PhD Dissertation

Stereology and Spatio-Temporal Models

Numerical Integration Methods for Volume Estimation and Extremes for Lévy-based Models

Mads Stehr

Department of Mathematics
Aarhus University
2020

Stereology and spatio-temporal models
Numerical integration methods for volume estimation and extremes for Lévy-based models

PhD dissertation by
Mads Stehr

Department of Mathematics, Aarhus University
Ny Munkegade 118, 8000 Aarhus C, Denmark

Supervised by
Associate Professor Markus Kiderlen, Aarhus University, main supervisor
Associate Professor Anders Rønn-Nielsen, Copenhagen Business School, co-supervisor

Submitted to Graduate School of Natural Sciences, Aarhus, July 20, 2020

Dissertation was typesat in kpfonts with
pdfLATEX and the memoir class

AARHUS
UNIVERSITY
DEPARTMENT OF MATHEMATICS

Contents

Preface

This dissertation concludes my PhD studies at the Department of Mathematics, Aarhus University, carried out from February 2017 to July 2020 under the supervision of Associate Professor Markus Kiderlen (main supervisor) and Associate Professor Anders Rønn-Nielsen (co-supervisor). My studies were funded jointly by the Graduate School of Natural Sciences (GSNS formerly known as GSST) and the Centre for Stochastic Geometry and Advanced Bioimaging (CSGB) through a grant from the Villum Foundation.

The dissertation consists of the following four self-contained papers in addition to a note to Paper A:

Paper A Asymptotic variance of Newton-Cotes quadratures based on randomized sampling points. To appear in *Advances in Applied Probability*.

Paper B Improving the Cavalieri estimator under non-equidistant sampling and dropouts. Submitted to *Image Analysis and Stereology*.

Paper C Tail asymptotics of an infinitely divisible space-time model with convolution equivalent Lévy measure. To appear in *Journal of Applied Probability*.

Paper D Extreme value theory for spatial random fields – with application to a Lévy-driven field. Submitted to *Extremes*.

Besides layout and minor insignificant adjustments, all papers correspond to their submitted version. However, in contrast to their appearance in this dissertation, Papers A and C will be published without supplementary materials, but with these available online. Furthermore, to ease reading, references between papers will be with respect to the enumeration of the dissertation and not the enumeration of the submitted papers.

An initial draft of Paper A was included in my progress report used for the qualification examination in June 2018. The paper has however improved significantly since then. Papers B–D are the result of the last two years of my studies and have thus not appeared in the progress report in any form. I have contributed extensively in both the research phase and the writing of all papers A–D with the exception of (Section B.5, Paper B), which is the work of my supervisor Markus Kiderlen.

The first chapter is an introduction to the topics of the dissertation. The purpose of the introduction is twofold. Firstly, it provides the reader with some basic theory, which is beneficial when reading the papers of the dissertation. Secondly, it gives an overview of results from the literature and relates these to the results of the papers.

I owe my supervisors Markus and Anders a dept of thanks for giving me the opportunity to do a PhD, for the many enriching discussions, for bearing with my

– at times – stupid questions, and for the general interest in my well-being. Due to their help and guidance, both professionally and personally, these last three and a half years have been an absolute pleasure. I am deeply grateful to Markus and Eva B. Vedel for wanting me as a PhD student in such a stimulating and interdisciplinary research environment that is CSGB.

I would also like to thank Professor Gennady Samorodnitsky from Cornell University. In addition to being a great host when I visited him and his department in Ithaca (NY), his helpful comments and suggestions concerning Paper D are greatly appreciated. Furthermore, I am grateful to all colleagues of Aarhus University and CSGB for creating such a nice working environment, and to the local LaTeXpert Lars Madsen for the immense help in setting up this dissertation. In particular I thank Julie Thøgersen, Mikkel Slot Nielsen, Thorbjørn Grønbæk, Victor Rohde, Mathias Ljungdahl, Louis Gammelgaard Jensen, Stine Hasselholt, Mathias Højgaard Jensen, and, last but not least, my office mate through three years Rikke Krog Eriksen for laughs, beers, housing and discussions.

Finally, my family and friends, and especially my girlfriend Sara, deserve enormous gratitude for their amazing support and willingness to listen to my mathematical blabbering. Most importantly, I thank you all for simply being the best anyone could ask for – and for accompanying me on Roskilde Festival every year!

Mads Stehr
Copenhagen, July 2020

Summary

Using the so-called Cavalieri estimator, the volume of a bounded 3-dimensional object can be unbiasedly estimated from area measurements on parallel section profiles. Mathematically, this corresponds to the problem of numerical integration on $\mathbb{R}$ when the integrand is known at a the points of a stationary and equidistant point process. Previously it has been shown that the variance of the Cavalieri estimator, which is simply a Riemann sum approximation, exhibits a slower decrease rate when the sampling nodes are not equidistant.

The first part of the dissertation concerns an alternative numerical estimation approach suitable for stationary and non-equidistant sampling. By using the information on the increments of the sampling points, we suggest using Newton-Cotes quadrature rules to approximate the integrand by a polynomial of a certain degree. It turns out that the variance inflation of the Cavalieri estimator under non-equidistant sampling can be avoided by this estimation approach. We investigate its variance properties under different sampling models, and we also suggest estimates of its variance based on sampling in a bounded interval in $\mathbb{R}$.

The second part of the dissertation mainly concerns spatial Lévy-driven moving average fields, which are integrals of a deterministic kernel function with respect to a Lévy basis. Throughout the dissertation we assume that the underlying Lévy measure is convolution equivalent, and we investigate the extremal behavior of the fields relative to that of the Lévy measure.

Firstly, we consider a space-time model, where the field is thought of as having a space and time component. Under reasonable assumptions on the Lévy measure and the integration kernel, we show that the tail of certain functionals (applied to the field) asymptotically equals the tail of the Lévy measure. As a complementary result we obtain that the field is continuous in the space component and càdlàg in time.

Secondly, we generalize classical extreme value results in $\mathbb{R}$ to higher-dimensional Euclidean space. In particular, assuming mixing conditions for a stationary random field, we show that an extremal types theorem holds for its normalized maximum on a large class of expanding sets. Furthermore, using similar proof techniques, we show that the normalized supremum of the Lévy-driven field converges to the Gumbel distribution as the index sets increase.

Resumé

Ved at bruge den såkaldte Cavalieri-estimator kan man estimere rumfanget af et begrænset, 3-dimensionelt objekt ud fra arealmålinger i parallele snitflader. Matematisk set svarer dette til et spørgsmål om numerisk integration, hvor integranden er kendt i punkterne fra en stationær og ækvidistant punktproces. Det er tidligere vist, at variansen af Cavalieri-estimatoren – som i øvrigt er en Riemann sum approksimation – aftager langsommere, når positionerne, der udgør punktprocessen, ikke er ækvidistante.

Den første del af afhandlingen omhandler en alternativ estimationsprocedure, der er tilpasset situationen med ikke-ækvidistante positioner. Ved at bruge deres indbyrdes afstande foreslår vi at anvende såkaldte Newton-Cotes kvadraturregler til at approksimere integranden med et polynomium af en passende grad. Det viser sig, at den forøgede varians af Cavalieri estimatoren kan undgås med denne estimationsprocedure. Vi undersøger Newton-Cotes-estimatorens variansegenskaber under to realistiske punktprocesmodeller, og vi foreslår variansestimater baseret på positioner i et begrænset interval.

Afhandlingens anden del er koncentreret om rumlige Lévy-drevne felter, der er givet som integralet af en deterministisk kernefunktion med hensyn til en Lévy basis. I afhandlingen antager vi, at det underliggende Lévy mål er foldningsækvivalent, og vi undersøger ekstremhændelser for felterne relativt til ekstremhændelser for deres Lévy mål.

Først undersøger vi et stokastisk felt, der både har en rum- og tidskomponent. Under rimelige antagelser om Lévy målet og kernefunktionen viser vi, at halen af visse funktionaler (anvendt på feltet) er asymptotisk ækvivalent med halen af Lévy målet. Vi viser desuden, at feltet er kontinuert i rumkomponenten og càdlàg i tid.

Efterfølgende generaliserer vi klassiske endimensionelle ekstremværdi resultater til højdimensionelle rum. Under visse uafhængighedslignende antagelser viser vi, at det normaliserede maximum af et stationært felt konvergerer mod en ekstremværdifordeling, når indeksmængden vokser passende. Ved at bruge lignende bevisteknikker viser vi tilsvarende, at det normaliserede supremum af et Lévy-drevet felt konvergerer mod Gumbel-fordelingen.

Introduction

This chapter serves as an introduction to the topics and main results of the dissertation. Seen as a whole, the dissertation lies within the area of spatial statistics, however it constitutes two separate and unrelated parts: Volume estimation in stereology and extremal probabilities for Lévy-based random fields. In Section 1 we present results from the literature on volume estimation and (more precisely) numerical integration in general, which in particular include results on the so-called (classical) Cavalieri estimator. This naturally leads to a presentation of the main findings of Papers A and B along with a conclusion of their applications. In Section 2 we give a short introduction to Lévy-based random fields, including the definition of infinitely divisible distributions and Lévy bases, before presenting results from the literature on tail asymptotics of such fields. After a description of the main results of Paper C, we introduce basic concepts and results from classical extreme value theory, relate these to results for Lévy-based fields, and conclude with a presentation of Paper D.

1 Numerical integration with application to volume estimation

Systematic sampling is a widely used technique in classical stereology to estimate for instance geometric quantities of biological tissues with an inhomogeneous structure; see [2] and the references therein. In particular, uniform systematic sampling on $\mathbb{R}$ can be used to unbiasedly estimate the volume of a bounded object $Y \subseteq \mathbb{R}^3$ through area measurements on parallel sections normal to some convenient sampling axis: Let the unit vector $v \in S^2$ be fixed and define $f(x) = \text{area}(Y \cap H_x)$ as the area of the intersection of Y with the plane $H_x = \{z \in \mathbb{R}^3 : z \cdot v = x\}$, which is normal to v and positioned at a signed distance x from the origin along v. For instance, if $Y = \{z \in \mathbb{R}^3 : |z| \leq 1\}$ is the unit ball in $\mathbb{R}^3$ centered at the origin and $v = (1,0,0)$, $f(x)$ is the area of the circle obtained as the intersection of the ball and the plane $H_x = \{(x, y_1, y_2) \in \mathbb{R}^3 : y_1, y_2 \in \mathbb{R}\}$, and hence

$$f(x) = \pi(1 - x^2)\mathbf{1}_{[-1,1]}(x) \tag{1.1}$$

in this case. We will return to this example later in the introduction. With f defined as above for general Y,

$$\Lambda = \int_{\mathbb{R}} f(x)\mathrm{d}x \tag{1.2}$$

coincides with the volume of $Y \subseteq \mathbb{R}^3$, and, if U is a uniform random variable on the interval $(0,1)$, it can be unbiasedly estimated using

$$\hat{\Lambda} = t \sum_{k \in \mathbb{Z}} f(tU + tk) \tag{1.3}$$

with the (section-)spacing $t > 0$. This estimator is commonly known as the Cavalieri estimator, and it serves as an important application of the estimation of quantities given as (1.2) for some integrable function $f : \mathbb{R} \to \mathbb{R}$, which is usually referred to as the measurement function.

The first part of this dissertation is devoted to the estimation of $\Lambda = \int_{\mathbb{R}} f(x)\mathrm{d}x$ and, to be more accurate, the precision of such estimators. The variance $\mathbb{V}\mathrm{ar}\hat{\Lambda}$ of $\hat{\Lambda}$ given by (1.3) is not trivially derived due to the correlation between observations, i.e $f(tU + tk)$ for varying k. In the classical work [21] and the subsequent adaption to stereology in e.g. [6], the transitive theory of Matheron was used to give a representation of the variance

$$\mathbb{V}\mathrm{ar}\hat{\Lambda} = t\sum_{k\in\mathbb{Z}} g(kt) - \int_{\mathbb{R}} g(y)\mathrm{d}y \tag{1.4}$$

based on the Euler-MacLaurin formula. Above, $g : y \mapsto \int_{\mathbb{R}} f(x+y)f(x)\mathrm{d}x$ is the so-called (geometric) covariogram of f, which when evaluated at $y \in \mathbb{R}$ intuitively describes the correlation between observations a distance y apart. In [16] the expression (1.4) is refined to a variance representation involving jumps and jump-locations of derivatives of f and consequently of g; see also [17]. It is assumed that f is (m,p)-piecewise smooth, meaning that f is C^{m-1}, compactly supported and that the kth derivative $f^{(k)}$ for $k = m,\dots,m+p$ only has finitely many discontinuities with finite associated jumps (see Sections A.1 and A.2 of Paper A for more details). If $f^{(m)}$ in fact is discontinuous, it is shown that

$$\mathbb{V}\mathrm{ar}\hat{\Lambda} = \mathbb{V}\mathrm{ar}_E\hat{\Lambda} + Z(t) + r(t), \tag{1.5}$$

where the extension term $\mathbb{V}\mathrm{ar}_E\hat{\Lambda}$ of order t^{2m+2} explains the overall behavior of the variance, the Zitterbewegung $Z(t)$ of order $O(t^{2m+2})$ fluctuates around 0, and the remainder $r(t)$ is of order $o(t^{2m+2})$ as $t \downarrow 0$. All three terms in (1.5) depend on the derivatives $f^{(m)}$, $f^{(m+1)}$ and their jumps; see Figure 1 for a depiction of the variance.

There are multiple extensions to the classical formulation just described. One obvious example is to allow for measurement functions in (1.3) which are not (m,p)-piecewise smooth for instance due to infinite jumps of the mth derivative. In [11, 13] such functions are considered, where the smoothness condition m-smooth is defined for all real $m \geq 0$ with (m,p)-piecewise smoothness as a special case. The theory makes use of fractional calculus, and, using a refined version of the Euler-MacLaurin formula, it is shown that the trend of the variance is of order t^{2m+2} (typically with $0 < m < 1$ in practice).

Another extension to the classical case is to allow for non-equidistant sampling nodes. Replacing $\{tU + tk\}_{k\in\mathbb{Z}}$ in (1.3) with a stationary point process $X \subseteq \mathbb{R}$ with intensity (expected number of points per unit interval) $1/t$, the unbiased estimator $\hat{\Lambda}$ is commonly known as the generalized Cavalieri estimator; see [1]. In [29, 30] the estimator is considered under some, practically realistic, sampling models inspired by studies on the structure of the brain; see e.g. [7] and Paper A for definitions of the models. It turns out that the precision of $\hat{\Lambda}$ based on non-equidistant sampling is inferior to that of equidistant sampling, as the variance is shown to decrease at a slower rate. For instance, if the sampling process X is derived from the equidistant process $\{tU + tk\}_{k\in\mathbb{Z}}$ by having small independent and non-degenerate perturbations of each point, the variance $\mathbb{V}\mathrm{ar}\hat{\Lambda}$ is of order t^3 for all (m,p)-piecewise smooth functions f

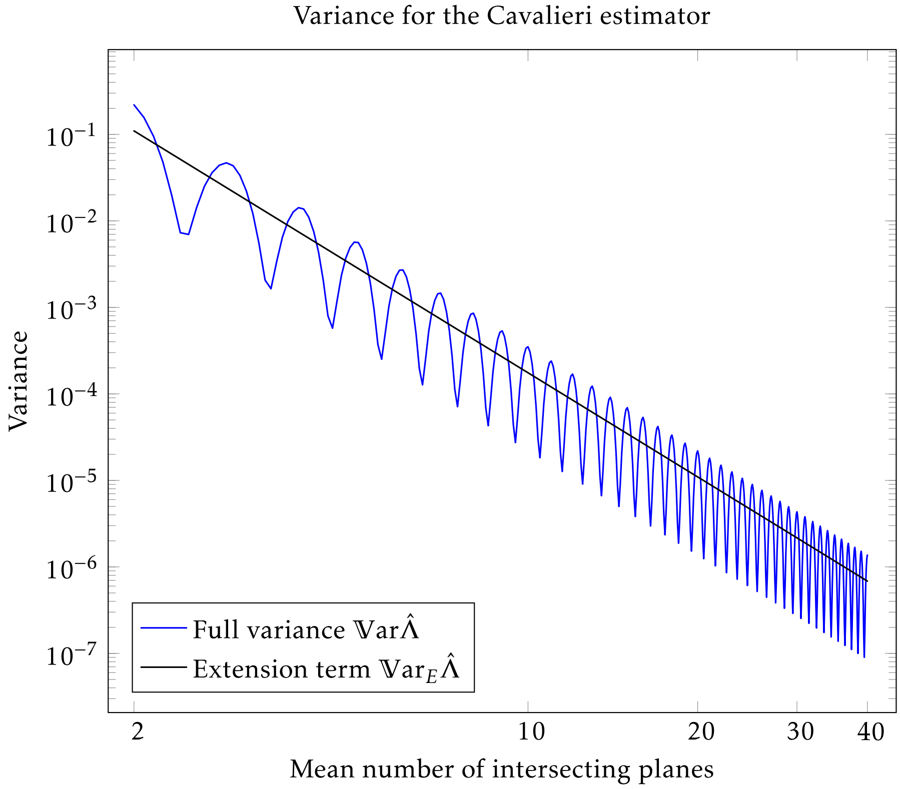


Figure 1: Variance of the classical Cavalieri estimator $\hat{\Lambda}$ for the volume estimation of the unit ball in $\mathbb{R}^3$ on a log−log scale. The measurement function is given by (1.1), and it is $(1,\infty)$-piecewise smooth. Furthermore, $\mathbb{V}\mathrm{ar}_E\hat{\Lambda} = t^4\pi^2/90$; see e.g. [16, Exercise 9].

with $m \geq 1$. This should be compared with the order $t^{2m+2} = o(t^3)$ as $t \downarrow 0$ under equidistant sampling. Moreover, the order of the variance depends on the underlying sampling model, as, for instance, sampling based on a renewal process on $\mathbb{R}$ yields an order of t for all $m \geq 0$. Throughout this introduction, the two mentioned sampling models will be referred to as the perturbed and cumulative models, respectively.

