Asymptotic Analysis of Stationary Random Tessellations with Applications to Network Modelling

Dissertation

zur Erlangung des Doktorgrades Dr. rer. nat.
der Fakultät für Mathematik und Wirtschaftswissenschaften
der Universität Ulm

vorgelegt von
Hendrik Schmidt
aus
Heidenheim an der Brenz

Juli 2006
Amtierender Dekan:  Prof. Dr. Ulrich Stadtmüller
1. Gutachter:       Prof. Dr. Volker Schmidt
2. Gutachter:       Prof. Dr. Ulrich Stadtmüller
3. Gutachter:       Prof. Dr. Lothar Heinrich
Acknowledgement

First and foremost, I would like to express my gratitude to Prof. Dr. V. Schmidt for giving me the opportunity to pursue my research interests at the Department of Stochastics in the framework of a joint research project with France Télécom R&D Division, Paris, on the modelling of telecommunication networks. I am deeply indebted to him for his guidance throughout this process as well as for his constant availability and his advice during the whole project.

My thanks go also to Prof. Dr. U. Stadtmüller for his interest in my work and his kind agreement to report about it.

Furthermore, I am very grateful to Prof. Dr. L. Heinrich from the University of Augsburg for many agreeable and fruitful discussions, for the collaboration that led to the results on central limit theorems, and his availability to also write a report about this thesis.

I would like to express my gratitude to Dr. C. Gloaguen for giving me the opportunity to learn more about telecommunication during my internship at France Télécom R&D Division in Paris.

This is also the place where I would like to say thank you to Jun.-Prof. Dr. E. Spodarev for interesting conversations and advice as well as to all my colleagues for the friendly atmosphere. Especially, I would like to mention Dipl.-Math.oec. F. Fleischer and thank him for numerous enjoyable discussions during our study and work together. Furthermore, I thank S. Eckel for her assistance with some simulations and for proof reading. With regard to the latter, I also acknowledge the help of Dr. S. Klenk and S. Lück. Furthermore, I appreciate the pleasant cooperation with several students at the Department of Stochastics who prepared their Diploma Thesis on aspects of network modelling during the recent years, in particular Dipl.-Math.oec. S. Hörner, Dipl.-Math.oec. A. Hornischer, and Dipl.-Math.oec. R. Thiedmann.

Last but not least, the warmest thank you of all goes to my parents for their love, support, and encouragement.
Contents

1 Introduction .......................................... 7
   1.1 Stochastic Network Modelling ......................... 8
       1.1.1 Telecommunication Access Networks ............... 9
       1.1.2 Stochastic Subscriber Line Model .............. 10
   1.2 Motivation .................................. 14
   1.3 Organisation of this Thesis ..................... 15
   1.4 The GeoStoch Library .......................... 17

2 Concepts of Stochastic Geometry ..................... 19
   2.1 Random Closed Sets ................................ 20
       2.1.1 The Space of Closed Sets ................... 20
       2.1.2 Definition and Properties .................. 20
   2.2 Random Point Processes .......................... 21
       2.2.1 Definition and Representation ................ 22
       2.2.2 Distribution and Basic Properties .......... 23
       2.2.3 Intensity Measure ......................... 23
       2.2.4 Palm Probabilities ........................ 24
       2.2.5 Poisson Point Processes .................. 25
   2.3 Random Marked Point Processes ................... 26
       2.3.1 Definition and Basic Properties .............. 26
       2.3.2 Intensity Measure and Palm Probabilities .... 27
       2.3.3 Ergodicity and Mixing .................... 28
2.4 Point Processes of Closed Sets ........................................... 29
2.5 Flat Processes .................................................................. 31
   2.5.1 Processes of $k$-Flats ............................................. 31
   2.5.2 Hyperplane Processes ........................................... 32
   2.5.3 Poisson Hyperplane Processes .............................. 33

3 Random Tessellations ......................................................... 35
   3.1 Introduction ................................................................. 35
      3.1.1 Polytopes and their Faces .................................. 35
      3.1.2 Tessellations in $\mathbb{R}^d$ ............................... 36
   3.2 Random Tessellations and their Representations ............... 37
      3.2.1 Definition and Basic Properties .............................. 37
      3.2.2 Marked Point Process Representation and Typical Cell .... 38
      3.2.3 Facet Processes ................................................. 39
   3.3 Random Tessellation Models ......................................... 40
      3.3.1 Poisson-Voronoi Tessellation ............................. 40
      3.3.2 Poisson-Delaunay Tessellation ............................ 41
      3.3.3 Poisson Hyperplane Tessellation ....................... 43
   3.4 Stationary Iterated Tessellations ..................................... 44
      3.4.1 Description of the Model .................................... 44
      3.4.2 Mean Value Relationships ................................... 48
   3.5 Asymptotic Behavior of Cells of Stationary Tessellations ........ 49
      3.5.1 Asymptotic Negligibility of Partially Observed Cells .... 50
      3.5.2 A View on the Assumption of Theorem 3.2 ........... 54

4 CLTs for Poisson Hyperplane Processes ................................ 55
   4.1 CLTs for Intersection Nodes of Poisson Hyperplanes .......... 57
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.4</td>
<td>A CLT with Random Normalization</td>
<td>68</td>
</tr>
<tr>
<td>4.2</td>
<td>CLTs for Flat Processes induced by Poisson Hyperplanes</td>
<td>69</td>
</tr>
<tr>
<td>4.2.1</td>
<td>CLTs for $k$-Flat Processes</td>
<td>70</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Explicit Moment Formulae</td>
<td>74</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Concluding Remarks</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>Multivariate Extensions</td>
<td>78</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Multivariate CLTs and Moment Formulae</td>
<td>79</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Explicit Formulae for Asymptotic Covariance Matrices</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td><strong>Statistical Inference for Tessellation Intensities</strong></td>
<td>85</td>
</tr>
<tr>
<td>5.1</td>
<td>Estimators and Tests for $k$-Flat Intensities</td>
<td>85</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Definition; Unbiasedness and Consistency</td>
<td>85</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Asymptotic Second-Order Properties</td>
<td>86</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Asymptotic Normality and Variance Stabilization</td>
<td>87</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Asymptotic Confidence Intervals and Tests</td>
<td>89</td>
</tr>
<tr>
<td>5.2</td>
<td>Some Extensions</td>
<td>90</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Confidence Intervals for Facet Intensities</td>
<td>90</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Multivariate Inference for Intensity Vectors</td>
<td>92</td>
</tr>
<tr>
<td>5.3</td>
<td>Inference for Poisson–Voronoi Tessellations</td>
<td>94</td>
</tr>
<tr>
<td>5.3.1</td>
<td>A CLT for the Vertices of Poisson–Voronoi Tessellations</td>
<td>95</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Asymptotic Confidence Intervals and Tests</td>
<td>95</td>
</tr>
<tr>
<td>5.4</td>
<td>Numerical Evaluations</td>
<td>97</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Goodness-of-Fit Tests for Poisson Line Processes</td>
<td>97</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Evaluation of Asymptotic Tests</td>
<td>102</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Goodness-of-Fit Tests for Poisson–Voronoi Tessellations</td>
<td>104</td>
</tr>
<tr>
<td>5.5</td>
<td>Applications in Network Modelling</td>
<td>104</td>
</tr>
</tbody>
</table>
6 CLTs for Stationary Tessellations with Random Inner Cell Structures 107
   6.1 Moments of Cumulative Random Functionals .................................. 110
       6.1.1 Expectation Vector and Covariance Matrix ................................. 110
       6.1.2 Integrability Conditions ......................................................... 115
   6.2 Laws of Large Numbers ................................................................. 117
       6.2.1 Deterministic Functionals .......................................................... 117
       6.2.2 Random Functionals ................................................................. 120
   6.3 Multivariate CLT .............................................................................. 125
       6.3.1 Some Auxiliary Results ............................................................... 125
       6.3.2 Proof of Theorem 6.4 .................................................................. 127
   6.4 Numerical Examples ......................................................................... 130
       6.4.1 Evaluation of Integrability Conditions ......................................... 131
       6.4.2 Simulation of Asymptotic Covariance Matrices ............................ 132
       6.4.3 Comparison of Empirical and Asymptotic Distributions ................ 133
   6.5 Applications to Network Modelling .................................................... 137

7 Statistical Model Fitting of Random Tessellations 141
   7.1 Some Preliminaries ........................................................................... 142
       7.1.1 Choice of Model Characteristics ................................................. 142
       7.1.2 Sampling Techniques and Unbiased Estimators ............................. 143
       7.1.3 Distance Functions ..................................................................... 145
   7.2 Optimal Model Choice for Non-iterated Tessellations ......................... 146
       7.2.1 Minimization Problem .................................................................. 146
       7.2.2 Minimization Techniques ............................................................. 147
       7.2.3 Numerical Examples .................................................................. 148
   7.3 Optimal Model Choice for Iterated Tessellations ................................. 150
       7.3.1 Minimization Problem .................................................................. 150
       7.3.2 Minimization Techniques ............................................................. 151
       7.3.3 Numerical Examples .................................................................. 153
   7.4 Model Check by Monte Carlo Tests ...................................................... 156
List of Tables 203
Nomenclature 205
Index 211
Chapter 1

Introduction

The present thesis arose out of an ongoing joint research project between the Department of Stochastics at the University of Ulm and France Télécom R&D Division, Paris. In particular, it was stimulated by questions regarding statistical model fitting of random tessellation models to real infrastructure data which exhibit a network structure similar to the infrastructure data of Paris in Figure 1.1. In the framework of this joint research project, the Stochastic Subscriber Line Model (SSLM) has been developed and is still extended in various directions. The SSLM is a flexible and general model for telecommunication networks, in particular for access networks in urban areas, which applies tools and methods from stochastic geometry.

In Section 1.1 we provide insight into stochastic network modelling, introduce briefly the SSLM, and give a short example of network analysis in the SSLM, where this analysis is based
on a particular underlying geometric model. In Section 1.2 we then motivate this thesis and
describe its goal before we give an overview of the contents in Section 1.3. We close this
chapter by some remarks on the GeoStoch software library in Section 1.4.

1.1 Stochastic Network Modelling

In order to adequately represent telecommunication networks, *spatial stochastic models* have
been developed and investigated during recent years. Such models are regarded as an alterna-
tive to more traditional approaches for cost measurement and strategic planning. They allow
for the incorporation of both stochastic and geometric features inherent in telecommunication
networks and provide therefore an accurate view both on location dependent network
characteristics and on the network’s variability in space and time.

There exist a number of popular examples in the field of telecommunication, *switching net-
works, multi–cast networks*, and *mobile telecommunication systems* to name just a few, where
stochastic–geometric models have been applied so far. Models that have been studied are
Poisson–Voronoi aggregated tessellations ([5], [36], and [118]), superpositions of Poisson–
Voronoi tessellations ([4]), spanning trees ([6], [7]), and coverage processes ([3]). Recently,

![Population density in Germany for the year 1996](image1)

![Realization of a modulated Poisson–Voronoi tessellation](image2)

(a) Population density in Germany for the year 1996

(b) Realization of a modulated Poisson–Voronoi tessellation

Figure 1.2: Nationwide telecommunication network modelling

time into account the spatial inhomogeneity of population densities as displayed in Figure
1.1. Stochastic Network Modelling

1.2 (a) for Germany. Adequate models for such large-scale networks seem to be modulated Poisson–Voronoi tessellations as shown in Figure 1.2 (b). Further descriptions of this modelling approach can be found in [18] and [19] as well as in [35], [97], and [120]. In the following we restrict ourselves to the case of urban access networks in telecommunication. Before introducing the SSLM in Section 1.1.2 we take a brief look at access networks in Section 1.1.1.

1.1.1 Telecommunication Access Networks

Telecommunication is a fastly evolving field of study. On the one hand this is due to the process of digitalization of the networks, replacing copper based cables by fibre optic technologies which results in improved operation and maintenance possibilities, high quality transmission, and flexibility in bandwidth for future demands. On the other hand, copper based networks have seen a remarkable comeback in recent years, especially in local networks, due to the introduction of the xDSL family of technologies, where xDSL stands for digital subscriber line, and due to the costs of digitalization.

A telecommunication network is a system to provide interconnections of subscribers for the exchange of information by voice, text, image, and video, to send and receive signals to and from the subscribers, respectively, in order to establish, maintain, and dismantle connections, as well as to provide additional services like billing or wake-up calls. The access network

**(a) Hierarchical physical link between a subscriber and its Wire Center Station**

**(b) Tree structure of a WCS subnetwork (without links between ND and CS)**

Figure 1.3: Hierarchical structure of access networks

or local loop is the part of the telecommunication network that connects a subscriber to its corresponding Wire Center Station (WCS). The hierarchical physical link shown in Figure 1.3 (a) is set up via network components like the Network Interface Device (ND), secondary and primary cabinets (CS and CP), and a Service Area Interface (SAI). In this way, a serving zone is associated to each WCS. The subnetwork that gathers all the links between the WCS and the subscribers in the corresponding serving zone displays a tree structure as shown in Figure 1.3 (b).

From the economic point of view, the (urban) access network can be considered as the place
where the telecommunication network depends strongly on the architecture of the (urban) infrastructure.

1.1.2 Stochastic Subscriber Line Model

The Stochastic Subscriber Line Model (SSLM) is a model for (urban) access networks, developed in recent years ([38] and [69]) in the framework of a joint research project between the Department of Stochastics at the University of Ulm and France Télécom R&D Division, Paris.

The SSLM is a spatial stochastic model that offers tools from stochastic geometry in order to describe the spatial irregularity as well as the geometric features of access networks. Based for example on the computation of mean shortest path lengths or mean subscriber line lengths ([41], [42]), the SSLM allows for stochastic econometrical analysis of network characteristics like connection costs, but also offers flexible and tractable models for strategic network planning. In particular, simple mean value formulae provided for network characteristics add to a reduction of complexity of the models and serve as benchmarks for cost evaluation, performance analysis, and strategic planning.

As Figure 1.4 shows, the modelling framework of the SSLM is subdivided into the Network Geometry Model, the Network Component Model, and the Network Topology Model. The

<table>
<thead>
<tr>
<th>Stochastic Subscriber Line Model (SSLM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geometry Model</td>
</tr>
<tr>
<td>Infrastructure system (roads, ...)</td>
</tr>
<tr>
<td>urban and rural</td>
</tr>
<tr>
<td>• Non-iterated tessellations (PLT,PVT,PDT)</td>
</tr>
<tr>
<td>• Nestings</td>
</tr>
<tr>
<td>• Multi-type nestings</td>
</tr>
<tr>
<td>Network Model</td>
</tr>
<tr>
<td>Placement of components</td>
</tr>
<tr>
<td>2-level models:</td>
</tr>
<tr>
<td>Higher level comp. (HLC) and lower level comp. (LLC)</td>
</tr>
<tr>
<td>• WCS (wire center stations, repartiteur) and SAI (service area interface, sous-repartiteur)</td>
</tr>
<tr>
<td>• SAI and subscriber</td>
</tr>
<tr>
<td>Topology Model</td>
</tr>
<tr>
<td>Topology of connections (trec, ...)</td>
</tr>
<tr>
<td>2-level models:</td>
</tr>
<tr>
<td>• Service zones:</td>
</tr>
<tr>
<td>WCS are nuclei of Voronoi cells</td>
</tr>
<tr>
<td>SAI is connected to WCS of its serving zone</td>
</tr>
</tbody>
</table>

Figure 1.4: The Stochastic Subscriber Line Model

Network Geometry Model is supposed to represent the cable trench system, which is located along the infrastructure system of a city or of a country. Random tessellations as well as iterated random tessellations are employed in the SSLM to capture the geometric structure of the network ([69]). Figure 1.5 shows some examples of both kinds of tessellations. The cells formed by these tessellations are the locations where subscribers are situated and the skeleton induced by the cell boundaries of the (iterated) tessellations forms the (hierarchical)
road system. The Network Component Model localizes the technical network components on the underlying geometric structure provided by the Network Geometry Model. For example it does not only provide models for the placement of network nodes along the roads where cables merge and split, but also models for the spatial placement of subscribers. Popular and tractable choices for network components in terms of stochastic geometry are (Poisson) point processes placed randomly along the skeleton induced by the underlying random tessellations, Figure 1.6 (a), whereas subscribers can be modeled by (Poisson) point processes in the plane, Figure 1.6 (b). To complete the picture, the Network Topology Model describes the nature of the link between lower level network components (for example SAI s or the subscribers themselves) and the WCS of the corresponding serving zone. Figure 1.7 shows two examples of topologies. The star topology depicted in Figure 1.7 (a), where lower level components are connected directly to the WCS through the (preferably shortest) path along the network geometry, is widely used ([33]).
To make the Stochastic Subscriber Line Model and the ideas of stochastic (cost) analysis of the network more transparent, consider the following example with two levels of network components that has been studied thoroughly in [41] and [42]. Assume that the network geometry is modelled by some Poisson line tessellation with intensity $\gamma$ as shown in Figure 1.8 (a), where $\gamma$ can be interpreted as the mean total length of induced lines per unit area. Furthermore, assume that higher level network components (like WCS) are placed randomly along the lines, Figure 1.8 (b), according to a Poisson law with intensity $\lambda_h$, which is the mean number of higher level network components per unit area. This results in a new subdivision of the plane into serving zones with respect to the higher level components, depicted in Figure 1.8 (c). Note that such Cox-Voronoi tessellations are studied in detail in [39].

Imagine that lower level network components are placed randomly either along the lines, Figure 1.9 (a), like the SAI, or are placed spatially within the cells formed by the underlying Poisson line tessellation, Figure 1.9 (b), like the subscribers.

Two characteristics of interest with regard to cost analysis within the SSLM are the mean shortest path length (MSPL) and the mean subscriber line length (MSLL). The MSPL is the
1.1. Stochastic Network Modelling

average length of the shortest path between a lower level network component or its projection on the lines and its nearest (in the Euclidean sense) higher level component, i.e., the nucleus of the serving zone in which the lower level component is situated. Likewise, the MSLL is the average length of the shortest path from a spatially placed lower level component to its nearest (again in the Euclidean sense) higher level component. In both cases, the average is taken over all locations of lower level components.

Applying concepts of stochastic geometry as well as of spatial statistics, the crucial outcome of the analysis done in [41] and [42] is summarized in Figure 1.10. It can be shown that both

![Figure 1.9: Placement of lower level network components](image)

(a) Linear placement of SAI.s (b) Spatial placement of subscribers

Figure 1.9: Placement of lower level network components

![Figure 1.10: Estimates for m(β) and m'(β); fitted curves m(β) = a β^b and m'(β) = a' β'^b](image)

(a) MSPL (a = 0.774, b = 0.450) (b) MSLL (a' = 1.124, b' = 0.425)

Figure 1.10: Estimates for m(β) and m'(β); fitted curves m(β) = a β^b and m'(β) = a' β'^b

the MSPL and the MSLL are linear functions in γ^{-1}, where the corresponding slopes m(β) and m'(β), respectively, turn out to be dependent only on the quotient β = γ/λ_h. Figure 1.10 depicts estimated values Ŧm(β) and Ŧm'(β), respectively, for these slopes as well as curves fitted to the obtained values by means of the least squares method. Hence, knowing γ as well
as $\lambda_h$, both MSPL and MSLL can be recovered easily from Figure 1.10 without any additional simulation.

1.2 Motivation

Reviewing the example of network analysis in Section 1.1.2, this analysis was based on the examination of the graph structure obtained after linear or spatial placement of network components. The results for the MSPL and the MSLL as summarized in Figure 1.10 hold for the particular network geometry of the Poisson line tessellation model. Similar results can be obtained for other underlying tessellation models leading most likely to a different behavior of the MSPL and the MSLL.

Thus, the need arises to study more precisely the possible network geometry models, i.e., the underlying tessellation models, and their structural behavior. In particular this means to analyse functionals induced by these tessellations as well as to examine their (asymptotic) distributional behavior. Typically, this study has to be based on the observation of a single realization of the tessellation in a (convex compact) sampling window. Examples of considered functionals are the number of nodes or the total length of edges, where the latter coincides with the total length of network cables. Finally, statistical inference with respect to such functionals and, more generally, with respect to the geometric models of the network is possible. For example, knowledge of the mean number of network nodes or the mean total length of cables in a certain area is of value for cost analysis of the whole network. Moreover, civil engineers are often interested in getting lower and upper bounds for such values, complementing or even replacing their own predictions based on mere rules of thumb.

Another problem in network modelling is to appropriately represent the geometric network structure. With regard to telecommunication networks this means to identify possible models in civil engineering, i.e., models for the actual placement of cables. At least in urban areas this coincides with the question how to decide between several competing (tessellation) models and to identify the model that fits certain (urban) infrastructure data best. An example of such an infrastructure system is the road system of Paris as displayed in Figure 1.1.

Both of these problems, statistical inference and model choice, initiated the research that lead to this thesis. Thus, we present some methodologies useful for statistical model fitting as well as statistical analysis of random tessellations. Our approach is two-fold.

On the one hand, we present central limit theorems (CLTs) for stationary tessellation models in the $d$-dimensional Euclidean space $\mathbb{R}^d$, showing that the asymptotic distribution of (suitably centered and normalized) functionals connected to these tessellations is a normal distribution. As it is quite common in spatial statistics, we assume that only a single realization of a certain tessellation model is available and that this realization is observable through a (convex compact) sampling window. Hence, the term asymptotic used earlier means that the distribution of the functionals is studied as the sampling window expands unboundedly but uniformly in all directions. The CLTs are a basis for (asymptotic) statistical inference of tessellation models.
and their related functionals. In particular, asymptotic confidence intervals and tests with respect to mean values, so-called intensities, of these functionals can be derived. However note that CLTs require certain conditions to hold and that they allow for statistical analysis within a specified tessellation model provided that (asymptotic) second order moments are known or can be estimated.

The presented CLTs for stationary random tessellation models complement existing literature that studies asymptotic distribution laws of a variety of stochastic–geometric models and deals with spatial statistics. To name some examples, [8] derives a limit theorem for statistics of spatial data, [21] and [72] study central limit theorems for random fields, and [89] considers limit theorems with applications to statistical physics. In [20] and the references therein, asymptotic properties of estimators for volume fractions as well as kernel estimation of spectral densities of stationary random closed sets are studied. A central limit theorem for a class of random measures associated with germ–grain models is investigated in [49] and also [93] derives central limit theorems for functionals of germ–grain models. Both [2] and [95] consider central limit theorems for two–dimensional Poisson–Voronoi from the perspective of computational geometry. Finally, [94] studies central limit theorems for Poisson line tessellations in $\mathbb{R}^2$ and points out further references.

On the other hand, we shed light on the problem of identifying appropriate tessellation models for network data by deciding between certain available and suitable models. In particular, we present a statistical model fitting procedure based on simulation techniques. Such methodologies are often applied if derivation and statement of explicit mathematical formulae is not possible and are, compared to the CLTs for example, less restrictive in the conditions that have to be met before application is admissible. Note that such fitting procedures and subsequent tests are discussed in [116] and [117] for example, where also further references can be found.

1.3 Organisation of this Thesis

We will start by introducing concepts of stochastic geometry in $\mathbb{R}^d$ in Chapters 2 and 3. The idea is to provide the reader with some foundations of this specific mathematical discipline in order to be able to understand the results of this thesis later on. After the introduction of basic concepts of random closed sets and random point processes (both unmarked and marked) we will briefly review random point processes of closed sets which will then allow for the definition of flat processes and random tessellations. Due to the particular importance of the latter models throughout this thesis, random tessellations will be discussed separately in Chapter 3. In particular, it will turn out that for the derivation of the results of this thesis it is useful to represent stationary random tessellations, being defined as special kind of point process of closed sets, as stationary marked point processes in $\mathbb{R}^d$. At the end of Chapter 3, we will prove the first result, a general theorem about the asymptotic negligibility of cells hitting the boundary of a sampling window $W^d \in \mathbb{R}^d$ that expands unboundedly but uniformly in all directions as the scaling factor $\varrho \uparrow \infty$.

Chapter 4 will be dedicated to CLTs for $k$–flat processes ($0 \leq k \leq d − 1$) induced by intersections of $d − k$ Poisson hyperplanes of some stationary, however not necessarily isotropic,
$d$-dimensional Poisson hyperplane process. More precisely, we will assume that we can observe a realization of a Poisson hyperplane process in $\mathbb{R}^d$ through a $d$-dimensional spherical sampling window $B_r^d$ centered at the origin with radius $r > 0$. We will show that the (suitably centered and normalized) functionals $\eta_k(B_r^d)$ and $\zeta_k(B_r^d)$, denoting the number and the total $k$-volume, respectively, of the induced $k$-flats in $B_r^d$ for any $k = 0, \ldots, d - 1$, converge to a Gaussian mean zero vector with some asymptotic variance as $r \to \infty$. It will turn out that if the Poisson hyperplane process inducing $\eta_k(B_r^d)$ and $\zeta_k(B_r^d)$ is isotropic, this asymptotic variance can be given explicitly. By our result we generalize a result obtained earlier by K. Pardoix ([94]) in the two-dimensional case where the number $\eta_0(B_r^2)$ of intersections of Poisson lines in $B_r^2$ is additionally equipped with a mark imposing restrictions on the directions of the intersecting lines. Representing $\eta_k(B_r^d)$ and $\zeta_k(B_r^d)$ as $U$-statistics and applying Hoeffding’s decomposition, a technique which seems to be more efficient in the $d$-dimensional case compared to the method of moments used in [94], we are in the position to provide a relatively simple and transparent proof of our CLTs. Furthermore, our proving technique also allows to treat the multivariate case, i.e., to show that the vectors having as entries the (centered and normalized) functionals $\eta_0(B_r^d), \ldots, \eta_{d-1}(B_r^d)$ and $\zeta_0(B_r^d), \ldots, \zeta_{d-1}(B_r^d)$, respectively, are normally distributed as $r \to \infty$.

In Chapter 5 we will apply the CLTs of Chapter 4 in order to allow for statistical inference by deriving asymptotic confidence intervals and asymptotic tests for the intensities of the induced $k$-flat processes. In a practical part, goodness-of-fit test techniques will be applied to check for normality.

Chapter 6 will consider stationary (and ergodic) random tessellations $\Psi = \{\Xi_n\}_{n \geq 1}$ in $\mathbb{R}^d$. The CLTs that will be derived in this chapter for functionals $J_n$ induced by some random structure acting in the interior of the cells $\Xi_n$ of $\Psi$ are much more general compared to the results in Chapter 4 and Chapter 5, respectively. Examples of such random structures include point configurations, fibre systems, or again tessellations nested in the cells $\Xi_n$ of $\Psi$. We will assume that the random inner cell structures are generated both independently of each other and independently of $\Psi$ by generic stationary random sets related to a stationary random vector measure $J_0$ on $\mathbb{R}^d$. In particular, we will assume that we can observe a single realization of $\Psi$ along with the structure embedded in the cells $\Xi_n$ of $\Psi$ within a convex compact $d$-dimensional sampling window $W_{\rho}^d \subset \mathbb{R}^d$. Then we will study the asymptotic behavior of a multivariate random functional as $W_{\rho}^d$ expands unboundedly but uniformly in all directions as $\rho \uparrow \infty$, where this functional is determined both by $\Psi$ as well as by the cell structures contained in $W_{\rho}^d$. It will turn out that the considered vector functional provides an unbiased estimator for the intensity vector associated with $J_0$. Moreover, under quite natural restrictions, we will provide strong laws of large numbers as well as a multivariate CLT for the normalized version of the vector functional. Finally, some numerical applications will be discussed.

Chapter 7 finally has to be considered in the framework of statistical inference between tessellation models. We will present a statistical fitting approach for random tessellation models to input data exhibiting network structure, where the model fitting procedure is based on simulation techniques. In particular, we will make use of mean value relationships explicitly known for specific functionals of certain tessellation models in order to minimize distances between
theoretically computed values of these functionals and estimated characteristics of input data. Our results can be checked using tests based on Monte Carlo techniques.

1.4 The GeoStoch Library

The GeoStoch library is a joint research project of the Department of Stochastics and the Department of Applied Information Processing at the University of Ulm ([75] and [76]). This library, based on the JAVA™ programming language, comprises methods from stochastic geometry, spatial statistics, and image analysis.

The webpage www.geostoch.de lists several projects for which the GeoStoch library is currently used. The idea behind the library is to provide a core of general methods that can be used for different applications without having to program everything anew each time. Furthermore, the library is consistently extended by methods developed in the course of ongoing projects, like the joint research project with France Télécom R&D Division.
Chapter 2

Concepts of Stochastic Geometry

Since its early beginnings in the 18th century, we recall as an example Buffon’s needle problem, stochastic geometry has developed into a rather widespread mathematical discipline today with applications to a variety of fields, ranging from biology, medicine, and forestry to material sciences and telecommunication for example. Despite the work and contributions of many mathematicians, it was however not until Matheron’s book ([73]) in 1975 that stochastic geometry became known to a broader mathematical audience as a discipline in its own right.

We refer to [112] for a more detailed introduction to the development of stochastic geometry as well as an attempt to position it amongst closely related fields like integral and convex geometry, spatial statistics, image analysis, and stereology.

In the following, and also in Chapter 3, we present some concepts of stochastic geometry necessary to understand the results presented in the subsequent chapters. Far from being comprehensive, our introduction follows a rather classical path. In Section 2.1, we introduce random closed sets in the $d$-dimensional Euclidean space $\mathbb{R}^d$, before we study random point processes as well as their generalization, random marked point processes, in Section 2.2 and Section 2.3, respectively. Finally, in Section 2.4 we consider point processes of closed sets in order to introduce flat processes in Section 2.5 as a particular example thereof and also to lay the ground for the description of random tessellations in Chapter 3.

A short review of basic notions of the underlying $d$-dimensional Euclidean space $\mathbb{R}^d$ can be found in Section A.1 of the appendix. A general introduction, in particular also to measure theory, is given in [11], [16], and [32] for example.

For a detailed discussion of the topics from stochastic geometry presented in this chapter the reader is referred to the existing literature. In particular, random closed sets are studied in [85], whereas point processes are covered in [31, 64, 66, 74, 123]. A profound discussion of stochastic geometry and further models, also in a more general setting, can be found in [1, 14, 46, 61, 73, 79, 103, 112, 116, 117]. All these monographs contain also a lot of other references as well as numerous applications of stochastic-geometric concepts to various fields.
2.1 Random Closed Sets

2.1.1 The Space of Closed Sets

Let $\mathcal{F}$, $\mathcal{K}$, and $\mathcal{C}$ denote the families of closed sets, compact sets, and convex bodies (convex and compact sets), respectively, in the $d$-dimensional Euclidean space $\mathbb{R}^d$. In particular, put $\mathcal{F}' = \mathcal{F} \setminus \{\emptyset\}$, $\mathcal{K}' = \mathcal{K} \setminus \{\emptyset\}$, and $\mathcal{C}' = \mathcal{C} \setminus \{\emptyset\}$, respectively, where $\emptyset$ is the empty set in $\mathbb{R}^d$.

Consider the sets

$$\mathcal{F}_A = \{F \in \mathcal{F} : F \cap A \neq \emptyset\} \quad \text{and} \quad \mathcal{F}^A = \{F \in \mathcal{F} : F \cap A = \emptyset\}, \quad A \subset \mathbb{R}^d,$$

and assume that $\mathcal{F}$ is endowed with the topology that is generated by

$$\{\mathcal{F}_K : K \in \mathcal{K}\} \cup \{\mathcal{F}_G : G \in \mathcal{G}\},$$

where $\mathcal{G}$ is the family of open sets in $\mathbb{R}^d$. Furthermore, we introduce the Borel $\sigma$-algebra $\mathcal{B}(\mathcal{F})$ on $\mathcal{F}$, where $\mathcal{B}(\mathcal{F})$ is generated by any of the families

$$\{\mathcal{F}_K : K \in \mathcal{K}\}, \quad \{\mathcal{F}_K : K \in \mathcal{K}\}, \quad \{\mathcal{F}_G : G \in \mathcal{G}\}, \quad \text{or} \quad \{\mathcal{F}_G : G \in \mathcal{G}\}.$$

Note that it can be shown (Satz 1.3.2 and Satz 1.3.3 in [112]) that both $\mathcal{C}$ and $\mathcal{K}$ are Borel sets in $\mathcal{F}$.

2.1.2 Definition and Properties

A random closed set (RACS) $\Xi$ in $\mathbb{R}^d$ is a measurable mapping $\Xi : \Omega \rightarrow \mathcal{F}$ from some probability space $(\Omega, \sigma(\Omega), \mathbb{P})$ into $(\mathcal{F}, \mathcal{B}(\mathcal{F}))$ with probability distribution $P_\Xi : \mathcal{B}(\mathcal{F}) \rightarrow [0, 1]$ given by

$$P_\Xi(B) = \mathbb{P}(\Xi \in B) = \mathbb{P}(\{\omega \in \Omega : \Xi(\omega) \in B\}), \quad B \in \mathcal{B}(\mathcal{F}).$$

In particular, $\Xi$ is called random compact set and random convex body in $\mathbb{R}^d$ if $P_\Xi(\mathcal{K}) = 1$ and $P_\Xi(\mathcal{C}) = 1$, respectively.

The distribution of a RACS $\Xi$ is uniquely determined by Choquet’s capacity functional $T_\Xi : \mathcal{K} \rightarrow [0, 1]$, which in turn is defined as

$$T_\Xi(K) = P_\Xi(\mathcal{F}_K) = \mathbb{P}(\Xi \cap K \neq \emptyset), \quad K \in \mathcal{K},$$

and can be interpreted as analogon to the distribution function of a real-valued random variable.

**Theorem 2.1** (Choquet’s theorem) Let $T : \mathcal{K} \rightarrow \mathbb{R}$ be a functional that fulfills

(i) $T(\emptyset) = 0$ and $0 \leq T(K) \leq 1$ for $K \in \mathcal{K}$.
(ii) $K_n \downarrow K$ implies $T(K_n) \downarrow T(K)$ for $K, K_1, K_2, \ldots \in \mathcal{K}$, where $K_n \downarrow K$ means that $K_1 \supset K_2 \supset \cdots$ and $K = \bigcap_{n \geq 1} K_n$.

(iii) For all $K, K_1, K_2, \ldots, K_n \in \mathcal{K}$ and $n \in \mathbb{N}$,

$$\sum_{j=0}^{n} (-1)^{j-1} \sum_{1 \leq i_1 < \cdots < i_j \leq n} T(K \cup K_{i_1} \cup \cdots \cup K_{i_j}) \geq 0, \quad (2.1)$$

where the inner sum in (2.1) is defined as $T(K)$ for $j = 0$ and where $\mathbb{N} = \{1, 2, \ldots\}$ denotes the set of positive integers.

Then there exists a uniquely determined probability measure $P$ on $\mathcal{F}$ such that

$$T(K) = P(\mathcal{F}_K), \quad K \in \mathcal{K}.$$ 

The proof of Theorem 2.1 can be found for example in Section 2–2 of [73]; cf. also [27]. Satz 1.4.1 on p. 27 in [112] shows that conditions (i) to (iii) are fulfilled in particular for $T = T_\Xi$. Hence, Choquet’s theorem shows the existence of the (uniquely defined) probability measure $P_\Xi$ on $\mathcal{B}(\mathcal{F})$. Furthermore, the sum (2.1) is then the probability that the RACS $\Xi$ hits $K_1, \ldots, K_n$ but does not hit $K$; cf. Section 2.2 of [73] or Section 1.4 and Section 2.2 of [112].

A RACS $\Xi$ in $\mathbb{R}^d$ is called stationary if

$$\Xi \overset{d}{=} \Xi + x, \quad \text{or, equivalently,} \quad T_\Xi(K) = T_{\Xi+x}(K), \quad K \in \mathcal{K}$$

for any $x \in \mathbb{R}^d$, where $\overset{d}{=} \Xi$ denotes equality in distribution, and where we refer to Section A.1.2 for the definition of the translation of a subset of $\mathbb{R}^d$. Also, $\Xi$ is called isotropic if

$$\Xi \overset{d}{=} \theta_R \Xi, \quad \text{or, equivalently,} \quad T_\Xi(K) = T_{\theta_R \Xi}(K), \quad K \in \mathcal{K},$$

where we refer once again to Section A.1.2 for the definition of the rotation (around the origin) of a subset of $\mathbb{R}^d$.

In view of the set operations defined in Section A.1.2 for subsets of the $\mathbb{R}^d$, the following theorem allows for the construction of several new RACSs out of given RACSs. The proof can be found for example on p. 23 in [112].

**Theorem 2.2** Let $\Xi$ and $\Xi'$ be two RACSs in $\mathbb{R}^d$. Then, $\Xi \cap \Xi'$, $\Xi \cup \Xi'$, $\rho \Xi$ for any $\rho \geq 0$, $\Xi + x$ for any $x \in \mathbb{R}^d$, and $\theta_R \Xi$ for any rotation $\theta_R$ around the origin are RACSs in $\mathbb{R}^d$. Furthermore, the closure $\text{cl}(\Xi \oplus \Xi')$ of the Minkowski addition $\Xi \oplus \Xi'$ of $\Xi$ and $\Xi'$, the closure $\text{cl} \Xi^c$ of the complement $\Xi^c$ of $\Xi$, and the boundary $\partial \Xi$ of $\Xi$ are RACSs in $\mathbb{R}^d$.

### 2.2 Random Point Processes

In this section we introduce the fundamental concept of random point processes in $\mathbb{R}^d$. Figure 2.1 displays some examples for $d = 2$. 
2.2.1 Definition and Representation

A (simple and locally finite) random point process on \( \mathbb{R}^d \) is a measurable mapping \( X : \Omega \to M(\mathbb{R}^d) \) from an arbitrary probability space \((\Omega, \sigma(\Omega), \mathbb{P})\) into the measurable space \((M(\mathbb{R}^d), M(\mathbb{R}^d))\), where \( M(\mathbb{R}^d) \) denotes the set of all simple and locally finite counting measures \( \tilde{\mu}_c : B(\mathbb{R}^d) \to \mathbb{N}_0 \cup \{\infty\} \); cf. also Formula (A.2) in Section A.1.3. Note that \( B(\mathbb{R}^d) \) denotes the Borel \( \sigma \)-algebra on \( \mathbb{R}^d \) and that \( \mathbb{N}_0 = \mathbb{N} \cup \{0\} \) is the set of non-negative integers. Furthermore, \( M(\mathbb{R}^d) \) is the \( \sigma \)-algebra of subsets of \( M(\mathbb{R}^d) \), generated by the sets \( \{\tilde{\mu}_c \in M(\mathbb{R}^d) : \tilde{\mu}_c(B) = k\} \) for \( k \in \mathbb{N}_0 \) and \( B \in B_0(\mathbb{R}^d) \), where \( B_0(\mathbb{R}^d) \) denotes the family of bounded Borel sets on \( \mathbb{R}^d \).

Note that a random point process \( X \) on \( \mathbb{R}^d \) can be represented as a sequence \( \{X_n\}_{n \geq 1} \) of random vectors \( X_n : \Omega \to \mathbb{R}^d \), expressed by the notation \( X = \{X_n\}_{n \geq 1} \); Section 11.3 of [64]. In particular, the canonical representation of \( X = \{X_n\}_{n \geq 1} \) is obtained if \((\Omega, \sigma(\Omega))\) is identified with \((M(\mathbb{R}^d), M(\mathbb{R}^d))\) and the mapping \( X(\tilde{\mu}_c) = \tilde{\mu}_c \) is considered. Hence, \( X \) may be seen as random element of the measurable space \((M(\mathbb{R}^d), M(\mathbb{R}^d))\) and can be written as

\[
X(B) = \sum_{x \in \text{supp } X} \delta_x(B) = \sum_{n \geq 1} \delta_{X_n}(B), \quad B \in B(\mathbb{R}^d),
\]

where \( \delta_x \) denotes the Dirac measure, cf. also Formula (A.1) in Section A.1.3, and \( \text{supp } X = \{x \in \mathbb{R}^d : X(\{x\}) > 0\} \) is the support of \( X \). Therefore, the random variable \( X(B) : \Omega \to \mathbb{N}_0 \cup \{\infty\} \) is the number of points of \( X \) which are located in \( B \in B(\mathbb{R}^d) \).

Finally, \( X \) can also be identified with its support \( \text{supp } X \) and hence be seen as a special kind of RACS, Section 11.3 of [64], since \( X \) is assumed to be locally finite and simple.
2.2. Random Point Processes

2.2.2 Distribution and Basic Properties

The distribution $P_X : \mathcal{M} (\mathbb{R}^d) \to [0, 1]$ of $X$ is given by

$$P_X (A) = \mathbb{P} (X \in A) = \mathbb{P} (\{ \omega \in \Omega : X (\omega) \in A \}) , \quad A \in \mathcal{M} (\mathbb{R}^d).$$

The following theorem provides a measurability statement as well as the uniqueness of the point process $X$. Its proof can be found for example on p. 65 in [112].

**Theorem 2.3** The mapping $X : (\Omega, \sigma (\Omega), \mathbb{P}) \to (\mathcal{M} (\mathbb{R}^d), \mathcal{M}(\mathbb{R}^d))$ is a point process in $\mathbb{R}^d$ if and only if $\{ X (K) = 0 \}$ is measurable for all $K \in \mathcal{K}$. Moreover, let $X$ and $X'$ be two (simple) point processes in $\mathbb{R}^d$, where

$$\mathbb{P} (X (K) = 0) = \mathbb{P} (X' (K) = 0) , \quad \text{for all } K \in \mathcal{K}.$$

Then $X \overset{d}{=} X'$, i.e., the two point processes $X$ and $X'$ are equal in distribution.

A random point process $X$ is called **stationary** if

$$X \overset{d}{=} X + x \quad \text{for all } x \in \mathbb{R}^d ,$$

where $(X + x)(B) = X (B - x)$ for any $B \in \mathcal{B} (\mathbb{R}^d)$, and $X$ is called **isotropic** if

$$X \overset{d}{=} \theta_R X \quad \text{for any rotation } \theta_R \text{ around the origin} ,$$

where $(\theta_R X)(B) = X (\theta_{R^{-1}} B)$ for any $B \in \mathcal{B} (\mathbb{R}^d)$.

2.2.3 Intensity Measure

The **intensity measure** $\Lambda_X : \mathcal{B} (\mathbb{R}^d) \to [0, \infty]$ of $X$ is defined by

$$\Lambda_X (B) = \mathbb{E} (X (B)) , \quad B \in \mathcal{B} (\mathbb{R}^d) ,$$

and, analogous to real-valued random variables, is the expected number of points in $B$. In case of a stationary point process $X$ we have that

$$\Lambda_X (B) = \mathbb{E} (X (B)) = \mathbb{E} ((X + x)(B)) = \mathbb{E} (X (B - x)) = \Lambda_X (B - x) , \quad x \in \mathbb{R}^d , B \in \mathcal{B} (\mathbb{R}^d) ,$$

and due to this translation invariance of $\Lambda_X (\cdot)$, Haar’s lemma (cf. Lemma A.1) implies that

$$\Lambda_X (B) = \lambda_X \nu_d (B) , \quad B \in \mathcal{B} (\mathbb{R}^d) ,$$

for some constant $\lambda_X \in [0, \infty]$. This constant is called the **intensity** of $X$ and can be interpreted as the expected number of points of $X$ per unit volume, i.e., $\lambda_X = \mathbb{E} X ([0, 1]^d)$. To this end we assume that $0 < \lambda_X < \infty$.

The following theorem is fundamental when working with point processes and is, basically, an application of Fubini’s Theorem. The proof can be found for instance in Chapter 3 of [112] or in Section 4.1 of [116].
Theorem 2.4 (Campbell’s theorem) Let $X$ be a random point process in $\mathbb{R}^d$ and let $f : \mathbb{R}^d \to [0, \infty)$ be a measurable function. Then $\sum_{x \in \text{supp} X} f(x) : \Omega \to [0, \infty]$ is measurable and

$$E \sum_{x \in \text{supp} X} f(x) = \int_{\mathbb{R}^d} f(x)X(dx) = \int_{\mathbb{R}^d} f(x)\Lambda_X(dx).$$

If $X$ is additionally stationary, i.e., $\Lambda_X(\cdot) = \lambda_X \nu_d(\cdot)$, then

$$E \sum_{x \in \text{supp} X} f(x) = \lambda_X \int_{\mathbb{R}^d} f(x)dx.$$

2.2.4 Palm Probabilities

When regarding a random point process $X$, it is often interesting to examine this point process and its distribution under the condition that a point of $X$ is situated at a fixed location $x \in \mathbb{R}^d$, which can be interpreted as regarding the point process from that location $x$. This (conditional) distribution of $X$ can then be used to consider the notion of Palm probabilities, cf. for example [64].

Let $X$ be a stationary random point process in $\mathbb{R}^d$ with intensity $\lambda_X \in (0, \infty)$ and let $B \in \mathcal{B}(\mathbb{R}^d)$ with $0 < \nu_d(B) < \infty$. Then, $P^o_X : \mathcal{M}(\mathbb{R}^d) \to [0, 1]$ defined as

$$P^o_X(A) = \frac{1}{\lambda_X \nu_d(B)} E \sum_{x \in \text{supp} X} \mathbb{1}_B(x) \mathbb{1}_A(X-x), \quad A \in \mathcal{M}(\mathbb{R}^d)$$

is called the Palm distribution of $X$ with respect to the origin $o$, where $\mathbb{1}_A(\cdot)$ and $\mathbb{1}_B(\cdot)$ denote the indicator function with respect to the sets $A$ and $B$, respectively. It can be shown that $P^o_X$ fulfills the properties of a distribution and, due to stationarity, does not depend on the particular choice of $B \in \mathcal{B}(\mathbb{R}^d)$; Section 12.2 of [64]. Note that $P^o_X$ can be interpreted as the conditional distribution of $X$ that there is a point of $X$ at the origin $o$.

The Palm distribution $P^o_X$ of a stationary random point process $X$ in $\mathbb{R}^d$ can be used for example to introduce the second reduced moment measure $\mathfrak{r}$ of $X$ by

$$\lambda_X \mathfrak{r}(B) = \int_{\mathcal{M}(\mathbb{R}^d)} \mu(B \setminus \{o\}) P^o_X(d\mu), \quad B \in \mathcal{B}(\mathbb{R}^d),$$

where $\lambda_X \in (0, \infty)$ is the intensity of $X$. Hence, $\lambda_X \mathfrak{r}(B)$ can be interpreted as the expected number of points in $B \setminus \{o\}$ under the condition that a point of $X$ is located at the origin $o$, where the latter point of $X$ at $o$ is not counted itself. Note that $\mathfrak{r}(B) = \mathfrak{r}(\bar{B})$ for any $B \in \mathcal{B}(\mathbb{R}^d)$, where $\bar{B} = \{-x : x \in B\}$ is the reflection of $B$, and that $K(r) = \mathfrak{r}(B^d_r)$ for $r \geq 0$ is called Ripley’s K-function; cf. Section 4.5 in [116].

Assume now that $X$ is additionally isotropic. The pair correlation function $g : [0, \infty] \to [0, \infty]$ of $X$ is given by

$$g(r) = \frac{1}{d \kappa_d r^{d-1}} K'(r), \quad r \geq 0,$$  

(2.3)
2.2. Random Point Processes

Figure 2.2: Realization of a clustered point process and the theoretical pair correlation function

where $K'$ denotes the first derivative of the $K$–function, provided that it exists.

The pair correlation function is an important tool in spatial statistics of point processes since it
describes their second-order behavior. Generally, a large value of $g(r)$ at $r_0$, say, means
that pairs of points having distance $r_0$ occur frequently, whereas a small value of $g(r)$ at $r_0$
means that pairs of points with distance $r_0$ show up only rarely. In case $g(r_0) = 0$, the
interpoint distance $r_0$ between pairs of points appears only with probability zero. Figure 2.2
(a) shows the realization of some clustered point process in $d = 2$ and Figure 2.2 (b) depicts
the theoretical pair correlation function for this type of point process. Note that Figure 4.10 on
p. 130 in [116] provides further pair correlation functions for different types of point processes.
Moreover, in Section 4.4.2 of [117] an extensive discussion of the pair correlation function and
its interpretation along with further figures can be found.

2.2.5 Poisson Point Processes

As an example of random point processes, we consider the Poisson point process in $\mathbb{R}^d$. This
type of random process often serves as a reference point process model, but also as a basis for
the construction of more sophisticated stochastic geometric models.

A random point process $X$ with some diffuse and locally finite intensity measure $\Lambda_X : \mathcal{B}(\mathbb{R}^d) \to
[0, \infty]$ is called a (non–stationary) Poisson point process in $\mathbb{R}^d$ if

(i) $X(B_1), \ldots, X(B_n)$ are independent for $B_1, \ldots, B_n \in \mathcal{B}_0(\mathbb{R}^d)$ with $B_i \cap B_j \neq \emptyset$ for $i \neq j$,
and if

(ii) it holds that

$$
P(X(B) = k) = \frac{(\Lambda_X(B))^k}{k!} \exp\{-\Lambda_X(B)\}, \quad k \in \mathbb{N}_0, \ B \in \mathcal{B}_0(\mathbb{R}^d).$$
In case of stationarity, the intensity measure \( \Lambda_X \) can be decomposed, i.e., \( \Lambda_X(B) = \lambda_X \mu_d(B) \) for any \( B \in \mathcal{B}(\mathbb{R}^d) \). Also, we have that \( X \) is automatically isotropic; Satz 3.3.1 on p. 80 in [112]. The \( K \)-function of a stationary Poisson point process \( X \) is given by \( K(r) = \kappa_d r^d \), which in turn yields that the pair correlation function \( g(r) = 1 \) for all \( r \geq 0 \), cf. Figure 2.2 (b). A standard algorithm for simulating random Poisson point processes is described in [98]. Figure 2.3 shows two realizations of a stationary Poisson point process with different intensity values \( \lambda_X \).

![Realizations of stationary random Poisson point processes](image)

Figure 2.3: Realizations of stationary random Poisson point processes

Finally, we mention the following theorem, where we identify a (simple) stationary Poisson point process \( X \) with its support \( \text{supp} X \). Hence, also the Palm distribution \( P^o_X \) of \( X \), introduced in 2.2.4, is interpreted as a measure on \( \mathcal{B}(\mathcal{F}) \); cf. p. 87 in [112]. The proof of the subsequent theorem can be found on p. 87 in [112].

**Theorem 2.5** (Slivnyak’s theorem) Let \( X \) be a stationary Poisson point process in \( \mathbb{R}^d \). Then,

\[
P^o_X(A) = \mathbb{P}(X \cup \{o\} \in A),
\]

for all \( A \in \mathcal{B}(\mathcal{F}) \).

### 2.3 Random Marked Point Processes

Random marked point processes in \( \mathbb{R}^d \) can be seen as a generalization of random point processes introduced in Section 2.2. Each point of the process is additionally equipped with a mark taken from some mark space \( \mathcal{D} \).

#### 2.3.1 Definition and Basic Properties

The mapping \( X_D : \Omega \to M(\mathbb{R}^d \times \mathcal{D}) \) from some probability space \( (\Omega, \sigma(\Omega), \mathbb{P}) \) into the measurable space \( (M(\mathbb{R}^d \times \mathcal{D}), \mathcal{M}(\mathbb{R}^d \times \mathcal{D})) \) is called a (simple and locally finite) random
2.3. Random Marked Point Processes

A marked point process in $\mathbb{R}^d$ with mark space $\mathcal{D}$. This mark space is assumed to be a Polish space, i.e., a complete and separable metric space, equipped with the Borel $\sigma$-algebra $\mathcal{B}(\mathcal{D})$. Furthermore, in the above definition of a random marked point process, $M(\mathbb{R}^d \times \mathcal{D})$ denotes the set of all counting measures $\mu'_c : \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathcal{D}) \to \mathbb{N}_0 \cup \{\infty\}$ which are simple and locally finite in the first component, and $\mathcal{M}(\mathbb{R}^d \times \mathcal{D})$ is the $\sigma$-algebra of subsets of $M(\mathbb{R}^d \times \mathcal{D})$ generated by the sets $\{\mu'_c \in M(\mathbb{R}^d \times \mathcal{D}) : \mu'_c(B \times G) = k\}$ for $B \in \mathcal{B}_0(\mathbb{R}^d)$, $G \in \mathcal{B}(\mathcal{D})$, and $k \in \mathbb{N}_0$. In analogy to an unmarked point process, the random marked point process $X_D$ may be represented as a sequence of random marked points $\{[X_n, D_n]\}_{n \geq 1}$, where both $X_n : \Omega \to \mathbb{R}^d$ and $D_n : \Omega \to \mathcal{D}$ are measurable mappings for every $n \geq 1$. Note that $X_D$ is called independently marked if the sequences $\{X_n\}_{n \geq 1}$ and $\{D_n\}_{n \geq 1}$ are independent and if $\{D_n\}_{n \geq 1}$ consists of independent and identically distributed (iid) random variables.

Thus, $X_D = \{[X_n, D_n]\}_{n \geq 1}$ can be written as

$$X_D(B \times G) = \sum_{[x,n] \in \text{supp} X_D} \delta_{[x,n]}(B \times G) = \sum_{n \geq 1} \delta_{[X_n, D_n]}(B \times G), \quad B \in \mathcal{B}(\mathbb{R}^d), G \in \mathcal{B}(\mathcal{D}),$$

where $X_D(B \times G)$ is the random number of points located in $B$ with mark in $G$.

The distribution $P_{X_D} : \mathcal{M}(\mathbb{R}^d \times \mathcal{D}) \to [0,1]$ of $X_D$ is defined by

$$P_{X_D}(A) = \mathbb{P}(X_D \in A) = \mathbb{P}\{\omega \in \Omega : X_D(\omega) \in A\}, \quad A \in \mathcal{M}(\mathbb{R}^d \times \mathcal{D}).$$

Stationarity and isotropy are introduced with respect to the first component of the points of $X_D$, i.e., $X_D$ is called stationary if $X_D \overset{d}{=} \{[X_n + x, D_n]\}$ for all $x \in \mathbb{R}^d$ and isotropic if $X_D \overset{d}{=} \{|R \times X_n, D_n]\}$ for any rotation $R \times X_n$ of $X_n$ around the origin, where $R$ is an orthogonal matrix with $det R = 1$.

2.3.2 Intensity Measure and Palm Probabilities

The intensity measure $\Lambda_D : \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathcal{D}) \to [0, \infty]$ of $X_D$ is defined by

$$\Lambda_D(B \times G) = \mathbb{E}X_D(B \times G), \quad B \in \mathcal{B}(\mathbb{R}^d), G \in \mathcal{B}(\mathcal{D}),$$

i.e., $\Lambda_D(B \times G)$ is the expected number of points in $B$ with mark in $G$. In the stationary case, analogous to Section 2.2.3, the intensity measure $\Lambda_D$ admits the decomposition (p. 89 in [112])

$$\Lambda_D(B \times G) = \lambda_D \int_{\mathbb{R}^d} \int_{\mathcal{D}} \mathbb{1}_B(x) \mathbb{1}_G(m) P_D(dm)dx, \quad B \in \mathcal{B}(\mathbb{R}^d), G \in \mathcal{B}(\mathcal{D}),$$

where the constant $\lambda_D \in [0, \infty]$ is the intensity of $X_D$ and where to this end we assume that $0 < \lambda_D < \infty$. The distribution $P_D : \mathcal{B}(\mathcal{D}) \to [0,1]$ is called the Palm mark distribution of $X_D$, given by

$$P_D(G) = \frac{1}{\lambda_D \nu_d(B)} \mathbb{E} \sum_{[x,n] \in \text{supp} X_D} \mathbb{1}_B(x) \mathbb{1}_G(m), \quad G \in \mathcal{B}(\mathcal{D}), \quad (2.4)$$
for any $B \in \mathcal{B}(\mathbb{R}^d)$ with $0 < \nu_d(B) < \infty$.

This allows to formulate a version of Campbell’s theorem for marked point processes, where the proof is in analogy to the proof of Theorem 2.4.

**Theorem 2.6** (Campbell’s theorem) Let $X_D$ be a random marked point process in $\mathbb{R}^d$ and let $f : \mathbb{R}^d \times D \to [0, \infty)$ be a measurable function. Then, $\sum_{[x,m] \in \text{supp } X_D} f(x, m) : \Omega \to [0, \infty]$ is measurable and

$$
\mathbb{E} \sum_{[x,m] \in \text{supp } X_D} f(x, m) = \mathbb{E} \int_{\mathbb{R}^d \times D} f(x, m) X_D(d(x, m)) = \int_{\mathbb{R}^d \times D} f(x, m) \Lambda_D(d(x, m)).
$$

If $X_D$ is additionally stationary,

$$
\mathbb{E} \sum_{[x,m] \in \text{supp } X_D} f(x, m) = \lambda_D \int_{\mathbb{R}^d} \int_D f(x, m) P_D(dm)dx,
$$

where $P_D$ denotes the Palm mark distribution as introduced in (2.4).

Note that, alternatively, random marked point processes can be introduced as random counting measures on the product $\sigma$-algebra $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(D)$, i.e., as unmarked point processes in $\mathbb{R}^d \times D$; Chapter 13 of [64] or Section 4.2 of [116].

### 2.3.3 Ergodicity and Mixing

We consider an important subclass of random (marked) point processes, namely *mixing* and *ergodic* (marked) point processes in $\mathbb{R}^d$. Vaguely spoken and seen from the applicational point of view, ergodicity means that all the information about the point process is inherent in one sample. Hence, ergodicity allows for example to replace averages over different realizations of the underlying random process by spatial averages with respect to a single realization of the process.

In order to derive theoretical results, one assumes that a realization of some random process, for example a random marked point process, is observable through a sequence of expanding sampling windows. Often, one considers a *convex averaging sequence* $\{W_n^d\}_{n \geq 1}$ of bounded Borel sets $W_n^d \in \mathcal{B}_0(\mathbb{R}^d)$, cf. p. 332 in [31], with the property that

(i) $W_n^d$ is convex for all $n \in \mathbb{N}$,

(ii) $W_n^d \subseteq W_{n+1}^d$ for $n \in \mathbb{N}$, and

(iii) $r(W_n^d) \to \infty$ as $n \to \infty$, where $r(B) = \sup \{r : B_r(x) \subseteq B, x \in \mathbb{R}^d \}$ for $B \in \mathcal{B}_0(\mathbb{R}^d)$. 

The following definitions are given for the case of random marked point processes, the respective specialisation for unmarked point processes is obvious.

Following Definition 10.3.I on p. 341 in [31], a stationary random (marked) point process \( X_D = \{ [X_n, D_n] \}_{n \geq 1} \) with distribution \( P_{X_D} \) is called

(i) **mixing** if for all \( A, A' \in \mathcal{M}(\mathbb{R}^d \times \mathcal{D}) \),

\[
P_{X_D}((A + x) \cap A') - P_{X_D}(A) P_{X_D}(A') \xrightarrow{\|x\| \to \infty} 0.
\]

(ii) **ergodic** if for all \( A, A' \in \mathcal{M}(\mathbb{R}^d \times \mathcal{D}) \),

\[
\frac{1}{\nu_d(W^n_d)} \int_{W^n_d} (P_{X_D}((A + x) \cap A') - P_{X_D}(A) P_{X_D}(A')) \, dx \xrightarrow{n \to \infty} 0.
\]

Note that any mixing (marked) point process in \( \mathbb{R}^d \) is also ergodic. An example of an (unmarked) mixing and ergodic random point process is given by a stationary Poisson point process as considered in Section 2.2.5. Note aside that by Propositions 10.3.II and 10.3.V on p. 342 and p. 343, respectively, in [31], ergodicity and mixing are characterized in terms of the trivality of the invariant \( \sigma \)-algebra and the tail \( \sigma \)-algebra, respectively.

The following lemma states a version of the **individual ergodic theorem** applied to a stationary and ergodic marked point process \( X_D \), where the proof can be found on p. 339 (Corollary 10.2.VII) in [31].

**Lemma 2.1** Let \( X_D = \{ [X_n, D_n] \}_{n \geq 1} \) be a stationary and ergodic marked point process in \( \mathbb{R}^d \) with finite intensity \( \lambda_D > 0 \) and with Palm mark distribution \( P_D \) introduced in (2.4). Consider the real-valued function \( h \in L^1(\mathcal{D}, \mathcal{B}(\mathcal{D}), P_D) \), where \( L^1(\mathcal{D}, \mathcal{B}(\mathcal{D}), P_D) \) denotes the family of all integrable functions defined on \( (\mathcal{D}, \mathcal{B}(\mathcal{D}), P_D) \). Then

\[
\frac{1}{\nu_d(W^n_d)} \sum_{\ell \geq 1} I_{W^n_d}(X_{\ell}) h(D_{\ell}) \xrightarrow{n \to \infty} \lambda_D \int_{\mathcal{D}} h(m) P_D(dm),
\]

where \( \{ W^n_d \}_{n \geq 1} \subset \mathbb{R}^d \) is an averaging sequence of convex bounded Borel sets.

Note that for example in [128] ergodic theorems for spatial processes are discussed in a general framework.

### 2.4 Point Processes of Closed Sets

In contrast to the preceding Sections 2.2 and 2.3, where point processes in \( \mathbb{R}^d \) are presented, i.e., random processes where the points (or elements) of the process are (random) vectors of point locations in the \( d \)-dimensional Euclidean space, this section considers a class of random point...
processes where the points are \((d\text{-dimensional})\) non-empty closed sets. Their introduction however, is rather similar to Sections 2.2 and 2.3, respectively.

Let \(\mathcal{B}(\mathcal{F}')\) denote the Borel \(\sigma\)-algebra on the family \(\mathcal{F}'\) of all non-empty closed sets in \(\mathbb{R}^d\) and let \(M(\mathcal{F}')\) be the family of simple and locally finite counting measures on \(\mathcal{B}(\mathcal{F}')\). A measurable mapping \(\Phi : \Omega \rightarrow M(\mathcal{F}')\) from some probability space \((\Omega, \sigma(\Omega), \mathbb{P})\) into the measurable space \((M(\mathcal{F}'), M(\mathcal{F}'))\) is called a \((\text{simple and locally finite})\) \textit{random point process} in \(\mathcal{F}'\) or a \textit{random point process of closed sets}, where \(\mathcal{M}(\mathcal{F}')\) denotes the \(\sigma\)-algebra on \(M(\mathcal{F}')\) generated similar as the respective \(\sigma\)-algebras in Sections 2.2.1 and 2.3.1.

In analogy to Section 2.2, a similar uniqueness statement holds as formulated in Theorem 2.3. It can be proven (Chapter 11 of [64] or Chapter 3 of [112]) that there exists a measurable sequence \(\{\Xi_n\}_{n \geq 1}\) of random closed sets \(\Xi_n : \Omega \rightarrow \mathcal{F}'\) such that \(\Phi = \{\Xi_n\}_{n \geq 1}\) can be written as \(\sum_{n \geq 1} \delta_{\Xi_n}\).

The random point process \(\Phi\) of closed sets is called \textit{stationary} if \(\Phi \overset{d}{=} \Phi + x\) for all \(x \in \mathbb{R}^d\), where \((\Phi + x)(B) = \Phi(B - x)\) for any \(B \in \mathcal{B}(\mathcal{F}')\) and \textit{isotropic} if \(\Phi \overset{d}{=} \theta_R \Phi\) for any rotation \(\theta_R\) around the origin, where \((\theta_R \Phi)(B) = \Phi(\theta_R^{-1}B)\) for any \(B \in \mathcal{B}(\mathcal{F}')\).

The \textit{intensity measure} \(\Lambda_\Phi : \mathcal{B}(\mathcal{F}') \rightarrow [0, \infty]\) of \(\Phi\) is defined by

\[\Lambda_\Phi(B) = \mathbb{E} \Phi(B), \quad B \in \mathcal{B}(\mathcal{F}'),\]

where \(\mathbb{E} \Phi(B)\) is the expectation of the random variable \(\Phi(B) : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}\).

In analogy to theorems 2.4 and 2.6, the following is true.

**Theorem 2.7** (Campbell’s theorem) Let \(\Phi = \{\Xi_n\}_{n \geq 1}\) be a random point process in \(\mathcal{F}'\) and let \(f : \mathcal{F}' \rightarrow [0, \infty]\) be a measurable function. Then \(\sum_{\Xi \in \text{supp } \Phi} f(\Xi) : \Omega \rightarrow [0, \infty]\) is measurable and

\[\mathbb{E} \sum_{\Xi \in \text{supp } \Phi} f(\Xi) = \int_{\mathcal{F}'} f(F) \Lambda_\Phi(dF),\]

where \(\text{supp } \Phi = \{F \in \mathcal{F}' : \Phi(\{F\}) > 0\}\) is the support of \(\Phi\).

Finally, the subsequent theorem, where the proof can be found on p. 100 in [112], connects a random point process of closed sets with RACSs.

**Theorem 2.8** Let \(\Phi = \{\Xi_n\}_{n \geq 1}\) be a random point process in \(\mathcal{F}'\). Then,

\[\bigcup_{\Xi \in \text{supp } \Phi} \Xi = \bigcup_{n \geq 1} \Xi_n\]

is a RACS, which is stationary (isotropic) if \(\Phi\) is stationary (isotropic).

Analogously to Section 2.2.5, Poisson point processes of closed sets can be introduced, cf. Chapter 3 of [112], and, analogously to Section 2.3, marked versions of point processes of closed sets can be regarded.
2.5 Flat Processes

2.5.1 Processes of \( k \)-Flats

For each \( k = 0, \ldots, d - 1 \) let \( \mathcal{A}_k^d \) denote the space of all affine \( k \)-dimensional subspaces in \( \mathbb{R}^d \) and let \( \mathcal{L}_k^d = \{ L \in \mathcal{A}_k^d : o \in L \} \) be the corresponding Grassmann manifold. A random point process \( \Phi_k : \Omega \rightarrow M(\mathcal{F}') \) of closed sets is said to be a random \( k \)-flat process in \( \mathbb{R}^d \) if the intensity measure \( \Lambda_k(\cdot) = \mathbb{E}\Phi_k(\cdot) \) of \( \Phi_k \) satisfies

\[
\Lambda_k(\mathcal{F}' \setminus \mathcal{A}_k^d) = 0,
\]

where we recall that \( \mathcal{F}' \) denotes the space of non-empty closed sets in \( \mathbb{R}^d \) and that \( M(\mathcal{F}') \) is the family of all (locally finite and simple) counting measures on the Borel \( \sigma \)-algebra \( \mathcal{B}(\mathcal{F}') \). The notions of stationarity and isotropy carry over from the respective notions in the case of point processes of closed sets in Section 2.4.

**Theorem 2.9** (Decomposition of the intensity measure) Let \( \Phi_k \) be a stationary \( k \)-flat process \( 0 \leq k \leq d - 1 \) in \( \mathbb{R}^d \). Provided that the intensity measure \( \Lambda_k(\cdot) \) of \( \Phi_k \) is locally finite and different from the zero measure, there exists a (uniquely defined) finite number \( \lambda_k > 0 \) and a (uniquely defined) probability measure \( \Theta_k \) on \( \mathcal{B}(\mathcal{A}_k^d) \) such that

\[
\Lambda_k(A) = \lambda_k \int_{\mathcal{L}_k^d} \int_{L^+} \mathbb{1}_A(L + x) \nu_{d-k}(dx) \Theta_k(dL) \tag{2.5}
\]

holds for any \( A \in \mathcal{B}(\mathcal{A}_k^d) \), where \( L^+ \in \mathcal{L}_{d-k}^d \) denotes the orthogonal complement of \( L \in \mathcal{L}_k^d \) and where \( \mathbb{1}_A(\cdot) \) stands for the indicator function of the set \( A \).

The proof of Theorem 2.9 can be found on pp. 107 in [112]. We call the number \( \lambda_k \) and the probability measure \( \Theta_k \) in Theorem 2.9 the intensity and the orientation distribution of \( \Phi_k \), respectively.

Formula (2.5) yields a simple interpretation of the intensity \( \lambda_k \) as ratio

\[
\lambda_k = \frac{1}{\kappa_{d-k} v_{d-k}} \mathbb{E}\Phi_k(\{ L \in \mathcal{A}_k^d : L \cap B_1^d \neq \emptyset \}) \quad \text{for all} \quad \kappa > 0. \tag{2.6}
\]

In other words, \( \lambda_k \kappa_{d-k} \) is the expected number of \( k \)-flats hitting the unit ball \( B_1^d \) in \( \mathbb{R}^d \). On the other hand, \( \lambda_k \) can be regarded as mean total \( k \)-volume of all \( k \)-flats in the unit cube \([0,1]^d\) as the following theorem shows.

**Theorem 2.10** Let \( \Phi_k \) be a stationary \( k \)-flat process \( 0 \leq k \leq d - 1 \) in \( \mathbb{R}^d \) with finite intensity \( \lambda_k > 0 \). Then,

\[
\mathbb{E} \sum_{L \in \text{supp} \Phi_k} \nu_k(B \cap L) = \lambda_k \nu_d(B), \quad B \in \mathcal{B}(\mathbb{R}^d), \tag{2.7}
\]

where \( \text{supp} \Phi_k = \{ L \in \mathcal{A}_k^d : \Phi_k(\{ L \}) > 0 \} \).
Proof Let $k \in \{0, \ldots, d-1\}$ and $B \in B_0(\mathbb{R}^d)$. Application of Campbell’s Theorem 2.7 for point processes of closed sets yields that

$$
\mathbb{E} \sum_{L \in \text{supp} \Phi_k} \nu_k(B \cap L) = \int_{A^d_k} \nu_k(B \cap L) \Lambda_k(dL).
$$

The integral on the right hand side of the latter equality can be rewritten by use of (2.5), i.e.,

$$
\int_{A^d_k} \nu_k(B \cap L) \Lambda_k(dL) = \lambda_k \int_{\mathbb{L}^d_k} \int_{L^\perp} \nu_k(B \cap (L + x)) \nu_{d-k}(dx) \Theta_k(dL).
$$

Finally, by Cavalieri’s principle, we obtain that

$$
\int_{L^\perp} \nu_k(B \cap (L + x)) \nu_{d-k}(dx) = \nu_d(B).
$$

Hence,

$$
\mathbb{E} \sum_{L \in \text{supp} \Phi_k} \nu_k(B \cap L) = \int_{\mathbb{L}^d_k} \nu_d(B) \Theta_k(dL),
$$

which allows to finish the proof by noting that $\int_{\mathbb{L}^d_k} \Theta_k(dL) = 1$. \hfill \qed

2.5.2 Hyperplane Processes

A $(d-1)$-flat in $\mathbb{R}^d$ is called hyperplane and can be parameterized by its signed perpendicular distance $p \in \mathbb{R}$ from the origin and its orientation vector $u \in S^{d-1}$, as depicted in Figure 2.4 for $d = 2$. In view of this parameterization we denote a hyperplane by $H(p, u) = \{x \in \mathbb{R}^d : \langle x, u \rangle = p\}$.

A $(d-1)$-flat process $\Phi_{d-1}$ is called a hyperplane process. Introducing the spherical orientation

![Figure 2.4: Parameterization of a line $H(p, u)$ in $\mathbb{R}^2$; cf. also Section 4.1.3](image-url)
distribution $\tilde{\Theta} : \mathcal{B}(S^d_+) \rightarrow [0, 1]$ defined by

$$\tilde{\Theta}(B) = \Theta_{d-1}(\{H(0, u) \in L^d_{d-1} : u \in B\}), \quad B \in \mathcal{B}(S^d_+),$$

we may rewrite (2.5) as

$$\Lambda_{d-1}(A) = \lambda_{d-1} \int_{S^d_+} \int_{\mathbb{R}} \mathbb{1}_A(H(p, u)) dp \tilde{\Theta}(du), \quad A \in \mathcal{B}(A^d_{d-1}).$$

Note that a spherical orientation distribution can be introduced alternatively as even (or symmetric) probability measure $\tilde{\Theta}^*$ on $\mathcal{B}(S^d_+)$, where $\tilde{\Theta}^*$ is connected with $\tilde{\Theta}$ by the relation

$$\tilde{\Theta}^*(B) = \frac{1}{2} \left( \tilde{\Theta}(B \cap S^d_+) + \tilde{\Theta}(\tilde{B} \cap S^d_+) \right), \quad B \in \mathcal{B}(S^d_+),$$

where $\tilde{B} = \{ -x : x \in B \}$ is the reflection of $B$. The symmetry condition $\tilde{\Theta}^*(B) = \tilde{\Theta}^*(\tilde{B})$ expresses the identification of hyperplanes with antipodal orientation vectors. A stationary hyperplane process $\Phi_{d-1}$ is said to be isotropic if $\tilde{\Theta}$ is the uniform distribution on $S^d_+$ or, equivalently, $\tilde{\Theta}^*$ is the uniform distribution on $S^d$.

In view of the spherical orientation distribution $\tilde{\Theta}$ introduced in (2.8), we can state a version of Crofton’s formula (cf. Section 3.2 of [111] and Chapter 7 of [112]) by

$$\int_{S^d_+} \int_{\mathbb{R}} V_j(C \cap H(p, u)) dp \tilde{\Theta}(du) = \frac{\kappa_{d-1}}{d \kappa_d} \frac{(j + 1) \kappa_d}{\kappa_j} V_{j+1}(C), \quad (2.9)$$

where $H(p, u)$ denotes a (random) hyperplane with orientation distribution $\tilde{\Theta}$ and $V_i(C)$ denotes the $i$th intrinsic volume in $\mathbb{R}^d$ for $i = 0, \ldots, d$; cf. Section A.4. Formula (2.9) holds for any $j = 0, 1, \ldots, d - 1$ and any convex compact set $C \in \mathcal{C}$; cf. Corollary 3.3.2 in [111].

Finally, we mention that it is well-known from convex geometry (Chapter 3.5 of [110] or Chapter 4.5 of [112]) that the spherical orientation distribution $\tilde{\Theta}$ on $S^d_+$ determines a unique centrally symmetric (p. 182 in [110] and [124]) convex body $\Pi(\tilde{\Theta})$ and vice versa. This convex body $\Pi(\tilde{\Theta})$ is called the associated zonoid or the Steiner convex set and is determined by

$$h(\Pi(\tilde{\Theta}), u) = \int_{S^d_+} |\langle u, v \rangle| \tilde{\Theta}(dv), \quad u \in \mathbb{R}^d,$$

where $h(C, u) = \max_{x \in C} \langle x, u \rangle$ denotes the support function of $C \in \mathcal{C}$; p. 3 in [14] for example.

2.5.3 Poisson Hyperplane Processes

A hyperplane process $\Phi_{d-1}$ is called Poisson hyperplane process if $\Phi_{d-1}$ is a Poisson point process of closed sets. The parameterization of hyperplanes introduced in Section 2.5.1 admits the representation of the (stationary) Poisson hyperplane process $\Phi_{d-1}$ (with finite intensity $\lambda_{d-1} > 0$ and spherical orientation distribution $\tilde{\Theta}$) as a stationary and independently marked
Poisson point process $\sum_{i \geq 1} \delta[p_i,u_i]$ on the real line with finite intensity $\lambda > 0$ and mark distribution $\tilde{\Theta}$ (on the mark space $S^{d-1}_+$. The intensity $\lambda$ is the expected number of one-dimensional points $P_i$ in the unit interval $[0,1)$ for $i \geq 1$, i.e., $\lambda = \mathbb{E}\#\{i : P_i \in [0,1)\}$, and in view of Formula (2.6) and the subsequent interpretation of $\lambda_{d-1}$ it can be seen that $\lambda$ and $\lambda_{d-1}$ coincide.

