. N., Zhu, S. C. and Liu, X. (2000). Equivalence of julész ensembles and frame models, *International Journal on Computer Vision* **38**(3): 247–265.
- Zhu, S. C., Liu, X. and Wu, Y. N. (2000). Exploring texture ensembles by efficient markov chain monte carlo - toward a trichromacy theory of texture, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(3): 554–569.

# Curriculum Vitae

- 1976 Born on December 17 in Olten (SO)
- 1985 – 1991 Primary school in Olten (SO)
- 1991 – 1993 Secondary school (Bezirksschule) in Olten (SO)
- 1993 – 1997 Kantonsschule (Matura Typus C) in Olten (SO)
- 1997 – 2002 Studies in mathematics at the Swiss Federal Institute of Technology (ETH) Zurich
- 2002 Diploma in mathematics  
Diploma thesis (master's thesis) in statistics,  
"Simulation von Diffusionsprozessen aus diskreten Beobachtungen",  
Seminar for Statistics, ETH Zurich
- 2003 – 2007 PhD Student and Teaching Assistant at the Seminar for Statistics, ETH Zurich