In addition to finding an asymptotic variance representation, it is naturally desirable to estimate the variance from observed data. Due to the decomposition (1.5), a widely used approach is simply to estimate the extension term $\mathbb{V}\mathrm{ar}_E\hat{\Lambda}$ as it may constitute a good approximation of the variance. In addition to depending on increments of the underlying sampling points, the extension term $\mathbb{V}\mathrm{ar}_E\hat{\Lambda}$ only depends on jumps of $f^{(m)}$ and not on their (relative) locations, making this estimation approach relatively easy when m is known; see e.g [2],[16] and [17], and [30] for non-equidistant sampling. When the smoothness constant m is unknown a similar estimation technique is described in [16] when the sampling points are equidistant. Due to the oscillation of the Zitterbewegung and the underlying structure of derivatives of f, it can be argued that an estimation of the variance based solely on the extension term can underestimate the variance substantially for a certain sample, in which case a better estimation can be obtained by taking some terms of the Zitterbewegung into account; see e.g. [12] and (Section B.5, Chapter B). Related to the question of variance estimation for the Cavalieri estimator, the paper [10] addresses the construction of confidence intervals for $\int_{\mathbb{R}} f\,\mathrm{d}x$ based on an equidistant sample, and it is among other results concluded that an asymptotic distribution of a standardized version of (1.3) exists when $f^{(m)}$ for $m \in \mathbb{N}$ has exactly one discontinuity.

Paper A

Motivated by the fact that $\mathbb{V}\mathrm{ar}\hat{\Lambda}$ has a substantially worse precision when $\hat{\Lambda}$ is based on non-equidistant sampling points, this paper investigates an alternative estimation procedure of $\int_{\mathbb{R}} f\,\mathrm{d}x$ adapted to such a scenario. Assuming that the measurement function f is measured at the points of a stationary point process $X \subseteq \mathbb{R}$, we suggest using Newton-Cotes quadrature rules (of sufficiently high order) to better suit the smoothness of f. The construction, which was initially suggested in [15], is as follows: Let $X = \{x_k\}_{k\in\mathbb{Z}}$ and let $n \in \mathbb{N}$ be fixed. On all intervals $[x_k, x_{k+n}]$, $k \in \mathbb{Z}$, the function f is approximated by a piecewise polynomial of degree at most n passing through the points $\{x_j, f(x_j)\}_{j=k}^{k+n}$. The nth order Newton-Cotes estimator $\hat{V}_n(f)$ is then defined as the sum of integrals of such approximating polynomials, averaged with respect to the chosen starting point.

In contrast to the generalized Cavalieri estimator, it turns out that the Newton-Cotes estimator exploits the smoothness of the integrand f to a degree similar as for equidistant sampling, as long as the order n of the estimator is sufficiently high relative to the smoothness of f. The theorem below constitutes (some of) the main findings of the paper. It is formulated in terms of $(m,1)$-piecewise smooth functions, but the result holds true for all (m,p)-piecewise smooth functions with $p \geq 1$.

Theorem (Theorems A.2.2 and A.2.8, Paper A). *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process with intensity $1/t > 0$ such that its typical point-increment has finite absolute moments of any order. Then $\hat{V}_n(f)$ is unbiased for $\int_{\mathbb{R}} f(x)\mathrm{d}x$ for all integrable, real-valued and compactly supported functions f. If f is $(m,1)$-piecewise smooth with $m \leq n$, the variance of $\hat{V}_n(f)$ obeys*

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq ct^{2m+2}$$

for some constant c, which does not depend on t.

If $m > n$ the theorem above only provides an order of $O(t^{2n+2})$ for the variance (this is due to the fact than an $(m+1,1)$-piecewise smooth function in particular is $(m,1)$-piecewise smooth), however, by introducing the covariance concept *strong admissibility*, we have shown the following additional result: If X is strongly admissible and $m > n$, then $\mathbb{V}\mathrm{ar}\hat{V}_n(f) = O(t^{2n+3})$ as $t \downarrow 0$. We show that both the perturbed and cumulative models are strongly admissible, and we also characterize the variance of the first-order Newton-Cotes estimator $\hat{V}_1(f)$ under these models explicitly.

In proving the theorem above we show a Peano kernel representation of the error of $\hat{V}_n(f)$ for piecewise smooth functions f. More precisely, let f be $(m,1)$-piecewise smooth, $D_{f^{(m)}}$ denote the set of jump-locations of the mth derivative $f^{(m)}$ of f, and let $J_{f^{(m)}}$ be the associated jumps; see (Section A.2, Paper A). Then, for any n, $m \leq n$ and realization X, there is function K_m such that

$$\hat{V}_n(f) - \int_{\mathbb{R}} f(x)\mathrm{d}x = \int_{\mathbb{R}} f^{(m+1)}(x)K_m(x)\mathrm{d}x + \sum_{a\in D_{f^{(m)}}} J_{f^{(m)}}(a)K_m(a). \tag{1.6}$$

The *Peano kernel* K_m is a piecewise polynomial of order at most $m+1$ with coefficients determined by X, and most importantly it does not depend on the function f. When X and thus K_m are considered random, the stationarity of X turns K_m into a stationary stochastic process. In fact, the admissibility property of X mentioned above is given by covariance conditions on its associated Peano kernel K_m.

Paper B

As a natural continuation of the work on Newton-Cotes estimators, this paper deals with the following topics: dropouts of the underlying point process, the construction of variance estimators, and the application to volume estimation of convex bodies. An observation showed in the appendix of the paper is the fact that the requirements on the measurement functions considered in Paper A can be relaxed a bit. We show that the results of Paper A hold under the assumption of *weakly* $(m,1)$-piecewise smooth functions, meaning (compared to $(m,1)$-piecewise smoothness) that the $(m+1)$st derivative is allowed to have infinite jumps.

In [29, 30] the generalized Cavalieri estimator based on sampling with dropouts was investigated. Dropouts are eliminated points of the point process modeled by independent thinning. It is among other things shown that perturbed sampling with dropouts results in an even worse decrease rate than perturbed sampling alone. One of the strengths of the Newton-Cotes estimators is the fact that they function with equal precision for all stationary point processes if $m \leq n$. Since a stationary process with dropouts is again stationary, the problem caused by the generalized Cavalieri estimator for sampling with dropouts does not apply to the Newton-Cotes estimators. In Paper B we give explicit expressions for the extension term $\mathbb{V}\mathrm{ar}_E \hat{V}_1(f)$ of the first-order Newton-Cotes estimator based on the perturbed and cumulative models with dropouts. The expression depends on model-specific parameters and on the thinning probability. Moreover, we show that these models are strongly admissible when combined with independent thinning, and thus the conclusions of the theorem above apply to these models as well.

In Paper A it is concluded that the variance of $\hat{V}_n(f)$ decomposes similarly as (1.5). We therefore approximate the variance by the extension term $\mathbb{V}\mathrm{ar}_E \hat{V}_n(f)$, which is a product of factors depending on the jumps of $f^{(m)}$ and on the point process, respectively. Based on information of the point process X in a bounded interval containing the support of f, we propose an estimation procedure for $\mathbb{V}\mathrm{ar}_E \hat{V}_n(f)$ with a bias of (approximately) order $o(1)$ as $t \downarrow 0$; see (Proposition B.4.1, Paper B). The strength of this estimation procedure is its generality, as it applies to any stationary point process. If the underlying sampling model is known, and in particular if it arises by dropouts from a known initial model with accessible increments, initial simulations indicate that a model-specific estimation may be preferred as it appears to have a smaller variance and to be more robust for varying intensity.

We apply our results to the stereological problem of estimating the volume of a strict convex body $Y \subseteq \mathbb{R}^3$ (compact, strictly convex set with non-empty interior) with sufficiently smooth boundary from parallel section profiles. For a fixed sampling axis, we show that the variance of $\hat{V}_n$ is given in terms of principal curvatures of Y. Moreover, we consider the situation where the direction of the sampling axis is uniformly randomized and show that the variance of $\hat{V}_n$ is then essentially proportional to the surface area of Y. This is well-known for the classical Cavalieri estimator (1.3) based on equidistant points (see [2, 6, 21]), but exact conditions on the convex body for this statement to hold have not been specified before.

Applications

As described from a theoretical perspective in Paper B, the Newton-Cotes estimators can be used to construct a superior estimate for the volume of a bounded object $Y \subseteq \mathbb{R}^3$, as compared to using the generalized Cavalieri estimator. However, in contrast to previous procedures, it requires information on the increments of the underlying sampling points. From a practical point of view, the object at hand is cut into slabs (sections), and the estimation procedure thus requires that the thickness of these slabs is accessible. Hence, if the slabs are measured and are of uneven thickness, the theory really finds its justification in both volume and associated variance estimation. If furthermore the (approximate) model of the location of the slabs is known, the Newton-Cotes theory can be used to achieve improved variance estimates (as indicated in Paper B).

The results also contribute to the theory of Monte Carlo integration on the real line $\mathbb{R}$. In particular, by following the proof in Paper A, the error representation (1.6) can be adopted to a wide range of approximation schemes by changing the kernel K_m appropriately.

2 Extremes for Lévy-based random fields

The second part of the dissertation concerns extremal probabilities of Lévy-based random fields, which are collections of random variables given as integrals with respect to so-called Lévy bases. Although the majority of Paper D deals with such fields it also extends to more general random objects as it shows classical extreme value results in a high-dimensional setting.

2.1 Infinitely divisible distributions and Lévy bases

A random variable Θ is said to be infinitely divisible if for all $n \in \mathbb{N}$ there are i.i.d. variables $\Theta_1, \ldots, \Theta_n$ such that $\sum_{i=1}^{n} \Theta_i$ equals Θ in distribution. Another very convenient characterization of infinitely divisible distributions is by the famous Lévy-Khintchine representation (see [27] for a detailed description of infinitely divisible distributions and processes): A random variable Θ is infinitely divisible if and only if the cumulant function $\lambda \mapsto C(\lambda \dagger \Theta) = \log \mathbb{E} e^{i\lambda\Theta}$ satisfies

$$C(\lambda \dagger \Theta) = ia\lambda - \tfrac{1}{2}\theta\lambda^2 + \int_{\mathbb{R}} \Big(e^{i\lambda x} - 1 - i\lambda x \mathbf{1}_{[-1,1]}(x) \Big) \rho(\mathrm{d}x) \tag{2.1}$$

for some triple (a, θ, ρ), where $a \in \mathbb{R}$, $\theta \geq 0$ and ρ is a measure on $\mathbb{R}$ satisfying $\rho(\{0\}) = 0$ and $\int_{\mathbb{R}} (1 \wedge x^2) \rho(\mathrm{d}x) < \infty$. The measure ρ is most often referred to as the Lévy measure. The class of infinitely divisible distributions covers many important distributions, for instance the Gaussian, lognormal, gamma, generalized inverse Gaussian, generalized hyperbolic and Gumbel, to name a few; see [18] and the references therein.

Any infinitely divisible process is associated to a Lévy process $L = (L_t)_{t \geq 0}$, which is a stochastic process possessing the following properties: $L_0 = 0$, the sample paths of L are càdlàg, and L has stationary and independent increments. These properties imply that the cumulant function of L satisfies $C(\lambda \dagger L_t) = tC(\lambda \dagger L_1)$, where L_1 is infinitely divisible and thus $C(\lambda \dagger L_1)$ is of the form (2.1); see e.g. [27, Corollary 11.6]. A natural extension of Lévy processes are the so-called Lévy bases, which for instance generalize the distributional properties to objects indexed by $\mathbb{R}^d$, $d \in \mathbb{N}$. A Lévy basis M on $\mathbb{R}^d$ is a random measure such that $(M(A_n))_{n\in\mathbb{N}}$ are independent, $M(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} M(A_n)$, and $M(A_1)$ is infinitely divisible for all disjoint Borelsets $A_1, A_2, \dots \subseteq \mathbb{R}^d$; see [26] for an exposition of infinitely divisible random measures and infinitely divisible stochastic integrals as in (2.2) below. In particular, if $d = 1$ and M on $[0,\infty)$ is stationary, we obtain a Lévy process as defined above by setting $L_t = M([0,t])$. In this dissertation we only consider stationary and isotropic Lévy bases, which are characterized by their Lévy-Khintchine representation

$$C(\lambda \dagger M(A)) = ia\lambda|A| - \tfrac{1}{2}\theta\lambda^2|A| + |A|\int_{\mathbb{R}} \Big(\mathrm{e}^{i\lambda x} - 1 - i\lambda x \mathbf{1}_{[-1,1]}(x)\Big)\rho(\mathrm{d}x)$$

for all $A \subseteq \mathbb{R}^d$, where $|\cdot|$ denotes the Lebesgue measure. A similar but more involved representation for bases which are not stationary and isotropic can be found in [26].

Lévy bases can be used to construct a wide range of random fields indexed by $\mathbb{R}^d$. If $f : \mathbb{R}^d \to \mathbb{R}$ is a measurable integration kernel, necessary and sufficient conditions for the Lévy-driven (moving average) field

$$X_v = \int_{\mathbb{R}^d} f(v-u)M(\mathrm{d}u) \qquad (v \in \mathbb{R}^d) \tag{2.2}$$

to exist are found in [22, Theorem 2.7]. Lévy-driven models as in (2.2) form a very rich modeling framework used for a variety of purposes, including modeling of financial assets ([4]), turbulent flows ([3]), brain imaging data ([14]) and wind power prices ([5]).

In [24, 25] such a field is studied, where the integration kernel is assumed to be positive and isotropic. Thus, the field under consideration is defined as

$$X_v = \int_{\mathbb{R}^d} f(|v-u|)M(\mathrm{d}u) \qquad (v \in \mathbb{R}^d), \tag{2.3}$$

where $f : [0,\infty) \to [0,\infty)$. The aim of the papers is to describe extremal probabilities of the field under the assumption that the Lévy measure ρ of M is convolution equivalent, meaning that it has an exponential (right) tail of strictly positive index, and that the tail of the convolution of ρ with itself is asymptotically of the same order as the tail of ρ; see ((C.2.2) and (C.2.3), Paper C) for the definition of a convolution equivalent measure. In this dissertation all Lévy measures are assumed convolution equivalent, and thus the Lévy measures mentioned in the remainder of this introduction will also be assumed to be convolution equivalent. For a comparison to results under the assumption of more general Lévy measures, see the introductory section of Paper C. Under mild assumptions on f it is shown in [24] that

$$\mathbb{P}\Big(\sup_{v\in B} X_v > x\Big) \sim K|B|\rho((x,\infty)) \qquad \text{as } x \to \infty, \tag{2.4}$$

for all compact and full-dimensional sets $B \subseteq \mathbb{R}^d$, where K is a computable constant. Here, $\sim$ denotes asymptotic equivalence, that is, $g(x) \sim h(x)$ as $x \to \infty$ if

$\lim_{x\to\infty} g(x)/h(x) = 1$. In [25] an equivalence similar to (2.4) is obtained, where the left-hand side is replaced with the probability that there is a translation and rotation of a fixed spatial object of indices in which (X_v) is larger than x. These papers constitute the motivation for Paper C, in which a time-dependence is added to the field.