Now, let $k \in \{0,\ldots,d-1\}$ be fixed. The intersections of $d-k$ (distinct) hyperplanes $H_1 = H(P_1,U_1),\ldots,H_{d-k} = H(P_{d-k},U_{d-k})$ belonging to the support of $\Phi_{d-1}$ induce a stationary $k$-flat process $\Phi_k$ which can be represented as multiple sum, i.e.,

$$
\Phi_k(B) = \frac{1}{(d-k)!} \sum_{H_1,\ldots,H_{d-k} \in \text{supp } \Phi_{d-1}} \mathbb{I}_B(H_1 \cap \ldots \cap H_{d-k}), \quad B \in \mathcal{B}(\mathbb{R}^d). \tag{2.10}
$$

If the indicator function $\mathbb{I}_B(\cdot)$ in (2.10) is replaced by $\nu_k(H_1 \cap \ldots \cap H_{d-k})$, we get a stationary random measure on $\mathcal{B}(\mathbb{R}^d)$. The intensity $\lambda_k$ and the orientation distribution $\Theta_k$ of $\Phi_k$ are connected with $\lambda$ and $\tilde{\Theta}$ by

$$
\lambda_k \Theta_k(B) = \frac{\lambda^{d-k}}{(d-k)!} \int_{S^{d-k}_+} \ldots \int_{S^{d-k}_+} \mathbb{I}_B(H(0,u_1) \cap \ldots \cap H(0,u_{d-k}))
\times \nu_{d-k}(u_1,\ldots,u_{d-k}) \tilde{\Theta}(du_1) \ldots \tilde{\Theta}(du_{d-k}), \tag{2.11}
$$

for $B \in \mathcal{B}(\mathbb{R}^{d-k}_+)$, where $\nu_{d-k}(u_1,\ldots,u_{d-k})$ denotes the $(d-k)$-dimensional volume of the parallelopode spanned by the vectors $u_1,\ldots,u_{d-k} \in S^{d-k}_+$, Chapter 4 of [112].

The intensity $\lambda_k$ of $\Phi_k$ can be expressed in terms of the $(d-k)$th intrinsic volume $V_{d-k}(\Pi(\tilde{\Theta}))$ of the Steiner convex set $\Pi(\tilde{\Theta})$ associated with $\tilde{\Theta}$, i.e.,

$$
\lambda_k = \lambda^{d-k} V_{d-k}(\Pi(\tilde{\Theta})), \quad k = 0,\ldots,d-1,
$$

cf. also p. 161 in [73] as well as Section A.4 for the introduction of the notion of intrinsic volumes. In the isotropic case, where $\tilde{\Theta}$ is the uniform distribution, we have that

$$
\Pi(\tilde{\Theta}) = \frac{\kappa_{d-1}}{d \kappa_d} B^d_1,
$$

which together with Formulae (A.22) and (A.23) of Theorem A.17 implies that

$$
\lambda_k = \binom{d}{k} \frac{\kappa_d}{\kappa_k} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{d-k} \lambda^{d-k} \quad \text{for } k = 0,1,\ldots,d-1. \tag{2.12}
$$

Finally, a Poisson hyperplane process is called non-degenerate if its spherical orientation distribution $\tilde{\Theta}$ is non-degenerate, i.e., if $\tilde{\Theta}(H(0,u) \cap S^{d-1}_+) < 1$ for any $u \in S^{d-1}_+$. 
Chapter 3

Random Tessellations

This chapter is devoted to the introduction of the notion of random tessellations in $\mathbb{R}^d$ and a brief discussion of some particular tessellation models as well as the study of their properties.

From the applicational point of view, (random) tessellations serve as models in a variety of different disciplines like biology, medicine, telecommunication as well as social sciences and economics. However, not only the planar case $d = 2$ is of practical interest, but also the spatial case $d = 3$ is extensively considered, for example in material sciences and astronomy. The monograph [92] provides a lot of examples of models as well as many fields of applications and contains an extensive list of references. Further general information on random tessellations can be found for example in [73, 79, 83, 84, 112, 116].

We first introduce deterministic tessellations in Section 3.1 where some main geometric features are explained. Afterwards, we show that random tessellations can be defined in terms of the models of stochastic geometry discussed in Chapter 2 and in particular we show in Section 3.2 how (stationary) random tessellations can be represented as (stationary) random marked point processes in $\mathbb{R}^d$. In Section 3.3 we present some particular models of random tessellations which are of interest later. We conclude this chapter by studying stationary and ergodic random tessellations in $\mathbb{R}^d$, more precisely a single realization of them, in some sequence of convex sampling windows $W$. We show that the influence of cells hitting the boundary $\partial W$ of $W$ becomes irrelevant as $W \uparrow \mathbb{R}^d$, i.e., as $W$ expands unboundedly but uniformly in all directions.

3.1 Introduction

3.1.1 Polytopes and their Faces

We consider a $d$-dimensional polytope $\mathfrak{p}$ in $\mathbb{R}^d$, called $d$-polytope for short. More precisely, and following the definition in Section 3.1 of [43], $\mathfrak{p}$ is a convex body in $\mathbb{R}^d$ whose extreme points...
form a finite set. Note that for any (non-empty) convex subset $B \subset \mathbb{R}^d$, a point $x \in B$ is called extreme point of $B$ if

$$y, z \in B, \alpha \in (0, 1) \quad \text{and} \quad x = \alpha y + (1 - \alpha) z \implies x = y = z.$$  

For further details cf. Section 3.1 of [43] for example. Additionally, we assume in the following that $p$ has non-empty interior. The family of all such $d$-polytopes is denoted by $\mathcal{P}$.

Any $d$-polytope $p \in \mathcal{P}$ induces a hierarchy of $k$-dimensional polytopes ($k$-polytopes) in $\mathbb{R}^d$, called the $k$-faces of $p$, for $k = 0, \ldots, d - 1$, where a 0-polytope in $\mathbb{R}^d$ is identified with the set \( \{x\} \) containing only a single point $x \in \mathbb{R}^d$. Every $k$-face is the intersection of $p$ with $d - k$ of the hyperplanes that are the boundary of the closed halfspaces constituting $p$; $k = 0, \ldots, d - 1$. The 0-faces are called the vertices of $p$, whereas the $(d - 1)$-faces are called the facets. Given the notion of $(d - 1)$-faces as $(d - 1)$-polytopes in the boundary $\partial p$ of $p$, we may define $k$-faces recursively for $k = 0, \ldots, d - 2$ as $k$-polytopes in the relative boundaries of the $(k + 1)$-faces, where relative means with respect to the affine hull. By $\mathcal{S}_k(p)$ we denote the family of all $k$-faces of $p$ for $k = 0, \ldots, d - 1$.

### 3.1.2 Tessellations in $\mathbb{R}^d$

A (deterministic) **tessellation** in $\mathbb{R}^d$ is a subdivision of the $d$-dimensional Euclidean space $\mathbb{R}^d$ into $d$-dimensional non-overlapping closed sets, called the **cells** of the tessellation, with non-empty interior. Any pair of cells shares common points on their boundaries.

Following [112], we confine ourselves to introduce a tessellation as a locally finite family of convex bodies. Hence, a tessellation $\tau$ in $\mathbb{R}^d$ is a countable family $\tau = \{C_n\}_{n \geq 1}$ of subsets $C_n \in \mathcal{C}$ for $n \geq 1$, such that

1. the interior int $C_n \neq \emptyset$ for all $n \geq 1$,
2. int $C_n \cap$ int $C_m = \emptyset$ for $n \neq m$,
3. $\bigcup_{n=1}^{\infty} C_n = \mathbb{R}^d$, and
4. $\sum_{n=1}^{\infty} 1_{\mathcal{F}_K}(C_n) < \infty$ for every $K \in \mathcal{K}$, where $\mathcal{F}_K$ is a hitting set (Section 2.1.1).

By $\mathcal{T}$ we denote the family of all tessellations in $\mathbb{R}^d$. Note that $\mathcal{T}$ is a Borel set in the family of all closed subsets of $\mathcal{F}'$ and that it can be proven that the cells of a tessellation $\tau \in \mathcal{T}$ are $d$-polytopes, i.e., $C_n \in \mathcal{P}$ for $n \geq 1$; Lemma 6.1.1 in [112].

For a tessellation $\tau = \{C_n\}_{n \geq 1}$, let $\bigcap_{n: x \in C_n} C_n \neq \emptyset$ be the finite intersection of (neighboring) cells of $\tau$ which contain $x$, i.e., $\bigcap_{n: x \in C_n} C_n$ is a $k$-polytope for some $k = 0, \ldots, d$ and let

$$\Upsilon_k(\tau) = \left\{ \bigcap_{n: x \in C_n} C_n : x \in \mathbb{R}^d, \dim \bigcap_{n: x \in C_n} C_n = k \right\}, \quad k = 0, \ldots, d,$$
where \( \dim \mathbf{p} \) denotes the dimension of the polytope \( \mathbf{p} \). An element of \( \Upsilon_k(\tau) \) is called a \( k \)-facet of the tessellation, in particular for \( k = d \) the elements of \( \Upsilon_d(\tau) \) are the cells of \( \tau \). Note that for \( k = 0 \) and \( k = d \),
\[
\Upsilon_k(\tau) = \bigcup_{n \geq 1} S_k(C_n).
\]

If the latter identity holds additionally for all \( k = 1, \ldots, d-1 \), then \( \tau \) is called regular (Section 3 of [83]), or, alternatively, face-to-face (p. 235 in [112]). Moreover, a tessellation \( \tau \) is called normal or, alternatively, ordinary if every \( k \)-facet is situated in the boundary of exactly \( d-k+1 \) cells for \( k = 0, \ldots, d-1 \). Note that this condition is always fulfilled for \( k = d-1 \). More generally each \( k \)-facet of a normal tessellation \( \tau \) is situated in
\[
\binom{d-k+1}{l-k}
\]
\( l \)-facets of \( \tau \) for \( 0 \leq k \leq l \leq d \). Figure 3.1 illustrates the geometrical notions of regularity and normality for tessellations in \( \mathbb{R}^2 \).

![Figure 3.1: Examples of non-regular, regular, and normal tessellations](image)

### 3.2 Random Tessellations and their Representations

#### 3.2.1 Definition and Basic Properties

A random tessellation \( \Psi \) in \( \mathbb{R}^d \) can be introduced as a special kind of point process in \( \mathcal{F}' \). More precisely, \( \Psi \) can be defined as a random point process of (non-empty) closed sets \( \Phi \) (cf. Section 2.4) whose intensity measure \( \Lambda_\Phi \) is concentrated on \( \mathcal{C}' \), i.e., \( \Lambda_\Phi(\mathcal{F}' \setminus \mathcal{C}') = 0 \), and where additionally \( \mathbb{P}(\text{supp} \Phi \in \mathcal{T}) = 1 \). With this definition we follow [112], where the notion of a random particle process is introduced in order to denote random point processes of closed sets with intensity measure concentrated on subsets of \( \mathcal{F}' \). The additional condition \( \mathbb{P}(\text{supp} \Phi \in \mathcal{T}) = 1 \) allows then to interpret a random tessellation \( \Psi \) as a point process of \( d \)-polytopes \( \Xi_n \), expressed by the notation \( \Psi = \{ \Xi_n \}_{n \geq 1} \), where these \( d \)-polytopes for \( n \geq 1 \) are pairwise non-overlapping with non-empty interior and fill the whole space. An alternative
but equivalent definition of random tessellations in $\mathbb{R}^d$ without using the notion of particle processes explicitly can be found on p. 46 in [83]. In accordance with Section 2.4, the notions of stationarity and isotropy conveniently carry over. To this end we assume that $\Psi$ is stationary.

The random tessellation $\Psi$ is called normal and regular if $\mathbb{P}$-a.s. all realizations of $\Psi$ are normal and regular, respectively. Note that there are important classes of stationary tessellations in $\mathbb{R}^d$ whose cells are constructed (realizationwise) according to specific geometric rules from the atoms of a stationary point process in $\mathbb{R}^d$. Among them are Voronoi and Laguerre tessellations as described in [92] in detail, which turn out to be normal if the generating point process is a stationary Poisson process ([83]). Furthermore, note also that this fact seems to remain preserved for a large class of “smooth” (even non-stationary) generating point processes. For example it is shown in [50] that Voronoi tessellations in $\mathbb{R}^d$ are normal if the $(d+2)$-th order product density of the generating stationary point process exists.

### 3.2.2 Marked Point Process Representation and Typical Cell

For many applications it is often quite convenient to represent a (stationary) random tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ as a (stationary) random marked point process in $\mathbb{R}^d$ by introducing the concept of particle centroids, cf. Chapter 4 in [112] and in particular Satz 4.3.1 on p. 136 in [112]).

Let $\alpha : C' \to \mathbb{R}^d$ be a measurable mapping that satisfies

$$\alpha(C) \in C' \quad \text{and} \quad \alpha(C + x) = \alpha(C) + x$$

for any $C \in C'$ and $x \in \mathbb{R}^d$. For a stationary random tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$, we call $\alpha(\Xi_n)$ the associated point or centroid of the nth cell $\Xi_n$, $n \geq 1$. Examples of choices for $\alpha(\Xi_n)$ are the center of the ball circumscribing $\Xi_n$ ([112]) or the center of gravity of the vertices of $\Xi_n$ ([83]). The crucial result of this section is that any stationary tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ can be represented as stationary marked point process

$$X_\Psi = \sum_{n \geq 1} \delta_{\alpha(\Xi_n), \Xi_n},$$

where $\Xi_n^o = \Xi_n - \alpha(\Xi_n)$ is called the $n$th centered cell containing the origin $\alpha$ and where the mark space $\mathcal{P}^o$ is the family of all $d$-polytopes with their associated point at the origin. The intensity $\gamma$ of $X_\Psi$ is given by

$$\gamma = \mathbb{E}\# \{n : \alpha(\Xi_n) \in [0, 1)^d\},$$

where we assume that $0 < \gamma < \infty$ in the following and where we also call $\gamma$ the intensity of the stationary tessellation $\Psi$. Moreover, the Palm mark distribution $P^o$ of $X_\Psi$ is given by

$$P^o(B) = \frac{1}{\gamma} \mathbb{E}\# \{n : \alpha(\Xi_n) \in [0, 1)^d, \Xi_n^o \in B\}$$
for $B \in \mathcal{B}(F') \cap \mathcal{P}^o$.

The latter distribution can also be used to characterize the typical cell of $\Psi$. More precisely, we call a random polytope $\Xi^* : \Omega \rightarrow \mathcal{P}^o$ the \textit{typical cell} of the stationary random tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ if the distribution of $\Xi^*$ coincides with $P^o$ defined in (3.4). Then, $\Xi^*$ is related with the intensity $\gamma > 0$ of $\Psi$ by

$$
\gamma^{-1} = \mathbb{E} \nu_d(\Xi^*) = \int_{\mathcal{P}^o} \nu_d(C) P^o(dC),
$$

(3.5)

cf. also p. 238 in [112], where the relation in (3.5) states that the mean volume of the typical cell is the reciprocal of the intensity $\gamma$ of $\Psi$.

### 3.2.3 Facet Processes

Any random tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ in $\mathbb{R}^d$ induces lower dimensional point processes $\Psi^{(k)}$ of $k$-polytopes as well as lower dimensional manifold processes $\tilde{\Psi}^{(k)}$ for $k = 0, \ldots, d-1$. More precisely, the process $\Psi^{(k)}$ defined by

$$
\text{supp} \Psi^{(k)} = \Upsilon_k(\Psi), \quad k = 0, \ldots, d-1,
$$

(3.6)

is called the \textit{point process of $k$-facets} of the random tessellation $\Psi$, p. 237 in [112], whereas the process

$$
\tilde{\Psi}^{(k)} = \bigcup_{F \in \Upsilon_k(\Psi)} F, \quad k = 0, \ldots, d-1
$$

(3.7)

may be interpreted as the $k$-dimensional skeleton induced by $\Psi$ in the $d$-dimensional Euclidean space and is called the $k$-\textit{facet process} of $\Psi$. Both processes $\Psi^{(k)}$ and $\tilde{\Psi}^{(k)}$ are stationary (isotropic) if the inducing random tessellation $\Psi$ is stationary (isotropic). Note that $\Psi^{(0)}$ and $\tilde{\Psi}^{(0)}$ coincide and are called the \textit{point process of vertices} of $\Psi$, whereas $\tilde{\Psi}^{(1)}$ is the \textit{process of edges} of $\Psi$. Moreover, if we allow for $k = d$, we have that $\Psi^{(d)}$ coincides with $\Psi$ itself.

Note that the point process of $k$-facets $\Psi^{(k)}$ can be represented as a marked point process in $\mathbb{R}^d$, cf. Chapter 4 and 6 of [112] as well as [71], analogously to the representation of $\Psi$ itself in (3.2). Hence, since we assume $\Psi$ to be stationary, we may introduce the intensity $\gamma^{(k)}$ of $\Psi^{(k)}$ for any $k = 0, \ldots, d-1$ as the mean number of $k$-facets per unit volume, i.e.,

$$
\gamma^{(k)} = \mathbb{E} \# \{\alpha(F) \in [0,1]^d : F \in \Upsilon_k(\Psi)\}, \quad k = 0, \ldots, d-1.
$$

(3.8)

Here, $\alpha : C' \rightarrow \mathbb{R}^d$ is a measurable mapping satisfying (3.1) and $\alpha(F)$ is called the associated point or the centroid of the $k$-facet $F \in \Upsilon_k(\Psi)$ for any $k = 0, \ldots, d-1$; Section 4 of [83].

Likewise, we introduce the intensity $\tilde{\gamma}^{(k)}$ of the $k$-facet process $\tilde{\Psi}^{(k)}$ for $k = 0, \ldots, d-1$ by

$$
\tilde{\gamma}^{(k)} = \mathbb{E} \nu_k(\tilde{\Psi}^{(k)} \cap [0,1]^d), \quad k = 0, \ldots, d-1,
$$

(3.9)

where the measure $\nu_k(\cdot)$ is explained in Section A.1.3 for $k = 0, \ldots, d-1$. In the following we assume that $0 < \tilde{\gamma}^{(k)} \leq \gamma^{(k)} < \infty$ for $k = 0, \ldots, d-1$ and we note aside that there exists a connection between both intensities; Formula (6.3) on p. 237 in [112].
3.3 Random Tessellation Models

In this section we review three specific (stationary) tessellation models which play a central role in many applications, for example in network modelling within the framework of the SSLM. In particular, we begin with the Poisson–Voronoi tessellation (PVT) in Section 3.3.1, continue with the Poisson–Delaunay tessellation (PDT) in Section 3.3.2 and close with the Poisson hyperplane tessellation (PHT) in Section 3.3.3.

Further information about these tessellation models, their non-stationary versions, as well as about alternative and more general models like Laguerre or Johnson–Mehl tessellations with curved boundaries can be found in [83, 92, 112, 116] as well as the references therein. In particular we would like to point to the work of R. Miles, for example [68, 81, 82]. Mean value relationships for the considered models, in particular in $d = 2$, are investigated in [63, 78, 87, 126] for example, while typical cells are studied in [24, 25, 28, 29, 77]. The Poisson–Voronoi tessellation is analysed in detail in [84] and particular considerations about its edges are given in [86] and [106]. Second-order quantities of random tessellations in $d$ dimensions ($d \geq 2$) are studied in [125] as well as for $d = 2$ and Poisson–Voronoi tessellations in [48].

3.3.1 Poisson–Voronoi Tessellation

Let $B = \{x_1, x_2, \ldots \} \subset \mathbb{R}^d$ be a locally finite set of points $x_1, x_2, \ldots$ in $\mathbb{R}^d$ whose convex hull $\text{conv} \ B$ is the whole Euclidean space $\mathbb{R}^d$, and consider the set

$$C_{\text{Vor}}(x_n, B) = \{y \in \mathbb{R}^d : \|x_n - y\| \leq \|x_m - y\| \text{ for all } x_m \in B \setminus \{x_n\} \}, \quad x_n \in B$$

that contains all points of $\mathbb{R}^d$ which are closer to $x_n \in B$ (in the Euclidean sense) than to $x_m \in B$ for all $m \neq n$. Alternatively, we may put

$$C_{\text{Vor}}(x_n, B) = \bigcap_{m \geq 1 \atop m \neq n} \left\{ y \in \mathbb{R}^d : \langle y, x_m - x_n \rangle \leq \frac{1}{2} \left( \|x_n\|^2 + \|x_m\|^2 \right) \right\},$$

where the intersection is over all closed halfspaces which contain $x_n$ for $m \neq n$ and which are bounded by the hyperplane perpendicular to $x_m - x_n$ and going through $(x_n + x_m)/2$. Immediately it can be observed that $C_{\text{Vor}}(x_n, B)$ is a closed and convex set with int $C_{\text{Vor}}(x_n, B) \neq \emptyset$ for $n \geq 1$. Moreover, since $\text{conv} \ B = \mathbb{R}^d$ is assumed, the Voronoi cells $C_{\text{Vor}}(x_n, B)$ are bounded for any $n \geq 1$ and $\tau = \{C_{\text{Vor}}(x_n, B)\}_{n \geq 1}$ is a regular tessellation; p. 256 in [112].

The family $\tau = \{C_{\text{Vor}}(x_n, B)\}_{n \geq 1}$ is called a Voronoi tessellation in $\mathbb{R}^d$ (with respect to $B$), where the sets $C_{\text{Vor}}(x_n, B)$ are called the Voronoi cells of $\tau$, and each $x_n \in B$, $n \geq 1$, is called the nucleus of $C_{\text{Vor}}(x_n, B)$.

Let $X = \{X_n\}_{n \geq 1}$ be a stationary Poisson point process in $\mathbb{R}^d$ as defined in Section 2.2.5 with intensity $\lambda_X$. Then $\Psi = \{\Xi_n\}_{n \geq 1}$ with $\Xi_n = C_{\text{Vor}}(X_n, X)$ for $n \geq 1$ is called a stationary Poisson–Voronoi tessellation (PVT) in $\mathbb{R}^d$. It can be shown that $\Psi$ is a normal tessellation.
P-a.s. (Satz 6.2.3 on p. 258 in [112]) and that $\Psi$ is both mixing and ergodic (Satz 6.4.1 on p. 287 in [112]). Apart from the intensity $\gamma$ introduced in (3.3), we consider the parameter $\gamma_{\text{PVT}}$, which is also called intensity (of the PVT) in the following and which coincides with both $\lambda_X$ and $\gamma$. Hence, $\gamma_{\text{PVT}}$ is the mean number of Voronoi cells per unit volume and whenever the intensity of a PVT is mentioned in the following, this is done with $\gamma_{\text{PVT}}$ in mind. Figure 3.2 shows realizations of stationary PVTs for $d = 2$ with different intensity values $\gamma_{\text{PVT}}$.

![Figure 3.2: Poisson–Voronoi tessellations with different intensities $\gamma_{\text{PVT}}$](image)

The following mean value relationships are important in many applications. The intensity $\tilde{\gamma}_{\text{PVT}}^{(k)}$ of the $k$-facet process $\tilde{\Psi}^{(k)}$ is related to the intensity $\gamma_{\text{PVT}}$ by

$$
\tilde{\gamma}_{\text{PVT}}^{(k)} = \frac{2^{d-k+1} \pi(d-k)/2}{d(d - k + 1)!} \frac{\Gamma(d^2-k(d-1)+1)}{\Gamma(d^2-k/2)^{d-k} \Gamma(1 + \frac{d}{2})^{d-k+k/d} \Gamma(d - k + \frac{k}{d})} \gamma_{\text{PVT}}^{1-k/d},
$$

for any $k = 0, \ldots, d - 1$, cf. Satz 6.2.4 on p. 258 in [112], where $\Gamma(\cdot)$ denotes the Gamma function. Note that in case $k = 0$ the latter formula coincides with Formula (7.7) in [83] for $\gamma_{\text{PVT}}^{(0)}$, i.e.,

$$
\gamma_{\text{PVT}}^{(0)} = \frac{2^{d+1} \pi(d-1)/2}{d^2 (d + 1) \Gamma(d + 2)} \left( \frac{\Gamma(d + 1)}{\Gamma(d + 2)} \right)^d \gamma_{\text{PVT}},
$$

$$
= c_d \gamma_{\text{PVT}} \quad \text{with} \quad c_d = \frac{2^d \pi^{d-1} \kappa_{d-1}}{d + 1 \kappa_d} \left( \frac{\kappa_{d-1}}{\kappa_d} \right)^d.
$$

In the particular case $d = 2$ we get that

$$
\tilde{\gamma}_{\text{PVT}}^{(0)} = \gamma_{\text{PVT}}^{(0)} = 2 \gamma_{\text{PVT}}, \quad \tilde{\gamma}_{\text{PVT}}^{(1)} = 2 \sqrt{\gamma_{\text{PVT}}}, \quad \gamma_{\text{PVT}}^{(1)} = 3 \gamma_{\text{PVT}}, \quad \gamma_{\text{PVT}}^{(2)} = \gamma_{\text{PVT}}.
$$

### 3.3.2 Poisson–Delaunay Tessellation

As pointed out in [92] in detail, Delaunay tessellations are dual to Voronoi tessellations. Let $B = \{x_1, x_2, \ldots\} \subset \mathbb{R}^d$ be a locally finite set whose convex hull is the whole Euclidean space

$$
\Psi(x) = \frac{1}{\lambda_X} = \frac{1}{\gamma_{\text{PVT}}}.
$$
Consider the Voronoi tessellation $\tau = \{C_{\text{Vor}}(x_n, B)\}_{n \geq 1}$ with respect to $B$. For a vertex of $\tau$, i.e., a point $x_0 \in \mathcal{T}_0(\tau)$, let

$$C_{\text{Del}}(x_0, B) = \operatorname{conv} \{x \in B : x_0 \in \mathcal{S}_0(C_{\text{Vor}}(x, B))\},$$

where $\operatorname{conv} \{\cdot\}$ denotes the convex hull of the set $\{\cdot\}$. Then the family $\tau' = \{C_{\text{Del}}(x_0, B) : x_0 \in \mathcal{T}_0(\tau)\}$ is a regular tessellation, called the Delaunay tessellation (with respect to $B$), consisting of the Delaunay cells $C_{\text{Del}}(x_0, B)$, $x_0 \in \mathcal{T}_0(\tau)$, which are closed and convex as well as bounded since we assume $\operatorname{conv} B = \mathbb{R}^d$.

Let $\Psi$ be a stationary Poisson–Voronoi tessellation induced by the stationary Poisson point process $X = \{X_n\}_{n \geq 1}$ with intensity $\lambda_X$. The dual tessellation $\Psi' = \{\Xi_n\}_{n \geq 1}$ with $\Xi_n = C_{\text{Del}}(X_n^{(0)}, X)$ and $X_n^{(0)} \in \mathcal{T}_0(\Psi)$ is called a Poisson–Delaunay tessellation (PDT). This tessellation $\Psi'$ can be shown to be normal $\mathcal{P}$-a.s. (p. 264 in [112]) as well as both mixing and ergodic (Satz 6.4.2 on p. 289 in [112]). Apart from the intensity $\gamma$ introduced in (3.3), we consider the intensity $\gamma_{\text{PDT}}$ of the PDT $\Psi'$ in the following, where $\gamma_{\text{PDT}} = \lambda_X$. Hence, $\gamma_{\text{PDT}}$ can be interpreted as the mean number of vertices of $\Psi'$ per unit volume.

The duality of a Poisson–Voronoi tessellation $\Psi$ and a Poisson–Delaunay tessellation $\Psi'$ is also expressed through a simple relationship between the intensity $\gamma_{\text{PVT}}^{(k)}$ of the point process of $k$-facets of $\Psi$ with the respective intensity $\gamma_{\text{PDT}}^{(k)}$ of $\Psi'$ for $k = 0, \ldots, d - 1$ and also $k = d$. More precisely, it can be shown that

$$\gamma_{\text{PDT}}^{(k)} = \gamma_{\text{PVT}}^{(d-k)}, \quad k = 0, \ldots, d, \quad (3.12)$$

where $\gamma_{\text{PDT}}^{(d)} = \gamma_{\text{PVT}}^{(0)} = \gamma$; Satz 6.2.9 on p. 265 in [112].

For the particular case $d = 2$, Figure 3.3 displays examples of stationary PDTs with different values $\gamma_{\text{PDT}}$. Furthermore, usage of (3.12) together with the mean value relationships in (3.11)

![Figure 3.3: Poisson–Delaunay tessellations with different intensities $\gamma_{\text{PDT}}$](image)

leads to mean value relationships for PDTs, i.e.,

$$\gamma_{\text{PDT}}^{(0)} = \gamma_{\text{PDT}}, \quad \gamma_{\text{PDT}}^{(1)} = 3\gamma_{\text{PDT}}, \quad \gamma_{\text{PDT}}^{(2)} = 2\gamma_{\text{PDT}}, \quad \text{and} \quad \gamma_{\text{PDT}}^{(1)} = \frac{32}{3\pi} \sqrt{\gamma_{\text{PDT}}}. \quad (3.13)$$
3.3.3 Poisson Hyperplane Tessellation

Recall that in Section 2.5.1 and in particular in Section 2.5.3 we introduced (stationary) Poisson hyperplane processes $\Phi_d$ as (stationary) point processes in $\mathcal{F}$ with intensity measure concentrated on the space $\mathcal{A}_d$ of $(d-1)$-dimensional affine subspaces of $\mathbb{R}^d$.

Let $\Phi_d$ be a stationary and non-degenerate Poisson hyperplane process in $\mathbb{R}^d$ with intensity $\lambda_d > 0$. Recall that non-degenerate means $\Theta(H(0,u) \cap \mathbb{S}^{d-1}_+^0) < 1$ for any $u \in \mathbb{S}^{d-1}_+$, where $\Theta$ is the spherical orientation distribution of $\Phi_d$.

Then, $\Phi_d$ induces a (stationary) Poisson hyperplane tessellation (PHT) $\Psi$ consisting $\mathbb{P}$-a.s. of bounded cells $\Xi_n$ for $n \geq 1$, where the intensity $\gamma_{\text{PHT}}$ of $\Psi$ coincides with the intensity $\lambda_d$ of the generating Poisson hyperplane process $\Phi_d$. Hence, the intensity $\gamma_{\text{PHT}}$ can be interpreted according to Section 2.5.3 as the mean total $(d-1)$-volume of the $(d-1)$-facets of $\Psi$ in $[0,1)^d$. By Satz 6.4.3 on p. 289 in [112] we have that $\Psi$ is mixing and ergodic.

Similarly to the formulae for the intensity $\lambda_k$ of the $k$-flat intersection processes in Section 2.5.3, it can be shown that the intensity $\gamma_{\text{PHT}}^{(k)}$ of the induced point process of $k$-facets as well as the intensity $\tilde{\gamma}_{\text{PHT}}^{(k)}$ of the $k$-facet process of $\Psi$ is related to the intensity $\gamma_{\text{PHT}}$ for any $k = 0, \ldots, d-1$; Satz 6.3.3 on p. 275 in [112]. More precisely, for a stationary PHT $\Psi$ it holds that

$$\tilde{\gamma}_{\text{PHT}}^{(k)} = V_{d-k}(\Pi(\tilde{\Theta})) \gamma_{\text{PHT}}^{d-k} \quad \text{and} \quad \gamma_{\text{PHT}}^{(k)} = \binom{d}{k} V_d(\Pi(\tilde{\Theta})) \gamma_{\text{PHT}}^d, \quad k = 0, \ldots, d-1,$$

cf. formulae (6.45) and (6.46) on p. 275 in [112], respectively, where $V_d(\Pi(\tilde{\Theta}))$ and $V_{d-k}(\Pi(\tilde{\Theta}))$ denote the $d$th intrinsic volume and the $(d-k)$-th intrinsic volume of the Steiner convex set $\Pi(\tilde{\Theta})$ associated with the spherical orientation distribution $\Theta$ of the generating Poisson hyperplane process $\Phi_d$. If $\Psi$ is additionally isotropic, i.e., $\Theta$ is the uniform distribution, we have that

$$\Pi(\tilde{\Theta}) = \frac{K_{d-1}}{d \kappa_d} B_1^d,$$

which together with formulae (A.22) and (A.23) of Theorem A.17 implies that

$$\tilde{\gamma}_{\text{PHT}}^{(k)} = \binom{d}{k} \frac{K_d}{\kappa_d} \left(\frac{K_{d-1}}{d \kappa_d}\right)^{d-k} \gamma_{\text{PHT}}^{d-k}, \quad k = 0, \ldots, d-1, \quad (3.14)$$

cf. Formula (6.47) on p. 275 in [112]. Likewise, we get that

$$\gamma_{\text{PHT}}^{(k)} = \binom{d}{k} \kappa_d \left(\frac{K_{d-1}}{d \kappa_d}\right)^d \gamma_{\text{PHT}}^d, \quad k = 0, \ldots, d-1, \quad (3.15)$$

cf. Formula (6.48) on p. 275 in [112].

In the two-dimensional case $d = 2$, we call a Poisson hyperplane tessellation a Poisson line tessellation (PLT), where Figure 3.4 shows realizations of PLTs for different values of $\gamma_{\text{PLT}}$. Moreover, for $d = 2$ and in view of (3.14),

$$\gamma_{\text{PLT}}^{(0)} = \frac{1}{\pi} \gamma_{\text{PLT}}^{(0)}, \quad \gamma_{\text{PLT}}^{(1)} = \frac{2}{\pi} \gamma_{\text{PLT}}^{(1)}, \quad \gamma_{\text{PLT}}^{(2)} = \frac{1}{\pi} \gamma_{\text{PLT}}^{(2)}, \quad \text{and} \quad \tilde{\gamma}_{\text{PLT}}^{(1)} = \gamma_{\text{PLT}}^{(1)}. \quad (3.16)$$
3.4 Stationary Iterated Tessellations

In this section we introduce the notion of a stationary iterated tessellation. Especially in their particular form of $k$-fold nestings, $k \in \mathbb{N}_0$, iterated tessellations play an important role in many applications. In the SSLSM they are used to obtain more realistic models of the geometric network structure compared to non-iterated tessellations.

Stationary iterated tessellations have been studied thoroughly in \(d=2\), in particular with respect to mean value relationships, in [63], [87], and [126] as well as in [69] and [71] for higher dimensions. Note that in [88] a class of tessellations is considered where the examined tessellations are stable with respect to iteration. In [58] and [70], functionals of the typical cell of special kinds of stationary iterated tessellations are examined.

3.4.1 Description of the Model

Let \(\tau_0 = \{C_n\}_{n \geq 1}\) be a tessellation in \(\mathbb{R}^d\) and consider a sequence \(\{\tau_n\}_{n \geq 1}\) of further tessellations \(\tau_n = \{C_{n\ell}\}_{\ell \geq 1}\) in \(\mathbb{R}^d\). Then,

\[
\tau = \{ C_{n\ell} \cap C_n : \text{int} (C_{n\ell} \cap C_n) \neq \emptyset \}_{n,\ell \geq 1}
\]

is an iterated tessellation in \(\mathbb{R}^d\) with initial tessellation \(\tau_0\) and component tessellations \(\tau_1, \tau_2, \ldots\), where each cell of \(\tau\) is the intersection of a cell of \(\tau_0\) with a cell of \(\tau_{n'}\) for some \(n' \geq 1\).

We recall the following theorem, the proof of which can be found in [69].

**Theorem 3.1** Let \(\Psi_0 = \{\Xi_n\}_{n \geq 1}\) be a random tessellation in \(\mathbb{R}^d\) and let \(M(\mathcal{F}')\) be the family of all locally finite (and simple) counting measures on \(\mathcal{B}(\mathcal{F}')\).
3.4. Stationary Iterated Tessellations

(a) Let $\Xi$ be a random convex body in $\mathbb{R}^d$ with $\text{int}(\Xi) \neq \emptyset$. Then the mapping $\Psi(\cdot \mid \Xi) : \Omega \rightarrow M(\mathcal{F}')$ defined by

$$\Psi(\cdot \mid \Xi) = \sum_{n \geq 1} \delta_{\Xi_n \cap \Xi(\cdot)} \mathbb{1}_{(\text{int}(\Xi_n \cap \Xi) \neq \emptyset)}$$

is a point process in $\mathcal{C}'$.

(b) Let $\{\Psi_n\}_{n \geq 1}$ be a sequence of random tessellations $\Psi_n = \{\Xi_{nt}\}_{t \geq 1}$ in $\mathbb{R}^d$. Then the mapping $\Psi : \Omega \rightarrow M(\mathcal{F}')$ defined by

$$\Psi(B) = \sum_{n \geq 1} \Psi_n(B \mid \Xi_n), \quad B \in \mathcal{B}(\mathcal{F}')$$

(3.17)

is a random tessellation in $\mathbb{R}^d$.

The random tessellation $\Psi$ introduced in (3.17) of Theorem 3.1 is called an iterated random tessellation in $\mathbb{R}^d$ with random initial tessellation $\Psi_0$ and random component tessellation $\{\Psi_n\}_{n \geq 1}$. A random iterated tessellation $\Psi$ with a stationary (isotropic) initial tessellation $\Psi_0$ and a sequence $\{\Psi_n\}_{n \geq 1}$ of component tessellations $\Psi_n$ which are independent of $\Psi_0$ for $n \geq 1$ and where the distribution of $\{\Psi_n\}_{n \geq 1}$ is invariant under arbitrary translations (rotations around the origin) is stationary (isotropic) itself. Further details can be found in [71].

There are several particular kinds of random iterated tessellations.

(i) Formally, each random tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ itself can be interpreted as a random iterated tessellation in $\mathbb{R}^d$ by putting $\Psi_n = \Psi$ for $n \in \mathbb{N}$, meaning that $\Psi_n(\cdot \mid \Xi_n) = \delta_{\Xi_n}$ for all $n \in \mathbb{N}$.

(ii) Consider an iterated tessellation $\Psi$ with initial tessellation $\Psi_0 = \{\Xi_n\}_{n \geq 1}$ and a sequence $\{\Psi_n\}_{n \geq 1}$ of component tessellations which are independent of $\Psi_0$ and where $\Psi_1 = \ldots = \Psi_n$ for all $n \in \mathbb{N}$, meaning that $\Psi_n(\cdot \mid \Xi_n) = \Psi_1(\cdot \mid \Xi_n)$ for all $n \in \mathbb{N}$. Then, $\Psi$ is called superposition of $\Psi_0$ and $\Psi_1$, denoted as $\Psi_0/\Psi_1$-superposition; cf. Figure 3.5.

(iii) Consider an iterated tessellation $\Psi$ with initial tessellation $\Psi_0 = \{\Xi_n\}_{n \geq 1}$ and a sequence $\{\Psi_n\}_{n \geq 1}$ of component tessellations, where $\Psi_1, \Psi_2, \ldots$ are assumed to be iid random tessellations which are independent of $\Psi_0$. Then, $\Psi$ is called nesting of $\Psi_0$ and $\Psi_1$, denoted as $\Psi_0/\Psi_1$-nesting; cf. Figure 3.6.

Any $\Psi_0/\Psi_1$-nesting can be generalized by introducing the notion of Bernoulli thinning. This means that an additional parameter $p \in [0, 1]$ is introduced which is the probability of iteration in each cell. Let $\{Z_n\}_{n \geq 1}$ be a sequence of iid random variables $Z_n : \Omega \rightarrow \{0, 1\}$ for $n \geq 1$ with $\mathbb{P}(Z_n = 0) = 1 - p$ and $\mathbb{P}(Z_n = 1) = p$. Assume that $\Psi_0 = \{\Xi_n\}_{n \geq 1}$ is a stationary random tessellation in $\mathbb{R}^d$ and recall that $\Psi_0$ is by definition a (special kind of) point process.
of closed sets $\sum_{n \geq 1} \delta_{\Xi_n}$; Section 2.4. In the following we consider the independently marked version $\Psi'_0 : \Omega \rightarrow \bar{M}(\mathcal{F}') \times \{0,1\}$ of $\Psi_0$ given by

$$\Psi'_0 = \sum_{n \geq 1} \delta_{[\Xi_n, Z_n]}.$$ 

Furthermore, let $\{\{\Psi_{0,n}, \Psi_{1,n}\}\}_{n \geq 1}$ be a family of random tessellations in $\mathbb{R}^d$ such that $\Psi_{0,n} = \Psi_0$ for all $n \geq 1$ and $\{\Psi_{1,n}\}_{n \geq 1}$ is a sequence of iid random tessellations $\Psi_{1,n} = \{\Xi_{n,\ell}\}_{\ell \geq 1}$ independent of $\Psi_0$ and with a distribution that is invariant under arbitrary translations. Then it can be shown that the mapping $\Psi : \Omega \rightarrow \bar{M}(\mathcal{F}')$, given by

$$\Psi(B) = \sum_{[\Xi_n, Z_n] \in \text{supp} \Psi'_0} \Psi_0,n(B \mid \Xi_n) \mathbb{I}_{[0]}(Z_n) + \sum_{[\Xi_n, Z_n] \in \text{supp} \Psi'_0} \Psi_{1,n}(B \mid \Xi_n) \mathbb{I}_{[1]}(Z_n).$$
for $B \in \mathcal{B}(\mathcal{F}')$ is a random tessellation in $\mathbb{R}^d$. More precisely, we call $\Psi$ a $\Psi_0/p\Psi_1$-nesting with Bernoulli thinning. Figure 3.7 displays some examples for $d = 2$. Note that such a nesting can be even further generalized to multi-type nestings ([58]), where Figure 3.8 displays some examples for $d = 2$, or clustered iterated tessellations ([69]). Figure 3.8 (a) shows a PLT with intensity 0.1 as initial tessellation, where the nesting is a PLT of intensity 0.1 with probability 0.5 and a PVT of intensity 0.01 with probability 0.5, respectively. Figure 3.8 (b) shows a PVT with intensity 0.001 as initial tessellation $\Psi_0$, where the nested tessellation is a PLT of intensity 0.1 with probability 0.3 and a PVT of intensity 0.01 with probability 0.4, respectively. Hence, with probability 0.3 a cell of the initial tessellation remains empty. Figure 3.8 (c) shows another example with $\Psi_0$ being a PDT.

Finally, we introduce $k$-fold iterated random tessellations. Let $\Psi_0$ be a random tessellation in $\mathbb{R}^d$ and let $\{\Psi_{k,\ell}\}_{\ell \geq 1}$ be a family of random tessellations for $k \in \mathbb{N}_0$. Assume that $\Psi_{[k]} =$
\[ \{\Xi_n\}_{n \geq 1} \text{ is a } k\text{-fold iterated tessellation in } \mathbb{R}^d. \text{ Then,} \\
\Psi_{k+1}(B) = \sum_{n \geq 1} \Psi_{k,n}(B \mid \Xi_n), \quad B \in \mathcal{B}(\mathcal{F}) \]

with \( \Psi_0 = \Psi_0 \) is called an \((k + 1)\text{-fold iterated random tessellation}\) in \( \mathbb{R}^d \) with random initial tessellation \( \Psi_0 \) and \( i \) level random component tessellation \( \{\Psi_{i,n}\}_{n \geq 1} \) for \( i = 1, \ldots, k \). In particular, if \( \{\Psi_{i,n}\}_{n \geq 1} \) forms a sequence of iid random tessellations which are independent of \( \Psi_0 \), we get the notion of a \( \Psi_0/\Psi_1/\ldots/\Psi_k \)-nesting. Figure 3.9 displays some examples of two-fold tessellations in \( d = 2 \) dimensions.

![Two-fold \( \Psi_0/\Psi_1/\Psi_2 \)-nestings (intensities \( \gamma_0, \gamma_1, \gamma_2 \), respectively)](image)

3.4.2 Mean Value Relationships

Mean value relationships for facet processes of stationary iterated tessellations as already presented in the formulae (3.11), (3.13), and (3.16) for special kinds of stationary (and isotropic) non-iterated tessellation models are provided in two dimensions in [87] and [126] and are extended to higher dimensions \( d \geq 2 \) in [69] and [71].

In particular for \( d = 2 \), let \( \Psi \) be a stationary \( \Psi_0/\rho\Psi_1 \)-nesting with Bernoulli thinning. Assume that the initial tessellation \( \Psi_0 \) is stationary and that the sequence \( \{\Psi_n\}_{n \geq 1} \) of component tessellations \( \Psi_n \) is invariant under arbitrary translations with \( \Psi_1, \Psi_2, \ldots \) being iid and independent of \( \Psi_0 \). If \( \Psi_0 \) or \( \Psi_1 \) is isotropic, the intensities \( \gamma^{(0)}, \gamma^{(1)} \), and \( \gamma^{(2)} \) of the point processes
\( \Psi(0), \Psi(1), \) and \( \Psi(2), \) respectively, induced by \( \Psi \) can be expressed in the following way,

\[
\begin{align*}
\gamma^{(0)} &= \gamma_0^{(0)} + p \, \gamma_1^{(0)} + \frac{4p}{\pi} \tilde{\gamma}_0^{(1)} \tilde{\gamma}_1^{(1)}, \\
\gamma^{(1)} &= \gamma_0^{(1)} + p \, \gamma_1^{(1)} + \frac{6p}{\pi} \tilde{\gamma}_0^{(1)} \tilde{\gamma}_1^{(1)}, \\
\gamma^{(2)} &= \gamma_0^{(2)} + p \, \gamma_1^{(2)} + \frac{2p}{\pi} \tilde{\gamma}_0^{(1)} \tilde{\gamma}_1^{(1)},
\end{align*}
\]  

(3.18)  
(3.19)  
(3.20)

where \( \tilde{\gamma}_0^{(1)} \) and \( \tilde{\gamma}_1^{(1)} \) denote the intensity of the process of edges induced by \( \Psi_0 \) and \( \Psi_1 \), respectively. Moreover, the intensity \( \tilde{\gamma}^{(1)} \) of the process of edges of \( \Psi \) can also be expressed in terms of \( \tilde{\gamma}_0^{(1)} \) and \( \tilde{\gamma}_1^{(1)} \) by

\[
\tilde{\gamma}^{(1)} = \frac{\tilde{\gamma}_0^{(1)}}{\gamma_0} + p \tilde{\gamma}_1^{(1)}.
\]  

(3.21)

Using the mean value relationships in (3.18) to (3.21) together with the respective mean value relationships for non-iterated stationary PVT (3.11), stationary PDT (3.13), and stationary (and isotropic) PLT (3.16), respectively, we are in the position to express the intensities \( \gamma^{(0)}, \gamma^{(1)}, \gamma^{(2)} \), and \( \tilde{\gamma}^{(1)} \) in terms of the intensity \( \gamma_0 \) of \( \Psi_0 \), the intensity \( \gamma_1 \) of \( \Psi_1 \), and the Bernoulli thinning parameter \( p \). Tables 3.1, 3.2, and 3.3, respectively, show the resulting formulae if we choose for \( \Psi_0 \) and \( \Psi_1 \) either of the stationary (and isotropic) tessellation models PVT, PDT, and PLT. Note that the intensities \( \gamma_0 \) and \( \gamma_1 \) have to be interpreted with the meaning of \( \gamma_{\text{PVT}}, \gamma_{\text{PDT}}, \) and \( \gamma_{\text{PLT}}, \) respectively.

| Table 3.1: Mean value relationships for a PVT/p \( \Psi_1 \)-nesting |
|-----------------|-----------------|-----------------|
| \( \Psi_1 = PLT \) | \( \Psi_1 = PVT \) | \( \Psi_1 = PDT \) |
| \( \gamma^{(0)} \) | \( 2\gamma_0 + \frac{2p}{\pi} \gamma_1^2 + \frac{8p}{\pi} \gamma_1 \sqrt{\gamma_0} \) | \( 2\gamma_0 + 2p \gamma_1 + \frac{16p}{\pi} \sqrt{\gamma_0} \gamma_1 \) | \( 2\gamma_0 + p \gamma_1 + \frac{256p}{3\pi^2} \sqrt{\gamma_0} \gamma_1 \) |
| \( \gamma^{(1)} \) | \( 3\gamma_0 + \frac{2p}{\pi} \gamma_1^2 + \frac{12p}{\pi} \gamma_1 \sqrt{\gamma_0} \) | \( 3\gamma_0 + 3p \gamma_1 + \frac{24p}{\pi} \sqrt{\gamma_0} \gamma_1 \) | \( 3\gamma_0 + 3p \gamma_1 + \frac{128p}{3\pi^2} \sqrt{\gamma_0} \gamma_1 \) |
| \( \gamma^{(2)} \) | \( \gamma_0 + \frac{2p}{\pi} \gamma_1^2 + \frac{4p}{\pi} \gamma_1 \sqrt{\gamma_0} \) | \( \gamma_0 + p \gamma_1 + \frac{8p}{\pi} \sqrt{\gamma_0} \gamma_1 \) | \( \gamma_0 + 2p \gamma_1 + \frac{128p}{3\pi^2} \sqrt{\gamma_0} \gamma_1 \) |
| \( \tilde{\gamma}^{(1)} \) | \( 2\sqrt{\gamma_0} + p \gamma_1 \) | \( 2\sqrt{\gamma_0} + 2p \sqrt{\gamma_1} \) | \( 2\sqrt{\gamma_0} + \frac{32p}{3\pi^2} \sqrt{\gamma_1} \) |

3.5 Asymptotic Behavior of Cells of Stationary Tessellations

In this section we present an auxiliary result that has been obtained in [51] and that is essential for the proof of Theorem 6.2 in Chapter 6. However, it is also of interest in its own right for applications in spatial statistics. In fact, the following Theorem 3.2 states that if we observe a single realization of some stationary random tessellation through an averaging sequence of convex sampling windows, the cells hitting the boundary of this sampling window, i.e., the cells
that are only partially observable, become asymptotically negligible. More precisely, assume that we observe a realization of some stationary (and ergodic) tessellation $\Psi = \{ \Xi_n \}_{n \geq 1}$ in $\mathbb{R}^d$ through a sequence $\{W^d_\varrho\}_{\varrho > 0}$ of convex compact sampling windows $W^d_\varrho \subset \mathbb{R}^d$, where the $(d$-dimensional) convex body $W \subset \mathbb{R}^d$ satisfies the inclusion $B^d_{r_1} \subset W \subset B^d_{r_2}$ for some fixed radii $r_1, r_2$ with $0 < r_1 < r_2 < \infty$. Recall that $\Psi$ may be represented alternatively as stationary (and ergodic) random marked point process $\sum_{n \geq 1} \delta_{\alpha(\Xi_n), \Xi_n}$, where $\Xi_n = \Xi_n - \alpha(\Xi_n)$ is the $n$th cell of $\Psi$ centered at the origin $\alpha$ and $\alpha(\Xi_n)$ denotes the associated point of the $n$th cell $\Xi_n$, $n \geq 1$. The intensity of $\sum_{n \geq 1} \delta_{\alpha(\Xi_n), \Xi_n}$ and hence of $\Psi$ is given by $\gamma = \mathbb{E}\#\{ n : \alpha(\Xi_n) \in [0, 1]^d \}$.

### Table 3.2: Mean value relationships for a $PDT/p\Psi_1$-nesting

<table>
<thead>
<tr>
<th>$\Psi_1 = PLT$</th>
<th>$\Psi_1 = PVT$</th>
<th>$\Psi_1 = PDT$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma^{(0)}$</td>
<td>$\gamma_0 + \frac{p}{\pi} \frac{\gamma^2}{2} + \frac{128 p^2}{3\pi^2} \gamma_0 \sqrt{\gamma_0}$</td>
<td>$\gamma_0 + 2p\gamma_1 + \frac{256 p^2}{3\pi^2} \sqrt{\gamma_0 \gamma_1}$</td>
</tr>
<tr>
<td>$\gamma^{(1)}$</td>
<td>$3\gamma_0 + \frac{2p}{\pi} \frac{\gamma^2}{2} + \frac{64 p^2}{3\pi^2} \gamma_0 \sqrt{\gamma_0}$</td>
<td>$3\gamma_0 + 3p\gamma_1 + \frac{128 p^2}{3\pi^2} \sqrt{\gamma_0 \gamma_1}$</td>
</tr>
<tr>
<td>$\gamma^{(2)}$</td>
<td>$2\gamma_0 + \frac{p}{\pi} \frac{\gamma^2}{2} + \frac{64 p^2}{3\pi^2} \gamma_0 \sqrt{\gamma_0}$</td>
<td>$2\gamma_0 + 2p\gamma_1 + \frac{2048 p^2}{3\pi^2} \sqrt{\gamma_0 \gamma_1}$</td>
</tr>
<tr>
<td>$\gamma^{(1)}$</td>
<td>$\frac{32}{3\pi^2} \sqrt{\gamma_0} + p\gamma_1$</td>
<td>$\frac{32}{3\pi^2} \sqrt{\gamma_0} + 2p\gamma_1$</td>
</tr>
</tbody>
</table>

### Table 3.3: Mean value relationships for a $PLT/p\Psi_1$-nesting

<table>
<thead>
<tr>
<th>$\Psi_1 = PLT$</th>
<th>$\Psi_1 = PVT$</th>
<th>$\Psi_1 = PDT$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma^{(0)}$</td>
<td>$\frac{1}{\pi}(\frac{\gamma^2}{2} + 4p\gamma_0 \gamma_1)$</td>
<td>$\frac{1}{\pi} \frac{\gamma^2}{2} + 2p\gamma_1 + \frac{8p}{\pi} \gamma_0 \sqrt{\gamma_1}$</td>
</tr>
<tr>
<td>$\gamma^{(1)}$</td>
<td>$\frac{1}{\pi}(2\gamma_0^2 + 2p\gamma_1^2 + 6p\gamma_0 \gamma_1)$</td>
<td>$\frac{1}{\pi} \frac{\gamma^2}{2} + 3p\gamma_1 + \frac{12p}{\pi} \gamma_0 \sqrt{\gamma_1}$</td>
</tr>
<tr>
<td>$\gamma^{(2)}$</td>
<td>$\frac{1}{\pi}(\gamma_0^2 + 2p\gamma_1^2 + 4p\gamma_0 \gamma_1)$</td>
<td>$\frac{1}{\pi} \frac{\gamma^2}{2} + 2p\gamma_1 + \frac{64p}{3\pi^2} \gamma_0 \sqrt{\gamma_1}$</td>
</tr>
<tr>
<td>$\gamma^{(1)}$</td>
<td>$\gamma_0 + p\gamma_1$</td>
<td>$\gamma_0 + 2p\gamma_1$</td>
</tr>
</tbody>
</table>

3.5.1 **Asymptotic Negligibility of Partially Observed Cells**

We define the families of events $\{ A_\varrho \}_{\varrho > \varrho_0}$ and $\{ B_\varrho \}_{\varrho > \varrho_0}$ by

$$A_\varrho = \bigcap_{n \geq 1} \left( \{ (\Xi_n^0 + \alpha(\Xi_n)) \cap W^d_\varrho = \emptyset \} \cup \{ \alpha(\Xi_n) \in W^d_{\varrho + q(\varrho)} \} \right)$$

and

$$B_\varrho = \bigcap_{n \geq 1} \left( \{ (\Xi_n^0 + \alpha(\Xi_n)) \subseteq W^d_\varrho \} \cup \{ \alpha(\Xi_n) \notin W^d_{\varrho - q(\varrho)} \} \right).$$
where \( q_0 \geq 0 \) is some constant and the function \( q : (q_0, \infty) \to (0, \infty) \) is such that \( q(\varrho) < \varrho \) for each \( \varrho > q_0 \). Note that the complement \( A^c_\varrho \) of the event \( A_\varrho \) for \( \varrho > q_0 \) is the event that at least one cell \( \Xi_n \) of \( \Psi \) exists for \( n \geq 1 \) that hits the sampling window \( W^d_\varrho \) but has its associated point located outside of \( W^d_\varrho + q(\varrho) \). Likewise, \( B^c_\varrho \) is the event that at least one cell \( \Xi_n \) of \( \Psi \) exists for \( n \geq 1 \) that is not completely contained in \( W^d_\varrho \) but has its associated point located in \( W^d_\varrho - q(\varrho) \).

**Theorem 3.2** Let \( \Psi \) be a stationary (and ergodic) tessellation in \( \mathbb{R}^d \). Under the assumption that \( \mathbb{E}D^d(\Xi^*) < \infty \), where \( D(\Xi^*) \) is the diameter of the typical cell \( \Xi^* \) of \( \Psi \), there exists a non-decreasing function \( q : (q_0, \infty) \to (0, \infty) \) for some constant \( q_0 > 0 \) satisfying \( q(\varrho) < \varrho \) for \( \varrho > q_0 \), \( q(\varrho) \to \infty \), and \( q(\varrho) / \varrho \downarrow 0 \) as \( \varrho \to \infty \) such that

\[
\lim_{\ell \to \infty} \mathbb{P} \left( \bigcup_{k \geq \ell} A^c_k \right) = 0 \quad \text{and} \quad \lim_{\ell \to \infty} \mathbb{P} \left( \bigcup_{k \geq \ell} B^c_k \right) = 0 . \tag{3.24}
\]

Before we state the proof of Theorem 3.2, we note that the properties \( q(\varrho) < \varrho \) for \( \varrho > q_0 \) and \( q(\varrho) / \varrho \downarrow 0 \) as \( \varrho \to \infty \) of the function \( \varrho \mapsto q(\varrho) \) mean that \( q(\varrho) \) increases weaker than \( \varrho \). Furthermore, consider again the interpretation of \( A^c_\varrho \) and \( B^c_\varrho \) for \( \varrho > q_0 \). Then, both statements in (3.24) can be interpreted in the sense that if we observe a realization of \( \Psi \) through \( W^d_\varrho \), it becomes less and less likely, as \( \varrho \to \infty \), that cells of \( \Psi \) occur which hit \( W^d_\varrho \) but have their associated point located outside of the enlarged sampling window \( W^d_{\varrho + q(\varrho)} \) and that cells of \( \Psi \) occur which hit \( W^d_\varrho \) but have their associated point located in the shrunk sampling window \( W^d_{\varrho - q(\varrho)} \). Hence, for a fixed sampling window \( W^d_\varrho \), Theorem 3.2 states that the influence of cells of \( \Psi \) which are only partially observable is negligible if \( W^d_\varrho \) is (suitably) large.

**Proof** To begin with we recall that the integrability of \( D^d(\Xi^*) \) implies the existence of a convex function \( H : [0, \infty) \to [0, \infty) \) strictly increasing on its support \( (x_0, \infty) \) (for some \( x_0 \geq 0 \)) such that \( H(x)/x \) is non-decreasing for \( x > 0 \) with \( \lim_{x \to \infty} H(x)/x = \infty \) and \( \mathbb{E}H(D^d(\Xi^*)) < \infty \); cf. for example Theorem II.22 in [30]. Then, we define the function \( \varrho \mapsto q(\varrho) \) implicitly, i.e., for some fixed \( r > 0 \) such that \( B^d_r \subseteq W \), the value \( q(\varrho) ( > x_0 r^d / r ) \) is given as the unique solution of the equation

\[
q^d = H(r^d q^d(\varrho)) \quad \text{for any} \quad \varrho > 0 , \tag{3.25}
\]

where \( r^d q^d(\varrho) > x_0 \). Due to (3.25) and the properties of the function \( H \), it is easily checked that the function \( \varrho \mapsto q(\varrho) \) possesses the required properties for \( \varrho > q_0 = \inf \{ x > 0 : H(r^d q^d(x)) > q^d(x) \} \). Note that \( B^d_r \subseteq W \) implies that \( B^d_{r q(\varrho)} \subseteq W^d_{q(\varrho)} \), which, together with the convexity of \( W \), yields that

\[
W^d_\varrho \oplus B^d_{r q(\varrho)} \subseteq W^d_\varrho \oplus W^d_{q(\varrho)} \subseteq W^d_{\varrho + q(\varrho)} . \tag{3.26}
\]

Regarding the latter set relations in (3.26) we immediately see that for any \( C \in \mathcal{C} \) with \( o \in C \), we have \( D(C) > r q(\varrho) \) if \( (W^d_{\varrho + q(\varrho)})^c \cap (W^d_\varrho \oplus C) \neq \emptyset \). Together with the definition of \( A_\varrho \) in
(3.22) and, hence, by reconsideration of $A_k$, we get that
\[
P\left( \bigcup_{k \geq \ell} A_k^c \right) = P\left( \bigcup_{n \geq 1} \bigcup_{k \geq \ell} \{ \alpha(\Xi_n) \in (W_k^{d+q(k)})^c \cap (W_k^{d} \oplus \Xi_n^c) \} \right),
\]
where $\Xi_n^c = \{ -x : x \in \Xi_n \}$ denotes the reflection of the cell $\Xi_n$ containing the origin $o$. The latter probability can be estimated immediately in view of the implication $D(\Xi_n^c) > r(q(k))$ stated above. Therefore,
\[
P\left( \bigcup_{k \geq \ell} A_k^c \right) \leq P\left( \bigcup_{n \geq 1} \bigcup_{k \geq \ell} \left( \{ \alpha(\Xi_n) \in W_k^{d} \oplus B^d_D(\Xi_n^c) \} \cap \{ D(\Xi_n^c) > r(q(k)) \} \right) \right)
\]
\[
= P\left( \bigcup_{n \geq 1} \bigcup_{k \geq \ell} \left( \{ \alpha(\Xi_n) \in W_k^{d} \oplus B^d_D(\Xi_n^c) \} \cap \{ r(q(k)) < D(\Xi_n^c) \leq r(q(k+1)) \} \right) \right).
\]
The last equality becomes clear by the fact that for any two sequences of events $\{E_k\}_{k \geq \ell}$ and $\{E'_k\}_{k \geq \ell}$ with $E_k \subseteq E_{k+1}$ and $E'_k \supseteq E'_{k+1}$ for $k \geq \ell$, the identity
\[
\bigcup_{k \geq \ell} E_k \cap E'_k = \bigcup_{k \geq \ell} E_k \cap (E'_k \setminus E_{k+1})
\]
holds, where in our situation we may choose $E_k = \{ \alpha(\Xi_n) \in W_k^{d} \oplus B^d_D(\Xi_n^c) \}$ and $E'_k = \{ D(\Xi_n^c) > r(q(k)) \}$. Hence, the subadditivity of $P$ and the inequality $P(Y \geq 1) \leq E[Y]$, which holds for any random variable $Y$, imply the following two estimates, i.e.,
\[
P\left( \bigcup_{k \geq \ell} A_k^c \right) \leq \sum_{k \geq \ell} P\left( \sum_{n \geq 1} \mathbb{I}_{W_k^d \oplus B^d_D(\Xi_n^c)}(\alpha(\Xi_n)) \mathbb{I}_{(r(q(k),r(q(k+1)))}(D(\Xi_n^c)) \geq 1) \right)
\]
\[
\leq \sum_{k \geq \ell} P\left( \sum_{n \geq 1} \mathbb{I}_{W_k^d \oplus B^d_D(\Xi_n^c)}(\alpha(\Xi_n)) \mathbb{I}_{(r(q(k),r(q(k+1))}(D(\Xi_n^c)) \right).
\]
Applying Campbell’s theorem (cf. Theorem 2.6) for stationary marked point processes together with Fubini’s theorem, we get that the latter estimate is equal to
\[
\gamma \sum_{k \geq \ell} \int_{P^o} \int_{\mathbb{R}^d} \mathbb{I}_{W_k^d \oplus B^d_D(C)}(x) \mathbb{I}_{(r(q(k),r(q(k+1))}(D(C)) \, dx \, P^o(dC)
\]
\[
= \gamma \sum_{k \geq \ell} \int_{P^o} \nu_d(W_k^d \oplus B^d_D(C)) \mathbb{I}_{(r(q(k),r(q(k+1))}(D(C)) \, P^o(dC),
\]
where the latter equality is obtained since, obviously,
\[
\int_{\mathbb{R}^d} \mathbb{I}_{W_k^d \oplus B^d_D(C)}(x) \, dx = \nu_d(W_k^d \oplus B^d_D(C)), \quad C \in \mathcal{C}.
\]
Also, since $W^d_k = k W \in C$, we get by Steiner’s formula, cf. Formula (A.20), together with the homogeneity relation $V_s(W^d_k) = V_s(k W) = k^s V_s(W)$, cf. Formula (A.23), for the intrinsic volumes $V_s$, $s = 0, \ldots , d$ that

$$
\nu_d\left(W^d_k \oplus B^d_{D(C)}\right) = \sum_{s=0}^{d} \kappa_{d-s} D^{d-s}(C) V_s(W^d_k) = \sum_{s=0}^{d} k^s \kappa_{d-s} D^{d-s}(C) V_s(W)
$$

for $C \in C$. Thus, we may write

$$
P\left( \bigcup_{k \geq \ell} A^c_k \right) \leq \gamma \sum_{k \geq \ell} \sum_{s=0}^{d} \kappa_{d-s} k^s V_s(W) \mathbb{E}\left( (D^{d-s}(\Xi^s)) \mathbb{I}_{(r q(k), r q(k+1))}(D(\Xi^s)) \right)$$

$$
\leq \gamma \sum_{k \geq \ell} \sum_{s=0}^{d} \kappa_{d-s} V_s(W) \frac{k^s r^{d-s} q^{d-s}(k+1)}{H(r^d q^d(k))} \times \mathbb{E}\left( H(D^{d}(\Xi^s)) \mathbb{I}_{(r q(k), r q(k+1))}(D(\Xi^s)) \right),
$$

where the latter estimate is obtained due to the properties of the function $H$, especially due to its monotonicity, i.e., $H(D^{d-s}(\Xi^s)) \leq H(D^{d}(\Xi^s))$ for $s = 0, \ldots , d$. By (3.25), we have $H(r^d q^d(k)) = k^d$ and $q(k+1) \leq k+1$ for all $k \geq \ell$ (with $\ell$ large enough) and therefore,

$$
\frac{k^s r^{d-s} q^{d-s}(k+1)}{H(r^d q^d(k))} \leq \frac{k^s r^{d-s} (k+1)^{d-s}}{k^d} = \left(1 + \frac{1}{k}\right)^{d-s} r^{d-s} \leq 2^{d-s} r^{d-s} \quad \text{if} \quad k \geq \ell.
$$

This yields that

$$
P\left( \bigcup_{k \geq \ell} A^c_k \right) \leq \gamma \sum_{s=0}^{d} \kappa_{d-s} (2r)^{d-s} V_s(W) \mathbb{E}(H(D^{d}(\Xi^s)) \mathbb{I}_{(r q(\ell), \infty)}(D(\Xi^s))) \to 0 \quad \ell \to \infty.
$$

To show the second assertion in (3.24) we first realize that

$$
P\left( \bigcup_{k \geq \ell} B^c_k \right) = P\left( \bigcup_{n \geq 1} \bigcup_{k \geq \ell} \{ \alpha(\Xi_n) \in W^d_{k-q(k)} \cap (W^d_k \oplus \Xi_n) \} \right).
$$

In analogy to the considerations in (3.26) above, we can show that

$$
(W^d_\theta)^c \oplus B^d_{q(\theta)} \subseteq (W^d_\theta)^c \oplus W^d_{q(\theta)} = (W^d_\theta \oplus W^d_{q(\theta)})^c = (W^d_{\theta-q(\theta)})^c,
$$

and hence, for any $C \in C$ with $o \in C$, we have $D(C) > r q(\theta)$ if $((W^d_\theta)^c \oplus C) \cap W^d_{\theta-q(\theta)} \neq \emptyset$. Together with $W^d_{k-q(k)} \subseteq W^d_k$, this gives that

$$
P\left( \bigcup_{k \geq \ell} B^c_k \right) \leq P\left( \bigcup_{n \geq 1} \bigcup_{k \geq \ell} \{ \alpha(\Xi_n) \in W^d_k \cap \{ D(\Xi_n) > r q(k) \} \} \right).
$$

Since the right hand side of the latter inequality is not larger than the bound on the right-hand side of (3.27), we immediately get that $P\left( \bigcup_{k \geq \ell} B^c_k \right) \to 0$. \hfill \qed
3.5.2 A View on the Assumption of Theorem 3.2

In this section, we restrict ourselves to the case $d = 2$, which is relevant for most applications, and consider concrete examples of stationary (and ergodic) tessellations $\Psi = \{\Xi_n\}_{n \geq 1}$ which fulfill the condition mentioned in Theorem 3.2. More precisely, we consider tessellations $\Psi$ in $\mathbb{R}^2$, such that $\mathbb{E}D^2(\Xi^*) < \infty$. In fact, we note that all moments of $D(\Xi^*)$ exist if the distribution of $D(\Xi^*)$ has an exponentially bounded tail.

**Theorem 3.3** Let $\Psi = \{\Xi_n\}_{n \in \mathbb{N}}$ be a stationary tessellation in $\mathbb{R}^2$ with typical cell $\Xi^*$.

(i) If $\Psi$ is a PLT, which is additionally isotropic, then it holds for all $x > 0$ that

$$
\mathbb{P}(D(\Xi^*) > x) \leq \pi x e^{-x} \left(1 - \frac{\pi - 2}{2} e^{-x} + \frac{1}{6} (\pi - 3)^2 x^2 e^{-2x} + \frac{1}{\pi} e^{-(\pi - 1)x}\right). \quad (3.28)
$$

(ii) If $\Psi$ is a PVT, then it holds for all $x > 0$ that

$$
\mathbb{P}(D(\Xi^*) > x) \leq \frac{\pi}{2} x^2 e^{-\pi x^2} \left(2 - \frac{\pi}{2} x^2 e^{-\pi x^2} + \frac{(\pi x^2)^2}{48} e^{-\frac{\pi x^2}{2}} + \frac{2}{\pi x^2} e^{-\frac{\pi x^2}{4}}\right). \quad (3.29)
$$

**Proof** In order to show both (3.28) and (3.29), we make use of results derived earlier in [24]. Let $\Psi$ be an (isotropic) PLT. Then, the bound for $\mathbb{P}(D(\Xi^*) > x)$ in (3.28) can easily be deduced from Theorem 8 in [24], where an upper bound is given for the tail probability $\mathbb{P}(R(\Xi^*) > r)$, $r > 0$, of the radius $R(\Xi^*)$ of the minimal ball circumscribing $\Xi^*$. In fact, Theorem 8 in [24] states that

$$
\mathbb{P}(R(\Xi^*) > r) \leq 2 \pi r e^{-2r} \left(1 - (\pi - 2) e^{-2r} + \frac{2}{3} (\pi - 3)^2 r^2 e^{-4r} + \frac{1}{2 \pi r} e^{-2(\pi - 1)r}\right)
$$

for any $r > 0$. Hence, for $x > 0$, we immediately get that

$$
\mathbb{P}(D(\Xi^*) > x) \leq \mathbb{P}(R(\Xi^*) > \frac{x}{2}),
$$

and therefore (3.28) is proven. Likewise, (3.29) can be deduced in the same manner from Theorem 3 in [24].

We conclude with the remark that in [24] even the exact distribution of $D(\Xi^*)$ is obtained both for the Poisson-Voronoi and the Poisson line case. Explicit values of the distribution function however have then to be obtained using numerical integration. Moreover, note that for the Poisson-Voronoi case a weaker upper bound than the one in (3.29) has already been obtained in [36].
Chapter 4

Limit Theorems for Poisson Hyperplane Processes

We study the asymptotic distributional behavior of flat processes induced by the intersection of hyperplanes of a stationary, but not necessarily isotropic, Poisson hyperplane process. The presented material is based partly on results obtained in [52].

Figure 4.1: Poisson lines hitting $B^2_r$ and their intersections

More precisely, we begin with our presentation of central limit theorems (CLTs) in Section 4.1, where we study the limiting distributional behavior of the (random) number $n_0(B^d_r)$ of nodes within a $d$-dimensional ball $B^d_r$ centered at the origin as the radius $r > 0$ tends to infinity. The nodes are assumed to be induced by intersections of $d$ hyperplanes of a stationary, however not necessarily isotropic, Poisson hyperplane process in $\mathbb{R}^d$. Figure 4.1 displays Poisson lines hitting the ball $B^2_r$ and their induced intersections in the planar case $d = 2$. Additionally, restraining conditions on the orientation vectors of the intersecting hyperplanes are imposed.
The idea behind the derivation of a CLT for the number \( \eta_0(B^d_t) \) of intersecting nodes in \( B^d_t \) is that \( \eta_0(B^d_t) \) can be expressed in the form of a multiple (random) sum, where the summation is done over all \( d \)-tuples of distinct hyperplanes having a common point in \( B^d_t \); Section 4.1.1. Noting that this sum can be expressed through a \( U \)-statistic, conditional that the number of hyperplanes meeting \( B^d_t \) is fixed, we apply Hoeffding’s decomposition in order to get an alternative expression for \( \eta_0(B^d_t) \). By proving a CLT for sums over a Poisson distributed number of iid random variables, we obtain that the (suitably centered and normalized) number \( \eta_0(B^d_t) \) of intersection nodes follows asymptotically a normal distribution with mean zero; Section 4.1.2.

The CLT in Section 4.1 is inspired by similar results in the planar case \( d = 2 \) for Poisson line processes obtained earlier by K. Paroux, however with random normalization; [94]. In order to give a short comparison of our CLT with the one in [94], Section 4.1.3 is devoted to a review of the planar case. Applying Hoeffding’s CLT, cf. Formula (A.19), for \( U \)-statistics (with random normalization) instead of the method of moments as in [94], we are able to simplify the proof of K. Paroux considerably.

Section 4.2 expands our CLT for \( \eta_0(B^d_t) \) in certain directions. In particular, in Section 4.2.1 we consider CLTs for the number as well as the total \( k \)-volume of \( k \)-flats \((k = 0, \ldots, d-1) \) hitting \( B^d_t \). The \( k \)-flats are assumed to be induced by the intersection of \( d-k \) (distinct) hyperplanes belonging, once again, to a stationary, not necessarily isotropic, Poisson hyperplanes process in \( \mathbb{R}^d \). If additionally we assume isotropy and impose no further restrictions on the orientation vectors of the intersecting hyperplanes, we are able to derive explicit formulae for intensities and asymptotic variances of the considered random variables; Section 4.2.2.

Moreover, in Section 4.3, we study the joint asymptotic distributional behavior of \( d \)-dimensional vectors with entries being either the number of \( k \)-flats hitting \( B^d_t \) or their total \( k \)-volume in \( B^d_t \); \( k = 0, \ldots, d-1 \).

We note that the normalization in our CLTs is done, up to some constants, with respect to the \( d \)-dimensional volume of \( B^d_t \) raised to the power \( 1 - 1/(2d) \). This rather unexpected normalization expresses the presence of long range dependence effects, inherent to the hyperplane process. Furthermore, note that the choice of a spherical sampling window \( B^d_t \) simplifies the proofs considerably. However, most of the results remain valid for more general classes of unboundedly expanding convex sampling windows. In fact, for example [109] studies random hyperplanes meeting a convex body. Finally, we mention that the results of this chapter, together with the asymptotic boundary behavior studied in Section 3.5, allow to derive similar CLTs for \( k \)-facets of Poisson hyperplane tessellations; \( k = 1, \ldots, d \).

The CLTs in this chapter are the basis for statistical inference of Poisson hyperplane processes as well as Poisson hyperplane tessellations. We refer to Chapter 5, where we will derive asymptotic confidence intervals and asymptotic tests based on the results of the present chapter.
4.1 CLTs for Intersection Nodes of Poisson Hyperplanes

Recall from Section 2.5 that a stationary \((d-1)\)-flat process \(\Phi_{d-1}\) in \(\mathbb{R}^d\), i.e., a stationary hyperplane process, can be represented by an (independently) marked point process \(\sum_{i \geq 1} \delta_{(P_i, U_i)}\) on the real line, where \(P_i\) is the signed perpendicular distance from the origin and \(U_i\) is the orientation vector of a (random) hyperplane \(H(P_i, U_i)\) for \(i \geq 1\). In particular, in case of a Poisson hyperplane process \(\Phi_{d-1}\) we get the independently marked Poisson process \(\Phi'_{d-1} = \sum_{i \geq 1} \delta_{(P_i, U_i)}\). The \(P_i\)'s form a Poisson process with intensity \(\lambda\) on the real line, where \(\lambda\) coincides with the intensity \(\lambda_{d-1}\) of \(\Phi_{d-1}\); Section 2.5.3. Recall furthermore that \(\Phi_{d-1}\) is called non-degenerate if its spherical orientation distribution \(\tilde{\Theta}\) is non-degenerate, i.e., \(\tilde{\Theta}(H(0, u) \cap S^{d-1}_+) < 1\) for any \(u \in S^{d-1}_+\), where \(S^{d-1}_+\) denotes the \((d-1)\)-dimensional upper unit hemisphere. This non-degeneracy assumption assures that each of the (stationary) \(k\)-flat processes \(\Phi_k\) that can be generated by intersections of \(d-k\) hyperplanes from the (stationary) process \(\Phi_{d-1}\) has positive intensity \(\lambda_k\) for \(k = 0, \ldots, d-1\) and that the Poisson hyperplane tessellation induced by \(\Phi_{d-1}\) consists of bounded cells \(\mathbb{P}\text{-a.s.};\) compare with Chapter 6 of [112].

4.1.1 Point Processes of Intersection Nodes

We examine, for any \(r > 0\) and \(B \in B((S^{d-1}_+)^d)\), the number \(\eta_0(B^d_r \times B)\) of those intersection points \(H(P_{i1}, U_{i1}) \cap \ldots \cap H(P_{id}, U_{id})\) in \(B^d_r\) for which the corresponding orientation vectors \(U_{i1}, \ldots, U_{id}\) satisfy the condition \((U_{i1}, \ldots, U_{id}) \in B\). In this context, \((U_{i1}, \ldots, U_{id})\) is the reordered vector \((U_{i1}, \ldots, U_{id})\) such that \(U_{(n_1)} \leq \ldots \leq U_{(n_d)}\), according to an appropriate linear ordering \(\preceq\) in \(S^{d-1}_+\).

Asymptotic normality of \(\eta_0(B^d_r \times B)\) as \(r \to \infty\) is based on the following fundamental property of the Poisson process \(\Phi'_{d-1}\). Given the number \(N_r = \Phi'_{d-1}([-r, r] \times S^{d-1}_+)\) of hyperplanes hitting the sampling window \(B^d_r\), say \(N_r = n\) for \(n \in \mathbb{N}\), the random vectors \((P_i, U_i), i = 1, \ldots, n,\) are independent with independent components, where the \(P_i\)'s are (conditionally) uniformly distributed on \([-r, r]\) and where \(U_i\ has the distribution \(\tilde{\Theta}\). Moreover, \(N_r\ is a Poisson distributed random variable with mean \(2\lambda r,\) since the \(P_i\)'s form a Poisson process on the real line with mean \(\lambda;\) cf. also the interpretation (2.6) of \(\lambda_{d-1}\).

Hence, we get that

\[
\eta_0(B^d_r \times B) \overset{d}{=} \frac{1}{d!} \sum_{1 \leq i_1, \ldots, i_d \leq N_r} f_B(X_{i_1}, \ldots, X_{i_d}),
\]

(4.1)

where \(\overset{d}{=}\) means equality in distribution and where \(X_i = (Q_i, U_i), i \geq 1,\) are independent random vectors, independent of \(N_r,\) and with independent components such that \(Q_i\) is uniformly distributed on \([-r, r]\) and \(U_i\ has the distribution \(\tilde{\Theta}\). The sum \(\sum\) runs over pairwise distinct indices of the arguments of the function \(f_B : (\mathbb{R} \times S^{d-1}_+)^d \to \{0, 1\},\) which in turn is defined by

\[
f_B((p_1, u_1), \ldots, (p_d, u_d)) = \chi(H(p_1, u_1) \cap \ldots \cap H(p_d, u_d) \cap B^d_r) I_B(u_1, \ldots, u_d),
\]

(4.2)
where \( \chi(K) = 1 \) for \( K \neq \emptyset \) and \( \chi(\emptyset) = 0 \). Since \( f_B \) is measurable and symmetric, i.e.,
\[
f_B((p_1, u_1), \ldots, (p_d, u_d)) = f_B((p_i(1), u_i(1)), \ldots, (p_i(d), u_i(d)))
\]
for any permutation \( \iota : \{1, \ldots, d\} \to \{1, \ldots, d\} \), the right hand side of (4.1) divided by \( \binom{N_r}{d} \)
and conditioned on \( N_r = n \), is a \( U \)-statistic of order \( d \) with kernel function \( f_B \); Section A.3.7.

To simplify notation we put
\[
\sigma^{(j,d)}_B = \mathbb{E}(f_B(X_1, \ldots, X_d) f_B(X_{d-j+1}, \ldots, X_{2d-j}))
\]
\[
= \mathbb{E}g_B^2(X_{d-j+1}, \ldots, X_d),
\]
(4.3)

where \( g_B((p_1, u_1), \ldots, (p_j, u_j)) = f_B((p_j, u_j), X_{j+1}, \ldots, X_d) \) for \( j = 1, \ldots, d \).

Note that due to the uniform distribution of \( Q_i, i \geq 1 \), on \([-r, r]\) both the expectations
\( \mathbb{E}f_B(X_1, \ldots, X_d) \) and the second moments \( \sigma^{(j,d)}_B \) do not depend on \( r > 0 \).

**Lemma 4.1** For any \( B \in \mathcal{B}((S^d_+)^d) \),
\[
\mathbb{E}\eta_0(B^d_r \times B) = \frac{(2\lambda r)^d}{d!} \mathbb{E}f_B(X_1, \ldots, X_d)
\]
and
\[
\operatorname{Var} \eta_0(B^d_r \times B) = \sum_{j=1}^{d} \frac{(2\lambda r)^{2d-j}}{j!((d-j)!)^2} \sigma^{(j,d)}_B .
\]
(4.5)

**Proof** We easily see from (4.1) that
\[
\mathbb{E}\eta_0(B^d_r \times B) = \frac{1}{d!} \mathbb{E} \sum_{1 \leq i_1, \ldots, i_d \leq N_r} f_B(X_{i_1}, \ldots, X_{i_d}),
\]
where within the mean on the right hand side of the latter equation we have a random sum which has \( N_r(N_r - 1) \ldots (N_r - d + 1) \) summands. Thus, by the symmetry of the function \( f_B \)
defined in (4.2) combined with the independence between \( N_r \) and the iid sequence consisting of \( X_i = (Q_i, U_i) \) for \( i \geq 1 \), we get that
\[
\mathbb{E}\eta_0(B^d_r \times B) = \frac{1}{d!} \mathbb{E} \left( N_r(N_r - 1) \ldots (N_r - d + 1) \right) f_B(X_1, \ldots, X_d).
\]

Note that the \( d \)-th factorial moment of a Poisson distributed random variable is equal to the \( d \)-th power of its mean, a fact which may be easily seen by regarding the (probability) generating function in the Poisson case and the resulting derivatives. Thus
\[
\mathbb{E}(N_r(N_r - 1) \ldots (N_r - d + 1)) = (\mathbb{E}N_r)^d = (2\lambda r)^d,
\]
(4.6)
which completes the proof of (4.4). To derive a formula for the variance \( \text{Var} \eta_0(B_r^d \times B) \) we again utilize the symmetry of \( f_B \) and employ simple combinatorial arguments which lead to

\[
\mathbb{E}\left( \eta_0(B_r^d \times B) \right)^2 = \frac{1}{(d!)^2} \mathbb{E}\left( \sum_{1 \leq i_1, \ldots, i_d \leq N_r}^* f_B(X_{i_1}, \ldots, X_{i_d}) \right)^2
\]

\[
= \sum_{j=0}^d \frac{j!}{(d!)^2} \binom{d}{j} \binom{d}{j} \times \mathbb{E}\left( \sum_{1 \leq i_1, \ldots, i_{2d-j} \leq N_r}^* f_B(X_{i_1}, \ldots, X_{i_d}) f_B(X_{i_d-j+1}, \ldots, X_{i_{2d-j}}) \right)
\]

\[
= \sum_{j=0}^d \frac{\mathbb{E}(N_r(N_r - 1) \ldots (N_r - 2d + j + 1))}{j!((d-j)!)^2} \times \mathbb{E}(f_B(X_1, \ldots, X_d) f_B(X_{d-j+1}, \ldots, X_{2d-j})) .
\]

Finally, applying (4.6) with \( d \) replaced by \( 2d - j \) for \( j = 0, 1, \ldots, d \), and noting that the summand for \( j = 0 \) in the last line coincides with \( (\mathbb{E} \eta_0(B_r^d \times B))^2 \), we obtain (4.5), which completes the proof. \( \square \)

As an immediate consequence of Lemma 4.1 we get the following corollary.

**Corollary 4.2** For any \( B \in \mathcal{B}(\mathbb{S}^{d-1}_+) \),

\[
\lim_{r \to \infty} \frac{\text{Var} \eta_0(B_r^d \times B)}{r^{2d-1}} = \frac{(2\lambda)^{2d-1}}{((d-1)!)^2} \sigma_B^{(1,d)} .
\]  

(4.7)

**Proof** Since \( \sigma_B^{(j,d)} \) is independent of \( r > 0 \) for \( j = 1, \ldots, d \), we immediately obtain the relation in (4.7) by noting that

\[
\lim_{r \to \infty} \frac{\text{Var} \eta_0(B_r^d \times B)}{r^{2d-1}} = \lim_{r \to \infty} \sum_{j=1}^d \frac{(2\lambda)^{2d-j}}{j!((d-j)!)^2} r^{1-j} \sigma_B^{(j,d)},
\]

where we applied (4.5) to get the latter relation. \( \square \)

An alternative formulation of the limit relation (4.7) is given by

\[
\lim_{r \to \infty} \frac{\text{Var} \eta_0(B_r^d \times B)}{(\nu_d(B_r^d))^{2-1/d}} = \frac{(2\lambda)^{2d-1}}{((d-1)!)^2} \frac{\nu_d(B_r^d)^{1/d}}{\nu_d(B_r^d)^{2-1/d}} \sigma_B^{(1,d)} ,
\]  

(4.8)

where (4.8) involves the volume \( \nu_d(B_r^d) \) of the spherical sampling region \( B_r^d \) raised to the power \( 2 - 1/d \). Compared to the normalization in (4.7), this version of the overnormalization expresses more clearly the long range dependence inherent in a Poisson hyperplane process.
4.1.2 CLTs for Marked Intersections of $d$ Hyperplanes

Now we formulate and prove a CLT for the number $\eta_0(B^d_r \times B)$ of marked intersection points as $B^d_r$ expands undoundedly, i.e., as $r \uparrow \infty$. The centering and normalizing constants which are needed are derived in Lemma 4.1 and in Corollary 4.2, respectively.

**Theorem 4.1** Let $B \in B((S^{d-1}_+)^d)$ be chosen such that $\sigma^{(1,d)}_B > 0$. Then,

$$
\frac{(d-1)!}{(2 \lambda r)^{d-1/2}} \left( \eta_0(B^d_r \times B) - \frac{(2 \lambda r)^d}{d!} \mathbb{E} f_B(X_1, \ldots, X_d) \right) \xrightarrow{r \to \infty} \mathcal{N} \left( 0, \sigma^{(1,d)}_B \right),
$$

where $\mathcal{N}(0, \sigma^{(1,d)}_B)$ is a Gaussian random variable with mean zero and variance $\sigma^{(1,d)}_B$.

**Proof** The equality in distribution expressed in relation (4.1) is equivalent to the equality in distribution $\eta_0(B^d_r \times B) \overset{d}{=} \frac{N_r}{d} U^{(d)}_{N_r}(f_B)$ with

$$
U^{(d)}_{N_r}(f_B) = \left( \binom{N_r}{d} \right)^{-1} \frac{1}{d!} \sum_{1 \leq i_1, \ldots, i_d \leq N_r} f_B(X_{i_1}, \ldots, X_{i_d}),
$$

(4.9)

where for any fixed $N_r = n \geq d$ the random multiple sum $U^{(d)}_{N_r}(f_B)$ is a $U$-statistic as defined in (A.15), Section A.3.7, of order $d$ with kernel function $f_B$ defined in (4.2). For brevity, let $n_r$ denote the expected value $\mathbb{E} N_r = 2 \lambda r$ and let $\mu_B = \mathbb{E} f_B(X_1, \ldots, X_d)$. Then, Hoeffding’s decomposition, cf. Formula (A.16), together with (4.4) yields that

$$
\eta_0(B^d_r \times B) - \frac{(2 \lambda r)^d}{d!} \mu_B \overset{d}{=} \frac{N_r}{d} U^{(d)}_{N_r}(f_B) - \frac{n_r^d}{d!} \mu_B
$$

$$
= \left( \binom{N_r}{d} - \frac{n_r^d}{d!} \right) \mu_B + \frac{N_r}{d} \sum_{i=1}^{N_r} \left( g_B(X_i) - \mu_B \right) + \left( \binom{N_r}{d} \right) R^{(d)}_{N_r}(f_B).
$$

(4.9)

Note that the latter expression, after some simple manipulations, can be written alternatively as

$$
\left( \frac{N_r}{d} - N_r \left( \frac{N_r - 1}{d - 1} \right) + n_r \left( \frac{N_r - 1}{d - 1} \right) - \frac{n_r^d}{d!} \right) \mu_B + \frac{N_r}{d} R^{(d)}_{N_r}(f_B)
$$

$$
+ \left( \frac{N_r - 1}{d - 1} \right) \left( \sum_{i=1}^{N_r} g_B(X_i) - n_r \mu_B \right),
$$

where $g_B(x) = \mathbb{E} f_B(x, X_2, \ldots, X_d)$ for $x \in [-r, r] \times S^{d-1}_+$. Thus,

$$
\eta_0(B^d_r \times B) - \frac{(2 \lambda r)^d}{d!} \mu_B \overset{d}{=} \left( \binom{N_r}{d} - N_r \left( \frac{N_r - 1}{d - 1} \right) + n_r \left( \frac{N_r - 1}{d - 1} \right) - \frac{n_r^d}{d!} \right) \mu_B
$$

$$
+ \left( \binom{N_r}{d} \right) R^{(d)}_{N_r}(f_B) + \left( \frac{N_r - 1}{d - 1} \right) \left( \sum_{i=1}^{N_r} g_B(X_i) - n_r \mu_B \right).
$$

(4.10)
4.1. CLTs for Intersection Nodes of Poisson Hyperplanes

In the subsequently formulated and proven Lemmas 4.3 to 4.6, we will show that

\[(i) \quad \frac{1}{r^{d-1/2}} \left( \frac{N_r}{d} \right) R_{N_r}^{(d)}(f_B) \xrightarrow{r \to \infty} 0, \]

\[(ii) \quad \frac{1}{r^{d-1/2}} \left( \left[ \frac{N_r}{d} \right] - N_r \left[ \frac{N_r - 1}{d - 1} \right] + n_r \left( \frac{N_r - 1}{d - 1} \right) - \frac{n_r^d}{d!} \right) \xrightarrow{r \to \infty} 0, \]

\[(iii) \quad \frac{1}{r^{d-1}} \left( \frac{N_r - 1}{d - 1} \right) \xrightarrow{r \to \infty} \frac{(2\lambda)^{d-1}}{(d-1)!}, \text{ and} \]

\[(iv) \quad \frac{1}{\sqrt{2\lambda r}} \left( \sum_{i=1}^{N_r} g_B(X_i) - 2\lambda r E f_B(X_1, \ldots, X_d) \right) \xrightarrow{r \to \infty} \mathcal{N}(0, \sigma_B^{(1,d)}). \]

Hence, we complete the proof of Theorem 4.1 in view of the statements (i) to (iv) and by application of Slutsky’s theorem (cf. Theorem A.8) to the right hand side of (4.10).

\[\Box\]

**Lemma 4.3** Under the conditions of Theorem 4.1,

\[
\frac{1}{r^{d-1/2}} \left( \frac{N_r}{d} \right) R_{N_r}^{(d)}(f_B) \xrightarrow{r \to \infty} 0. \tag{4.11}
\]

**Proof** Since $N_r$ is independent of $X_1, X_2, \ldots$, the estimate (A.18), cf. Section A.3.7, for the remainder $R_{N_r}^{(d)}(f_B)$ of the (random) $U$-statistic considered in (4.9) implies that

\[
\mathbb{E}\left( \left( R_{N_r}^{(d)}(f_B) \right)^2 \mid N_r = n \right) \leq \frac{c_d}{n^2} \mathbb{E}f_B^2(X_1, \ldots, X_d) \quad \text{for} \quad n \geq d.
\]

Hence, we conclude that

\[
\mathbb{E}\left( \left( \frac{N_r}{d} \right) R_{N_r}^{(d)}(f_B) \right)^2 = \sum_{n \geq d} \binom{n}{d}^2 \mathbb{E}\left( \left( R_{N_r}^{(d)}(f_B) \right)^2 \mid N_r = n \right) \mathbb{P}(N_r = n)
\]

\[
\leq c_d \mathbb{E}f_B^2(X_1, \ldots, X_d) \sum_{n \geq d} \binom{n}{d}^2 \frac{1}{n^2} \mathbb{P}(N_r = n).
\]

Since \(\binom{n}{d}^2 = \binom{n}{d} \frac{(n-d)!}{(n-1)!}^2\), we obtain the inequality

\[
\sum_{n \geq d} \binom{n}{d}^2 \frac{1}{n^2} \mathbb{P}(N_r = n) \leq \frac{1}{d^2} \sum_{n \geq d} \binom{n-1}{d-1}^2 \mathbb{P}(N_r = n),
\]

which finally leads to

\[
\mathbb{E}\left( \left( \frac{N_r}{d} \right) R_{N_r}^{(d)}(f_B) \right)^2 \leq \frac{c_d \mathbb{E}f_B^2(X_1, \ldots, X_d)}{d^2} \mathbb{E}\left( \left( \frac{N_r - 1}{d - 1} \right) \right)^2. \tag{4.12}
\]
To determine the second moment of the binomial coefficient \( \binom{N_r - 1}{d - 1} \) appearing on the right hand side of the estimate in (4.12) we use the expansions

\[
x(x - 1) \cdots (x - k + 1) = \sum_{j=1}^{k} s_{j,k}^{(1)} x^j \quad \text{and} \quad x^k = \sum_{j=1}^{k} s_{j,k}^{(2)} x(x - 1) \cdots (x - j + 1),
\]

where \( s_{j,k}^{(1)} \) and \( s_{j,k}^{(2)} \) denote the Stirling numbers of the first and second kind, respectively; cf. [99] for example. Combining (4.6) with the second one of the latter expansions, it is easily seen that \( \mathbb{E} N_r^k \) is a polynomial of degree \( k \in \mathbb{N} \) in \( n_r \) and that furthermore, \( \mathbb{E} ((N_r - 1)(N_r - 2) \cdots (N_r - d + 1))^2 \) can be expressed as a polynomial of degree \( 2d - 2 \) in \( n_r \). Therefore we immediately get that

\[
\frac{1}{r^{2d-1}} \mathbb{E} \left( \binom{N_r - 1}{d - 1} \right)^2 \xrightarrow{r \to \infty} 0,
\]

which, due to the estimate in (4.12), implies (4.11).

**Lemma 4.4** Under the conditions of Theorem 4.1,

\[
\frac{1}{r^{d-1/2}} \left( \binom{N_r}{d} - N_r \binom{N_r - 1}{d - 1} + n_r \binom{N_r - 1}{d - 1} - \frac{n_r^d}{d!} \right) \xrightarrow{r \to \infty} 0,
\]

where \( n_r \) denotes the expectation \( \mathbb{E} N_r = 2 \lambda r \).

**Proof** Note that the term

\[
\binom{N_r}{d} - N_r \binom{N_r - 1}{d - 1} + n_r \binom{N_r - 1}{d - 1} - \frac{n_r^d}{d!},
\]

which occurs in (4.13), can be rewritten as

\[
\binom{N_r}{d} - \frac{N_r^d}{d!} - (N_r - n_r) \left( \binom{N_r - 1}{d - 1} - \frac{N_r^{d-1}}{(d-1)!} \right) + \frac{1}{d!} \left( N_r^d - n_r^d - d \left( N_r - n_r \right) N_r^{d-1} \right).
\]

Since both

\[
\binom{N_r}{d} = \frac{1}{d!} N_r(N_r - 1) \cdots (N_r - d + 1) \quad \text{and} \quad \binom{N_r - 1}{d - 1} = \frac{1}{(d-1)!} (N_r - 1) \cdots (N_r - d + 1),
\]

the differences \( \binom{N_r}{d} - \frac{N_r^d}{d!} \) and \( \binom{N_r - 1}{d - 1} - \frac{N_r^{d-1}}{(d-1)!} \) can be written as polynomials in \( N_r \) of degree \( d - 1 \) and degree \( d - 2 \), respectively, with coefficients not depending on \( r > 0 \); cf. also the proof of Lemma 4.3. Therefore, the behavior of these polynomials as \( r \to \infty \) is determined by \( N_r^{d-1} \) and \( N_r^{d-2} \), respectively. Thus, we obtain immediately that both

\[
\frac{1}{r^{d-1/2}} \left( \binom{N_r}{d} - \frac{N_r^d}{d!} \right) \xrightarrow{r \to \infty} 0 \quad \text{and} \quad \frac{n_r}{r^{d-1/2}} \left( \binom{N_r - 1}{d - 1} - \frac{N_r^{d-1}}{(d-1)!} \right) \xrightarrow{r \to \infty} 0.
\]
since $\mathbb{E}N_r^{2d-2}$ and $\mathbb{E}N_r^2$ are polynomials in $n_r$ of degree $2d-2$ and $2d-4$, respectively. Furthermore, using the binomial formula together with $\mathbb{E}N_r^2 = n_r + n_r^2$ and $\mathbb{E}N_r = n_r$, we get that

$$\mathbb{E}\left( \frac{N_r}{r} - 2\lambda \right)^2 = \frac{n_r}{r^2} + \frac{n_r^2}{r^2} - \frac{4\lambda n_r}{r} + 4\lambda^2 \xrightarrow{r \to \infty} 0,$$

which implies that $\frac{N_r}{r} \xrightarrow{P} 2\lambda = \frac{\lambda}{r}$. Together with

$$\frac{1}{r^{d-3/2}} \left( \binom{N_r}{d-1} - \frac{N_r^{d-1}}{(d-1)!} \right) \xrightarrow{r \to \infty} 0,$$

which is obtained similarly as shown above, we get that

$$\frac{N_r}{r^{d-1/2}} \left( \binom{N_r}{d-1} - \frac{N_r^{d-1}}{(d-1)!} \right) \xrightarrow{r \to \infty} 0.$$

Finally, we note that

$$N_r^d - n_r^d$$

$$= N_r^d - n_r N_r^{d-1} + n_r N_r^{d-1} - n_r^2 N_r^{d-2} + n_r^2 N_r^{d-2} - \ldots - n_r^d$$

$$= N_r^{d-1}(N_r - n_r) + n_r N_r^{d-2}(N_r - n_r) + n_r^2 N_r^{d-3}(N_r - n_r) + \ldots + n_r^{d-1}(N_r - n_r)$$

$$= (N_r - n_r) \sum_{k=0}^{d-1} n_r^k N_r^{d-1-k}$$

and also

$$\sum_{k=0}^{d-1} n_r^k N_r^{d-1-k} - dN_r^{d-1}$$

$$= (1 - d) N_r^{d-1} + (d - 1)n_r - (d - 2)n_r) N_r^{d-2}$$

$$+ ((d - 2)n_r^2 - (d - 3)n_r^2) N_r^{d-3} + \ldots + n_r^{d-1}$$

$$= -(d - 1) N_r^{d-1} - (d - 2) n_r N_r^{d-2} - \ldots - n_r^{d-2} N_r$$

$$+ (d - 1)n_r N_r^{d-2} + (d - 2) n_r^2 N_r^{d-3} + \ldots + n_r^{d-1}$$

$$= -\left( N_r \sum_{k=0}^{d-2} (d - k - 1) n_r^k N_r^{d-1-k} - n_r \sum_{k=0}^{d-2} (d - k - 1) n_r^k N_r^{d-1-k} \right).$$

Hence, we immediately get that

$$\frac{1}{r^{d-1/2}} \left( N_r^d - n_r^d - d (N_r - n_r) N_r^{d-1} \right)$$

$$= -\left( \frac{N_r - n_r}{r^{3/4}} \right)^2 \frac{1}{r^{d-2}} \left( \sum_{k=0}^{d-2} (d - k - 1) n_r^k N_r^{d-1-k} \right).$$
where \( r^{2-d} \sum_{k=0}^{d-2} (d-k-1) n_r^k N_r^{d-k-2} \) converges in probability to a positive but finite constant as \( r \to \infty \). The latter can be seen by arguing once again that second moments of the summands are polynomials in \( n_r \) of degree \( 2d - 2k - 4 \) for \( k = 0, \ldots, d - 2 \). Furthermore, \( \mathbb{E}(N_r - n_r)^2 \) is a polynomial in \( n_r \) of degree 1 and hence, \( r^{-3/2} (N_r - n_r)^2 \) converges to zero in probability as \( r \to \infty \), which yields

\[
- \left( \frac{N_r - n_r}{r^{3/4}} \right)^2 \frac{1}{r^{d-2}} \sum_{k=0}^{d-2} (d-k-1) n_r^k N_r^{d-k-2} \xrightarrow{r \to \infty} 0.
\]

This finally completes the proof by usage of Slutsky’s theorem; cf. Theorem A.8.

\[ \qquad \square \]

**Lemma 4.5** Under the conditions of Theorem 4.1,

\[
\frac{1}{r^{d-1}} \left( \frac{N_r - 1}{d - 1} \right) \xrightarrow{r \to \infty} \frac{(2\lambda)^{d-1}}{(d-1)!}.
\]

**Proof** Using a binomial formula, we immediately obtain that

\[
\mathbb{E} \left( \frac{1}{r^{d-1}} \left( \frac{N_r - 1}{d - 1} \right) - \frac{(2\lambda)^{d-1}}{(d-1)!} \right)^2 = \frac{1}{r^{2d-2}} \mathbb{E} \left( \left( \frac{N_r - 1}{d - 1} \right)^2 - 2 \frac{(2\lambda)^{d-1}}{(d-1)!} \mathbb{E} \left( \frac{N_r - 1}{d - 1} \right) + \frac{(2\lambda)^{2d-2}}{(d-1)!^2} \right).
\]

Similarly as in the proof of Lemma 4.3, we use the fact that \( \mathbb{E} N_r^k \) is a polynomial of degree \( k \in \mathbb{N} \) in \( n_r \) and that \( \mathbb{E}((N_r - 1)(N_r - 2) \cdots (N_r - d + 1)) \) is a polynomial in \( n_r \) of degree \( d - 1 \). Together with

\[
\left( \frac{N_r - 1}{d - 1} \right) = \frac{1}{(d-1)!} ((N_r - 1) \cdots (N_r - d + 1))
\]

we get that both

\[
\frac{1}{r^{2d-2}} \mathbb{E} \left( \left( \frac{N_r - 1}{d - 1} \right)^2 \right) \xrightarrow{r \to \infty} \frac{(2\lambda)^{2d-2}}{(d-1)!^2} \quad \text{and} \quad \frac{1}{r^{d-1}} \mathbb{E} \left( \frac{N_r - 1}{d - 1} \right) \xrightarrow{r \to \infty} \frac{(2\lambda)^{d-1}}{(d-1)!}.
\]

Hence, in view of (4.15), the statement in (4.14) is established.

\[ \qquad \square \]

**Lemma 4.6** Under the conditions of Theorem 4.1,

\[
\frac{1}{\sqrt{2\lambda r}} \left( \sum_{i=1}^{N_r} g_B(X_i) - 2\lambda r \mathbb{E} f_B(X_1, \ldots, X_d) \right) \xrightarrow{r \to \infty} \mathcal{N}(0, \sigma_B^{(1,d)})
\]

where \( g_B(x) = \mathbb{E} f_B(x, X_2, \ldots, X_d) \) for \( x \in [-r, r] \times S_+^{d-1} \).
4.1. CLTs for Intersection Nodes of Poisson Hyperplanes

Proof  Let again \( n_r = EN_r = 2\lambda r \) and let furthermore \( \mu_B = \mathbb{E}g_B(X_1) = \mathbb{E}f_B(X_1, \ldots, X_d) \). The characteristic function of

\[
\xi_r = n_r^{-1/2} \left( \sum_{i=1}^{N_r} g_B(X_i) - n_r \mu_B \right)
\]

is then given by

\[
\mathbb{E}e^{it\xi_r} = \exp(-it\sqrt{n_r}\mu_B) \mathbb{E} \exp\left( \frac{it}{\sqrt{n_r}} \sum_{i=1}^{N_r} g_B(X_i) \right)
\]

The expectation that shows up on the right hand side of the latter equation is the characteristic function of the compound \( n_r^{-1/2} \sum_{i=1}^{N_r} g_B(X_i) \) and can be simplified by utilizing the independence of \( N_r \) from the sequence \( X_1, X_2, \ldots \) and by using the fact that the probability generating function \( \mathbb{E}z^{N_r} \) takes the form \( \exp(n_r(z - 1)) \) for any complex \( z \). In fact, using a result similar to the one in Theorem 4.2.1 of [101], this characteristic function can be given by

\[
\mathbb{E} \exp\left( \frac{it}{\sqrt{n_r}} \sum_{i=1}^{N_r} g_B(X_i) \right) = \exp\left( n_r \left( \mathbb{E}e^{\frac{it}{\sqrt{n_r}}g_B(X_1)} - 1 \right) \right)
\]

Thus,

\[
\mathbb{E}e^{it\xi_r} = \exp\left( n_r \mathbb{E}\left( e^{\frac{it}{\sqrt{n_r}}g_B(X_1)} - 1 - \frac{it}{\sqrt{n_r}}g_B(X_1) \right) \right)
\]

or, equivalently, \( \log \mathbb{E}e^{it\xi_r} \) is given by

\[
n_r \mathbb{E}\left( \exp\left( \frac{it}{\sqrt{n_r}} g_B(X_1) \right) - 1 - \frac{it}{\sqrt{n_r}} g_B(X_1) - \frac{(it)^2}{2n_r} g_B^2(X_1) \right) - \frac{t^2 \sigma_B^{(1,d)}}{2}.
\]

The well-known inequality \( |e^{ix} - 1 - ix - \frac{|x|^2}{2}| \leq \frac{|x|^2}{6} \) for any \( x \in \mathbb{R} \) combined with \( n_r^{-1/2} \mathbb{E}[g_B^2(X_1)] \sim \mathcal{R} \) gives

\[
\log \mathbb{E}e^{it\xi_r} \sim -\frac{t^2 \sigma_B^{(1,d)}}{2} \quad \text{for} \ t \in \mathbb{R},
\]

which is equivalent to the assertion of Lemma 4.6. \( \square \)

4.1.3 Review of the Planar Case

Throughout this section we assume that \( d = 2 \) and that the underlying stationary Poisson line process is isotropic. A short review of Theorem 4.1 is presented for the case where the ordered angles of the orientation vectors of intersecting pairs of lines are situated within a specific rectangle \( B(a,b) = [0,a] \times [0,b] \), with \( 0 \leq a \leq b \leq \pi \) and \( B(a,b) \subseteq [0, \pi]^2 \).

Consider once again the marked–point–process representation \( \Phi'_1 = \sum_{i \geq 1} \delta_{[P_i, U_i]} \) of a planar stationary and isotropic Poisson line process \( \Phi_1 \) with intensity \( \lambda > 0 \). In this special case each
orientation vector \( U_i \in \mathbb{S}^1 \) is completely determined by the angle \( \Delta_i \) between the unit vector \( U_i \) and the \( x\)-axis measured in anti-clockwise direction. Owing to the isotropy of \( \Phi_1 \), the angles \( \Delta_1, \Delta_2, \ldots \) are independent and uniformly distributed on \([0, \pi]\). Therefore, \( \text{supp } \Phi_1 \) consists of parametrized lines \( \ell(p_i, \Delta_i) \) in \( \mathbb{R}^2 \), where \( \ell(p_i, \Delta_i) = \{(x, y) \in \mathbb{R}^2 : x \cos \Delta_i + y \sin \Delta_i = p_i\} \).

For any \( r > 0 \) fixed, \( \eta_0(B_r^2 \times B(a, b)) \) is the random number of those intersection points \( \ell(p_{i1}, \Delta_{n1}) \cap \ell(p_{i2}, \Delta_{n2}) \) in \( B_r^2 \) for which the ordered bivariate vector \((\Delta_{(n1)}, \Delta_{(n2)}) \in B(a, b)\). Recall that

\[
\eta_0(B_r^2 \times B(a, b)) = \frac{1}{2} \sum_{i,j=1}^{N_r} f_{B(a,b)}((Q_i, \Delta_i), (Q_j, \Delta_j)),
\]

where the symbol \( \sum^* \) means that the sum runs over all pairs of indices \( i \) and \( j \) with \( i \neq j \) and where, as in Sections 4.1.1 and 4.1.2, we have that the random vectors \((Q_1, \Delta_1), (Q_2, \Delta_2), \ldots : \Omega \rightarrow [-r, r] \times [0, \pi] \) are independent and uniformly distributed on \([-r, r] \times [0, \pi] \). Furthermore \( N_r : \Omega \rightarrow \{0, 1, \ldots \} \) is a Poisson distributed random variable having expectation \( 2\pi r \) and being independent of the \((Q_i, \Delta_i)\)'s. The function \( f_{B(a,b)} \) given by

\[
f_{B(a,b)}((p_1, \delta_1), (p_2, \delta_2)) = \chi(\ell(p_1, \delta_1) \cap \ell(p_2, \delta_2) \cap B_r^2) \mathbb{I}_{B(a,b)}(\delta_1, \delta_2))
\]

is symmetric since \((\delta_1, \delta_2)\) are lexicographically ordered, i.e., \((\delta_1, \delta_2) = (\delta_1 \wedge \delta_2, \delta_1 \vee \delta_2)\).

With the abbreviations

\[
\mu_{B(a,b)} = \mathbb{E} f_{B(a,b)}((Q_1, \Delta_1), (Q_2, \Delta_2))
\]

\[
= \mathbb{P}(\ell(p_{i1}, \Delta_{i1}) \cap \ell(p_{i2}, \Delta_{i2}) \cap B_r^2 \neq \emptyset, 0 \leq \Delta_1 \wedge \Delta_2 \leq a, 0 \leq \Delta_1 \vee \Delta_2 \leq b)
\]

and \( \sigma_{B(a,b)}^{(1,2)} \) defined by (4.3), we obtain the following corollary of Theorem 4.1.

**Corollary 4.7** Let \( B(a, b) = [0, a] \times [0, b] \) with \( 0 \leq a \leq b \leq \pi \). Then,

\[
\frac{1}{(2\lambda r)^{3/2}} \left( \eta_0(B_r^2 \times B(a, b)) - \frac{(2\lambda r)^2}{2} \mu_{B(a,b)} \right) \xrightarrow{d_{r \to \infty}} \mathcal{N}(0, \sigma_{B(a,b)}^{(1,2)}),
\]

where

\[
\mu_{B(a,b)} = \frac{1}{4\pi} \int_0^a \int_0^b |\sin(u - v)| \, du \, dv = \frac{1}{4\pi} \left( 2a - \sin a - \sin b + \sin(b - a) \right)
\]

and where \( \mathcal{N}(0, \sigma_{B(a,b)}^{(1,2)}) \) denotes a Gaussian random variable with mean zero and variance \( \sigma_{B(a,b)}^{(1,2)} \), given by

\[
\sigma_{B(a,b)}^{(1,2)} = \frac{2}{3\pi^2} \left( 5a + 7(\sin(b - a) - \sin b) \right.
\]

\[
+ (b + \sin b) \cos b \left( b - a + \sin(b - a) \right) \cos(b - a) \right).
\]
4.1. CLTs for Intersection Nodes of Poisson Hyperplanes

\textbf{Proof} \hspace{1em} Since by (4.20) we have that $\sigma^{(1,2)}_{B(a,b)} > 0$, we obtain (4.18) as an immediate consequence of Theorem 4.1 for $d = 2$ and $B = B(a,b)$. To verify (4.19) we use a direct approach. By definition (4.17), we obtain

$$
\mu_{B(a,b)} = \frac{1}{4\pi^2} \int_0^\pi \int_0^\pi \int_{-1}^1 \int_{-1}^1 \mathbb{I}_{(x^2 + y^2 \leq 1)}(p_1, \delta_1, p_2, \delta_2) \times \mathbb{I}_{[0,a]}(\delta_1 \wedge \delta_2) \mathbb{I}_{[0,b]}(\delta_1 \vee \delta_2) \, dp_1 \, dp_2 \, d\delta_1 \, d\delta_2,
$$

where the factor $1/(4\pi^2)$ occurs due to the uniform distribution of $Q_1$ and $Q_2$ as well as of the angles $\Delta_1$ and $\Delta_2$, and where $(x, y)^T$ denotes the coordinate vector of the intersection point of the two lines $\ell_{(p_1, \delta_1)}$ and $\ell_{(p_2, \delta_2)}$, i.e.,

$$
x = -\frac{p_1 \sin \delta_2 + p_2 \sin \delta_1}{\sin(\delta_2 - \delta_1)} \quad \text{and} \quad y = \frac{p_1 \cos \delta_2 + p_2 \cos \delta_1}{\sin(\delta_2 - \delta_1)}.
$$

Using these latter explicit expressions for $x$ and $y$, it is possible to see that the condition $x^2 + y^2 \leq 1$ is equivalent to the condition that $p_1$ is situated in an interval of length $2\sqrt{1 - p_2^2 |\sin(\delta_2 - \delta_1)|}$, where the lower and upper boundary of this interval is given by $p_2 \cos((\delta_2 - \delta_1) \pm \sqrt{1 - p_2^2 \sin^2(\delta_2 - \delta_1)})$, respectively. Hence we arrive at

$$
\mu_{B(a,b)} = \frac{1}{2\pi} \int_a^b \int_{-1}^1 \sqrt{1 - p_2^2 |\sin(\delta_2 - \delta_1)|} \, dp_2 \, d\delta_1 \, d\delta_2,
$$

which, combined with $\int_{-1}^1 \sqrt{1 - p^2} \, dp = \frac{\pi}{2}$, confirms (4.19). To determine the asymptotic variance $\sigma^{(1,2)}_{B(a,b)}$ appearing in (4.18), we argue similarly. By (4.3),

$$
\sigma^{(1,2)}_{B(a,b)} = \mathbb{E} \left( f_{B(a,b)}((Q_1, \Delta_1), (Q_2, \Delta_2)) \right) = \frac{1}{8\pi^3} \int_0^\pi \int_0^\pi \int_{-1}^1 \int_{-1}^1 \mathbb{I}_{(x^2 + y^2 \leq 1)}(p_1, \delta_1, p_2, \delta_2) \times \mathbb{I}_{(x^2 + y^2 \leq 1)}(p_2, \delta_2, p_3, \delta_3) \mathbb{I}_{[0,a]}(\delta_1 \wedge \delta_2 \wedge \delta_3) \mathbb{I}_{[0,b]}(\delta_1 \vee \delta_2 \vee \delta_3) \, dp_1 \, dp_2 \, dp_3 \, d\delta_1 \, d\delta_2 \, d\delta_3
$$

$$
= \frac{4}{8\pi^3} \int_a^b \int_0^b \int_0^b \int_{-1}^1 (1 - p_2^2) \, |\sin(\delta_1 - \delta_2)| \, |\sin(\delta_2 - \delta_3)| \, dp_2 \times \mathbb{I}_{[0,a]}(\delta_1 \wedge \delta_2 \wedge \delta_3) \, d\delta_1 \, d\delta_2 \, d\delta_3.
$$

Using that $\int_{-1}^1 (1 - p^2) \, dp = \frac{4}{3}$ and that $\mathbb{I}_{[0,a]}(\delta_1 \wedge \delta_2 \wedge \delta_3) = 1 - \prod_{i=1}^3 \mathbb{I}_{(a,\pi]}(\delta_i)$ we get that

$$
\sigma^{(1,2)}_{B(a,b)} = \frac{2}{3\pi^3} \left( \int_a^b \int_0^b \int_0^b |\sin(\delta_1 - \delta_2)| \, |\sin(\delta_2 - \delta_3)| \, d\delta_1 \, d\delta_2 \, d\delta_3 - \int_a^b \int_a^b \int_0^b |\sin(\delta_1 - \delta_2)| \, |\sin(\delta_2 - \delta_3)| \, d\delta_1 \, d\delta_2 \, d\delta_3 \right).
$$
Thus we need to solve
\[
\int_{a}^{b} \int_{a}^{b} \int_{a}^{b} |\sin(\delta_1 - \delta_2)| |\sin(\delta_2 - \delta_3)| \, d\delta_1 \, d\delta_2 \, d\delta_3 = \int_{b}^{a} \left( \int_{a}^{b} |\sin(\delta_1 - \delta_2)| \, d\delta_1 \right)^2 \, d\delta_2 ,
\]
since for \( a = 0 \) we also get the solution of the other 3-fold integral. Therefore we first solve the inner integral on the right hand side of the latter equation and get that
\[
\int_{a}^{b} |\sin(\delta_1 - \delta_2)| \, d\delta_1 = \int_{\delta_2}^{b} \sin(\delta_1 - \delta_2) \, d\delta_1 - \int_{a}^{\delta_2} \sin(\delta_1 - \delta_2) \, d\delta_1
\]
\[= 2 - \cos(b - \delta_2) - \cos(a - \delta_2) .
\]
Using this result we obtain that
\[
\int_{a}^{b} (2 - \cos(b - \delta_2) - \cos(a - \delta_2))^2 \, d\delta_2
\]
\[= 5(b - a) - 7 \sin(b - a) + \sin(b - a) \cos(b - a) + (b - a) \cos(b - a) .
\]
where we used the well-known relations
\[
2 \cos^2(s - t) = (\cos(2s - 2t) + 1) \quad \text{and} \quad 2 \cos(s - t) \cos(u - t) = (\cos(s - u) + \cos(s + u - 2t) .
\]
Hence, the other 3-fold integral (where \( a = 0 \)) is then immediately seen to equal
\[
5b - 7 \sin b + \sin b \cos b + b \cos b .
\]
Combining the results yields (4.20).

4.1.4 A CLT with Random Normalization

In this section we remain in the two-dimensional Euclidean space and we review briefly another type of a CLT for Poisson line processes, where the normalization is random. The proof of such a CLT goes back to Paroux ([94]), where the method of moments ([16]) has been used as proving technique, resulting in somewhat lengthy computations.

Using directly Hoeffding’s CLT for \( U \)-statistics, cf. Theorem A.19 in Section A.19, we provide a new proof of Paroux’s CLT with random normalization, where the advantage of our proof lies in its shortness.

\[\textbf{Theorem 4.2} \quad \text{For arbitrary fixed } a, b \in [0, \pi] \text{ with } a < b, \text{ let}
\]
\[
Z_{B(a,b)}^{(r)} = \frac{1}{(N_r(N_r - 1))^{3/4}} \sum_{i,j=1}^{N_r} \left( f_{B(a,b)}(X_i, X_j) - \mu_{B(a,b)} \right) . \tag{4.21}
\]

Then,
\[
Z_{B(a,b)}^{(r)} \xrightarrow{d} \mathcal{N}(0, \sigma_{B(a,b)}^{(1,2)} - \mu_{B(a,b)}^2) ,
\]
where \( \mu_{B(a,b)} \) and \( \sigma_{B(a,b)}^{(1,2)} \) are given by (4.19) and (4.20), respectively.
4.2. CLTs for Flat Processes induced by Poisson Hyperplanes

Proof. To begin with, we rewrite $Z_{B(a,b)}^{(r)}$ as

$$Z_{B(a,b)}^{(r)} = \frac{N_r^{1/2}}{2} \left( 1 - \frac{1}{N_r} \right)^{1/4} \left( \frac{2}{N_r(N_r - 1)} \sum_{i,j=1}^{N_r} (f_{B(a,b)}(X_i, X_j) - \mu_{B(a,b)}) \right),$$

where $f_{B(a,b)}(x_1, x_2)$ is a measurable, symmetric function in $x_1, x_2 \in [-r, r] \times S^1$ satisfying $\mathbb{E}f_{B(a,b)}(X_1, X_2) = \mu_{B(a,b)}$. Application of Hoeffding’s CLT for $U$-statistics, cf. Theorem A.19, yields that

$$\frac{[n_r]^{1/2}}{2} U^{(2)}_{[n_r]}(f_{B(a,b)}) \xrightarrow{d_{r \to \infty}} \mathcal{N}(0, \text{Var } g_{B(a,b)}(X_1)),$$

(4.22)

where $[n_r]$ stands for the integer part of $n_r = \mathbb{E}N_r = 2\lambda r$ and where the $U$-statistic $U^{(2)}_{[n_r]}(f_{B(a,b)})$ is given by

$$U^{(2)}_{[n_r]}(f_{B(a,b)}) = \frac{2}{[n_r]/([n_r] - 1)} \sum_{i,j=1}^{[n_r]} (f_{B(a,b)}(X_i, X_j) - \mu_{B(a,b)})$$

with $g_{B(a,b)}(x) = \mathbb{E}f_{B(a,b)}(x, X_2)$. Taking into account the abbreviating notation introduced in (4.3) for $d = 2, j = 1$ and Hoeffding’s decomposition for $U$-statistics given in (A.16), we get that

$$\text{Var } g_{B(a,b)}(X_1) = \sigma_{B(a,b)}^{(1,2)} - \mu_{B(a,b)}^2.$$ 

Since $N_r \xrightarrow{P_{r \to \infty}} \infty$ such that $N_r/[n_r] \xrightarrow{P_{r \to \infty}} 1$, Theorem VIII.7.1 in [99] tells us that in (4.22) the deterministic integer $[n_r]$ can be replaced by the random number $N_r$ without changing the limit $\mathcal{N}(0, \text{Var } g_{B(a,b)}(X_1))$. Finally, a straightforward application of Slutsky’s theorem (cf. Theorem A.8) completes the proof of Theorem 4.2.

Note that it may be regarded as an advantage of random normalization that then the asymptotic variance $\sigma_{B(a,b)}^{(1,2)} - \mu_{B(a,b)}^2$ of the number of intersection points in (4.21) is always smaller than the asymptotic variance $\sigma_{B(a,b)}^{(1,2)}$ obtained in (4.18) under deterministic normalization.

Furthermore, instead of imposing conditions on the angles of the orientation vectors of the intersecting lines $\ell_{(p_1,\delta_1)}$ and $\ell_{(p_2,\delta_2)}$, the two angles at the intersecting points $x_i$ and $x_j$ of $\ell_{(p_1,\delta_1)}$ and $\ell_{(p_2,\delta_2)}$, respectively, with the $x$-axis are considered in [94]. More precisely, the ordered pair of angles $(\alpha, \beta)$ is considered, where $\alpha$ denotes the angle at $x_i \land x_j$ between the $x$-axis and the half-line from $\ell_{(p_1,\delta_1)} \cap \ell_{(p_2,\delta_2)}$ to the intersection point with the $x$-axis and, likewise, $\beta$ is the angle at $x_i \lor x_j$.

4.2 CLTs for Flat Processes induced by Poisson Hyperplanes

The results in the preceding Section 4.1 can be generalized as follows. Recall again that $\Phi_{d-1}^{\prime} = \sum_{i \geq 1} \delta_{[P_i, \ell_{(i)]}}$ is the marked-point-process representation of a stationary non-degenerate Poisson hyperplane process $\Phi_{d-1}$ in $\mathbb{R}^d$ with intensity $\lambda > 0$. The non-degeneracy assumption
means that \( \tilde{\Theta}(H(0, u) \cap S_{d-1}^+) < 1 \) for any \( u \in S_{d-1}^+ \), where \( \tilde{\Theta} \) denotes the spherical orientation distribution. For notational ease, the results in this section are presented for the particular case \( B = (S_{d-1}^+)^d \) only. However, they remain valid for more general sets \( B \in \mathcal{B}((S_{d-1}^+)^d) \).

To be more precise, in this section we derive CLTs for the number \( \Phi_k(\{ L \in \mathcal{A}_d^d : L \cap B_r^d \neq \emptyset \}) \) of \( k \)-flats hitting the spherical sampling window \( B_r^d \) as well as for their total \( k \)-volume contained in \( B_r^d \) when the radius \( r \) tends to infinity; \( k = 0, \ldots, d - 1 \). Note that the particular case \( k = 0 \) has already been considered in the preceding Section 4.1.

Hence, instead of the intersection of \( d \) hyperplanes, we consider a generalized version of Theorem 4.1. In particular, for any fixed \( k \in \{0, \ldots, d - 1\} \) we define families of statistics \( \eta_k(B_r^d) \) and \( \zeta_k(B_r^d) \), denoting the number of \( k \)-flats hitting \( B_r^d \) and their total \( k \)-volume in \( B_r^d \), respectively, where the \( k \)-flats are induced by the intersection of \( d - k \) hyperplanes of the process \( \Phi_{d-1} \) represented by \( \Phi'_{d-1} \). These statistics, closely related to the \( k \)-flat intersection process \( \Phi_k \) defined in (2.10), have the form of (random) \( U \)-statistics of order \( d - k \) without normalizing factor \( (N_r)_{d-k} \), i.e.,

\[
\eta_k(B_r^d) = \Phi_k(\{ L \in \mathcal{A}_d^d : L \cap B_r^d \neq \emptyset \})
\]

\[
\frac{d}{(d - k)!} \sum_{1 \leq i_1, \ldots, i_{d-k} \leq N_r} \chi(\cap_{j=1}^{d-k} H(X_{i_j}) \cap B_r^d)
\]

and

\[
\zeta_k(B_r^d) = \sum_{L \in \text{supp} \Phi_k} \nu_k(B_r^d \cap L)
\]

\[
\frac{d}{(d - k)!} \sum_{1 \leq i_1, \ldots, i_{d-k} \leq N_r} \nu_k(\cap_{j=1}^{d-k} H(X_{i_j}) \cap B_r^d)
\]

where we recall that \( \sum^* \) denotes summation over pairwise distinct indices, and where we note that \( \nu_0(\cap_{j=1}^{d} H(X_{i_j}) \cap B_r^d) \) coincides with \( \chi(\cap_{j=1}^{d} H(X_{i_j}) \cap B_r^d) \); cf. also Section A.1.3. Also, recall that \( X_i = (Q_i, U_i) \) are independent random vectors, independent of \( N_r \), with independent components \( Q_i \) and \( U_i \) such that \( Q_i \) is uniformly distributed on \([-r, r] \) and \( U_i \) has the distribution \( \tilde{\Theta} \) for \( i \geq 1 \).

### 4.2.1 CLTs for \( k \)-Flat Processes

In analogy to Section 4.1.1 note that the expectations \( \mathbb{E} \chi(H(X_1) \cap \ldots \cap H(X_{d-k}) \cap B_r^d) \) and \( r^{-k} \mathbb{E} \eta_k(H(X_1) \cap \ldots \cap H(X_{d-k}) \cap B_r^d) \) do not depend on \( r > 0 \) since the first components \( Q_i \) of the iid random vectors \( X_i = (Q_i, U_i) \), \( i \geq 1 \), are uniformly distributed on \([-r, r] \).

As an extension of Lemma 4.1 for \( B = (S_{d-1}^+)^d \) we now determine the first-order and second-order moments of the random variables \( \eta_k(B_r^d) \) and \( \zeta_k(B_r^d) \) for \( k = 0, \ldots, d - 1 \) given in (4.23) and (4.24), respectively. Therefore, let

\[
\sigma_{\chi,k}^{(d-k)} = \mathbb{E} \left( \chi(\cap_{i=1}^{d-k} H(X_{i}) \cap B_r^d) \chi(\cap_{i=d-k+1}^{2d-k-1} H(X_{i}) \cap B_r^d) \right) = \mathbb{E} g_{\chi,k}^2(X_{d-k})
\]
and
\[
\sigma_{\nu,k}^{(1,d-k)} = r^{-2k} \mathbb{E}\left( \nu_k(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d) \nu_k(\cap_{i=d-k+1}^{2d-2k-1} H(X_i) \cap B_r^d) \right)
= \mathbb{E} g_{\nu,k}^2(X_{d-k}), \tag{4.26}
\]
where
\[
g_{\chi,k}(p,u) = \mathbb{E}\chi(H(X_1) \cap \ldots \cap H(X_{d-k-1}) \cap H(p,u) \cap B_r^d) \tag{4.27}
\]
and
\[
g_{\nu,k}(p,u) = r^{-k} \mathbb{E}\nu_k(H(X_1) \cap \ldots \cap H(X_{d-k-1}) \cap H(p,u) \cap B_r^d) \tag{4.28}
\]
for \((p,u) \in [-r,r] \times \mathbb{S}^{d-1}_{+}\). Note that the second moments \(\sigma_{\chi,k}^{(1,d-k)}\) and \(\sigma_{\nu,k}^{(1,d-k)}\) do also not depend on \(r > 0\). Using this notation, we can state the following moment formulae.

**Lemma 4.8** For each \(k = 0, \ldots, d-1\),
\[
\mathbb{E} \eta_k(B_r^d) = \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{P}(H(X_1) \cap \ldots \cap H(X_{d-k}) \cap B_r^d \neq \emptyset), \tag{4.29}
\]
\[
\mathbb{E} \zeta_k(B_r^d) = \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{E}\nu_k(H(X_1) \cap \ldots \cap H(X_{d-k}) \cap B_r^d), \tag{4.30}
\]
and
\[
\lim_{r \to \infty} \frac{\text{Var} \eta_k(B_r^d)}{r^{2d-2k-1}} = \frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2} \sigma_{\chi,k}^{(1,d-k)}, \tag{4.31}
\]
\[
\lim_{r \to \infty} \frac{\text{Var} \zeta_k(B_r^d)}{r^{2d-1}} = \frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2} \sigma_{\nu,k}^{(1,d-k)}. \tag{4.32}
\]

**Proof** Let \(k \in \{0, \ldots, d-1\}\). In analogy to the proof of Lemma 4.1 we get that
\[
\mathbb{E} \eta_k(B_r^d) = \mathbb{E}\left( \frac{N_r}{d-k} \right) \mathbb{E}\chi(H(X_1) \cap \ldots \cap H(X_{d-k}) \cap B_r^d)
\]
and
\[
\mathbb{E} \zeta_k(B_r^d) = \mathbb{E}\left( \frac{N_r}{d-k} \right) \mathbb{E}\nu_k(H(X_1) \cap \ldots \cap H(X_{d-k}) \cap B_r^d).
\]
Applying (4.6), i.e.,
\[
\mathbb{E}(N_r(N_r - 1) \ldots (N_r - (d-k) + 1)) = (2\lambda r)^{d-k},
\]
we obtain that
\[
\mathbb{E}\left( \frac{N_r}{d-k} \right) = \frac{(2\lambda r)^{d-k}}{(d-k)!},
\]
which gives both (4.29) and (4.30). Furthermore, along the lines of the proof of Lemma 4.1, we obtain that

\[
\mathbb{E} \eta_k^2(B_r^d) = \frac{1}{(d-k)!} \mathbb{E} \left( \sum_{1 \leq i_1, \ldots, i_d \leq N_r} \chi(H(X_{i_1}) \cap \ldots \cap H(X_{i_d-k}) \cap B_r^d) \right)^2 \\
= \sum_{j=0}^{d-k} \frac{(2\lambda r)^{2d-2k-j}}{j!((d-k-j))!^2} \\
\times \mathbb{E} \left( \chi(\cap_{p=1}^{d-k} H(X_p) \cap B_r^d) \chi(\cap_{q=d-k-j+1}^{2(d-k)-j} H(X_q) \cap B_r^d) \right) \\
\tag{4.33}
\]

and

\[
\mathbb{E} \zeta_k^2(B_r^d) = \frac{1}{(d-k)!} \mathbb{E} \left( \sum_{1 \leq i_1, \ldots, i_d \leq N_r} \nu_k(H(X_{i_1}) \cap \ldots \cap H(X_{i_d-k}) \cap B_r^d) \right)^2 \\
= \sum_{j=0}^{d-k} \frac{(2\lambda r)^{2d-2k-j}}{j!((d-k-j))!^2} \\
\times \mathbb{E} \left( \nu_k(\cap_{p=1}^{d-k} H(X_p) \cap B_r^d) \nu_k(\cap_{q=d-k-j+1}^{2(d-k)-j} H(X_q) \cap B_r^d) \right). \\
\tag{4.34}
\]

Hence, immediately,

\[
\text{Var} \eta_k(B_r^d) = \sum_{j=1}^{d-k} \frac{(2\lambda r)^{2d-2k-j}}{j!((d-k-j))!^2} \\
\times \mathbb{E} \left( \chi(\cap_{p=1}^{d-k} H(X_p) \cap B_r^d) \chi(\cap_{q=d-k-j+1}^{2(d-k)-j} H(X_q) \cap B_r^d) \right) \\
\]

and

\[
\text{Var} \zeta_k(B_r^d) = \sum_{j=1}^{d-k} \frac{(2\lambda r)^{2d-2k-j}}{j!((d-k-j))!^2} \\
\times \mathbb{E} \left( \nu_k(\cap_{p=1}^{d-k} H(X_p) \cap B_r^d) \nu_k(\cap_{q=d-k-j+1}^{2(d-k)-j} H(X_q) \cap B_r^d) \right). \\
\]

After dividing by \( r^{2d-2k-1} \) and \( r^{2d-1} \), respectively, and letting \( r \to \infty \), we get the desired relationships (4.31) and (4.32).

Recall now that the random variables \( \eta_k(B_r^d) \) and \( \zeta_k(B_r^d) \) given in (4.23) and (4.24), respectively, can be expressed as \( U \)-statistics conditional that \( N_r = n \), say, allowing to apply Hoeffding’s decomposition; cf. Formula A.16. Hence we can state the following CLTs, the proofs of which are in complete analogy to the proof of Theorem 4.1.

**Theorem 4.3** Let \( \Phi_{d-1} \) be a stationary non-degenerate Poisson hyperplane process with intensity \( \lambda > 0 \). Then, for \( k = 0, 1, \ldots, d-1 \),

\[
Z_{k,r}(x) = \frac{(d-k-1)!}{(2\lambda r)^{d-k-1/2}} \left( \eta_k(B_r^d) - \mathbb{E} \eta_k(B_r^d) \right) \\
\xrightarrow{r \to \infty} \mathcal{N}(0, \sigma_{\chi,k}^2) \\
\tag{4.35}
\]
and
\[ Z_{k,r}^{(d)}(\nu) = \frac{(d-k-1)!r^{-k}}{(2\lambda r)^{d-k-1/2}} \left( \zeta_k(B_r^d) - \mathbb{E} \zeta_k(B_r^d) \right) \xrightarrow{\text{d}} N(0, \sigma_{\chi,k}^{(1,d-k)}) , \tag{4.36} \]
where \( N(0, \sigma_{\chi,k}^{(1,d-k)}) \) and \( N(0, \sigma_{\nu,k}^{(1,d-k)}) \) are Gaussian random variables with mean zero and variance \( \sigma_{\chi,k}^{(1,d-k)} \) and \( \sigma_{\nu,k}^{(1,d-k)} \), respectively.

**Proof** Let \( k \in \{0, \ldots, d-1\} \). The equality in distribution expressed in formulae (4.23) and (4.24) is equivalent to the equality in distribution
\[ \eta_k(B_r^d) \overset{d}{=} \left( \binom{N_r}{d-k} \right)^{-1} U_{N_r}^{(d-k)}(\chi) \quad \text{and} \quad r^{-k} \zeta_k(B_r^d) \overset{d}{=} \left( \binom{N_r}{d-k} \right) U_{N_r}^{(d-k)}(\nu) , \]
respectively. Hence, with
\[ U_{N_r}^{(d-k)}(\chi) = \left( \binom{N_r}{d-k} \right)^{-1} \frac{1}{(d-k)!} \sum_{1 \leq i_1, \ldots, i_d \leq N_r} \chi(\cap_{j=1}^{d-k} H(X_{i_j}) \cap B_r^d) \tag{4.37} \]
and
\[ U_{N_r}^{(d-k)}(\nu) = \left( \binom{N_r}{d-k} \right)^{-1} \frac{1}{(d-k)!} \sum_{1 \leq i_1, \ldots, i_d \leq N_r} r^{-k} \nu_k(\cap_{j=1}^{d-k} H(X_{i_j}) \cap B_r^d) , \tag{4.38} \]
we get that for any fixed \( N_r = n \geq d-k \) the random multiple sums \( U_{N_r}^{(d-k)}(\chi) \) and \( U_{N_r}^{(d-k)}(\nu) \) are U-statistics as defined in (A.15) of order \( d-k \) with kernel function \( \chi(\cdot) \) and \( r^{-k} \nu_k(\cdot) \), respectively. With the abbreviating notation \( n_r = EN_r = 2\lambda r \) and by usage of (4.29), we get that
\[ \eta_k(B_r^d) - \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{P}(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d \neq \emptyset) \]
\[ \overset{d}{=} \left( \binom{N_r}{d-k} \right) U_{N_r}^{(d-k)}(\chi) - \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{P}(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d \neq \emptyset) \]
\[ = \left( \binom{N_r}{d-k} \right) - \frac{n_r^{d-k}}{(d-k)!} \mathbb{P}(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d \neq \emptyset) + \left( \binom{N_r}{d-k} \right) R_{N_r}^{(d-k)}(\chi) \]
\[ + \left( \binom{N_r}{d-k} \right) \frac{d-k}{N_r} \sum_{i=1}^{N_r} \left( g_{\chi,k}(X_i) - \mathbb{P}(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d \neq \emptyset) \right) , \]
where \( g_{\chi,k}(\cdot) \) is given by (4.27), and where we used Hoeffding’s decomposition of U-statistics, cf. Formula (A.16) in Section A.3.7, applied to \( U_{N_r}^{(d-k)}(\chi) \). Likewise, using (4.30), we get that
\[ r^{-k} \left( \zeta_k(B_r^d) - \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{E} \nu_k(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d) \right) \]
\[ \overset{d}{=} \left( \binom{N_r}{d-k} \right) U_{N_r}^{(d-k)}(\nu) - \frac{(2\lambda r)^{d-k}}{(d-k)!} r^{-k} \mathbb{E} \nu_k(\cap_{i=1}^{d-k} H(X_i) \cap B_r^d) . \]
With \( g_{\nu,k}(\cdot) \) given by (4.28) and by application of Hoeffding’s decomposition of \( U \)-statistics to \( U^{(d-k)}_{N_r}(\nu) \), we obtain that

\[
\frac{N_r}{d-k} U^{(d-k)}_{N_r}(\nu) - \frac{(2\lambda r)^{d-k}}{(d-k)!} r^{-k} \varepsilon_{\nu k}(\cap_{i=1}^{d-k} H(X_i) \cap B^{d}_{r})
\]

\[
= \left( \frac{N_r}{d-k} - \frac{n^{d-k}}{(d-k)!} \right) r^{-k} \varepsilon_{\nu k}(\cap_{i=1}^{d-k} H(X_i) \cap B^{d}_{r}) + \left( \frac{N_r}{d-k} \right) r^{-k} R^{(d-k)}_{N_r}(\nu)
\]

\[
+ \left( \frac{N_r}{d-k} \right) \sum_{i=1}^{N_r} (g_{\nu,k}(X_i) - r^{-k} \varepsilon_{\nu k}(\cap_{i=1}^{d-k} H(X_i) \cap B^{d}_{r})).
\]

Hence, the proof can now be finished analogously to the proof of Theorem 4.1 by noting that both \( \chi(\cdot) \) and \( r^{-k} \varepsilon_{\nu k}(\cdot) \) fulfill the properties of the kernel function \( f_B \) as employed in the latter theorem. \( \square \)

### 4.2.2 Explicit Moment Formulae

If the stationary Poisson hyperplane process \( \Phi_{d-1} \) is additionally isotropic, i.e., \( \Theta \) is the uniform distribution on \( S^{d-1}_{+} \), the intensity \( \lambda_{k} \) of the \( k \)-flat intersection process \( \Phi_{k} \) induced by \( \Phi_{d-1} \) is explicitly given by (2.12) for \( k = 0, \ldots, d - 1 \). In the following, we reconsider formulae (4.29) and (4.30) for \( \varepsilon_{\nu k}(B^{d}_{r}) \) and \( \varepsilon_{\eta k}(B^{d}_{r}) \), respectively, as well as formulae (4.25) and (4.26) for \( \sigma_{\chi,\nu,k}^{(1,d-k)} \) and \( \sigma_{\nu,k}^{(1,d-k)} \), respectively. The latter two expressions have been defined in Section 4.2.1 in terms of the functions \( g_{\chi,k}(\cdot) \) and \( g_{\nu,k}(\cdot) \), given in (4.27) and (4.28), respectively, for \( k = 0, \ldots, d - 1 \). In particular, we provide explicit expressions for the expectations and asymptotic variances occurring in the CLTs (4.35) and (4.36), respectively.

**Theorem 4.4** Let \( \Phi_{d-1} \) be a stationary non-degenerate Poisson hyperplane process with intensity \( \lambda > 0 \). Then, for \( k = 0, 1, \ldots, d - 1 \),

\[
\varepsilon_{\nu k}(B^{d}_{r}) = \lambda_{k} \kappa_{d-k} r^{d-k} \quad \text{and} \quad \varepsilon_{\eta k}(B^{d}_{r}) = \lambda_{k} \kappa_{d} r^{d},
\]

with \( \lambda_{k} \) defined by (2.5). If additionally \( \Phi_{d-1} \) is isotropic, \( \lambda_{k} \) is explicitly given by (2.12), i.e.,

\[
\lambda_{k} = \left( \frac{d}{k} \right) \frac{\kappa_{d}}{\kappa_{k}} \left( \frac{\kappa_{d-1}}{d \kappa_{d}} \right)^{d-k} \lambda^{d-k} \quad \text{for} \quad k = 0, 1, \ldots, d - 1.
\]

Moreover,

\[
\sigma_{\chi,\nu,k}^{(1,d-k)} = \frac{(\kappa_{d-k-1} (d-k-1)!)^2}{(2d-2k-1)!} \left( \frac{d! \kappa_{d}}{k! \kappa_{k}} \right)^2 \left( \frac{\kappa_{d-1}}{d \kappa_{d}} \right)^{2(d-k)}
\]

and

\[
\sigma_{\nu,k}^{(1,d-k)} = \frac{2k \kappa_{d-1} (d-1)!^2}{(2d-1)!} \left( \frac{d! \kappa_{d}}{k! \kappa_{k}} \right)^2 \left( \frac{\kappa_{d-1}}{d \kappa_{d}} \right)^{2(d-k)}.
\]
Before we state the proof of Theorem 4.4, note that the distribution of \( \eta_{d-1}(B_1^d) \) and \( \zeta_{d-1}(B_1^d) \) is independent of the spherical orientation distribution \( \Theta \). This is due to the fact that the number of hyperplanes \( H(P_i, U_i) \) hitting \( B_1^d \) for \( i \geq 1 \) as well as the total \((d-1)\)-volume of these hyperplanes in \( B_1^d \) do not depend on \( U_i \). Therefore, the formulae in (4.40) and (4.41), respectively, hold for \( k = d - 1 \) also in the anisotropic case.

**Proof** Let \( k \in \{0, \ldots, d - 1\} \). Both formulae for the mean values in (4.39) are an immediate consequence of (4.23) and (4.24), using (2.6) and (2.7), respectively. Clearly, in case \( \Phi_{d-1} \) is additionally isotropic, the intensities \( \lambda_k \) can be explicitly determined by (2.12). To show (4.40), we use the relation (4.25), i.e., \( \sigma_{(1,d-k)}^2 = \mathbb{E} g_{X,k}^2(X_{d-k}) \), where the function \( g_{X,k}(p,v) \) is defined in (4.27). Since the random vectors \( X_i = (Q_i, U_i) \), \( i = 1, \ldots, d - k \), are iid with independent components, where \( Q_i \) is uniformly distributed on \([-r, r] \) and where \( U_i \) has the uniform spherical orientation distribution \( \Theta \) on \( S_{d-1}^+ \), the function \( g_{X,k}(p,u) \) may be written in the form (Chapter 5 of [111])

\[
g_{X,k}(p,u) = \left( \prod_{i=1}^{d-k-1} \int_{S_{d-1}^+} \int_\mathbb{R} \chi(H(p_i, u_i) \cap B_1^d) \, dp_i \, \tilde{\Theta}(du_i) \right)^{-1} \times \int_{S_{d-1}^+} \cdots \int_{S_{d-1}^+} \int_\mathbb{R} \chi(\bigcap_{j=1}^{d-k-1} H(p_j, u_j)) \cap H(p,u) \cap B_1^d \, dp_1 \tilde{\Theta}(du_1) \cdots dp_{d-k-1} \tilde{\Theta}(du_{d-k-1}) .
\]

Since \( \int_\mathbb{R} \chi(B_r \cap H((p,u))) \, dp = 2r \), we obtain that \( \prod_{i=1}^{d-k-1} \int_{S_{d-1}^+} 2r \tilde{\Theta}(du) = (2r)^{d-k-1} \) and hence that

\[
g_{X,k}(p,u) = \frac{1}{(2r)^{d-k-1}} \int_{S_{d-1}^+} \cdots \int_{S_{d-1}^+} \int_\mathbb{R} \chi(\bigcap_{j=1}^{d-k-1} H(p_j, u_j)) \cap H(p,u) \cap B_1^d \, dp_1 \tilde{\Theta}(du_1) \cdots dp_{d-k-1} \tilde{\Theta}(du_{d-k-1}) .
\]

A closed expression for \( g_{X,k}(p,u) \) is obtained by a \((d-k-1)\)-fold application of Crofton’s formula, cf. Formula (2.9). Indeed, applying (2.9) successively for \( j = 0, \ldots, d - k - 1 \) we get that

\[
g_{X,k}(p,u) = \frac{(d-k-1)!}{(2r)^{d-k-1}} \left( \frac{\kappa_{d-1}}{\kappa_d} \right)^{d-k-1} V_{d-k-1}(H(p,u) \cap B_1^d) ,
\]

where \( V_{d-k-1}(H(p,u) \cap B_1^d) \) is the \((d-k-1)\)-th intrinsic volume of \( H(p,u) \cap B_1^d \). Since \( H(p,u) \cap B_1^d \) is a \((d-1)\)-dimensional ball in \( \mathbb{R}^d \) with radius \( \sqrt{r^2 - p^2} \), Chapter 4.5 of [110], we get that

\[
V_{d-k-1}(H(p,u) \cap B_1^d) = V_{d-k-1}(B_{r^2-p^2}^{d-1}) .
\]

The invariance and homogeneity properties (Satz 2.2.2 in [111] or Theorem A.17) of \( V_{d-k-1} \) yield that

\[
V_{d-k-1}(B_{r^2-p^2}^{d-1}) = V_{d-k-1}(\sqrt{r^2 - p^2} B_1^{d-1}) = (r^2 - p^2)^{(d-k-1)/2} V_{d-k-1}(B_1^{d-1}) ,
\]
and since \( V_j(B_1^{d-1}) = (d-1)^j \frac{\kappa_{d-1}}{\kappa_d} \) for \( j = 0, \ldots, d \), p. 79 in [111], we finally obtain that
\[
V_{d-k-1}(H(p, u) \cap B_d^d) = (r^2 - p^2)^{(d-k-1)/2} \left( \frac{d-1}{k} \right) \frac{\kappa_{d-1}}{\kappa_d}.
\]
Summarizing the above steps we arrive at
\[
g_{\chi,k}((p, u)) = \frac{\kappa_{d-k-1}}{2d-k-1} \frac{d!}{k! \kappa_k} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{d-k} \left( 1 - \frac{p^2}{r^2} \right)^{(d-k-1)/2} \tag{4.42}
\]
for \((p, u) \in [-r, r] \times S^{d-1} \) and
\[
\sigma_{\chi,k}^{(1,d-k)} = \frac{1}{2r} \int_{-r}^{r} \int_{S^{d-1}_+} g_{\chi,k}((p, u)) \Theta(du) \, dp
\]
\[
= \left( \frac{d!}{k! \kappa_k} \frac{\kappa_{d-k-1}}{2d-k-1} \right)^2 \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{2(d-k)} \int_0^1 (1 - p^2)^{d-k-1} \, dp.
\]
Finally, we get (4.40) by observing that
\[
\int_0^1 (1 - p^2)^s \, dp = \frac{\sqrt{\pi} \Gamma(s + 1)}{(2s + 1) \Gamma(s + 1/2)} = \frac{(s! \, 2^s)^2}{(2s + 1)!}, \quad s = 0, 1, \ldots \tag{4.43}
\]
To verify (4.41) we make use of Formula (6.1-10) on p. 164 in [73] for the second moment of \( \zeta_k(B_1^d) \),
\[
\mathbb{E} \zeta_k^2(B_1^d) = \sum_{j=0}^{d-k} \frac{d! \, (d-j)! \, r^{2d-j} \, \kappa_{2d-j} \, \kappa_d \, \kappa_{d-j}^3}{j! \, (d-k-j)!^2 \, \kappa_{2(d-j)} \, \kappa_k^2} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{2(d-k)-j}. \tag{4.44}
\]
From the second formula in (4.39) together with (2.12), it is seen that in (4.44) the summand for \( j = 0 \) equals \( (\mathbb{E} \zeta_k(B_1^d))^2 \) and therefore, in accordance with (4.32),
\[
\sigma_{\chi,k}^{(1,d-k)} = \frac{((d-k-1))!^2}{(2 \lambda)^{2d-2k-1}} \lim_{r \to \infty} \frac{\text{Var} \, \zeta_k(B_1^d)}{r^{2d-1}}
\]
\[
= \frac{d! \, (d-1)! \, \kappa_{2d-1} \, \kappa_d \, \kappa_{d-1}^3}{(k!)^2} \frac{\kappa_{d-1}^3}{\kappa_{2d-2} \, \kappa_k^2} \left( \frac{\kappa_{d-1}}{2 \kappa_d} \right)^{2d-2k-1}.
\]
Finally, we obtain (4.41) by taking into account the relation
\[
\frac{\kappa_{2d-1}}{\kappa_{2d-2}} = \frac{\Gamma(d) \sqrt{\pi}}{\Gamma(d + 1/2)} = \frac{2^{2d-1} \, ((d-1)!)^2}{(2d-1)!},
\]
where the first equality is clear by the definition of \( \kappa_{2d-1} \) and \( \kappa_{2d-2} \) in (A.3), and where the second equality follows from Legendre’s duplication formula given by
\[
g^{2s-1} \, \Gamma \left( s + \frac{1}{2} \right) \, \Gamma(s) = \sqrt{\pi} \, \Gamma(2s), \tag{4.45}
\]
applied to the integer \( s = d \). This completes the proof of Theorem 4.4. \( \square \)
4.2. CLTs for Flat Processes induced by Poisson Hyperplanes

4.2.3 Concluding Remarks

Before we close Section 4.2 about CLTs for flat processes, several remarks have to be made, especially in view of the explicit moment formulae obtained in Theorem 4.4.