In [9] a similar field

$$X_t = \int_{R_+\times\mathbb{R}} f(r, t-s)M(\mathrm{d}r, \mathrm{d}s) \qquad (t \geq 0) \tag{2.5}$$

is studied under the assumption of an underlying convolution equivalent Lévy measure. In contrast to the field in (2.3), the index set here is one-dimensional, and furthermore the kernel f is assumed to satisfy $f(r,s) = 0$ for all $s < 0$. It is shown that an equivalence similar to (2.4) holds true for $\mathbb{P}(\sup_{t\in[0,T]} X_t > x)$ in the limit $x \to \infty$. The proof is based on convergences of marked point processes, which, as an additional result, shows that the running supremum of (X_t) (when normalized) converges to the Gumbel distribution as $T \to \infty$. We will return to and elaborate on this result shortly.

Paper C

In this paper we extend the results of [24, 25] to apply to space-time fields of the form

$$X_{v,t} = \int_{\mathbb{R}^d\times\mathbb{R}} f(|v-u|, t-s)M(\mathrm{d}u, \mathrm{d}s),$$

where M is a Lévy basis on $\mathbb{R}^{d+1}$ with a convolution equivalent Lévy measure, and where we think of $v \in \mathbb{R}^d$ as the position in space and $t \geq 0$ as time. At a first glance, the field seems not to differ much from that defined in (2.3), however, by defining $f(r,s) = 0$ for all $s < 0$, the field $(X_{v,t})$ only depends on noise from the past, i.e. of noise accounted for by M up to time t. This is in contrast to the field in (2.3).

The main result of Paper C is a generalization of the asymptotic equivalence (2.4): If $B \subseteq \mathbb{R}^d$ and $[0,T]$ are compact and full-dimensional index sets, then, under mild assumptions on the kernel f and the underlying Lévy measure,

$$\mathbb{P}\Big(\Psi((X_{v,t})_{v\in B, t\in[0,T]}) > x\Big) \sim C\rho((x/c, \infty)) \qquad \text{as } x \to \infty, \tag{2.6}$$

for a large class of functionals Ψ, where the constants C, c depend on Ψ. In particular, if $\Psi((x_{v,t})) = \sup_{v\in B}\sup_{t\in[0,T]} x_{v,t}$, then (2.6) reads

$$\mathbb{P}\Big(\sup_{v\in B}\sup_{t\in[0,T]} X_{v,t} > x\Big) \sim KT|B|\rho((x,\infty)) \qquad \text{as } x \to \infty,$$

for some computable K.

Necessary to the proof of (2.6) is a result on continuity properties of the field $(X_{v,t})$ which also has a value of its own. Under certain continuity assumptions of the kernel f, we show that the field admits paths which are continuous in the space component and càdlàg in time.

2.2 Extreme value theory

In classical extreme value theory an i.i.d. sequence $(\xi_n)_{n\in\mathbb{N}}$ is considered, and the aim is to characterize the asymptotic distribution of $M_n = \max\{\xi_1,\dots,\xi_n\}$ as $n \to \infty$. One of the main results is the Extremal Types Theorem, stating that if there are sequences (a_n) and (b_n) such that

$$\mathbb{P}(M_n \leq a_n x + b_n) \to G(x) \qquad \text{for all } x \in \mathbb{R}, \tag{2.7}$$

where G is a non-degenerate distribution function, then G is one of three so-called extreme value distributions; see [8, 19, 23] for detailed treatments of classical extreme value theory. One such distribution is the Gumbel distribution, which is also infinitely divisible. Very conveniently, it is possible to conclude the asymptotic distribution of the maximum from the distribution of the individual ξ-variables: The convergence (2.7) is satisfied if and only if

$$n\mathbb{P}(\xi_1 > a_n x + b_n) \to -\log G(x) \qquad \text{for all } x \in \mathbb{R},$$

as $n \to \infty$. This equivalence between the tail of a single variable and the running maximum holds for more general limits and when replacing $a_n x + b_n$ with more general sequences (see [19, Theorem 1.5.1]). Results as the one just mentioned can also be obtained for dependent sequences. In particular, if a *stationary* sequence (ξ_n) satisfies certain mixing and anti-clustering conditions, then the conclusions above remain true; see e.g. [8, Section 4.4] for the definition of these conditions and for examples of stationary sequences satisfying one or both conditions.

In [20] a spatial random field indexed by $\mathbb{R}^2$ is considered with the purpose of obtaining an extremal types theorem similar as above. More specifically, if a stationary (continuously indexed) field $(Z_v)_{v\in\mathbb{R}^2}$ satisfies some coordinate-wise mixing condition, then the convergence

$$\mathbb{P}\Big(M(C_n) \leq a_n x + b_n\Big) \to G(x) \qquad \text{for all } x \in \mathbb{R},$$

where $M(C_n) = \sup_{v\in C_n} Z_v$ for an an increasing sequence of boxes $C_n \subseteq \mathbb{R}^2$, implies that G is of extreme value type.

As mentioned above, it is shown in [9] that the running supremum of the Lévy-based field $(X_t)_{t\geq 0}$ defined in (2.5) converges to the Gumbel distribution: There are sequences (a_T) and (b_T) such that

$$\mathbb{P}\Big(\sup_{t\in[0,T]} X_t \leq a_T x + b_T\Big) \to \exp\Big(\mathrm{e}^{-x} K\Big) \qquad \text{for all } x \in \mathbb{R},$$

as $T \to \infty$, where K is a computable constant. Generalizing this result to spatial Lévy-based objects as defined by (2.3) was the original motivation of Paper D.

Paper D

The aim of this paper is firstly to expand classical, one-dimensional extreme value results to the d-dimensional Euclidean space. Secondly, we apply similar techniques to describe the distribution function of (a normalization of) $\sup_{v\in C_n} X_v$, where (X_v) is a stationary Lévy-driven moving average given by (2.3) and (C_n) is a sequence of sufficiently nice increasing sets in $\mathbb{R}^d$.

Before turning to the probabilistic setup, we present conditions under which the deterministic sets (C_n) increase appropriately. We need the sets to expand in a particularly nice way such that they can be approximated by a certain class of expanding cubes. We require that C_n is a union of a fixed number of connected convex bodies (convex and compact sets with non-empty interior), which, due to the Steiner formula ([28]) from classical convex geometry, guarantees the appropriate set-expansion if only the so-called intrinsic volumes of C_n are sufficiently bounded relative to the volume of C_n; see (Assumption D.2.4, Paper D) for details. This is very desirable as the assumption (as expected) is then associated to the geometry of the sets at hand.

The first part of the paper concerns the distribution of $M(D_n) = \max_{v\in D_n} \xi_v$, where $(\xi_v)_{v\in\mathbb{Z}^d}$ is a stationary random field and (D_n) is a sequence of sets in $\mathbb{Z}^d$ increasing appropriately (meaning that there are continuous index sets C_n as above such that $D_n = C_n \cap \mathbb{Z}^d$). More precisely, for a given real sequence (x_n), we formulate mixing and anti-clustering conditions such that, when satisfied,

$$|D_n|\mathbb{P}(\xi_v > x_n) \to \tau \quad \text{if and only if} \quad \mathbb{P}(M(D_n) \le x_n) \to \mathrm{e}^{-\tau}$$

for all $0 \le \tau < \infty$, where $|D_n|$ denotes the number of points in D_n. Furthermore, assuming the mixing and anti-clustering conditions, an extremal types theorem holds under the additional assumption $|D_{n+1}|/|D_n| \to 1$, which ensures that the growth of the sequence (D_n) is not too explosive. All of the assumptions will in particular be satisfied if $C_n = r_n C$ for a union of convex bodies C and a sequence (r_n), where $r_n \to \infty$ and $r_{n+1}/r_n \to 1$. Hence, if there are sequences (a_n) and (b_n) such that

$$\mathbb{P}\Big(M(D_n) \le a_n x + b_n\Big) \to G(x) \qquad \text{for all } x \in \mathbb{R},$$

for a non-degenerate distribution G, then G is an extreme value distribution. Assuming only the mixing condition, we show that the same conclusion holds if in fact $|D_n| \sim a \cdot n$ for a finite constant a.

In the second part of the paper, we consider the Lévy-driven field $(X_v)_{v\in\mathbb{R}^d}$ given by (2.3), and we show that $\sup_{v\in C_n} X_v$ (when normalized) converges to the Gumbel distribution, i.e.

$$\mathbb{P}\Big(\sup_{v\in C_n} X_v \le a_n x + b_n\Big) \to \exp\Big(\mathrm{e}^{-x} K\Big) \qquad \text{for all } x \in \mathbb{R}, \tag{2.8}$$

for appropriate norming constants (a_n), (b_n), where K is again a computable constant. Actually, the proof techniques used to show this immediately gives the following complementary result: If $(Y_v)_{v\in\mathbb{R}^d}$ is a stationary and ergodic field independent of $(X_v)_{v\in\mathbb{R}^d}$ with sufficiently light tail, then the normalization of the running supremum $\sup_{v\in C_n}(X_v + Y_v)$ converges to the Gumbel distribution (in the sense of (2.8)).

References

[1] Baddeley, A.J., K.-A. Dorph-Petersen and E.B.V. Jensen (2006). A note on the stereological implications of irregular spacing of sections. *J. Microsc.* 222(3), 177–181.

[2] Baddeley, A.J. and E.B.V. Jensen (2004). *Stereology for Statisticians*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press, Boca Raton.

[3] Barndorff-Nielsen, O.E. and J. Schmiegel (2004). Lévy-based Tempo-Spatial Modelling with Applications to Turbulence. *Uspekhi Mat. Nauk* 159, 63–90.

[4] Barndorff-Nielsen, O.E. and N. Shephard (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B* 63(2), 167–241.

[5] Benth, F.E. and A. Pircalabu (2018). A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures. *Appl. math. finance* 25(1), 36–65.

[6] Cruz-Orive, L.M (1989). On the precision of systematic sampling: a review of Matheron's transitive methods. *J. Microsc.* 153, 315–333.

[7] Dorph-Petersen, K.-A., J.N. Pierri, Q. Wu, A.R. Sampson and D.A. Lewis (2007). Primary visual cortex volume and total neuron number are reduced in schizophrenia. *J. Comp. Neurol.* 501(2), 290–301.

[8] Embrechts, P., C. Klüppelberg and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Applications of mathematics. Springer, New York.

[9] Fasen, V. (2008). Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. *Extremes* 12(3), 265–296.

[10] García-Fiñana, M. (2006). Confidence intervals in Cavalieri sampling. *J. Microsc.* 222(3), 146–157.

[11] García-Fiñana, M. and L.M Cruz-Orive (2000a). Fractional trend of the variance in Cavalieri sampling. *Image Anal. Stereol.* 19(2), 71–79.

[12] García-Fiñana, M. and L.M Cruz-Orive (2000b). New approximations for the variance in Cavalieri sampling. *J. Microsc.* 199(3), 224–238.

[13] García-Fiñana, M. and L.M Cruz-Orive (2004). Improved variance prediction for systematic sampling on $\mathbb{R}$. *Statistics* 38, 243–272.

[14] Jónsdóttir, K.Ý., A. Rønn-Nielsen, K. Mouridsen and E.B.V. Jensen (2013). Lévy-based Modelling in Brain Imaging. *Scand. Stat. Theory Appl.* 40(3), 511–529.

[15] Kiderlen, M. and K.-A. Dorph-Petersen (2017). The Cavalieri estimator with unequal section spacing revisited. English. *Image Anal. Stereol.* 36(2), 135–141.

[16] Kiêu, K. (1997). *Three Lectures on Systematic Geometric Sampling*. Technical Report, Aarhus University.

[17] Kiêu, K., S. Souchet and J. Istas (1999). Precision of systematic sampling and transitive methods. *J. Stat. Plan. Inference* 77(2), 263–279.

[18] Kyprianou, A.E. (2006). *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. First edition. Springer, Berlin.

[19] Leadbetter, M.R., G. Lindgren and H. Rootzen (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer, New York.

[20] Leadbetter, M.R. and H. Rootzén (1998). "On Extreme Values in Stationary Random Fields". *Karatzas I., Rajput B.S., Taqqu M.S. (eds) Stochastic Processes and Related Topics*. Trends in Mathematics. Birkhäuser, Boston.

[21] Matheron, G. (1965). *Les Variables régionalisées et leur Estimation*. Masson et Cie, Paris.

[22] Rajput, B.S. and J. Rosinski (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Relat. Fields* 82, 451–487.

[23] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

[24] Rønn-Nielsen, A. and E.B.V. Jensen (2016). Tail asymptotics for the supremum of an infinitely divisible field with convolution equivalent Lévy measure. *J. Appl. Probab.* 53(1), 244–261.

[25] Rønn-Nielsen, A. and E.B.V. Jensen (2017). Excursion sets of infinitely divisible random fields with convolution equivalent Lévy measure. *J. Appl. Probab.* 54(3), 833–851.

[26] Samorodnitsky, G. (2016). *Stochastic Processes and Long Range Dependence*. Springer International Publishing.

[27] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.

[28] Schneider, R. (2014). *Convex Bodies: The Brunn–Minkowski Theory*. Second edition. Cambridge University Press, Cambridge.

[29] Ziegel, J., A.J. Baddeley, K.-A. Dorph-Petersen and E.B.V. Jensen (2010). Systematic sampling with errors in sample locations. *Biometrika* 97, 1–13.

[30] Ziegel, J., E.B.V. Jensen and K.-A. Dorph-Petersen (2011). Variance estimation for generalized Cavalieri estimators. *Biometrika* 98, 187–198.

Paper A

Asymptotic Variance of Newton-Cotes Quadratures based on Randomized Sampling Points

Mads Stehr and Markus Kiderlen

Accepted for publication by the Applied Probability Trust
(http://www.appliedprobability.org) in *Advances in Applied Probability* 52.4 (December 2020)

Abstract

We consider the problem of numerical integration when the sampling nodes form a stationary point process on the real line. In previous papers it was argued that a naïve Riemann sum approach can cause a severe variance inflation when the sampling points are not equidistant. We show that this inflation can be avoided using a higher order Newton-Cotes quadrature rule which exploits smoothness properties of the integrand. Under mild assumptions, the resulting estimator is unbiased and its variance asymptotically obeys a power law as a function of the mean point distance. If the Newton-Cotes rule is of sufficiently high order, the exponent of this law turns out to only depend on the point process through its mean point distance. We illustrate our findings with the stereological estimation of the volume of a compact object, suggesting alternatives to the well-established Cavalieri estimator.

MSC: 65D30; 60G55; 60G10; 60K05; 65D32

Keywords: Point process; stationary stochastic process; randomized Newton-Cotes quadrature; numerical integration; asymptotic variance bounds; renewal process

A.1 Introduction

Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function with compact support. We intend to approximate the integral of f based on its values at finitely many random sampling points. If $X \subset \mathbb{R}$ is a stationary point process with intensity $1/t > 0$, the random variable

$$\hat{V}_0(f) = t \sum_{x \in X} f(x), \tag{A.1.1}$$

is unbiased for $\int f(x)\mathrm{d}x$ due to Campbell's theorem; see, e.g. [10, Chap. 3]. The simplest situation is that of *equidistant* sampling points. In this case, we can write $X = t(U + \mathbb{Z})$, where U is a uniform random variable in the interval $(0,1)$, and the estimator becomes

$$\hat{V}(f) = t \sum_{k \in \mathbb{Z}} f\Big(t(U+k)\Big). \tag{A.1.2}$$

The estimator (A.1.2) corresponds to systematic sampling with randomized start. Its variance behavior is well understood; see e.g. [7]. It was remarked in [2] and [12] that the variance of (A.1.1) can be substantially larger in the non-equidistant case. The purpose of the present paper, following an idea in [6], is to show and quantify that this variance increase can be reduced – essentially to the level of the equidistant case – using Newton-Cotes quadrature approximations of sufficiently high order $n \in \mathbb{N}$ instead of the crude sum (A.1.1).