First, note that (4.29), (4.30), and (4.39) yield, for \( k = 0, 1, \ldots, d-1 \), the simple relationships

\[
\mathbb{P}(H(X_1) \cap \cdots \cap H(X_{d-k}) \cap B_r^d \neq \emptyset) = \frac{(d-k)!}{(2\lambda)^{d-k}} \lambda_k \kappa_{d-k}
\]

\[
\mathbb{E} \nu_k(H(X_1) \cap \cdots \cap H(X_{d-k}) \cap B_r^d) = \frac{(d-k)!}{(2\lambda)^{d-k}} \kappa_k \kappa_{d-k} r^k.
\]

In case \( \Phi_{d-1} \) is additionally isotropic, one can use (2.12) to get the explicit expressions

\[
\mathbb{P}(H(X_1) \cap \cdots \cap H(X_{d-k}) \cap B_r^d \neq \emptyset) = \frac{2^{k-d} d!}{k!} \kappa_k (\frac{\kappa_{d-1}}{d \kappa_d})^{d-k} \left( 1 - \frac{p^2}{r^2} \right)^{(d-1)/2}.
\]

These formulae comply with results on p. 160/161 in [73].

If in the proof of (4.40) we replace the function \( g_{\chi,k}(\cdot) \) by \( g_{\nu,k}(\cdot) \) defined in (4.28) (with \( V_k = \nu_k \) instead of \( V_0 = \chi \)) it is easily seen that

\[
g_{\nu,k}(p, u) = \frac{\kappa_{d-1}}{2^{d-k-1} k!} \frac{d! \kappa_d}{\kappa_k} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{d-k} \left( 1 - \frac{p^2}{r^2} \right)^{(d-1)/2} \tag{4.46}
\]

for \((p, u) \in [-r, r] \times S_{d-1}^d\) after iterated application of Crofton’s formula, cf. Formula (2.9), which confirms once more (4.41) without using Matheron’s formula (4.44).

However, in case \( \Phi_{d-1} \) is stationary but not necessarily isotropic, formulae for the functions \( g_{\chi,k}(\cdot) \) and \( g_{\nu,k}(\cdot) \) defined in (4.27) and (4.28), respectively, can also be obtained, which then in turn allow for quite explicit expressions of \( \sigma_{\chi,k}^{(1,d-k)} \) and \( \sigma_{\nu,k}^{(1,d-k)} \) for \( k = 0, \ldots, d-1 \) in terms of the Steiner convex set \( \Pi(\overline{\Theta}) \) associated with the spherical orientation distribution \( \Theta \) of \( \Phi_{d-1} \); cf. Section 2.5.2. The following theorem summarizes results obtained in [54] to where we also refer for the proof.

**Theorem 4.5** Let \( \Phi_{d-1} \) be a stationary non-degenerate Poisson hyperplane process with (finite) intensity \( \lambda > 0 \) and spherical orientation distribution \( \Theta \) with associated zonoid \( \Pi(\Theta) \). Then it holds for \( k = 0, 1, \ldots, d-1 \) and \((p, u) \in [-r, r] \times S_{d-1}^d\) that

\[
g_{\chi,k}(p, u) = \frac{(d-k-1)! \kappa_{d-k-1}}{2^{d-k-1}} \left( 1 - \frac{p^2}{r^2} \right)^{(d-k-1)/2} V_{d-k-1}(\Pi'(\overline{\Theta})) \tag{4.47}
\]

and

\[
g_{\nu,k}(p, u) = \frac{(d-k-1)! \kappa_{d-1}}{2^{d-k-1}} \left( 1 - \frac{p^2}{r^2} \right)^{(d-1)/2} V_{d-k-1}(\Pi'(\overline{\Theta})) \tag{4.48}
\]
where \( C^u \) denotes the image of \( C \in C' \) under orthogonal projection onto \( H(0, u) \) and \( V_j^{(d-1)}(C^u) \) is the \( j \)th intrinsic volume of the non-empty convex body \( C^u \) in the \((d-1)\)-dimensional Euclidean space \( \mathbb{R}^{d-1} \).

Finally, regarding once again the second-order moment formula (4.44) for \( k = 0 \) and subtracting \((E \eta_0(B^d_r))^2 \) or \((E \zeta_0(B^d_r))^2 \), where \( E \eta_0(B^d_r) = E \zeta_0(B^d_r) = \lambda_0 \kappa_d r^d \) by Formula (4.39), we get, in case of isotropy of \( \Phi_{d-1} \), an explicit expression for \( \text{Var} \eta_0(B^d_r) \). The latter explicit expression is a sum that is similar to the sum on the right hand side of (4.44), however with summation from \( j = 1 \) to \( d \), and that can be interpreted as a sum of monomials in \( \lambda r \). This allows to determine the pair correlation function \( g_0(r) \), cf. Section 2.2.4, of the stationary and isotropic point process \( \eta_0 \) as a polynomial of degree \( d - 1 \) in \((\lambda r)^{-1} \). More precisely, putting \( g_0(r) = 1 + \sum_{j=1}^{d-1} c_{dj} (\lambda r)^{-j} \) and utilizing the relationship

\[
\text{Var} \eta_0(B^d_r) = d \kappa_d \lambda_0^2 \int_0^{2r} \nu_d(B^d_r \cap (B^d_r + (u, 0, \ldots, 0))) (g_0(u) - 1) u^{d-1} du + \lambda_0 \kappa_d r^d,
\]

p. 131 in [116], we get that \( c_{dj} = \binom{d-1}{j} \left( \frac{\kappa_{d-j}}{\kappa_d} \right)^2 \left( \frac{d \kappa_d}{\kappa_{d-1}} \right)^j \) by comparison of coefficients. Here, we used (2.12) for \( k = 0 \) together with

\[
\int_0^2 \nu_d(B^d_r \cap (B^d_r + (u, 0, \ldots, 0))) u^{d-j-1} du = \frac{1}{d \kappa_d} \int_{B^d_r} \int_{B^d_r} \frac{dx \, dy}{\|x - y\|^j} = \frac{2^{d-j+1} \kappa_{2d-j}}{(d-j)(d-j+1) \kappa_{d-j}}
\]

for \( j = 1, \ldots, d-1 \); cf. p. 177 in [111].

### 4.3 Multivariate Extensions

In this section we extend the results of the preceding Section 4.2 by establishing multivariate CLTs which describe the joint asymptotic behavior (as \( r \to \infty \)) of the closely correlated random variables \( Z^{(d)}_{0,r}(\chi), \ldots, Z^{(d)}_{d-1,r}(\chi) \) as well as of \( Z^{(d)}_{0,r}(\nu), \ldots, Z^{(d)}_{d-1,r}(\nu) \), where we recall that

\[
Z^{(d)}_{k,r}(\chi) = \frac{(d-k-1)!}{(2 \lambda r)^{d-k-1/2}} \left( \eta_k(B^d_r) - E \eta_k(B^d_r) \right), \quad k = 0, \ldots, d-1
\]

and

\[
Z^{(d)}_{k,r}(\nu) = \frac{(d-k-1)!}{(2 \lambda r)^{d-k-1/2}} \left( \zeta_k(B^d_r) - E \zeta_k(B^d_r) \right), \quad k = 0, \ldots, d-1,
\]

as already defined in (4.35) and (4.36), respectively. To begin with, analogously to (4.25) and (4.26) in Section 4.2.1, we define the mixed second moments

\[
\sigma^{(1, d-k, d-l)}_{\chi, k, l} = E \left( \chi \left( \cap_{i=1}^{d-k} H(X_i) \cap B^d_r \right) \chi \left( \cap_{i=d-k-l+1}^{2d-k-l-1} H(X_i) \cap B^d_r \right) \right)
\]

\[
= E(g_{\chi,k}(X_{d-k})g_{\chi,l}(X_{d-k})) \quad (4.49)
\]
and
\[
\sigma^{(1,d-k,d-l)}_{\nu,k,l} = r^{-(k+l)} \mathbb{E} \left( \nu_k \left( \bigcap_{i=1}^{d-k} H(X_i) \cap B^d_r \right) \nu_l \left( \bigcap_{i=d-k}^{2d-k-l-1} H(X_i) \cap B^d_r \right) \right) \\
= \mathbb{E} (g_{\nu,k}(X_{d-k}) \nu_l(X_{d-k})) \tag{4.50}
\]
for \(k, l = 0, \ldots, d - 1\), where the functions \(g_{\chi,k}(\cdot)\) and \(g_{\nu,l}(\cdot)\) are given by (4.27) and (4.28), respectively.

### 4.3.1 Multivariate CLTs and Moment Formulae

The following result generalizes the second part of Lemma 4.8.

**Lemma 4.9** For any \(k, l = 0, \ldots, d - 1\),
\[
\lim_{r \to \infty} \mathbb{E}(Z^{(d)}_{k,r}(\chi) Z^{(d)}_{l,r}(\chi)) = \sigma^{(1,d-k,d-l)}_{\chi,k,l} \tag{4.51}
\]
and
\[
\lim_{r \to \infty} \mathbb{E}(Z^{(d)}_{k,r}(\nu) Z^{(d)}_{l,r}(\nu)) = \sigma^{(1,d-k,d-l)}_{\nu,k,l}. \tag{4.52}
\]

**Proof** The relations (4.51) and (4.52) can be shown by similar combinatorial arguments as employed in the proof of Lemma 4.1 to verify (4.5). Using the definition of \(Z^{(d)}_{k,r}(\chi)\) as given in (4.35) we easily get that
\[
\mathbb{E}(Z^{(d)}_{k,r}(\chi) Z^{(d)}_{l,r}(\chi)) = \frac{(d - k - 1)! (d - l - 1)!}{(2 \lambda r)^{2d-k-l-1}} \times \left( \mathbb{E}(\eta_k(B^d_r)) \eta_l(B^d_r) - \mathbb{E} \eta_k(B^d_r) \mathbb{E} \eta_l(B^d_r) \right).
\]

According to (4.23), we may write for \(0 \leq k \leq l \leq d - 1\) that
\[
\mathbb{E}(\eta_k(B^d_r) \eta_l(B^d_r)) = \sum_{j=0}^{d-l} \frac{\mathbb{E}(N_r(N_r - 1) \cdots (N_r - (2d - k - l - j + 1)))}{j! (d - k - j)! (d - l - j)!} \times \mathbb{E} \left( \chi(\bigcap_{p=1}^{d-k} H(X_p) \cap B^d_r) \chi(\bigcap_{q=d-k-l-j+1}^{2d-k-l-j} H(X_q) \cap B^d_r) \right).
\]

Using (4.6) and the fact that the summand for \(j = 0\) in the sum on the right hand side of the latter equation equals \(\mathbb{E} \eta_k(B^d_r) \mathbb{E} \eta_l(B^d_r)\) we arrive at
\[
\mathbb{E}(Z^{(d)}_{k,r}(\chi) Z^{(d)}_{l,r}(\chi)) = \frac{(d - k - 1)! (d - l - 1)!}{(2 \lambda r)^{2d-k-l-1}} \sum_{j=1}^{d-l} \frac{(2 \lambda r)^{2d-k-l-j}}{j! (d - k - j)! (d - l - j)!} \times \mathbb{E} \left( \chi(\bigcap_{p=1}^{d-k} H(X_p) \cap B^d_r) \chi(\bigcap_{q=d-k-l-j+1}^{2d-k-l-j} H(X_q) \cap B^d_r) \right).
\]

Hence, dividing by \(r^{2d-k-l-1}\) and letting \(r \to \infty\) immediately yields (4.51). The proof of (4.52) is completely analogous. \(\square\)
Theorem 4.6 Let the assumptions of Theorem 4.3 be satisfied. Then,

\[ (Z_{k,r}^{(d)}(\chi))_{k=0}^{d-1} \xrightarrow{\text{d}} \mathcal{N}(0, \Sigma(\chi)) \]  

and

\[ (Z_{k,r}^{(d)}(\nu))_{k=0}^{d-1} \xrightarrow{\text{d}} \mathcal{N}(0, \Sigma(\nu)) , \]  

where \( \mathcal{N}(0, \Sigma(\chi)) \) and \( \mathcal{N}(0, \Sigma(\nu)) \) are \( d \)-dimensional Gaussian random vectors with mean zero and covariance matrix \( \Sigma(\chi) = (\sigma_{\chi, k, l})_{k,l=0}^{d-1} \) and \( \Sigma(\nu) = (\sigma_{\nu, k, l})_{k,l=0}^{d-1} \), respectively.

The entries of \( \Sigma(\chi) \) and \( \Sigma(\nu) \) are given by the limits in (4.51) and in (4.52), respectively.

Proof We first recall that due to the well-known Cramér–Wold device, cf. Theorem A.10, the multivariate CLT (4.53) is equivalent to the one-dimensional CLT

\[ \sum_{k=0}^{d-1} t_k Z_{k,r}^{(d)}(\chi) \xrightarrow{\text{d}} \mathcal{N}(0, t^T \Sigma(\chi)t) \]  

for any \( t = (t_0, \ldots, t_{d-1})^T \in \mathbb{R}^d \setminus \{0\} \). Hence, the proof of (4.53) can be put down to the case of the (one-dimensional) CLTs considered in Theorems 4.1 and 4.3. First, using (4.35) together with (4.27) and (4.29) we are able to rewrite the linear combination \( \sum_{k=0}^{d-1} t_k Z_{k,r}^{(d)}(\chi) \), i.e.,

\[ \sum_{k=0}^{d-1} t_k Z_{k,r}^{(d)}(\chi) = \sum_{k=0}^{d-1} t_k \frac{(d-k-1)!}{(2\lambda r)^{d-k-1/2}} (\eta_k(B_{r}^d) - \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{E} g_{\chi,k}(X_1)) \]

Using the definition of a \( U \)-statistic given in (A.15), together with (4.23) we get that

\[ \sum_{k=0}^{d-1} t_k Z_{k,r}^{(d)}(\chi) = \sum_{k=0}^{d-1} t_k \frac{(d-k-1)!}{(2\lambda r)^{d-k-1/2}} \times \left( U_{N_r}^{(d-k)}(\chi) \left( \frac{N_r}{d-k} \right) - \frac{(2\lambda r)^{d-k}}{(d-k)!} \mathbb{E} g_{\chi,k}(X_1) \right) \]

where for any fixed \( N_r = n \geq d-k \), the random multiple sum \( U_{N_r}^{(d-k)}(\chi) \) given in (4.37) is a \( U \)-statistic with kernel function \( \chi_{(d-k)} \mathcal{H}(X_j) \cap B_{r}^d \). Next, we apply Hoeffding’s decomposition, cf. Formula (A.16), to \( U_{N_r}^{(d-k)}(\chi) \) and proceed in the same manner as in the proof of Theorem 4.1. In view of the limiting relations (4.13) and (4.11) in Lemma 4.4 and Lemma 4.3, respectively, with \( d \) replaced by \( d-k \) for \( k = 0, 1, \ldots, d-1 \) and combined with Slutsky’s theorem (cf. Theorem A.8) we recognize that the weak limit of \( \sum_{k=0}^{d-1} t_k Z_{k,r}^{(d)}(\chi) \) coincides with that of

\[ \sum_{k=0}^{d-1} t_k \frac{(d-k-1)!}{(2\lambda r)^{d-k-1/2}} \left( \frac{N_r}{d-k} - \frac{1}{d-k-1} \left( \sum_{i=1}^{N_r} g_{\chi,k}(X_i) - 2\lambda r \mathbb{E} g_{\chi,k}(X_1) \right) \right) \]

as \( r \to \infty \). Finally, by means of (4.14) with \( d \) again replaced by \( d-k \) for \( k = 0, 1, \ldots, d-1 \) it follows that

\[ \frac{1}{r^{d-k-1}} \left( \frac{N_r}{d-k} - \frac{1}{d-k-1} \right) \xrightarrow{\text{P}} \frac{(2\lambda)^{d-k-1}}{(d-k-1)!} \]
and hence, it remains to show that
\[
\frac{1}{\sqrt{2\lambda r}} \sum_{k=0}^{d-1} t_k \left( \sum_{i=1}^{N_r} g_{\chi,k}(X_i) - 2 \lambda r \mathbb{E}g_{\chi,k}(X_1) \right)
\]
\[\xrightarrow{d, r \to \infty} \mathcal{N}\left( 0, \sum_{k,l=0}^{d-1} t_k t_l \mathbb{E}(g_{\chi,k}(X_1)g_{\chi,l}(X_1)) \right).\]

However, the latter CLT is obtained by proving Lemma 4.6 once more for the function \(\sum_{k=0}^{d-1} t_k g_{\chi,k}()\) instead of \(g_B()\). To show the second assertion (4.54) one only needs to repeat the just finished proof with the kernel function \(r^{-k} \nu_k(\cap_{j=1}^{d-k} H(X_j) \cap B_d^1)\) and with \(g_{\nu_1}()\) instead of \(g_{\chi,1}()\). This completes the proof of Theorem 4.6. \(\Box\)

### 4.3.2 Explicit Formulae for Asymptotic Covariance Matrices

As in Lemma 4.4, the additional assumption of isotropy allows to compute explicit formulae for the entries \(\sigma_{\chi,k,l}^{(1,d-k,d-l)}\) and \(\sigma_{\nu,1}^{(1,d-k,d-l)}\), introduced in (4.49) and (4.50), of the asymptotic covariance matrices \(\Sigma(\chi)\) and \(\Sigma(\nu)\), respectively.

**Lemma 4.10** Let \(\Phi_{d-1}\) be a stationary and isotropic Poisson hyperplane process. Then for \(k, l = 0, \ldots, d - 1\),
\[
\sigma_{\chi,k,l}^{(1,d-k,d-l)} = \frac{(d! \kappa_d)^2}{k! l! \kappa_k \kappa_l} \frac{\kappa_{d-k-l-1} \kappa_{2d-k-l-1}}{\kappa_{2d-k-l-2}} \left( \frac{k_{d-1}}{d \kappa_d} \right)^{2d-k-l}
\]
(4.56)
\[
= \sqrt{\sigma_{\chi,k}^{(1,d-k)}} \sigma_{\chi,l}^{(1,d-l)} \frac{B\left(\frac{2d-k-l}{2}, \frac{2d-k-l}{2}\right)}{\sqrt{B(d-k, d-k)B(d-l, d-l)}},
\]
where \(\sigma_{\chi,k}^{(1,d-k)}\) is given in (4.25) for \(k = 0, \ldots, d-1\), and where \(B(s, t) = \int_0^1 x^{s-1}(1-x)^{t-1} dx = \Gamma(s)\Gamma(t)/\Gamma(s+t)\) denotes Euler’s Beta function. Furthermore,
\[
\sigma_{\nu,k,l}^{(1,d-k,d-l)} = \frac{(\kappa_d \kappa_{d-1} d! (d-1)!)^2 2^{k+l}}{k! l! \kappa_k \kappa_l (2d-1)!} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{2d-k-l}
\]
(4.57)
\[
= \sqrt{\sigma_{\nu,k}^{(1,d-k)}} \sigma_{\nu,l}^{(1,d-l)} ,
\]
where \(\sigma_{\nu,k}^{(1,d-k)}\) is given in (4.26) for \(k = 0, \ldots, d-1\).

**Proof** Let \(k, l \in \{0, \ldots, d-1\}\). Both (4.56) and (4.57) can be obtained using the shape of the functions \(g_{\chi,k}()\) and \(g_{\nu,k}()\) derived in the proof of Theorem 4.4. By (4.42) we have that
\[
g_{\chi,k}(p, u) = \frac{\kappa_{d-k-1} d! \kappa_d}{k! \kappa_k} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{d-k} \left(1 - \frac{p^2}{r^2}\right)^{(d-k-1)/2},
\]
and together with the definition of $\sigma^{(1,d-k,d-l)}_{\chi,k,l}$ in (4.49), we get that

$$
\sigma^{(1,d-k,d-l)}_{\chi,k,l} = \frac{1}{2r} \int_{-r}^{r} \int_{S_{2d-1}} g_{\chi,k}((p,u)) g_{\chi,l}((p,u)) \tilde{\Theta}(du) dp
$$

$$
= \frac{d!\kappa_d \kappa_{d-1}}{k! \kappa_k 2^d} \frac{d!\kappa_d \kappa_{d-1}}{l! \kappa_l 2^d} \left( \frac{\kappa_{d-1}d}{d \kappa_d} \right)^{2d-k-l} \int_{0}^{1} (1-p^2)^{2d-k-l-2} dp .
$$

Thus, by noting that

$$
\int_{0}^{1} (1-p^2)^{2d-k-l-2} dp = \frac{\kappa_{2d-k-l-1}}{2 \kappa_{2d-k-l-2}} = 2^{2d-k-l-2} B\left( \frac{2d-k-l}{2}, \frac{2d-k-l}{2} \right),
$$

we obtain the first part of (4.56), where the second identity in the previous equation turns out to be a simple consequence of (4.45) for $s = (2d-k-l)/2$ and the definition of the Beta function. The second part of (4.56) is seen by inserting the variances $\sigma^{(1,d-k)}_{\chi,k}$ given by (4.40) combined with $B(d-k,d-k) = ((d-k-1)!/(2d-2k)!$, Likewise, using (4.46), we get that

$$
\sigma^{(1,d-k,d-l)}_{\nu,k,l} = \frac{1}{2r} \int_{-r}^{r} \int_{S_{2d-1}} g_{\nu,k}((p,u)) g_{\nu,l}((p,u)) \tilde{\Theta}(du) dp
$$

$$
= \frac{d!\kappa_d \kappa_{d-1}}{k! \kappa_k 2^d} \frac{d!\kappa_d \kappa_{d-1}}{l! \kappa_l 2^d} \left( \frac{\kappa_{d-1}d}{d \kappa_d} \right)^{2d-k-l} \int_{0}^{1} (1-p^2)^{d-k-1} dp .
$$

Hence, taking (4.43) for $s = d - 1$, the first part of (4.57) is shown and the second equality is immediately seen from Theorem 4.4.

Note that if we use the formulae for $g_{\chi,k}(\cdot)$ and $g_{\nu,k}(\cdot)$ given by (4.47) and (4.48), respectively, for $k = 0, \ldots, d-1$ in Theorem 4.5, it is possible to obtain also rather explicit expressions for the entries of the covariance matrices $\Sigma(\chi)$ and $\Sigma(\nu)$, respectively, in the case where $\Phi_{d-1}$ is stationary but not necessarily isotropic.

To conclude, we examine the covariance matrices $\Sigma(\chi)$ and $\Sigma(\nu)$ appearing in Theorem 4.6 and get the following corollary.

**Corollary 4.11** The covariance matrix $\Sigma(\chi)$ always possesses full rank $d$, whereas the rank of the covariance matrix $\Sigma(\nu)$ equals 1 for any dimension $d \geq 1$. Moreover,

$$
\frac{Z^{(d)}_{r}(\nu)}{\sqrt{\sigma^{(1,d-l)}_{\nu,k,l}}} = \frac{Z^{(d)}_{k,l}(\nu)}{\sqrt{\sigma^{(1,d-l)}_{\nu,k,l}}} \xrightarrow{P} 0 \quad \text{for} \quad 0 \leq k < l \leq d-1 .
$$

**Proof** Note that $\Sigma(\chi)$ possesses full rank if this matrix with entries given in (4.49) (and in (4.56)) is strictly positive. This however can be seen by the fact that

$$
\sum_{k,l=0}^{d-1} t_k t_l B\left( \frac{2d-k-l}{2}, \frac{2d-k-l}{2} \right) = \int_{0}^{1} \left( \sum_{k=0}^{d-1} t_k \left( x(1-x) \right)^{\frac{d-k-1}{2}} \right)^2 dx > 0
$$
for any \((t_0, \ldots, t_{d-1})^\top \in \mathbb{R}^d \setminus \{0\}\), which means that the symmetric matrix with entries \(B((2d-k-l)/2, (2d-k-l)/2)\) for \(k, l = 0, \ldots, d - 1\) is strictly positive. In order to show that the rank of \(\Sigma(\nu)\) is 1 for any \(d \geq 1\), we only need to observe that the second equality in (4.57) implies that each entry of the asymptotic covariance matrix of the normalized random vector 
\[
\left(\frac{Z_{k,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,k}}} \right)_{k=0}^{d-1}
\]
equals 1. Moreover, for any \(k, l \in \mathbb{N}\) with \(0 \leq k < l \leq d - 1\) consider

\[
\mathbb{E} \left( \frac{Z_{l,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,l}}} - \frac{Z_{k,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,k}}} \right)^2 = \text{Var} \left( \frac{Z_{l,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,l}}} - \frac{Z_{k,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,k}}} \right)
\]

\[
= \frac{1}{{\sigma_{\nu,l}^{(1,d-l)}}} \mathbb{E} \left( Z_{l,r}^{(d)}(\nu) \right)^2 - \frac{2}{{\sigma_{\nu,l}^{(1,d-l)}}{\sigma_{\nu,k}^{(1,d-k)}}} \mathbb{E} \left( Z_{l,r}^{(d)}(\nu) Z_{k,r}^{(d)}(\nu) \right) + \frac{1}{{\sigma_{\nu,k}^{(1,d-k)}}} \mathbb{E} \left( Z_{k,r}^{(d)}(\nu) \right)^2.
\]

Note that

\[
\mathbb{E} \left( Z_{l,r}^{(d)}(\nu) \right)^2 = \text{Var} Z_{l,r}^{(d)}(\nu) \xrightarrow{r \to \infty} {\sigma_{\nu,l}^{(1,d-l)}} ,
\]

and that by (4.51) together with the second equality in (4.57),

\[
\mathbb{E} \left( Z_{l,r}^{(d)}(\nu) Z_{k,r}^{(d)}(\nu) \right) \xrightarrow{r \to \infty} \sqrt{\frac{\sigma_{\nu,l}^{(1,d-l)}}{\sigma_{\nu,k}^{(1,d-k)}}}.
\]

Putting things together, the latter two limiting relations yield that

\[
\mathbb{E} \left( \frac{Z_{l,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,l}^{(1,d-l)}}} - \frac{Z_{k,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,k}^{(1,d-k)}}} \right)^2 \xrightarrow{r \to \infty} 0.
\]

Hence, (4.58) is established. \(\square\)

In view of the latter proof, the somewhat surprising result (4.58) can be interpreted in the sense that the variance of the difference of any two components of \(\left(\frac{Z_{k,r}^{(d)}(\nu)}{\sqrt{\sigma_{\nu,k}^{(1,d-k)}}}\right)_{k=0}^{d-1}\) tends to zero as \(r \to \infty\). Together with Slutsky’s theorem (cf. Theorem A.8), this allows for the conclusion that the normal convergence in Theorem 4.3 for a single component, \(Z_{0,r}^{(d)}(\nu)\) say, implies asymptotic normality of the other components. Thus, relation (4.58) can be interpreted as a kind of asymptotic second-order relationship for \(k\)-flat processes induced by stationary and isotropic Poisson hyperplane processes.

Note that if we drop the assumption of isotropy of \(\Phi_{d-1}\), the following theorem, provided in [54], examines the rank of the covariance matrices \(\Sigma(\chi)\) and \(\Sigma(\nu)\) in a more general context.

**Theorem 4.7** Let \(\Phi_{d-1}\) be a stationary Poisson hyperplane process with (finite) intensity \(\lambda > 0\) and with spherical orientation distribution \(\tilde{\Theta}\).

(i) The covariance matrix \(\Sigma(\chi)\) possesses full rank \(d\) if and only if \(\tilde{\Theta}\) is non-degenerate.
\[
\tilde{\Theta}\left(\left\{ u \in S_{d+1}^{-1} : \sum_{j=0}^{d-1} t_j V_j^{(d-1)}(\Pi u(\tilde{\Theta})) = 0 \right\} \right) = 1,
\]

where \( C^u \) denotes the image of \( C \in \mathcal{C} \) under orthogonal projection onto \( H(0, u) \) and \( V_j^{(d-1)}(C^u) \) is the \( j \)th intrinsic volume of the non-empty convex body \( C^u \) in the \( (d-1) \)-dimensional Euclidean space \( \mathbb{R}^{d-1} \).

(iii) The covariance matrix \( \Sigma(\nu) \) possesses rank 1 if and only if
\[
\tilde{\Theta}\left(\left\{ u \in S_{d+1}^{-1} : V_{k-1}^{(d-1)}(\Pi u(\tilde{\Theta})) = k V_k(\Pi(\tilde{\Theta})) \right\} \right) = 1, \quad k = 2, \ldots, d.
\]
Chapter 5

Statistical Inference for Tessellation Intensities

In this chapter we review and extend the results of Chapter 4 under the view point of statistical applications and statistical analysis of intensities of $k$-flat processes induced by intersections of $d-k$ hyperplanes; $k = 0, \ldots, d - 1$. More precisely, we use the CLTs in the Theorems 4.1 and 4.3, as well as their multivariate extensions in Theorem 4.6 to derive (asymptotically exact) confidence intervals (CIs) as well as asymptotic tests for the intensities $\lambda_k$ of $k$-flat intersection processes $\Phi_k$ induced by some stationary and isotropic Poisson hyperplane process (PHP) $\Phi_{d-1}$ with finite intensity $\lambda > 0$; $k = 0, \ldots, d - 1$. The statistical inference is based on a single realization of $\Phi_k$ observed through a spherical sampling window $B_r^d$.

5.1 Estimators and Tests for $k$-Flat Intensities

We start our presentation by regarding two unbiased and (strongly) consistent estimators for the intersection intensities $\lambda_k$, based on the statistics $\eta_k$ and $\zeta_k$ defined in (4.23) and (4.24), respectively, i.e.,

\[ \eta_k(B_r^d) = \Phi_k(\{L \in A_k^d : L \cap B_r^d \neq \emptyset\}) \quad \text{and} \quad \zeta_k(B_r^d) = \sum_{L \in \text{supp } \Phi_k} \nu_k(B_r^d \cap L), \]

where $k = 0, \ldots, d - 1$.

5.1.1 Definition; Unbiasedness and Consistency

For any $k = 0, \ldots, d - 1$ consider the estimators

\[
\hat{\lambda}_{k,r} = \frac{\eta_k(B_r^d)}{\nu_{d-k}(B_r^{d-k})} \quad \text{and} \quad \tilde{\lambda}_{k,r} = \frac{\zeta_k(B_r^d)}{\nu_d(B_r^d)} \tag{5.1}
\]

for the intensity $\lambda_k$. 
Lemma 5.1 Let $\Phi_{d-1}$ be a stationary non-degenerate PHP with (finite) intensity $\lambda > 0$. Then for $k = 0, \ldots, d-1$ the estimators $\hat{\lambda}_{k,r}$ and $\bar{\lambda}_{k,r}$ defined in (5.1) are unbiased estimators for the intensity $\lambda_k$ of the stationary $k$-flat intersection process $\Phi_k$. Moreover, both $\hat{\lambda}_{k,r}$ and $\bar{\lambda}_{k,r}$ are strongly consistent.

Proof Let $k \in \{0, \ldots, d-1\}$. Then, applying the mean value formulae in (4.39), the unbiasedness of both $\hat{\lambda}_{k,r}$ and $\bar{\lambda}_{k,r}$ is obtained immediately. Since $\Phi_{d-1}$ is ergodic and even mixing (Chapter 6.4 of [112]) the estimators $\hat{\lambda}_{k,r}$ and $\bar{\lambda}_{k,r}$ are strongly consistent.

Recall formulae (4.33) and (4.34) for the second moments $E\eta_k^2(B_r^d)$ and $E\zeta_k^2(B_r^d)$, where the expectations of the functions $\chi(\cdot)$ and $\nu_k(\cdot)$ involved in the latter formulae do not depend on the radius $r > 0$. Then, it is possible to show that

$$E\bigg(\frac{\eta_k(B_r^d)}{\nu_{d-k}(B_r^{d-k})} - \lambda_k\bigg)^2 \xrightarrow{r \to \infty} 0 \quad \text{and} \quad E\bigg(\frac{\zeta_k(B_r^d)}{\nu_{d-k}(B_r^{d-k})} - \lambda_k\bigg)^2 \xrightarrow{r \to \infty} 0.$$ 

Note that the estimators $\hat{\lambda}_{k,r}$ and $\bar{\lambda}_{k,r}$ have been considered earlier in literature. For example in [9] and [90], stationary and isotropic Poisson line processes $\Phi_1$ in $\mathbb{R}^2$ are examined in a compact convex sampling window $W \subset \mathbb{R}^2$ and it is shown that $\text{Var} \hat{\lambda}_1(W) < \text{Var} \bar{\lambda}_1(W)$, where $\hat{\lambda}_1(W) = \frac{n}{\sum_{i=1}^n} \nu_0(\Phi_1 \cap \partial W)$ and $\bar{\lambda}_1(W) = \frac{1}{\nu_0(W)} \nu_1(\Phi_1 \cap W)$, respectively, with $\partial W$ denoting the boundary of $W$. In case $W = B_r^2$, it holds that $\hat{\lambda}_1(B_r^2) = \hat{\lambda}_{1,r}$ and $\bar{\lambda}_1(B_r^2) = \bar{\lambda}_{1,r}$, where $\hat{\lambda}_{1,r}$ and $\bar{\lambda}_{1,r}$ have been introduced in (5.1). In [105], the latter results are generalized to unbiased estimators for the intensity of stationary $k$-flats processes $(k = 1, \ldots, d-1)$ in $\mathbb{R}^d$ for specified types of orientation distributions. In particular, for example in the case of Poisson $k$-flats processes, sufficient conditions that the derived estimators are the uniformly best unbiased estimators are given and efficiency of estimators is examined. An overview of these and further estimation problems from the perspective of stereology can also be found in [10] and the references therein as well as in [121].

5.1.2 Asymptotic Second-Order Properties

In the following lemma we compare asymptotic variances of the estimators $\hat{\lambda}_{k,r}$ and $\bar{\lambda}_{k,r}$. Our results supplement results as described above at the end of the last section.

Lemma 5.2 Let $\Phi_{d-1}$ be a stationary and isotropic Poisson hyperplane process in $\mathbb{R}^d$ with intensity $\lambda > 0$. For $k = 0, \ldots, d-1$ and $j = 1, 2, \ldots$, let

$$a_{d,k} = \binom{d}{k} \frac{k_d}{\lambda_k} \left(\frac{k_{d-1}}{d \lambda_{d-k}}\right)^{d-k} \quad \text{and} \quad b_j = \frac{2^{j-1} (j-1)! k_{d-1}}{\sqrt{2} (2j-1)! \lambda^{j-1}}.$$  

(5.2)

Then, for $k = 0, \ldots, d-1$,

$$\lim_{r \to \infty} \frac{r}{4(d-k)^2} \text{Var} \hat{\lambda}_{k,r} = \lambda^{2d-2k-1} a_{d,k}^2 b_{d-k}^2 \quad \text{and} \quad \lim_{r \to \infty} \frac{r}{4(d-k)^2} \text{Var} \bar{\lambda}_{k,r} = \lambda^{2d-2k-1} a_{d,k}^2 b_{d-k}^2.$$  

(5.3)
5.1. Estimators and Tests for \(k\)-Flat Intensities

\[ \lim_{r \to \infty} r \Var \hat{\lambda}_{k,r} \leq \lim_{r \to \infty} r \Var \tilde{\lambda}_{k,r}. \]  
\(\text{(5.4)}\)

**Proof** Let \(k \in \{0, \ldots, d-1\} \). Since,

\[ \frac{r \Var \hat{\lambda}_{k,r}}{4(d-k)^2} = \frac{1}{4(d-k)^2 \kappa^2_{d-k} r^{2d-2k-1}} \Var \eta_k(B^d_r) \]

and

\[ \frac{r \Var \tilde{\lambda}_{k,r}}{4(d-k)^2} = \frac{1}{4(d-k)^2 \kappa^2_{d-k} r^{2d-1}} \Var \zeta_k(B^d_r), \]

we obtain the limit formulae in (5.3) by regarding (4.31) together with (4.40) and (4.32) together with (4.41), respectively. By means of Legendre’s duplication formula (4.45) one can verify the inequality \(b_j \leq b_{j+1}\) for \(j = 1, 2, \ldots\), which in turn implies (5.4). \(\Box\)

Note that by the definition of \(a_{d,k}\) in (5.2) and in view of (2.12), we get that

\[ \lambda_k = a_{d,k} \lambda^{d-k}, \quad k = 0, \ldots, d-1. \]  
\(\text{(5.5)}\)

5.1.3 Asymptotic Normality and Variance Stabilization

In this section, we reconsider Theorem 4.3 under the assumption of isotropy of the Poisson hyperplane process \(\Phi_{d-1}\) in order to show that the distribution of the estimators \(\hat{\lambda}_{k,r}\) and \(\tilde{\lambda}_{k,r}\) converges to normal mean zero limits as \(r \to \infty\). Indeed, we immediately obtain the following corollary of Theorem 4.3.

**Corollary 5.3** Let \(\Phi_{d-1}\) be a stationary and isotropic PHP with intensity \(\lambda > 0\). For \(k = 0, \ldots, d-1\) it holds that

\[ \sqrt{r} \left( \hat{\lambda}_{k,r} - \lambda_k \right) \xrightarrow{d} \mathcal{N} \left( 0, 4(d-k)^2 \lambda_k^{2-k} \frac{1}{a_{d,k}^2} b_{d-k}^2 \right) \]  
\(\text{(5.6)}\)

and

\[ \sqrt{r} \left( \tilde{\lambda}_{k,r} - \lambda_k \right) \xrightarrow{d} \mathcal{N} \left( 0, 4(d-k)^2 \lambda_k^{2-k} \frac{1}{a_{d,k}^2} b_d^2 \right). \]  
\(\text{(5.7)}\)

**Proof** Let \(k \in \{0, \ldots, d-1\} \). We recall from Theorem 4.3 that

\[ \frac{\sqrt{r}}{r^{d-k}} (\eta_k(B^d_r) - \mathbb{E} \eta_k(B^d_r)) \xrightarrow{d} \mathcal{N} \left( 0, \frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2} \sigma_{1,d-k} \right) \]

and

\[ \frac{\sqrt{r}}{r^{d}} (\zeta_k(B^d_r) - \mathbb{E} \zeta_k(B^d_r)) \xrightarrow{d} \mathcal{N} \left( 0, \frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2} \sigma_{d,k} \right). \]
where $\sigma_{\chi,k}^{(1,d-k)}$ and $\sigma_{\nu,k}^{(1,d-k)}$ are explicitly given by (4.40) and (4.41), respectively. Hence, together with the expectation formulae in (4.39), we get that

$$
\sqrt{r} \left( \hat{\lambda}_{k,r} - \lambda_k \right) \xrightarrow{d \to \infty} \mathcal{N}(0, \frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2 \kappa_{d-k}^2 \sigma_{\chi,k}^{(1,d-k)}})
$$

and

$$
\sqrt{r} \left( \hat{\lambda}_{k,r} - \lambda_k \right) \xrightarrow{d \to \infty} \mathcal{N}(0, \frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2 \kappa_d^2 \sigma_{\nu,k}^{(1,d-k)}}).
$$

Using the abbreviations $a_{d,k}$ for $k = 0, \ldots, d - 1$ and $b_j$ for $j = 1, 2, \ldots$ introduced previously in (5.2), we find, after some elementary manipulations, that

$$
\frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2 \kappa_{d-k}^2} = 4(d-k)^2 \lambda_k^{2-1/(d-k)} a_{d,k}^{1/(d-k)} b_{d-k}^2
$$

and

$$
\frac{(2\lambda)^{2d-2k-1}}{((d-k-1)!)^2 \kappa_d^2} = 4(d-k)^2 \lambda_k^{2-1/(d-k)} a_{d,k}^{1/(d-k)} b_d^2.
$$

\[\square\]

In order to construct asymptotic confidence intervals and asymptotic tests for the intensity $\lambda_k$ of the $k$-flat intersection process $\Phi_k$ for any $k = 0, \ldots, d - 1$, we apply a variance stabilizing transformation, cf. Section A.3.6, to the CLTs of Corollary 5.3. Then we obtain normal mean zero limits with variance one.

**Theorem 5.1** Let $\Phi_{d-1}$ be a stationary and isotropic PHP with intensity $\lambda > 0$. Then, for $k = 0, \ldots, d - 1$, it holds that

$$
\hat{Z}_{k,r}^{(d)} = \frac{\sqrt{r}}{b_{d-k}} a_{d,k}^{-\frac{1}{2(d-k)}} \left( \hat{\lambda}_{k,r}^{\frac{1}{2(d-k)}} - \lambda_k^{\frac{1}{2(d-k)}} \right) \xrightarrow{d \to \infty} \mathcal{N}(0,1)
$$

and

$$
\tilde{Z}_{k,r}^{(d)} = \frac{\sqrt{r}}{b_d} a_{d,k}^{-\frac{1}{2(d-k)}} \left( \hat{\lambda}_{k,r}^{\frac{1}{2(d-k)}} - \lambda_k^{\frac{1}{2(d-k)}} \right) \xrightarrow{d \to \infty} \mathcal{N}(0,1).
$$

**Proof** Let $k \in \{0, \ldots, d-1\}$. We define the functions

$$
f_\chi(x) = \frac{1}{b_{d-k}} \left( \frac{x}{a_{d,k}} \right)^{\frac{1}{2(d-k)}}\frac{1}{2(d-k)-1} \quad \text{and} \quad f_\nu(x) = \frac{1}{b_d} \left( \frac{x}{a_{d,k}} \right)^{\frac{1}{2(d-k)}}\frac{1}{2(d-k)-1},
$$

where the first derivatives $f'_\chi(x)$ and $f'_\nu(x)$ with respect to $x$ of the latter functions are given by

$$
f'_\chi(x) = \frac{(d-k)^{-1}}{2 b_{d-k} a_{d,k}} \left( \frac{x}{a_{d,k}} \right)^{\frac{1}{2(d-k)-1}} \quad \text{and} \quad f'_\nu(x) = \frac{(d-k)^{-1}}{2 b_d a_{d,k}} \left( \frac{x}{a_{d,k}} \right)^{\frac{1}{2(d-k)-1}},
$$

respectively.
respectively. Together with Theorem A.16 these derivatives imply in turn that both
\[
\sqrt{r} \left( f_x(\tilde{\lambda}_{k,r}) - f_x(\lambda_k) \right) \xrightarrow{d} \mathcal{N} \left( 0, \frac{1}{2(d-k)} \frac{\lambda_k^2}{a_{d,k}^{1-\frac{k}{d}}} b_{d-k}^2 \right) = \mathcal{N}(0,1)
\]
and
\[
\sqrt{r} \left( f_{\nu}(\lambda_{k,r}) - f_{\nu}(\lambda_k) \right) \xrightarrow{d} \mathcal{N} \left( 0, \frac{1}{2(d-k)} \frac{\lambda_k^2}{a_{d,k}^{1-\frac{k}{d}}} b_{d}^2 \right) = \mathcal{N}(0,1).
\]
Here, we used (5.6) and (5.7), respectively, of Corollary 5.3 and the fact that \( \tilde{\lambda}_{k,r} \) and \( \lambda_{k,r} \) are consistent estimators of \( \lambda_k \); cf. Lemma 5.1.

5.1.4 Asymptotic Confidence Intervals and Tests

Theorem 5.1 gives rise to \( 100(1-\alpha)\% \) (asymptotically exact) confidence intervals \( I_{k,r}^{(d)}(\alpha) \) and \( \tilde{I}_{k,r}^{(d)}(\alpha) \), respectively, for \( \lambda_k \), \( k = 0, \ldots, d-1 \).

**Theorem 5.2** Let \( \Phi_{d-1} \) be a stationary and isotropic PHP with (finite) intensity \( \lambda > 0 \). Then, \( \mathbb{P}(\lambda_k \in I_{k,r}^{(d)}(\alpha)) \xrightarrow{r \to \infty} 1 - \alpha \) and \( \mathbb{P}(\lambda_k \in \tilde{I}_{k,r}^{(d)}(\alpha)) \xrightarrow{r \to \infty} 1 - \alpha \), respectively, for any \( \alpha \in (0,1) \) and \( k = 0, \ldots, d-1 \), where
\[
I_{k,r}^{(d)}(\alpha) = \left[ b_r^{-}(I_{k,r}^{(d)}), b_r^{+}(I_{k,r}^{(d)}) \right] \quad \text{and} \quad \tilde{I}_{k,r}^{(d)}(\alpha) = \left[ b_r^{-}(\tilde{I}_{k,r}^{(d)}), b_r^{+}(\tilde{I}_{k,r}^{(d)}) \right]
\]
with
\[
b_r^{\pm}(I_{k,r}^{(d)}(\alpha),\alpha) = \left( \tilde{\lambda}_{k,r} \right)^{\frac{1}{2(d-k)}} \pm \frac{1}{\sqrt{r}} \frac{a_{d,k}^{1-\frac{k}{d}} b_{d-k}}{\sqrt{\alpha}} z_{1-\alpha/2}^{(2(d-k))}
\]
and
\[
b_r^{\pm}(\tilde{I}_{k,r}^{(d)}(\alpha),\alpha) = \left( \tilde{\lambda}_{k,r} \right)^{\frac{1}{2(d-k)}} \pm \frac{1}{\sqrt{r}} \frac{a_{d,k}^{1-\frac{k}{d}} b_{d}}{\sqrt{\alpha}} z_{1-\alpha/2}^{(2(d-k))},
\]
respectively, and \( z_{1-\alpha/2} \) denotes the \( (1-\alpha/2) \)-quantile of the standard normal distribution.

Moreover, by relation (5.5), the intervals \( I_{k,r}^{(d)}(\alpha) \) and \( \tilde{I}_{k,r}^{(d)}(\alpha) \) can be transformed into
\[
\tilde{J}_{k,r}^{(d)}(\alpha) = \left[ b_r^{-}(J_{k,r}^{(d)}), b_r^{+}(J_{k,r}^{(d)}) \right] \quad \text{and} \quad \tilde{J}_{k,r}^{(d)}(\alpha) = \left[ b_r^{-}(\tilde{J}_{k,r}^{(d)}), b_r^{+}(\tilde{J}_{k,r}^{(d)}) \right],
\]
respectively, which cover the intensity \( \lambda \) of the PHP \( \Phi_{d-1} \) with probability \( 1 - \alpha \) as \( r \to \infty \), where
\[
b_r^{\pm}(J_{k,r}^{(d)},\alpha) = \left( \frac{\lambda_{k,r}}{a_{d,k}} \right)^{\frac{1}{2(d-k)}} \pm \frac{b_{d-k}}{\sqrt{r}} z_{1-\alpha/2}^{2(d-k)}
\]
and
\[
b_r^{\pm}(\tilde{J}_{k,r}^{(d)},\alpha) = \left( \frac{\lambda_{k,r}}{a_{d,k}} \right)^{\frac{1}{2(d-k)}} \pm \frac{b_{d}}{\sqrt{r}} z_{1-\alpha/2}^{2(d-k)}.
\]
5.2 Some Extensions

5.2.1 Confidence Intervals for Facet Intensities

In this section, we show that in particular cases it is possible to derive asymptotic confidence intervals and asymptotic tests for the intensities of facets of Poisson hyperplane tessellations.

More precisely, we assume that we observe a realization of an $s$-stationary and isotropic Poisson hyperplane tessellation, \( (\psi, \Psi) \in \mathbb{R}^d \), with (finite) intensity \( \lambda > 0 \). The values \( \psi_j(\ell) \) and \( \psi_j(\ell') \) are given in (3.12).

Note that for \( k = d - 1 \) we obtain (asymptotic) tests for the mean total \((d - 1)\)-volume in \( [0, 1] \) of the Poisson hyperplanes themselves, induced by the process \( \Phi_{d-1} \).

In case of one-sided tests we consider, for some fixed \( k = 0, \ldots, d - 1 \), the null hypothesis \( H_0 : \lambda_k \leq \hat{\lambda}_k \) (versus \( H_1 : \lambda_k > \hat{\lambda}_k \)), whereas \( H_0 : \lambda_k \leq \bar{\lambda}_k \) (versus \( H_1 : \lambda_k > \bar{\lambda}_k \)). Then we reject

\[
\begin{align*}
\hat{\lambda}_k > \psi_j(\ell, k, 2a) \\
\bar{\lambda}_k > \psi_j(\ell', k, 2a)
\end{align*}
\]

and

\[
\begin{align*}
\hat{\lambda}_k < \psi_j(\ell, k, 2a) \\
\bar{\lambda}_k < \psi_j(\ell', k, 2a)
\end{align*}
\]

respectively, depending on whether the estimator \( \hat{\lambda}_k \) or \( \bar{\lambda}_k \) is used. Likewise, \( H_0 : \lambda_k \geq \hat{\lambda}_k \) (versus \( H_1 : \lambda_k < \hat{\lambda}_k \)). Then we reject

\[
\begin{align*}
\hat{\lambda}_k < \psi_j(\ell, k, 2a) \\
\bar{\lambda}_k < \psi_j(\ell', k, 2a)
\end{align*}
\]

and

\[
\begin{align*}
\hat{\lambda}_k > \psi_j(\ell, k, 2a) \\
\bar{\lambda}_k > \psi_j(\ell', k, 2a)
\end{align*}
\]

respectively, depending on whether the estimator \( \hat{\lambda}_k \) or \( \bar{\lambda}_k \) is used. Recall that \( \psi_j(\ell, k, 2a) \) and \( \psi_j(\ell', k, 2a) \), respectively, are given in (3.11), whereas \( b\) and \( b' \) are given in (3.8) and (3.9), respectively. We reject the null hypothesis whenever

\[
\begin{align*}
\hat{\lambda}_k > \psi_j(\ell, k, 2a) & \quad \text{and} \quad \psi_j(\ell, k, 2a) > \hat{\lambda}_k, \\
\bar{\lambda}_k > \psi_j(\ell', k, 2a) & \quad \text{and} \quad \psi_j(\ell', k, 2a) > \bar{\lambda}_k.
\end{align*}
\]

Since the (asymptotic) confidence intervals considered in Theorem 5.2 are closely related to corresponding (asymptotic) tests, we now turn our attention to some test problems. In the following, let \( \alpha \in (0, 1) \) be the (asymptotic) significance level, i.e., the (maximal) probability to reject a null hypothesis despite its correctness.
the underlying PHP. Using formulae (3.14) and (3.15) in Section 3.3.3, we get that \( \lambda \) can be expressed in terms of \( \widetilde{\gamma}^{(t)} \) and \( \gamma^{(t)} \) by
\[
\lambda = \left( \frac{d \kappa_d}{\kappa_d - 1} \right) \left( \frac{d \kappa_d}{\kappa_d - 1} \right)^{1/(d - \ell)} \quad \text{and} \quad \lambda = \left( \frac{d \kappa_d}{\kappa_d - 1} \right) \left( \frac{\gamma^{(t)}}{\gamma^{(t)}} \right)^{1/d}, \quad \ell = 0, \ldots, d - 1,
\]
respectively. In view of the confidence intervals \( J_{k,r}^{(d)}(\alpha) \) and \( \overline{J}_{k,r}^{(d)}(\alpha) \) derived in Theorem 5.2, we have that
\[
P(\widetilde{\gamma}^{(t)} \in \widetilde{K}_{\ell}^{(d)}(\widetilde{\lambda}_{k,r}, \alpha)) \longrightarrow 1 - \alpha \quad \text{and} \quad P\left(\gamma^{(t)} \in K_{\ell}^{(d)}(\lambda_{k,r}, \alpha)\right) \longrightarrow 1 - \alpha,
\]
where both \( \widetilde{K}_{\ell}^{(d)}(\widetilde{\lambda}_{k,r}, \alpha) \) and \( \widetilde{K}_{\ell}^{(d)}(\widetilde{\lambda}_{k,r}, \alpha) \) as well as both \( K_{\ell}^{(d)}(\lambda_{k,r}, \alpha) \) and \( K_{\ell}^{(d)}(\lambda_{k,r}, \alpha) \) are the following 100(1 - \( \alpha \))\% (asymptotically exact) confidence intervals for \( \widetilde{\gamma}^{(t)} \) and \( \gamma^{(t)} \), respectively. Introducing
\[
b_{\ell}^{\pm}(\overline{K}_{\ell}^{(d)}, \widetilde{\lambda}_{k,r}, \alpha) = \left( \frac{\lambda_{k,r}}{a_{d,k}} \right)^{2(d - \ell) \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}} \left( \frac{d}{\ell} \right)^{d \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}},
\]
and
\[
b_{\ell}^{\pm}(\overline{K}_{\ell}^{(d)}, \widetilde{\lambda}_{k,r}, \alpha) = \left( \frac{\lambda_{k,r}}{a_{d,k}} \right)^{2(d - \ell) \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}} \left( \frac{d}{\ell} \right)^{d \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}},
\]
as well as
\[
b_{\ell}^{\pm}(K_{\ell}^{(d)}, \lambda_{k,r}, \alpha) = \left( \frac{\lambda_{k,r}}{a_{d,k}} \right)^{2d \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}} \left( \frac{d}{\ell} \right)^{d \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}},
\]
and
\[
b_{\ell}^{\pm}(K_{\ell}^{(d)}, \lambda_{k,r}, \alpha) = \left( \frac{\lambda_{k,r}}{a_{d,k}} \right)^{2d \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}} \left( \frac{d}{\ell} \right)^{d \left( 1 \pm \frac{k_{d-1}}{d \kappa_{d}} \right)^{\frac{1}{z_1 - 1/2}}},
\]
the intervals \( \overline{K}_{\ell}^{(d)}(\widetilde{\lambda}_{k,r}, \alpha) \) and \( \overline{K}_{\ell}^{(d)}(\lambda_{k,r}, \alpha) \) are given by
\[
\left[ b_{\ell}^{-}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha), b_{\ell}^{+}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha) \right] \text{ and } \left[ b_{\ell}^{-}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha), b_{\ell}^{+}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha) \right],
\]
respectively, for \( k, \ell = 0, \ldots, d - 1 \). Likewise, \( K_{\ell}^{(d)}(\widetilde{\lambda}_{k,r}, \alpha) \) and \( K_{\ell}^{(d)}(\lambda_{k,r}, \alpha) \) are given by
\[
\left[ b_{\ell}^{-}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha), b_{\ell}^{+}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha) \right] \text{ and } \left[ b_{\ell}^{-}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha), b_{\ell}^{+}(\overline{K}_{\ell}^{(d)}, \lambda_{k,r}, \alpha) \right],
\]
respectively, for \( k, \ell = 0, \ldots, d - 1 \). Note that \( K_{\ell}^{(d)}(\lambda_{0,r}, \alpha) = \overline{K}_{\ell}^{(d)}(\lambda_{0,r}, \alpha) = I_{0,r}(\alpha) \) and \( K_{\ell}^{(d)}(\lambda_{0,r}, \alpha) = \overline{K}_{\ell}^{(d)}(\lambda_{0,r}, \alpha) = I_{0,r}(\alpha) \).

Of course, it is fairly easy to derive (asymptotic) tests for null hypotheses like \( H_0: \gamma^{(t)} = \gamma_*^{(t)} \) and \( H_0: \gamma^{(t)} = \gamma_*^{(t)} \) for \( \ell = 0, \ldots, d - 1 \), where \( \gamma_*^{(t)} \) and \( \gamma_*^{(t)} \) are fixed specified values.
5.2.2 Multivariate Inference for Intensity Vectors

Now we reconsider the multivariate extensions of the CLTs for \( k \)-flat processes, \( k = 0, \ldots, d - 1 \), in Section 4.3. We concentrate on the counting case, i.e., we consider

\[
(Z^{(d)}_{k,r}(\chi))_{k=0}^{d-1} \xrightarrow{d_{r \to \infty}} \mathcal{N}(0, \Sigma(\chi)),
\]

which is the statement (4.53) in Theorem 4.6, where \( \mathcal{N}(0, \Sigma(\chi)) \) denotes a \( d \)-dimensional Gaussian random vector with mean \( o = (0, \ldots, 0)^\top \) and covariance matrix \( \Sigma(\chi) = (\sigma_{\chi,k,l}^{(1,d-k,d-l)})_{k,l=0}^{d-1} \)
and where we recall that

\[
Z^{(d)}_{k,r}(\chi) = \frac{(d - k - 1)!}{(2\lambda r)^{d-k-1/2}} (\eta_k(B^d) - \mathbb{E} \eta_k(B^d)), \quad k = 0, \ldots, d - 1 .
\]

Moreover, we again additionally assume throughout this section that \( \Phi_{d-1} \) is isotropic, hence, by (4.56) in Lemma 4.10,

\[
\sigma_{\chi,i,j}^{(1,d-i,d-j)} = \frac{(d!)^2}{i! j! k_i k_j} \frac{\kappa_{d-i-1} \kappa_{d-j-1} \kappa_{2d-i-j-1} \kappa_{2d-i-j-2}}{d \kappa_d} \lambda^{2d-i-j},
\]

for \( i, j = 0, \ldots, d - 1 \). Furthermore, we recall from Corollary 4.11 that \( \Sigma(\chi) \) possesses full rank \( d \), i.e., equivalently, \( \det \Sigma(\chi) \neq 0 \) (even \( \det \Sigma(\chi) > 0 \)).

For the intensity \( \lambda > 0 \) of the underlying PHP \( \Phi_{d-1} \), let \( D_\lambda \) be a \( d \times d \) diagonal matrix such that

\[
D_\lambda = \text{diag}\left( \frac{(2\lambda)^{d-k-1/2}}{\kappa_{d-k} (d-k-1)!} \right)_{k=0}^{d-1},
\]

where we note that \( \det D_\lambda > 0 \) since \( \lambda > 0 \). Then,

\[
Z^{(d)}_r(\chi) = D_\lambda \left( (Z^{(d)}_{k,r}(\chi))_{k=0}^{d-1} \right)^\top \sim \sqrt{r} \left( \lambda_k - \lambda d_{k=0}^{d-1} \right)^\top \xrightarrow{d_{r \to \infty}} \mathcal{N}(0, \Sigma_\chi(\chi)), \quad (5.13)
\]

where \( \mathcal{N}(o, \Sigma_\chi(\chi)) \) is a Gaussian random vector with mean zero and \( d \times d \) covariance matrix \( \Sigma_\chi(\chi) = D_\lambda \Sigma(\chi) D_\lambda \). The \( (i,j) \)-th entry of \( \Sigma_\chi(\chi) \) is given by

\[
\frac{(d!)^2 \kappa_{d-i-1} \kappa_{d-j-1} \kappa_{2d-i-j-1} \kappa_{2d-i-j-2} \lambda^{2d-i-j}}{i! j! k_i k_j k_{d-i} k_{d-j}} \frac{\kappa_{d-i-1} \kappa_{d-j-1}}{d \kappa_d} \frac{1}{(d-i-1)! (d-j-1)!} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{2d-i-j},
\]

which implies that \( \det \Sigma_\chi(\chi) > 0 \).

By relation (5.5), which, recall, states that \( \lambda_k = a_{d,k} \lambda_{d-k} \) for \( k = 0, \ldots, d - 1 \), we get that

\[
Z^{(d)}_r(\chi) \xrightarrow{d_{r \to \infty}} \mathcal{N}(0, \Sigma_k(\chi)), \quad k = 0, \ldots, d - 1 , \quad (5.15)
\]

with \( Z^{(d)}_r(\chi) \) defined by (5.13) and where the \( (i,j) \)-th entry of the covariance matrix \( \Sigma_k(\chi) \) is given by

\[
\frac{(d!)^2 \kappa_{d-i-1} \kappa_{d-j-1} \kappa_{2d-i-j-1} \kappa_{2d-i-j-2} \lambda^{2d-i-j}}{i! j! k_i k_j k_{d-i} k_{d-j}} \left( \frac{\kappa_{d-1}}{d \kappa_d} \right)^{2d-i-j} \frac{1}{(d-i-1)! (d-j-1)!} \left( \frac{\lambda_k}{a_{d,k}} \right)^{2d-i-j-1},
\]

as desired.

\[\Box\]
for $i, j = 0, \ldots, d - 1$. Replacing $\lambda_k$ by the estimator $\hat{\lambda}_{k,r}$ for all $k = 0, \ldots, d - 1$ in the latter covariance matrices $\Sigma_k(\chi)$ yields matrices $\hat{\Sigma}_{k,r}(\chi)$ with $\det \hat{\Sigma}_{k,r}(\chi) > 0$ (assumed to this end). By the consistency of the estimator $\hat{\lambda}_{k,r}$ discussed in Section 5.1.1, we get that

$$
\|\hat{\Sigma}_{k,r}(\chi) - \Sigma_k(\chi)\| \xrightarrow{P} 0, \quad k = 0, \ldots, d - 1,
$$

where $\|A\| = (\sum_{i,j=1}^d a_{ij}^2)^{1/2}$ for any $d \times d$ matrix $A = (a_{ij})_{i,j=1}^d$. This latter convergence in probability implies

$$
\det \hat{\Sigma}_{k,r}(\chi) \xrightarrow{P} \det \Sigma_k(\chi) > 0, \quad k = 0, \ldots, d - 1,
$$

such that

$$
\hat{\Sigma}_{k,r}^{-1/2}(\chi) \Sigma_k^{1/2}(\chi) \xrightarrow{P} I_d,
$$

where $I_d$ denotes the $d \times d$ identity matrix, and where $\hat{\Sigma}_{k,r}^{-1/2}(\chi)$ and $\Sigma_k^{1/2}(\chi)$ denote a square root of $\hat{\Sigma}_{k,r}(\chi)$ and $\Sigma_k(\chi)$, respectively, for $k = 0, \ldots, d - 1$.

Note that $\hat{\Sigma}_{k,r}^{-1/2}(\chi)$ can be computed by the fact that

$$
\hat{\Sigma}_{k,r}(\chi) = O_{k,r} \left( \text{diag}(\hat{\epsilon}_{k,1}, \ldots, \hat{\epsilon}_{k,d}) \right)^\top O_{k,r}^\top, \quad k = 0, \ldots, d - 1,
$$

where $\hat{\epsilon}_{k,1}, \ldots, \hat{\epsilon}_{k,d}$ are the positive eigenvalues of the matrix $\hat{\Sigma}_{k,r}(\chi)$ and where $O_{k,r}$ is an orthogonal $d \times d$ matrix containing as column vectors the eigenvectors $\hat{v}_{k,1}, \ldots, \hat{v}_{k,d}$ of $\hat{\Sigma}_{k,r}(\chi)$ corresponding to $\hat{\epsilon}_{k,1}, \ldots, \hat{\epsilon}_{k,d}$. Then,

$$
\hat{\Sigma}_{k,r}^{-1}(\chi) = O_{k,r} \left( \text{diag}(\hat{\epsilon}_{k,1}^{-1/2}, \ldots, \hat{\epsilon}_{k,d}^{-1/2}) \right)^\top O_{k,r}^\top, \quad k = 0, \ldots, d - 1.
$$

Hence, under the null hypothesis $H_0 : (\lambda_0, \ldots, \lambda_{d-1})^\top = (\lambda_0^*, \ldots, \lambda_{d-1}^*)^\top$, where $\lambda_0^*, \ldots, \lambda_{d-1}^*$ are some fixed specified values for $\lambda_0, \ldots, \lambda_{d-1}$ we get by (5.15) together with Slutsky's theorem, cf. Theorem A.8, that

$$
\hat{Z}^{(d)}(\chi) = \hat{\Sigma}_{k,r}^{-1/2}(\chi) \sqrt{r} \left( \begin{array}{c} \hat{\lambda}_{0,r} - \lambda_0^* \\ \vdots \\ \hat{\lambda}_{d-1,r} - \lambda_{d-1}^* \end{array} \right) \xrightarrow{r \to \infty} N(0, I_d).
$$

Application of the continuous mapping theorem yields that

$$
\|\hat{Z}^{(d)}(\chi)\|^2 \xrightarrow{r \to \infty} \chi^2_d.
$$

Hence, the null hypothesis $H_0$ is rejected at the significance level $\alpha \in (0, 1)$ if the computed value of $\|\hat{Z}^{(d)}(\chi)\|^2$ is greater than the $(1 - \alpha)$-quantile $\chi^2_{d,1-\alpha}$ of the $\chi^2$-distribution with $d$ degrees of freedom. Note that for $d = 2$ we have $\chi^2_{d,1-\alpha} = -2 \log(\alpha)$ since $\|\hat{Z}^{(2)}(\chi)\|^2$ is exponentially distributed with expectation 2. This in turn implies that the null hypothesis
$H_0 : (\lambda_0, \lambda_1)^T = (\lambda^*_0, \lambda^*_1)^T$ is rejected at the significance level $\alpha \in (0, 1)$, if $|\hat{Z}_r^{(2)}(\chi)|^2 > \sqrt{-2\log(\alpha)}$.

Another possibility is to consider tests for linear hypotheses. Let $b = (b_0, \ldots, b_{d-1})^T \in \mathbb{R}^d$ be a $d$-dimensional vector such that $\sigma^2_{d, \lambda}(b, \chi) = b^T \Sigma_\lambda(b) > 0$. With $Z_r^{(d)}(\chi)$ given by (5.13) we obtain that

$$b^T Z_r^{(d)}(\chi) = \sqrt{r} \sum_{k=0}^{d-1} b_k (\hat{\lambda}_{k,r} - \lambda_k) \xrightarrow{d \to \infty} \mathcal{N}(0, \sigma^2_{d, \lambda}(b, \chi)),$$

where $\sigma^2_{d, \lambda}(b, \chi)$ can be given explicitly by

$$\sigma^2_{d, \lambda}(b, \chi) = \sum_{i,j=1}^{d-1} \frac{b_i b_j d!^2 \lambda^{2d-i-j-1}}{i! j! (d-i-1)! (d-j-1)!} \kappa_{2d-i-j-1} \kappa_{d-i-1} \kappa_{d-j-1} \left( \kappa_{d-1} \right)^{2d-i-j} \frac{d}{d \kappa_d}.$$

Altogether, by Slutsky’s theorem (cf. Theorem A.8), we get that

$$\frac{\sqrt{r}}{\hat{\sigma}_{d, r}(b, \chi)} \sum_{k=0}^{d-1} b_k (\hat{\lambda}_{k,r} - \lambda_k) \xrightarrow{d \to \infty} \mathcal{N}(0, 1),$$

where $\hat{\sigma}_{d, r}(b, \chi)$ coincides with $\sigma^2_{d, \lambda}(b, \chi)$ except that $\lambda$ is replaced by $\hat{\lambda}_{d-1,r}$, i.e., $\hat{\sigma}_{d, r}(b, \chi)$ equals

$$\hat{\sigma}_{d, r}(b, \chi) = \sum_{i,j=1}^{d-1} \frac{b_i b_j d!^2 \lambda^{2d-i-j-1}}{i! j! (d-i-1)! (d-j-1)!} \kappa_{2d-i-j-1} \kappa_{d-i-1} \kappa_{d-j-1} \left( \kappa_{d-1} \right)^{2d-i-j} \frac{d}{d \kappa_d}.$$

By the consistency properties of $\hat{\lambda}_{d-1,r}$ discussed in Section 5.1.1, the estimator $\hat{\sigma}_{d, r}(b, \chi)$ is (strongly) consistent. Furthermore note that, alternatively, we could also consider an estimator $\tilde{\sigma}_{d, r}(b, \chi)$ depending on $\lambda_{d-1,r}$. To provide an example of a test, assume that $d = 2$ and that we want to test the null hypothesis $H_0 : b_0 \lambda_0 + b_1 \lambda_1 = b_0 \lambda^*_0 + b_1 \lambda^*_1$ versus the alternative hypothesis $H_1 : b_0 \lambda_0 + b_1 \lambda_1 \neq b_0 \lambda^*_0 + b_1 \lambda^*_1$. Note that $\lambda^*_0$ and $\lambda^*_1$ are fixed specified values for $\lambda_0$ and $\lambda_1$, respectively, whereas $b_0$ and $b_1$ may be interpreted as weights. Then, $H_0$ is rejected if

$$\sqrt{r} \frac{b_0 \lambda_{0,r} + b_1 \lambda_{1,r} - b_0 \lambda^*_0 - b_1 \lambda^*_1}{\sqrt{\frac{b_0^2}{3\pi^2} \lambda_{0,r}^2 + \frac{b_1^2}{3\pi^2} b_1^2 \lambda_{1,r}^2 + \frac{1}{2\pi^2} b_1^2 \lambda_{1,r}^2}} > z_{1-\alpha/2},$$

where $z_{1-\alpha/2}$ denotes the $(1 - \alpha/2)$-quantile of the standard normal distribution.

### 5.3 Inference for Poisson–Voronoi Tessellations

In order to show that statistical inference is not restricted to the class of (non-degenerate) Poisson hyperplane processes, we review a CLT for PVTs in $\mathbb{R}^d$ in this section. The reported
results have been obtained in [50]. Note that an alternative and general approach to CLTs in the Poisson–Voronoi case is given in [95]. There, from the view point of computational geometry, a CLT for a quite large class of functionals related to PVTs is obtained among other results.

5.3.1 A CLT for the Vertices of Poisson–Voronoi Tessellations

Consider a stationary PVT $\Psi$ with (finite) intensity $\gamma > 0$, where $\Psi$ is generated by a stationary Poisson point process in $\mathbb{R}^d$ with some positive intensity, which coincides with the intensity $\gamma$ of the PVT. Recall from Section 3.3.1 that each cell of the Poisson–Voronoi tessellation $\Psi$ is defined as the closure of the set of all points in $\mathbb{R}^d$ which are closest (in the Euclidean sense) to the nucleus of this cell, where the nucleus is a point of the generating Poisson point process. Moreover, recall that each cell of $\Psi$ is a $d$-polytope.

For the stationary (and isotropic) point process $\Psi^{(0)}$ of vertices of $\Psi$ with intensity $\gamma^{(0)}$ we recall from (3.10) in Section 3.3.1 that

$$\gamma^{(0)} = c_d \gamma \quad \text{with} \quad c_d = \frac{\pi^{d-1}}{d+1} \frac{\kappa_d}{\kappa_d-1} \left( \frac{\kappa_d-1}{\kappa_d} \right)^d .$$

Consider an averaging sequence of convex sampling windows $W_n^d \subset \mathbb{R}^d$, i.e., recall, the sets $W_n^d$ are compact convex, increasing, and contain a ball with unboundedly growing radius (p. 332 in [31]). Since the $\beta$-mixing coefficient of a stationary PVT is exponentially decaying ([47]), a CLT for the random number of vertices $\Psi^{(0)}(W_n^d)$ of a stationary PVT $\Psi$ within $W_n^d$ can be derived as $n \to \infty$. More precisely, in [50] it is shown that

$$\left( \frac{\nu_d(W_n^d)}{\nu_d(W_n^d)} - \gamma^{(0)} \nu_d(W_n^d) \right) \xrightarrow{n \to \infty} \mathcal{N}\left(0, \gamma^{(0)} \left(1 + c_d \sigma_n^2\right)\right) , \quad (5.17)$$

where $\sigma_n^2$ is a constant depending only on the dimension $d$ and which is expressible in terms of multiple integrals. A detailed discussion including numerical computations for the cases of $d = 2$ and $d = 3$ can be found in [48] and [50], respectively. The rounded values obtained there are $\sigma_2^2 = 0.5$ and $\sigma_3^2 = 5.084$, together with $c_2 = 2$ and $c_3 = 6.768$ from (5.16).

5.3.2 Asymptotic Confidence Intervals and Tests

By the stationarity of $\Psi^{(0)}$, we have that

$$\hat{\gamma}_n^{(0)} = \frac{\Psi^{(0)}(W_n^d)}{\nu_d(W_n^d)}$$

is an unbiased estimator for the intensity $\gamma^{(0)}$ and, by (5.17),

$$\left( \frac{\nu_d(W_n^d)}{\nu_d(W_n^d)} \right)^{1/2} \left( \hat{\gamma}_n^{(0)} - \gamma^{(0)} \right) \xrightarrow{n \to \infty} \mathcal{N}\left(0, \gamma^{(0)} \left(1 + c_d \sigma_n^2\right)\right) , \quad (5.18)$$
Furthermore, ergodicity of $\Psi^{(0)}$ implies that

$$\hat{\gamma}_n^{(0)} \xrightarrow{a.s.} \gamma^{(0)} ~. \quad (5.19)$$

A simple transformation using (5.17) and (5.19) yields that

$$\hat{\gamma}_{0,n}^{(d)} = 2 \sqrt{\frac{\nu_d(W_n^d)}{1 + c_d \sigma_d^2}} \left( \sqrt{\hat{\gamma}_n^{(0)}} - \sqrt{\hat{\gamma}_n^{(d)}} \right) \xrightarrow{n \to \infty} N(0,1). \quad (5.20)$$

Similar to the confidence intervals derived in Theorem 5.2 for Poisson hyperplane processes, the CLT (5.20) allows for the construction of a $100(1-\alpha)\%$ (asymptotically exact) confidence interval for $\gamma$. Indeed, for any $\alpha \in (0,1)$ and for a large enough sampling window $W_n^d$, one has $\gamma \in J_n^{(d)}(\alpha)$ with probability $1-\alpha$, where $J_n^{(d)}(\alpha)$ is given by

$$J_n^{(d)}(\alpha) = \left[ \frac{1}{\nu_d(W_n^d)} b_n^-(J_n^{(d)}, \alpha), \frac{1}{c_d \nu_d(W_n^d)} b_n^+(J_n^{(d)}, \alpha) \right],$$

with

$$b_n^\pm(J_n^{(d)}, \alpha) = \left( \sqrt{\Psi^{(0)}(W_n^d)} \pm \frac{z_{1-\alpha/2}}{2} \sqrt{1 + c_d \sigma_d^2} \right)^2.$$

Moreover, however not necessarily useful, also for $\gamma^{(0)}$ itself we can construct a $100(1-\alpha)\%$ (asymptotically exact) confidence interval $I_n^{(d)}(\alpha)$, where

$$I_n^{(d)}(\alpha) = \left[ \frac{1}{\nu_d(W_n^d)} b_n^-(I_n^{(d)}, \alpha), \frac{1}{c_d \nu_d(W_n^d)} b_n^+(I_n^{(d)}, \alpha) \right],$$

and $b_n^\pm(I_n^{(d)}, \alpha) = b_n^\pm(J_n^{(d)}, \alpha)$, respectively.

Immediately, asymptotic tests can be given. Assume for example that we want to test the null hypothesis $H_0 : \gamma = \gamma^*$ for the significance level $\alpha \in (0,1)$, where $\gamma^*$ is some fixed specified value of $\gamma$. Based on counting the vertices $\Psi^{(0)}(W_n^d)$ of a single realization of the PVT $\Psi$ in a (sufficiently large) sampling window $W_n^d$, the null hypothesis $H_0$ is rejected if $\gamma^* \notin J_n^{(d)}(\alpha)$.