The resulting Newton-Cotes estimator $\hat{V}_n(f)$ is unbiased under mild assumptions; see Theorem A.2.2. To analyze the variance of $\hat{V}_n(f)$, the refined Euler-MacLaurin theory in [7] appears no longer to be sufficient, and we therefore extend the classical Peano kernel theorem ([11, Theorem 3.2.3]) to locally finite, not necessarily equidistant sets of nodes in Theorem A.2.3. This allows us to give explicit variance bounds in Theorem A.2.8 depending on the smoothness of the measurement function f. These bounds follow a power law as functions of t. Interestingly, if the order of the Newton-Cotes estimator is large enough (compared to the smoothness of f), the exponent of this power law coincides with the exponent in the equidistant case – independently of the covariance structure of X. However, if the Newton-Cotes order is too small, the exponent may be worse than in the equidistant case, and it may depend on X. We introduce the notion of strongly n-admissible point processes (see Definition A.2.7 for details) and show that the exponent for the variance bound is better for such point processes when n is not large enough; see (A.2.7). In Theorems A.2.9 and A.2.10 this general theory is applied to particular point process models: a model with i.i.d. perturbations of the equidistant case and a renewal process. In both cases, explicit variance expansions are derived for Newton-Cotes estimators of order $n = 1$ showing in particular that the power law exponents in Theorem A.2.8 in general can not be improved.

The paper is organized as follows. The main results, as outlined above, are stated rigorously in the next section. In Section A.3 more relevant notation is introduced, the nth order Newton-Cotes estimator is formally derived, and the refined Peano kernel theorem is proven. In Section A.4 we derive integrability statements which will be of relevance when proving the main results, Theorems A.2.2 and A.2.8, in Sections A.5 and A.6, respectively. In Section A.7 we show that point processes from the perturbed and cumulative model (renewal process) are strongly n-admissible for all $n \in \mathbb{N}$, and we derive the exact variance expressions presented in Theorems A.2.9 and A.2.10. Section A.8 applies our findings to the stereological problem of volume estimation of a compact set in $\mathbb{R}^3$ and contains a simulation study. Conclusions and ideas for future work can be found in Section A.9.

A.2 Main results

As for the estimators (A.1.1) and (A.1.2), throughout this paper we consider a point process X with intensity $1/t$ and we apply Newton-Cotes quadratures to functions evaluated at the points of X.

Let $n \in \mathbb{N}$ be given. We recall the definition of the *nth order Newton-Cotes estimator* $\hat{V}_n(f)$ from [6] for a fixed realization of X. On the interval from a point $x_0 \in X$ to its nth right neighbour in X, say x_n, the function f is approximated by a polynomial of degree at most $n \in \mathbb{N}$ passing through the points $\{x_j, f(x_j)\}_{j=0}^n$, where $x_1 < \cdots < x_{n-1}$ are the ordered points in $X \cap (x_0, x_n)$. $\hat{V}_n(f)$ is then an average of the integral of the concatenation of such approximations (composite rule) with respect to the starting point chosen. The estimator $\hat{V}_n(f)$ turns out to be a weighted average of f over all points in X,

$$\hat{V}_n(f) = \sum_{x \in X} \alpha(x;X) f(x), \tag{A.2.1}$$

where the weights satisfy $\alpha(x;X) = t\alpha(x/t;X/t)$ for all $x \in X$, where (when considered random) X/t is of unit intensity; see (A.3.4) for details. We will see in Remark A.5.1 that $\alpha(x;X) = t$ when $X = t(U + \mathbb{Z})$ is an equidistant process, and therefore, Newton-Cotes estimators of any order coincide with (A.1.2) in the equidistant case.

When applying the estimator on randomized sampling points, we work under the general assumption that a typical distance between two consecutive points has finite positive and negative moments of all orders:

Assumption A.2.1.

$$\mathbb{E}^0 h_1^j < \infty \qquad \textit{for all } j \in \mathbb{Z}. \tag{A.2.2}$$

Here $\mathbb{E}^0$ is the expectation under the Palm-distribution of X, that is, the distribution of X given that $0 \in X$ (see e.g. [10, Sec. 3.3]), and h_1 is the lag between 0 and its right neighbor in X. Note that (A.2.2) holds for X if and only if it holds for aX for all $a > 0$. Assumption A.2.1 is certainly not necessary for the results to hold for a given n, but finding a necessary and sufficient condition appears to be quite technical. Our first result shows the unbiasedness of $\hat{V}_n(f)$.

Theorem A.2.2. *Let $n \in \mathbb{N}$ and $t > 0$ be given and assume that X is a stationary point process such that Assumption A.2.1 is satisfied. Then $\hat{V}_n(f)$ is unbiased:*

$$\mathbb{E}\hat{V}_n(f) = \int_{\mathbb{R}} f(x)\mathrm{d}x$$

for all integrable and real-valued functions f with compact support.

This will be shown in Section A.5, where we also argue that Assumption A.4.1 below, which is weaker than Assumption A.2.1, is sufficient to ensure unbiasedness. We also remark that unbiasedness is known to hold for $n = 1$ without integrability conditions and for $n = 2$ under a condition weaker than Assumption A.4.1; see [6, Ex. 1 and Cor. 3].

Like in the case of classical quadrature, high order quadrature is reducing the discretization error when the measurement function is smooth. We adopt a smoothness condition which is in widespread use in stereological applications. For $m, p \in \mathbb{N}_0 \cup \{\infty\}$, we say that a measurable function f with compact support is (m, p)-piecewise smooth if it is in $C^k(\mathbb{R})$ for $k = \max\{m-1, 0\}$, and all derivatives up to order $m+p$ exist and are continuous except in at most finitely many points, where they may have finite jumps. Hence, if f is (m, p)-piecewise smooth, m is the smallest order of derivative of f which may have jumps; see e.g. [7] for details on such functions. For our results to hold, we require that $p \geq 1$, however, the exact value of p is otherwise irrelevant. We therefore state all results for $(m, 1)$-piecewise smooth functions. We say that a function f is *exactly* $(m, 1)$*-piecewise smooth* if it is $(m, 1)$-piecewise smooth with discontinuous mth derivative. We let $D_{f^{(m)}}$ denote the finite set of discontinuity points of $f^{(m)}$, with

$$a \mapsto J_{f^{(m)}}(a) = \lim_{x \to a^+} f^{(m)}(x) - \lim_{x \to a^-} f^{(m)}(x)$$

denoting the corresponding jump-function.

Our second result expresses the discretization error

$$R^{(n)}(f) = \hat{V}_n(f) - \int_{\mathbb{R}} f(x) \mathrm{d}x \tag{A.2.3}$$

in terms of higher order derivatives of f. We state it for a realization of X, that is, we consider X as a deterministic, locally finite set of distinct points with convex hull $\mathrm{Conv}(X) = \mathbb{R}$.

Theorem A.2.3 (Refined Peano kernel theorem for Newton-Cotes estimation). *Let $n \in \mathbb{N}$ be fixed. Given X and $m \leq n$ there exists a function K_m such that*

$$R^{(n)}(f) = \int_{\mathbb{R}} f^{(m+1)}(r) K_m(r) \mathrm{d}r + \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) K_m(a)$$

for all $(m, 1)$-piecewise smooth functions $f : \mathbb{R} \to \mathbb{R}$.

Remark A.2.4. The function K_m will be called the *mth Peano kernel*. It is a piecewise polynomial of order at most $m + 1$ with coefficients given in terms of X. The mth Peano kernel is explicitly given by (A.3.7), below. It is shown in Lemma A.4.2 that for a stationary point process X satisfying Assumption A.2.1, K_m is a stationary stochastic process on the real line with finite absolute moments of any (positive) order. In particular, the mean $\mathbb{E}K_m(0) = \mathbb{E}K_m(r)$ and the covariance function $H_m(s) = \mathrm{Cov}(K_m(r), K_m(s + r))$ are both finite and independent of $r \in \mathbb{R}$.

As initially considered and shown in [9], the variance of (A.1.2) in the equidistant case depends on jumps of high order derivatives of the measurement function; see also [7, Chapter 5]. This is outlined in the following for comparison with the general case. Let $*$ denote the convolution operator and let the reflection $\check{f}$ of f be defined as $\check{f}(x) = f(-x)$. When the measurement function f is $(m, 1)$-piecewise smooth, it can be shown [7, Corollary 5.8] that the so-called *covariogram* $g = f * \check{f}$ of f is $(2m+1,1)$-piecewise smooth. When f is exactly $(m,1)$-piecewise smooth, one usually decomposes the variance of $\hat{V}(f)$ as

$$\mathbb{V}\mathrm{ar}(\hat{V}(f)) = \mathbb{V}\mathrm{ar}_E(\hat{V}(f)) + Z(t) + o(t^{2m+2}) \tag{A.2.4}$$

when $t \downarrow 0$. The *Zitterbewegung* $Z(t)$, which is of order t^{2m+2}, is a finite sum of terms oscillating around 0, $o(t^{2m+2})$ is a low-order remainder and the *extension term*

$$\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = t^{2m+2} g^{(2m+1)}(0^+) c_m \tag{A.2.5}$$

explains the overall trend of the variance. Here $c_m = -\frac{2B_{2m+2}}{(2m+2)!} \neq 0$, where B_k is the kth Bernoulli number (see Section A.6 below), and as such c_m does not depend on t or the function f, other than through its order of smoothness.

Motivated by stereological applications, and adopting the naming from [12], we will mainly work with two classes of point process models. Both models are defined as scalings of unit-intensity processes.

Example A.2.5 (Perturbed model). A stationary point process X with intensity $1/t$ is from the perturbed model if it is derived from equidistant points by having i.i.d. perturbations tE_k, $k \in \mathbb{Z}$, of every point, i.e. $X = \{t(U + k + E_k)\}_{k \in \mathbb{Z}}$; see Section A.7.1. Note that the perturbations may have a degenerate distribution concentrated at 0, and hence the equidistant model is a particular instance of the perturbed model.

Example A.2.6 (Model with cumulative errors). A stationary point process X with intensity $1/t$ is from the model with cumulative errors if $X = tX^u$, where X^u is a unit-intensity two-sided stationary renewal process on the real line with holding times ω_i, $i \in \mathbb{Z}$. In particular, the holding times $\{t\omega_i\}$ between two consecutive points of X form an i.i.d. sequence; see Section A.7.2.

If X is from the perturbed model (with non-degenerate perturbations), the variance of (A.1.1) satisfies $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = t^2 c' + Z_0(t) + o(t^2)$ when $m = 0$ and $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = t^3 c'' + o(t^3)$ when $m \geq 1$ as $t \downarrow 0$. This was shown in [12, Prop. 1] apart from the missing Zitterbewegung term $Z_0(t)$ of order t^2 in the first equation, which was omitted there as it was erroneously claimed that the last term in [12, Eq. (A3)] is of order $o(t^{2m+2})$. Hence, for all $m \geq 1$, the rate of decrease of $V_0(f)$ in the non-equidistant case is strictly smaller than in the equisdistant case; cf. (A.2.5) for the latter. The behavior is even worse in the model with cumulative errors, as $\mathbb{V}\mathrm{ar}(\hat{V}_0(f)) = tc''' + o(t)$ for all $m \geq 0$; see [12, Prop. 2].

In order to formulate corresponding rates of decrease for Newton-Cotes estimators, we need the notion of an admissible point process. The Peano kernel in the definition of an admissible point process is explicitly given in (A.3.7) with $m = n$.

Definition A.2.7 (Admissible point process). *Let X be a stationary point process satisfying Assumption A.2.1. For $n \in \mathbb{N}$ let H_n be the covariance function of K_n. Then X is called strongly n-admissible if $\int_0^z H_n(s)\mathrm{d}s$ is uniformly bounded in $z \geq 0$. X is called weakly n-admissible if $\lim_{z\to\infty} \frac{1}{z} \int_0^z H_n(s)\mathrm{d}s = 0$.*

From the definition (A.3.7) of K_m it is easily seen that X is weakly/strongly admissible if and only if aX is weakly/strongly admissible for all constants $a > 0$. Admissibility is closely related to ergodicity properties of the stationary field K_n, and hence to those of X. In fact, if K_n has an exponentially decaying α-mixing coefficient (see, for instance, [5, Subsection 1.3.2] for the definition of this coefficient), then [5, Theorem 3.(1), p. 9] and the fact that $\mathbb{E}K_n(0)^{2+\varepsilon} < \infty$, $\varepsilon > 0$, imply that $H_n(s)$ is exponentially decaying, and hence, X is strongly n-admissible for all $n \in \mathbb{N}$.

The covariance function H_n need not be decaying for X to be strongly n-admissible. When X is from the perturbed model, the covariance function is closely related to Bernoulli functions (see Section A.6), which are 1-periodic functions integrating to 0 on an interval of unit length. This is used in Lemma A.7.1 to show that X is strongly n-admissible for all $n \in \mathbb{N}$. Concerning the model with cumulative errors, we show in Lemma A.7.3 that H_n is indeed exponentially decaying when assuming that the holding times of the process have finite exponential moments. The proof relies on a result from [1] concerning the convergence rate of convolutions of the renewal measure of a pure renewal process.

Theorem A.2.8. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process with intensity $1/t > 0$ such that Assumption A.2.1 holds. If f is $(m,1)$-piecewise smooth and $k = \min\{m,n\}$, the variance of the estimator* (A.2.1) *obeys*

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq ct^{2k+2} \tag{A.2.6}$$

for some constant c, which does not depend on t.

If $m > n$ and X is strongly n-admissible, then

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq c' t^{2n+3} \tag{A.2.7}$$

for some constant c', which does not depend on t.

If f is exactly $(m,1)$-piecewise smooth with $m < n$, the decrease rate in (A.2.6) is optimal. This is also true in the case $m = n$ if X is weakly n-admissible; see Remark A.6.2.

When using the trapezoidal estimator, that is $n = 1$, we have exact expressions of the asymptotic behavior of the variance when X is from the perturbed model and the model with cumulative errors. In the perturbed case, the rate of decrease of the upper bound in (A.2.7) is optimal if the perturbations E_i are non-degenerate.

Theorem A.2.9. *Let X be from the perturbed model with intensity 1/t, and let μ_k be the kth moment of the perturbations E_i. Assume that the measurement function f is exactly $(m,1)$-piecewise smooth with covariogram $g = f * \check{f}$. Then, for $t \downarrow 0$,*

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = -t^2 g'(0^+)(\mu_2 + \tfrac{1}{6}) + Z_0(t) + o(t^2), \qquad \textit{for } m = 0, \tag{A.2.8}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}(2\mu_2 + 2\mu_4 + \tfrac{1}{30}) + Z_1(t) + o(t^4), \qquad \textit{for } m = 1, \tag{A.2.9}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 g^{(4)}(0)\tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2) + o(t^5), \qquad \textit{for } m \geq 2, \tag{A.2.10}$$

where the Zitterbewegung $Z_m(t)$ is given by (A.7.7). *It is of order t^{2m+2}, and it is a finite sum of terms oscillating around 0. Moreover, if E_i has a density with a finite number of* *finite jumps and $m \geq 2$, the remainder $o(t^5)$ is explicitly given by*

$$\begin{aligned} t^6 g^{(5)}(0^+)\tfrac{1}{720}\Big(&-34\mu_2 - 90\mu_2^2 + 110\mu_4 + 180\mu_2\mu_4 \\ &- 180\mu_2^3 - 170\mu_3^2 + 8\mu_6 - \tfrac{1}{21}\Big) + Z_2(t) + o(t^6). \end{aligned} \tag{A.2.11}$$

We compare these findings with the equidistant case. The Zitterbewegung in (A.2.4) is not present in the decomposition of Theorem A.2.9 when $m \geq 2$, or rather it is of lower order and thus part of the low-order remainder. As the Bernoulli numbers satisfy $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$ and $B_6 = \frac{1}{42}$, the extension term (A.2.5) becomes $\mathbb{V}\mathrm{ar}_E \hat{V}(f) = -t^2 g'(0^+)\frac{1}{6}$, $\mathbb{V}\mathrm{ar}_E \hat{V}(f) = t^4 g^{(3)}(0^+)\frac{1}{12}\frac{1}{30}$ and $\mathbb{V}\mathrm{ar}_E \hat{V}(f) = -t^6 g^{(5)}(0^+)\frac{1}{21}\frac{1}{720}$ for $m = 0,1,2$, respectively. Hence, the extension term of the trapezoidal estimator with perturbed sampling can come arbitrarily close to (A.2.5) if the errors E_i are sufficiently small. Under the model with cumulative errors, a corresponding statement holds as a consequence of the following result.