Alternatively, but equivalently, $H_0$ is rejected whenever

$$\hat{\gamma}_n^0 \begin{cases} < \left( -\frac{z_{1-\alpha/2}}{2} \sqrt{\frac{1 + c_d \sigma_d^2}{\nu_d(W_n^d)}} + \sqrt{\nu_d(W_n^d) \gamma^*} \right)^2 \quad \text{or} \quad \hat{\gamma}_n^0 > \left( \frac{z_{1-\alpha/2}}{2} \sqrt{\frac{1 + c_d \sigma_d^2}{\nu_d(W_n^d)}} + \sqrt{\nu_d(W_n^d) \gamma^*} \right)^2. \end{cases}$$

Also, tests of the null hypothesis $H_0 : \gamma \leq \gamma^*$ or the null hypothesis $H_0 : \gamma \geq \gamma^*$ are possible. In the latter case for example, $H_0$ is rejected whenever

$$\gamma^* > \frac{1}{c_d \nu_d(W_n^d)} b_n^+(J_n^{(d)}, 2\alpha)$$

or, equivalently, whenever

$$\hat{\gamma}_n^0 \begin{cases} < \left( -\frac{z_{1-\alpha}}{2} \sqrt{\frac{1 + c_d \sigma_d^2}{\nu_d(W_n^d)}} + \sqrt{c_d \gamma^*} \right)^2. \end{cases}$$
5.4 Numerical Evaluations

In this section we assume $d = 2$ and examine briefly the CLTs, confidence intervals, and tests presented in the previous sections. In particular, we study the weak convergence of $\hat{Z}_{0,r}^{(2)}$, $\hat{Z}_{1,r}^{(2)}$, and $\bar{Z}_{1,r}^{(2)}$, given in (5.8) and (5.9), respectively, as $r \to \infty$, as well as of $\bar{Y}_{0,n}^{(2)}$ given in (5.20) as $n \to \infty$. We therefore observe realizations of simulated underlying Poisson line processes and Poisson–Voronoi tessellations in a sequence of sampling windows. For Poisson line processes, we also consider the bivariate case of the vector

$$Z_r^{(2)}(\chi) = (Z_{0,r}^{(2)}(\chi), Z_{1,r}^{(2)}(\chi))^\top;$$

cf. also (4.53) in Section 4.3.1 as well as Section 5.2.2. We use a bivariate goodness–of–fit test for normality proposed by Henze and Zirkler; cf. [55] and [56] as well as Section A.3.8). The applied goodness–of–fit tests have been implemented in JAVA™ and are incorporated in the GeoStoch library. Beyond that, in particular for the graphical evaluations, S–Plus™ has been used.

5.4.1 Goodness–of–Fit Tests for Poisson Line Processes

Assume that we observe realizations of a stationary and isotropic Poisson line process $\Phi_1$ with intensity $\lambda = 0.1$ through a sequence $\{B_r^2\}_{r \geq 1}$ of spherical sampling windows $B_r^2 \subset B_{r+1}^2 \subset \ldots \subset B_n^2$ for $i \in \mathbb{N}$. Recall from Section 4.1.3 that $\Phi_1$ admits the marked–point–process representation $\sum_{i \geq 1} \delta_{[P_i,U_i]}$ with the unit vector $U_i$ being completely determined by the angle $\Delta_i$ measured in anti-clockwise direction between $U_i$ and the $x$–axis. The angles $\Delta_1, \Delta_2, \ldots$ are independent and uniformly distributed on $[0, \pi]$.

Tables 5.1 and 5.2 show the performance of the estimators $\hat{\lambda}_{0,r}$, $\hat{\lambda}_{1,r}$, and $\bar{\lambda}_{1,r}$. In view of the results presented in these tables we note that the estimated values correspond well to the theoretical values if the number of employed realizations grows. In particular the estimated values in Table 5.1 for a single realization ($m = 1$) of $\Phi_1$ emphasize the large variation between different realizations of Poisson line processes.

Table 5.1: Estimation of $\lambda_0$ by $\hat{\lambda}_{0,r}$ ($m$ realizations of $\Phi_1$ in $B_r^2$ with $\lambda = 0.1$; theoretical value $\lambda_0 = 0.003183$)

<table>
<thead>
<tr>
<th>$r \setminus m$</th>
<th>1</th>
<th>1000</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>0.003233</td>
<td>0.003175</td>
<td>0.003179</td>
</tr>
<tr>
<td>600</td>
<td>0.003115</td>
<td>0.003165</td>
<td>0.003179</td>
</tr>
<tr>
<td>900</td>
<td>0.003605</td>
<td>0.003179</td>
<td>0.003189</td>
</tr>
<tr>
<td>1200</td>
<td>0.003115</td>
<td>0.003206</td>
<td>0.003176</td>
</tr>
</tbody>
</table>
Table 5.2: Estimation of $\lambda_1$ ($m$ realizations of $\Phi_1$ in $B_r^2$ with $\lambda = 0.1$; theoretical value $\lambda_1 = \lambda$)

\[
\begin{array}{cccc}
 r \backslash m & 1 & 5000 \\
 300 & 0.098333 & 0.100010 \\
 600 & 0.095833 & 0.099861 \\
 900 & 0.108889 & 0.100001 \\
 1200 & 0.103750 & 0.099890 \\
\end{array}
\]

\[
\begin{array}{cccc}
 r \backslash m & 1 & 5000 \\
 300 & 0.100630 & 0.099955 \\
 600 & 0.098021 & 0.099896 \\
 900 & 0.106488 & 0.100057 \\
 1200 & 0.104984 & 0.099883 \\
\end{array}
\]

For the goodness-of-fit tests we concentrate on $\hat{Z}_{0,r}^{(2)}$ and choose $r_1 = 300$ and $r_j = j r_1$ for $j = 2, 3, 4$. Each time, we test the hypothesis

\[ H_0^{(r)} : F_{\hat{Z}_{0,r}^{(2)}} = F_{N(0,1)}, \]

where $F_{\hat{Z}_{0,r}^{(2)}}$ denotes the cumulative distribution function (cdf) of $\hat{Z}_{0,r}^{(2)}$ and $F_{N(0,1)}$ is the cdf of the standard normal distribution; cf. Section A.3.1. Figure 5.1 displays histograms, quantile plots, and empirical distribution functions (edfs) for $m = 1000$ realizations of $\Phi_1$ in $B_{300}^2$ and $m = 5000$ realizations of $\Phi_1$ in $B_{5000}^2$. As we can notice, the respective figures for the case of 5000 realizations (of $\Phi_1$) show a better fit compared to the figures for the case of 1000 realizations. Note that in case of 5000 realizations we also observe $\Phi_1$ through a larger (circular) sampling window. Table 5.3 summarizes the results of the goodness-of-fit tests for the (null) hypothesis $H_0^{(r)}$, where the significance level, i.e., the (maximal) probability of rejecting the hypothesis $H_0^{(r)}$ despite its correctness, is $\alpha = 0.05$.

Table 5.3: Goodness-of-fit test evaluating $H_0^{(r)}$ for $m$ simulations of $\Phi_1$ in $B_r^2$; functional $\hat{Z}_{0,r}^{(2)}$

Rejection of $H_0^{(r)}$ is denoted by $\ast$.

\[
\begin{array}{cccc}
 (a) \text{Pearson test} & (b) \text{Kolmogorov-Smirnov test} \\
 r & T_{1000} & p_{1000} & T_{5000} & p_{5000} & r & T_{1000}' & p_{1000}' & T_{5000}' & p_{5000}' \\
 300 & 42.50 & 0.051 & 126.10^* & < 10^{-3} & 300 & 1.58^* & 0.014 & 2.85^* & < 10^{-3} \\
 600 & 32.18 & 0.312 & 116.16^* & < 10^{-3} & 600 & 1.72^* & 0.006 & 2.47^* & < 10^{-3} \\
 900 & 34.88 & 0.209 & 78.44 & 0.204 & 900 & 1.33 & 0.057 & 1.76^* & 0.004 \\
 1200 & 28.52 & 0.490 & 70.13 & 0.440 & 1200 & 0.64 & 0.800 & 2.28^* & < 10^{-3} \\
\end{array}
\]

In case of the Pearson test, with results displayed in Table 5.3 (a), the degrees of freedom 29 and 69 are chosen, since we use 30 classes in the case of 1000 simulations and 70 classes
Figure 5.1: Histograms, quantile plots (q-plots), and empirical distribution functions for $\hat{Z}_{0,r}^{(2)}$ (m realizations of $\Phi_I$ with $\lambda = 0.1$ in $B_r^2$)
in the case of 5000 simulations, respectively, following a rule of thumb provided in Section 10.2 of [22]. Note that the 0.95-quantiles $\chi^2_{29;0.95}$ and $\chi^2_{69;0.95}$ of the $\chi^2$-distribution and the $\chi^2$-distribution, respectively, are given by $\chi^2_{29} = 42.56$ and $\chi^2_{69} = 89.39$. Thus, we reject $H_0^{(r)}$ if for the computed values $T_{1000}$ and $T_{5000}$ of Pearson’s test statistic $T_m$, cf. Section 10.2 of [22] for example, we have that $T_{1000} > 42.56$ and $T_{5000} > 89.39$, respectively. Alternatively, we can also consider the p-values $p_{1000}$ and $p_{5000}$. Generally, the p-value of a sample is the smallest value of the significance level $\alpha$ for which this sample will lead to rejection of the null hypothesis. Note however that the p-value itself is not a significance level and that p-values are data dependent; cf. Section 8.3 of [26]. Thus, small values of $p_{1000}$ and $p_{5000}$ indicate that $H_0^{(r)}$ should be rejected, more precisely, for this test and for all subsequent tests, we reject $H_0^{(r)}$ if $p_{1000} \leq 0.05$ and $p_{5000} \leq 0.05$, respectively.

Likewise, Table 5.3 (b) shows the results of Kolmogorov–Smirnov’s test. The 0.95-quantile $\alpha_{0.95}$ of the (asymptotic) Kolmogorov–Smirnov distribution is given by $\alpha_{0.95} = 1.36$. Therefore, we reject $H_0^{(r)}$ if for the calculated values $T'_{1000}$ and $T'_{5000}$ of Kolmogorov–Smirnov’s test statistic $T'_m$, cf. Section 10.6 of [22] for example, we have that $T'_{1000} > 1.36$ and $T'_{5000} > 1.36$, respectively.

We first simulate $\Phi_1$ with intensity $\lambda = 0.1$ a certain number of times, 1000 say, and test $H_0^{(r)}$ in $B_{r1}^2 \subset B_{r2}^2 \subset \ldots$ until $H_0^{(r)}$ is not rejected. In Table 5.3, as well as in all subsequent tables that contain results of goodness-of-fit tests, the values of the computed test statistics of Pearson and Kolmogorov–Smirnov, respectively, along with the respective p-values are printed in bold. After $H_0^{(r)}$ has not been rejected for 1000 simulations for the first time, we increase the number of simulations to 5000, say, in order to study the influence of the number of simulations on the test result. Again in Table 5.3, as well as in all subsequent tables that contain results of goodness-of-fit tests, we print then all values of computed test statistics along with the respective p-values in bold. Values which are not printed in bold are provided to give additional information. Rejection of the null hypothesis $H_0^{(r)}$ is denoted by an asterisk * behind the value of Pearson’s test statistic $T_m$ or the value of Kolmogorov–Smirnov’s test statistic $T'_m$.

If we review the results, we see that at least for a radius $r = 900$ of $B_2^2$ both tests do not reject $H_0^{(900)}$ for $m = 1000$ which complies well to the graphical results of Figure 5.1. Note that the hypothesis $H_0^{(900)}$ is not rejected by Pearson’s test for $m = 5000$, but that Kolmogorov–Smirnov’s test rejects $H_0^{(900)}$, whereas the graphical evaluation in Figure 5.1 looks pretty convincing in favor of $H_0^{(900)}$. However, also for $r = 300$, Pearson’s test does not reject $H_0^{(300)}$, whereas the graphical results (cf. once again Figure 5.1) are not as convincing as in the case of radius $r = 900$. Moreover, the results of Table 5.3 also show the effect that for 5000 realizations of $\Phi_1$ we need larger sampling windows compared to the case of 1000 realizations such that the null hypothesis is rejected for the first time. This effect is also visible in subsequent tables that report about goodness-of-fit results in this chapter, as well as in the respective tables of Section 6.4.3 in Chapter 6. For fixed size of the sampling window, the test based on the larger number of simulations (5000) uses estimated values of the underlying random functionals where the sampling bias is reduced compared to estimated values in the case of fewer (1000)
simulations; cf. also the discussion of the performance of our estimators in Tables 5.1 and 5.2. Hence, the test in the case of 5000 simulations is based on improved estimated values and its decision can be seen to be more critical towards the assumed normal distribution in the null hypothesis compared to the test based on 1000 simulations in a sampling window of the same size.

Finally, we note that further examples and also results for \( \hat{Z}_{1,r}^{(2)} \) and \( \tilde{Z}_{1,r}^{(2)} \) are given in [58]. Here we restrict ourselves to present Kolmogorov–Smirnov’s test in Table 5.4 (a) for \( \hat{Z}_{1,r}^{(2)} \) and in Table 5.4 (b) for \( \tilde{Z}_{1,r}^{(2)} \), where in both cases the significance level is \( \alpha = 0.05 \). Note that Pearson’s test showed similar results; cf. [58]. The hypothesis of (standard) normality of \( \hat{Z}_{1,r}^{(2)} \) is not rejected for \( r = 900 \) and \( m = 1000 \). In contrast, for \( \tilde{Z}_{1,r}^{(2)} \) the hypothesis \( H_0^{(3000)} \) is not rejected both for \( m = 1000 \) and \( m = 5000 \). Hence, we may conclude that (standard) normality can be assumed in case of \( \hat{Z}_{1,r}^{(2)} \) for much smaller (circular) sampling windows \( B^2_r \) than for the functionals \( \hat{Z}_{0,r}^{(2)} \) and \( \tilde{Z}_{1,r}^{(2)} \). The latter two functionals are based on the estimators \( \hat{\lambda}_{k,r} \) for \( k = 0, 1 \) introduced in (5.1), whereas \( \tilde{Z}_{1,r}^{(2)} \) is based on \( \tilde{\lambda}_{1,r} \) introduced also in (5.1). In view of the properties of these estimators (cf. Sections 5.1.1 and 5.1.2 and especially Lemma 5.2) the goodness-of-fit behavior of \( \hat{Z}_{0,r}^{(2)} \) and \( \tilde{Z}_{1,r}^{(2)} \) on the one hand and of \( \tilde{Z}_{1,r}^{(2)} \) on the other hand may be interpreted that there exists a trade-off between smaller asymptotic variance of \( \hat{\lambda}_{k,r} \) for \( k = 0, 1 \) and a somewhat lower rate of convergence of both \( \hat{Z}_{0,r}^{(2)} \) and \( \tilde{Z}_{1,r}^{(2)} \).

Table 5.4: Kolmogorov–Smirnov test evaluating \( H_0^{(r)} \) for \( m \) simulations of \( \Phi_1 \) in \( B^2_r \) (test statistic \( T_m \), p-value \( p_m \)). Rejection of \( H_0^{(r)} \) is denoted by *.

<table>
<thead>
<tr>
<th>( r )</th>
<th>( T'_{1000} )</th>
<th>( p_{1000} )</th>
<th>( T'_{5000} )</th>
<th>( p_{5000} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>1.61*</td>
<td>0.011</td>
<td>2.80*</td>
<td>&lt;10(^{-3})</td>
</tr>
<tr>
<td>600</td>
<td>1.57*</td>
<td>0.014</td>
<td>2.21*</td>
<td>&lt;10(^{-3})</td>
</tr>
<tr>
<td>900</td>
<td>1.13</td>
<td>0.157</td>
<td>2.08*</td>
<td>&lt;10(^{-3})</td>
</tr>
<tr>
<td>1200</td>
<td>1.29</td>
<td>0.072</td>
<td>2.17*</td>
<td>&lt;10(^{-3})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( r )</th>
<th>( T'_{1000} )</th>
<th>( p_{1000} )</th>
<th>( T'_{5000} )</th>
<th>( p_{5000} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>0.87</td>
<td>0.432</td>
<td>1.33</td>
<td>0.059</td>
</tr>
<tr>
<td>600</td>
<td>0.92</td>
<td>0.366</td>
<td>1.26</td>
<td>0.081</td>
</tr>
<tr>
<td>900</td>
<td>0.91</td>
<td>0.384</td>
<td>0.62</td>
<td>0.842</td>
</tr>
<tr>
<td>1200</td>
<td>1.03</td>
<td>0.240</td>
<td>1.17</td>
<td>0.131</td>
</tr>
</tbody>
</table>

Now, we evaluate the distribution of the vector

\[
Z_r^{(2)}(\chi) = (Z_{0,r}^{(2)}(\chi), Z_{1,r}^{(2)}(\chi))^\top
\]

as \( r \to \infty \); cf. Section 5.2.2. Table 5.5 shows the results of the Henze–Zirkler bivariate goodness-of-fit test applied to the test problem

\[
\tilde{H}_0^{(r)}: \text{The distribution of } (Z_{0,r}^{(2)}(\chi), Z_{1,r}^{(2)}(\chi))^\top \text{ is } \mathcal{N}(\cdot, \cdot),
\]
Table 5.5: Bivariate goodness–of–fit test evaluating $H_0^{(r)}$ (results based on 1000 realizations of $\Phi_1$ with $\lambda = 0.1$ in $B_2^2$; Henze–Zirkler’s test statistic $T_{1000,\beta}^{(r)}$ and 0.95-quantiles $q_{\beta}(0.95)$; rejection of $H_0^{(r)}$ indicated by *).

<table>
<thead>
<tr>
<th>$\beta \setminus r$</th>
<th>600</th>
<th>900</th>
<th>1200</th>
<th>1500</th>
<th>1800</th>
<th>$q_{\beta}(0.95)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.39143*</td>
<td>0.50179*</td>
<td>0.25392*</td>
<td>0.41030*</td>
<td>0.17658*</td>
<td>0.07901</td>
</tr>
<tr>
<td>1.0</td>
<td>1.63606*</td>
<td>2.03362*</td>
<td>1.08289*</td>
<td>1.91461*</td>
<td>0.81809*</td>
<td>0.55424</td>
</tr>
<tr>
<td>$\beta(1000)$</td>
<td>1.15312*</td>
<td>1.82195*</td>
<td>1.10032*</td>
<td>1.80803*</td>
<td>1.15967*</td>
<td>1.14666</td>
</tr>
<tr>
<td>3.0</td>
<td>1.32289*</td>
<td>1.58560*</td>
<td>1.02650*</td>
<td>1.56528*</td>
<td>1.15341*</td>
<td>1.18628</td>
</tr>
<tr>
<td>5.0</td>
<td>1.06130*</td>
<td>1.23920*</td>
<td>0.96203*</td>
<td>1.21500*</td>
<td>1.12642*</td>
<td>1.16945</td>
</tr>
<tr>
<td>7.0</td>
<td>1.01446</td>
<td>1.11670</td>
<td>0.95943</td>
<td>1.09605</td>
<td>1.11769</td>
<td>1.13660</td>
</tr>
</tbody>
</table>

for different values of $\beta$; cf. Section A.3.8. Figure 5.2 (a) shows a bivariate histogram based on 1000 realizations of the random vector in (5.21), whereas Figure 5.2 (b) shows the density of the bivariate Gaussian random vector $N(o, \Sigma(\chi))$ with mean $o = (0, 0)^T$ and covariance matrix $\Sigma(\chi)$ whose entries are given by (4.56) in Lemma 4.10. If we interpret briefly the bivariate goodness–of–fit evaluation of the vector $Z^{(2)}_r(\chi)$ given in (5.21), we see from Table 5.5 that values of $\beta$ greater than 3.0 seem to be appropriate. In view of the description of Henze–Zirkler’s test in Section A.3.8, this means that the kernel density estimator involved in the computation of $D_{1000,\beta}$ (cf. Section A.3.8) should have a rather small bandwidth. Also, the choice $\beta = \beta(1000)$, i.e., the choice of $\beta$ adapted to the number of simulations, cf. Section A.3.8, seems not an appropriate choice here. The bivariate histogram in Figure 5.2 (a), based on the observation of 1000 realizations of $Z^{(2)}_r(\chi)$ in $B_1^{200}$, represents the corresponding density of $N(o, \Sigma(\chi))$ as depicted in Figure 5.2 (b) rather well.

5.4.2 Evaluation of Asymptotic Tests

We consider a stationary and isotropic Poisson line process $\Phi_1$ in $\mathbb{R}^2$ with intensity $\lambda > 0$. Assuming that we can observe the intersection points of Poisson lines induced by realizations of $\Phi_1$ through the circular sampling window $B_{900}^2$, we want to estimate the power function $\lambda \mapsto \text{pow}(\lambda)$, where $\text{pow}(\lambda) \in [0, 1]$, of a test with hypothesis

$$H_0 : \lambda \leq \lambda^* \quad \text{versus} \quad H_1 : \lambda > \lambda^* .$$

(5.22)

From the description of (asymptotic) tests for $k$-flat intensities, together with the (asymptotic) confidence intervals $J_{d,r}^{(d)}(\alpha)$ in Theorem 5.2, we get for $d = 2$ and $k = 0$ that $H_0$ is rejected at the significance level $\alpha \in (0, 1)$ whenever $\lambda^* < b_{\tau}^{-1}(J_{0,r}^{(2)}, 2\alpha)$. More explicitly, $H_0$ is rejected if

$$\hat{\lambda}_{0,r} > \frac{1}{\pi} \left( \sqrt{\lambda^*} + \frac{2}{\pi \sqrt{3} \tau} z_{1-\alpha} \right)^4 .$$
Generally, a power function describes the probability of rejecting a null hypothesis. In case of the one-sided hypothesis $H_0 : \lambda \leq \lambda^*$ we note that for any $\lambda \leq \lambda^*$ the value of $\text{pow}(\lambda)$ coincides with the probability of rejecting $H_0$ despite this hypothesis is true. Let $\alpha = 0.05$ and $\lambda^* = 0.1$. Based on 1000 realizations of $\Phi_1$, Figure 5.3 shows the estimated power function of the test described above. The curve in this figure, obtained by connecting the different estimates of $\text{pow}(\lambda)$ at discrete values of $\lambda > 0$, shows a typical behavior. For $\lambda \leq \lambda^*$, the estimates are below the significance level $\alpha$ which is depicted in Figure 5.3 as a horizontal line. For values $\lambda > \lambda^*$ the estimates of $\text{pow}(\lambda)$ tend to one rather quickly.

Figure 5.3: Estimated power function of the test (5.22)
5.4.3 Goodness-of-Fit Tests for Poisson–Voronoi Tessellations

Assume that we observe realizations of stationary and isotropic Poisson–Voronoi tessellations $\Psi$ with intensity $\gamma_{\text{PVT}} = 0.01$. More precisely, analogously to Section 5.4.1, we evaluate the realizations of the random variable $\hat{Y}_{0,n}^{(2)}$, given in (5.20), for realizations of $\Psi$. In accordance with Section 5.3.1, we consider unboundedly increasing sequences

$$\{W_n^2\}_{n \geq 1} = \{B_n^2\}_{n \geq 1} \quad \text{and} \quad \{W_n^2\}_{n \geq 1} = \{[-n,n]^2\}_{n \geq 1}$$

of circular sampling windows $W_n^2 = B_n^2$ and rectangular sampling windows $\tilde{W}_n^2 = [-n,n]^2$, respectively. In view of (5.20), we examine the hypothesis

$$H_0^{(n)} : F_{\hat{Y}_0,n} = F_N(0,1) ,$$

where $F_{\hat{Y}_0,n}$ denotes the cdf of $\hat{Y}_{0,n}^{(2)}$. Figure 5.4 shows histograms, quantile plots, and empirical distribution functions for $\hat{Y}_{0,n}^{(2)}$ observed through both types of sampling windows $W_n^2$ and $\tilde{W}_n^2$, where $n = 100$. Furthermore, Table 5.6 shows the corresponding results of Kolmogorov–Smirnov’s goodness-of-fit test for the significance level $\alpha = 0.05$. Note that the 0.95-quantile $s_{0.95}$ of the (asymptotic) Kolmogorov–Smirnov distribution is $s_{0.95} = 1.36$. Hence, $H_0^{(n)}$ is rejected if the calculated values $T_{1000}^n > 1.36$ and $T_{5000}^n > 1.36$. Alternatively, we reject $H_0^{(n)}$ if $p_{1000} < 0.05$ and $p_{5000} < 0.05$. As we can see by the results given in Table 5.6, $H_0^{(n)}$ is not rejected already for $n = 100$ based on $m = 1000$ realizations of $\hat{Y}_{0,n}^{(2)}$.

Table 5.6: Kolmogorov–Smirnov test based on $m$ simulations of $\hat{Y}_{0,n}^{(2)}$ (test statistic $T_m^n$ and p-values $p_m$; rejection of $H_0^{(n)}$ is denoted by *)

<table>
<thead>
<tr>
<th>$n$</th>
<th>$T_{1000}^n$</th>
<th>$p_{1000}$</th>
<th>$T_{5000}^n$</th>
<th>$p_{5000}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1.09</td>
<td>0.188</td>
<td>2.67*</td>
<td>&lt; 10$^{-3}$</td>
</tr>
<tr>
<td>100</td>
<td>0.99</td>
<td>0.284</td>
<td>1.80*</td>
<td>0.003</td>
</tr>
<tr>
<td>150</td>
<td>0.49</td>
<td>0.972</td>
<td>1.11</td>
<td>0.173</td>
</tr>
<tr>
<td>200</td>
<td>0.69</td>
<td>0.728</td>
<td>1.33</td>
<td>0.059</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n$</th>
<th>$T_{1000}^n$</th>
<th>$p_{1000}$</th>
<th>$T_{5000}^n$</th>
<th>$p_{5000}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1.10</td>
<td>0.178</td>
<td>2.75*</td>
<td>&lt; 10$^{-3}$</td>
</tr>
<tr>
<td>100</td>
<td>1.06</td>
<td>0.214</td>
<td>1.66*</td>
<td>0.008</td>
</tr>
<tr>
<td>150</td>
<td>0.57</td>
<td>0.905</td>
<td>1.40*</td>
<td>0.040</td>
</tr>
<tr>
<td>200</td>
<td>0.61</td>
<td>0.851</td>
<td>1.16</td>
<td>0.134</td>
</tr>
</tbody>
</table>

5.5 Applications in Network Modelling

The results summarized in Sections 5.1 and 5.2 for Poisson line processes (and their induced Poisson line tessellations) as well as in Section 5.3 for Poisson–Voronoi tessellations can be...
Figure 5.4: Histograms, quantile plots and empirical distribution functions for $\hat{Y}_{0,n}^{(2)}$ based on $m = 5000$ simulations of $\Psi$ with intensity $\gamma = 0.01$ in $W_{100}^2$ and $\tilde{W}_{100}^2$. 

(a) Histogram ($W_{100}^2$)  
(b) Histogram ($\tilde{W}_{100}^2$)  
(c) Quantile plot ($W_{100}^2$)  
(d) Quantile plot ($\tilde{W}_{100}^2$)  
(e) Edf ($W_{100}^2$)  
(f) Edf ($\tilde{W}_{100}^2$)
applied to network modelling. For example in the context of telecommunication, the results may serve as tools for statistical analysis of both tessellation models.

From the civil engineering point of view, it is desirable to use tractable models instead of too sophisticated ones in order to get information about the geometric structure and the properties of access networks for example. Both Poisson line tessellations and Poisson–Voronoi tessellations can serve as such models. Beyond, as we will show in Chapter 7 of the present thesis (cf. also [40] and [57]), these models are also realistic with respect to real data, at least as models for main road systems.

Often, engineers prefer to get such geometric structural information in the form of worst case and best case scenario values and they want to test their predictions, which are often obtained first hand by some rules of thumb. Thus, the confidence intervals and tests derived earlier in this chapter contribute to such a risk analysis of structural parameters within the network model.

To name an example, assume that we observe (urban) infrastructure data in a sufficiently large (circular) sampling window $B_r$. Recall that the geometry of telecommunication networks depends strongly on the geometry of such data, especially with respect to the placement of cables. We are in a position to provide lower and upper bounds, $b^{-}_r(\alpha)$ and $b^{+}_r(\alpha)$ say, for the mean total length $\lambda$ of roads, and thus of cables, per unit area. To this end, consider the confidence interval $J_{0,r}^{(2)}(\alpha) = [b^{-}_r(\alpha), b^{+}_r(\alpha)]$ obtained from Theorem 5.2 for $d = 2$ and $k = 0$, where

$$b^{+}_r(\alpha) = \frac{1}{r} \left( \left( \eta_0(B^2_r) \right)^{1/4} + \frac{2 z_{1-\alpha/2}}{\pi \sqrt{3}} \right)^2.$$ 

Therefore, based only on the knowledge of the number of road crossings within a large (circular) region $B^2_r$, the values $b^{-}_r(\alpha)$ and $b^{+}_r(\alpha)$ provide lower and upper bounds for the mean total length of the main road system and thus the cables intersecting a certain subregion $W$ of $B^2_r$. In subsequent steps, such bounds can be used to determine costs as well as capacities of connection.

Moreover, we can also provide a test for the null hypothesis $H_0 : \lambda = \lambda^*$, where $\lambda^*$ denotes some specified value of the mean total length of lines per unit area, versus the alternative hypothesis $H_1 : \lambda \neq \lambda^*$. In the context of telecommunication, $\lambda^*$ can be interpreted as the ratio of the total length of cables and the area (of an urban district) in which the cables are placed. Such values for the cable length are often determined by civil engineers according to their own rules of thumb. Based on the confidence interval $J_{0,r}^{(2)}(\alpha)$ for $r$ large enough, $H_0$ would be rejected if

$$\hat{\lambda}_{0,r} < \frac{1}{\pi} \left( \sqrt{\lambda^*} - z_{1-\alpha/2} \frac{2}{\pi \sqrt{3r}} \right)^4 \quad \text{or} \quad \hat{\lambda}_{0,r} > \frac{1}{\pi} \left( \sqrt{\lambda^*} + z_{1-\alpha/2} \frac{2}{\pi \sqrt{3r}} \right)^4,$$

where $\alpha$ is a suitable significance level.
Chapter 6

Limit Theorems for Stationary Tessellations with Random Inner Cell Structures

Based on results obtained in [51], we consider stationary and ergodic random tessellations $\Psi = \{\Xi_n\}_{n \geq 1} \in \mathbb{R}^d$, where the cells $\Xi_n$ contain some random inner structures. Particularly, we study the asymptotic behavior of (random) functionals acting on the interior of the cells $\Xi_n$, $n \geq 1$. Asymptotic in this context means that the sampling window $W^d_\rho$, $\rho > 0$, through which a realisation of $\Psi$ is observed expands unboundedly as $\rho \uparrow \infty$.

![Cells of $\Psi$ and their inner structure induced by a PLT](image)

More precisely, for each $n \geq 1$, we consider a vector $J_n = (J_n^{(1)}, \ldots, J_n^{(m)})^\top$ of $m \geq 1$ jointly stationary random measures (Chapter 1 of [61]) on $\mathcal{B}(\mathbb{R}^d)$, where the $i$th component $J_n^{(i)}$ of $J_n$ for $i = 1, \ldots, m$ is a functional that describes the inner structure of the $n$th cell $\Xi_n$ of the
tessellation \( \Psi \). We assume that the sequence \( J = \{J_n\}_{n \geq 1} \) is independent of \( \Psi \) and that it consists of iid copies of some generic random vector measure \( J_0 = (J_0^{(1)}, \ldots, J_0^{(m)})^\top \).

For example in the planar case \( d = 2 \), one may study \( J_n = (J_n^{(1)}, J_n^{(2)})^\top \), where \( J_n^{(1)} \) counts the vertices and \( J_n^{(2)} \) measures the (total) length of the edges of some component tessellation \( \Psi_n \), generated independently within each cell \( \Xi_n \) of \( \Psi \). Figure 6.1 shows two examples where \( \Psi \) is a PLT with the inner structure induced by another PLT (Figure 6.1 (a)) and where \( \Psi \) is a PVT with the inner structure induced by a PLT (Figure 6.1 (b)), respectively. In the context of the SLM mentioned in Section 1.1.2 the inner cellular linear network can be thought to model side streets within the cells formed by the main road system.

However, the inner structure in the cells \( \Xi_n, n \geq 1 \), of \( \Psi \) need not necessarily be induced by (component) tessellations. Figure 6.2 shows two examples where the inner structure of the cells of \( \Psi \) is induced by a Poisson point process (Figure 6.2 (a)) and by a cluster process (Figure 6.2 (b)), respectively, i.e., a point process with points of a daughter process being clustered around the points of a parent process. Note the resemblance of these two figures to the spatial placement of subscribers in the SLM as described in Chapter 1.

![Cells of a PLT and their inner structure induced by point processes](image)

Figure 6.2: Cells of a PLT and their inner structure induced by point processes

Suppose that only a single realization of the tessellation \( \Psi \) as well as of the vector measures \( J_1, J_2, \ldots \) is available. This realization is observed through some (presumably large) sampling window \( W_0^d \) and \( J_n \) is observable only in the interior \( \text{int}(\Xi_n \cap W_0^d) \) of \( \Xi_n \cap W_0^d \) for \( n \geq 1 \), where to this end the notation \( \Xi_n \cap W_0^d \) is used also if we mean to denote \( \text{int}(\Xi \cap W_0^d) \).

The main subject we investigate in the present chapter is the vector of cumulative functionals \( Z_{\varrho}^{(d)}(J) = (Z_{1,\varrho}^{(d)}(J), \ldots, Z_{m,\varrho}^{(d)}(J))^\top \), where the components \( Z_{i,\varrho}^{(d)}(J) \) of \( Z_{\varrho}^{(d)}(J) \) are given by

\[
Z_{i,\varrho}^{(d)}(J) = \sum_{n \geq 1} J_n^{(i)}(\Xi_n \cap W_0^d), \quad i = 1, \ldots, m. \tag{6.1}
\]
In a first step, in Section 6.1, we examine first- and second-order moments for the considered functionals. More precisely, we derive the expectation vector $E Z_0^{(d)}(J)$ and the covariance matrix $Cov(Z_0^{(d)}(J))$ of $Z_0^{(d)}(J)$ as well as the asymptotic covariance matrix $\Sigma(J) = \lim_{\varrho \to \infty} Cov(Z_0^{(d)}(J))$ of the vector of normalized functionals

$$
\mathcal{Z}_0^{(d)}(J) = \begin{pmatrix}
\lambda_{j_1}^{(1)}(J) - \lambda_{j_1}^{(1)}(J) / \nu_d(W_0^{d}) \\
\lambda_{j_2}^{(2)}(J) - \lambda_{j_2}^{(2)}(J) / \nu_d(W_0^{d}) \\
\vdots \\
\lambda_{j_m}^{(m)}(J) - \lambda_{j_m}^{(m)}(J) / \nu_d(W_0^{d}) 
\end{pmatrix} \cdot \quad (6.2)
$$

Here, $\lambda_{j}^{(i)} = E J_0^{(i)}([0,1]^d)$ is the intensity of the stationary random measure $J_0^{(i)}$ for $i = 1, \ldots, m$.

In Section 6.2, under some mild integrability conditions, we show strong laws of large numbers, i.e.,

$$
\frac{1}{\nu_d(W_0^{d})} \sum_{n \geq 1} J_0^{(i)}(\Xi_n \cap W_0^{d}) \xrightarrow{a.s.} \lambda_{j}^{(i)} \quad \text{for} \quad i = 1, \ldots, m,
$$

cf. Theorem 6.3. Thus, the vector $Z_0^{(d)}(J)/\nu_d(W_0^{d})$ is a strongly consistent (and unbiased) estimator for the intensity vector $(\lambda_{j_1}^{(1)}, \ldots, \lambda_{j_m}^{(m)})^\top$ of the stationary vector measure $J_0$. The proof of Theorem 6.3 relies on the ergodicity of the tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ and the conditional independence of the random vectors $J_1(\Xi_1 \cap W_0^{d})$, $J_2(\Xi_2 \cap W_0^{d})$, \ldots, given the tessellation $\Psi$. Note that, in order to deal with boundary effects, Theorem 3.2 is needed, which shows that the contribution of those cells of $\Psi$ hitting the boundary $\partial W_0^{d}$ of the sampling window $W_0^{d}$ is asymptotically negligible as $\varrho \uparrow \infty$.

Finally, in Section 6.3, we derive the following multivariate CLT

$$
\mathcal{Z}_0^{(d)}(J) \xrightarrow{d} \mathcal{N}(0, \Sigma(J)), \quad \text{i.e.,} \quad \lim_{\varrho \to \infty} \sup_{x \in \mathbb{R}^m} \left| \mathbb{P}(\mathcal{Z}_0^{(d)}(J) \leq x) - F_{\mathcal{N}(0, \Sigma(J))}(x) \right| = 0,
$$

cf. Theorem 6.4, where $F_{\mathcal{N}(0, \Sigma(J))}$ denotes the cdf of the $(m$-dimensional) Gaussian vector $\mathcal{N}(0, \Sigma(J))$ with zero mean components and covariance matrix $\Sigma(J)$.

We conclude this chapter by Section 6.4, where we consider numerical examples for the case that $\Psi$ is a PLT and where we assume that the inner structure in the cells of $\Psi$ is induced by certain types of tessellations. Note that before we present the numerical examples we show that the integrability conditions which we have to impose throughout this chapter are fulfilled for this specific choice of $\Psi$. Furthermore, in Section 6.5 we point out possible applications in network modelling.

Note that besides the investigations of inner cellular structures as reported in this chapter, random structures can be analysed that are induced on the skeleton of a stationary random (initial) tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$ by nested tessellations; [53]. Particularly in the two-dimensional case $d = 2$ the T-crossings can be studied, i.e., the points where the edges of $\Psi$ are met by the edges of a tessellation nested in each cell $\Xi_n$ for $n \geq 1$. Hence, the results in [53] complement the present results in this chapter since together they describe the behavior of random functionals both in the interior of the cells $\Xi_n$ of $\Psi$ and on their boundaries.
6.1 Moments of Cumulative Random Functionals

Let $\Psi = \{\Xi_n\}_{n \geq 1}$ be an arbitrary stationary and ergodic random tessellation in $\mathbb{R}^d$. Recall from Section 3.2.2 that the tessellation $\Psi$ can be equivalently described as a stationary and ergodic marked point process $X_\Psi = \sum_{n \geq 1} \delta_{\alpha(\Xi_n), \Xi_n}$, where $\Xi_n^d = \Xi_n - \alpha(\Xi_n)$ is the centered $n$th cell of $\Psi$ containing the origin and $\alpha(\Xi_n)$ denotes the associated point of the $n$th cell $\Xi_n$, $n \geq 1$. The intensity $\lambda_\Psi = \mathbb{E}\#\{n: \alpha(\Xi_n) \in [0,1]^d\}$ of $X_\Psi$ is the mean number of associated points per unit area and is assumed to be finite and positive. For each individual cell $\Xi_n$ of $\Psi$, we consider an $m$-dimensional vector $J_n = (J_n^{(1)}, \ldots, J_n^{(m)})^\top$ of jointly stationary random measures in $\mathbb{R}^d$, which describe the inner structure of $\Xi_n$. We assume that the sequence of functionals $\{J_n\}_{n \geq 1}$ is independent of $\Psi$ and consists of iid copies of a generic stationary random vector measure $J_0 = (J_0^{(1)}, \ldots, J_0^{(m)})^\top$, where stationarity implies in particular that $J_0(C + x)^d = J_0(C)$ for any $C \in \mathcal{C}$ and $x \in \mathbb{R}^d$.

We assume that only a single realization of the tessellation $\Psi$ and of the random vectors $J_n(\Xi_n \cap W^d_\varrho)$, $n \geq 1$, can be observed through an averaging sequence $\{W^d_\varrho\}_{\varrho > 0}$ of unboundedly expanding convex sampling windows $W^d_\varrho$ with

$$W^d_\varrho = \varrho W \uparrow \mathbb{R}^d,$$  \hspace{1cm} (6.3)

where $W \subset \mathbb{R}^d$ is a ($d$-dimensional) convex body and $\varrho > 0$ is a scaling factor. Furthermore, we assume that for some fixed radii $r_1, r_2 > 0$ the convex body $W$ satisfies the inclusion

$$B^d_{r_1} \subseteq W \subseteq B^d_{r_2}, \quad 0 < r_1 < r_2 < \infty.$$  \hspace{1cm} (6.4)

The latter condition (6.4) justifies the notation $\uparrow \mathbb{R}^d$ in (6.3), meaning that $W^d_\varrho$ grows uniformly in all directions as $\varrho \uparrow \infty$.

6.1.1 Expectation Vector and Covariance Matrix

To start with, we consider the random vector $Z^{(d)}_\varrho(J) = (Z^{(d)}_\varrho(J), \ldots, Z^{(d)}_\varrho(J))^\top$ with entries being defined by (6.1) and we determine its expectation vector $\mathbb{E}Z^{(d)}_\varrho(J)$. Furthermore, we derive conditions under which the covariance matrix $\text{Cov}(Z^{(d)}_\varrho(J))$ and the limit $\Sigma(J) = \lim_{\varrho \to \infty} \text{Cov}(Z^{(d)}_\varrho(J))$ exist, where $Z^{(d)}_\varrho(J)$ is the normalized vector of functionals introduced in (6.2).

Recall that $P^\alpha$ denotes the family of $d$-polytopes in $\mathbb{R}^d$ with associated point at the origin and $P^\alpha$ is the Palm mark distribution of the marked point process $\sum_{n \geq 1} \delta_{\alpha(\Xi_n), \Xi_n}$ as given in (3.4). To this end we use the notation $\mathbb{E}_\Psi$ in order to denote the (conditional) expectation given a realization of the tessellation $\Psi$. Likewise, the notations $\text{Var}_\Psi$ and $\text{Cov}_\Psi$ are employed; cf. also the proof of the subsequent theorem.
6.1. Moments of Cumulative Random Functionals

**Theorem 6.1** If $\lambda_j^{(i)} = \mathbb{E} J_0^{(i)}([0,1]^d) < \infty$ for each $i = 1, \ldots, m$, then

$$
\mathbb{E} Z_{ij}^{(d)}(J) = \nu_d(W_{ij}^d)(\lambda_j^{(i)}, \ldots, \lambda_j^{(m)})^\top.
$$

(6.5)

Under the additional assumption that

$$
\int_{\mathbb{R}^d} \mathbb{E}(J_0^{(i)}(C))^2 P^o(dC) < \infty, \quad i = 1, \ldots, m,
$$

(6.6)

the covariance matrix $\text{Cov}(Z_{ij}^{(d)}(J)) = \left( \text{Cov}(Z_{i,j}^{(d)}(J), Z_{j,j}^{(d)}(J)) \right)_{i,j=1}^m$ exists with

$$
\text{Cov}(Z_{i,j}^{(d)}(J), Z_{j,j}^{(d)}(J)) = \lambda_j \int_{\mathbb{R}^d} \text{Cov} \left( J_0^{(i)}(C \cap (W_{ij}^d - x)), J_0^{(j)}(C \cap (W_{ij}^d - x)) \right) dx P^o(dC).
$$

(6.7)

Moreover, the asymptotic covariance matrix $\Sigma(J) = \lim_{\rho \to 0} \mathbb{E}(\tilde{Z}_{ij}^{(d)}(J)) = (\sigma_{ij}^2)_{i,j=1}^m$ exists with

$$
\sigma_{ij}^2 = \lambda_j \int_{\mathbb{R}^d} \text{Cov} \left( J_0^{(i)}(C), J_0^{(j)}(C) \right) P^o(dC), \quad i, j = 1, \ldots, m.
$$

(6.8)

**Proof** Let $i \in \{1, \ldots, m\}$. Since the tessellation $\Psi$ and the sequence $\{J_n\}_{n \geq 1}$ of functionals are assumed to be independent, using the notation $\mathbb{E}_\Psi$ introduced just before Theorem 6.1, we have that

$$
\mathbb{E} Z_{i,j}^{(d)}(J) = \mathbb{E} \left( \sum_{n \geq 1} J_n^{(i)}(\Xi_n \cap W_{ij}^d) \right) = \mathbb{E} \left( \sum_{n \geq 1} \mathbb{E}_\Psi \left( J_0^{(i)}(\Xi_n \cap W_{ij}^d) \right) \right).
$$

Thus, $\mathbb{E}_\Psi(J_0^{(i)}(\Xi_n \cap W_{ij}^d))$ is the expectation with respect to $J_0^{(i)}$ when $\Xi_n$ is assumed to be fixed, and because of the stationarity of the random measure $J_0^{(i)}$, the expectation $\mathbb{E}_\Psi(J_0^{(i)}(\Xi_n \cap W_{ij}^d))$ equals $\lambda_j^{(i)} \nu_d(\Xi_n \cap W_{ij}^d)$ by application of Haar’s lemma; cf. Lemma A.1. Since the interior of the cells $\Xi_n$, $n \geq 1$ fills the space $\mathbb{R}^d$ up to a set of Lebesgue measure zero, we have that

$$
\sum_{n \geq 1} \nu_d(\Xi_n \cap W_{ij}^d) = \nu_d(W_{ij}^d),
$$

(6.9)

which proves (6.5). To derive (6.7) we first carry out all rearrangements without paying attention to the existence of the involved integrals and expectations. Questions of existence are examined later on. Let $i, j \in \{1, \ldots, m\}$, then,

$$
\text{Cov} \left( Z_{i,j}^{(d)}(J), Z_{j,j}^{(d)}(J) \right) = \text{Cov} \left( \sum_{n \geq 1} J_n^{(i)}(\Xi_n \cap W_{ij}^d), \sum_{n \geq 1} J_n^{(j)}(\Xi_n \cap W_{ij}^d) \right).
$$

Hence, using once again the (conditional) independence of $\Psi$ and the sequence $\{J_n\}_{n \geq 1}$, we get that

$$
\text{Cov} \left( Z_{i,j}^{(d)}(J), Z_{j,j}^{(d)}(J) \right) = \mathbb{E} \left( \sum_{n \geq 1} \mathbb{E}_\Psi \left( J_n^{(i)}(\Xi_n \cap W_{ij}^d) J_n^{(j)}(\Xi_n \cap W_{ij}^d) \right) \right)
$$

$$
- \mathbb{E} \left( \sum_{n \geq 1} \mathbb{E}_\Psi J_0^{(i)}(\Xi_n \cap W_{ij}^d) \right) \mathbb{E} \left( \sum_{n \geq 1} \mathbb{E}_\Psi J_0^{(j)}(\Xi_n \cap W_{ij}^d) \right).
$$
Splitting up the double sum on the right hand side of the latter equality and combining this with (6.5), we get that
\[
\text{Cov}(Z^{(d)}_{i,\varrho}(J), Z^{(d)}_{j,\varrho}(J)) = \mathbb{E} \left( \sum_{n \geq 1} \text{Cov}\left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon), J_0^{(j)}(\Xi_n \cap W_d^\varepsilon) \right) \right) \\
+ \mathbb{E} \left( \sum_{n, \ell \geq 1}^* \text{Cov}\left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon), J_{\ell}^{(j)}(\Xi_\ell \cap W_d^\varepsilon) \right) \right) \\
- \lambda_j^{(i)} \lambda_j^{(j)} \nu_d^{(j)}(W_d^\varepsilon),
\]
where, recall, \( \sum^* \) denotes summation over all pairs of indices \( n, \ell \) with \( n \neq \ell \). For such pairs, we have that
\[
\mathbb{E} \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon) J_{\ell}^{(j)}(\Xi_\ell \cap W_d^\varepsilon) \right) = \mathbb{E} \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon) \mathbb{E} J_{\ell}^{(j)}(\Xi_\ell \cap W_d^\varepsilon) \right)
\]
because of the assumed independence of \( J_n^{(i)} \) and \( J_{\ell}^{(j)} \) for all \( i, j = 1, \ldots, m \) and \( n \neq \ell \). Therefore,
\[
\mathbb{E} \left( \sum_{n, \ell \geq 1}^* \mathbb{E} \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon) J_{\ell}^{(j)}(\Xi_\ell \cap W_d^\varepsilon) \right) \right) \\
= \mathbb{E} \left( \sum_{n, \ell \geq 1} \mathbb{E} \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon) \mathbb{E} J_{\ell}^{(j)}(\Xi_\ell \cap W_d^\varepsilon) \right) \right) \\
- \mathbb{E} \left( \sum_{n \geq 1} \mathbb{E} \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon) \mathbb{E} J_0^{(j)}(\Xi_n \cap W_d^\varepsilon) \right) \right),
\]
where
\[
\mathbb{E} \left( \sum_{n, \ell \geq 1} \mathbb{E} \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon) \mathbb{E} J_{\ell}^{(j)}(\Xi_\ell \cap W_d^\varepsilon) \right) \right) = \lambda_j^{(i)} \lambda_j^{(j)} \nu_d^{(j)}(W_d^\varepsilon).
\]
Thus, using the notation \( \text{Cov}_\Psi \) introduced before Theorem 6.1 we find that
\[
\text{Cov}(Z^{(d)}_{i,\varrho}(J), Z^{(d)}_{j,\varrho}(J)) = \mathbb{E} \left( \sum_{n \geq 1} \text{Cov}_\Psi \left( J_0^{(i)}(\Xi_n \cap W_d^\varepsilon), J_0^{(j)}(\Xi_n \cap W_d^\varepsilon) \right) \right), \tag{6.10}
\]
Finally, writing \( \Xi_n = \Xi_n^0 + \alpha(\Xi_n) \) and applying Campbell’s theorem, cf. Theorem 2.6, to the stationary marked point process \( \sum_{n \geq 1} \delta_{\alpha(\Xi_n), \Xi_n^0} \), we get that
\[
\text{Cov}(Z^{(d)}_{i,\varrho}(J), Z^{(d)}_{j,\varrho}(J)) \\
= \lambda_\varrho \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \text{Cov} \left( J_0^{(i)}((C + x) \cap W_d^\varepsilon), J_0^{(j)}((C + x) \cap W_d^\varepsilon) \right) P^\varrho(dx) \, \text{d}C \\
= \lambda_\varrho \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \text{Cov} \left( J_0^{(i)}((C \cap W_d^\varepsilon - x)), J_0^{(j)}((C \cap W_d^\varepsilon - x)) \right) \, \text{d}x \, P^\varrho(dC),
\]
where in order to get the latter equality we used Fubini’s theorem, cf. Theorem A.2, and the fact that the covariance \( \text{Cov}(J_0^{(i)}(C_1), J_0^{(j)}(C_2)) \) is invariant under diagonal shifts by the stationarity of the random vector measure \( J_0 = (J_0^{(1)}, \ldots, J_0^{(m)})^\top \), i.e.,
\[
\text{Cov}(J_0^{(i)}(C_1 + x), J_0^{(j)}(C_2 + x)) = \text{Cov}(J_0^{(i)}(C_2), J_0^{(j)}(C_2))
\]
for any \( C_1, C_2 \in \mathcal{C} \) and \( x \in \mathbb{R}^d \). What is left to complete the proof of (6.7) is to justify the steps and changes of integration above. This can be achieved by showing that our integrability condition (6.6) ensures the existence of the second moment
\[
\mathbb{E} \sum_{n \geq 1} \left( J_n^{(i)}(\Xi_n \cap W_e^d) \right)^2 = \lambda \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{E}(J_0^{(i)}(C \cap (W_e^d - x)))^2 \, dx \, P^o(dC)
\]
for each \( i = 1, \ldots, m \). By Fubini’s theorem, cf. Theorem A.2, we have that
\[
\int_{\mathbb{R}^d} (J_0^{(i)}(C \cap (W_e^d - x)))^2 \, dx = \int_{\mathbb{C}} \int_{\mathbb{C}} \nu_d((W_e^d - y) \cap (W_e^d - z)) J_0^{(i)}(dy) J_0^{(i)}(dz)
\]
and therefore,
\[
\mathbb{E} \sum_{n \geq 1} \left( J_n^{(i)}(\Xi_n \cap W_e^d) \right)^2 \leq \lambda \nu_d(W_e^d) \int_{\mathbb{R}^d} \mathbb{E}(J_0^{(i)}(C))^2 \, P^o(dC) < \infty,
\]
for \( i = 1, \ldots, m \), where the existence of the integral of the right hand side is ensured by (6.6). The existence of the other expressions can be obtained by proceeding analogously, i.e., by applying Campbell’s theorem (cf. Theorem 2.6) and using the inequality of Cauchy–Schwarz, cf. Formula (A.6) in Theorem A.5, which allows to show that
\[
\mathbb{E} \sum_{n \geq 1} J_n^{(i)}(\Xi_n \cap W_e^d) J_n^{(j)}(\Xi_n \cap W_e^d)
\]
\[
= \lambda \nu_d(W_e^d) \int_{\mathbb{R}^d} \mathbb{E}(J_0^{(i)}(C \cap (W_e^d - x))) J_0^{(j)}(C \cap (W_e^d - x))) \, dx \, P^o(dC)
\]
\[
\leq \lambda \nu_d(W_e^d) \int_{\mathbb{R}^d} \mathbb{E}(J_0^{(i)}(C)) J_0^{(j)}(C)) \, P^o(dC) < \infty.
\]
Therefore, (6.7) is proven. In order to show (6.8), we first note that the properties (6.3) and (6.4) of the sequence \( \{W_e^d\}_{e>0} \) of convex sampling windows \( W_e^d \) entail that
\[
\lim_{e \to \infty} \frac{\nu_d((W_e^d - y) \cap (W_e^d - z))}{\nu_d(W_e^d)} = 1 \quad \text{for any fixed } y, z \in \mathbb{R}^d.
\]
Furthermore, note that the mixed second moments
\[
\mathbb{E}(J_0^{(i)}(C \cap (W_e^d - x))) J_0^{(j)}(C \cap (W_e^d - x)))
\]
can be bounded from above, similarly as it was done in (6.12) for the second moments. Hence, for any \( C \in \mathcal{P}^\circ \) we get that

\[
\lim_{\rho \to \infty} \int_{\mathbb{R}^d} \frac{J_0^{(i)}(C \cap (W^d_\rho - x)) J_0^{(j)}(C \cap (W^d_\rho - x))}{\nu_d(W^d_\rho)} \; dx \\
= \lim_{\rho \to \infty} \int_{C} \int_{C} \frac{\nu_d((W^d_\rho - y) \cap (W^d_\rho - z))}{\nu_d(W^d_\rho)} \; j_0^{(i)}(dy) j_0^{(j)}(dz) \\
= \int_{C} \int_{C} \lim_{\rho \to \infty} \frac{\nu_d((W^d_\rho - y) \cap (W^d_\rho - z))}{\nu_d(W^d_\rho)} \; j_0^{(i)}(dy) j_0^{(j)}(dz),
\]

where we applied the dominated convergence theorem (cf. Theorem A.1) to interchange integration and the limit. Thus,

\[
\lim_{\rho \to \infty} \int_{\mathbb{R}^d} \frac{\mathbb{E}(J_0^{(i)}(C \cap (W^d_\rho - x)) J_0^{(j)}(C \cap (W^d_\rho - x)))}{\nu_d(W^d_\rho)} \; dx = \mathbb{E}(j_0^{(i)}(C) j_0^{(j)}(C)).
\]

We can see that the latter limit does not depend on \( \rho > 0 \) and that this latter limiting relation remains valid for the first moments resulting from the covariance formula (6.10). By the previously proven result (6.7), we get that

\[
\sigma_{ij}^2 = \lim_{\rho \to \infty} \frac{\text{Cov}(Z^{(d)}_{i,\rho}(J), Z^{(d)}_{j,\rho}(J))}{\nu_d(W^d_\rho)} \\
= \lambda \psi \lim_{\rho \to \infty} \int_{\mathcal{P}^\circ} \int_{\mathbb{R}^d} \frac{\text{Cov}(j_0^{(i)}(C \cap (W^d_\rho - x)), j_0^{(j)}(C \cap (W^d_\rho - x)))}{\nu_d(W^d_\rho)} \; dx \; \rho^\circ(dC),
\]

where the latter expression on the right hand side does not depend on \( \rho > 0 \). If we consider

\[
\left| \int_{\mathbb{R}^d} \frac{\text{Cov}(j_0^{(i)}(C \cap (W^d_\rho - x)), j_0^{(j)}(C \cap (W^d_\rho - x)))}{\nu_d(W^d_\rho)} \; dx \right| \\
\leq \frac{1}{\nu_d(W^d_\rho)} \int_{\mathbb{R}^d} \sqrt{\text{Var}(j_0^{(i)}(C \cap (W^d_\rho - x))) \; \text{Var}(j_0^{(j)}(C \cap (W^d_\rho - x)))} \; dx,
\]

and use the inequalities

\[
\left( \text{Var}X \; \text{Var}Y \right)^{1/2} \leq \frac{1}{2} \left( \text{Var}X + \text{Var}Y \right) \quad \text{and} \quad \text{Var}X \leq \mathbb{E}X^2,
\]

which hold for (\( \mathbb{P} \)-a.s. non-negative) random variables \( X \) and \( Y \), we get that

\[
\left| \int_{\mathbb{R}^d} \frac{\text{Cov}(j_0^{(i)}(C \cap (W^d_\rho - x)), j_0^{(j)}(C \cap (W^d_\rho - x)))}{\nu_d(W^d_\rho)} \; dx \right| \\
\leq \frac{1}{2 \nu_d(W^d_\rho)} \left( \int_{\mathbb{R}^d} \mathbb{E}(j_0^{(i)}(C \cap (W^d_\rho - x)))^2 \; dx + \int_{\mathbb{R}^d} \mathbb{E}(j_0^{(j)}(C \cap (W^d_\rho - x)))^2 \; dx \right) < \infty.
\]
The existence of the latter two integrals is ensured by (6.11) and our integrability condition (6.6). Thus, we get an integrable upper bound for the inner integral of
\[
\lim_{q \to \infty} \int_{p_0} \int_{\mathbb{R}^d} \frac{\operatorname{Cov}(J_0^{(i)}(C \cap (W_q^d - x)), J_0^{(j)}(C \cap (W_q^d - x)))}{\nu_d(W_q^d)} \, dx \, P^o(dC),
\]
and therefore, applying once again the dominated convergence theorem (cf. Theorem A.1),
\[
\sigma_{ij}^2 = \lambda \Psi \int_{p_0} \lim_{q \to \infty} \int_{\mathbb{R}^d} \frac{\operatorname{Cov}(J_0^{(i)}(C \cap (W_q^d - x)), J_0^{(j)}(C \cap (W_q^d - x)))}{\nu_d(W_q^d)} \, dx \, P^o(dC).
\]
Finally, we are in the position to finish the proof of Theorem 6.1 by noting that in view of the latter relation for \( \sigma_{ij}^2 \) in combination with the before shown, we get
\[
\sigma_{ij}^2 = \lambda \Psi \int_{p_0} \operatorname{Cov}(J_0^{(i)}(C), J_0^{(j)}(C)) \, P^o(dC),
\]
which is just assertion (6.8).

Note that the second part of Theorem 6.1 implies that the asymptotic variance of the scalar product \( t^T \tilde{Z}_q^{(d)}(J) = \sum_{i=1}^m t_i \tilde{Z}_q^{(d)}(J) \) exists for any \( t = (t_1, \ldots, t_m)^T \in \mathbb{R}^m \), more precisely, that
\[
\lim_{q \to \infty} \operatorname{Var}(t^T \tilde{Z}_q^{(d)}(J)) = \lim_{q \to \infty} \mathbb{E}(t^T \tilde{Z}_q^{(d)}(J))^2 = \sigma^2(t), \quad (6.13)
\]
where
\[
\sigma^2(t) = t^T \Sigma(J) t = \sum_{i,j=1}^m t_i t_j \sigma_{ij}^2, \quad (6.14)
\]
with \( \Sigma(J) = (\sigma_{ij}^2)_{i,j=1}^m \) given by (6.8).

### 6.1.2 Integrability Conditions

In the following we discuss the integrability conditions in (6.6) with the aim to rewrite them and to have separate conditions on the random measure \( J_0^{(i)} \) for \( i = 1, \ldots, m \) as well as on the Palm mark distribution \( P^o \) of \( X_\Psi = \sum_{n \geq 1} \delta_{[\alpha (\xi_n), \xi_n]} \), i.e., on the typical cell \( \Xi^* \) of the tessellation \( \Psi \). The idea behind is that these separate conditions are easier to check and together imply the conditions in (6.6).

**Lemma 6.1** The following inequalities
\[
(\lambda_i^j)^2 \mathbb{E} \nu_d^2(\Xi^*) \leq \int_{p_0} \mathbb{E}(J_0^{(i)}(C))^2 \, P^o(dC) \leq \mathbb{E} \nu_d^2(\Xi^* \oplus B_d^d) \mathbb{E}(J_0^{(i)}([0,1]^d))^2 \quad (6.15)
\]
hold for each \( i = 1, \ldots, m \). Consequently, (6.6) is satisfied whenever
\[
\mathbb{E} V_k^2(\Xi^*) < \infty \quad \text{and} \quad \mathbb{E}(J_0^{(i)}([0,1]^d))^2 < \infty \quad (6.16)
\]
for any \( k = 1, \ldots, d \) and \( i = 1, \ldots, m \), where \( V_k(\Xi^*) \) denotes the kth intrinsic volume of the typical cell \( \Xi^* \) of \( \Psi \).
Proof  Let $i \in \{1, \ldots, m\}$. Taking into account that $EJ_0^{(i)}(C) = \lambda_j^{(i)}\nu_d(C)$ for any $C \in \mathcal{C}$, the first inequality in (6.15) immediately follows from Jensen’s inequality, cf. Formula (A.7) in Theorem A.5, which reads as $(EJ_0^{(i)}(C))^2 \leq E((J_0^{(i)}(C))^2)$. Namely,

$$(\lambda_j^{(i)})^2 E\nu_d^2(\Xi^*) = \int_{p_o} (EJ_0^{(i)}(C))^2 P_o(dC) \leq \int_{p_o} E(J_0^{(i)}(C))^2 P_o(dC).$$

To prove the second inequality in (6.15) consider the obvious set-theoretic inclusions

$$C \subseteq \bigcup_{z \in \mathbb{Z}^d : C \cap ([0,1]^d + z) \neq \emptyset } ([0,1]^d + z) \subseteq C + B_{\sqrt{d}}^d,$$

which are true for any subset $C$ of $\mathbb{R}^d$, where $\mathbb{Z}^d = \mathbb{Z} \times \ldots \times \mathbb{Z}$ with $\mathbb{Z} = \{\ldots, -1, 0, 1, 2, \ldots \}$ denotes the set of integers. Hence,

$$\#\{z \in \mathbb{Z}^d : C \cap ([0,1]^d + z) \neq \emptyset \} \leq \nu_d(C \oplus B_{\sqrt{d}}^d) = \sum_{k=0}^{d} \kappa_{d-k} V_k(C) d^{(d-k)/2},$$

for each $C \in \mathcal{C}$, where $\kappa_{d-k}$ is defined by (A.3) in Section A.1.3 and where the latter equality is a consequence of Steiner’s formula; cf. Formula (A.20). Thus, using the monotonicity of the set function $E(J_0^{(i)}(\cdot))^2$ and applying subsequently the inequality $(a_1 + \cdots + a_s)^2 \leq s (a_1^2 + \cdots + a_s^2)$ for $s \in \mathbb{N}$, we get that

$$E(J_0^{(i)}(C))^2 \leq E\left(J_0^{(i)}\left(\bigcup_{z \in \mathbb{Z}^d : C \cap ([0,1]^d + z) \neq \emptyset} ([0,1]^d + z)\right)\right)^2 \leq \sum_{z \in \mathbb{Z}^d : C \cap ([0,1]^d + z) \neq \emptyset} E(J_0^{(i)}([0,1]^d + z))^2.$$ 

By the translation invariance of $E(J_0^{(i)}(\cdot))^2$, we can further estimate the latter bound and obtain that

$$E(J_0^{(i)}(C))^2 \leq (\# \{z \in \mathbb{Z}^d : C \cap ([0,1]^d + z) \neq \emptyset \})^2 E(J_0^{(i)}([0,1]^d))^2 \leq \nu_d^2(C \oplus B_{\sqrt{d}}^d) E(J_0^{(i)}([0,1]^d))^2,$$

where the latter and final estimate is due to the estimate in (6.17). This proves the second inequality in (6.15). Notice that by Steiner’s formula, cf. Formula (A.20),

$$E\int_{p_o} \nu_d^2(\Xi^* \oplus B_{\sqrt{d}}^d) = E\left(\sum_{k=0}^{d} \kappa_{d-k} V_k(\Xi^*) d^{(d-k)/2}\right)^2,$$

and hence, together with the inequality of Cauchy–Schwarz, cf. Formula (A.6), (or by application of $(a_1 + \cdots + a_s)^2 \leq s (a_1^2 + \cdots + a_s^2)$ for $s = d + 1$) we have that $E\nu_d^2(\Xi^* \oplus B_{\sqrt{d}}^d) < \infty$ is true if the first integrability condition in (6.16) is satisfied. □
We conclude this section by pointing out that often it might not be possible to directly check whether the first integrability condition in (6.16) is met or not, since this involves determining the second moment \( \mathbb{E} V_k^2(\Xi) \) of the \( k \)th intrinsic volume \( V_k(\Xi^*) \) with respect to the typical cell \( \Xi^* \) of \( \Psi \). However, the isodiametric inequality, cf. Formula (A.21) in Theorem A.17, implies that \( \mathbb{E} V_k^2(\Xi^*) < \infty \) holds for each \( k = 1, \ldots, d \) provided that
\[
\mathbb{E} D^{2d}(\Xi^*) < \infty. \tag{6.18}
\]

### 6.2 Laws of Large Numbers

Recall that the individual ergodic theorem as introduced in Lemma 2.1 in Section 2.3.3 applied to the (stationary and ergodic) marked point process \( \sum_{n \geq 1} \delta_{[\alpha(\Xi_n), \Xi_n]} \), which represents the tessellation \( \Psi \), reads as follows; cf. Proposition 10.2.VI and Corollary 10.2.VII in [31]. For any real-valued measurable function \( h : \mathcal{B}(\mathcal{F}) \cap \mathcal{P}^o \to \mathbb{R} \) with \( h \in L^1(\mathcal{P}^o, \mathcal{B}(\mathcal{F}) \cap \mathcal{P}^o, \mathcal{P}^o) \), cf. Section A.21, we have that
\[
\frac{1}{\nu_d(W_d^\circ)} \sum_{n \geq 1} \mathbb{I}_{W^d_\circ(\alpha(\Xi_n))} h(\Xi^*_n) \xrightarrow{a.s.} \lambda_{\Psi} \mathbb{E} h(\Xi^*) = \lambda_{\Psi} \int_{\mathcal{P}^o} h(C) \mathcal{P}^o(dC). \tag{6.19}
\]

However, as it is always the case in spatial statistics, a single realization of the tessellation \( \Psi = \{ \Xi_n \}_{n \geq 1} \) is observed through a (convex compact) sampling window \( W_\circ^d \). Hence, when we consider spatial averages over cells \( \Xi_n \) of \( \Psi \), we obviously have to deal with boundary effects. This means that we observe cells \( \Xi_n \) which intersect the boundary \( \partial W^d_\circ \) of \( W^d_\circ \) and can therefore be observed only partially. In view of Theorem 3.2 established previously in Section 3.5.1, such boundary effects are taken into account by the laws of large numbers which we are going to present subsequently.

#### 6.2.1 Deterministic Functionals

The main result of this section is Theorem 6.2 which provides the strong consistency of a (not necessarily unbiased) estimator for \( \lambda_{\Psi} \mathbb{E} g(\Xi^*) \) in the case of a non-random, translation invariant, and isotonic functional \( g \) on \( \mathcal{C} \). Note that similar results have already been applied earlier in the literature, for example in the case \( d = 2 \) in [28].

\[ \textbf{Theorem 6.2} \]

Let \( g : \mathcal{C} \to [0, \infty) \) be a \( \mathcal{B}(\mathcal{F}) \)-measurable, nonnegative set function such that \( g(C_1) \leq g(C_2) \) for \( C_1, C_2 \in \mathcal{C} \) with \( C_1 \subseteq C_2 \) and \( g(C) = g(C + x) \) for any \( C \in \mathcal{C} \) and \( x \in \mathbb{R}^d \). If the typical cell \( \Xi^* \) of \( \Psi \) satisfies
\[
\mathbb{E} D^d(\Xi^*) < \infty \quad \text{and} \quad \mathbb{E} g(\Xi^*) < \infty, \tag{6.20}
\]
then
\[
\frac{1}{\nu_d(W_d^\circ)} \sum_{n \geq 1} \mathbb{I}_{\Xi_n \cap W_d^\circ \neq \emptyset} g(\Xi_n \cap W_d^\circ) \xrightarrow{a.s.} \lambda_{\Psi} \mathbb{E} g(\Xi^*). \tag{6.21}
\]
Proof} For notational simplicity we put
\[
T_k = \frac{1}{\nu_I(W^k)} \sum_{n \geq 1} \mathbb{I}_{((\Xi^k_n + \alpha(\Xi_n)) \cap W^k_\ell \neq \emptyset)} g((\Xi^k_n + \alpha(\Xi_n)) \cap W^d_\ell) \quad \text{for} \quad k > 0,
\]
in order to denote the left hand side of (6.21). By rewriting the almost sure convergence, the assertion of Theorem 6.2 is equivalent to
\[
\mathbb{P}
\left(
\sup_{k \geq \ell} |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta
\right)
= \mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta \right\}
\right) \xrightarrow{\ell \to \infty} 0
\] (6.22)
for any \( \delta > 0 \); cf. Theorem A.7. Recall the definition of the two families of events \( \{A_q\}_{q > \delta_0} \) and \( \{B_q\}_{q > \delta_0} \) introduced already in (3.22) and (3.23), respectively, of Section 3.5, i.e.,
\[
A_q = \bigcap_{n \geq 1} \left( \left\{ (\Xi^q_n + \alpha(\Xi_n)) \cap W^d_q = \emptyset \right\} \cup \left\{ \alpha(\Xi_n) \in W^d_{q + q(\ell)} \right\} \right)
\]
and
\[
B_q = \bigcap_{n \geq 1} \left( \left\{ (\Xi^q_n + \alpha(\Xi_n)) \in W^d_q \right\} \cup \left\{ \alpha(\Xi_n) \notin W^d_{q - q(\ell)} \right\} \right).
\]
In this context, \( q(\ell) \) is the value of the function \( q : \ell \mapsto q(\ell) \) at \( q > 0 \), where \( q \mapsto q(\ell) \) is defined implicitly by (3.25) in the proof of Theorem 3.2. Recall that \( q \mapsto q(\ell) \) is non-decreasing with \( q(\ell) \to \infty \) and \( q(\ell)/\ell \downarrow 0 \) as \( \ell \to \infty \). Then, using the standard estimate \( \mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B) \) for any events \( A, B \in \sigma(\Omega) \), we immediately obtain that
\[
\mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta \right\}
\right) \leq \mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta \right\} \cap A_k \cap B_k
\right)
+ \mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta \right\} \cap (A_k \cap B_k)^c
\right).
\]
Now we consider both probabilities on the right hand side of the latter estimate separately.
First, applying again the same standard estimate as before together with the fact that
\[
\mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta \right\} \cap A_k \cap B_k
\right)
= \mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ T_k \geq \lambda_I \mathbb{E}g(\Xi^*) + \delta \right\} \cap A_k \cap B_k
\right) \cup \bigcup_{k \geq \ell} \left\{ T_k \leq \lambda_I \mathbb{E}g(\Xi^*) - \delta \right\} \cap A_k \cap B_k,
\]
we get that
\[
\mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ |T_k - \lambda_I \mathbb{E}g(\Xi^*)| \geq \delta \right\} \cap A_k \cap B_k
\right) \leq \mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ T_k \geq \lambda_I \mathbb{E}g(\Xi^*) + \delta \right\} \cap A_k
\right)
+ \mathbb{P}
\left(
\bigcup_{k \geq \ell} \left\{ T_k \leq \lambda_I \mathbb{E}g(\Xi^*) - \delta \right\} \cap B_k
\right).
\]
Second, since \((A_k \cap B_k)^c = A_k^c \cup B_k^c\) we immediately have that
\[
P\left( \bigcup_{k \geq \ell} \left\{ \left| T_k - \lambda \psi E g(\Xi^*) \right| \geq \delta \right\} \cap (A_k \cap B_k)^c \right) \leq P\left( \bigcup_{k \geq \ell} A_k^c \right) + P\left( \bigcup_{k \geq \ell} B_k^c \right).
\]
Collecting the estimates obtained above, we obtain altogether that
\[
P\left( \bigcup_{k \geq \ell} \left\{ \left| T_k - \lambda \psi E g(\Xi^*) \right| \geq \delta \right\} \right) \leq P\left( \bigcup_{k \geq \ell} \left\{ T_k \geq \lambda \psi E g(\Xi^*) + \delta \right\} \right) + P\left( \bigcup_{k \geq \ell} \left\{ T_k \leq \lambda \psi E g(\Xi^*) - \delta \right\} \right) + P\left( \bigcup_{k \geq \ell} A_k^c \right) + P\left( \bigcup_{k \geq \ell} B_k^c \right).
\]
Taking into account the properties of the functional \(g : C \rightarrow [0, \infty)\), we can verify on \(A_k\) the inequality
\[
T_k \leq \frac{1}{\nu_d(W_k^d)} \sum_{n \geq 1} \mathbb{I}_{W_{k+\nu(q)}^d}(\alpha(\Xi_n)) g(\Xi_n^d)
\]
for each \(k \geq \ell\) with \(\ell\) sufficiently large. Likewise, for \(k \geq \ell\), we have on \(B_k\) that
\[
T_k \geq \frac{1}{\nu_d(W_k^d)} \sum_{n \geq 1} \mathbb{I}_{W_{k-\nu(q)}^d}(\alpha(\Xi_n)) g(\Xi_n^d).
\]
Hence,
\[
P\left( \bigcup_{k \geq \ell} \left\{ T_k \geq \lambda \psi E g(\Xi^*) + \delta \right\} \right) + P\left( \bigcup_{k \geq \ell} \left\{ T_k \leq \lambda \psi E g(\Xi^*) - \delta \right\} \right) \leq \sum_{\kappa \in \{-1, +1\}} P\left( \sup_{k \geq \ell} \frac{1}{\nu_d(W_k^d)} \sum_{n \geq 1} \mathbb{I}_{W_{k+\nu(q)}^d}(\alpha(\Xi_n)) g(\Xi_n^d) - \lambda \psi E g(\Xi^*) \geq \delta \right).
\]
However, the latter sum tends to zero as \(\ell \rightarrow \infty\), again by Theorem A.7, since the spatial ergodic theorem (6.19) states that
\[
\frac{1}{\nu_d(W_{\ell+\nu(q)}^d)} \sum_{n \geq 1} \mathbb{I}_{W_{\ell+\nu(q)}^d}(\alpha(\Xi_n)) g(\Xi_n^d) \xrightarrow{a.s. \ \ell \rightarrow \infty} \lambda \psi E g(\Xi^*)
\]
(6.23)
where the asymptotic behavior of \(q(\rho)\), in particular the fact that \(\frac{q(\rho)}{\rho} \downarrow 0\) as \(\rho \rightarrow \infty\), implies
\[
\frac{\nu_d(W_{\ell+\nu(q)}^d)}{\nu_d(W_{\ell}^d)} = \frac{(\rho + \kappa q(\rho))^{d\nu_d(W_{\ell}^d)}}{\rho^{d\nu_d(W_{\ell}^d)}} = \left( 1 + \kappa \frac{q(\rho)}{\rho} \right)^d \xrightarrow{\rho \rightarrow \infty} 1
\]
(6.24)
for \(\kappa = \pm 1\). Therefore, combining (6.23) and (6.24) yields
\[
\frac{1}{\nu_d(W_{\ell}^d)} \sum_{n \geq 1} \mathbb{I}_{W_{\ell+\nu(q)}^d}(\alpha(\Xi_n)) g(\Xi_n^d) \xrightarrow{a.s. \ \ell \rightarrow \infty} \lambda \psi E g(\Xi^*)
\]
Finally, Lemma 3.2 shows the validity of (6.22). \(\square\)

The following corollary is an immediate consequence of the preceding Theorem 6.2 and will be useful later on.
Corollary 6.2 Let \( J_0 = (J_0^{(1)}, \ldots, J_0^{(m)})^\top \) be a vector of jointly stationary random measures on \( \mathbb{R}^d \) which are independent of \( \Psi \) and satisfy the integrability condition \( (6.6) \) (or \( (6.16) \)) together with \( \mathbb{E}D^d(\Xi^*) < \infty \). Then, for any \( t = (t_1, \ldots, t_m)^\top \in \mathbb{R}^m \),
\[
\frac{1}{\nu_d(W_{\theta}^d)} \sum_{n \geq 1} \mathbb{E}_\Psi(t^\top J_0(\Xi_n \cap W_{\theta}^d))^2 \xrightarrow{a.s.} \theta \rightarrow \infty \lambda_\Psi \mathbb{E}(t^\top J_0(\Xi^*))^2 \quad (6.25)
\]
and
\[
\frac{1}{\nu_d(W_{\theta}^d)} \sum_{n \geq 1} \nu_d^2(\Xi_n \cap W_{\theta}^d) \xrightarrow{a.s.} \theta \rightarrow \infty \lambda_\Psi \nu_d^2(\Xi^*). \quad (6.26)
\]

Proof Both \( (6.25) \) and \( (6.26) \) can be proven by a straightforward application of Theorem 6.2 to each of the choices
\[
g_1(C) = \mathbb{E}
\left(J_0^{(i)}(C) J_0^{(j)}(C)\right) \quad \text{and} \quad g_2(C) = \nu_d^2(C), \quad 1 \leq i, j \leq m, \ C \in \mathcal{C},
\]
respectively, for \( g(\cdot) \).

\[\Box\]

6.2.2 Random Functionals

It turns out that the result of Theorem 6.2 remains true if the (non-random) functional \( g : \mathcal{C} \rightarrow [0, \infty) \) is replaced by the stationary random measures \( J_n^{(i)} \) for \( i = 1, \ldots, m \).

Theorem 6.3 Assume that \( \mathbb{E}D^d(\Xi^*) < \infty \) and \( \lambda_j^{(i)} = \mathbb{E}J_0^{(i)}([0,1)^d) < \infty \) for \( i = 1, \ldots, m \). Then,
\[
\frac{1}{\nu_d(W_{\theta}^d)} \sum_{n \geq 1} J_n^{(i)}(\Xi_n \cap W_{\theta}^d) \xrightarrow{a.s.} \theta \rightarrow \infty \lambda_j^{(i)}, \quad i = 1, \ldots, m. \quad (6.27)
\]

Note that the limit \( \lambda_j^{(i)} \) in \( (6.27) \) corresponds to that of \( (6.21) \), i.e.,
\[
\lambda_j^{(i)} = \lambda_\Psi \nu_{J_0^{(i)}}(\Xi^*), \quad i = 1, \ldots, m,
\]
where \( \lambda_\Psi \) is the intensity of \( \Psi \). The reason for this is that by \( (3.5) \), i.e., \( \lambda_\Psi = (\mathbb{E}\nu_d(\Xi^*))^{-1} \), and due to the stationarity of \( J_0^{(i)} \) and Haar's lemma (cf. Lemma A.1) consequently, we get that
\[
\lambda_\Psi \nu_{J_0^{(i)}}(\Xi^*) = \lambda_\Psi \mathbb{E}J_0^{(i)}([0,1)^d) \mathbb{E}\nu_d(\Xi^*) = \lambda_j^{(i)} \quad \text{for } i = 1, \ldots, m,
\]
provided that the random (generic) vector measure \( J_0 \) and the typical cell \( \Xi^* \) of \( \Psi \) are independent. In accordance with the notation for the (conditional) expectation \( \mathbb{E}_\Psi \) we use, to this end, the notation \( \mathbb{P}_\Psi \) in order to denote the (conditional) probability given the tessellation \( \Psi \).
Proof Let \( i \in \{1, \ldots, m\} \). Using the same arguments as in the proof of Theorem 6.2 together with Lemma 3.2, we get that (6.27) is equivalent to

\[
\frac{1}{\nu_d(W^d)} \sum_{n \geq 1} \mathbb{1}_{W^d}(\alpha(\Xi_n)) J_n^{(i)}(\Xi_n) \xrightarrow{a.s.} \lambda^{(i)}_{\pi}, \tag{6.28}
\]

On the other hand, a simple application of the ergodic average (6.19) to the function \( h(C) = \nu_d(C) \) for \( C \in \mathcal{P}^\circ \) yields that

\[
\frac{1}{\nu_d(W^d)} \sum_{n \geq 1} \mathbb{1}_{W^d}(\alpha(\Xi_n)) \nu_d(\Xi_n) \xrightarrow{\mathbb{P}} \lambda_\Psi \mathbb{E} \nu_d(\Xi^\sigma) = 1, \tag{6.29}
\]

where the latter equality follows from (3.5). Thus, multiplying (6.29) by \( \lambda^{(i)}_{\pi} \) on both sides and combining the result with (6.28), we get that (6.27) is equivalent to \( S_N/\nu_d(W^{(d)}_N) \xrightarrow{a.s.} 0 \), where \( S_N \) can be written as partial sum \( S_N = U_1 + \cdots + U_N \) for \( N \in \mathbb{N} \) with

\[
U_k = \sum_{n \geq 1} \mathbb{1}_{W^d \setminus W^{d}_{k-1}}(\alpha(\Xi_n)) \left( J_n^{(i)}(\Xi_n) - \lambda^{(i)}_{\pi} \nu_d(\Xi_n) \right), \quad k \in \mathbb{N}.
\]

Thus, by Theorem A.7, we have that (6.30) in Lemma 6.3, which is shown and proven subsequently, provides a necessary and sufficient condition for \( S_N/\nu_d(W^{(d)}_N) \xrightarrow{a.s.} 0 \) to hold. Hence, the proof of Theorem 6.3 is finished. \( \Box \)

Before we prove the following lemma which is applied in the proof of Theorem 6.3, we note that in (6.28) the summation is over all cells \( \Xi_n \) of \( \Psi \) with associated point \( \alpha(\Xi_n) \in W^d \). Therefore, in contrast to (6.27), we consider also cells \( \Xi_n \) of \( \Psi \) whose interior int \( \Xi_n \) hits the boundary \( \partial W^d_\pi \) of \( W^d_\pi \). Cells \( \Xi_n \) with \( \alpha(\Xi_n) \notin W^d_\pi \) and int \( (\Xi_n \cap W^d_\pi) \neq \emptyset \) however are not considered. In this regard, the left hand side of (6.28) is an estimator for \( \lambda^{(i)}_{\pi} \) in the sense of the well-known associated point rule of spatial statistics; cf. also Section 7.1.2. Furthermore note that (6.28) is similar to the expression appearing in the individual ergodic theorem as stated in (6.19), except that \( J_n^{(i)} \) is also random.

**Lemma 6.3** Let \( S_k = \sum_{j=1}^{k} U_j \) for \( k \in \mathbb{N} \), where

\[
U_j = \sum_{n \geq 1} \mathbb{1}_{W^d \setminus W^{d}_{j-1}}(\alpha(\Xi_n)) \left( J_n^{(i)}(\Xi_n) - \lambda^{(i)}_{\pi} \nu_d(\Xi_n) \right), \quad j \in \mathbb{N}.
\]

for \( 1 \leq i \leq m \). Under the conditions of Theorem 6.3, there exists an integer \( \ell_0 = \ell_0(\delta, \delta^\prime) \) for any given \( \delta, \delta^\prime > 0 \) such that

\[
\mathbb{P}\left( \sup_{k \geq \ell} \nu_d(W^d_k) \geq \delta \right) \leq \delta^\prime \quad \text{for any } \ell \geq \ell_0. \tag{6.30}
\]
Proof Let $i \in \{1, \ldots, m\}$ and, for brevity, $a_k = \nu_d(W_k^d)$, where $k \in \mathbb{N}$. For $\varepsilon > 0$, below chosen as function of $\delta$ and $\delta'$, we introduce the truncated random variables

$$U_k^{(\varepsilon)} = \sum_{n \geq 1} 1_{W_k^d \setminus W_{k-1}^d}(\alpha(\Xi_n)) \left( J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n) \right) 1_{\{ |J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n)| \leq \varepsilon (a_k \vee a_\ell) \}}$$

for $k = 1, 2, \ldots$, and their partial sums $S_N^{(\varepsilon)} = U_1^{(\varepsilon)} + \cdots + U_N^{(\varepsilon)}$ for $N \geq 1$. Note that the random variables $U_k^{(\varepsilon)}$ are conditionally independent given the tessellation $\Psi = \{\Xi_n\}_{n \geq 1}$. Since $S_N^{(\varepsilon)}(\omega) = S_N(\omega)$ for any $N \geq 1$, whenever

$$\omega \in A_{\varepsilon, \ell}(\Psi) = \bigcap_{k \geq 1} \bigcap_{n, \alpha(\Xi_n) \in W_k^d \setminus W_{k-1}^d} \left\{ |J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n)| \leq \varepsilon (a_k \vee a_\ell) \right\}, \quad (6.31)$$

it follows by a standard estimate that, for any fixed $\delta > 0$,

$$\left| \mathbb{P}\left( \sup_{k \geq \ell} \frac{S_k}{\alpha_k} \geq \delta \right) - \mathbb{P}\left( \sup_{k \geq \ell} \frac{S_k^{(\varepsilon)}}{\alpha_k} \geq \delta \right) \right| \leq \mathbb{P}\left( A_{\varepsilon, \ell}^{c}(\Psi) \right). \quad (6.32)$$

Using the definition of $A_{\varepsilon, \ell}(\Psi)$ in (6.31) in order to rewrite its complement $A_{\varepsilon, \ell}^{c}(\Psi)$ reveals that due to the subadditivity of probability measures,

$$\mathbb{P}\left( A_{\varepsilon, \ell}^{c}(\Psi) \right) \leq \sum_{k \geq 1} \mathbb{P}\left( \sum_{n \geq 1} 1_{W_k^d \setminus W_{k-1}^d}(\alpha(\Xi_n)) 1_{\{ |J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n)| \geq \varepsilon (a_k \vee a_\ell) \}} \right). \quad (6.33)$$

Note that the sum

$$\sum_{n \geq 1} 1_{W_k^d \setminus W_{k-1}^d}(\alpha(\Xi_n)) 1_{\{ |J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n)| \geq \varepsilon (a_k \vee a_\ell) \}}$$

that occurs in (6.33) is a discrete random variable taking values in $\mathbb{N}_0$. Hence,

$$\mathbb{P}\left( \sum_{n \geq 1} 1_{W_k^d \setminus W_{k-1}^d}(\alpha(\Xi_n)) 1_{\{ |J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n)| \geq \varepsilon (a_k \vee a_\ell) \}} \geq 1 \right)
\leq \mathbb{E}\left( \sum_{n \geq 1} 1_{W_k^d \setminus W_{k-1}^d}(\alpha(\Xi_n)) 1_{\{ |J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n)| \geq \varepsilon (a_k \vee a_\ell) \}} \right)$$

Let $T = T(\Xi^*)$ be the non-negative random variable $T = |J_0^{(i)}(\Xi^*) - \lambda_j^{(i)} \nu_d(\Xi^*)|$ for fixed $i$ with expectation $\mathbb{E}T < \infty$. Applying Campbell’s theorem (cf. Theorem 2.6) to the expectation on the right-hand side of the latter estimate, we get that

$$\mathbb{P}(A_{\varepsilon, \ell}^{c}(\Psi)) \leq \lambda_{\Psi} \sum_{k \geq 1} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 1_{W_k^d \setminus W_{k-1}^d}(x) 1_{\{T(C) > \varepsilon (a_k \vee a_\ell)\}} \, dx \, P^\alpha(dC),$$

where $P^\alpha$ is the Palm mark distribution of $X_\Psi$ and $P^\alpha$ is the family of $d$-polytopes with associated point centered at the origin. Note that $\int_{\mathbb{R}^d} 1_{W_k^d \setminus W_{k-1}^d}(x) \, dx = a_k - a_{k-1}$ and $\mathbb{P}(A) = \mathbb{E}I_A$ for any $A \in \sigma(\Omega)$. Then, we get that

$$\mathbb{P}(A_{\varepsilon, \ell}^{c}(\Psi)) \leq \lambda_{\Psi} \alpha_\ell \mathbb{P}(T > \varepsilon a_\ell) + \lambda_{\Psi} \sum_{k > \ell} (a_k - a_{k-1}) \mathbb{P}(T > \varepsilon a_k).$$
by splitting the sum \( \sum_{k \geq 1} \) and by using
\[
\sum_{k=1}^{\ell} \int_{\mathbb{R}^d} \mathbb{I}_{W_k \backslash W_{k-1}}(x) dx = \int_{\mathbb{R}^d} \mathbb{I}_{W_\ell}(x) dx, \quad \ell \geq 1.
\]
Finally, we obtain
\[
P(A_{\varepsilon, \ell}^c(\Psi)) \leq \frac{\lambda \Psi}{\varepsilon} \mathbb{E}(T \mathbb{I}_{\{T > \varepsilon a_k\}})
\]
and thus, \( P(A_{\varepsilon, \ell}^c(\Psi)) \to 0 \) as \( \ell \to \infty \). Furthermore, we have that
\[
P\left( \sup_{k \geq \ell} \frac{|S_k^{(e)}|}{a_k} \geq \delta \right) \leq \mathbb{E}\left( P \left( \sup_{k \geq \ell} \frac{|S_k^{(e)} - \mathbb{E} \Psi S_k^{(e)}|}{a_k} \geq \frac{\delta}{2} \right) \right) + \mathbb{E}\left( P \left( \sup_{k \geq \ell} \frac{\mathbb{E} \Psi S_k^{(e)}}{a_k} \geq \frac{\delta}{2} \right) \right),
\]
(6.34)
where \( P \Psi \) is the notation introduced immediately after Theorem 6.3, i.e., \( P \Psi \) denotes the probability measure that refers to \( J_n^{(i)} \) given the tessellation \( \Psi \). To estimate the first term on the right-hand side of (6.34), we make use of the Hájek–Rényi inequality, cf. Theorem A.6, applied to the conditionally independent random variables \( T_k = U_k^{(e)} - \mathbb{E} \Psi U_k^{(e)} \) which have conditional mean zero, i.e., \( \mathbb{E} \Psi T_k = 0 \). In view of Theorem A.6 with \( m = \ell \) and \( c_k = 1/a_k \) for \( k \geq \ell \) we have that
\[
P\left( \sup_{k \geq \ell} \frac{|S_k^{(e)} - \mathbb{E} \Psi S_k^{(e)}|}{a_k} \geq \delta \right) \leq \frac{4}{\delta^2} \frac{\mathbb{E} \text{Var} \Psi (S_k^{(e)})}{a_k^2} + \frac{4}{\delta^2} \sum_{k > \ell} \mathbb{E} \text{Var} \Psi (U_k^{(e)})
\]
(6.35)
Having in mind that, given the tessellation \( \Psi = \{\Xi_n\}_{n \geq 1} \), the random variables \( J_n^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n) \) with mean zero for \( n \geq 1 \) are mutually independent, it is clear that
\[
\text{Var} \Psi (S_k^{(e)}) \leq \sum_{n \geq 1} \mathbb{I}_{W_k^{(e)}}(\alpha(\Xi_n)) \mathbb{E} \Psi \left( (J_0^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n))^2 \mathbb{I}_{\{J_0^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n) \leq \varepsilon a_k\}} \right).
\]
and, for \( k > \ell \),
\[
\text{Var} \Psi (U_k^{(e)}) \leq \sum_{n \geq 1} \mathbb{I}_{W_k^{(e)} \backslash W_{k-1}^{(e)}}(\alpha(\Xi_n)) \mathbb{E} \Psi \left( (J_0^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n))^2 \mathbb{I}_{\{J_0^{(i)}(\Xi_n) - \lambda_j^{(i)} \nu_d(\Xi_n) \leq \varepsilon a_k\}} \right).
\]
Recall again that \( T = T(\Xi^*) \) denotes the non–negative random variable \( T = |J_0^{(i)}(\Xi^*) - \lambda_j^{(i)} \nu_d(\Xi^*)| \) for fixed \( i \) with \( \mathbb{E} T < \infty \). In view of the latter two bounds for \( \text{Var} \Psi (S_k^{(e)}) \) and \( \text{Var} \Psi (U_k^{(e)}) \), respectively, and by application of Campbell’s theorem (cf. Theorem 2.6), we get that
\[
\frac{\mathbb{E} \text{Var} \Psi (S_k^{(e)})}{a_k^2} + \sum_{k > \ell} \frac{\mathbb{E} \text{Var} \Psi (U_k^{(e)})}{a_k^2} \leq \frac{\lambda \Psi}{\varepsilon} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{W_k^{(e)}}(x) \mathbb{E} \Psi (T(C))^2 \mathbb{I}_{\{T(C) \leq \varepsilon a_k\}} dx \, P^o(dC)
\]
\[
+ \frac{\lambda \Psi}{\varepsilon} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{W_k^{(e)} \backslash W_{k-1}^{(e)}}(x) \mathbb{E} \Psi (T(C))^2 \mathbb{I}_{\{T(C) \leq \varepsilon a_k\}} dx \, P^o(dC).
\]
Noting that \( \int_{\mathbb{R}^d} \mathbb{I}_{W_k^d}(x) dx = a_\ell \) and \( \int_{\mathbb{R}^d} \mathbb{I}_{W_k^{d_1}W_{k-1}^{d_1}}(x) dx = a_k - a_{k-1} \), we get
\[
\mathbb{E}\text{Var}_\psi(S^{(e)}_\ell) + \sum_{k>\ell} \frac{\mathbb{E}\text{Var}_\psi(U^{(e)}_k)}{a_k^2} \leq \frac{\lambda_\psi}{a_\ell} \mathbb{E}(T^2 \mathbb{I}_{(T \leq \varepsilon a_\ell)}) + \lambda_\psi \sum_{k>\ell} \frac{a_k - a_{k-1}}{a_k a_{k-1}} \mathbb{E}(T^2 \mathbb{I}_{(T \leq \varepsilon a_k)}),
\]
where we used additionally the inequality \( a_k^2 \geq a_k a_{k-1} \) which holds for any \( k \geq 1 \). Regarding the estimate on the right hand side of the latter relation, we obtain that
\[
\mathbb{E}\text{Var}_\psi(S^{(e)}_\ell) + \sum_{k>\ell} \frac{\mathbb{E}\text{Var}_\psi(U^{(e)}_k)}{a_k^2}
\leq \lambda_\psi \varepsilon ET + \lambda_\psi \frac{a_{\ell+1}}{a_\ell} \mathbb{E}(T \mathbb{I}_{(T \leq \varepsilon a_\ell)}) + \lambda_\psi \sum_{k>\ell} \frac{a_{k+1}}{a_k} \mathbb{E}(T \mathbb{I}_{(\varepsilon a_k < T \leq \varepsilon a_{k+1})})
\leq \lambda_\psi \varepsilon \left(1 + \frac{a_{\ell+1}}{a_\ell}\right) ET.
\]
Since, for any \( \ell \geq 1 \),
\[
\frac{a_{\ell+1}}{a_\ell} = \frac{\nu_d(W_{\ell+1}^d)}{\nu_d(W_{\ell}^d)} = \frac{(\ell+1)^d}{\ell^d} = (1 + 1/\ell)^d \leq 2^d
\]
we get, with the estimate obtained before, that
\[
\lambda_\psi \varepsilon \left(1 + \frac{a_{\ell+1}}{a_\ell}\right) ET \leq \frac{\delta \delta'}{12},
\]
provided that we put \( \varepsilon = \delta'/(12(1 + 2^d) \lambda_\psi ET) \). Thus, the right hand side of (6.35) does not exceed \( \delta'/3 \) for \( \ell \geq 1 \). Furthermore, to consider the second term on the right hand side of (6.34), note that \( |\mathbb{E}\psi S^{(e)}_k| \leq Z^{(e)}_{k,\ell} \) for any \( k \geq \ell \), where
\[
Z^{(e)}_{k,\ell} = \sum_{n \geq 1} \mathbb{I}_{W_k^d(a(\Xi_n))} \mathbb{E}_\psi \left( |J^{(i)}_0(\Xi_n) - \lambda^{(i)}_j \nu_d(\Xi_n)| \mathbb{I}_{|\lambda^{(i)}_j(\Xi_n) - \lambda^{(i)}_j \nu_d(\Xi_n)| > \varepsilon a_\ell} \right). \tag{6.36}
\]
With the above choice of \( \varepsilon > 0 \), take \( \ell_1 = \ell_1(\delta, \delta') \) to be the smallest integer such that \( \lambda_\psi ET \mathbb{I}_{(T > \varepsilon a_\ell)} \leq (\delta \wedge \delta')/3 \). This implies that max\( \ell \geq \ell_1 \) \( \mathbb{P}(A_{\varepsilon,\ell}(\Psi)) \leq \delta'/3 \) and, for \( \ell \geq \ell_1 \),
\[
\mathbb{P} \left( \sup_{k \geq \ell} \frac{|\mathbb{E}_\psi S^{(e)}_k|}{a_k} \geq \frac{\delta}{2} \right) \leq \mathbb{P} \left( \sup_{k \geq \ell} \frac{Z^{(e)}_{k,\ell_1}}{a_k} \geq \frac{\delta}{2} \right) \leq \mathbb{P} \left( \sup_{k \geq \ell} \frac{|Z^{(e)}_{k,\ell_1} - a_k \lambda_\psi ET \mathbb{I}_{(T > \varepsilon a_\ell)}|}{a_k} \geq \frac{\delta}{6} \right). \tag{6.37}
\]
However, it is easy to rewrite the expectation \( \mathbb{E}_\psi(\cdot) \) occurring in (6.36) as a function depending on the (centered) cells \( \Xi_n \). Therefore, the ergodic theorem (6.19) can be applied and yields \( (Z^{(e)}_{N,\ell_1}/a_N) \xrightarrow{a.s.} \lambda_\psi ET \mathbb{I}_{(T > \varepsilon a_\ell)} \) as \( N \to \infty \). Hence, there exists an integer \( \ell_0 = \ell_0(\delta, \delta') \) (larger than \( \ell_1 \)) such that the right hand side of (6.37) becomes smaller than \( \delta'/3 \) for any \( \ell \geq \ell_0 \). Together with the other estimates above and combined with (6.32)–(6.35), this yields (6.30). □
6.3 Multivariate CLT

In this section we prove a central limit theorem, which states asymptotic normality of the normalized random vector \( \tilde{Z}^{(d)}_{\varrho}(J) = (\tilde{Z}^{(d)}_{1,\varrho}(J), \ldots, \tilde{Z}^{(d)}_{m,\varrho}(J))^{\top} \) as \( \varrho \uparrow \infty \), where \( \tilde{Z}^{(d)}_{\varrho}(J) \) has been defined in (6.2).

**Theorem 6.4** Suppose that the conditions (6.6) (or (6.16)) as well as \( \mathbb{E}D^2(\Xi^*) < \infty \) are satisfied. Moreover, assume that the asymptotic covariance matrix \( \Sigma(J) = (\sigma_{ij}^2)_{i,j=1}^m \) given in (6.8) is distinct from the null matrix, i.e., \( \max_{1 \leq i \leq m} \sigma_{ii}^2 > 0 \). Then,

\[
\tilde{Z}^{(d)}_{\varrho}(J) \xrightarrow{d, \varrho \to \infty} N(o, \Sigma(J)),
\]

where \( N(o, \Sigma(J)) \) is an \((m\text{-dimensional})\) Gaussian vector with mean zero and covariance matrix \( \Sigma(J) \).

To make the proof of Theorem 6.4 more transparent, we first collect some auxiliary results in Section 6.3.1 and postpone the main part of the proof to Section 6.3.2 below.

In the particular case \( d = 2 \), the isodiametric inequality, cf. Formula A.21, implies that the second moment of the perimeter \( V_1(\Xi^*) \) of the typical cell \( \Xi^* \) of \( \Psi \) exists if and only if the second moment of the diameter \( D(\Xi^*) \) of \( \Xi^* \) exists. Therefore, in this case, \( \mathbb{E}D^2(\Xi^*) < \infty \) and \( \mathbb{E}d^2(\Xi^*) < \infty \) are the only conditions on the typical cell \( \Xi^* \) of \( \Psi \), which are needed to show (6.38).

### 6.3.1 Some Auxiliary Results

For any fixed vector \( t = (t_1, \ldots, t_m)^{\top} \in \mathbb{R}^m \), consider \( \sigma^2(t) = t^{\top} \Sigma(J) t \geq 0 \) as given in (6.14) and let

\[
B^2_{\varrho}(t, \Psi) = \sum_{n \geq 1} \mathbb{E}_\Psi \left( \sum_{i=1}^m t_i \left( j_{0}^{(i)}(\Xi_n \cap W^d_{\varrho}) - \lambda_j^{(i)} \nu_d(\Xi_n \cap W^d_{\varrho}) \right) \right)^2
\]

in order to introduce, for any \( \varepsilon > 0 \), the event

\[
E_\varrho(t, \varepsilon) = \left\{ \left| B^2_{\varrho}(t, \Psi) - \sigma^2(t) \nu_d(W^d_{\varrho}) \right| < \varepsilon \nu_d(W^d_{\varrho}) \right\}.
\]

**Lemma 6.4** Under the conditions of Theorem 6.4, for any \( t \in \mathbb{R}^m \) and \( \varepsilon > 0 \),

\[
\lim_{\varrho \to \infty} \mathbb{P}(E_\varrho^c(t, \varepsilon)) = 0.
\]

**Proof** First we note that (6.39) states convergence in probability of the ratio \( B^2_{\varrho}(t, \Psi)/\nu_d(W^d_{\varrho}) \) to \( \sigma^2(t) \) as \( \varrho \to \infty \). Using the fact that \( \mathbb{E}_\Psi \left( j_{0}^{(i)}(\Xi_n \cap W^d_{\varrho}) \right) = \lambda_j^{(i)} \nu_d(\Xi_n \cap W^d_{\varrho}) \) for any \( i = 1, \ldots, m \), cf. also the proof of Theorem 6.1, it is easy to see that the identity

\[
\mathbb{E}_\Psi \left( \sum_{i=1}^m t_i j_{0}^{(i)}(\Xi_n \cap W^d_{\varrho}) \right) \nu_d(\Xi_n \cap W^d_{\varrho})^{\top} \mathbb{E}_\Psi \left( \sum_{i=1}^m t_i \lambda_j^{(i)} \nu_d(\Xi_n \cap W^d_{\varrho}) \right) \nu_d(\Xi_n \cap W^d_{\varrho})
\]

is valid.
leads to
\[
\mathbb{E}_\Psi\left( \sum_{i=1}^{m} t_i (J_0^{(i)}(\Xi_n \cap W^d_{\theta}) - \lambda_j^{(i)} \nu_d(\Xi_n \cap W^d_{\theta})) \right)^2
\]
\[
= \mathbb{E}_\Psi\left( \sum_{i=1}^{m} t_i (J_0^{(i)}(\Xi_n \cap W^d_{\theta})) \right)^2 - \nu_d^2(\Xi_n \cap W^d_{\theta}) \left( \sum_{i=1}^{m} t_i \lambda_j^{(i)} \right)^2,
\]
which holds for any cell \( \Xi_n \) of \( \Psi \). Thus, by the relations (6.25) and (6.26) of Corollary 6.2,
\[
\frac{B^2_{\theta}(t, \Psi)}{\nu_d(W^d_{\theta})} \xrightarrow{a.s.} \sqrt{\lambda} \mathbb{E}_\Psi\left( \sum_{i=1}^{m} t_i (J_0^{(i)}(\Xi^*)) \right)^2 - \sqrt{\lambda} \mathbb{E}_\Psi \left( \sum_{i=1}^{m} t_i \lambda_j^{(i)} \right)^2 \nu_d^2(\Xi^*) = \sum_{i,j=1}^{m} t_i t_j \sigma_{ij}^2,
\]
where \( \sigma_{ij}^2 = \sqrt{\lambda} (\mathbb{E}_\Psi(J_0^{(i)}(\Xi^*))J_0^{(j)}(\Xi^*) - \lambda_j^{(i)} \lambda_j^{(j)} \nu_d^2(\Xi^*)) \) for \( i, j = 1, \ldots, m \). But these quantities coincide with the entries of the matrix \( \Sigma(J) \) as given in (6.8). In other words, the ratio \( B^2_{\theta}(t, \Psi)/\nu_d(W^d_{\theta}) \) converges almost surely to \( \sigma^2(t) \) as \( \theta \to \infty \). This implies the convergence in probability.

For the rest of this section and in the subsequent Section 6.3.2 we use the abbreviation
\[
\tilde{J}_0(t, C) = t^T J_0(C) - t^T \lambda_j \nu_d(C) = \sum_{i=1}^{m} t_i (J_0^{(i)}(C) - \lambda_j^{(i)} \nu_d(C))
\]
for any \( t = (t_1, \ldots, t_m)^T \in \mathbb{R}^m \) and \( C \in \mathcal{C} \), where \( \lambda_j = (\lambda_j^{(1)}, \ldots, \lambda_j^{(m)})^T \).

**Lemma 6.5** Under the condition (6.6) of Theorem 6.1, for any \( \delta > 0 \),
\[
\frac{1}{\nu_d(W^d_{\theta})} \mathbb{E}\left( \sum_{n \geq 1} \mathbb{E}_\Psi\left( (\tilde{J}_0(t, \Xi_n \cap W^d_{\theta}))^2 \mathbb{I}\{ |\tilde{J}_0(t, \Xi_n \cap W^d_{\theta})| \geq \delta \sqrt{\nu_d(W^d_{\theta})} \} \right) \right) \xrightarrow{\theta \to \infty} 0.
\]

**Proof** Since \( \tilde{J}_0(t, C) = \tilde{J}_0(t, C + x) \) for any \( C \in \mathcal{C} \) and \( x \in \mathbb{R}^d \), by means of Campbell’s theorem for marked point processes, cf. Theorem 2.6, we get that
\[
\mathbb{E}\left( \sum_{n \geq 1} \mathbb{E}_\Psi\left( (\tilde{J}_0(t, \Xi_n \cap W^d_{\theta}))^2 \mathbb{I}\{ |\tilde{J}_0(t, \Xi_n \cap W^d_{\theta})| \geq \delta \sqrt{\nu_d(W^d_{\theta})} \} \right) \right)
\]
\[
= \lambda \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{E}\left( (\tilde{J}_0(t, C \cap (W^d_{\theta} - x)))^2 \mathbb{I}\{ |\tilde{J}_0(t, C \cap (W^d_{\theta} - x))| \geq \delta \sqrt{\nu_d(W^d_{\theta})} \} \right) \, dx \, P^0(dC).
\]
By the definition of \( \tilde{J}_0(t, C) \) introduced in (6.41) together with Cauchy-Schwarz’s inequality, cf. Formula A.6 in Theorem A.5, we get that
\[
|\tilde{J}_0(t, C)| \leq \|t\| (\|J_0(C)\| + \|\lambda_j\| \nu_d(C))
\]
6.3. Multivariate CLT

The latter result immediately allows to estimate \((\tilde{J}_0(\cdot))^2\) such that

\[
(\tilde{J}_0(t, C \cap (W^d_q - x)))^2 \leq 2 \|t\|^2 \left( \sum_{i=1}^{m} (J^{(i)}_0(C \cap (W^d_q - x)))^2 + \|\lambda_J\|^2 \nu^2_q(C \cap (W^d_q - x)) \right),
\]

since

\[
(\|J_0(C \cap (W^d_q - x))\| - \|\lambda_J\| \nu^2_q(C \cap (W^d_q - x)))^2 \geq 0.
\]

Hence,

\[
\int_{\mathbb{R}^d} (\tilde{J}_0(t, C \cap (W^d_q - x)))^2 \, dx \leq 2 \|t\|^2 \left( \sum_{i=1}^{m} \int_{\mathbb{R}^d} (J^{(i)}_0(C \cap (W^d_q - x)))^2 \, dx \right.
\]

\[
\left. + \|\lambda_J\|^2 \int_{\mathbb{R}^d} \nu^2_q(C \cap (W^d_q - x)) \, dx \right),
\]

where the sum on the right hand side can be again estimated by using (6.11). Thus,

\[
\int_{\mathbb{R}^d} (\tilde{J}_0(t, C \cap (W^d_q - x)))^2 \, dx \leq 2 \|t\|^2 \nu_q(W^d_q) \left( \sum_{i=1}^{m} (J^{(i)}_0(C))^2 + \|\lambda_J\|^2 \nu^2_q(C) \right).
\]

Summarizing the above estimates shows that the left hand side of (6.42) is bounded by the product of \(2 \lambda_q \|t\|^2\) and

\[
\int_{\mathbb{R}^d} \mathbb{E} \left( \sum_{i=1}^{m} \left( (J^{(i)}_0(C))^2 + \|\lambda_J\|^2 \nu^2_q(C) \right) \mathbb{I} \left\{ \|t\| \left( \sum_{i=1}^{m} J^{(i)}_0(C) + \|\lambda_J\| \nu_q(C) \right) \geq \delta \sqrt{\nu_q(W^d_q)} \right\} \right) \, dC.
\]

By the integrability conditions in (6.6), the latter expression converges to zero as \(q \to \infty\). \(\square\)

6.3.2 Proof of Theorem 6.4

Recall that \(t^\top \tilde{Z}^{(d)}_q(J) = \sum_{i=1}^{m} t_i \tilde{Z}^{(d)}_q(J, i)\) and \(\sigma^2(t) = t^\top \Sigma(J) t\) for \(t = (t_1, \ldots, t_m)^\top \in \mathbb{R}^m\). The Cramér-Wold device, cf. Theorem A.10, states that the \(m\)-variate CLT in (6.38) can be put down to the (one-dimensional) CLT

\[
t^\top \tilde{Z}^{(d)}_q(J) \underset{q \to \infty}{\overset{d}{\to}} \mathcal{N}(0, \sigma^2(t)). \tag{6.43}
\]

Equivalently, cf. Section A.3.5, the latter CLT in (6.43) can be formulated as

\[
\lim_{q \to \infty} \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( t^\top \tilde{Z}^{(d)}_q(J) \leq x \right) - F_{\mathcal{N}(0,1)} \left( \frac{x}{\sigma(t)} \right) \right| = 0, \tag{6.44}
\]

for all \(t \in \mathbb{R}^m\) with \(\sigma^2(t) > 0\), and \(t^\top \tilde{Z}^{(d)}_q(J) \overset{\mathbb{P}}{\to} 0\) if \(\sigma^2(t) = 0\). The latter holds since

\[
\mathbb{E}(t^\top \tilde{Z}^{(d)}_q(J))^2 \underset{q \to \infty}{\to} \sigma^2(t) \text{ by (6.13)}. \]

Thus, let \(t \in \mathbb{R}^m\) be fixed such that \(\sigma^2(t) > 0\). Since the
random vector measures \( J_n = (J_n^{(1)}, \ldots, J_n^{(m)})^\top \) are mutually independent and independent of the tessellation \( \Psi = \{\Xi_n\}_{n \geq 1} \), we may write that

\[
P \left( t^\top \tilde{Z}_d^n(J) \leq x \right) = \mathbb{E} \left( \mathbb{P}_\Psi \left( \sum_{n \geq 1} t^\top (J_n(\Xi_n \cap W_\theta^d) - \lambda_J \nu_d(\Xi_n \cap W_\theta^d)) \leq x \sqrt{\nu_d(W_\theta^d)} \right) \right),
\]

where \( \lambda_J = (\lambda_J^{(1)}, \ldots, \lambda_J^{(m)})^\top \) and \( \mathbb{P}_\Psi \) is the (conditional) probability given the tessellation \( \Psi \), as introduced after Theorem 6.3 in Section 6.2.2. Recall that we introduced the abbreviation

\[
B_\psi^2(t, \Psi) = \sum_{n \geq 1} \mathbb{E}_\Psi \left( \sum_{i=1}^m t_i (J_i^{(i)}(\Xi_n \cap W_\theta^d) - \lambda_J^{(i)} \nu_d(\Xi_n \cap W_\theta^d)) \right)^2
\]

at the beginning of Section 6.3.1, where we can assume that \( B_\psi^2(t, \Psi) > 0 \) since \( \sigma^2(t) > 0 \) and due to the convergence in (6.40). Furthermore, recall that

\[
\tilde{J}_0(t, C) = t^\top J_0(C) - t^\top \lambda_J \nu_d(C) = \sum_{i=1}^m t_i (J_i^{(i)}(C) - \lambda_J^{(i)} \nu_d(C)), \quad C \subset \mathbb{R}^m
\]

as given in (6.41). The latter abbreviating notations \( B_\psi^2(t, \Psi) \) and \( \tilde{J}_0(t, C) \), together with a generalized version of the Berry–Esseen inequality for sums of independent random variables (cf. Corollary A.2), yield the estimate

\[
\sup_{x \in \mathbb{R}} \left| \mathbb{P}_\Psi \left( \sum_{n \geq 1} t^\top (J_n(\Xi_n \cap W_\theta^d) - \lambda_J \nu_d(\Xi_n \cap W_\theta^d)) \leq x \sqrt{\nu_d(W_\theta^d)} \right) - \Phi \left( \frac{x}{B_\psi(t, \Psi)} \sqrt{\nu_d(W_\theta^d)} \right) \right| \leq a_1 \varepsilon + \frac{a_2}{B_\psi^2(t, \Psi)} \sum_{n \geq 1} \mathbb{E}_\Psi \left( (\tilde{J}_0(t, \Xi_n \cap W_\theta^d))^2 \mathbb{I}_{|\tilde{J}_0(t, \Xi_n \cap W_\theta^d)| \geq x} \right),
\]

where \( \varepsilon \in (0, 1) \) can be chosen arbitrarily small such that \( \varepsilon \leq \frac{3}{4} \sigma^2(t) \) and \( a_1, a_2 > 0 \) are absolute constants. For almost all realizations of (6.40) and by the inequalities

\[
\left| \frac{B_\psi^2(t, \Psi)}{\nu_d(W_\theta^d)} - \sigma^2(t) \right| \leq \varepsilon \leq \frac{3}{4} \sigma^2(t)
\]

we get that

\[
\frac{\sqrt{\nu_d(W_\theta^d)}}{B_\psi(t, \Psi)} \leq \frac{2}{\sigma(t)}
\]

(6.46)

and (by assuming that \( \varepsilon > 0 \) can be chosen appropriately) that

\[
\left| \frac{\sqrt{\nu_d(W_\theta^d)}}{B_\psi(t, \Psi)} - \frac{1}{\sigma(t)} \right| \leq \frac{2 \varepsilon}{\sigma^3(t)}
\]

(6.47)
for all $\varrho > 0$ sufficiently large. The mean value theorem, together with $\max_{x \in \mathbb{R}} F'_{N(0,1)}(x) = 1/\sqrt{2\pi}$, where $F'_{N(0,1)}(\cdot)$ is the first derivative of $F_{N(0,1)}(\cdot)$, yields that

$$
\left| F_{N(0,1)} \left( x \frac{\sqrt{\nu_d(W^d)}^d}{B_\varrho(t, \Psi)} \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right| \leq \frac{|x|}{\sqrt{2\pi}} \left| \frac{\sqrt{\nu_d(W^d)}}{B_\varrho(t, \Psi)} - \frac{1}{\sigma(t)} \right|. \tag{6.48}
$$

Furthermore, by the results derived above, we have that

$$
\left| \mathbb{P} \left( t^T \tilde{Z}^{(d)}_\varrho (J) \leq x \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right| = \left| \mathbb{E} \left( \mathbb{P}_{\Psi} \left( \sum_{n \geq 1} t^T (J_n(\Xi_n \cap W^d) - \lambda_J \nu_d(\Xi_n \cap W^d)) \leq x \sqrt{\nu_d(W^d)} \right) \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right|.
$$

Using the event $E_\varrho(t, \varepsilon)$ that was introduced at the beginning of Section 6.3.1 as

$$
E_\varrho(t, \varepsilon) = \left\{ \left| B^2_\varrho(t, \Psi) - \sigma^2(t) \nu_d(W^d) \right| < \varepsilon \nu_d(W^d) \right\},
$$

the right hand side of the latter equation is bounded from above by the sum of $\mathbb{E} \left( \mathbb{I}_{E_\varrho(t, \varepsilon)} \right) = \mathbb{P} \left( E_\varrho(t, \varepsilon) \right)$ and

$$
\mathbb{E} \left( \mathbb{I}_{E_\varrho(t, \varepsilon)} \right) \left| \mathbb{E}_{\varrho(t, \varepsilon)} \left( \sum_{n \geq 1} t^T (J_n(\Xi_n \cap W^d) - \lambda_J \nu_d(\Xi_n \cap W^d)) \leq x \sqrt{\nu_d(W^d)} \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right|.
$$

Using the triangular inequality however, we are in the position to estimate the latter expression by

$$
\mathbb{E} \left( \mathbb{I}_{E_\varrho(t, \varepsilon)} \right) F_{N(0,1)} \left( x \frac{\sqrt{\nu_d(W^d)}}{B_\varrho(t, \Psi)} \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) + \mathbb{E} \left( \mathbb{I}_{E_\varrho(t, \varepsilon)} \right) \mathbb{P}_{\Psi} \left( \sum_{n \geq 1} t^T (J_n(\Xi_n \cap W^d) - \lambda_J \nu_d(\Xi_n \cap W^d)) \leq x \sqrt{\nu_d(W^d)} \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right|.
$$

Hence, altogether, by the inequalities (6.45), (6.46), (6.47), and (6.48), we get that

$$
\left| \mathbb{P} \left( t^T \tilde{Z}^{(d)}_\varrho (J) \leq x \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right| \leq \mathbb{P}(E_\varrho(t, \varepsilon)) + \frac{2 \varepsilon |x|}{\sqrt{2\pi} \sigma^3(t)} + a_1 \varepsilon + \frac{4 a_2}{\sigma^2(t) \nu_d(W^d)} \mathbb{E} \left( \sum_{n \geq 1} \mathbb{E}_{\varrho(t, \varepsilon)} \left( (\tilde{J}_0(t, \Xi_n \cap W^d))^2 \mathbb{I}_{\left\{ |\tilde{J}_0(t, \Xi_n \cap W^d)| \geq \frac{x}{a_2} \sqrt{\nu_d(W^d)} \right\}} \right) \right).
$$

Assume now that $|x| \geq 1/\sqrt{\varrho}$. Then,

$$
\left| \mathbb{P} \left( t^T \tilde{Z}^{(d)}_\varrho (J) \leq x \right) - F_{N(0,1)} \left( x \frac{\sigma(t)}{\sigma(t)} \right) \right| \leq \mathbb{P}(X > x \frac{\sigma(t)}{\sigma(t)}),
$$
where $X \sim \mathcal{N}(0, 1)$, i.e., $X$ is a Gaussian random variable with mean zero and variance one, having the cdf $F_{\mathcal{N}(0,1)}$. Hence,

$$\left| \mathbf{P}(t^\top \tilde{Z}_d(J) \leq x) - F_{\mathcal{N}(0,1)} \right| \leq \mathbf{P}(t^\top \tilde{Z}_d(J) \geq x) + \mathbf{P}(X > \frac{x}{\sigma(t)})$$

$$\leq \mathbf{P}(\left| t^\top \tilde{Z}_d(J) \right| \geq x) + \mathbf{P}(\left| X \right| > \frac{x}{\sigma(t)}) ,$$

where we used that $\mathbf{P}(Y > t) \leq \mathbf{P}(\left| Y \right| > t)$ for any real-valued random variable $Y$. Thus,

$$\mathbf{P}(\left| t^\top \tilde{Z}_d(J) \right| \geq x) + \mathbf{P}(\left| X \right| > \frac{x}{\sigma(t)}) \leq \mathbf{P}(\left| t^\top \tilde{Z}_d(J) \right| \geq \frac{1}{\sqrt{\varepsilon}}) + \mathbf{P}(\left| X \right| > \frac{x}{\sqrt{\varepsilon} \sigma(t)})$$

by the assumption that $|x| \geq 1/\varepsilon$. Applying Chebyshev’s inequality, cf. Formula A.8 in Theorem A.5, to the latter estimate, we obtain that

$$\left| \mathbf{P}(t^\top \tilde{Z}_d(J) \leq x) - F_{\mathcal{N}(0,1)} \right| \leq \varepsilon \left( \mathbb{E}(t^\top \tilde{Z}_d(J)^2) + \sigma^2(t) \right). \quad (6.49)$$

Hence, estimating

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P}(t^\top \tilde{Z}_d(J) \leq x) - F_{\mathcal{N}(0,1)} \left( \frac{x}{\sigma(t)} \right) \right|$$

$$\leq \sup_{|x| < \frac{1}{\sqrt{\varepsilon}}} \left| \mathbf{P}(t^\top \tilde{Z}_d(J) \leq x) - F_{\mathcal{N}(0,1)} \left( \frac{x}{\sigma(t)} \right) \right|$$

$$+ \sup_{|x| \geq \frac{1}{\sqrt{\varepsilon}}} \left| \mathbf{P}(t^\top \tilde{Z}_d(J) \leq x) - F_{\mathcal{N}(0,1)} \left( \frac{x}{\sigma(t)} \right) \right|$$

we conclude from (6.13), Lemma 6.4 and Lemma 6.5 that

$$\lim_{\varepsilon \to 0} \sup_{x \in \mathbb{R}} \left| \mathbf{P}(t^\top \tilde{Z}_d(J) \leq x) - F_{\mathcal{N}(0,1)} \left( \frac{x}{\sigma(t)} \right) \right| \leq a_1 \varepsilon + \frac{\sqrt{2 \varepsilon}}{\sqrt{\pi} \sigma^3(t)} + 2 \varepsilon \sigma^2(t) ,$$

where the term $2 \varepsilon \sigma^2(t)$ in the latter bound follows from (6.49) if Theorem 6.1 in connection with (6.13) is used. This proves (6.44) and therefore (6.38) since $\varepsilon \in (0, 1)$ can be chosen arbitrarily small.

### 6.4 Numerical Examples

Throughout this section let $d = 2$. As an example we consider the case where the inner cellular structure of $\Psi = \{\Xi_n\}_{n \geq 1}$ is induced again by a tessellation. More precisely, let $\Psi = \Psi_0$ be a stationary and isotropic (initial) PLT with intensity $\lambda_{\Psi_0} = 0.01$. Assume that within the cells of $\Xi_n$ of $\Psi_0$ a component tessellation is nested, which is also a PLT with intensity $\gamma = \gamma_{\text{PLT}} = 0.17725$ for each $n \in \mathbb{N}$. The latter value of $\gamma_{\text{PLT}}$ is due to the choice $\gamma_1^{(2)} = 0.01$ for the mean number of cells per unit area of the nested PLT. Figure 6.3 shows a realization of
the considered PLT/PLT-nesting. Furthermore, suppose that \( m = 2 \) and that in each cell \( \Xi_n \) of \( \Psi_0 \) the (random) vector \( J_n = (J_n^{(1)}, J_n^{(2)})^\top \) of functionals \( J_n^{(1)} \) and \( J_n^{(2)} \), respectively, can be observed, where \( J_n^{(1)} \) counts the vertices of \( \Psi_n \) and \( J_n^{(2)} \) measures the total length of edges of \( \Psi_n \) for \( n \in \mathbb{N} \). In the following we study realizations of the PLT/PLT-nesting in the sampling window \( W_\varrho^2 \), where \( W_\varrho^2 = \varrho W \) with \( W \) being a rectangle.

### 6.4.1 Evaluation of Integrability Conditions

In the example introduced above, the integrability conditions of (6.16), cf. also (6.18) and the remark after Theorem 6.4, are obviously satisfied whenever

\[
\mathbb{E} D^4(\Xi^*) < \infty \quad \text{and} \quad \mathbb{E} \left( J_0^{(i)}([0,1]^2) \right)^2 < \infty, \quad i = 1, 2. \tag{6.50}
\]

The first condition of (6.50) is true because for our particular choice of \( \Psi_0 \), the random variable \( D(\Xi^*) \) has an exponentially bounded tail function, cf. (3.28) in Theorem 3.3. The second condition of (6.50) also holds. For showing this, note that for both (generic) functionals \( J_0^{(1)} \) and \( J_0^{(2)} \) we obviously have that

\[
\mathbb{E} \left( J_0^{(i)}([0,1]^2) \right)^2 \leq \mathbb{E} \left( J_0^{(i)}(B_1^2) \right)^2, \quad i = 1, 2.
\]

The specific choice of the functionals \( J_0^{(1)} \) and \( J_0^{(2)} \) yields that

\[
\mathbb{E} \left( J_0^{(1)}(B_1^2) \right)^2 = \mathbb{E} \left( \eta_0(B_1^2) \right)^2 \quad \text{and} \quad \mathbb{E} \left( J_0^{(2)}(B_1^2) \right)^2 = \mathbb{E} \left( \zeta_1(B_1^2) \right)^2,
\]

where \( \eta_0(\cdot) \) and \( \zeta_1(\cdot) \) are the functionals for the component tessellation \( \Psi_n \) introduced in (4.23) and (4.24), respectively. Regarding once more the second moment formulae (4.33) and (4.34).
of η0(·) and ζ1(·) as stated in the proof of Lemma 4.8, we get that
\[
\mathbb{E}(J_0^{(1)}(B_1^2))^2 = \sum_{j=0}^{2} \frac{(2\gamma)^{4-j}}{j!(2-j)!} \mathbb{E}\left(\chi(H(X_1) \cap H(X_2) \cap B_1^2) \chi(\cap_{i=3-j}^{4-j} H(X_i) \cap B_1^2)\right)
\leq 4\gamma^2(4\gamma^2 + 2\gamma + 1) < \infty.
\]
Likewise,
\[
\mathbb{E}(J_0^{(2)}(B_1^2))^2 = \sum_{j=0}^{1} \frac{(2\gamma)^{2-j}}{j!(1-j)!} \mathbb{E}\left(\nu_1(H(X_1) \cap B_1^2) \nu_1(H(X_{2-j}) \cap B_1^2)\right)
\leq 8\gamma(2\gamma + 1) < \infty.
\]

### 6.4.2 Simulation of Asymptotic Covariance Matrices

The simulation of the asymptotic covariance matrix Σ(J) = (σ^2_{ij})_{i,j=1}^2 with entries σ^2_{ij} given by (6.8) for i, j = 1, 2 can be done using a simulation algorithm consisting of two steps.

First we fix a number n_1 ∈ \mathbb{N}, say, and simulate for each ℓ ∈ \mathbb{N} with 1 ≤ ℓ ≤ n_1 the Palm mark distribution P_0 of the marked-point process representation X_{Ψ0} = \sum_{n \geq 1} \delta_{[\alpha(\Xi_n),\Xi_n]} of the initial tessellation Ψ_0 = {\Xi_n}_{n \geq 1}. This is possible in case Ψ_0 is either a PVT or a PLT since in these cases, the Palm distributions of the generating Poisson processes are known explicitly. More precisely, in both cases Slivnyak’s theorem (cf. Theorem 2.5) can be applied and provides the theoretical basis for the construction of sampling algorithms for the typical cell Ξ^* of Ψ_0. For a detailed discussion of simulation algorithms, we refer to [92] in case Ψ_0 is a PVT, where an algorithm can be based on radial simulation, and to [37] for example in case Ψ_0 is a PLT; cf. also [69] and [70] for further details.

For each of the n_1 realizations ξ^*_ℓ of the typical cell Ξ^* of Ψ_0, denoted by ξ^*_ℓ for ℓ = 1, . . . , n_1, we do simulations of the inner cellular structure. More precisely, we fix a number n_2(ℓ) ∈ \mathbb{N}, say, where n_2(ℓ) is chosen as n_2(ℓ) = n_2 for all ℓ = 1, . . . , n_1 in the following. Then we simulate n_2 times the random measure J_0^{(1)}(ξ^*_ℓ) and J_0^{(2)}(ξ^*_ℓ) in the interior of ξ^*_ℓ by simulating the inner cellular tessellation Ψ_1. This results in two vectors of size n_2 containing sample values of J_0^{(1)}(ξ^*_0) and J_0^{(2)}(ξ^*_0), respectively. Note that by this approach we do not simulate the typical cell of a Ψ_0/Ψ_1-nesting, but rather measure (random) functionals in realizations of the typical cell Ξ^* of the initial tessellation Ψ_0.

We now obtain the simulated covariance estimates \hat{σ}^{(ℓ)}_{XY} for i, j = 1, 2 by using the natural approach
\[
\hat{σ}^{(ℓ)}_{XY} = \frac{1}{n_2 - 1} \left(\sum_{k=1}^{n_2} X_k^{(ℓ)} Y_k^{(ℓ)} - n_2 \overline{X}_n^{(ℓ)} \overline{Y}_n^{(ℓ)}\right)
\]
where (X_1^{(ℓ)}, . . . , X_{n_2}^{(ℓ)}) and (Y_1^{(ℓ)}, . . . , Y_{n_2}^{(ℓ)}) denote two vectors of sample variables and \overline{X}_n^{(ℓ)} and \overline{Y}_n^{(ℓ)} denote the sample mean for ℓ = 1, . . . , n_1. Hence, the (i, j)-th entry σ^2_{ij} of Σ(J),
6.4. Numerical Examples

based on simulation of the typical cell $\Xi^*$ of $\Psi_0$, is obtained by

$$\lambda_{ij} = \frac{1}{n_1} \sum_{\ell=1}^{n_1} s_{ij} \hat{J}(\ell), \quad i, j = 1, 2.$$  

For the PLT/PLT-nesting example, we choose $n_1 = 100000$ as well as $n_2 = 1000$ and get

$$\Sigma(J) = \begin{pmatrix} 0.05006 & 0.35518 \\ 0.35518 & 3.14866 \end{pmatrix}.$$  

(6.51)

Note that the functionals $J_n^{(1)}$ and $J_n^{(2)}$ with the specific choice described above are positively correlated since the (estimated) coefficient of correlation resulting from (6.51) is 0.8946.

6.4.3 Comparison of Empirical and Asymptotic Distributions

In the following, we examine the CLT derived in Theorem 6.4, i.e., we examine the distribution of the bivariate (centered and normalised) functional

$$\tilde{Z}_{\vartheta}^{(2)}(J) = (\tilde{Z}_{1,\vartheta}^{(2)}(J), \tilde{Z}_{2,\vartheta}^{(2)}(J))^\top$$

in the rectangular sampling window $W^2 = \vartheta W$ with $W = [-250,250]^2$. In view of (6.2) we have that

$$\tilde{Z}_{\vartheta}^{(2)}(J) = \left( \frac{1}{500 \vartheta} Z_{1,\vartheta}^{(2)}(J) - 500 \vartheta \lambda_1^{(1)}(J), \frac{1}{500 \vartheta} Z_{2,\vartheta}^{(2)}(J) - 500 \vartheta \lambda_1^{(2)}(J) \right)^\top,$$

with

$$Z_{1,\vartheta}^{(2)}(J) = \sum_{n \geq 1} J_n^{(1)}(\Xi_n \cap W^2_{\vartheta})$$

and

$$Z_{2,\vartheta}^{(2)}(J) = \sum_{n \geq 1} J_n^{(2)}(\Xi_n \cap W^2_{\vartheta}),$$

respectively, by (6.1).

First, we consider both functionals $\tilde{Z}_{1,\vartheta}^{(2)}(J)$ and $\tilde{Z}_{2,\vartheta}^{(2)}(J)$ separately, i.e., we apply univariate goodness-of-fit techniques in order to test the hypothesis

$$H_{0,i}^\vartheta : F_{Z_{i,\vartheta}^{(2)}(J)} = F_N(0, \sigma_{ni}^2)$$

for $i = 1, 2$,

where $\sigma_{ni}^2$ is taken from (6.51).

Figure 6.4 and Figure 6.5 show the results of different graphical examination techniques for $\vartheta = 3$. Tables 6.1 (a) and 6.1 (b) summarize the results using Pearson’s goodness-of-fit test for $\vartheta = 1, 2$, and Tables 6.2 (a) and 6.2 (b) show the results using Kolmogorov–Smirnov’s goodness-of-fit test for $\vartheta = 1, 2, 3$, where in both cases the significance level is $\alpha = 0.05$. The 0.95-quantiles $\chi^2_{29,0.95}$ and $\chi^2_{69,0.95}$ of the $\chi^2_{29}$-distribution and the $\chi^2_{69}$-distribution, respectively are $\chi^2_{29} = 42.56$ and $\chi^2_{69} = 89.39$, whereas the 0.95-quantile $q_{0.95}$ of the (asymptotic) Kolmogorov–Smirnov distribution is given by $q_{0.95} = 1.36$. 
Figure 6.4: Histograms, quantile plots and empirical distribution functions for $\bar{Z}_{1,3}^{(2)}(J)$ and $n$ simulations.
Figure 6.5: Histograms, quantile plots and empirical distribution functions for $\tilde{Z}_{2,3}^{(2)}(J)$ and $n$ simulations
Table 6.1: Values of Pearson’s test statistic \( T_k \) and \( p \)-values \( p_k \) for \( k \) simulations of the vector \( \tilde{Z}^{(2)}_\varphi (J) \). Rejection of \( H^\varphi_{0,i} \), \( i = 1, 2 \) is denoted by *.

(a) Functional \( \tilde{Z}^{(1)}_{1,\varphi}(J) \)

<table>
<thead>
<tr>
<th>( \varphi )</th>
<th>( T_{1000} )</th>
<th>( p_{1000} )</th>
<th>( T_{5000} )</th>
<th>( p_{5000} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43.34*</td>
<td>0.042</td>
<td>99.58*</td>
<td>0.009</td>
</tr>
<tr>
<td>2</td>
<td>36.62</td>
<td>0.156</td>
<td>74.10</td>
<td>0.315</td>
</tr>
</tbody>
</table>

(b) Functional \( \tilde{Z}^{(2)}_{2,\varphi}(J) \)

<table>
<thead>
<tr>
<th>( \varphi )</th>
<th>( T_{1000} )</th>
<th>( p_{1000} )</th>
<th>( T_{5000} )</th>
<th>( p_{5000} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23.60</td>
<td>0.748</td>
<td>83.18</td>
<td>0.117</td>
</tr>
<tr>
<td>2</td>
<td>25.04</td>
<td>0.676</td>
<td>71.00</td>
<td>0.411</td>
</tr>
</tbody>
</table>

Therefore, the decision rule is as follows: In case of Pearson’s test, we reject \( H^\varphi_{0,i} \) for \( i = 1, 2 \) if for the computed values \( T_{1000} \) and \( T_{5000} \) of Pearson’s test statistic \( T_k \) it holds that \( T_{1000} > 42.56 \) and \( T_{5000} > 89.39 \), respectively. Alternatively, we may decide in view of the \( p \)-value \( p_k \), i.e., \( H^\varphi_{0,i} \) is rejected if \( p_{1000} \leq 0.05 \) and \( p_{5000} \leq 0.05 \), respectively. Likewise, in case of Kolmogorov–Smirnov’s test, we reject \( H^\varphi_{0,i} \) for \( i = 1, 2 \) if for the computed values \( T'_{1000} \) and \( T'_{5000} \) of Kolmogorov–Smirnov’s test statistic \( T'_k \) it holds that \( T'_{1000} > 1.36 \) and \( T'_{5000} > 1.36 \), respectively. Again alternatively, \( H^\varphi_{0,i} \) is rejected for \( i = 1, 2 \) if \( p_{1000} \leq 0.05 \) and \( p_{5000} \leq 0.05 \), respectively.

Table 6.2: Values of Kolmogorov–Smirnov’s test statistic \( T'_k \) and \( p \)-values \( p_k \) for \( k \) simulations of the vector \( \tilde{Z}^{(2)}_\varphi (J) \). Rejection of \( H^\varphi_{0,i} \), \( i = 1, 2 \) is denoted by *.

(a) Functional \( \tilde{Z}^{(1)}_{1,\varphi}(J) \)

<table>
<thead>
<tr>
<th>( \varphi )</th>
<th>( T'_{1000} )</th>
<th>( p_{1000} )</th>
<th>( T'_{5000} )</th>
<th>( p_{5000} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.88</td>
<td>0.417</td>
<td>1.15</td>
<td>0.115</td>
</tr>
<tr>
<td>2</td>
<td>1.23</td>
<td>0.097</td>
<td>0.77</td>
<td>0.596</td>
</tr>
<tr>
<td>3</td>
<td>1.20</td>
<td>0.113</td>
<td>0.81</td>
<td>0.532</td>
</tr>
</tbody>
</table>

(b) Functional \( \tilde{Z}^{(2)}_{2,\varphi}(J) \)

<table>
<thead>
<tr>
<th>( \varphi )</th>
<th>( T'_{1000} )</th>
<th>( p_{1000} )</th>
<th>( T'_{5000} )</th>
<th>( p_{5000} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.74</td>
<td>0.638</td>
<td>0.66*</td>
<td>0.027</td>
</tr>
<tr>
<td>2</td>
<td>1.55*</td>
<td>0.016</td>
<td>1.63*</td>
<td>0.010</td>
</tr>
<tr>
<td>3</td>
<td>1.54*</td>
<td>0.017</td>
<td>1.07</td>
<td>0.170</td>
</tr>
</tbody>
</table>

Regarding the results in Tables 6.1 and 6.2 we see that for both functionals \( \tilde{Z}^{(2)}_1(J) \) and \( \tilde{Z}^{(2)}_2(J) \) the null hypotheses \( H^\varphi_{0,1} \) and \( H^\varphi_{0,2} \) are not rejected even for relatively small values of \( \varphi \), i.e., for relatively small sampling windows \( W^2 \). Additionally, also the graphical analysis in Figures 6.4 and 6.5 shows that the distribution of both \( \tilde{Z}^{(2)}_1(J) \) and \( \tilde{Z}^{(2)}_2(J) \) can be assumed to be normal for relatively small sampling windows. Hence, we may assume that the functionals considered in the CLTs of Chapter 6, compared to the functionals in the CLTs of Chapters 4 and 5, possess a faster rate of convergence to their asymptotic normal distribution. This may be explained by the fact that in Chapter 6 we have some additional randomness, since the functionals act on the interior of the cells of some (initial) tessellation, whereas in Chapters 4 and 5 we just consider functionals of a tessellation itself.

Finally, we consider the bivariate case and use Henze–Zirkler’s test, Section A.3.8, in order to
test the distribution of the vector $\tilde{Z}_2(\mathbf{J})$ for bivariate normality, where the significance level is $\alpha = 0.05$. More precisely, we test the hypothesis $H_0^\theta$ stating that $\tilde{Z}_2(\mathbf{J})$ is a Gaussian random vector. Table 6.3 summarizes the results of Henze–Zirkler’s test and Figure 6.6 shows the bivariate histogram as well as the density function of a Gaussian bivariate vector with mean zero and covariance matrix $\Sigma(\mathbf{J})$ given in (6.51). Also in view of Figure 6.6, the histogram of $\tilde{Z}_1(\mathbf{J}) = (\tilde{Z}_{1,1}(\mathbf{J}), \tilde{Z}_{2,1}(\mathbf{J}))^\top$ represents the theoretical density as depicted in Figure 6.6 (b) pretty well. The results point in the same direction as the goodness-of-fit results in the univariate case, i.e., asymptotic normality of the vector $\tilde{Z}_2(\mathbf{J})$ can be assumed for rather small sampling windows $W_\theta$. Additionally, almost all choices for $\beta$ seem to provide a reasonable outcome.

### 6.5 Applications to Network Modelling

As already mentioned at the beginning of this chapter, the obtained results contribute to the analysis of inner cellular structures in the context of the SSLM that has been introduced in Section 1.1.2.

Assume for example that we want to model a two-level hierarchical road system consisting of main roads and side streets. The side streets subdivide the cells (or blocks) formed by the main roads. Then the SSLM provides the $\Psi_0/p\Psi_1$-nesting with Bernoulli thinning as class of tessellation models that reflect the morphological structure of such a two-level hierarchical road system appropriately.

In civil engineering, one is often interested in a block-wise analysis and a block-wise modelling of cable trenches and network equipment. One possibility that adds to such an analysis and
modelling approach is to consider a vector functional \( J_n = (J_n^{(1)}, J_n^{(2)})^\top \) in each cell \( \Xi_n \) of the initial tessellation \( \Psi_0 = \{ \Xi_n \}_{n \geq 1} \) of the \( \Psi_0/p\Psi_1 \)-nesting. Suppose that \( J_n^{(1)} \) and \( J_n^{(2)} \) are the functionals that have been considered already before and recall that \( J_n^{(1)} \) counts the vertices and \( J_n^{(2)} \) measures the total length of edges of \( \Psi_1 \) in \( \Xi_n \) for \( n \geq 1 \). To provide an interpretation, \( J_n^{(1)} \) counts the crossings of side streets within a block formed by the main roads and \( J_n^{(2)} \) measures their total length. Hence, large values of \( J_n^{(1)} \) indicate that the block exhibits a morphological structure with many street crossings and in turn one can expect also many side streets. However, this results also in a larger value of \( J_n^{(2)} \). Therefore, both functionals \( J_n^{(1)} \) and \( J_n^{(2)} \) seem to be strongly correlated. In practice, large values of \( J_n^{(1)} \) or \( J_n^{(2)} \) for a certain block indicate that more cables are required and that more network equipment has to be placed in this block and that therefore the costs of civil engineering are higher than in a block where both functional values are low.

Assume now that we identified a certain type of \( \Psi_0/p\Psi_1 \)-nesting as best fitting model for a given set of real network data, like the data of Paris shown in Figure 1.1; cf. Chapter 7 for details. Suppose for example that these two-level data, observed through some (convex compact) sampling window \( W \), are best described by a PLT/PLT-nesting, a model type that is widely used in practice. Recall that the CLT in Theorem 6.4 states that

\[
\tilde{Z}^{(2)}_\vartheta (J) = \left( \tilde{Z}^{(2)}_{1,\vartheta} (J), \tilde{Z}^{(2)}_{2,\vartheta} (J) \right)^\top \xrightarrow{d} \mathcal{N}(o, \Sigma(J)). \quad (6.52)
\]

Hence we may assume that the centered and normalised vector functional \( \tilde{Z}^{(2)}_\vartheta (J) \) is normally distributed with mean zero and covariance matrix \( \Sigma(J) \) if the sampling window \( W^2_\vartheta \) is sufficiently large. The examination in Section 6.4.3 allows for the conclusion that (6.52) can be

---

Table 6.3: Henze–Zirkler’s test statistic \( T_{k, \beta} \) for \( k \) simulations

<table>
<thead>
<tr>
<th>( k )</th>
<th>( \beta / \varrho )</th>
<th>( T_{k, \beta} )</th>
<th>( q_{0.95}(\beta) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.5</td>
<td>0.057743</td>
<td>0.01648</td>
</tr>
<tr>
<td>1000</td>
<td>1.0</td>
<td>0.478170</td>
<td>0.18863</td>
</tr>
<tr>
<td>1000</td>
<td>3.0</td>
<td>0.926241</td>
<td>0.75513</td>
</tr>
<tr>
<td>1000</td>
<td>( \beta(1000) )</td>
<td>0.943753</td>
<td>0.669137</td>
</tr>
<tr>
<td>5000</td>
<td>0.5</td>
<td>0.088243*</td>
<td>0.10140*</td>
</tr>
<tr>
<td>5000</td>
<td>1.0</td>
<td>0.514686</td>
<td>0.49845</td>
</tr>
<tr>
<td>5000</td>
<td>3.0</td>
<td>0.922674</td>
<td>0.93909</td>
</tr>
<tr>
<td>5000</td>
<td>( \beta(5000) )</td>
<td>0.922551</td>
<td>0.94357</td>
</tr>
<tr>
<td>10000</td>
<td>0.5</td>
<td>0.255134*</td>
<td>0.031513</td>
</tr>
<tr>
<td>10000</td>
<td>1.0</td>
<td>1.227606*</td>
<td>0.220107</td>
</tr>
<tr>
<td>10000</td>
<td>3.0</td>
<td>1.156124</td>
<td>0.995328</td>
</tr>
<tr>
<td>10000</td>
<td>( \beta(10000) )</td>
<td>1.077985</td>
<td>1.049832</td>
</tr>
</tbody>
</table>
assumed for rather small sampling windows. In practice, the size of the sampling window is often much larger and despite the fact that intensities of initial and component tessellation are often smaller than the chosen values in our example, we may assume that $\hat{Z}_0^{(2)}(J)$ is indeed (approximately) normally distributed. In Section 6.4.2 we simulated the (asymptotic) covariance matrix $\Sigma(J)$, the result is given in (6.51). As we can observe, the correlation between $J_0^{(1)}$ and $J_0^{(2)}$ is given by 0.8946. In view of the practical interpretation of both functionals this means that, quite naturally, the more side street crossings are located in a block the larger the total length of side streets and also the more side streets are located in that block. This in turn leads to the interesting conclusion for this PLT / PLT-model that with respect to civil engineering of cable trenches, counting the number of street crossings per block admits a prediction for the required cable length. This again provides a mathematical basis for rules of thumb used by civil engineers for cost calculation and strategic planning.

Now we turn our attention to statistical inference for the vector $J_n = (J_n^{(1)}, J_n^{(2)})^\top$ of (random) inner cellular functionals $J_n^{(1)}$ and $J_n^{(2)}$. Assume that we have complete information about the (initial) tessellation $\Psi_0$, in particular assume that we know its intensity $\lambda_{\Psi_0}$, i.e., the mean number of cells per unit area. Thus, whenever we can simulate the Palm mark distribution of the marked-point process representation $X_{\Psi_0} = \sum_{n \geq 1} \delta_{\xi_n} \xi(\Xi)$ of $\Psi_0$, i.e., whenever we can simulate the typical cell $\Xi^*$ of $\Psi_0$, we get a value for $\lambda_{J_n}^{(i)}$ (where $i = 1, 2$) by

$$\lambda_{J_n}^{(i)} = \lambda_{\Psi_0} \mathbb{E} J_n^{(i)}(\Xi^*), \quad i = 1, 2,$$

in view of Formula (6.2.2). In practice this leads to a two-step algorithm to compute $\lambda_{J_n}^{(i)}$ which is completely analogous to the algorithm described in Section 6.4.2 for the simulation of the covariance matrix $\Sigma(J)$. In case $\Psi_0$ is of a tessellation type where it is not obvious how to obtain $\Xi^*$ by simulation or in case we face computational limitations for such a simulation, we may estimate $\lambda_{J_n}^{(i)}$ for $i = 1, 2$ from a single realization of $\Psi_0$ and $J_n = (J_n^{(1)}, J_n^{(2)})^\top$. To do so, we consider Formula (6.5) in Theorem 6.1 and Formula (6.27) in Theorem 6.3, which state that

$$\hat{\lambda}_{J_n}^{(i)} = \frac{1}{\nu_2(W^2_s)} \sum_{n \geq 1} J_n^{(i)}(\Xi_n \cap W^2_s), \quad i = 1, 2$$

is an unbiased and strongly consistent estimator (as $\varrho \to \infty$) for $\lambda_{J_n}^{(i)}$. Note that in view of Theorem 3.2 this estimator uses only informations that is contained in the (sufficiently large) sampling window $W^2_s$. More precisely, $\hat{\lambda}_{J_n}^{(i)}$ uses also information of such cells $\Xi_n$ of $\Psi_0 = \{\Xi_n\}_{n \geq 1}$ that are observable only partially in the interior of $W^2_s$.

In analogy to Chapter 5 it is also possible to use (6.52) in order to derive (asymptotically exact) confidence intervals for the intensity vector $\lambda_J = (\lambda_{J_n}^{(1)}, \lambda_{J_n}^{(2)})^\top$ of the vector functional $J_n = (J_n^{(1)}, J_n^{(2)})^\top$ as well as to derive tests. Note however that in contrast to the CLTs of Chapters 4 and 5, no explicit formula is known for the entries of $\Sigma(J)$, which are given in (6.8). Hence, estimators $\hat{\Sigma}(J)$ for the asymptotic covariance matrix $\Sigma(J)$ have to be constructed and their properties have to be studied. For this, it is expected that similar approaches as already considered in [20] and [108] will turn out to be useful.
Chapter 7

Statistical Model Fitting of Random Tessellations

In this chapter, we describe an approach to statistical model fitting of random tessellations to real data. More precisely, we consider data that exhibit a network structure in the sense that they consist of connected segments forming cells. Figure 7.1 shows two examples of such data sets. In particular, Figure 7.1 (a) displays a magnified part of Figure 1.1 in Chapter 1, i.e., a magnified part of the road system of Paris. Figure 7.1 (b) shows the graph structure of keratin filament networks observed via scanning electron microscopy. In both cases the data are displayed after preprocessing of the raw data. Such preprocessing steps applied to the observed raw data are often necessary and useful. The road data in Figure 7.1 (a) for example are originally given in Lambert 2 coordinates and have to be transformed to Cartesian coordi-
nates ([107]). The data displayed in Figure 7.1 (b) are originally based on scanning electron microscope (SEM) images and tools from image analysis ([59], [114], [115]), in particular image segmentation algorithms, are used in [13] to extract and recover the graph structure.

The model fitting procedure described in this chapter is based on the minimization of distances between characteristics of the given network data and characteristics of the tessellation models that are to be fitted to these data. In principle there is a wide range of possible tessellation models as well as choices of such characteristics available. However, we will restrict ourselves to such stationary (and isotropic) tessellation models and characteristics, where theoretical mean value relationships are known explicitly. In particular, we consider stationary (and isotropic) PVTs, PDTs, and PLTs as well as the models obtained thereof through (one-fold) nesting and Bernoulli thinning as described in Section 3.4. Note that similar approaches to model fitting have been described earlier in the literature, in particular for random point processes in [117], but also briefly for tessellations in [116] for example. In [34] our approach is modified to include also tessellation models obtained by superposition.

This chapter is structured as follows. In Section 7.1 we consider some preliminaries, namely the employed model characteristics in Section 7.1.1, estimators for these characteristics in Section 7.1.2, and possible choices of distance functions in Section 7.1.3. In the subsequent sections, we describe the model fitting procedure in detail. In order to validate the correctness of our model fitting procedure and in order to analyse its performance, extensive simulation studies have to be done where the given network data are obtained by simulation of theoretical tessellation models. In Section 7.4 we present a test based on Monte Carlo simulation to evaluate the model fitting procedure. We conclude by regarding examples where we fit tessellation models to the road system of Paris. In [12], examples can be found where iterated tessellation models are fitted to the graph structure induced by keratin filament networks. In particular, it is shown that a slight modification of our model fitting algorithm applied to the keratin network data reveals fundamental morphological reorganisation of keratin filament networks in pancreatic cancer cells in response to transforming growth factor $\alpha$ (TGF$\alpha$), which in turn is involved in pancreatic cancer progression.