Theorem A.2.10. *Let X be from the model with cumulative errors with intensity $1/t$ and let the i.i.d. holding times ω_i (of the unit-intensity process) satisfy $\mathbb{E}e^{\eta\omega_1} < \infty$ for some $\eta > 0$. Define ν_k as the kth moment of ω_1. Let the measurement function f be exactly $(m,1)$-piecewise smooth with covariogram $g = f * \check{f}$. Then, for $t \downarrow 0$,*

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = -t^2 g'(0^+)\tfrac{1}{6}\nu_3 + o(t^2), \qquad \text{for } m = 0, \tag{A.2.12}$$

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}\tfrac{1}{30}(6\nu_5 - 5\nu_3^2) + o(t^4), \qquad \text{for } m = 1\,. \tag{A.2.13}$$

A.3 The Peano kernel representation

In this section we consider a locally finite set $X \subset \mathbb{R}$ such that $\mathrm{Conv}(X) = \mathbb{R}$, and an integrable function $f : \mathbb{R} \to \mathbb{R}$ with compact support which is known at all points in X. For any $x \in X$ and $j \in \mathbb{Z}$ we define $s_j(x) = s_j(x;X)$ as the jth successor (predecessor for $j < 0$) of x in X, with $s_0(x) = x$ by definition. Hence, for $j \geq 1$, $s_j(x)$ and $s_{-j}(x)$ are the unique points in $X \cap (x,\infty)$ and $X \cap (-\infty,x)$, respectively, such that $\#(X \cap (x, s_j(x)]) = \#(X \cap [s_{-j}(x), x)) = j$. Note that

$$s_j(x+t; X+t) = s_j(x;X) + t \tag{A.3.1}$$

for all $t \in \mathbb{R}$. For all $x \in X$ and $j \in \mathbb{Z}$ we let $h_j(x) = h_j(x;X) = s_j(x;X) - s_{j-1}(x;X)$ be the distance from the jth successor (predecessor) of x to its left neighbour in X. By (A.3.1),

$$h_j(x+t; X+t) = h_j(x;X) \tag{A.3.2}$$

for all $t \in \mathbb{R}$. We now recall the principle of Newton-Cotes quadrature, adapted to an infinite set of nodes; see [6] for details. On the interval $[x, s_n(x)]$, $x \in X$, the function f is approximated by a polynomial of degree at most $n \in \mathbb{N}$ passing through the points $\{s_j(x), f(s_j(x))\}_{j=0}^n$. The integral of this polynomial on $[x, s_n(x)]$ is

$$I_x^{(n)}(f) = I_x^{(n)}(f;X) = \sum_{j=0}^{n} \beta_j^{(n)}(x) f(s_j(x))$$

where
$$\beta_j^{(n)}(x) = \beta_j^{(n)}(x;X) = \int_x^{s_n(x)} \prod_{\substack{k=0\\k\neq j}}^{n} \frac{y - s_k(x)}{s_j(x) - s_k(x)} \mathrm{d}y \tag{A.3.3}$$

for $x \in X$. The approximation $\hat{V}_n(f) = \frac{1}{n}\sum_{x\in X} I_x^{(n)}(f) = \sum_{x\in X}\alpha(x)f(x)$ is then an average of the sum of the integral-approximations $I_x^{(n)}$ with respect to the starting point chosen. Here
$$\alpha(x) = \alpha(x;X) = \frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(s_{-j}(x)). \tag{A.3.4}$$

Remark A.3.1. From [11, Theorem 2.1.1.1] the integral approximation on an interval $[x, s_n(x)]$ is exact whenever $f = p$ is a polynomial of degree at most n. That is, $R_x^{(n)}(p) = 0$, with the discretization error $R_x^{(n)}$ defined by

$$R_x^{(n)}(f) = R_x^{(n)}(f;X) = I_x^{(n)}(f) - \int_x^{s_n(x)} f(y)\mathrm{d}y, \tag{A.3.5}$$

$x \in X$.

As shown in Lemma A.I.1, $\beta_j^{(n)}$ is a rational function of point-increments, and (A.3.2) then implies that

$$\beta_j^{(n)}(x+t;X+t) = \beta_j^{(n)}(x;X) \quad \text{and} \quad \alpha(x+t;X+t) = \alpha(x;X) \tag{A.3.6}$$

for all $t \in \mathbb{R}$ and $x \in X$.

We are now ready to prove the refined Peano kernel theorem as stated in Theorem A.2.3. Given n, X, and $m \in \mathbb{N}_0$, the mth Peano kernel from Theorem A.2.3 is defined as
$$K_m(r) = K_m(r;X) = \frac{1}{m!n}\sum_{x\in X}\mathbf{1}_{(x,s_n(x)]}(r)R_x^{(n)}((\cdot - r)_+^m). \tag{A.3.7}$$

The mapping $x \mapsto (x-r)_+^m$ should be understood as

$$(x-r)_+^m = \begin{cases} (x-r)^m & \text{for} \quad x > r, \\ 0 & \text{for} \quad x \leq r. \end{cases}$$

Hence, K_m is a piecewise polynomial of degree at most $m+1$ with coefficients determined by X.

Proof of Theorem A.2.3. Fix $n \in \mathbb{N}$ and note that $nR^{(n)}(f) = \sum_{x\in X} R_x^{(n)}(f)$ due to (A.2.3) and (A.3.5). For all $x \in X$ and $y \in [x, s_n(x)]$, an induction argument using the refined partial integration formula [7, Lemma 4.1] yields

$$\begin{aligned} f(y^-) = &\sum_{k=0}^{m} \frac{f^{(k)}(x^+)}{k!}(y-x)^k \\ &+ \frac{1}{m!}\sum_{a\in D_{f^{(m)}}\cap(x,y)} J_{f^{(m)}}(a)(y-a)^m + \frac{1}{m!}\int_x^y f^{(m+1)}(t)(y-t)^m \mathrm{d}t, \end{aligned}$$

for all $(m,1)$-piecewise smooth functions f, $m \in \mathbb{N}_0$. We now assume $m \le n$. Using the linearity of $R_x^{(n)}$, the fact that all polynomials of order at most n are integrated exactly, and the fact that $R_x^{(n)}$ commutes with integration, we find that (with all expressions considered as functions of y)

$$\begin{aligned} &m! R_x^{(n)}(f) \\ &= R_x^{(n)}\Big(\sum_{a \in D_{f^{(m)}} \cap (x,y)} J_{f^{(m)}}(a)(y-a)^m + \int_x^y f^{(m+1)}(t)(y-t)^m \mathrm{d}t \Big) \\ &= \sum_{a \in D_{f^{(m)}} \cap (x, s_n(x)]} J_{f^{(m)}}(a) R_x^{(n)}\big((y-a)_+^m\big) + \int_x^{s_n(x)} f^{(m+1)}(t) R_x^{(n)}\big((y-t)_+^m\big) \mathrm{d}t. \end{aligned}$$

Changing the summation order, (A.3.7) implies that

$$R^{(n)}(f) = \frac{1}{n} \sum_{x \in X} R_x^{(n)}(f) = \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a) K_m(a) + \int_{\mathbb{R}} f^{(m+1)}(t) K_m(t) \mathrm{d}t,$$

as claimed. □

Before proceeding, we state a useful lemma on continuity properties of the Peano kernel. For $r \in \mathbb{R}$ we have

$$K_m(r) = \frac{1}{m!n} \sum_{x \in X} \mathbf{1}_{(x, s_1(x)]}(r) \sum_{i=1-n}^{0} R_{s_i(x)}^{(n)}\big((\cdot - r)_+^m\big).$$

The following result is a simple consequence of this representation and the fact that polynomials of degree at most n are approximated exactly.

Lemma A.3.2. *Fix $n \in \mathbb{N}$ and a locally finite point-set X with* $\mathrm{Conv}(X) = \mathbb{R}$*. Then, for all $x \in X$ and $m \in \mathbb{N}$, the function K_m is differentiable on $(x, s_1(x))$ with derivative $-K_{m-1}$ and jump*

$$J_{K_m}(x) = \frac{1}{m!n} R_x^{(n)}\big((\cdot - x)^m\big).$$

In particular, K_m is $(m-1)$-times continuously differentiable for all $1 \le m \le n$.

A.4 Integrability properties

To argue that $\hat{V}_n(f)$ is an unbiased estimator for $\int f(x)\mathrm{d}x$ when applied to randomized sampling points, we recall the notion of the Palm distribution of a stationary point process $X \subset \mathbb{R}$. It can be interpreted as the conditional distribution of X given that $0 \in X$. We denote it by $\mathbb{P}^0$ with the corresponding expectation denoted by $\mathbb{E}^0$. When considering the point process X under its Palm distribution, we will often suppress the dependence on the point $0 \in X$ in the various expression, i.e. under $\mathbb{P}^0$ we for instance write

$$s_i = s_i(0), \qquad h_i = h_i(0), \qquad \beta_j^{(n)} = \beta_j^{(n)}(0)$$

for all $i \in \mathbb{Z}$ and $j = 0, \dots, n$. In addition we write $\mathbf{h} = (h_1, \dots, h_n)$ and, for $i \in \mathbb{Z}$, $\mathbf{h}(s_i) = (h_1(s_i), \dots, h_n(s_i)) = (h_{i+1}, \dots, h_{i+n})$ under $\mathbb{P}^0$. As mentioned in Section A.1, a weaker assumption than (A.2.2) is sufficient to ensure the unbiasedness of the estimator.

Assumption A.4.1. *For a given $n \in \mathbb{N}$ we assume that*

$$\mathbb{E}^0\left[\frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{h}^{\mathbf{m}'}}\right] < \infty \tag{A.4.1}$$

for all multi-indices $\mathbf{m}, \mathbf{m}' \in \mathbb{N}_0^n$ with $|\mathbf{m}| \in \{n+1, n+2\}$ and $|\mathbf{m}'| = n$, where $|\mathbf{m}| = |(m_1, \ldots, m_n)| = \sum_{k=1}^n m_k$.

Using Hölder's inequality and [6, Eq. (13)], one shows that Assumption A.2.1 is stronger than Assumption A.4.1. In Lemma A.I.1 it is shown that the weight $\beta_j^{(n)}(x)$ is a rational function of the point-increments $(h_1(x), \ldots, h_n(x))$, $x \in X$, where the numerator is a homogeneous polynomial of degree $n+1$, and the denominator is a non-vanishing homogeneous polynomial of degree n with non-negative coefficients. From the fact that the Palm distribution is invariant under bijective point shifts [6, Eq. (13)], it is easily seen that $\mathbb{E}^0|\beta_j^{(n)}(s_{-j})| < \infty$ for all $j \in \{0, \cdots, n\}$ when Assumption A.4.1 is satisfied, and consequently

$$\mathbb{E}^0|\alpha(0)| < \infty, \tag{A.4.2}$$

see Lemma A.I.2. We conclude that either of the two assumptions is sufficient to guarantee the Palm-integrability of $\alpha(0)$, which will be used in the proof of Theorem A.2.2.

To argue for the variance bounds presented in Theorem A.2.8 we need higher-order moment and translation invariance properties of the Peano kernel K_m defined in (A.3.7).

Lemma A.4.2. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process. Then, for all $m \in \mathbb{N}_0$, K_m is a stationary stochastic process. If Assumption A.2.1 holds, $K_m(0)$ has finite absolute moments of all (positive) orders. Moreover, if X has intensity γ, K_m satisfies*

$$\mathbb{E}K_m(0) = \gamma \mathbb{E}^0 J_{K_{m+1}}(0) = \frac{\gamma}{(m+1)!n}\mathbb{E}^0 R_0^{(n)}\left((\,\cdot\,)^{m+1}\right) \tag{A.4.3}$$

for all $m \in \mathbb{N}_0$. In particular, $\mathbb{E}K_m(0) = 0$ for all $m < n$.

Proof. Fix $n \in \mathbb{N}$. For any $r, s \in \mathbb{R}$ and any locally finite pointset X, the Peano kernel satisfies

$$K_m(r+s; X) = K_m(r; X - s). \tag{A.4.4}$$

This follows from the definition of K_m and

$$R_x^{(n)}((\cdot - (r+s))_+^m; X) = R_{x-s}^{(n)}((\cdot - r)_+^m; X - s), \qquad x \in X,$$

which in turn is a consequence of (A.3.1) and (A.3.6). Due to (A.4.4) the stationarity of K_m is inherited from the stationarity of the point process X.

We now prove that $K_m(0)$ has finite absolute moments. Let $k \in \mathbb{N}$ be given. For arbitrary $r \in \mathbb{R}$ put $I_r = \{x \in X : r \in (x, s_n(x)]\}$. Using Hölder's inequality and some rather crude upper bounds we obtain from (A.3.5) and (A.3.7)

$$\left|K_m^k(0)\right| \le \sum_{x \in I_0}\left|R_x^{(n)}((\cdot)_+^m)\right|^k \le \sum_{x \in I_0}(s_n(x))^{km}\left(\sum_{j=0}^n|\beta_j^{(n)}(x)| + s_n(x)\right)^k.$$

By the refined Campbell Theorem [10, Theorem 3.5.3], (A.3.1) and (A.3.6) it follows that

$$\mathbb{E}\left|K_m^k(0)\right| \le \gamma \mathbb{E}^0 \int_0^{s_n} x^{km}\Big(\sum_{j=0}^n |\beta_j^{(n)}| + x\Big)^k \mathrm{d}x \le \gamma \mathbb{E}^0 s_n^{km+1}\Big(\sum_{j=0}^n |\beta_j^{(n)}| + s_n\Big)^k,$$

where γ is the intensity of X. By Lemma A.I.1, Assumption A.2.1 and the fact that $s_n = \sum_{j=1}^n h_j$ under $\mathbb{P}^0$, the variables s_n and $\beta_j^{(n)}$ have finite absolute moments of all orders under $\mathbb{P}^0$. This implies that $\mathbb{E}|K_m^k(0)| < \infty$.

Equation (A.4.3) is a simple consequence of the refined Campbell Theorem [10, Theorem 3.5.3], Lemma A.3.2 and [6, Eq. (13)]. □

A.5 Unbiasedness of Newton-Cotes estimators

Proof of Theorem A.2.2. Fix $n \in \mathbb{N}$ and let $X \subseteq \mathbb{R}$ be a stationary point process with finite and positive intensity γ. As α satisfies (A.3.6) and $\alpha(0)$ is Palm-integrable by (A.4.2), [6, Theorem 1] can be applied. It states that

$$\mathbb{E}\hat{V}_n(f) = \gamma \mathbb{E}^0[\alpha(0)] \int_{\mathbb{R}} f(x)\mathrm{d}x \tag{A.5.1}$$

holds for all integrable functions $f : \mathbb{R} \to \mathbb{R}$ with compact support. Hence, if we can show that $\mathbb{E}^0[\alpha(0)] = \gamma^{-1}$, we have shown that $\hat{V}_n(f)$ is unbiased.

For $s \in \mathbb{R}$ reuse the notation I_s from the end of the previous section. When f is an integrable function and $|f| \le 1$, (A.3.5) implies

$$\sum_{x \in I_s} |R_x^{(n)}(f)| \le \sum_{x \in I_s} \Big(\sum_{j=0}^n |\beta_j^{(n)}(x)| + (s_n(x) - x)\Big).$$

The refined Campbell theorem [10, Theorem 3.5.3], (A.3.1) and (A.3.6) imply

$$\mathbb{E}\sum_{x \in I_s} |R_x^{(n)}(f)| \le \gamma \mathbb{E}^0 \int_{s-s_n}^{s} \Big(\sum_{j=0}^n |\beta_j^{(n)}| + s_n\Big)\mathrm{d}x = \gamma \mathbb{E}^0\Big[s_n \sum_{j=0}^n |\beta_j^{(n)}| + s_n^2\Big] < \infty,$$

where the finiteness follows from Lemma A.I.1 and Assumption A.4.1, which is weaker than Assumption A.2.1. Note that the finite upper bound is independent of s.