7.1 Some Preliminaries

7.1.1 Choice of Model Characteristics

The model fitting procedure we propose is based on minimizing the values of a distance function between two vectors of characteristics. The idea is that the entries of these vectors are characteristics that describe the spatial-geometric structure of the given network data and of the tessellation models, respectively. To be more precise, we consider characteristics $z_i$ for $i \geq 1$ which are measured per unit area $[0, 1)^2$, i.e., we consider intensity values. To this end, we regard the vector

$$z = (z_1, z_2, z_3, z_4)^T \quad \text{(7.1)}$$

of the four characteristics $z_1, \ldots, z_4$, where $z_1$ is the mean number of vertices, $z_2$ is the mean number of edges, $z_3$ is the mean number of cells, and $z_4$ is the mean total length of edges. In
case of tessellation models the entries of this vector coincide with the intensities $\gamma^{(0)}, \gamma^{(1)}, \gamma^{(2)}$, and $\tilde{\gamma}^{(1)}$ of facet processes introduced in Section 3.2.3. From the perspective of model fitting, the characteristics $z_1, \ldots, z_4$ can be regarded as global characteristics. They are chosen for two reasons. First, they reflect spatial-geometric structures of networks and, second, in case of our choice of particular tessellation models PVT, PDT, and PLT, mean value relationships of the facet intensities $\gamma^{(0)}, \gamma^{(1)}, \gamma^{(2)}$, and $\tilde{\gamma}^{(1)}$ are relatively simple and can be given explicitly; [71], [126]. Besides the global characteristics it is also possible to examine another category, so-called local characteristics, which refer to single cells of the tessellation structure. Examples of local characteristics comprise the mean edge-length per cell, the mean perimeter per cell, and the mean area per cell. However, it turns out that such characteristics are less useful since unbiased estimators are not always obvious and since these characteristics cannot reflect the structure of the whole tessellation ([57]). Therefore, in the following, we concentrate on the (global) characteristics $z_1, \ldots, z_4$ summarized in the vector $z = (z_1, \ldots, z_4)^T$.

### 7.1.2 Sampling Techniques and Unbiased Estimators

In general, observing samples of spatial data through a (compact) sampling window $W \subset \mathbb{R}^2$, not necessarily convex, leads to the problem of deciding how to count sample profiles. In our context, the profiles are characteristics of the cellular structure with intensities summarized in the vector $z = (z_1, \ldots, z_4)^T$. The difficulty is to find sampling techniques which deal appropriately with boundary effects that occur whenever sample profiles cross the boundary $\partial W$ of the sampling window $W$. In other words, the difficulty is to find unbiased sampling rules. A comprehensive overview of sampling techniques in spatial statistics can be found in [10] and [91], while [116] is a good reference for statistics with respect to spatial random tessellations; cf. also the discussion on asymptotic insignificance of partially observed cells in Section 3.5.1.

![Diagram](a) Plus sampling  
(b) Minus sampling

Figure 7.2: Sampling techniques leading to biased estimation
Generally, we may differentiate between two rather naive sampling approaches. One approach requires more information than is given in the sampling window, whereas the other approach uses less information than is actually provided by the sampling window. The sampling techniques that correspond to these two approaches are called plus sampling and minus sampling, illustrated in Figures 7.2 (a) and 7.2 (b), respectively. Using plus sampling, all profiles that intersect with the sampling window \( W \) are taken into consideration, whereas minus sampling samples only those profiles situated completely inside \( W \). It is clear however, cf. also § 3.5 of [46], that plus sampling systematically overcounts the number of profiles (per unit area), whereas minus sampling undercounts the number of profiles (per unit area). Hence, both sampling techniques introduce a bias to the estimation.

To overcome this sampling bias, several more sophisticated sampling techniques are proposed in literature, like the associated point rule described in § 3.5 of [46] and illustrated by Figure 7.3 (a) or the tiling rule described in [44] and [45] and illustrated by Figure 7.3 (b). The latter technique is suited for sampling windows \( W \) with rectangular shape, where only those profiles are taken into consideration that intersect the upper and/or right boundary of \( W \) or are completely contained in \( W \). To this end, we use the associated point rule, where a (unique) point is associated with each profile and only those profiles with associated point located in the interior of \( W \) are taken into account. For example if the observed data exhibit cellular network structure and we want to sample the cells of this network, a practical choice for an associated point is the lexicographically smallest vertex.

![Figure 7.3: Sampling techniques leading to unbiased estimation](image)

We now turn to the problem of constructing concrete unbiased sampling rules, called estimators in the sequel, for the entries \( z_1, \ldots, z_4 \) of the vector \( z = (z_1, \ldots, z_4)^T \). Hence, in the following, we are going to look for a vector

\[
\hat{z} = \hat{z}(W) = (\hat{z}_1(W), \hat{z}_2(W), \hat{z}_3(W), \hat{z}_4(W))^T,
\]

where \( \hat{z}_1(W), \ldots, \hat{z}_4(W) \) are unbiased estimators for the characteristics \( z_1, \ldots, z_4 \), based on information provided by the sampling window \( W \). Applying the associated point rule, we get
that
\[ \hat{z}_1(W) = \frac{\Psi^{(0)}(W)}{\nu_2(W)}, \quad \hat{z}_2(W) = \frac{\Psi^{(1)}(W)}{\nu_2(W)}, \quad \hat{z}_3(W) = \frac{\Psi^{(2)}(W)}{\nu_2(W)}, \quad \hat{z}_4(W) = \frac{\tilde{\Psi}^{(1)}(W)}{\nu_2(W)}, \]
where, quite in analogy to Section 3.2.3, we let \( \Psi^{(0)}(W) \) denote the number of nodes in \( W \) and \( \Psi^{(1)}(W) \) and \( \Psi^{(2)}(W) \) denote the number of edges with lexicographically smallest endpoint in \( W \) and the number of cells with lexicographically smallest vertex in \( W \), respectively. Furthermore, \( \tilde{\Psi}^{(1)}(W) \) is the total length of edges observed in \( W \).

### 7.1.3 Distance Functions

In order to compare the estimated vector of characteristics of the given network data with the corresponding vector of computed intensity values for the investigated tessellation models, different distance functions can be considered.

However, reasonable choices for such distance functions are far from being obvious. In the following we display several possibilities. Let \( x = (x_1, \ldots, x_n)^T \) and \( y = (y_1, \ldots, y_n)^T \) be two vectors in \( \mathbb{R}^n \) for \( n \in \mathbb{N} \). Then we consider the

- **Euclidean distance function**
  \[ d_e(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad \text{(absolute)} \]
  \[ d_e(x, y) = \sqrt{\sum_{i=1}^{n} \left( \frac{x_i - y_i}{y_i} \right)^2} \quad \text{(relative)} \]

- **Absolute value distance function**
  \[ d_a(x, y) = \sum_{i=1}^{n} |x_i - y_i| \quad \text{(absolute)} \]
  \[ d_a(x, y) = \sum_{i=1}^{n} \left| \frac{x_i - y_i}{y_i} \right| \quad \text{(relative)} \]

- **Maximum norm distance function**
  \[ d_m(x, y) = \max_{i=1, \ldots, n} |x_i - y_i| \quad \text{(absolute)} \]
  \[ d_m(x, y) = \max_{i=1, \ldots, n} \left| \frac{x_i - y_i}{y_i} \right| \quad \text{(relative)} \]

We mention that the absolute distance functions \( d_e, d_a, \) and \( d_m \) can be influenced strongly by single components of the vectors \( x \) and \( y \) with possibly extreme values, whereas relative distance functions like \( d_{re}, d_{ra}, \) and \( d_{rm} \) are preferable for our purposes since an averaging effect occurs and since they are scale-invariant. The latter property of scale invariance allows for the application of the model fitting procedure to data on different scales as well as to the simultaneous consideration of differently scaled characteristics in the vector \( z \) given in (7.1).

In view of these advantages of relative distance functions over absolute distance functions, and also in view of simulation results obtained in [40] and [57], where both absolute and relative distance functions are used in practical minimization examples, we will concentrate on relative distance functions in the following. However note that the relative distance functions are not symmetric in their argument vectors \( x \) and \( y \) anymore. Thus, this kind of distance functions should be handled with care when used for the fitting procedure. The scaling needs to be done always with respect to the same reference argument.
7.2 Optimal Model Choice for Non–iterated Tessellations

In this section, we describe how an optimal or best fitting non–iterated stationary random tessellation model $\Psi^*$ along with its optimal intensity parameter $\gamma^*$ can be found in order to describe the given network data with respect to a specified distance function. In particular, the potential choices of tessellation models $\Psi$ that are considered for model fitting are any of the three stationary (and isotropic) models PVT, PDT, and PLT, i.e., $\Psi$ can be $\Psi_{\text{PVT}}$, $\Psi_{\text{PDT}}$, or $\Psi_{\text{PLT}}$. Each of these models depends on a single parameter only, the intensity $\gamma_{\text{PVT}}$, $\gamma_{\text{PDT}}$, and $\gamma_{\text{PLT}}$, respectively, where we recall from Section 3.3 that each intensity has to be interpreted differently. More precisely, recall that $\gamma_{\text{PVT}}$ is the mean number of cells per unit area, $\gamma_{\text{PDT}}$ is the mean number of vertices per unit area, and $\gamma_{\text{PLT}}$ coincides with the intensity of the generating Poisson line process and is the mean total length of edges per unit area.

7.2.1 Minimization Problem

We define a function $w : [0, \infty) \to [0, \infty)$ by

$$w(\gamma) = d(z(\gamma), \hat{z}(W)), \quad \gamma > 0,$$

where $z(\gamma) = (z_1(\gamma), \ldots, z_4(\gamma))^\top$ denotes the vector of tessellation characteristics with each entry $z_1(\gamma), \ldots, z_4(\gamma)$ depending on $\gamma$. More precisely, for each tessellation model under consideration, PVT, PDT, and PLT, we are able to express all four characteristics $z_1(\gamma), \ldots, z_4(\gamma)$ in terms of $\gamma = \gamma_{\text{PVT}}$, $\gamma = \gamma_{\text{PDT}}$, and $\gamma = \gamma_{\text{PLT}}$, respectively, by usage of the mean value relationships for facet intensities given in (3.11) for a PVT, in (3.13) for a PDT, and in (3.16) for a PLT. The vector $\hat{z}(W) = (\hat{z}_1(W), \ldots, \hat{z}_4(W))^\top$ contains the four estimated values $\hat{z}_1(W), \ldots, \hat{z}_4(W)$ of the characteristics $z_1, \ldots, z_4$, where the corresponding estimators are given in (7.2). For $d(\cdot, \cdot)$ any of the (relative) distance functions of Section 7.1.3 can be used.

Thus we have to solve the minimization problem

$$\left\{ \begin{array}{l} w(\gamma) \longrightarrow \min \\ \gamma > 0 \end{array} \right. \quad (7.3)$$

for each of the considered models PVT, PDT, and PLT, respectively, resulting in best fitting models within each model class, i.e., in best fitting models $\Psi^*_{\text{PVT}}$, $\Psi^*_{\text{PDT}}$, and $\Psi^*_{\text{PLT}}$, where

$$\gamma^*_{\text{PLT}} = \arg\min_{\gamma_{\text{PVT}}} w(\gamma_{\text{PVT}}), \quad \gamma^*_{\text{PVT}} = \arg\min_{\gamma_{\text{PDT}}} w(\gamma_{\text{PDT}}), \quad \gamma^*_{\text{PDT}} = \arg\min_{\gamma_{\text{PLT}}} w(\gamma_{\text{PLT}}).$$

One possibility to identify the overall best fitting random tessellation model $\Psi^*$ along with the optimal intensity $\gamma^*$ among the different models is to consider

$$w(\gamma^*) = \min\{w(\gamma^*_{\text{PVT}}), w(\gamma^*_{\text{PDT}}), w(\gamma^*_{\text{PLT}})\},$$

and

$$\gamma^* = \arg\min \{w(\gamma^*_{\text{PVT}}), w(\gamma^*_{\text{PDT}}), w(\gamma^*_{\text{PLT}})\}.$$
7.2.2 Minimization Techniques

In order to solve the minimization problem in (7.3), several possibilities exist. Apart from using mathematical software packages like MAPLE™ or Mathematica™, one can also consider methods that can be incorporated in the GeoStoch library. This is especially useful if estimation of characteristics and subsequent model fitting has to be done very often, for example in order to get type-intensity maps ([119]). Such maps visualize for example for each location of a certain region both the best fitting tessellation model making transparent the borderline between different model regions, and the corresponding optimal intensity. Therefore, they provide both qualitative (tessellation model) and quantitative (intensity) information with respect to location.

The simplest minimization method is to discretize the non-negative real line by some grid. Then (7.3) may be rewritten as

\[
\begin{align*}
\{ \ w(\gamma) \rightarrow \min \\ \gamma \in G([0, \infty)) \} ,
\end{align*}
\]

(7.4)

where the notation \( \gamma \in G([0, \infty)) \) is used to express that \( w(\gamma) \) is minimized with respect to discrete values of \( \gamma \) which in turn are obtained by traversing through a grid \( G([0, \infty)) \) with a certain step-width on the non-negative real line. Note that in practice this grid is bounded, i.e., one considers only finitely many discrete values for \( \gamma \) in a range \([\gamma_{\text{min}}, \gamma_{\text{max}}]\) with \( \gamma_{\text{min}}, \gamma_{\text{max}} > 0 \). Apart from this computationally challenging traversing search, other numerical methods can be considered. However, for specific choices of distance functions, (7.3) can also be solved analytically. In [119] both for the Euclidean distance function and for the absolute value distance function an analytical solution of (7.3) is examined. The solution is based on a reconsideration of the explicit form of the function \( w(\gamma) \) for \( \gamma > 0 \) for both distance functions. As an example consider the case of the relative Euclidean distance function and assume that the estimated vector \( \hat{\gamma} \) of characteristics is given by \( \hat{\gamma} = (1.8, 0.03, 0.02, 0.53) \). Figure 7.4 shows a plot of \( w(\gamma) \) as function in \( \gamma = \gamma_{\text{PLT}} \) for the case of a PLT. In view of the formula for the relative Euclidean distance function in Section 7.1.3 and with \( z_1(\gamma), \ldots, z_4(\gamma) \) expressed in terms of \( \gamma \) by usage of the mean value relationships in (3.16) for facet intensities of a stationary (and isotropic) PLT, we get that \( w(\gamma) \) is a polynomial in \( \gamma \) of order four and the first derivative \( w'(\gamma) \) is a cubic polynomial. Hence, this leads to the problem of solving

\[ w'(\gamma) = 0, \]

i.e., to the problem of finding the roots of a cubic equation. Note that also in case of a PVT or a PDT we get that \( w'(\gamma) \) is a cubic polynomial in \( \gamma \), where \( \gamma = \gamma_{\text{PVT}} \) and \( \gamma = \gamma_{\text{PDT}} \), respectively. Application of Cardano’s formula, which provides explicit solutions of cubic equations, cf. pp. 90, pp. 106, and pp. 414 in [17], leads to an optimal value \( \gamma^* \). Checking the second derivative of \( w(\gamma) \) for example approves that \( \gamma^* \) is indeed the (global) minimum of \( w(\gamma) \). In case of a PLT and \( \hat{\gamma} = (1.8, 0.03, 0.02, 0.53)^T \) this can also be seen in Figure 7.4.
7.2.3 Numerical Examples

We consider three different examples, where the input data are taken from single realizations of stationary (and isotropic) non-iterated tessellation models, namely a PVT with intensity $\gamma_{\text{PVT}} = 0.001$, a PDT with intensity $\gamma_{\text{PDT}} = 0.001$, and a PLT with intensity $\gamma_{\text{PLT}} = 0.1$, observed through a rectangular sampling window $W = [0, 500]^2$.

Table 7.1: Estimated characteristics $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_4)^\top$ for different tessellation models

<table>
<thead>
<tr>
<th>Model</th>
<th>$\gamma$</th>
<th>$\hat{z}_1$</th>
<th>$\hat{z}_2$</th>
<th>$\hat{z}_3$</th>
<th>$\hat{z}_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVT</td>
<td>0.001</td>
<td>0.001844</td>
<td>0.002764</td>
<td>0.000920</td>
<td>0.060509</td>
</tr>
<tr>
<td>PDT</td>
<td>0.001</td>
<td>0.000948</td>
<td>0.002860</td>
<td>0.001912</td>
<td>0.105125</td>
</tr>
<tr>
<td>PLT</td>
<td>0.100</td>
<td>0.002896</td>
<td>0.005792</td>
<td>0.002896</td>
<td>0.095825</td>
</tr>
</tbody>
</table>

Table 7.1 presents estimates of the characteristics $z = (z_1, \ldots, z_4)^\top$ for this single realization and Figure 7.5 displays a realization of the respective models. Tables 7.2 to 7.4 show the results of the model fitting procedure applied to find the best fitting model among PVT, PDT, and PLT.

First of all, we note that the fitting results do not depend strongly on the choice of the particular (relative) distance functions. In all three cases, the decision of the fitting procedure is in favor of the model that was used to obtain the input network data through simulation.

Note that the distance value of the best fitting model is, again in all three cases, remarkably lower than the distance of the second best fitting model. However, not only does the procedure recognize the correct model but also the intensity of the input model is reproduced very well.
7.2. Optimal Model Choice for Non-iterated Tessellations

![Figure 7.5: Realizations of the considered tessellation models in Section 7.2.3](image)

<table>
<thead>
<tr>
<th>Best fitting model</th>
<th>Model</th>
<th>Euclidean distance function $d_{re}$</th>
<th>PVT</th>
<th>PVT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>absolute distance function $d_{ra}$</td>
<td>0.003401</td>
<td>0.005438</td>
</tr>
<tr>
<td>Second best fitting model</td>
<td>Model</td>
<td>PLT</td>
<td>PLT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.480518</td>
<td>0.791463</td>
<td></td>
</tr>
<tr>
<td>Third best fitting model</td>
<td>Model</td>
<td>PDT</td>
<td>PDT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.890474</td>
<td>1.454736</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Best fitting model</th>
<th>Model</th>
<th>Euclidean distance function $d_{re}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>absolute distance function $d_{ra}$</td>
</tr>
<tr>
<td>Second best fitting model</td>
<td>Model</td>
<td>PLT</td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.632374</td>
</tr>
<tr>
<td>Third best fitting model</td>
<td>Model</td>
<td>PVT</td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.968856</td>
</tr>
</tbody>
</table>
Table 7.4: Model fitting results; input PLT ($\gamma_{\text{PLT}} = 0.1$)

<table>
<thead>
<tr>
<th>Best fitting model</th>
<th>Model</th>
<th>PLT</th>
<th>PLT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intensity</td>
<td>0.095417</td>
<td>0.095384</td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.004429</td>
<td>0.004608</td>
</tr>
<tr>
<td>Second best fitting model</td>
<td>Model</td>
<td>PVT</td>
<td>PVT</td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.4698281</td>
<td>0.749595</td>
</tr>
<tr>
<td>Third best fitting model</td>
<td>Model</td>
<td>PDT</td>
<td>PDT</td>
</tr>
<tr>
<td></td>
<td>Distance</td>
<td>0.643859</td>
<td>1.0098289</td>
</tr>
</tbody>
</table>

7.3 Optimal Model Choice for Iterated Tessellations

In this section we turn to the problem of finding a best fitting iterated tessellation model. More precisely, we consider the class of $\Psi_0/p\Psi_1$-nestings, where the stationary (and isotropic) and independent tessellations $\Psi_0$ and $\Psi_1$ can be a PVT, a PDT, and a PLT, respectively, and where $p$ is the Bernoulli thinning parameter introduced in Section 3.4. We proceed similarly as for model fitting of non-iterated tessellations described in Section 7.2. In particular note that $\Psi_0$ depends on the intensity $\gamma_0$ and $\Psi_1$ depends on the intensity $\gamma_1$ which are both assumed positive (and finite) in the following and which have to be interpreted according to the specific choice of $\Psi_0$ and $\Psi_1$, respectively. Note that besides the consideration of $\Psi_0/p\Psi_1$-nestings, also superpositions of $\Psi_0$ and $\Psi_1$ can be considered ([34]). The fitting procedure itself remains the same as will be described subsequently for nestings. Merely the mean value relationships given in Tables 3.1 to 3.3 have to be replaced by the respective relationships for superpositions; [69], [104].

7.3.1 Minimization Problem

We introduce the function $w : [0, \infty)^2 \times [0, 1] \rightarrow [0, \infty)$ by

$$w(\gamma_0, \gamma_1, p) = d(z(\gamma_0, \gamma_1, p), \hat{z}), \quad \gamma_0, \gamma_1 > 0, p \in [0, 1],$$

where $z(\gamma_0, \gamma_1, p)$ denotes the vector of (theoretical) intensities depending on the parameter $\gamma_0$ of $\Psi_0$, the parameter $\gamma_1$ of $\Psi_1$, and on the Bernoulli thinning parameter $p$. Note that for our particular choices of $\Psi_0$ and $\Psi_1$, respectively, this vector is determined through the mean value relationships in Tables 3.1 to 3.3. Furthermore, $\hat{z}$ is the vector of estimated characteristics of the given network data, where once again the estimators given in (7.2) are used for estimation. In principle for the (relative) distance function $d(\cdot, \cdot)$ all choices presented in Section 7.1.3 are possible. In the sequel, we restrict ourselves to the case of the relative Euclidean distance function, i.e., we consider

$$w(\gamma_0, \gamma_1, p) = d_{re}(z(\gamma_0, \gamma_1, p), \hat{z}).$$
Thus the minimization problem

\[
\begin{align*}
&\min w(\gamma_0, \gamma_1, p) \\
&\text{subject to } \gamma_0, \gamma_1 > 0, p \in [0, 1]
\end{align*}
\]  

(7.5)

has to be solved for each possible choice of \(\Psi_0\) and \(\Psi_1\). If \(p = 0\), we refer to the minimization techniques discussed in Section 7.2. If \(p = 1\), to obtain a \(\Psi_0/\Psi_1\)-nesting, nine combinations of the basic non-iterated stationary (and isotropic) tessellation models PVT, PDT, and PLT are possible. A choice of \(p \in (0, 1)\) allows for \(\Psi_0/p \Psi_1\)-nestings which can be expected to describe a wider range of network input data. In practice, only some discrete values of \(p \in [0, 1]\) are considered, resulting in infinitely many possible models for the minimization procedure. Beyond, after exploratory data analysis, it is often appropriate to consider values of \(p\) which are close to one. In case of the SSLM, such a choice of \(p\) can be interpreted to model a small perturbation in the iterated two-level model of roads, for example parks or graveyards where no side streets pass through. Hence, it is appropriate to consider values of \(p\) with \(p \in [p_{min}, 1]\) and \(p_{min} > 0.5\) or even \(p_{min} > 0.75\).

### 7.3.2 Minimization Techniques

As already described for the case of non-iterated tessellations in Section 7.2, the minimization problem (7.5) can be solved using MAPLE™ or Mathematica™. However it is again favorable to consider methods which may be integrated in the GeoStoch system; Section 7.2.2.

A first idea is to consider the discretized minimization problem

\[
\begin{align*}
&\min w(\gamma_0, \gamma_1, p) \\
&\text{subject to } (\gamma_0, \gamma_1) \in \mathbb{G}([0, \infty)^2), p \in \mathbb{G}([0, 1])
\end{align*}
\]  

(7.6)

where \(\mathbb{G}([0, \infty)^2\) is a grid on the non-negative plane \([0, \infty)^2\) and \(\mathbb{G}([0, 1])\) is a (finite) grid on \([0, 1]\). One possibility to find an optimal parameter vector \((\gamma^*_0, \gamma^*_1, p^*)\) is to bound \(\mathbb{G}([0, \infty)^2)\) by choosing \(\gamma^*_0, \gamma^*_1 > 0\) as well as \(\gamma^*_0, \gamma^*_1 > 0\) as shown in Figure 7.6. Then, we can evaluate \(\gamma_0\) and \(\gamma_1\) on the bounded grid \(\mathbb{G}([0, \infty)^2)\cap [\gamma^*_0, \gamma^*_0] \times [\gamma^*_1, \gamma^*_1]\) consisting only of finitely many grid points. In order to do so and to additionally evaluate \(p\) on \(\mathbb{G}([0, 1])\) the traversing search can be used. The advantage is that always an optimal solution \((\gamma^*_0, \gamma^*_1, p^*)\) is found.

However, there are certain disadvantages. First, the quality of the found optimal parameter vector \((\gamma^*_0, \gamma^*_1, p^*)\) depends strongly on the discretization. If the step width, i.e., the distance between the grid points where \(\gamma_0\), \(\gamma_1\) and \(p\) are evaluated, respectively, is chosen to be small, the optimal solution \((\gamma^*_0, \gamma^*_1, p^*)\) is quite good. The run time of the traversing search however is unacceptable in most cases. If on the other hand the step width of the grid points is chosen to be large, the quality of the approximated solution is fairly poor. Besides, there is always the problem of deciding which step width is to be considered small or large, depending on the particular data situation. A second disadvantage that adds to the run time of the traversing search considerably is the grid size, i.e., more precisely, the appropriate choice of \(\gamma^*_0, \gamma^*_1\)
and $\gamma_1^{\min}, \gamma_1^{\max}$ as well as also $p_{\min}$ and $p_{\max}$. In both cases, some preliminary exploratory
data analysis can help to choose an appropriate step width and grid size.

Having in mind the disadvantages of discretization and the application of numerical search
methods like the traversing search, one can utilize the fact that (7.5) can be reduced from
a three–dimensional minimization problem to a two–dimensional one. Indeed, reviewing
the function $w(\cdot, \cdot, \cdot)$ involved in (7.5), we see that this function depends explicitly on $\gamma_0, \gamma_1,$ and
$p$ via usage of the relative Euclidean distance function defined in Section 7.1.3, and the mean
value relationships in Tables 3.1 to 3.3. If we solve

$$\frac{\partial}{\partial p} w(\cdot, \cdot, p) = 0,$$

where $\partial/\partial p$ denotes the partial derivative of $w(\cdot, \cdot, p)$ with respect to $p$, we get an explicit
optimal value $p^*$ for $p$. This value $p^* = p^*(\gamma_0, \gamma_1)$ depends only on $\gamma_0$ and $\gamma_1$ and can be
plugged into $w(\cdot, \cdot, p)$ in order to get

$$w(\gamma_0, \gamma_1, p^*(\gamma_0, \gamma_1)),$$

such that the latter function depends only on the parameters $\gamma_0$ and $\gamma_1$, respectively. Thus,
to solve (7.5) with $w(\gamma_0, \gamma_1, p)$ replaced by $w(\gamma_0, \gamma_1, p^*)$, we need to solve a two–dimensional
minimization problem with respect to $\gamma_0$ and $\gamma_1$. In the following, apart from solving the
discretized version (7.6) of (7.5) with the traversing search, we consider the Nelder–Mead
algorithm (NMA), cf. Chapter 17 of [60], as well as its combination with a projected gradient
method (NMAPG), cf. Chapter 10 of [60]. The Nelder–Mead algorithm provides a direct search
method for the minimization of non–linear functions. Its advantage is its relative simplicity
compared to alternative methods and the fact that no derivatives are used. Therefore, it is
one of the most widely used minimization techniques in practice, despite the drawback that
no theoretical basis is given that states convergence of NMA to an optimal solution. As the following examinations will show, we found that for our purposes the NMA performs at least not worse then the traversing search (with relatively small step width), in most cases even better, with the big advantage of remarkably reduced run time. In particular, if we use the NMA results to obtain a starting point for the subsequently applied projected gradient method, the results of the model fitting are convincingly good. A proceeding vice versa is also possible.

7.3.3 Numerical Examples

In this section we present some examples for the fitting of one-fold nestings. Figure 7.7 shows

![Image of a realization of PLT/PVT-nesting](image)

Estimated characteristics:
\[ \hat{\gamma}_1 = 0.027738 \]
\[ \hat{\gamma}_2 = 0.043243 \]
\[ \hat{\gamma}_3 = 0.015505 \]
\[ \hat{\gamma}_4 = 0.229280 \]

Figure 7.7: Realization of PLT/PVT-nesting \((\gamma_0 = 0.1, \gamma_1 = 0.004, p = 1.0)\) and estimated characteristics \(\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_4)^T\)

a realization of a PLT/PVT-nesting with no Bernoulli thinning applied, i.e., \(p = 1\), and intensities \(\gamma_0 = \gamma_{\text{PLT}} = 0.1\) of the PLT and \(\gamma_1 = \gamma_{\text{PVT}} = 0.004\) of the PVT. The sampling window \(W\) is chosen to be of rectangular shape, more precisely \(W = [0, 1000]^2\). With respect to the examples in the non-iterated case, the sampling window is enlarged. This can be reasoned by the fact that for iterated tessellations more information is needed in order for the fitting procedure to work well. Note that \(\gamma_0\) has the same value as the intensity of the PLT-input in the non-iterated example. Table 7.5 summarizes the results of the different minimization techniques for these input data.

Note that there exist however examples where, after the fitting procedure has been applied, the intensity values \(\gamma_0\) and \(\gamma_1\) are found to be switched, i.e., the decision in favor of a certain one-fold tessellation model is correct, but the intensities of initial and component tessellation are not recognized properly. This effect may be accounted to the structure of the mean value relationships in Tables 3.1 to 3.3 whenever \(\Phi_0\) and \(\Phi_1\) are of the same model type and \(p = 1\). Here we get an additional argument in favor of regarding \(p \in [p_{\text{min}}, 1]\) with \(p_{\text{min}} > 0.5\) as already mentioned above. In Table 7.5, we can see that for PLT/PVT input data both the
traversing search and NMA/NMAPG reproduce the input model pretty well. As expected, the run time of the traversing search is much higher compared to NMA and NMAPG. Note that the values in the column entitled ‘run time’ in Table 7.5, as well as in all subsequent tables containing results of the fitting procedure, have to be interpreted in terms of the NMA run time on a Sun Fire V40z server with four AMD Opteron 850 CPUs, each with 2.4 GHz and 16 GB main memory.

A second example is shown in Figure 7.8 where we consider a PDT/pPVT-nesting with \( \gamma_0 = \gamma_{\text{PDT}} = 0.001 \) and \( \gamma_1 = \gamma_{\text{PVT}} = 0.005 \) as well as with Bernoulli thinning parameter \( p = 0.9 \) in a rectangular sampling window \( W = [0,1000]^2 \). Here, compared to the first example, the intensity values are reduced considerably as can be seen in Figure 7.8. Therefore, less information about the PDT/pPVT-nesting is contained in \( W \). The results of the model fitting procedure are summarized in Table 7.6, where the admissible range for the Bernoulli thinning parameter \( p \) is restricted to \([0.7, 1]\). As it can be observed, all three techniques deliver

![Table 7.5: Fitting of a \( \Psi_0/\Psi_1 \)-nesting to the data of Figure 7.7](image)

<table>
<thead>
<tr>
<th>Minimization method</th>
<th>Best fitting model</th>
<th>Rel. Eucl. distance</th>
<th>( \gamma_0 )</th>
<th>( \gamma_1 )</th>
<th>( p )</th>
<th>Run time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traversing</td>
<td>PLT/PVT</td>
<td>0.003640</td>
<td>0.101500</td>
<td>0.004032</td>
<td>1.0</td>
<td>460</td>
</tr>
<tr>
<td>NMA</td>
<td>PLT/PVT</td>
<td>0.003628</td>
<td>0.101475</td>
<td>0.004033</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>NMAPG</td>
<td>PLT/PVT</td>
<td>0.003628</td>
<td>0.101475</td>
<td>0.004033</td>
<td>1.0</td>
<td>( \approx 1 )</td>
</tr>
</tbody>
</table>

Estimated characteristics:

\[
\hat{z}_1 = 0.027007 \\
\hat{z}_2 = 0.041356 \\
\hat{z}_3 = 0.014791 \\
\hat{z}_4 = 0.221451
\]

Figure 7.8: Realization of PDT/pPVT-nesting \( (\gamma_0 = 0.001, \gamma_1 = 0.005, \ p = 0.9) \) and estimated characteristics \( \hat{z} = (\hat{z}_1, \ldots, \hat{z}_4)^\top \) about the same result. One disadvantage is that the Bernoulli thinning parameter \( p \) is not reproduced properly in this example. A possible explanation for the latter is that \( W \) is (still) too small in order to get enough information about \( p \).
We close this section by a final example of a PLT/PLT-nesting with $\gamma_0 = \gamma_{\text{PLT}} = 0.001$ and $\gamma_1 = \gamma_{\text{PLT}} = 0.015$ observed through a rectangular sampling window $W = [0,2000]^2$. Here, once again the sampling window $W$ is enlarged and the intensity values $\gamma_0$ and $\gamma_1$ are relatively low. Figure 7.9 depicts the PLT/PLT-nesting and shows the estimated values $\hat{\varphi} = (\hat{z}_1, \ldots, \hat{z}_4)$. For those input data we use only NMA and NMAPG. Both techniques decide in favor of a PLT/PLT-nesting with intensities $\gamma_0^* = 0.009915$ and $\gamma_1^* = 0.014875$. Hence, the chosen input intensities are reproduced very well despite the fact that the estimated values $\hat{z}_1, \ldots, \hat{z}_4$ are rather small, meaning that not much information about the input data is contained in the sampling window. For reasons of completeness, we note that the minimal distance is 0.014399.

In view of the presented examples as well as further examples discussed in [57] and [119] one can draw the conclusion that fitting of iterated tessellations is much more challenging compared to fitting non-iterated tessellation models. Both the traversing search and NMA as well as NMAPG seem to work well. Note the drastic reduction of run time in case of NMA and NMAPG.
7.4 Model Check by Monte Carlo Tests

The following *Monte Carlo technique* is a general test principle based on simulations and is widely used in different fields of applications. A description of this technique, along with examples in stochastic geometry, can be found for example in [116] and [117]. Here, we use Monte Carlo tests (MC tests) in order to examine the performance of the model fitting method explained in Sections 7.2 and 7.3 in the case where the given network data are taken from realizations of the tessellation models as well as to validate the decision in favor of a best fitting model in the case of real network data.

7.4.1 Description of the Monte Carlo Test

First we establish a (simple) null hypothesis $H_0$ stating that the given network data can be described by the tessellation model $\Psi(H_0)$ depending on the (hypothetical) parameter $\gamma(H_0)$ in the case of non-iterated tessellation models and $\gamma_0(H_0)$ and $\gamma_1(H_0)$ as well as possibly a Bernoulli thinning parameter $p(H_0)$ in the case of one-fold nestings, respectively.

In order to assess the fit of a tessellation model chosen as best fitting model $\Psi^*$ by the minimization procedure described in Sections 7.2 and 7.3, respectively, we choose $H_0 : \Psi(H_0) = \Psi^*, \gamma(H_0) = \gamma^*$ (in case of non-iterated tessellation models) or $H_0 : \Psi(H_0) = \Psi^*, \gamma_0(H_0) = \gamma_0^*, \gamma_1(H_0) = \gamma_1^*, p(H_0) = p^*$ (in case of one-fold nestings). The alternative hypothesis $H_1$ states that the given network data can be described by the other tessellation models under consideration (with intensity parameters obtained through the minimization procedure). For example assume we consider non-iterated tessellations and $H_0$ states that $\Psi(H_0) = \Psi^*$, where $\Psi^*$ is a PLT, say, with intensity $\gamma(H_0) = \gamma^*$. Then, $H_1$ states that the input data can be described by a PVT and a PDT, respectively, where the intensity of these two models is obtained by the minimization procedure. Subsequently, we will only state the null hypothesis $H_0$ for short.

Note that in case of real network data it is useful not only to test for $H_0$ with $\Psi(H_0) = \Psi^*$, but also for $\Psi(H_0)$ to go through all tessellation models under consideration. Hence, if $\Psi^*$ is a PLT for example, we also do tests with the second-best and third-best model, i.e., $\Psi(H_0)$ is chosen to be a PVT and a PDT with some fixed intensity parameters $\gamma(H_0)$.

The significance level $\alpha$ of the just described MC test, i.e., the maximal probability of rejecting $H_0$ despite its correctness, has to be chosen. Popular choices are $\alpha = 0.05$ or $\alpha = 0.01$. Subsequently, the tessellation model $\Psi(H_0)$ with intensity $\gamma(H_0)$ (or $\gamma_0(H_0)$, $\gamma_1(H_0)$, and $p$) is simulated $n$ times. In [117] it is suggested to use $n = 99$ if $\alpha = 0.05$ or $n = 999$ if $\alpha = 0.01$.

Then, we choose a distance function in order to compare the data obtained by the $n$ simulations of $\Psi(H_0)$ and the vector of characteristics obtained via theoretical calculation for $\Psi(H_0)$ using the parameter $\gamma(H_0)$ (or $\gamma_0(H_0)$, $\gamma_1(H_0)$, and $p$, respectively). Eventually, we obtain $n$ distance values $d_1, \ldots, d_n$ as well as one further value $d_{n+1} = d^*$ as distance value between the vector of characteristics for $\Psi(H_0)$ and the (estimated) vector of characteristics of input data. The distance function can again be taken from the selection in Section 7.1.3, its choice however
is independent from the choice of a distance function used for the minimization procedure. Relative distance functions \(d(\cdot, \cdot)\) can be expected to perform better than absolute ones due to the advantages described in Section 7.1.3, however then the order of the arguments of \(d(\cdot, \cdot)\) has to be kept in mind. Complying to the definition in Section 7.1.3, the vector of characteristics \(z(\gamma(H_0))\) calculated for \(\gamma(H_0)\) (or \(z(\gamma_0(H_0), \gamma_1(H_0), p(H_0))\) calculated for \(\gamma_0(H_0), \gamma_1(H_0),\) and \(p(H_0),\) respectively) would be the second argument of the distance function \(d(\cdot, \cdot)\).

Subsequently, the \(n + 1\) distance values \(d_1, \ldots, d_n\) and \(d^*\) are ordered in ascending order, which leads to a sequence \(d_1, \ldots, d_{(n+1)}\), where \(d_{(i)}\) denotes the \(i\)th smallest distance value for \(1 \leq i \leq n + 1\). The null hypothesis \(H_0\) is rejected if the position \(i^*\) of \(d^*\) in this ordered sequence is contained in the rejection region \(R_\alpha = [n - \alpha(n + 1) + 2, \ldots, n + 1]\). Alternatively, we may consider the p-value, given by \(1 - (i^* - 1)/(n + 1)\), which is a number taking values in the interval \([1/(n + 1), 1]\). As always, a rule is that \(H_0\) should be rejected if the obtained p-value is smaller than (or equal to) the chosen significance level \(\alpha\).

### 7.4.2 Numerical Examples

In this section we provide some examples of MC tests. Further examples can be found in [40].

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(n)</th>
<th>(R_\alpha)</th>
<th>(d^*)</th>
<th>(d_{(1)})</th>
<th>(d_{(n+1)})</th>
<th>(i^*)</th>
<th>p-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>([96, 100])</td>
<td>0.162</td>
<td>0.006</td>
<td>1.669</td>
<td>36</td>
<td>0.680</td>
<td>no</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>([991, 1000])</td>
<td>0.162</td>
<td>0.002</td>
<td>2.170</td>
<td>263</td>
<td>0.738</td>
<td>no</td>
</tr>
</tbody>
</table>

We start being regarding non-iterated tessellation models, in particular, the same examples as already studied in Section 7.2.3. Throughout this section we use the (relative) Euclidean distance function. Table 7.7 shows the results of the MC test for PLT input data with intensity \(\gamma_{\text{PLT}} = 0.1\), where the results of the fitting procedure are summarized in Table 7.4. As intensity value \(\gamma(H_0)\) we chose the theoretical value 0.1. In contrast, Table 7.8 shows the results of the same MC test for PLT input if \(\gamma(H_0) = 0.095417\) as obtained by the model fitting procedure; cf. Table 7.4.

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(n)</th>
<th>(R_\alpha)</th>
<th>(d^*)</th>
<th>(d_{(1)})</th>
<th>(d_{(n+1)})</th>
<th>(i^*)</th>
<th>p-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>([96, 100])</td>
<td>0.004</td>
<td>0.004</td>
<td>1.146</td>
<td>1</td>
<td>1.000</td>
<td>no</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>([991, 1000])</td>
<td>0.004</td>
<td>0.002</td>
<td>1.656</td>
<td>3</td>
<td>0.998</td>
<td>no</td>
</tr>
</tbody>
</table>
In Table 7.8 the position $i^*$ of the optimal distance value $d^*$ is remarkably reduced compared to Table 7.7. Moreover, the value $d^*$ itself is much lower for the test results in Table 7.8. This corresponds quite well to the discussion in [117] that for optimized parameters the MC test is more friendly towards the null hypothesis.

As a second example of non-iterated tessellations consider a PVT-input with intensity $\gamma_{\text{PVT}} = 0.001$. Such input data were considered in Section 7.2.3, the results of the fitting procedure are summarized in Table 7.2. Tables 7.9 and 7.10 show the results of the MC tests and have to be interpreted analogously to Tables 7.7 and 7.8, respectively.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>p-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>[96, 100]</td>
<td>0.143</td>
<td>0.004</td>
<td>0.299</td>
<td>78</td>
<td>0.230</td>
<td>no</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>[991, 1000]</td>
<td>0.143</td>
<td>0.003</td>
<td>0.347</td>
<td>804</td>
<td>0.197</td>
<td>no</td>
</tr>
</tbody>
</table>

Table 7.10: MC test for PVT input ($\gamma_{\text{PVT}} = 0.001$), $\Psi(H_0) = \text{PVT}$ with $\gamma(H_0) = 0.000921$

As an example of a one-fold nesting we consider the PLT/PLT-nesting, already considered as last example of Section 7.3.3. Table 7.11 shows the results of the MC test for the significance level 0.05. The null hypothesis states that $\Psi(H_0)$ is a PLT/PLT-nesting with $\gamma_0(H_0) = 0.001$ and $\gamma_1(H_0) = 0.015$. As we can see, $H_0$ is not rejected, however $d^*$ and $i^*$ take larger values compared to the non-iterated examples in Tables 7.7 to 7.10.

<table>
<thead>
<tr>
<th>Intensities</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>p-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>optimized</td>
<td>[96, 100]</td>
<td>0.445</td>
<td>0.013</td>
<td>1.183</td>
<td>64</td>
<td>0.37</td>
<td>no</td>
</tr>
<tr>
<td>theoretical</td>
<td>[96, 100]</td>
<td>0.423</td>
<td>0.021</td>
<td>1.631</td>
<td>60</td>
<td>0.41</td>
<td>no</td>
</tr>
</tbody>
</table>

7.4.3 Performance Analysis

In case of simulated input data the power of MC tests can be estimated and together with this the probability of the error to accept the null hypothesis $H_0$ despite it is not true. The
procedure is in principle analogous to the proceeding described in Section 7.4.1, except that we replace \( \Psi = \Psi(H_0) \) by one of the tessellation models stated in the alternative hypothesis \( H_1 \), where an intensity parameter of \( \Psi \) can be obtained via the minimization procedure. This can be interpreted in the sense that we examine the performance under \( H_1 \). Finally, we repeat the procedure \( i \) times, \( i \geq 1 \), and for each \( \ell = 1, \ldots, i \) we are able to report whether the position \( i^*_\ell \) of the optimal distance value \( d^*_\ell \) within the ordered sequence \( d_{\ell,1}, \ldots, d_{\ell,(n+1)} \) of distance values is in \( R_\alpha \) or not. Here, the distance value \( d^*_\ell \) is calculated between the vector of characteristics of the input data, where we know that these data are not taken from the model \( \Psi \), and the vector of characteristics obtained with the help of the theoretical mean value relationships for the tessellation model \( \Psi \).

An estimate of the power of the MC test is obtained by regarding the estimator

\[
\frac{1}{i} \# \{ i^*_\ell \in R_\alpha, \ell = 1, \ldots, i \}.
\]

(7.7)

The power of the Monte Carlo test can be considered to be high if the estimate obtained through (7.7) takes values which are close to one.

If we regard the examples of non-iterated tessellations in Section 7.2.3 once more, the estimated power is often high, for example in the case of PVT input and \( \Psi(H_0) \) chosen as a PDT with intensity \( \gamma(H_0) = 0.000552 \) as given in Table 7.2 reaching even 1 for \( i = 1000 \) in (7.7). This complies also to results of other examples considered in [40].

Still, there are certain subtleties which render more difficult the performance analysis. First there exist also examples in the non-iterated case where the power is remarkably low. In the case of iterated tessellation models, reconsidering the examples in Section 7.3.3, the estimated power is higher for those alternative tessellation models that show a distance value much higher than the minimal distance value of the identified best fitting model. For example in the case of the PLT/PLT-nesting example at the end of Section 7.3.3, the PDT/PVT-model has an estimated power of 0.99 (for \( i = 1000 \) in (7.7)), whereas models like PLT/PDT and PDT/PLT, whose geometrical structure resembles the one of a PLT/PLT-nesting, show very low estimated power values. However, regarding Figure 7.9, we observe that in this PLT/PLT-nesting example the sampling window carries much less information compared to the other two fitting examples in Section 7.3.3. This is due to the fact that despite the rather large sampling window, the four estimated values \( \hat{\gamma}_1, \ldots, \hat{\gamma}_4 \) are much smaller than in the other examples of Section 7.3.3, where additionally the sampling window is smaller, too. Therefore, \( \hat{\gamma}_1, \ldots, \hat{\gamma}_4 \) are based on relatively few measurement values.

The main reason for results of the kind described above is the fact that when considering the alternative hypothesis, the only reasonable intensity values that can be used to formulate the hypothesis are the intensity values that are obtained through the preliminary fitting procedure. Thus, any test is once again more friendly to the null hypothesis than theoretically allowed. Nevertheless, a performance analysis like this is useful, leading to the recommendation that for real input data also those models cannot be excluded as potentially appropriate models that do not show the minimal distance value; Section 7.5 below. Therefore, subsequently to the fitting procedure, an MC test should be conducted for any model. Another recommendation
is to have enough information in the sampling window or, if this is not the case, the sampling window should be enlarged if practically feasible.

7.5 Application to Telecommunication Data

In this section, we consider real road data. More precisely, we consider some regions of the infrastructure data of Paris (Figure 1.1), where these data are given in geodesic coordinates and have to be transformed to Cartesian coordinates. After preprocessing, the data are given in the form of connected line segments, where each line segment represents a part of a road. Additionally, there is a mark attached to each line segment, indicating the type of road it belongs to. In [107] seven types of roads were identified, ranking from highways down to side streets. For the fitting procedure, only main roads and side streets are taken into account, other types of roads like dead ends are discarded. In [107] and subsequently in [57], algorithms are shown for the removal of dead ends and for the reconstruction of a network forming (non-convex) cells.

In the following, we present two examples of real road data. Further examples of real road data are examined and presented in [41].

Figure 7.10: Road system of Paris in a square of side length 3000 (meters)

7.5.1 Hierarchical Fitting of Iterated Tessellations

In the first example we observe real road data displayed in Figure 7.10 through a sampling window $W = [0,3000]^2$, chosen to capture information about the inner city of Paris.
the real data display a hierarchy of main roads and side streets, one-fold tessellation models seem to give realistic models that reflect the morphological structure of the data. Regarding the main roads contained in Figure 7.10 (b) separately and by applying the fitting procedure for non-iterated tessellation models to these data, a PLT with optimal intensity $\gamma^* = \gamma_{\text{PLT}} = 0.0024$ is obtained as best fitting model, where the relative Euclidean distance is used. Comparing however the minimal distance 0.211 with the minimal distance 0.297 of the second best model, a PVT with intensity $\gamma_{\text{PVT}} = 0.000001$, also this model may not be excluded as appropriate model, whereas the minimal distance 0.734 indicates that a PDT with intensity $\gamma_{\text{PDT}} = 0.000001$ is not suitable. Thus, as described in Section 7.4.3, we employ an MC test for all considered tessellation models to check whether a certain model can be taken as a (sufficiently good) representation of the real data. Tables 7.12 to 7.14 show the results of these tests in case of different null hypotheses $H_0$.

Table 7.12: MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PLT with $\gamma(H_0) = 0.0024$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>$p$-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>[96, 100]</td>
<td>0.246</td>
<td>0.038</td>
<td>5.396</td>
<td>14</td>
<td>0.860</td>
<td>no</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>[991, 1000]</td>
<td>0.246</td>
<td>0.031</td>
<td>12.038</td>
<td>126</td>
<td>0.875</td>
<td>no</td>
</tr>
</tbody>
</table>

As expected, Table 7.12 shows that the identified best fitting model PLT cannot be rejected.

Table 7.13 underlines that also a PVT may be suitable, since here the null hypothesis is also not rejected. Note however that $i^*$ is considerably greater than in the respective cases in Table 7.12.

Table 7.13: MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PVT with $\gamma(H_0) = 10^{-6}$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>$p$-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>[96, 100]</td>
<td>0.372</td>
<td>0.025</td>
<td>1.018</td>
<td>64</td>
<td>0.370</td>
<td>no</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>[991, 1000]</td>
<td>0.372</td>
<td>0.025</td>
<td>1.682</td>
<td>583</td>
<td>0.418</td>
<td>no</td>
</tr>
</tbody>
</table>

Table 7.14 finally shows that a PDT indeed is not suitable, since the results presented there allow for the conclusion that $H_0$ will be rejected for significance levels greater than 0.02.

Table 7.14: MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PDT with $\gamma(H_0) = 10^{-6}$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>$p$-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>[96, 100]</td>
<td>1.440</td>
<td>0.047</td>
<td>1.667</td>
<td>99</td>
<td>0.020</td>
<td>yes</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>[991, 1000]</td>
<td>1.440</td>
<td>0.047</td>
<td>2.037</td>
<td>982</td>
<td>0.019</td>
<td>no</td>
</tr>
</tbody>
</table>
To get a best fitting one-fold nesting model we fix the optimal initial tessellation. Since the distance is minimal for a PLT with parameter $\gamma_0 = \gamma_{PLT} = 0.0024$ we concentrate on this model in the following to describe $\Psi_0$. Hence, we finally only need to distinguish between PLT/PLT, PLT/PVT, and PLT/PDT, where in this example we assume $p = 1.0$ fixed. Table 7.15 shows the results if we apply our minimization procedure with fixed initial tessellation model PLT of intensity $\gamma_0 = 0.0024$ to these three models. The results, depicted also in Figure 7.11, indicate that a PLT/PLT-nesting would be the best fitting model, where for the nested PLT we have the intensity parameter $\gamma_1 = \gamma_{PLT} = 0.013906$.

Table 7.15: Distance values $d_{\text{rec, min}}$ for data of Figure 7.11; fixed initial PLT; nested $\Psi_1$ (intensity $\gamma_1$)

<table>
<thead>
<tr>
<th>$\Psi_0/\Psi_1$</th>
<th>$d_{\text{rec, min}}$</th>
<th>$\gamma_0$</th>
<th>$\gamma_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLT/PLT</td>
<td>0.152</td>
<td>0.0024</td>
<td>0.013906</td>
</tr>
<tr>
<td>PLT/PVT</td>
<td>0.205</td>
<td>0.0024</td>
<td>0.000044</td>
</tr>
<tr>
<td>PLT/PDT</td>
<td>0.366</td>
<td>0.0024</td>
<td>0.000028</td>
</tr>
</tbody>
</table>

Again we consider MC tests, for all tessellation models in Table 7.15, i.e., for $H_0$ we choose $\Psi(H_0)$ to be a PLT/PLT, PLT/PVT, or PLT/PDT with intensity parameters $\gamma_0(H_0) = \gamma_0$ and $\gamma_1(H_0) = \gamma_1$, respectively, as given in Table 7.15. We restrict ourselves to the presentation of two examples. First we consider the case where $H_0$ states that $\Psi(H_0)$ is a PLT/PDT with intensity parameters $\gamma_0(H_0) = 0.0024$ and $\gamma_1(H_0) = 0.000028$. Table 7.16 shows the results for the relative Euclidean distance function.
Table 7.16: MC test for data of Figure 7.10; $Ψ(H_0)$ is a PLT/PDT with $\gamma_0(H_0) = 0.0024$ and $\gamma_1(H_0) = 0.000028$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>p-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>[96, 100]</td>
<td>0.00313</td>
<td>0.00003</td>
<td>0.00313</td>
<td>100</td>
<td>0.00</td>
<td>yes</td>
</tr>
<tr>
<td>0.01</td>
<td>[991, 1000]</td>
<td>0.00313</td>
<td>1.15251</td>
<td>0.00355</td>
<td>997</td>
<td>0.004</td>
<td>yes</td>
</tr>
</tbody>
</table>

Finally, Table 7.17 shows the results of the MC test for the null hypothesis $H_0$ with $Ψ(H_0) = Ψ^*$ and $Ψ^*$ being a PLT/PLT as obtained according to the minimization procedure with optimal intensity parameters $\gamma_0^* = 0.0024$ and $\gamma_1^* = 0.013906$. In this case, using the (relative) Euclidean distance function, we cannot reject this null hypothesis.

Table 7.17: MC test for main roads in Figure 7.10; $Ψ(H_0)$ is a PLT/PLT with $\gamma_0(H_0) = 0.0024$ and $\gamma_1(H_0) = 0.013906$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>$R_\alpha$</th>
<th>$d^*$</th>
<th>$d_{(1)}$</th>
<th>$d_{(n+1)}$</th>
<th>$i^*$</th>
<th>p-value</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99</td>
<td>[96, 100]</td>
<td>0.15327</td>
<td>0.00968</td>
<td>1.22830</td>
<td>30</td>
<td>0.71</td>
<td>no</td>
</tr>
<tr>
<td>0.01</td>
<td>999</td>
<td>[991, 1000]</td>
<td>0.15327</td>
<td>0.00966</td>
<td>1.04893</td>
<td>306</td>
<td>0.695</td>
<td>no</td>
</tr>
</tbody>
</table>

7.5.2 Immediate Fitting of Iterated Tessellations

As a second example, we consider another region of Paris; Figure 7.12 (a). Apart from the fact that we could proceed completely analogously as in the first example considered in Section 7.5.1, we use another fitting approach this time. This alternative immediate approach consists in fitting immediately a one-fold nesting to the data in Figure 7.12 (a). In this example, we additionally consider a Bernoulli thinning with parameter $p \in [0.7, 1.0]$. The restriction of the range of $p$ itself is accounted to the fact that most of the cells in Figure 7.12 (a) contain smaller side streets. We get the estimated vector of characteristics by $\hat{z} = (0.000080, 0.000131, 0.000052, 0.014339)^T$ and NMA and NMAPG suggest a PVT/pPDT model with intensity $\gamma_0 = 0.0000062$, $\gamma_1 = 0.0000086$, and $p = 0.94$, where a realization of this best fitting model is shown in Figure 7.12 (b). If we compare however Figure 7.12 (a) with Figure 7.12 (b), we can see that, despite the PVT/pPDT–nesting is the best fitting model among all considered tessellation models, it cannot capture the slight spatial inhomogeneities observable in Figure 7.12 (a). Thus, this is a reason to further investigate the fitting procedure presented in this chapter. On the one hand, it might be worth to investigate alternative characteristics of tessellation models beyond the considered four characteristics $z_1, \ldots, z_4$ as introduced in Section 7.1.1. Furthermore, also other and in particular more general tessellation models can be considered. There is however a drawback that the more general a tessellation
Figure 7.12: Network data in $W = [1500,4500]^2$ and best fitting tessellation

model is, the less characteristics can be described by explicit formulae. This may result in the necessity of extensive simulation studies.
Appendix A

Mathematical Addendum

This appendix aims at providing a short overview of some mathematical concepts and notions which are important in this thesis but which are not presented in the main chapters in order to not disturb the flow of reading.

For a more detailed discussion of the presented topics, the reader is referred to the literature, in particular to [11], [16], [26], [32], [62], and [113]. Further literature is pointed out in the respective sections.

A.1 The d–dimensional Euclidean Space

Consider the set \( \mathbb{R} \) of real numbers on the line and define \( \mathbb{Z} = \{ \ldots, -1, 0, 1, 2, \ldots \} \) to be the set of integers. Moreover, \( \mathbb{N} = \{ 1, 2, 3, \ldots \} \) and \( \mathbb{N}_0 = \mathbb{N} \cup \{ 0 \} \) denote the set of positive integers and the set of non-negative integers, respectively.

The underlying space considered in this thesis is the d–dimensional Euclidean space \( \mathbb{R}^d \), obtained as d-fold Cartesian product \( \mathbb{R}^d = \mathbb{R} \times \ldots \times \mathbb{R} \). Hence, in set notation, \( \mathbb{R}^d = \{(x_1, \ldots, x_d)\mathsf{T} : x_1, \ldots, x_d \in \mathbb{R}\} \), where a point \( x \in \mathbb{R}^d \) consists of \( d \) real–valued components, i.e., \( x = (x_1, \ldots, x_d)\mathsf{T} \) is a \( d \)-dimensional vector with \( x_i \in \mathbb{R} \) for all \( i = 1, \ldots, d \). The origin \( o \) is given by \( o = (0, \ldots, 0)^\mathsf{T} \). Note that \( \mathbb{R}^2 \) and \( \mathbb{R}^3 \) are referred to as plane and space, respectively.

A.1.1 Topology

The Euclidean space \( \mathbb{R}^d \) is equipped with the standard scalar product \( \langle \cdot, \cdot \rangle : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \),

\[
(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{d} x_i y_i ,
\]
where \( x = (x_1, \ldots, x_d)\T \) and \( y = (y_1, \ldots, y_d)\T \), respectively. Note that \( x \in \mathbb{R}^d \) is called a unit vector if \( \langle x, x \rangle = 1 \) and that \( x, y \in \mathbb{R}^d \) are called orthogonal if \( \langle x, y \rangle = 0 \). The induced Euclidean norm \( \| \cdot \| : \mathbb{R}^d \to \mathbb{R} \) is given by
\[
\| x \| = \sqrt{\langle x, x \rangle}, \quad x \in \mathbb{R}^d.
\]
This allows for the definition of the \( d \)-dimensional (closed) ball \( B^d_r(x) = \{ y \in \mathbb{R}^d : \| x - y \| \leq r \} \) centered at \( x \in \mathbb{R}^d \) with radius \( r > 0 \). For brevity, let \( B^d_r(o) = B^d_r \).

For an arbitrary set \( B \subset \mathbb{R}^d \), the interior, the closure, the boundary, and the complement of \( B \) are denoted by \( \text{int} B, \text{cl} B, \partial B, \text{and} B^c \), respectively. The sphere with center \( x \in \mathbb{R}^d \) and radius \( r > 0 \) is given by \( \partial B^d_r(x) = \{ y \in \mathbb{R}^d : \| x - y \| = r \} \). In particular, \( S^{d-1} = \partial B^d_1(o) \) is called unit sphere and \( S^d_{+} = \{ (x_1, \ldots, x_d)^T \in S^{d-1} : x_d \geq 0 \} \) denotes the upper unit hemisphere.

### A.1.2 Operations on Subsets

The vector space structure of \( \mathbb{R}^d \) allows for addition of \( x, y \in \mathbb{R}^d \) as well as for scalar multiplication of \( x \in \mathbb{R}^d \) by \( \rho \in \mathbb{R} \), i.e.,
\[
x + y = (x_1 + y_1, \ldots, x_d + y_d)^T \quad \text{and} \quad \rho x = (\rho x_1, \ldots, \rho x_d)^T.
\]
For subsets of \( \mathbb{R}^d \), a number of set operations may be defined. In particular, for \( B, B' \subset \mathbb{R}^d \) and \( x \in \mathbb{R}^d \) consider

- the translation \( B + x = \{ y + x : y \in B \} \);
- the multiplication \( \rho B = \{ \rho x : x \in B \} \) for \( \rho \in \mathbb{R} \) and the reflection (about the origin) \( \hat{B} = -B \), where \( B \) is called symmetric if \( B = \hat{B} \);
- the rotation \( \theta_R B = \{ Rx : x \in B \} \) for any rotation \( Rx \) of \( x \in \mathbb{R}^d \) around the origin, where \( R \) denotes an orthogonal \( d \times d \) matrix with determinant \( \det R = 1 \);
- the Minkowski addition \( B \oplus B' = \{ x + y : x \in B, y \in B' \} \) and the Minkowski subtraction \( B \ominus B' = (B^c \oplus B')^c \).

### A.1.3 Measures

Together with the Borel \( \sigma \)-algebra \( \mathcal{B}(\mathbb{R}^d) \) of \((d\text{-dimensional})\) Borel sets, i.e., the smallest \( \sigma \)-algebra on \( \mathbb{R}^d \) that contains all open subsets of \( \mathbb{R}^d \), the \((d\text{-dimensional})\) Euclidean space forms the measurable space \( (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \). The family of bounded Borel sets on \( \mathbb{R}^d \) is denoted by \( \mathcal{B}_0(\mathbb{R}^d) \). Let \( \mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty] \) be a measure on \((\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \). We call \( x \in \mathbb{R}^d \) an atom of \( \mu \) if \( \mu(\{x\}) > 0 \), whereas \( \mu \) is called diffuse if \( \mu(\{x\}) = 0 \) for all \( x \in \mathbb{R}^d \). Moreover, \( \mu \) is locally finite if \( \mu(B) < \infty \) for all \( B \in \mathcal{B}_0(\mathbb{R}^d) \). Finally, \( \mu \) is translation invariant and rotation
invariant if \( \mu(B + x) = \mu(B) \) and \( \mu(\theta_R B) = \mu(B) \), respectively, for any Borel set \( B \in \mathcal{B}(\mathbb{R}^d) \) as well as any orientation vector \( x \in \mathbb{R}^d \) and any rotation \( \theta_R \) around \( o \).

**Counting Measures**

On \((\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\) consider the Dirac measure \( \delta_x : \mathcal{B}(\mathbb{R}^d) \to \{0, 1\} \) for \( x \in \mathbb{R}^d \), where

\[
\delta_x(B) = \begin{cases} 
0 & \text{if } x \notin B \\
1 & \text{if } x \in B
\end{cases}, \quad B \in \mathcal{B}(\mathbb{R}^d)
\]  

(A.1)

which is a probability measure. A counting measure \( \mu_c : \mathcal{B}(\mathbb{R}^d) \to \mathbb{N}_0 \cup \{\infty\} \) is a non-negative, integer-valued measure on \( \mathcal{B}(\mathbb{R}^d) \) that can be represented by finite or countable sums of Dirac measures, i.e.,

\[
\mu_c(B) = \sum_{x \in \text{supp} \mu_c} a_x \delta_x(B), \quad B \in \mathcal{B}(\mathbb{R}^d),
\]  

(A.2)

where \( a_x \in \mathbb{N} \) are some weights and \( \text{supp} \mu_c \) denotes the support of \( \mu_c \), given by \( \text{supp} \mu_c = \{x \in \mathbb{R}^d : \mu_c(x) > 0\} \). Note that we call \( \mu_c \) simple if \( \mu_c(x) = 1 \) for all \( x \in \text{supp} \mu_c \).

**Lebesgue Measure**

The Lebesgue measure \( \nu_d : \mathcal{B}(\mathbb{R}^d) \to [0, \infty] \) is uniquely defined by

\[
\nu_d([a_1, b_1] \times \ldots \times [a_d, b_d]) = \prod_{i=1}^{d} (b_i - a_i), \quad d \geq 1,
\]

where \([a_i, b_i]\) are half open intervals with \( a_i \leq b_i \) for \( i = 1, \ldots, d \). Note that \( \nu_d \) is diffuse and locally finite as well as both translation and rotation invariant. Moreover, it holds that \( \nu_d([0,1]^d) = 1 \).

**Lemma A.1** (Haar’s lemma) Let \( \mu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty] \) be a translation invariant measure. Then there exists a constant \( \rho \in [0, \infty] \) such that \( \mu(B) = \rho \nu_d(B) \) for all \( B \in \mathcal{B}(\mathbb{R}^d) \). If \( \mu \) is additionally locally finite then \( \rho < \infty \).

**Hausdorff Measure**

For brevity put

\[
\kappa_k = \nu_k(B_1^k) = \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2} + 1\right)}, \quad k = 1, \ldots, d,
\]

(A.3)

where \( \Gamma(s) = \int_{0}^{\infty} e^{-y} y^{s-1} \, dy \) for \( s > 0 \) denotes the Gamma function. Note that \( \kappa_k \) denotes the volume of the unit ball in \( \mathbb{R}^k \) for \( k = 1, \ldots, d \) and that we let \( \kappa_0 = 1 \).

The Hausdorff measure of order \( k \) in \( \mathbb{R}^d \) can be introduced as a measure \( \mathcal{H}_k : \mathcal{B}(\mathbb{R}^d) \to [0, \infty] \), which is, for \( k = 1, \ldots, d \), defined by

\[
\mathcal{H}_k(B) = \sup_{\delta > 0} \inf \left\{ \sum_{n \geq 1} \kappa_k \left( \frac{D(B_n)}{2} \right)^k : B \subset \bigcup_{n \geq 1} B_n, D(B_n) \leq \delta \right\}, \quad B \in \mathcal{B}(\mathbb{R}^d),
\]
cf. p. 77 in [32], p. 247 in [16], and [100], where \( \kappa_k = \nu_k(B^k) \) is given by (A.3). The infimum is taken over all countable coverings of \( B \) by sets \( B_n \in \mathcal{B}(\mathbb{R}^d) \), \( n \geq 1 \), which have a diameter \( D(B_n) = \sup \{ \| x - y \| : x, y \in B_n \} \) less than or equal to \( \delta > 0 \). Note that the Hausdorff measure \( \mathcal{H}_d \) of order \( d \) coincides with the Lebesgue measure \( \nu_d \). Generally, the Hausdorff measure of order \( k \) can be interpreted to measure the trace of \( k \)-dimensional objects in \( d \)-dimensional space.

If we admit \( k = 0 \) in the definition of the Hausdorff measure, \( \mathcal{H}_0 \) is the counting measure that counts the elements of a set \( B \in \mathcal{B}(\mathbb{R}^d) \). We denote this measure by \( \nu_0 \), i.e., \( \nu_0(B) = \#B \) for \( B \in \mathcal{B}(\mathbb{R}^d) \), where \( \nu_0(B) \) is finite if \( B \) contains finitely many elements and infinite otherwise.

**Spherical Lebesgue measure**

Finally, define the *spherical Lebesgue measure* \( \tilde{\nu}_d : \mathcal{B}(S^{d-1}) \to [0, \infty) \) by

\[
\tilde{\nu}_d(A) = d \nu_d(\{ \rho x \in \mathbb{R}^d : x \in A, \rho \in [0, 1] \}), \quad A \in \mathcal{B}(S^{d-1}),
\]

where

\[
\tilde{\nu}_d(S^{d-1}) = \omega_d = d \kappa_d = \frac{2 \pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)},
\]

denotes the *surface area of the unit ball* \( B_1^d \).

### A.2 Integration Theory

#### A.2.1 Measurable and Integrable Functions

Let \( (\Omega, \sigma(\Omega)) \) and \( (\Omega', \sigma(\Omega')) \) be two measurable spaces, where \( \sigma(\Omega) \) and \( \sigma(\Omega') \) denote a \( \sigma \)-algebra on the (non-empty) sets \( \Omega \) and \( \Omega' \), respectively. Then we call the mapping \( f : \Omega \to \Omega' \) a \( (\sigma(\Omega), \sigma(\Omega')) \)-measurable mapping if

\[
f^{-1}(A') = \{ \omega \in \Omega : f(\omega) \in A' \} \in \sigma(\Omega), \quad \text{for all } A' \in \sigma(\Omega').
\]

Furthermore, \( f : \Omega \to \mathbb{R} \) is called \( \sigma(\Omega) \)-measurable function, or *measurable function*, if \( f \) is a \( (\sigma(\Omega), \mathcal{B}(\mathbb{R})) \)-measurable mapping. Note that if additionally \( (\Omega, \sigma(\Omega)) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \), we call \( f \) a *Borel measurable function*.

Let \( (\Omega, \sigma(\Omega), \mu) \) be a measure space, i.e., a measurable space \( (\Omega, \sigma(\Omega)) \) equipped with a measure \( \mu : \sigma(\Omega) \to [0, \infty] \) and let \( f : \Omega \to \mathbb{R} \) be a measurable function. Then \( f \) is called \( \mu \)-*integrable* if \( \int f^+ d\mu \) and \( \int f^- d\mu \) are real numbers, where \( f^+(\omega) = \max\{0, f(\omega)\} \) and \( f^- (\omega) = \max\{0, -f(\omega)\} \) for \( \omega \in \Omega \), and

\[
\int f d\mu = \int f^+ d\mu - \int f^- d\mu
\]

is called \( \mu \)-*integral* of \( f \) on \( \Omega \); cf. [11] for example for further details.

Note that \( L^1 = L^1(\Omega, \sigma(\Omega), \mu) \) and \( L^2 = L^2(\Omega, \sigma(\Omega), \mu) \) denote the families of \( \mu \)-integrable and quadratic \( \mu \)-integrable functions on \( (\Omega, \sigma(\Omega), \mu) \), respectively.
A.3. Concepts of Probability Calculus

A.2.2 Integration Theorems

Consider a measure space \((\Omega, \sigma(\Omega), \mu)\). Some set \(A \in \sigma(\Omega)\) with \(\mu(A) = 0\) is called a null set and \((\Omega, \sigma(\Omega), \mu)\) is called a complete measure space if for a subset \(B \subset A\) of some null set \(A \in \sigma(\Omega)\) it holds that \(B \in \sigma(\Omega)\). Furthermore, \((\Omega_1 \times \Omega_2, \sigma(\Omega_1) \otimes \sigma(\Omega_2), \mu_1 \times \mu_2)\) is called product space of \((\Omega_1, \sigma(\Omega_1), \mu_1)\) and \((\Omega_2, \sigma(\Omega_2), \mu_2)\) equipped with the product measure \(\mu_1 \times \mu_2\).

Now we are in the position to recall some essential theorems from integration theory.

**Theorem A.1** (Lebesgue’s theorem/Dominated convergence theorem) Let \(\{f_n\}_{n \geq 1}\) be a sequence of measurable functions \(f_n : \Omega \rightarrow \mathbb{R}\) and let \(f : \Omega \rightarrow \mathbb{R}\) be integrable functions where \(f_n \rightarrow f\) almost everywhere on \(\Omega\) and \(|f_n| \leq g\) for all \(n \in \mathbb{N}\). Then,

\[
\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu
\]

**Theorem A.2** (Fubini’s theorem) Let \((\Omega_1, \sigma(\Omega_1), \mu_1)\) and \((\Omega_2, \sigma(\Omega_2), \mu_2)\) be two complete measure spaces and let \((\Omega_1 \times \Omega_2, \sigma(\Omega_1) \otimes \sigma(\Omega_2), \mu_1 \times \mu_2)\) be the product space. Furthermore, let \(f : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}\) be a \(\mu_1 \times \mu_2\)-integrable function. Then,

(i) The function \(f(\cdot, \omega_2)\) is \(\mu_1\)-integrable for almost all \(\omega_2 \in \Omega_2\) and the function \(f(\omega_1, \cdot)\) is \(\mu_2\)-integrable for almost all \(\omega_1 \in \Omega_1\).

(ii) The function \(\int_{\Omega_2} f(\cdot, \omega_2) \, d\mu_2(\omega_2)\) is \(\mu_1\)-integrable and the function \(\int_{\Omega_1} f(\omega_1, \cdot) \, d\mu_1(\omega_1)\) is \(\mu_2\)-integrable.

(iii) It holds that

\[
\int_{\Omega_1 \times \Omega_2} f \, (\mu_1 \times \mu_2) = \int_{\Omega_2} \left( \int_{\Omega_1} f \, d\mu_1 \right) \, d\mu_2 = \int_{\Omega_1} \left( \int_{\Omega_2} f \, d\mu_2 \right) \, d\mu_1.
\]

A.3 Concepts of Probability Calculus

A.3.1 Random Variables and Some Properties

Let \(X : \Omega \rightarrow \mathbb{R}\) be a real-valued random variable defined on some probability space \((\Omega, \sigma(\Omega), \mathbb{P})\), where \(\sigma(\Omega)\) denotes a \(\sigma\)-algebra on \(\Omega\) and \(\mathbb{P} : \sigma(\Omega) \rightarrow [0, 1]\) is a probability measure. Hence, \(X\) is a mapping from the non-empty set \(\Omega\) into the real line, fulfilling

\[
\{ \omega \in \Omega : X(\omega) \in B \} \in \sigma(\Omega),
\]

for each \(B \in \mathcal{B}(\mathbb{R}^d)\), where \(\mathcal{B}(\mathbb{R})\) denotes the family of Borel sets on the real line. Likewise, the mapping \((X_1, \ldots, X_d) : \Omega \rightarrow \mathbb{R}^d\) is called \(d\)-dimensional random vector in \(\mathbb{R}^d\) if \(X_1, \ldots, X_d : \Omega \rightarrow \mathbb{R}\) are real-valued random variables.
Let $X_1, \ldots, X_n$ be real-valued random variables for $n \in \mathbb{N}$. They are called independent if
\[
P(X_1 \in B_1, \ldots, X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i), \quad B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}).
\]

**Distribution and Distribution Function**

Let $X : \Omega \to \mathbb{R}^d$ be a $d$-dimensional random vector. We call $P_X : \mathcal{B}(\mathbb{R}^d) \to [0, 1]$ with
\[
P_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\}), \quad B \in \mathcal{B}(\mathbb{R}^d)
\]
the distribution of $X = (X_1, \ldots, X_d)^\top$ and the function $F_X : \mathbb{R}^d \to [0, 1]$ with
\[
F_X(x_1, \ldots, x_d) = P(X_1 \leq x_1, \ldots, X_d \leq x_d) = P_X((-\infty, x_1] \times \cdots \times (-\infty, x_d))
\]
for $x = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$ the (cumulative) distribution function (cdf) of $X$. Two random vectors $X$ and $Y$ are said to be identically distributed, denoted by $X \overset{d}{=} Y$, if $P_X = P_Y$ or, equivalently, $F_X \equiv F_Y$.

Let $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ be independent and identically distributed (iid) random variables for some $n \in \mathbb{N}$. The mapping $\hat{F}_n : \mathbb{R} \times \Omega \to [0, 1]$ given by
\[
\hat{F}_n(x, \omega) = \frac{1}{n} \#\{i : 1 \leq i \leq n, X_i(\omega) \leq x\}
\]
is called the empirical distribution function of $X_1, \ldots, X_n$. The following theorem presents some of its properties.