Now let $r > 0$ be given and consider the function $f_r = \mathbf{1}_{[0,r]}$. Recall that

$$R^{(n)}(f_r) = \hat{V}_n(f_r) - \int_{\mathbb{R}} f_r(x)\mathrm{d}x = \frac{1}{n}\sum_{x \in X} R_x^{(n)}(f_r)$$

is the error of the nth Newton-Cotes estimator. The Newton-Cotes approximation on an interval $[x, s_n(x)]$ is exact for all polynomials of degree at most n, and in particular, it is exact for constant functions. Hence, $R_x^{(n)}(f_r) = 0$ whenever $[x, s_n(x)] \cap \{0, r\} = \emptyset$. This implies

$$|\mathbb{E}R^{(n)}(f_r)| \le \mathbb{E}\sum_{x \in I_0} |R_x^{(n)}(f)| + \mathbb{E}\sum_{x \in I_r} |R_x^{(n)}(f)| \le 2C,$$

for some finite $C \in \mathbb{R}$ which is independent of r. Equation (A.5.1) now implies

$$0 = \lim_{r \to \infty} \tfrac{1}{r}\mathbb{E}R^{(n)}(f_r) = \gamma \mathbb{E}^0 \alpha(0) - 1,$$

so $\mathbb{E}^0\alpha(0) = 1/\gamma$ as asserted. □

Remark A.5.1. If $X = t\big(U + \mathbb{Z}\big)$ is the equidistant point process, $\alpha(x) = t$ for all $x \in X$. In fact, the Palm version of X is the deterministic set $t\mathbb{Z}$, and (A.3.6) yields

$$\alpha(x; X) = \alpha(x - x; X - x) = \alpha(0; t\mathbb{Z})$$

for all $x \in X$. Hence $\alpha = \alpha(x)$ is deterministic, and $\hat{V}_n(f) = \alpha \sum_{x \in X} f(x)$. Assumption A.2.1 is trivially satisfied, so Theorem A.2.2 implies the well-known fact that $\hat{V}_n(f)$ is unbiased for $\int f \mathrm{d}x$. This is equivalent to $\alpha = t$.

A.6 Asymptotic variance behavior of Newton-Cotes estimators

In this and the following section we derive variance expressions showing the exact dependence on the mean point distance $t > 0$. To this end, we will consider the Peano kernel and its associated covariance function applied to the unit-intensity scaling of the process X: Let the point process of interest X have intensity $1/t$ and define its unit-intensity scaled process by $X^u = X/t$. Hence the Peano kernel $K_m(\cdot\,; X^u)$ with respect to X^u remains unchanged when X is rescaled. To avoid intricate notation, we will write $\mathbb{E}_u$, Cov_u, $\mathbb{V}\mathrm{ar}_u$ when expectation, covariance and variance are understood with respect to X^u; for instance $\mathbb{E}_u K_m(r) = \mathbb{E} K_m(r; X^u)$. Similarly, the Palm-distribution of X^u and its expectation are denoted $\mathbb{P}^0_u$ and $\mathbb{E}^0_u$, respectively. Lastly, the covariance function of K_m applied to X^u is denoted by $H^u_m(s) = \mathrm{Cov}_u\big(K_m(s + r), K_m(r)\big)$ for all $s, r \in \mathbb{R}$.

Before proving Theorem A.2.8, we recall the variance decomposition of the estimator (A.1.2) in the equidistant case, as it shows great resemblance to the new non-equidistant set-up. First we introduce the periodic Bernoulli functions P_m, which we define as in [8, Paragraph 297]: Let $(\tilde{P}_m)_{m=0}^{\infty}$ be the sequence of rescaled Bernoulli polynomials, which are defined inductively by $\tilde{P}_0(x) = 1$, $\tilde{P}_1(x) = x - \frac{1}{2}$ and $\tilde{P}'_{m+1} = \tilde{P}_m$, $\tilde{P}_{m+1}(0) = \tilde{P}_{m+1}(1) = \frac{1}{(m+1)!} B_{m+1}$, for $m \in \mathbb{N}$, where B_m is the mth Bernoulli number. This normalization is chosen as in [7] in order to ease comparison with the results there. Then $P_m(x) = \tilde{P}_m(x - \lfloor x \rfloor)$ is the mth Bernoulli polynomial, evaluated at the fractional part of $x \in \mathbb{R}$. Note that P_m is continuous for all $m \neq 1$. When the measurement function f is $(m, 1)$-piecewise smooth, the variance decomposes as [7, Chap. 5]

$$\mathbb{V}\mathrm{ar}(\hat{V}(f)) = -t^{2m+2} \sum_{a \in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a) P_{2m+2}(\tfrac{a}{t}) + o(t^{2m+2}) \tag{A.6.1}$$

as $t \downarrow 0$. Here, $g = f * \check{f}$ is the covariogram of f, and the term $o(t^{2m+2})$ can explicitly be given as $-t^{2m+2} \int_{\mathbb{R}} g^{(2m+2)}(s) P_{2m+2}(\frac{s}{t}) \mathrm{d}s$. When the point process X is not equidistant, we find a similar variance representation involving the Peano kernels instead of the periodic Bernoulli functions.

Proposition A.6.1. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process with intensity $1/t$ such that Assumption A.2.1 holds. If f is $(m, 1)$-piecewise smooth and $k = \min\{m, n\}$, then*

$$\begin{aligned}(-1)^{k+1} \mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = t^{2k+2} &\sum_{a \in D_{g^{(2k+1)}}} J_{g^{(2k+1)}}(a) H^u_k(\tfrac{a}{t}) \\ &+ t^{2k+2} \int_{\mathbb{R}} g^{(2k+2)}(s) H^u_k(\tfrac{s}{t}) \mathrm{d}s.\end{aligned} \tag{A.6.2}$$

If $k = m < n$ or X is weakly n-admissible, the variance behavior is determined by the first term, as

$$\int_{\mathbb{R}} g^{(2k+2)}(s)H_k^u(\tfrac{s}{t})\mathrm{d}s = o(1)$$

for $t \downarrow 0$.

Proof. The definition of $\alpha(x)$ and elementary calculations give

$$\alpha(x;X) = t\alpha(x/t;X^u)$$

for $x \in X$, so putting $f_t(x) = f(tx)$ we see that

$$\hat{V}_n(f) = t\hat{V}_n(f_t;X^u), \tag{A.6.3}$$

where the latter estimator is given in terms of the unit-intensity process X^u. As $k \leq n$, Theorem A.2.3 implies

$$R^{(n)}(f_t;X^u) = \int_{\mathbb{R}} f_t^{(k+1)}(s)K_k(s;X^u)\mathrm{d}s + \sum_{a\in D_{f_t^{(k)}}} J_{f_t^{(k)}}(a)K_k(a;X^u).$$

Using $f_t'(x) = tf'(tx)$ whenever the derivative is defined, we arrive at

$$R^{(n)}(f_t;X^u) = t^k\int_{\mathbb{R}} f^{(k+1)}(s)K_k(\tfrac{s}{t};X^u)\mathrm{d}s + t^k\sum_{a\in D_{f^{(k)}}} J_{f^{(k)}}(a)K_k(\tfrac{a}{t};X^u).$$

Hence, using (A.6.3) and the unbiasedness of $\hat{V}_n$, we get

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) &= t^2\mathbb{V}\mathrm{ar}_u(\hat{V}_n(f_t)) = t^2\mathbb{E}_u(R^{(n)}(f_t))^2\\ &= t^{2k+2}\mathbb{E}_u\Big(\int_{\mathbb{R}} f^{(k+1)}(s)K_k(\tfrac{s}{t})\mathrm{d}s + \sum_{a\in D_{f^{(k)}}} J_{f^{(k)}}(a)K_k(\tfrac{a}{t})\Big)^2.\end{aligned} \tag{A.6.4}$$

An application of [7, Prop. 5.7] yields

$$\begin{aligned}f^{(k+1)} * \check{f}^{(k+1)}(x) &= (-1)^{k+1}g^{(2k+2)}(x)\\ &\quad - \sum_{a\in D_{f^{(k)}}} J_{f^{(k)}}(a)f^{(k+1)}(a-x) - \sum_{a\in D_{f^{(k)}}} J_{f^{(k)}}(a)f^{(k+1)}(a+x),\end{aligned} \tag{A.6.5}$$

and furthermore the jumps of $g^{(2k+1)}$ are given by $J_{g^{(2k+1)}} = (-1)^{k+1}J_{f^{(k)}} * \check{J}_{f^{(k)}}$, see [7, Eq. (5.12)]. The stationarity and square integrability of K_k from Lemma A.4.2 implies that $\mathbb{E}_u K_k(r)$ and $H_k^u(s) = \mathrm{Cov}_u(K_k(r), K_k(s+r))$ are both finite and independent of $r \in \mathbb{R}$. Equation (A.6.2) now follows by expanding (A.6.4), applying (A.6.5), and using the structure of $J_{g^{(2k+1)}}$ together with Fubini's theorem. The latter may be applied due to the square integrability of K_k and the fact that $f^{(k+1)}$ is bounded with compact support.

We now show $\lim_{t\downarrow 0}\int_{\mathbb{R}} g^{(2k+2)}(s)H_k^u(\tfrac{s}{t})\mathrm{d}s = 0$ if $k = m < n$ or X is weakly n-admissible. The weak admissibility assumption yields

$$\lim_{t\downarrow 0}\int_0^1 H_k^u(\tfrac{s}{t})\mathrm{d}s = 0 \tag{A.6.6}$$

for $k = n$. Equation (A.6.6) also holds for $k = m < n$ without additional assumptions. In fact, for $k < n$ we have $K'_{k+1} = -K_k$ by Lemma A.3.2 and thus using Fubini's Theorem,

$$\left|\int_0^t H_k^u(s)\mathrm{d}s\right| = \left|\mathrm{Cov}_u\Big(K_k(0), K_{k+1}(0)\Big) - \mathrm{Cov}_u\Big(K_k(0), K_{k+1}(t)\Big)\right| \le c < \infty,$$

where Hölder's inequality and the stationarity of the Peano kernels have been used to show that the constant c is independent of t. A substitution allows to derive (A.6.6) from this.

Now fix $k \le n$ and let $\epsilon > 0$ be given. As $g^{(2k+2)}$ is integrable and bounded, there is a simple function ϕ such that $\phi \le g^{(2k+2)}$ and

$$0 \le \int_{\mathbb{R}} g^{(2k+2)}(s)\mathrm{d}s - \int_{\mathbb{R}} \phi(s)\mathrm{d}s < \frac{\epsilon}{2C},$$

where the finite constant $C > 0$ satisfies $\sup_{s\in\mathbb{R}} |H_k^u(s)| \le C$. This implies that

$$\left|\int_{\mathbb{R}} g^{(2k+2)}(s)H_k^u(\tfrac{s}{t})\mathrm{d}s - \int_{\mathbb{R}} \phi(s)H_k^u(\tfrac{s}{t})\mathrm{d}s\right| < \frac{\epsilon}{2}.$$

As ϕ is simple, (A.6.6) implies that $\lim_{t\downarrow 0}\int_{\mathbb{R}} \phi(s)H_k^u(\frac{s}{t})\mathrm{d}s = 0$. We conclude that $|\int_{\mathbb{R}} \phi(s)H_k^u(\frac{s}{t})\mathrm{d}s| < \frac{\epsilon}{2}$ for sufficiently small $t > 0$, and hence

$$\left|\int_{\mathbb{R}} g^{(2k+2)}(s)H_k^u(\tfrac{s}{t})\mathrm{d}s\right| < \epsilon$$

for such small $t > 0$. □

Proof of Theorem A.2.8. Recall that $k = \min\{m, n\}$. Due to Lemma A.4.2 there exists $C < \infty$ such that $\sup_{s\in\mathbb{R}} |H_k^u(s)| \le C$, and we immediately see from (A.6.2) that

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \le t^{2k+2}\Big(C\|g^{(2k+2)}\|_\infty \lambda(\operatorname{supp} g) + C\sum_{a\in D_{g^{(2k+1)}}} |J_{g^{(2k+1)}}(a)|\Big),$$

where $\lambda(\operatorname{supp} g) < \infty$ is the Lebesgue measure of the support of g. As g is $(2k+1, 1)$-piecewise smooth by [7, Corollary 5.8], the t-independent constant is finite, and (A.2.6) therefore follows.

For the stronger result (A.2.7), note that $m > n$ and hence g is $(2n+3, 1)$-piecewise smooth, and in particular $g^{(2n+2)}$ is continuous. An application of Proposition A.6.1 to the $(n, 1)$-piecewise smooth function f and a substitution gives

$$(-1)^{n+1}\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = t^{2n+3}\int_{\mathbb{R}} g^{(2n+2)}(st)H_n^u(s)\mathrm{d}s. \tag{A.6.7}$$

Let $b > 0$ satisfy $\operatorname{supp} g \subset [-b, b]$. As $g^{(2n+3)}$ is bounded and measurable, $g^{(2n+2)}$ is absolutely continuous. As H_n^u is bounded and hence integrable on $[-b/t, b/t]$ for any $t > 0$, also the function V given by $V(s) = \int_{-b/t}^s H_n^u(y)\mathrm{d}y$ is absolutely continuous on $[-b/t, b/t]$ with derivative H_n^u almost everywhere; see e.g. [4, Section 9.3] for details on absolutely continuous functions. Furthermore, as X and therefore X^u are assumed strongly n-admissible, V is bounded by a t-independent constant C', say. Partial integration for absolutely continuous functions [4, Theorem 9] shows

$$\int_{\mathbb{R}} g^{(2n+2)}(st)H_n^u(s)\mathrm{d}s = \int_{-b/t}^{b/t} g^{(2n+2)}(st)H_n^u(s)\mathrm{d}s = -t\int_{-b/t}^{b/t} g^{(2n+3)}(st)V(s)\mathrm{d}s,$$

where we used that $g^{(2n+2)}$ vanishes at $\pm b$. Returning to (A.6.7) we find

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) \leq t^{2n+3} t \int_{-b/t}^{b/t} |g^{(2n+3)}(st)||V(s)|\mathrm{d}s \leq t^{2n+3} 2b \|g^{(2n+3)}\|_\infty C'.$$

This proves the assertion. □

Remark A.6.2. If f is exactly $(m,1)$-piecewise smooth with $m \leq n$ and (A.6.6) is satisfied with $k = m$, the variance

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = (-1)^{m+1} t^{2m+2} \sum_{a \in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a) H_m^u(\tfrac{a}{t}) + o(t^{2m+2})$$

is exactly of order t^{2m+2}. This is easily seen by assuming that

$$\sum_{a \in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a) H_m^u(\tfrac{a}{t}) \to 0$$

as $t \to 0$, and using that $g^{(2m+1)}$ has a jump at 0. Applying (A.6.6) yields a contradiction. In particular, the decrease rate in (A.2.6) is optimal if $m < n$ or X is weakly n-admissible.

A.7 Variance behavior under perturbed and cumulative sampling

In this section the general findings will be exemplified and made more explicit for the perturbed model and the model with cumulative errors introduced in Examples A.2.5 and A.2.6, respectively.

A.7.1 Perturbed sampling

To construct the perturbed model, we let U be uniform on $(0,1)$ and independent of the sequence of i.i.d. variables $\{E_i\}_{i\in\mathbb{Z}}$, where $|E_i| < \frac{1}{2}$ almost surely and $\mathbb{E}E_1 = 0$. The perturbed model is the stationary point process $X = \{x_i\}_{i\in\mathbb{Z}}$ for which $x_i = t\Big(U + i + E_i\Big)$, for all $i \in \mathbb{Z}$. Under its Palm distribution, we have

$$h_k = t\Big(1 + E_k - E_{k-1}\Big) \leq 2t, \tag{A.7.1}$$

$k \in \mathbb{Z}$, so (A.2.2) is equivalent to

$$\mathbb{E}(1 + E_1 - E_0)^{-j} < \infty \tag{A.7.2}$$

for all $j \in \mathbb{N}$. For instance, (A.7.2) holds if there is $\epsilon > 0$ such that $|E_0| \leq \frac{1}{2} - \epsilon$ almost surely. For the scaled perturbed model X^u we define the shifted kernel K_m^* by

$$K_m^*(r) = K_m(r + U; X^u) = K_m(r; X^u - U)$$

for $m \in \mathbb{N}_0$. Note that it only depends on the perturbations $\{E_i\}$ and not on the initial uniform translation, and thus it is not (necessarily) a stationary process. However, by the i.i.d. structure of $\{E_i\}$ and the fact that $\beta_j^{(n)}$ is a rational function of point-increments, we see that

$$K_m^*(r) \stackrel{\mathcal{D}}{=} K_m^*(r + k) \tag{A.7.3}$$

for all $k \in \mathbb{Z}$. This can be used to show that X^u (and equivalently X) is strongly admissible.