**Theorem A.3** Let $x \in \mathbb{R}$. Then,

(i) For the random variable $n\hat{F}_n(x)$ it holds that
\[
P(n\hat{F}_n(x) = k) = \binom{n}{k} (F(x))^k (1 - F(x))^{n-k}
\]
for all $k = 0, \ldots, n$. In particular,
\[
\mathbb{E}\hat{F}_n(x) = F(x) \quad \text{and} \quad \text{Var}\hat{F}_n(x) = \frac{1}{n} F(x) (1 - F(x)).
\]

(ii) It holds that
\[
P\left(\lim_{n \to \infty}\hat{F}_n(x) = F(x)\right) = 1.
\]

(iii) If $F(x) \in (0, 1)$ then it holds for every $y \in \mathbb{R}$ that
\[
\lim_{n \to \infty} P\left(\sqrt{n} \frac{\hat{F}_n(x) - F(x)}{\sqrt{F(x)(1 - F(x))}} \leq y\right) = F_{\mathcal{N}(0,1)}(y).
\]
Finally, we mention Glivenko–Cantelli’s theorem, which extends the point–wise convergence mentioned in Theorem A.3 (ii).

**Theorem A.4** (Glivenko–Cantelli’s theorem) It holds that
\[
P\left( \lim_{n \to \infty} \sup_{x \in \mathbb{R}} |\bar{F}_n(x) - F(x)| = 0 \right) = 1.
\]

**Quantile Function**

Let \( F : \mathbb{R} \to \mathbb{R} \) be some non–decreasing and right–continuous function. Then, \( F^{-1} : \mathbb{R} \to \mathbb{R} \) given by
\[
F^{-1}(y) = \inf \{ x : F(x) \geq y \}
\]
(A.5)
is called the **generalized inverse function** of \( F \), where \( \inf \emptyset = \infty \). Consider a random variable \( X \) on \((\Omega, \sigma(\Omega), \mathbb{P})\) with cdf \( F_X \). Then the function \( F_X^{-1} \) is called the **quantile function** of \( X \).

**Moments**

The \( k \)th moment \( \mathbb{E}X^k \) for \( k \in \mathbb{N} \) of a (real–valued) random variable \( X \) with \( \int_{-\infty}^{\infty} |x|^k dF_X(x) < \infty \), where \( F_X \) is the cdf of \( X \), is given by
\[
\mathbb{E}X^k = \int_{\Omega} X^k(\omega) \mathbb{P}(d\omega) = \int_{-\infty}^{\infty} x^k dF_X(x), \quad k \in \mathbb{N}.
\]
In particular, \( \mathbb{E}X \) is called the **expectation** (or the **mean**) of \( X \). Moreover,
\[
\text{Var}X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2
\]
is called the **variance** of \( X \).

**Normal Distribution**

A random variable \( X \) is said to be **normally distributed** with mean \( \mu \in \mathbb{R} \) and variance \( \sigma^2 > 0 \), denoted by \( X \sim \mathcal{N}(\mu, \sigma^2) \), if
\[
F_X(x) = \frac{1}{\sqrt{2\pi \sigma^2}} \int_{-\infty}^{x} \exp\left\{ -\frac{1}{2} \left( \frac{t - \mu}{\sigma} \right)^2 \right\} dt, \quad x \in \mathbb{R}.
\]
In this case, the latter cdf \( F_X \) is denoted by \( F_{\mathcal{N}(\mu, \sigma^2)} \) and in particular in the case of the **standard normal distribution** with mean \( \mu = 0 \) and variance \( \sigma^2 = 1 \) we use the notation \( F_{\mathcal{N}(0,1)} \). Likewise, a \( d \)-dimensional random vector \( X = (X_1, \ldots, X_d)^\top \) is said to be **(regularly) \( d \)-variate normally distributed** with expectation vector \( \mu = (\mu_1, \ldots, \mu_d)^\top \) and (symmetric and positive definite) \( d \times d \) covariance matrix \( \Sigma = (\sigma_{ij})_{i,j=1}^{d} \), denoted by \( X \sim \mathcal{N}(\mu, \Sigma) \), if its \( (d \text{-dimensional}) \) distribution function \( F_X : \mathbb{R}^d \to [0,1] \) is given by
\[
F(x_1, \ldots, x_d) = \frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{\det \Sigma}} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} \exp\left\{ -\frac{1}{2} (t - \mu)^\top \Sigma (t - \mu) \right\} dt_1 \cdots dt_d
\]
for any vector \((x_1, \ldots, x_d)^\top \in \mathbb{R}^d\). This can be generalized along the following lines; pp. 397 in [16]. Note that for any symmetric and non-negative definite \(d \times d\) matrix \(\Sigma\) with rank \(r \leq d\) there exists a \(d \times r\) matrix \(K\) with rank \(r\) such that \(\Sigma = K K^\top\). Let \(Y = (Y_1, \ldots, Y_d)^\top\) be a \(d\)-dimensional random vector with expectation vector \(\mu = (\mu_1, \ldots, \mu_d)^\top\) and (symmetric and non-negative definite) \(d \times d\) covariance matrix \(\Sigma\) with rank \(r \leq d\). Then, \(Y\) is said to be (singularly) \(d\)-variate normally distributed, denoted by \(Y \sim \mathcal{N}(\mu, \Sigma)\), if \(Y = \mu + K X\) (and \(r < d\)), where \(K\) is a \(d \times r\) matrix with rank \(r\) such that \(\Sigma = K K^\top\). Also, \(X\) is an \(r\)-dimensional random vector with \(X \sim \mathcal{N}(0, I_r)\), where the matrix \(I_r\) is the \(r\)-dimensional identity matrix and \(o\) is the \(r\)-dimensional zero vector. Both in the case of the regular and the singular \(d\)-variate normal distribution with mean vector \(\mu\) and covariance matrix \(\Sigma\), we denote the \(d\)-variate cumulative distribution function by \(F_{\mathcal{N}(\mu, \Sigma)}\).

Note that the notation \(\mathcal{N}(\mu, \sigma^2)\) is also used to denote a real-valued Gaussian random variable with mean \(\mu\) and variance \(\sigma^2\). Likewise, \(\mathcal{N}(\mu, \Sigma)\) denotes also a \(d\)-dimensional Gaussian random vector with expectation vector \(\mu \in \mathbb{R}^d\) and covariance matrix \(\Sigma\).

\(\chi^2\)-Distribution

Let \(X = (X_1, \ldots, X_d)^\top \sim \mathcal{N}(0, I_d)\), where \(I_d\) is the \(d\)-dimensional identity matrix and \(o\) the \(d\)-dimensional zero vector. Then, \(X^\top X = \sum_{i=1}^d X_i^2 \sim \chi_d^2\) is (centrally) \(\chi^2\)-distributed with \(d\) degrees of freedom, denoted by \(X^\top X \sim \chi_d^2\).

### A.3.2 Inequalities

In this section we consider several inequalities for (real-valued) random variables. The following theorem presents several standard inequalities. The proofs of the different inequalities can be found in Section 4.4 of [62].

**Theorem A.5** Let \(X\) and \(Y\) be two real-valued random variables on \((\Omega, \sigma(\Omega), \mathbb{P})\) with cdf \(F_X\) and \(F_Y\), respectively.

(i) (Inequality of Cauchy–Schwarz) Assume that both \(\int_{\mathbb{R}} x^2 dF_X(x) < \infty\) and \(\int_{\mathbb{R}} x^2 dF_Y(x) < \infty\). Then it holds for the random variable \(Z = XY\) that \(\int_{\mathbb{R}} |x| dF_Z(x) < \infty\), where \(F_Z\) is the cdf of \(Z\), and

\[
\mathbb{E}|XY| \leq \sqrt{\mathbb{E}X^2 \mathbb{E}Y^2}. \tag{A.6}
\]

(ii) (Jensen’s inequality) Let \(g: \mathbb{R} \to \mathbb{R}\) be a convex function and assume that both

\[
\int_{\mathbb{R}} |x| dF_X(x) < \infty \quad \text{and} \quad \int_{\mathbb{R}} |g(x)| dF_X(x) < \infty.
\]

Then,

\[
g(\mathbb{E}X) \leq \mathbb{E}g(X). \tag{A.7}
\]
(iii) (Čebyshev’s inequality) Assume that \( \int_{\mathbb{R}} x^2 \, dF_X(x) < \infty \). Then for each \( \varepsilon > 0 \),
\[
P( |X - \mathbb{E}X| \geq \varepsilon ) \leq \frac{\text{Var}\,X}{\varepsilon^2}.
\]  
(A.8)

The following inequality is known as the Hâjk–Rényi inequality in literature. The proof can be found on pp. 53 in [96].

**Theorem A.6** For some \( n \in \mathbb{N} \) let \( X_1, \ldots, X_n \) be real-valued and independent random variables with mean zero and assume that \( \mathbb{E}X_k^2 < \infty \) for \( k = 1, \ldots, n \). Furthermore, let \( S_k = \sum_{i=1}^{k} X_i \) for \( k = 1, \ldots, n \) and assume that \( c_1, c_2, \ldots, c_n \) are positive constants satisfying \( c_1 \geq c_2 \geq \ldots \geq c_n \). Then,
\[
P \left( \max_{m \leq k \leq n} c_k |S_k| \geq x \right) \leq \frac{1}{x^2} \left( \frac{c_1^2}{m} \sum_{k=1}^{m} \mathbb{E}X_k^2 + \sum_{k=m+1}^{n} c_k^2 \mathbb{E}X_k^2 \right)
\]  
(A.9)

for every \( x > 0 \) and any \( m \in \mathbb{N} \) with \( m < n \).

### A.3.3 Convergence of Random Variables

Let \( X_1, X_2, \ldots \) and \( X \) be real-valued random variables on \( (\Omega, \sigma(\Omega), \mathbb{P}) \). Then, as \( n \to \infty \),

(i) \( X_n \) converges in probability to \( X \), denoted by \( X_n \xrightarrow{p} X \), if \( \lim_{n \to \infty} \mathbb{P}( |X_n - X| < \varepsilon ) = 1 \) for all \( \varepsilon > 0 \),

(ii) \( X_n \) converges almost surely (a.s.) to \( X \), denoted by \( X_n \xrightarrow{a.s.} X \), if \( \mathbb{P}( \lim_{n \to \infty} X_n = X ) = 1 \),

(iii) \( X_n \) converges to \( X \) in mean and in quadratic mean if \( \lim_{n \to \infty} \mathbb{E}( |X_n - X| ) = 0 \) and \( \lim_{n \to \infty} \mathbb{E}( (X_n - X)^2 ) = 0 \), respectively, and

(iv) \( X_n \) converges weakly to \( X \) (convergence in distribution), denoted by \( X_n \xrightarrow{d} X \), if \( \lim_{n \to \infty} F_{X_n}(x) = F_X(x) \) at each continuity point \( x \in \mathbb{R} \) of \( F_X \).

Note that \( X_n \xrightarrow{a.s.} X \) implies \( X_n \xrightarrow{p} X \) as \( n \to \infty \) and that both convergence in mean and in quadratic mean imply convergence in probability. Furthermore, the following theorem provides a necessary and sufficient condition for the almost sure convergence of \( X_n \) to \( X \) as \( n \to \infty \). The proof of this theorem can be found on p. 137 in [62] or on p. 204 in [96].

**Theorem A.7** Let \( X_1, X_2, \ldots \) be a sequence of real-valued random variables. Then it holds that
\[
X_n \xrightarrow{a.s.} X
\]  
as \( n \to \infty \).
if and only if
\[ \lim_{n \to \infty} P \left( \sup_{k \geq n} |X_k - X| > \varepsilon \right) = \lim_{n \to \infty} P \left( \bigcup_{k \geq n} \{|X_n - X| > \varepsilon\} \right) = 0 \]
for all \( \varepsilon > 0 \).

Often, the following theorem is useful, for example in order to prove central limit theorems.

**Theorem A.8** (Slutsky’s theorem) Let \( X_1, X_2, \ldots \) and \( Y_1, Y_2, \ldots \) be two sequences of real-valued random variables such that \( X_n \xrightarrow{d} X \) and \( Y_n \xrightarrow{P} c \), where \( c \) is a finite constant. Then,

(i) \( X_n + Y_n \xrightarrow{n \to \infty} X + c \),

(ii) \( X_n Y_n \xrightarrow{n \to \infty} cX \),

(iii) \( X_n/Y_n \xrightarrow{n \to \infty} X/c \) if \( c \neq 0 \).

The above introduced convergence concepts of random variables can easily be extended to the same notions with respect to \((d\text{-dimensional})\) random vectors in \( \mathbb{R}^d \) by the following theorem.

**Theorem A.9** A sequence of random vectors \( X_1, X_2, \ldots \) in \( \mathbb{R}^d \) converges to a random vector \( X \) almost surely, in probability, in mean, and in quadratic mean, respectively, if and only if each component of the vector \( X_n \) converges to the same component of the vector \( X \) almost surely, in probability, in mean, and in quadratic mean, respectively, as \( n \to \infty \). Moreover, if the sequence \( X_1, X_2, \ldots \) of random vectors converges weakly to \( X \), then each component of the vector \( X_n \) converges weakly to the same component of the vector \( X \) as \( n \to \infty \).

The following result 'completes' the 'incomplete' statement of the Theorem A.9 in the case of weak convergence of random vectors.

**Theorem A.10** (Cramér–Wold device) Let \( X_1, X_2, \ldots, \) and \( X \) be random vectors in \( \mathbb{R}^d \). Then \( X_n \) converges weakly to \( X \) as \( n \to \infty \) if and only if each linear combination of the components of \( X_n \) converges weakly to the same linear combination of the components of \( X \) as \( n \to \infty \).

### A.3.4 Laws of Large Numbers

In the following, we mention a weak and a strong law of large numbers in the classical sense. Detailed information can be found in Chapters 4 and 6 of [96] or in Section 1.8 of [113].
A.3. Concepts of Probability Calculus

Theorem A.11 Let $X_1, X_2, \ldots$ be a sequence of real-valued and iid random variables with mean $E X_1 = \mu$ and (positive) variance $\text{Var} X_1 = \sigma^2 < \infty$. Then,

(a) $\overline{X}_n \xrightarrow{P} n \rightarrow \infty \mu$ (weak law of large numbers), and

(b) $\overline{X}_n \xrightarrow{a.s.} n \rightarrow \infty \mu$ (strong law of large numbers),

where $\overline{X}_n = \sum_{i=1}^n X_i$.

A.3.5 Classical Central Limit Theorems; Rates of Convergence

Let $X_1, X_2, \ldots$ and $X$ be real-valued random variables on the probability space $(\Omega, \sigma(\Omega), \mathbb{P})$ and let $X \sim \mathcal{N}(\mu, \sigma^2)$ with mean $\mu \in \mathbb{R}$ and (finite) variance $\sigma^2 > 0$. Then, in order to express the weak convergence $X_n \xrightarrow{d} n \rightarrow \infty X$, we also use the notation $X_n \xrightarrow{d} n \rightarrow \infty \mathcal{N}(\mu, \sigma^2)$, which means that $X_n$ converges (weakly) to a Gaussian random variable with mean $\mu$ and variance $\sigma^2$ as $n \rightarrow \infty$. Alternatively, but equivalently, the convergence in the supremum norm, i.e.,

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |\mathbb{P}(X_n \leq x) - F_{\mathcal{N}(\mu, \sigma^2)}(x)| = 0,$$

is used to express the weak convergence of $X_n$, where $F_{\mathcal{N}(\mu, \sigma^2)}(\cdot)$ denotes the cumulative distribution function of a Gaussian random variable $\mathcal{N}(\mu, \sigma^2)$ with mean $\mu \in \mathbb{R}$ and finite variance $\sigma^2 > 0$.

Analogously, for the sequence $X_1, X_2, \ldots$ of random vectors in $\mathbb{R}^d$, we write $X_n \xrightarrow{d} n \rightarrow \infty \mathcal{N}(\mu, \Sigma)$ in order to express (weak) convergence of $X_n$ to the Gaussian random vector $\mathcal{N}(\mu, \Sigma)$ with mean $\mu \in \mathbb{R}^d$ and $d \times d$ covariance matrix $\Sigma$ as $n \rightarrow \infty$. Again alternatively, but equivalently, the convergence in the supremum norm, i.e.,

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} |\mathbb{P}(X_n \leq x) - F_{\mathcal{N}(\mu, \Sigma)}(x)| = 0,$$

is used to express the weak convergence of the $d$-dimensional random vector $X_n$ to a Gaussian random vector $\mathcal{N}(\mu, \Sigma)$ with mean $\mu \in \mathbb{R}^d$ and $d \times d$ covariance matrix $\Sigma$ whose cumulative distribution function is $F_{\mathcal{N}(\mu, \Sigma)}(\cdot)$.

The following classical central limit theorem is sometimes called the theorem of Lindeberg–Lévy; p. 28 in [113]. Its proof can be found in Chapter VIII of [99] for example; cf. also Section 4.2 of [96].

Theorem A.12 Let $X_1, X_2, \ldots$ be an iid sequence of (real-valued) random variables having mean $\mu$ and finite variance $\sigma^2 > 0$. Then,

$$\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n X_i - \mu \right) \xrightarrow{d} n \rightarrow \infty \mathcal{N}(0, \sigma^2).$$
Likewise, for an iid sequence $X_1, X_2, \ldots$ of $d$-dimensional random vectors having mean vector $\mu \in \mathbb{R}^d$ and $d \times d$ covariance matrix $\Sigma$, 

$$\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right) \xrightarrow{\mathcal{D}} N(0, \Sigma),$$

where $o$ is the $d$-dimensional zero vector.

The subsequent central limit theorem generalizes Theorem A.12 since it considers random sums of (real-valued) random variables and states their (weak) convergence to normal limits. The proof can be found for example in Chapter VIII of [99].

**Theorem A.13** Let $X_1, X_2, \ldots$ be an iid sequence of random variables having mean $\mu$ and finite variance $\sigma^2 > 0$. Let $N_1, N_2, \ldots : \Omega \to \mathbb{N}$ be a sequence of integer-valued random variables and consider the sequence $c_1, c_2, \ldots$ of positive constants with $c_n \to \infty$ as $n \to \infty$ such that

$$\frac{N_n}{c_n} \xrightarrow{P} c,$$

where $c$ denotes some positive constant. Then,

$$\left( N_n \right)^{-1/2} \sum_{i=1}^{N_n} (X_i - \mu) \xrightarrow{\mathcal{D}} N(0, \sigma^2).$$

We now turn briefly to the issue of rates of convergence in the central limit theorems. For a detailed discussion of this topic, we refer to Chapter 5 of [96] for example.

**Theorem A.14** Let $X_1, X_2, \ldots, X_n$ for $n \in \mathbb{N}$ be real-valued random variables with mean zero.

(a) Assume that $X_1, X_2, \ldots, X_n$ are independent and that $\mathbb{E}|X_i|^3 < \infty$ for all $i = 1, 2, \ldots n$.

Let 

$$B^2_n = \sum_{i=1}^{n} \mathbb{E}X_i^2 \quad \text{and} \quad L_n = B_n^{-3} \sum_{i=1}^{n} \mathbb{E}|X_i|^3.$$

Then there exists a finite constant $c > 0$ such that for any $n \in \mathbb{N}$,

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \sum_{i=1}^{n} X_i \leq x B_n \right) - F_{N(0,1)}(x) \right| \leq c L_n. \quad (A.10)$$

(b) Assume that $X_1, X_2, \ldots, X_n$ are iid with $\mathbb{E}|X_i|^3 < \infty$ and variance $\sigma^2 > 0$. Then there exists a finite constant $c > 0$ such that for any $n \in \mathbb{N}$,

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( n^{-1/2} \sum_{i=1}^{n} X_i \leq x \sigma \right) - F_{N(0,1)}(x) \right| \leq c n^{-1/2} \frac{\mathbb{E}|X_i|^3}{\sigma^3}. \quad (A.11)$$
Note that (A.10) is called Esseen’s inequality, whereas (A.11) is called the Berry–Esseen inequality. However, both for (A.10) and (A.11) to hold, the existence of the third-order moments is needed. The following theorem provides an estimate of the difference of the cdf of some (normalized) sum of independent (real-valued) random variables and the cdf of the standard normal distribution under weaker conditions on the moments.

**Theorem A.15** Assume that the function \( g : \mathbb{R} \to [0, \infty) \) satisfies the following conditions,

(i) \( g(-x) = g(x) \) for \( x \in \mathbb{R} \),

(ii) \( x \mapsto g(x) \) is non-decreasing for all \( x > 0 \),

(iii) \( x \mapsto x/g(x) \) is non-decreasing for all \( x > 0 \).

Let \( X_1, X_2, \ldots, X_n \) be independent real-valued random variables for \( n \in \mathbb{N} \) with mean zero and

\[
\mathbb{E}(X_i^2 g(X_i)) < \infty \quad \text{for all} \quad i = 1, 2, \ldots, n.
\]

Put \( B_n^2 = \sum_{i=1}^{n} \mathbb{E}X_i^2 \). Then, there exists a (finite) constant \( c > 0 \) such that for any \( n \in \mathbb{N} \),

\[
\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left( \sum_{i=1}^{n} X_i \leq x B_n \right) - F_{N(0,1)}(x) \right| \leq \frac{c}{B_n^2} \sum_{i=1}^{n} \mathbb{E}(X_i^2 g(X_i)). \tag{A.12}
\]

The proof of Theorem A.15 can be found on p. 151 in [96]. Immediately, we get the following corollary.

**Corollary A.2** Let \( X_1, X_2, \ldots, X_n \) for \( n \in \mathbb{N} \) be a independent real-valued random variables with mean zero. Then there exist two (finite) constants \( c_1, c_2 > 0 \) such that, for any \( \varepsilon \in (0, 1) \) and \( n \in \mathbb{N} \),

\[
\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left( \sum_{i=1}^{n} X_i \leq x B_n \right) - F_{N(0,1)}(x) \right| \leq c_1 \varepsilon + \frac{c_2}{B_n^2} \sum_{i=1}^{n} \mathbb{E}X_i^2 \mathbb{I}\{|X_i| \geq \varepsilon B_n\},
\]

where \( B_n^2 = \sum_{i=1}^{n} \mathbb{E}X_i^2 \).

**Proof** Let \( n \in \mathbb{N} \). Put \( Y_1 = X_1/B_n \), \( Y_2 = X_2/B_n \), \ldots, \( Y_n = X_n/B_n \) and apply Theorem A.15 to the random variables \( Y_1, Y_2, \ldots, Y_n \). Consider the function \( g : \mathbb{R} \to [0, \infty) \) given by

\[
g(x) = \begin{cases} 
\varepsilon & \text{if } |x| < \varepsilon, \\
|x| & \text{if } \varepsilon \leq |x| < 1, \\
1 & \text{if } |x| \geq 1.
\end{cases}
\]

Then \( g \) is an even function, which is non-decreasing for all \( x > 0 \), and \( x \mapsto x/g(x) \) is also non-decreasing for all \( x > 0 \). Since

\[
\sum_{i=1}^{n} \mathbb{E}(Y_i^2 g(Y_i)) = \frac{1}{B_n^2} \sum_{i=1}^{n} \mathbb{E}\left(X_i^2 g\left(\frac{X_i}{B_n}\right)\right),
\]

the corollary is proven in view of Theorem A.15. \(\square\)
A.3.6 Variance Stabilizing Transformation

When regarding CLTs, often the asymptotic variance depends explicitly on the centering constant. This undesired effect with regard to the derivation of (asymptotic) confidence intervals for example can be accounted for by a so-called variance stabilizing transformation ([15] or [127]).

Theorem A.16 Let $X_1, X_2, \ldots$ be a sequence of $(k$-dimensional) random vectors in $\mathbb{R}^k$ for $k \in \mathbb{N}$ defined on some probability space $(\Omega, \sigma(\Omega), \mathbb{P})$. Assume that $X_n \xrightarrow{P_{n \to \infty}} c$ for some $c \in \mathbb{R}^k$, and let $f : \mathbb{R}^k \to \mathbb{R}^\ell$ be a measurable function for any $k, \ell \in \mathbb{N}$.

(i) If $f$ is continuous at $c$ we have that

$$f(X_n) \xrightarrow{P_{n \to \infty}} f(c). \quad (A.13)$$

(ii) If $f$ is differentiable at $c$ and if

$$\sqrt{n} \left( X_n - c \right) \xrightarrow{d_{n \to \infty}} \mathcal{N}(\mathbf{o}, \Sigma),$$

where $\mathcal{N}(\mathbf{o}, \Sigma)$ denotes a $k$-dimensional Gaussian vector with mean $\mathbf{o}$ and covariance matrix $\Sigma$. Then,

$$\sqrt{n} \left( f(X_n) - f(c) \right) \xrightarrow{d_{n \to \infty}} \mathcal{N}(\mathbf{o}, \nabla(c) \Sigma (\nabla(c))^\top), \quad (A.14)$$

where $\mathcal{N}(\mathbf{o}, \nabla(c) \Sigma (\nabla(c))^\top)$ denotes an $\ell$-dimensional Gaussian vector with mean $\mathbf{o}$ and covariance matrix $\nabla(c) \Sigma (\nabla(c))^\top$. The $(i, j)$-th entry of the $\ell \times k$ Jacobi matrix $\nabla(c)$ is given by the partial derivative $\partial f_i(c)/\partial x_j$ for $i = 1, \ldots, \ell$ and $j = 1, \ldots, k$.

A.3.7 $U$-Statistics and Hoeffding’s Decomposition

In the following a special class of statistics, so-called $U$-statistics, are introduced and some properties thereof are considered, in particular Hoeffding’s decomposition and Hoeffding’s CLT. A more detailed discussion of $U$-statistics can be found in [65] and [113].

Let $X_1, X_2, \ldots$ be a sequence of iid random vectors in $\mathbb{R}^d$ with cdf $F$ and let the function $f : \mathbb{R}^{md} \to \mathbb{R}$ be Borel-measurable and symmetric with $\mathbb{E}[f(X_1, \ldots, X_m)] < \infty$, where symmetric means that $f(x_{i(1)}, \ldots, x_{i(m)}) = f(x_1, \ldots, x_m)$ for any permutation $\iota : \{1, \ldots, m\} \to \{1, \ldots, m\}$.

Then we call

$$U_n^{(m)}(f) = \frac{1}{\binom{m}{n}} \sum_{1 \leq i_1 < \ldots < i_m \leq n} f(X_{i_1}, \ldots, X_{i_m}) \quad \text{for } n \geq m \quad (A.15)$$
a $U$-statistic of order $m$ with the kernel function $f$. Note that $U^{(m)}_n(f)$ is an unbiased estimator for $\mu = \mathbb{E}f(X_1, \ldots, X_n)$ since

$$
\mathbb{E}U^{(m)}_n(f) = \frac{1}{\binom{n}{m}} \sum_{1 \leq i_1 < \ldots < i_m \leq n} \mathbb{E}f(X_{i_1}, \ldots, X_{i_m}) = \mu,
$$

and that $U^{(m)}_n(f)$ gives rise to many estimators from classical statistical inference. Consider for example the case $m = 1$, where

$$
U^{(1)}_n(f) = \frac{1}{n} \sum_{i=1}^{n} f(X_i) \quad \text{and} \quad \mu = \mathbb{E}f(X_1) = \int f(x)dF(x).
$$

If $f(x) = x$ (and $d = 1$) the sample mean $U^{(1)}_n(f) = \overline{X}_n$ with $\mathbb{E}\overline{X}_n = \mu$ is obtained. Another well-known example is obtained for $m = 2$ and $f(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$ (again with $d = 1$), since then the sample variance $S^2_n$ is obtained by

$$
U^{(2)}_n(f) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \frac{1}{2}(X_i - X_j)^2 = \frac{1}{n(n-1)} \left( (n-1) \sum_{i=1}^{n} X_i^2 - \left( \sum_{i=1}^{n} X_i \right)^2 + \sum_{i=1}^{n} X_i^2 \right) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n) = S^2_n.
$$

By elementary rearrangements any $U$–statistic admits Hoeffding’s decomposition, i.e., $U^{(m)}_n(f)$ may be rewritten (Chapter 5.1.5 in [113]) as

$$
U^{(m)}_n(f) = \mu + \frac{m}{n} \sum_{i=1}^{n} (g(X_i) - \mu) + R^{(m)}_n(f),
$$

where $g(x) = \mathbb{E}f(x, X_2, \ldots, X_m)$ and

$$
R^{(m)}_n(f) = \sum_{k=2}^{m} \binom{m}{k} \left( \binom{n}{k} \right)^{-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \Phi_k(X_{i_1}, \ldots, X_{i_k}),
$$

with functions $\Phi_k : \mathbb{R}^k \rightarrow \mathbb{R}, 2 \leq k \leq m$, defined by

$$
\Phi_k(x_1, \ldots, x_k) = \sum_{j=1}^{k} (-1)^{k-j} \sum_{1 \leq i_1 < \ldots < i_j \leq k} \left( \mathbb{E}f(x_{i_1}, \ldots, x_{i_j}, X_{j+1}, \ldots, X_m) - \mu \right).
$$

Note that $\widetilde{U}^{(m)}_n(f) = U^{(m)}_n(f) - R^{(m)}_n(f)$ is called the projection of $U^{(m)}_n(f)$.

The crucial outcome of Hoeffding’s decomposition (A.16) can be summarized by

$$
\mathbb{E}\left( R^{(m)}_n(f) \right)^2 = O(n^{-2}, \quad n \rightarrow \infty),
$$

(A.17)
provided that $\mathbb{E} f^2(X_1, \ldots, X_m) < \infty$, cf. Chapter 5.3.2 in [113], which leads to the (almost
sure) representation of the $U$–statistic $U_{n}^{(m)}(f)$, $m \geq 2$, as mean of iid random variables. Indeed, it can be shown that

$$\text{Var} \, R_{n}^{(m)}(f) = \text{Var} \, U_{n}^{(m)}(f) - \text{Var} \, \widetilde{U}_{n}^{(m)}(f),$$

which, together with $\mathbb{E} R_{n}^{(m)}(f) = 0$, yields the estimate

$$\mathbb{E}(R_{n}^{(m)}(f))^2 \leq \frac{c_m}{n^2} \mathbb{E} f^2(X_1, \ldots, X_m) \quad \text{for} \quad n \geq m,$$  \hspace{1cm} (A.18)

where $c_m < \infty$ is a constant depending only on $m$. Thus, by (A.18), we have that (A.17) holds.

Finally, provided that $\mathbb{E} f^2(X_1, \ldots, X_m) < \infty$ (and $\mathbb{E} g^2(X_1) - \mu^2 > 0$), the latter result (A.18) immediately leads to Hoeffding’s CLT for $U$–statistics (Chapter 5.5 in [113]), i.e.,

$$\sqrt{n} \left( U_{n}^{(m)}(f) - \mu \right) \lim_{n \to \infty} \mathbb{N}(0, m^2(\mathbb{E} g^2(X_1) - \mu^2)).$$  \hspace{1cm} (A.19)

### A.3.8 Goodness–of–Fit Test

This section briefly presents some goodness–of–fit techniques and in particular goodness–of–fit tests. The latter are used to test the hypothesis whether the distribution of some sample data is a normal distribution. In the univariate case, popular choices for goodness–of–fit tests are the Pearson test, the Pearson-Fisher test, and the Kolmogorov-Smirnov test, where we refer to [22] for further information. Apart from these tests, also graphical examination of the given sample data helps to analyse the goodness–of–fit of certain models. As an example one may consider histograms, quantile plots, and plots of the empirical distribution function.

Let $X_1, \ldots, X_n$ be real–valued and iid random variables for some $n \in \mathbb{N}$ and let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ with $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ be the order statistic of $X_1, \ldots, X_n$. The general philosophy of a quantile plot, cf. Section 2.6.2 of [101], is to visualize the information given by (a realization of) the iid sample $X_1, \ldots, X_n$ in connection with the proposed type of distribution. For example in case of normality, cf. p. 71 in [101], we consider the plot

$$\left\{ \left( F_{\mathcal{N}(0,1)}^{-1}\left( \frac{k}{n+1} \right), X_{(k)} \right), 1 \leq k \leq n \right\},$$

where $F_{\mathcal{N}(0,1)}^{-1}$ denotes the quantile function of a Gaussian random variable with mean zero and variance one. If the resulting point plot is 'linear' this strongly indicates that the true distribution function of the particular feature (observed via the sample $X_1, \ldots, X_n$) can be assumed to be normal. Note that the quantile function $F_{\mathcal{N}(\mu, \sigma^2)}^{-1}$ of the normal distribution with mean $\mu$ and (finite and positive) variance $\sigma^2$ is given by $F_{\mathcal{N}(\mu, \sigma^2)}^{-1} = \mu + \sigma F_{\mathcal{N}(0,1)}^{-1}$. Thus, the slope of the regression line fitted via the least squares method to the points of the quantile plot is an estimator for $\sigma$ and the intercept at zero estimates $\mu$. 
We now discuss an example of a goodness-of-fit test provided in [56] on the multivariate normal distribution. In [80] an overview of alternative tests and a bibliography of further literature can be found. In many cases however, the test in [56] works well and even outperforms competing tests ([55] and [80]).

Let \( X_1, \ldots, X_n \) be random vectors with values in \( \mathbb{R}^d, \ d \geq 1 \), and assume that they are iid copies of some generic random vector \( X \), where

\[
\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{and} \quad S_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)(X_i - \bar{X}_n)^\top
\]
denote the sample mean and sample covariance matrix. Assume that we want to test the (composite) hypothesis

\[ H_0^{(d)} : \text{The distribution law of } X \text{ is } \mathcal{N}(\mu, \Sigma) \text{ for some mean vector } \mu \in \mathbb{R}^d \text{ and some symmetric and regular, i.e., positive definite, } d \times d \text{ matrix } \Sigma. \]

The test statistic \( T_{n,\beta} \) is given by

\[
T_{n,\beta} = T_{n,\beta}(X_1, \ldots, X_n) = 4n \mathbb{I}\{S_n \text{ singular}\} + n \, D_{n,\beta} \mathbb{I}\{S_n \text{ regular}\},
\]
where \( D_{n,\beta} \) is an \( L^2 \)-distance measure between a multivariate normal density and some kernel density estimator with bandwidth inversely proportional to \( \beta \). The value of \( \beta \) can be fixed or chosen to depend on \( n \), for example

\[
\beta = \beta_d(n) = \frac{1}{\sqrt{2}} \left( \frac{2d + 1}{4} - \frac{1}{n} \right)^{1/2}.
\]

For the \( L^2 \)-distance \( D_{n,\beta} \), it is shown in ([56]) that

\[
D_{n,\beta} = \frac{1}{n^2} \sum_{i,j=1}^{n} \exp\left( -\frac{\beta^2}{2} \| Y_i - Y_j \|^2 \right)
- \frac{2}{(1 + \beta^2)^{d/2}} \frac{1}{n} \sum_{i=1}^{n} \exp\left( -\frac{\beta^2}{2(1 + \beta^2)} \| Y_i \|^2 \right) + \frac{1}{(1 + 2\beta^2)^{d/2}},
\]
where \( Y_i = S_n^{-1/2}(X_i - \bar{X}_n) \) for \( i = 1, \ldots, n \). The hypothesis \( H_0^{(d)} \) is rejected at the significance level \( \alpha \) if

\[
T_{n,\beta} > q_{d,\beta}(1 - \alpha),
\]
where \( q_{d,\beta}(1 - \alpha) \) denotes the \((1 - \alpha)\)-quantile of the limiting distribution of \( T_{n,\beta} \) as \( n \to \infty \).

This latter \((1 - \alpha)\)-quantile \( q_{d,\beta}(1 - \alpha) \) is given by

\[
q_{d,\beta}(1 - \alpha) = \mu_{d,\beta} \left( 1 + \frac{\sigma_{d,\beta}^2}{\mu_{d,\beta}^2} \right)^{-1/2} \exp \left( z_{1-\alpha} \left( \log \left( 1 + \frac{\sigma_{d,\beta}^2}{\mu_{d,\beta}^2} \right) \right)^{1/2} \right),
\]
where \( z_{1-\alpha} \) is the \((1-\alpha)\)-quantile of the standard normal distribution. The symbols \( \mu_{d,\beta} \) and \( \sigma_{d,\beta}^2 \) are given by

\[
\mu_{d,\beta} = 1 - \frac{1}{(1 + 2\beta^2)^{d/2}} \left( 1 + \frac{d \beta^2}{1 + 2\beta^2} + \frac{d(d + 2) \beta^4}{2(1 + 2\beta^2)^2} \right)
\]
and

\[
\sigma_{d,\beta}^2 = \frac{2}{(1 + 4\beta^2)^{d/2}} + \frac{2}{(1 + 2\beta^2)^d} \left( 1 + \frac{2d \beta^4}{(1 + 2\beta^2)^2} + \frac{3d(d + 2) \beta^8}{4(1 + 2\beta^2)^4} \right)
\times \frac{4}{\omega(\beta)^{d/2}} \left( 1 + \frac{3d \beta^4}{2\omega(\beta)} + \frac{d(d + 2) \beta^8}{2\omega^2(\beta)} \right)
\]
respectively, where \( \omega(\beta) = (1 + \beta^2)(1 + 3\beta^2) \).

### A.4 Intrinsic Volumes

Consider the family \( \mathcal{C} \) of convex bodies in \( \mathbb{R}^d \) together with the functionals \( V_j : \mathcal{C} \to \mathbb{R} \) for \( j = 0, \ldots, d \), where \( V_j(C) \) is called the \( j \)th intrinsic volume of \( C \in \mathcal{C} \), defined for \( j = 0, \ldots, d \) implicitly through Steiner’s formula

\[
\nu_d(C \oplus B^d_r) = \sum_{j=0}^{d} r^{d-j} \kappa_{d-j} V_j(C), \quad C \in \mathcal{C},
\]
for the volume \( \nu_d(C \oplus B^d_r) \) of parallel convex bodies \( C \oplus B^d_r \) of \( C \), where, recall, \( B^d_r \) denotes the \( d \)-dimensional ball with radius \( r > 0 \) centered at the origin and where \( \kappa_{d-j} = \nu_{d-j}(B^d_1) \) for \( j = 0, \ldots, d-1 \) is given by (A.3). Note that in particular \( V_d(C) = \nu_d(C) \) and \( V_0(C) = \chi(C) \), where \( \chi(C) = 1 \) if \( C \neq \emptyset \) and zero otherwise.

Alternatively to \( V_j \) the \( j \)th quermassintegral or Minkowski functional \( M_j : \mathcal{C} \to \mathbb{R} \) for \( j = 0, \ldots, d \) can be regarded, where

\[
M_j(C) = \frac{\kappa_j}{\binom{d}{j}} V_{d-j}(C), \quad C \in \mathcal{C}, j = 0, \ldots, d.
\]

The following theorem summarizes some properties of intrinsic volumes, where \( D(C) = \sup\{||x - y|| : x, y \in C \} \) denotes the diameter of \( C \in \mathcal{C} \), which fulfills \( D(\rho C) = \rho D(C) \) for \( \rho > 0 \).

**Theorem A.17** Let \( V_j : \mathcal{C} \to \mathbb{R} \) be the \( j \)th intrinsic volume; \( j = 0, \ldots, d \). Then, for any \( j = 0, \ldots, d \) and \( C \in \mathcal{C} \),

\[
V_j(C) \leq \frac{\kappa_d}{\kappa_{d-j}} \left( \frac{d}{d-j} \right) \left( \frac{D(C)}{2} \right)^j.
\]

(A.21)
Furthermore,

\[ V_j(B_1^d) = \binom{d}{j} \frac{\kappa_d}{\kappa_{d-j}}, \quad j = 0, \ldots, d \]  \hspace{1cm} (A.22)

and

\[ V_j(\rho C) = \rho^j V_j(C), \quad \rho > 0 \]  \hspace{1cm} (A.23)

as well as \( V_j(\emptyset) = 0 \), \( V_j(C + x) = V_j(C) \) for any \( x \in \mathbb{R}^d \), and \( V_j(\theta_R C) = V_j(C) \) for any rotation \( \theta_R \) around the origin. Also, \( V_j(C_1) \leq V_j(C_2) \) for \( C_1, C_2 \in \mathcal{C} \) with \( C_1 \subseteq C_2 \).

The proof of Theorem A.17 immediately follows from Satz 2.2.2 in [111] except for the inequality (A.21) which in turn is a consequence of isodiametric inequalities in § 22 of [67]. For further information on convex and integral geometry, we refer to [110] and [111] for example.
Appendix B

Zusammenfassung


Komponenten der niederer Hierarchiestufe, die sich in der Einflusszone befinden, mit deren Zentrum verbunden werden.


Danach wenden wir uns dem Problem der statistischen Modellanpassung im zweidimensionalen Euklidischen Raum $\mathbb{R}^2$ zu. Hier untersuchen wir näher, wie wir angesichts gegebener netzwerkartiger Daten zwischen verschiedenen Mosaikmodellen entscheiden können, um das am besten passende Modell und seine entsprechenden Modellparameter zu finden. Das Vorgehen basiert hier im Wesentlichen auf Monte-Carlo-Simulationstechniken, welche im Vergleich zu den zentralen Grenzwertsätzen weniger restrictiv in ihrer Anwendung sind.

Im Einzelnen ergibt sich folgende Gliederung der vorliegenden Arbeit.

Nach einer Einleitung und Motivation in Kapitel 1, wobei insbesondere auch das SSLM genauer beschrieben wird, werden in den Kapiteln 2 und 3 einige Konzepte der stochastischen Geometrie behandelt, welche im weiteren Verlauf benötigt werden. Wegen ihrer besonderen Bedeutung bei den zentralen Grenzwertsätzen und der Modellierung von netzwerkartigen Strukturen werden zufällige Mosaik dabei separat in Kapitel 3 behandelt. Am Ende desselben Kapitels beweisen wir ein erstes Ergebnis, eine Aussage über die asymptotische Irrelevanz derjenigen Zellen eines stationären (und ergodischen) zufälligen Mosaiks, die nicht vollständig in einem konvexen und beschränkten Beobachtungsfensters $W_g^d \subset \mathbb{R}^d$ enthalten sind. Dabei bedeutet asymptotisch, dass $W_g^d$ unbeschränkt und gleichmäßig in alle Richtungen wächst.

In Kapitel 4 leiten wir einen zentralen Grenzwertsatz für die Anzahl der Schnittpunkte her, die durch Schnitte von Hyperebenen eines stationären Poissonischen Hyperebenenprozesses im $\mathbb{R}^d$ induziert werden. Wir nehmen dabei an, dass wir eine Realisierung dieses Prozesses in einer Folge von $d$-dimensionalen, im Ursprung zentrierten Kugeln $B_g^d$ mit wachsendem Radius
r > 0 beobachten können. Durch unser Resultat verallgemeinern wir ein Ergebnis von K. Paroux ([94]) für den Schnittpunktprozess von stationären und isotropen Poissonischen Geradenprozessen im $\mathbb{R}^2$. Unser Beweis des zentralen Grenzwertsatzes basiert auf Hoeffdings Zerlegung von $U$-Statistiken, einer Methode, die für den $d$-dimensionalen Fall weitaus effizienter und angenehrt erscheint als die „Momentenmethode“, die in [94] für den zweidimensionalen Fall verwendet wurde. Darüber hinaus erweitern wir unseren zentralen Grenzwertsatz in mehrere Richtungen. Zunächst betrachten wir $k$-Ebenenprozesse, die durch die Schnitte von $d - k$ Hyperebenen des zugrundeliegenden stationären Poissonischen Hyperebenenprozesses induziert werden ($0 \leq k \leq d - 1$) und leiten auch hier zentrale Grenzwertsätze sowohl für die Anzahl als auch für das $k$-Volumen dieser $k$-Ebenenprozesse in wachsenden Kugeln $B^d_r$ her. Schließlich betrachten wir den multivariaten Fall und beweisen zentrale Grenzwertsätze für die $d$-dimensionalen Vektoren der Anzahlen und des $k$-Volumens dieser induzierten $k$-Ebenenprozesse.

Im sich daran anschließenden Kapitel 5 werden die in Kapitel 4 hergeleiteten zentralen Grenzwertsätze verwendet, um asymptotische Konfidenzintervalle und asymptotische Tests für die Intensitäten der durch die Schnitte von Poissonischen Hyperebenen induzierten $k$-Ebenenprozesse herzuleiten. Dabei wird sowohl die univariate als auch der multivariate Fall behandelt. Außerdem wird in einem praktischen Teil die Hypothese der Normalverteiltheit der betrachteten Funktionale für wachsende Beobachtungsfenster $B^d_r$ mit Hilfe von Anpassungstests untersucht.

Kapitel 6 beschäftigt sich mit der Untersuchung von zufälligen stationären und ergodischen Mosaiken $\Psi = \{\Xi_n\}_{n \geq 1}$ im $\mathbb{R}^d$, wobei angenommen wird, dass eine Realisierung von $\Psi$ in einem konvexen und beschränkten Beobachtungsfenster $W^d_\phi \subset \mathbb{R}^d$ gegeben ist. Weiterhin wird angenommen, dass die Zellen $\Xi_n$ von $\Psi$ eine zufällige innere Struktur besitzen, zum Beispiel ein einbeschriebenes Punkt muster, Fasersysteme oder wiederum Mosaik. Dabei sollen diese inneren Strukturen unabhängig voneinander und unabhängig von $\Psi$ erzeugt werden und ein (generisches) zufälliges vektorielles Maß $J_0$ auf dem $\mathbb{R}^d$ erzeugen. Insbesondere untersuchen wir das asymptotische Verhalten eines multivariaten zufälligen Funktional $\text{d} \Psi$ (welches sich durch geeignetes Zentrieren und Normalisieren aus $J_0$ ergibt) falls $W^d_\phi$ unbeschränkt und gleichmäßig in all Richtungen wächst. Dabei ist das zufällige Funktional sowohl durch $\Psi$ als auch durch die individuellen Zellstrukturen in $W^d_\phi$ festgelegt. Es stellt sich heraus, dass dieses Funktional erwartungstreue Schätzer für den Intensitätsvektor von $J_0$ liefert. Außerdem werden starke Gesetze der großen Zahlen und ein multivariater zentraler Grenzwertsatz hergeleitet. Zum Schluss behandeln wir numerische Beispiele und untersuchen mit Hilfe von Anpassungstests die Hypothese der Normalverteiltheit der betrachteten Funktionale.

Bibliography


Bibliography


# List of Figures

1.1 Infrastructure system of Paris ........................................... 7
1.2 Nationwide telecommunication network modelling .......................... 8
1.3 Hierarchical structure of access networks .................................... 9
1.4 The Stochastic Subscriber Line Model ....................................... 10
1.5 Realizations of random tessellation models for the network’s geometry .... 11
1.6 Linear and spatial placement of network components ....................... 11
1.7 Types of network topologies ............................................... 12
1.8 Poisson line network and induced Cox–Voronoi serving zones .......... 12
1.9 Placement of lower level network components ............................ 13
1.10 Estimates for \(m(\beta)\) and \(m'(\beta)\); fitted curves \(m(\beta) = a\beta^b\) and \(m'(\beta) = a'\beta^{b'}\) .... 13

2.1 Realizations of random point processes ...................................... 22
2.2 Realization of a clustered point process and the theoretical pair correlation function .................................................. 25
2.3 Realizations of stationary random Poisson point processes ............... 26
2.4 Parameterization of a line \(H(p,u)\) in \(\mathbb{R}^2\); cf. also Section 4.1.3 ....... 32

3.1 Examples of non-regular, regular, and normal tessellations .......... 37
3.2 Poisson–Voronoi tessellations with different intensities \(\gamma_{\text{PVT}}\) ........ 41
3.3 Poisson–Delaunay tessellations with different intensities \(\gamma_{\text{PDT}}\) ........ 42
3.4 Poisson line tessellations with different intensities \(\gamma_{\text{PLT}}\) ............. 44
3.5 Superpositions of \(\Psi_0\) (intensity \(\gamma_0\)) and \(\Psi_1\) (intensity \(\gamma_1\)) ........ 46
3.6 \(\Psi_0/\Psi_1\)-nestings (intensities \(\gamma_0\) and \(\gamma_1\), respectively) .......... 46
3.7 $\Psi_0/p\Psi_1$-nestings (intensities $\gamma_0$ and $\gamma_1$, respectively) .................. 47
3.8 Multi type–nestings with initial tessellation $\Psi_0$ (intensity $\gamma_0$) .................. 47
3.9 Two–fold $\Psi_0/\Psi_1/\Psi_2$-nestings (intensities $\gamma_0$, $\gamma_1$, $\gamma_2$, respectively) ........ 48
4.1 Poisson lines hitting $B_d^2$ and their intersections .......................... 55
5.1 Histograms, quantile plots (q–plots), and empirical distribution functions for $\hat{Z}_{0,r}^{(2)}$ ($m$ realizations of $\Phi_1$ with $\lambda = 0.1$ in $B_d^2$) .................. 99
5.2 Comparison of bivariate histogram and density function .................. 103
5.3 Estimated power function of the test (5.22) .......................... 103
5.4 Histograms, quantile plots and empirical distribution functions for $\hat{Y}_{0,n}^{(2)}$ based on $m = 5000$ simulations of $\Psi$ with intensity $\gamma = 0.01$ in $W_{100}^2$ and $\bar{W}_{100}^2$ ........ 105
6.1 Cells of $\Psi$ and their inner structure induced by a PLT .................................. 107
6.2 Cells of a PLT and their inner structure induced by point processes ........ 108
6.3 PLT/PLT–nesting ........................................................................ 131
6.4 Histograms, quantile plots and empirical distribution functions for $\hat{Z}_{1,3}^{(2)} (J)$ and $n$ simulations .................................................. 134
6.5 Histograms, quantile plots and empirical distribution functions for $\hat{Z}_{2,3}^{(2)} (J)$ and $n$ simulations .................................................. 135
6.6 Comparison of bivariate histogram and density function .................. 137
7.1 Examples of data with network structure ............................................. 141
7.2 Sampling techniques leading to biased estimation .................................. 143
7.3 Sampling techniques leading to unbiased estimation ............................. 144
7.4 Plot of $w(\gamma)$ (PLT with $\hat{z} = (1.8, 0.03, 0.02, 0.53)\top$; Euclidean distance function) 148
7.5 Realizations of the considered tessellation models in Section 7.2.3 ........... 149
7.6 Discretization of $[\gamma_0^{\text{min}}, \gamma_0^{\text{max}}] \times [\gamma_1^{\text{min}}, \gamma_1^{\text{max}}]$ .................................. 152
7.7 Realization of PLT/PVT–nesting ($\gamma_0 = 0.1$, $\gamma_1 = 0.004$, $p = 1.0$) and estimated characteristics $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_4)\top$ .................................. 153
7.8 Realization of PDT/pPVT–nesting ($\gamma_0 = 0.001$, $\gamma_1 = 0.005$, $p = 0.9$) and estimated characteristics $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_4)\top$ .................................. 153
7.9 Realization of PLT/PLT–nesting ($\gamma_0 = 0.001$, $\gamma_1 = 0.015$, $p = 1.0$) and estimated characteristics $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_4)\top$ .................................. 155
7.10 Road system of Paris in a square of side length 3000 (meters) . . . . . . . . . . 160
7.11 Network data and best fitting tessellation model (cf. Table 7.15) . . . . . . 162
7.12 Network data in $W = [1500, 4500]^2$ and best fitting tessellation . . . . . 164
List of Tables

3.1 Mean value relationships for a $PVT/p \Psi_1$-nesting ................................. 49
3.2 Mean value relationships for a $PDT/p \Psi_1$-nesting ................................. 50
3.3 Mean value relationships for a $PLT/p \Psi_1$-nesting ................................. 50

5.1 Estimation of $\lambda_0$ by $\hat{\lambda}_{0,r}$ ($m$ realizations of $\Phi_1$ in $B_2^2$ with $\lambda = 0.1$; theoretical value $\lambda_0 = 0.003183$) ................................................................. 97
5.2 Estimation of $\lambda_1$ ($m$ realizations of $\Phi_1$ in $B_2^2$ with $\lambda = 0.1$; theoretical value $\lambda_1 = \lambda$) ................................................................. 98
5.3 Goodness-of-fit test evaluating $H_0^{(r)}$ for $m$ simulations of $\Phi_1$ in $B_2^2$; functional $\hat{Z}_{0,r}^{(2)}$. Rejection of $H_0^{(r)}$ is denoted by $\ast$. ......................... 98
5.4 Komogorov–Smirnov test evaluating $H_0^{(r)}$ for $m$ simulations of $\Phi_1$ in $B_2^2$ (test statistic $T_m$, $p$-value $p_m$). Rejection of $H_0^{(r)}$ is denoted by $\ast$. .......... 101
5.5 Bivariate goodness-of-fit test evaluating $\hat{H}_0^{(r)}$ (results based on 1000 realizations of $\Phi_1$ with $\lambda = 0.1$ in $B_2^2$; Henze–Zirkler’s test statistic $T_{1000,\beta}^{(r)}$ and 0.95-quantiles $q_{\beta}(0.95)$; rejection of $\hat{H}_0^{(r)}$ indicated by $\ast$). ......................... 102
5.6 Komogorov–Smirnov test based on $m$ simulations of $\hat{y}_{0,m}^{(2)}$ (test statistic $T_m$ and $p$-values $p_m$; rejection of $H_0^{(m)}$ is denoted by $\ast$) ......................... 104

6.1 Values of Pearson’s test statistic $T_k$ and $p$-values $p_k$ for $k$ simulations of the vector $\hat{Z}_o^{(2)}(J)$. Rejection of $H_{0,i}^\theta$, $i = 1, 2$ is denoted by $\ast$. ................. 136
6.2 Values of Kolmogorov–Smirnov’s test statistic $T_k'$ and $p$-values $p_k$ for $k$ simulations of the vector $\hat{Z}_o^{(2)}(J)$. Rejection of $H_{0,i}^\theta$, $i = 1, 2$ is denoted by $\ast$. ................. 136
6.3 Henze–Zirkler’s test statistic $T_{k,\beta}$ for $k$ simulations ................................. 138
7.1 Estimated characteristics $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_4)^\top$ for different tessellation models . . . 148
7.2  Model fitting results; input PVT ($\gamma_{\text{PVT}} = 0.001$)  
7.3  Model fitting results; input PDT ($\gamma_{\text{PDT}} = 0.001$)  
7.4  Model fitting results; input PLT ($\gamma_{\text{PLT}} = 0.1$)  
7.5  Fitting of a $\Psi_0/\Psi_1$-nesting to the data of Figure 7.7  
7.6  Fitting of a $\Psi_0/p\Psi_1$-nesting to the data of Figure 7.8  
7.7  MC test for PLT input ($\gamma_{\text{PLT}} = 0.1$), $\Psi(H_0) = \text{PLT}$ with $\gamma(H_0) = 0.1$  
7.8  MC test for PLT input ($\gamma_{\text{PLT}} = 0.1$), $\Psi(H_0) = \text{PLT}$ with $\gamma(H_0) = 0.095417$  
7.9  MC test for PVT input ($\gamma_{\text{PVT}} = 0.001$), $\Psi(H_0) = \text{PVT}$ with $\gamma(H_0) = 0.001$  
7.10 MC test for PVT input ($\gamma_{\text{PVT}} = 0.001$), $\Psi(H_0) = \text{PVT}$ with $\gamma(H_0) = 0.000921$  
7.11 MC test for PLT/PLT input ($\gamma_0 = 0.001$, $\gamma_1 = 0.015$), $\Psi(H_0) = \text{PLT/PLT}$  
7.12 MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PLT with $\gamma(H_0) = 0.0024$  
7.13 MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PVT with $\gamma(H_0) = 10^{-6}$  
7.14 MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PDT with $\gamma(H_0) = 10^{-6}$  
7.15 Distance values $d_{re,min}$ for data of Figure 7.11; fixed initial PLT; nested $\Psi_1$ (intensity $\gamma_1$)  
7.16 MC test for data of Figure 7.10; $\Psi(H_0)$ is a PLT/PDT with $\gamma_0(H_0) = 0.0024$ and $\gamma_1(H_0) = 0.000028$  
7.17 MC test for main roads in Figure 7.10; $\Psi(H_0)$ is a PLT/PLT with $\gamma_0(H_0) = 0.0024$ and $\gamma_1(H_0) = 0.013906$
Nomenclature

\((\Omega, \sigma(\Omega), \mathbb{P})\)  
\(B(\mathbb{R}^d)\)  
\(B(\mathbb{R}^d) \otimes B(D)\)  
\(B_0(\mathbb{R}^d)\)  
\(A_k^d\)  
\(B(A_k^d)\)  
\(B(D)\)  
\(B(\mathcal{F})\)  
\(B(\mathcal{F}')\)  
\(B(\mathcal{L}_k^d)\)  
\(C\)  
\(D\)  
\(\mathcal{F}\)  
\(\mathcal{G}\)  
\(\mathcal{H}_k\)  
\(\mathcal{K}\)  
\(\mathcal{L}_k^d\)  
\(\mathcal{M}(\mathcal{F}')\)  
\(\mathcal{M}(\mathbb{R}^d \times D))\)  
\(\mathcal{M}(\mathbb{R}^d)\)  

basic probability space, page 20  
Borel \(\sigma\)-algebra on \(\mathbb{R}^d\), page 166  
product \(\sigma\)-algebra on \(\mathbb{R}^d \times D\), page 27  
family of bounded Borel sets on \(\mathbb{R}^d\), page 166  
affine \(k\)-dimensional subspaces in \(\mathbb{R}^d\), \(k = 0, \ldots, d - 1\), page 31  
Borel \(\sigma\)-algebra on \(A_k^d\), \(k = 0, \ldots, d - 1\), page 31  
Borel \(\sigma\)-algebra on the mark space \(D\), page 27  
Borel \(\sigma\)-algebra on \(\mathcal{F}\), page 20  
Borel \(\sigma\)-algebra on \(\mathcal{F}'\), page 30  
Borel \(\sigma\)-algebra on \(\mathcal{L}_k^d\), page 31  
family of convex bodies in \(\mathbb{R}^d\), page 20  
mark space, page 27  
g family of closed sets in \(\mathbb{R}^d\), page 20  
g family of open sets in \(\mathbb{R}^d\), page 20  
Hausdorff measure on \(\mathbb{R}^d\) (of order \(k\)), page 167  
g family of compact sets in \(\mathbb{R}^d\), page 20  
Grassmann manifold (corresponding to \(A_k^d\), \(k = 0, \ldots, d - 1\), page 31  
\(\sigma\)-algebra on \(\mathcal{M}(\mathcal{F}')\), page 30  
\(\sigma\)-algebra of subsets of \(\mathcal{M}(\mathbb{R}^d \times D)\), page 27  
\(\sigma\)-algebra of subsets of \(\mathcal{M}(\mathbb{R}^d)\), page 22
**P**
family of \(d\)-polytopes (with non-empty interior), page 36

\(\mathcal{P}\)
family of \(d\)-polytopes with associated point at \(a\), page 38

\(S_k(p)\)
family of \(k\)-faces of \(p\), page 36

\(T\)
family of tessellations in \(\mathbb{R}^d\), page 36

\(\hat{B}\)
reflection of \(B\) (about the origin), page 166

\(\chi(\cdot)\)
set indicator function, page 58

\(\chi_{d,1-\alpha}\)
\((1-\alpha)\)-quantile of the \(\chi^2\)-distribution with \(d\) degrees of freedom, page 93

\(\text{cl}\, B\)
closure of \(B\), page 166

\(\text{Cov}\)
covariance matrix, page 109

\(\delta_x\)
Dirac measure, page 167

\(\emptyset\)
empty set in \(\mathbb{R}^d\), page 20

\(\Delta\)
equality in distribution, page 21

\(\frak{S}\)
second reduced moment measure, page 24

\(\Gamma(\cdot)\)
Gamma function, page 167

\(\mathbb{1}_S(\cdot)\)
indicator function with respect to \(S\), page 24

\(\text{int}\, B\)
interior of \(B\), page 166

\(\iota\)
permutation, page 58

\(\kappa_d\)
volume of the unit ball in \(\mathbb{R}^d\), page 167

\(\Lambda_D\)
intensity measure of \(X_D\), page 27

\(\lambda_k\)
intensity of \(\Phi_k\), \(k = 0, \ldots, d-1\), page 31

\(\Lambda_k(\cdot)\)
intensity measure of \(\Phi_k\), \(k = 0, \ldots, d-1\), page 31

\(\lambda_X(\cdot)\)
intensity of \(X\), page 23

\(\Lambda_X(\cdot)\)
intensity measure of \(X\), page 23

\(\Lambda_\Phi\)
intensity measure of \(\Phi\), page 30

\(\langle \cdot, \cdot \rangle\)
scalar product on \(\mathbb{R}^d\), page 165

\(M(\mathbb{R}^d)\)
family of simple and locally finite counting measures on \(\mathbb{R}^d\), page 22

\(G([0,\infty))\)
grid on the non-negative real line, page 147
\(\mu\) measure on \((\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\), page 166
\(\mu_c\) counting measure on \(\mathcal{B}(\mathbb{R}^d)\), page 167
\(N\) set of positive integers, page 165
\(N_0\) set of non-negative integers, page 165
\(\nu_d\) Lebesgue measure on \(\mathcal{B}(\mathbb{R}^d)\), page 167
\(\omega_d\) surface area of the unit ball in \(\mathbb{R}^d\), page 168
\(\ominus\) Minkowski subtraction, page 166
\(\oplus\) Minkowski addition, page 166
\(\partial B\) boundary of \(B\), page 166
\(\partial B_d(x)\) sphere centered at \(x\) with radius \(r\), page 166
\(\Phi\) point process of non-empty closed sets, page 30
\(\Phi_k\) \(k\)-flat process in \(\mathbb{R}^d\), \(k = 0, \ldots, d - 1\), page 31
\(\Pi(\tilde{\Theta})\) Steiner convex set/associated zonoid of \(\tilde{\Theta}\), page 33
\(p\) \(d\)-polytope, page 35
\(\preceq\) linear ordering (on \(S^{d-1}_+\)), page 57
\(\mathbb{R}\) set of real numbers on the line, page 165
\(\mathbb{R}^d\) \(d\)-dimensional Euclidean space, page 165
\(S^{d-1}\) sphere centered at \(o\) with radius \(r\), page 166
\(S^{d-1}_+\) upper unit hemisphere, page 166
\(\sum^*\) Sum over pairwise distinct indices, page 57
\(W^d_{\hat{e}}\) sampling window in \(\mathbb{R}^d\) (with scaling factor \(\varrho > 0\)), page 50
\(\tau\) deterministic tessellation in \(\mathbb{R}^d\), page 36
\(\Theta_k\) orientation distribution of \(\Phi_k\), \(k = 0, \ldots, d - 1\), page 31
\(\theta_R\) rotation (around the origin), page 166
\(\Upsilon_k(\tau)\) family of \(k\)-facets of \(\tau\), \(k = 0, \ldots, d\), page 36
\(\bar{\mu}_c\) simple and locally finite counting measure on \(\mathbb{R}^d \times \mathcal{D}\), page 27
\(\bar{\nu}_d(\cdot)\) spherical Lebesgue measure, page 168
\(\Psi^{(k)}\)  \(k\)-facet process induced by \(\Psi\), page 39

\(\Theta\)  spherical orientation distribution on \(\mathcal{B}(S_d)\), page 33

\(\Theta^*\)  spherical orientation distribution on \(\mathcal{B}(S_d)\), page 33

\(\Xi\)  random closed set (RACS), page 20

\(\Xi^*\)  typical cell of \(\Psi = \{\Xi_n\}_{n \geq 1}\), page 39

\(Z\)  set of integers, page 165

\(B^c\)  complement of \(B\), page 166

\(B^d_r(x)\)  closed ball centered at \(x\) with radius \(r\), page 166

\(B^d_r\)  closed ball centered at \(o\) with radius \(r\), page 166

\(C_{DeI}(x, B)\)  Delaunay cell with nucleus \(x\), page 42

\(C_{Vor}(x, B)\)  Voronoi cell with nucleus \(x\), page 40

\(D(B)\)  diameter of \(B\), page 168

\(F_X^{-1}\)  quantile function of the random variable \(X\), page 171

\(g_0(\cdot)\)  pair correlation function (of point process of nodes), page 78

\(h(C, u)\)  support function of \(C \in \mathcal{C}'\), page 33

\(H(p, u)\)  hyperplane in \(\mathbb{R}^d\), page 32

\(L^1(\Omega, \sigma(\Omega), \mu)\)  space of \(\mu\)-integrable functions, page 168

\(L^2(\Omega, \sigma(\Omega), \mu)\)  space of quadratic \(\mu\)-integrable functions, page 168

\(L^d_{\perp} \in \mathcal{L}_d^{d-k}\)  orthogonal complement of \(L \in \mathcal{L}_d^d\), page 31

\(M(\mathcal{F}')\)  family of simple and locally finite counting measures on \(\mathcal{B}(\mathcal{F}')\), page 30

\(M(\mathbb{R}^d \times \mathcal{D})\)  family of simple and locally finite counting measures on \(\mathbb{R}^d \times \mathcal{D}\), page 27

\(M_j\)  \(j\)th Minkowski functional, \(j = 0, \ldots, d\), page 182

\(P_D\)  Palm mark distribution of \(X_D\), page 27

\(P_\Xi\)  distribution of the RACS \(\Xi\), page 20

\(P_X\)  distribution of the random point process \(X\) in \(\mathbb{R}^d\), page 23

\(P_X^{pe}\)  Palm distribution of the random point process \(X\), page 24

\(P_{X_D}\)  distribution of the random marked point process \(X_D\) in \(\mathbb{R}^d\), page 27
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{\Xi}$</td>
<td>Choquet’s capacity functional, page 20</td>
</tr>
<tr>
<td>$V_j$</td>
<td>$j$th intrinsic volume, $j = 0, \ldots, d$, page 182</td>
</tr>
<tr>
<td>$W$</td>
<td>Sampling window in $\mathbb{R}^d$ ($W \in C$), page 50</td>
</tr>
<tr>
<td>$X$</td>
<td>Random point process in $\mathbb{R}^d$, page 22</td>
</tr>
<tr>
<td>$X_D$</td>
<td>Random marked point process in $\mathbb{R}^d$, page 27</td>
</tr>
</tbody>
</table>
Index

\( U \)-statistic, 56, 179
  - Hoeffding’s CLT, 180
  - Hoeffding’s decomposition, 179
  - kernel function, 179
  - projection, 179
\( \beta \)-mixing coefficient, 95
\( \sigma \)-algebra, 169

access network, 9
  - Network Interface Device, 9
  - primary cabinet, 9
  - secondary cabinet, 9
  - Service Area Interface, 9
  - serving zone, 9
  - Wire Center Station, 9
associated point, 38, 110
associated zonoid, 33, 77

ball
  - \( d \)-dimensional closed, 166
Bernoulli thinning, 45
  - parameter, 45
Borel \( \sigma \)-algebra, 166
Borel \( \sigma \)-algebra on \( \mathbb{R}^d \), 22
boundary of a set, 166
Buffon’s needle problem, 19

Campbell’s theorem, 52
  - for point processes of closed sets, 30
  - for random marked point processes, 28
  - for random point processes, 24
Cardano’s formula, 147
Cartesian product, 165
Cavalieri’s principle, 32
Choquet, 20
  - capacity functional, 20
  - theorem of, 20
  - family of, 20
  - random, 20
closure of a set, 166
compact sets
  - family of, 20
  - random, 20
complement of a set, 166
continuous mapping theorem, 93
convex averaging sequence, 28
convex bodies
  - centrally symmetric, 33
  - family of, 20
  - random, 20
convex geometry, 19
convex hull, 42
counting measure, 167
coverage process, 8
Cox–Voronoi tessellation, 12
Cramér–Wold device, 174
Crofton’s formula, 33, 75

Delaunay tessellation, 42
  - Delaunay cell, 42
diameter of a set, 168
digital subscriber line, 9
Dirac measure, 167
distance function
  - absolute value, 145
  - Euclidean, 145
  - maximum norm, 145
distribution
  - \( \chi^2 \), 172
  - exponentially bounded tail, 54
  - normal, 171
  - Poisson, 57
regular $d$–variate normal, 171
singular $d$–variate normal, 172
standard normal, 171, 182
distribution function
cumulative, 170
empirical, 170
dominated convergence theorem, 169
empty set in $\mathbb{R}^d$, 20
Euclidean norm, 166
Euclidean space, 165
2–dimensional, 165
3–dimensional, 165
d–dimensional, 165
affine subspace, 31
orthogonal vectors, 166
unit vector, 166
Euler’s Beta function, 81, 82
family of integrable functions, 29, 168
family of quadratic integrable functions, 168
flat process, 31
intensity, 31
isotropic, 31
orientation distribution, 31
stationary, 31
Fubini’s theorem, 52, 113, 169
function
Borel measurable, 168
Gamma, 167
integrable, 168
measurable, 58, 168
symmetric, 58
Gamma function, 167
generalized inverse function, 171
Glivenko–Cantelli’s theorem, 171
goodness–of–fit test, 180
Henze and Zirkler, 97
Kolmogorov–Smirnov, 180
Pearson, 180
Pearson–Fisher, 180
Grassmann manifold, 31
Haar’s lemma, 23, 167
halfspace
closed, 40
Hausdorff measure, 167
hitting set, 36
Hoeffding’s CLT, 180
Hoeffding’s decomposition, 179
hyperplane, 32
orientation vector, 32
parameterization, 32
signed perpendicular distance, 32
hyperplane process, 32
associated zonoid, 33
Poisson, 33
spherical orientation distribution, 33
Steiner convex set, 33
image analysis, 19
indicator function, 24
individual ergodic theorem, 29, 117
inequality
Čebyschev, 172
Berry–Esseen, 177
Cauchy–Schwarz, 172
Esseen, 177
Hájek–Rényi, 173
isodiametric, 182
Jensen, 172
integral geometry, 19
interior of a set, 166
intrinsic volume, 33, 182
iterated random tessellation, 45
$k$–fold nesting, 48
Bernoulli thinning, 45
component tessellation, 45
initial tessellation, 45
isotropic, 45
mean value relationships, 48
multi–type nesting, 47
nesting, 45
nesting with Bernoulli thinning, 47
stationary, 45
superposition, 45
iterated tessellation, 44
component tessellation, 44
initial tessellation, 44
kernel density estimator, 181
  bandwidth, 181
Lebesgue measure, 167
Lebesgue’s theorem, 169
Legendre’s duplication formula, 76, 87
local loop, 9
long range dependence, 56, 59
mapping
  measurable, 168
matrix
  asymptotic covariance matrix, 109
  covariance, 171
  covariance matrix, 109
  diagonal, 92
  identity, 93, 172
  orthogonal, 93, 166
mean shortest path length, 12
mean subscriber line length, 12
mean value theorem, 129
measurable mapping, 168
measurable space, 166
  product space, 169
measure, 166
  atom, 166
  counting, 167
  diffuse, 166
  Dirac, 167
  Hausdorff, 167
  Lebesgue, 167
  locally finite, 166
  product, 169
  rotation invariant, 167
  simple counting, 167
  spherical Lebesgue, 168
  translation invariant, 166
measure space, 168
  complete, 169
method of moments, 56
Minkowski addition, 166
Minkowski functional, 182
Monte Carlo simulation, 142
  Monte Carlo test, 156
    significance level, 156
    multi-cast network, 8
Nelder–Mead algorithm, 152
non-degenerate, 43
null set, 169
open sets
  family of, 20
ordering
  lexicographic, 66
  linear, 57
origin, 165
p-value, 100
pair correlation function, 78
Palm distribution, 24
Palm mark distribution, 38
Palm probability, 24
particle centroid, 38
permutation, 58, 178
Poisson hyperplane process, 33
  non-degenerate, 34
Poisson hyperplane tessellation, 43
Poisson point process, 25
Poisson–Delaunay tessellation, 42
Poisson–Voronoi tessellation, 40
Polish space, 27
polytope
  $d$-dimensional, 35
  $d$-polytope, 35
  $k$-dimensional (in $\mathbb{R}^d$), 36
  facet, 36
  vertex, 36
probability measure, 169
  even, 33
  symmetric, 33
probability space, 20, 169
product measure, 169
product space, 169
quantile function, 171
quantile plot, 180
quermassintegrale, 182

RACS, 20
random closed set, 20
  isotropic, 21
  stationary, 21
random compact set, 20
random convex body, 20, 45
random marked point process, 27
  distribution, 27
  ergodic, 29
  independently marked, 27
  intensity, 27
  intensity measure, 27
  isotropic, 27
  mark space, 27
  mixing, 29
  Palm mark distribution, 27
  stationary, 27
random particle process, 37
random point process, 22
  Campbell’s theorem, 24
  canonical representation, 22
  in $\mathbb{R}^d$, 22
  intensity, 23
  intensity measure, 23
  isotropic, 23
  marked, 27
  of closed sets, 30
  pair correlation function, 24
  Palm distribution, 24
  Poisson, 25
  Ripley’s $K$–function, 24
  second reduced moment measure, 24
  stationary, 23
random point process of closed sets, 30
  intensity measure, 30
  isotropic, 30
  stationary, 30
random tessellation, 37
  $k$–facet process, 39
  associated point, 38
  centered cell, 38
  centroid, 38
  isotropic, 38
  iterated, 45
  Johnson–Mehl, 40
  Laguerre, 40
  point process of $k$–facets, 39
  point process of vertices, 39
  Poisson hyperplane, 40, 43
  Poisson line, 43
  Poisson–Delaunay, 40, 42
  Poisson–Voronoi, 40
  stationary, 38
  typical cell, 39
random variable
  $k$th moment, 171
  almost surely convergence, 173
  convergence in distribution, 173
  convergence in mean, 173
  convergence in probability, 173
  convergence in quadratic mean, 173
  distribution, 170
  distribution function, 170
  expectation, 171
  identically distributed, 170
  independent, 170
  mean, 171
  variance, 171
  weak convergence, 173
random vector, 169
sample
  mean, 179
  variance, 179
sample covariance matrix, 181
sample mean, 181
sample profile, 143
sampling technique, 143
  associated point rule, 144
  minus sampling, 144
  plus sampling, 144
  tiling rule, 144
scalar multiplication, 166
scalar product, 165
Service Area Interface, 9
service zone, 9
set operation, 166
  Minkowski addition, 166
  Minkowski subtraction, 166
  multiplication, 166
  reflection, 166
  rotation, 166
  translation, 166
significance level, 98, 181
Slepnyak's theorem, 26
Slutsky's theorem, 174
spanning tree, 8
spatial statistics, 19
spatial stochastic model, 8, 10
sphere, 166
  unit, 166
spherical Lebesgue measure, 168
spherical orientation distribution, 33, 43
  non-degenerate, 34
Steiner convex set, 33, 43
Steiner's formula, 116, 182
stereology, 19
Stirling number
  first kind, 62
  second kind, 62
stochastic geometry, 19
Stochastic Subscriber Line Model, 7, 10
  Network Component Model, 11
  Network Geometry Model, 10
  Network Topology Model, 11
subscriber, 9
support function, 33
supremum norm, 175
switching network, 8
symmetric function, 178
symmetric set, 166

Telecommunication network, 9
tessellation, 36
  $k$-facet, 37
  aggregate, 8
  cell, 36
  component, 44
  Cox–Voronoi, 12
  Delaunay, 42
deterministic, 36
face-to-face, 37
family of, 36
initial, 44
iterated, 44
normal, 37
ordinary, 37
random, 37
regular, 37
Voronoi, 40
test
  power function, 103
unit sphere, 166
upper unit hemisphere, 57, 166
variance stabilizing transformation, 178
vector space, 166
  addition, 166
  scalar multiplication, 166
Voronoi tessellation, 40
  nucleus, 40
  Voronoi cell, 40

Wire Center Station, 9
Erklärung


Ulm, 11. Juli 2006