Lemma A.7.1. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process from the perturbed model such that* (A.7.2) *holds. Then, for all $m \in \mathbb{N}_0$ and $r \geq 2n+2$, $H_m^u(r) = H_m^u(r+1)$ and $\int_r^{r+1} H_m^u(s)\mathrm{d}s = 0$. In particular, X is strongly n-admissible.*

Proof. Fix $n \in \mathbb{N}$ and $r \geq 2n+2$. For such r, the independence between U and $\{E_i\}$ yields

$$\begin{aligned}\mathbb{E}_u[K_m(0)K_m(r)] &= \int_0^1 \mathbb{E}[K_m^*(-u)K_m^*(r-u)]\mathrm{d}u \\ &= \int_0^1 \mathbb{E}[K_m^*(-u)]\mathbb{E}[K_m^*(r-u)]\mathrm{d}u.\end{aligned} \tag{A.7.4}$$

Equations (A.7.3) and (A.7.4) imply that $\mathbb{E}_u[K_m(0)K_m(r)] = \mathbb{E}_u[K_m(0)K_m(r+1)]$ and by stationarity of K_m we conclude that $H_m^u(r) = H_m^u(r+1)$.

Returning to (A.7.4), we find by Fubini's theorem, a substitution and the stationarity of K_m that

$$\begin{aligned}&\int_r^{r+1} \mathbb{E}_u[K_m(0)K_m(s)]\mathrm{d}s \\ &\quad= \int_0^1 \mathbb{E}[K_m^*(-u)]\int_0^1 \mathbb{E}[K_m^*(r+1-u-s)]\mathrm{d}s\mathrm{d}u \\ &\quad= \int_0^1 \mathbb{E}[K_m^*(-u)]\mathbb{E}_u[K_m(r+1-u)]\mathrm{d}u = (\mathbb{E}_u K_m(0))^2,\end{aligned}$$

which yields the asserted properties of H_m^u. This clearly implies that X^u and hence also X are strongly n-admissible. □

In order to obtain explicit leading terms in Theorem A.2.9 we state in the following a connection between the covariance function H_m^u and certain periodic Bernoulli functions. For our purpose, it is enough to consider $m \in \{0,1\}$, but we also state that the result holds for all m, when no perturbations are present.

Lemma A.7.2. *Let $n = 1$ and let X be from the perturbed model. Then*

$$H_m^u(r) = (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)] \tag{A.7.5}$$

for $m = 0,1$ and all $|r| \geq 4$.

If $X = t\big(U + \mathbb{Z}\big)$ (hence $X^u = U + \mathbb{Z}$) is the equidistant model, then

$$H_m^u(r) = (-1)^m P_{2m+2}(r) \tag{A.7.6}$$

for all $n \in \mathbb{N}$, $m \leq n$ and $r \in \mathbb{R}$.

The proof of Lemma A.7.2 can be found in the supplementary material; see Corollaries A.II.2 and A.II.3. As a consequence of (A.7.6), for the equidistant model $X = t\big(U + \mathbb{Z}\big)$, the variance representation (A.6.2) found using the Peano kernels coincides with the classical variance representation (A.6.1) found using Euler-McLaurin formulae.

Before turning to the proof of Theorem A.2.9, we emphasize that the integrability condition (A.2.2), or equivalently, condition (A.7.2), was omitted in the statement of the Theorem as we work with the trapezoidal rule. In fact, the unbiasedness of $\hat{V}_1(f)$

for all stationary point processes and integrable, compactly supported functions f was already remarked in the paragraph following the statement of Theorem A.2.2. Due to (A.7.1), the weights satisfy

$$\beta_0^{(1)}(x) = \beta_1^{(1)}(x) = \tfrac{1}{2}h_1(x) \le t,$$

$x \in X$, which replaces condition (A.2.2) in all the arguments in Sections A.5 and A.6. The assumptions of Proposition A.6.1 are thus satisfied.

Proof of Theorem A.2.9. Let $m \in \{0,1\}$. The $(2m+1)$st derivative of the covariogram g is an odd function, implying $J_{g^{(2m+1)}}(0) = 2g^{(2m+1)}(0^+)$. As X is strongly admissible, Proposition A.6.1 in combination with (A.7.5) yields the variance decomposition

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = (-1)^{m+1}t^{2m+2}2g^{(2m+1)}(0^+)H_m^u(0) + Z_m(t) + o(t^{2m+2}),$$

where the Zitterbewegung $Z_m(t)$ is given by

$$Z_m(t) = -t^{2m+2}\sum_{a\in D_{g^{(2m+1)}}\setminus\{0\}} J_{g^{(2m+1)}}(a)\mathbb{E}[P_{2m+2}(\tfrac{a}{t} + E_1 - E_0)]. \tag{A.7.7}$$

The facts that Z_m is a finite sum of terms each oscillating around 0 and that it is of order t^{2m+2} follow from arguments similar to those of [7, Section 5.2] as f is assumed to be exactly $(m,1)$-piecewise smooth. By the refined Campbell Theorem [10, Theorem 3.5.3] and the facts that $\mathbb{E}_u K_0(0) = 0$ and $\mathbb{E}_u K_1(0) = \frac{1}{2}\mathbb{E}_u^0[R_0^{(1)}((\cdot)^2)]$ by (A.4.3), we find that $H_m^u(0) = \mathbb{V}\mathrm{ar}_u(K_m(0))$ satisfies

$$H_0^u(0) = \mathbb{E}_u^0\int_0^{h_1}(\tfrac{1}{2}h_1 - y)^2\mathrm{d}y = \tfrac{1}{12}\mathbb{E}_u^0 h_1^3\,, \tag{A.7.8}$$

$$H_1^u(0) = \mathbb{E}_u^0\int_0^{h_1}(\tfrac{1}{2}h_1 y - \tfrac{1}{2}y^2)^2\mathrm{d}y - (\tfrac{1}{12}\mathbb{E}_u^0 h_1^3)^2 = \tfrac{1}{120}\mathbb{E}_u^0 h_1^5 - \tfrac{1}{144}(\mathbb{E}_u^0 h_1^3)^2. \tag{A.7.9}$$

Using (A.7.1), it is elementary to conclude (A.2.8) and (A.2.9).

Now let $m \ge 2$ be given and define $\tilde{H}_1^u$ by $\tilde{H}_1^u(s) = H_1^u(s) + \mathbb{E}[P_4(s + E_1 - E_0)]$. Due to Lemma A.7.2, $\tilde{H}_1^u(s)$ vanishes for $|r| > 4$. Since $g^{(4)}$ is continuous, an application of Proposition A.6.1 to the $(1,1)$-piecewise smooth function f, Fubini's theorem and the refined partial integration formula [7, Lemma 4.1] yield

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5\int_{\mathbb{R}} g^{(4)}(st)\tilde{H}_1^u(s)\mathrm{d}s - t^6\int_{\mathbb{R}} g^{(6)}(s)\mathbb{E}[P_6(\tfrac{s}{t} + E_1 - E_0)]\mathrm{d}s \\ - t^6\sum_{a\in D_{g^{(5)}}} J_{g^{(5)}}(a)\mathbb{E}[P_6(\tfrac{a}{t} + E_1 - E_0)].\end{aligned} \tag{A.7.10}$$

As the last two terms in (A.7.10) are of order $o(t^5)$, we only have to simplify the first term.

For all sufficiently small $t > 0$ and all $s \in \mathbb{R}$ with $|s| \le 4$ the function $g^{(4)}$ is differentiable on the open interval with endpoints 0 and st, so there is a point ξ_{st} in this interval such that

$$g^{(4)}(st) = g^{(4)}(0) + g^{(5)}(\xi_{st})st$$

by the mean value theorem. Inserting this into the first term of (A.7.10), and using the fact that $g^{(5)}$ and $\tilde{H}_1^u$ are bounded, yields

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 g^{(4)}(0)\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s + o(t^5) \tag{A.7.11}$$

as $t \downarrow 0$.

Noting that P_4 integrates to 0 on each interval of unit length, another application of Fubini's theorem, the refined partial integration formula [7, Lemma 4.1] and Lemma A.3.2 gives

$$\begin{aligned}\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s &= \int_{-4}^{4} H_1^u(s)\mathrm{d}s \\ &= \mathbb{E}_u\Big(\Big[K_2(-4) + \sum_{x\in X^u\cap[-4,4]} J_{K_2}(x) - K_2(4)\Big]\Big(K_1(0) - \mathbb{E}_u K_1(0)\Big)\Big).\end{aligned} \tag{A.7.12}$$

The arguments that lead to (A.7.4) in combination with (A.7.3) where $r = -4$ and $k = 8$ imply $\mathbb{E}_u[K_2(-4)K_1(0)] = \mathbb{E}_u[K_2(4)K_1(0)]$, and the two marginal terms in the last expression of (A.7.12) cancel. Hence, by the refined Campbell Theorem [10, Theorem 3.5.3] and the translation covariance of J_{K_2},

$$\begin{aligned}\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s &= \mathbb{E}_u \sum_{x\in X^u\cap[-4,4]} J_{K_2}(x)\Big(K_1(0) - \mathbb{E}_u K_1(0)\Big) \\ &= \mathbb{E}_u^0 J_{K_2}(0)\int_{-4}^{4}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x = \sum_{j=-3}^{2}\theta_j + Q.\end{aligned} \tag{A.7.13}$$

where $\theta_j = \mathbb{E}_u^0 J_{K_2}(0)\int_{s_j}^{s_{j+1}}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x$ and

$$Q = \mathbb{E}_u^0 J_{K_2}(0)\int_{-4}^{s_{-3}}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x + \mathbb{E}^0 J_{K_2}(0)\int_{s_3}^{4}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x.$$

Here we have used the fact that $s_3 \le 4$, $s_5 \ge 4$, $s_{-3} \ge -4$ and $s_{-5} \le -4$ under $\mathbb{P}_u^0$ (the Palm-distribution of X^u). Using Lemma A.4.2, it is seen that $J_{K_2}(0) = (1/12)h_1^3$ and consequently that θ_j evaluates to $\mathbb{E}_u^0(1/144)h_1^3(h_{j+1}^3 - h_{j+1}\mathbb{E}_u^0 h_1^3)$. As h_{j+1} only depends on the perturbations E_{j+1} and E_j, we conclude by independence that $\theta_j = 0$ for all $|j| > 1$. Moreover, a coupling argument shows that $Q = \theta_3 + \theta_4 = 0$. The Palm expectation of $J_{K_2}(0)\int_{-4}^{s_{-3}}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x$ is unchanged when we put $E_{-3} = E_5$, $E_{-4} = E_4$ and $E_{-5} = E_3$. Under this coupling assumption, $s_{-3} = s_5 - 8$, $s_{-4} = s_4 - 8$, $s_{-5} = s_3 - 8$, $h_{-3} = h_5$, and $h_{-4} = h_4$, and hence

$$\begin{aligned}Q &= \mathbb{E}_u^0 J_{K_2}(0)\int_{-4}^{s_5-8}\Big(K_1(x+8) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x + \mathbb{E}_u^0 J_{K_2}(0)\int_{s_3}^{4}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x \\ &= \mathbb{E}_u^0 J_{K_2}(0)\int_{s_3}^{s_5}\Big(K_1(x) - \mathbb{E}_u K_1(0)\Big)\mathrm{d}x = \theta_3 + \theta_4.\end{aligned}$$

Summarizing, we obtain from (A.7.13) that

$$\begin{aligned}\int_{-4}^{4} \tilde{H}_1^u(s)\mathrm{d}s &= \tfrac{1}{144}\sum_{j=-1}^{1} \mathbb{E}_u^0 h_1^3(h_{j+1}^3 - h_{j+1}\mathbb{E}_u^0 h_1^3) \\ &= \tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2),\end{aligned}$$

where the last equality follows from lengthy and tedious –but elementary– calculations. Inserting this into (A.7.11) yields the assertion (A.2.10).

The expression (A.2.11) of the remainder is found by different arguments which will be detailed in the appendix following this paper. □

A.7.2 Cumulative sampling

Before turning to the proof of Theorem A.2.10, we state in Lemma A.7.3 below that the covariance function of the Peano kernel decreases exponentially, from which admissibility follows.

The unit-intensity scaled cumulative process X^u is a stationary point process with i.i.d. holding times $\{\omega_i\}_{i\in\mathbb{Z}}$. We assume that ω_1 has cumulative distribution function F with density wrt. Lebesgue measure such that $F(0) = 0$. Moreover, since X^u has intensity 1, the holding times satisfy $\mathbb{E}\omega_i = 1$. To explicitly construct the point process X^u, the first point X_0 of $X^u \cap (0,\infty)$ is chosen with cumulative distribution function G,

$$G(x) = \int_0^x \bar{F}(y)\mathrm{d}y, \qquad x \geq 0,$$

where $\bar{F}(y) = 1 - F(y)$; see eg. [1, Chap. V: Cor. 3.6]. Note that the distribution G has density $\bar{F}$. Given X_0, the last point X_{-1} of $X \cap (-\infty, 0)$ (i.e. largest point) is chosen according to $X_{-1} = X_0 - \omega^*$, where ω^* is the conditional distribution of ω_0 given $\omega_0 > X_0$. This assures that $X_{-1} < 0$, and corrects [12], where ω_0 was used instead of ω^*. Having chosen increments $\{\omega_i\}_{i\neq 0}$ independent of X_{-1}, X_0, and setting $x_0 = X_0$, $x_i = X_0 + \sum_{\ell=1}^{i} \omega_\ell$ and $x_{-i} = X_{-1} + \sum_{\ell=1}^{i-1} \omega_{-\ell}$, for all $i \in \mathbb{N}$, we obtain a realization $X^u = \{x_i\}_{i\in\mathbb{Z}}$ of the cumulative point process. This construction implies that the point interval containing the origin has the length weighted distribution, as expected.

The following lemma is stated in terms of the scaled unit-intensity cumulative process X^u, but it is easily seen that it might as well have been formulated in terms of the process X with intensity $1/t$.

Lemma A.7.3. *Let $n \in \mathbb{N}$ be given, and let the unit-intensity process X^u be from the cumulative model such that $\mathbb{E}\mathrm{e}^{\eta\omega_1} < \infty$ for some $\eta > 0$, and such that Assumption A.2.1 is satisfied. Then*

$$H_m^u(s) = O(\mathrm{e}^{-\epsilon s}), \qquad s \to \infty, \tag{A.7.14}$$

for some $\epsilon > 0$. In particular, X^u and X are strongly n-admissible.

Proof. The admissibility claim obviously follows from (A.7.14).

In the case of the trapezoidal estimator we can state the theorem without the integrability assumption (A.2.2). This is because finite moments of the Peano kernel only require (A.2.2) to be true for $j \in \mathbb{N}$. As we assume that the increments have exponential moments, they in particular have finite moments of any positive order, and hence, all integrability results of the Peano kernels apply.

The proof relies on exponential decays in renewal theory, and we refer to [1, Chapter V] for an introduction. Moreover, for fixed $n \in \mathbb{N}$ and all $m \in \mathbb{N}_0$, we will explicitly use the fact that $K_m(s)$ depends on n points of the underlying point process to each side of s.

Let $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$ be given. Let $N = (N(s))_{s\geq 0}$ be a pure renewal process with increments $\{\omega_i\}_{i\in\mathbb{N}}$, and let U be the corresponding renewal measure. Also, let $y_0 = 0$ and $y_i = \sum_{\ell=1}^{i} \omega_\ell$ for $i \in \mathbb{N}$, that is, $X^u \cap (0,\infty) = \{X_0 + y_i\}_{i\in\mathbb{N}_0}$. Then $y_i \sim F^{*i}$ for all $i \in \mathbb{N}$. Define $\psi : [0,\infty) \to \mathbb{R}$ by $\psi(s) = \mathbb{E}[K_m(s;N)\mathbf{1}_{y_{n-1}\leq s}] = \mathbb{E}_u^0[K_m(s)\mathbf{1}_{s_{n-1}\leq s}]$, and let $\tilde{\psi} : [0,\infty) \to \mathbb{R}$ be given by

$$\tilde{\psi}(s) = \mathbb{E}[K_m(s;N)\mathbf{1}_{y_{n-1}\leq s}\mathbf{1}_{y_n > s}] = \mathbb{E}_u^0[K_m(s)\mathbf{1}_{s_{n-1}\leq s}\mathbf{1}_{s_n>s}].$$

Then $\psi(s) = U * \tilde{\psi}(s)$. This can be seen by a renewal argument obtaining the renewal equation $\psi = \tilde{\psi} + \psi * F$, which has the desired solution. Another rather intuitive approach is to condition on the nth to last point of N prior to s happening at time $x \in [0,s)$, which has probability $U(\mathrm{d}x)\bar{F}*F^{*(n-1)}(s-x)$. Now we initialize a new independent pure renewal process at time x and we obtain, integrating over $[0,s]$,

$$\begin{aligned}\psi(s) &= \int_0^s \mathbb{E}[K_m(s-x;N) \mid y_{n-1} \leq s-x,\, y_n > s-x]\bar{F} * F^{*(n-1)}(s-x)U(\mathrm{d}x)\\ &= \int_0^s \tilde{\psi}(s-x)U(\mathrm{d}x)\\ &= U * \tilde{\psi}(s).\end{aligned}$$

The exponential moment assumption implies $\bar{F}(s) = O(\mathrm{e}^{-\eta s})$, which in turn implies that also $\bar{G}(s) = O(\mathrm{e}^{-\eta s})$, as $s \to \infty$. Moreover, $\overline{G * F^{*i}}(s) = O(\mathrm{e}^{-\eta s})$, $s \to \infty$, for all $i \in \mathbb{N}$. We consider

$$\mathbb{E}_u K_m(0)K_m(s) = \mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}\leq s} + \mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}>s},$$

and an application of Cauchy-Schwarz inequality yields

$$\begin{aligned}\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}>s} &\leq [\mathbb{E}_u K_m^2(s)K_m^2(0)]^{1/2}\mathbb{P}(X_0 + y_{2n-2} > s)^{1/2}\\ &\leq C\left(\overline{G * F^{*(2n-2)}}(s)\right)^{1/2}\end{aligned}$$

for some finite C. Hence $\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}>s} = O(\mathrm{e}^{-\eta s/2})$ as $s \to \infty$, and (A.7.14) therefore follows once we show that

$$\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}\leq s} = (\mathbb{E}_u K_m(0))^2 + O(\mathrm{e}^{-\epsilon s}) \tag{A.7.15}$$

for some $\epsilon > 0$, as $s \to \infty$. We apply a renewal argument conditioning on the nth arrival in $X^u \cap (0,\infty)$, that is, conditioning on the value of $X_0 + y_{n-1} \sim G * F^{*(n-1)}$, and then initializing a new independent pure renewal process,

$$\begin{aligned}&\mathbb{E}_u K_m(s)K_m(0)\mathbf{1}_{X_0+y_{2n-2}\leq s}\\ &\quad= \int_0^s \mathbb{E}_u\Big[K_m(0)K_m(s)\mathbf{1}_{X_0+y_{2n-2}\leq s} \mid X_0 + y_{n-1} = v\Big]\Big(G * F^{*(n-1)}\Big)(\mathrm{d}v)\\ &\quad= \int_0^s \mathbb{E}\Big[K_m(s-v;N)\mathbf{1}_{y_{n-1}\leq s}\Big]\mathbb{E}_u\Big[K_m(0) \mid X_0 + y_{n-1} = v\Big]\Big(G * F^{*(n-1)}\Big)(\mathrm{d}v)\\ &\quad= \int_0^s \psi(s-v)\mathbb{E}_u\Big[K_m(0) \mid X_0 + y_{n-1} = v\Big]\Big(G * F^{*(n-1)}\Big)(\mathrm{d}v).\end{aligned}$$

Since $\mathbb{E}_u K_m(0) = \mathbb{E}_u^0 J_{K_{m+1}}(0)$ due to (A.4.3), an application of Fubini's theorem yields

$$\begin{aligned}\int_0^\infty \tilde{\psi}(s)\mathrm{d}s &= \mathbb{E}\int_{y_{n-1}}^{y_n} K_m(s;N)\mathrm{d}s \\ &= \mathbb{E}K_{m+1}(y_{n-1}^+;N) - \mathbb{E}K_{m+1}(y_n^-;N) = \mathbb{E}_u K_m(0),\end{aligned}$$

and consequently, by [1, Chapter VII: Thm. 2.10(iii)],

$$\psi(s) = U * \tilde{\psi}(s) = \mathbb{E}_u K_m(0) + O(\mathrm{e}^{-\epsilon' s}) \tag{A.7.16}$$

for some $0 < \epsilon' < \eta$, as $s \to \infty$. Furthermore, by another application of Cauchy-Schwarz inequality, we conclude that

$$\begin{aligned}\int_0^s \mathbb{E}_u\Big[K_m(0) \mid X_0 + y_{n-1} = v\Big]\Big(G * F^{*(n-1)}\Big)(\mathrm{d}v) \\ = \mathbb{E}_u K_m(0) - \mathbb{E}_u K_m(0)\mathbf{1}_{X_0+y_{n-1}>s} \\ = \mathbb{E}_u K_m(0) + O(\mathrm{e}^{-\eta s/2})\end{aligned} \tag{A.7.17}$$

as $s \to \infty$. Combining (A.7.16) and (A.7.17) yields (A.7.15). □

As for the perturbed model, Theorem A.2.10 is stated without Assumption A.2.1. This is because the strong admissibility and the variance decomposition are satisfied for the trapezoidal rule, when assuming (A.2.2) for $j \in \mathbb{N}$ only. This relaxed assumption is ensured by the finite exponential moments of the increments.

Proof of Theorem A.2.10. Let $m \in \{0,1\}$. As X is strongly admissible, we find by Proposition A.6.1 in combination with the decrease rate (A.7.14) that the variance decompose as

$$\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = (-1)^{m+1} t^{2m+2} 2g^{(2m+1)}(0^+) H_m^u(0) + o(t^{2m+2}).$$

From (A.7.8) and (A.7.9) and the fact that $\mathbb{E}_u^0 h_1^j = \nu_j$, we conclude that (A.2.12) and (A.2.13) are satisfied. □

A.8 An application in stereology

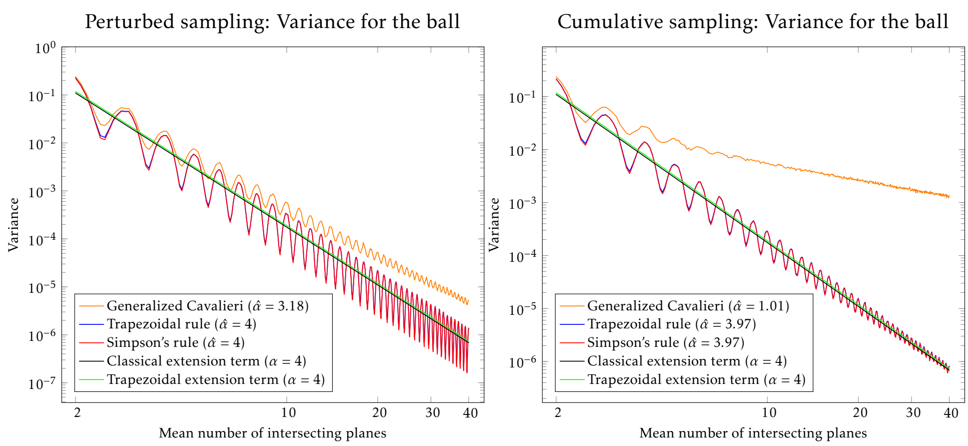
To illustrate the general theory, we describe a geometric application that also was the original motivation for this work. In stereology, the volume of a compact object $Y \subset \mathbb{R}^3$ can be approximated from sections with equidistant and parallel planes with joint normal direction v in the unit sphere S^2, if the area of each intersection profile is accessible; see [3, Chap. 7].

Formally, if $f(x)$ is the area of the intersection of Y with the plane $\{y \in \mathbb{R}^3 : v^T y = x\}$ positioned at a signed distance $x \in \mathbb{R}$ from the origin along v, the integral $\int f \mathrm{d}x$ coincides with the volume of Y by Fubini's theorem. If $f(x)$ is available at all points of the equidistant stationary point process $X = t(U + \mathbb{Z})$, the volume-estimator (A.1.2) can be used and is called the *(classical) Cavalieri estimator*. When $f(x)$ is known at the points of a stationary point process X with intensity $1/t$, the so-called *(generalized) Cavalieri estimator* (A.1.1) can be used. However, as outlined above, the generalized Cavalieri estimator does not exploit the smoothness of f and thus has a suboptimal decrease rate as $t \downarrow 0$. We therefore suggest employing Newton-Cotes estimators (of appropriate order) instead.

This section is devoted to Monte Carlo simulations illustrating the advantage of the new estimators when the sampling points are not equidistant. For illustration purposes we start considering the Euclidean unit ball $Y = \{z \in \mathbb{R}^3 : \|z\| \le 1\}$. In this case, the measurement function is

$$f(x) = \mathbf{1}_{[-1,1]}(x)\pi(1-x^2),$$

which is a $(1,\infty)$-piecewise smooth function as f' has jumps and is piecewise linear. Applying the classical Cavalieri estimator to such a function yields the extension term $\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = \frac{\pi^2}{90}t^4$ due to (A.2.5). Using sampling by the perturbed model or the model with cumulative errors we expect that the generalized Cavalieri estimator decreases at a rate of 3 and 1, respectively, whereas the trapezoidal estimator ($n = 1$) and Simpson's estimator ($n = 2$) decreases at a rate of 4 in both point-models, an asymptotic behavior visible in Figure 2 below. It shows the empirical variances of those three estimators based on 2000 Monte Carlo simulations as functions of the mean number of sections, that is $2/t$, with the variance plot including the extension term of the classical Cavalieri estimator and the extension term of the trapezoidal estimator as given by the dominating terms in (A.2.9) and (A.2.13) for the perturbed and cumulative model, respectively. The variances in this and the following figures are shown in a double-logarithmic scale with α and $\hat{\alpha}$ being the theoretical and approximate rates of decrease ($\hat{\alpha}$ has been found by the least squares method applied to the datapoints where $15 \le 2/t \le 40$).



(a) Perturbed sampling **(b)** Cumulative sampling

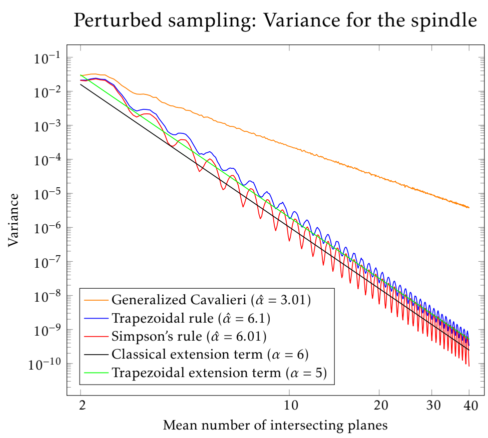
Figure 2: Empirical variance for the volume estimation of the unit ball in $\mathbb{R}^3$ based on perturbed sampling with $E_i \sim \mathrm{Unif}((-s,s))$ and sampling with cumulative errors with $\omega_i \sim \mathrm{Unif}((1-c,1+c))$. We choose s and c such that the average relative deviation (the coefficient of error) of the point-increment from the ideal increment 1 is 5%. In both figures, the graph of the trapezoidal estimator (blue) is almost completely hidden by the graph of Simpson's estimator (red), and the trapezoidal extension term (green) is almost identical to the classical extension term (black).

The graphs of Figure 2 are characteristic for the behavior of variances and extensions terms for objects with $(1,1)$-piecewise smooth measurement functions. For instance ellipsoids, or, more generally, strictly convex bodies lead to the same variance behavior apart from the facts that intercepts of these curves may be shifted and the Zitterbewegung may differ.

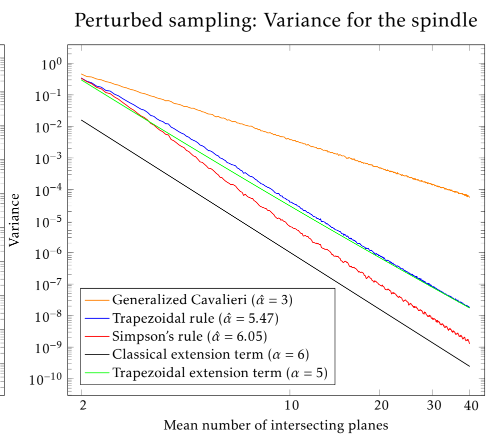
For comparison, we therefore give another example, where the measurement function exhibits a higher order of smoothness. The measurement function

$$f(x) = \mathbf{1}_{[-1,1]}(x)\frac{\pi}{2}(1+\cos\pi x),$$

is obtained from a spindle shaped body of revolution, if all section planes are orthogonal to the rotation axis. The corresponding convex body is illustrated in [6, Fig. 4]. The measurement function f is $(2,\infty)$-piecewise smooth. Using this measurement function, the extension term of the classical estimator is $\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = \frac{\pi^6}{60480}t^6$. Figure 3 shows empirical variances based on perturbed sampling with the two extension terms included, where the extension term of the trapezoidal estimator is given as the sum of the dominating terms in (A.2.10) and (A.2.11). In Figure 3a we use small perturbations to illustrates the fact that the dominating term in (A.2.10) can be made arbitrarily small. Hence, a decrease rate of 6 for the variance of the trapezoidal estimator can be a good approximation with small perturbations, as the trapezoidal extension term is approximately given by $1.7\cdot 10^{-4}t^5 + 3.0\cdot 10^{-2}t^6$ here. Even when we consider $100 \le 2/t \le 200$, we only obtain an approximate decrease rate of $\hat{\alpha} = 5.66$. For comparison, Figure 3b gives a better illustration of the actual asymptotic rate of decrease which corresponds to the bound from Theorem A.2.8, that is, $\alpha = 5$. Here we use larger perturbations, which in turn gives an approximate trapezoidal extension term of $0.043t^5 + 0.25t^6$. Increasing the number of intersecting planes to $2/t \le 100$ the actual rate is even more apparent, as we here obtain an approximate decrease rate of $\hat{\alpha} = 5.19$.



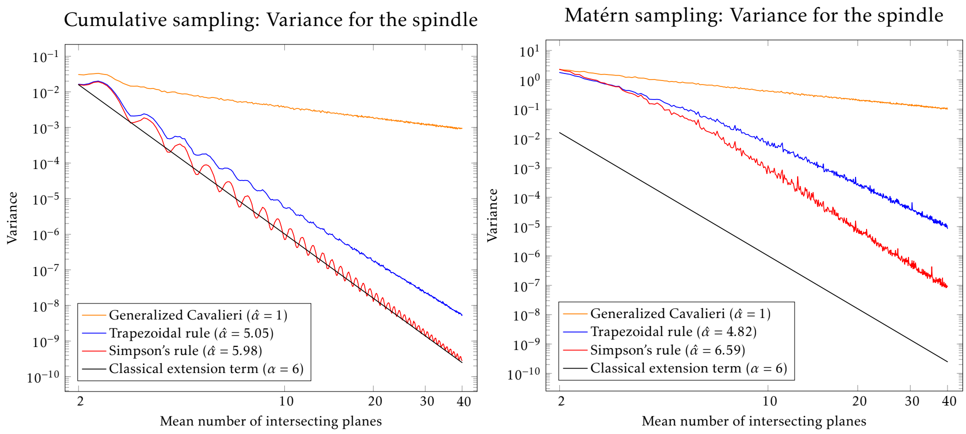
(a) 5% CE (small perturbations)



(b) 20% CE (large perturbations)

Figure 3: Empirical variance for the volume estimation of a spindle shaped body of revolution in $\mathbb{R}^3$ based on perturbed sampling with $E_i \sim \mathrm{Unif}((-s,s))$. We choose s such that the coefficient of error (CE) of the point-increments are 5% (left) and 20% (right).

The last two simulations are meant to illustrate the findings in Theorem A.2.8 for point process models where we do not have explicit formulae for the extension term. The first is the already discussed model with accumulated errors. To illustrate the wide range of point process models to which our results apply, we also simulated from the Matérn hard core process of type II; see [10, sec. 3.5 pp. 93-94], which satisfies the strong integrability assumption (A.2.2). The empirical variances for the aforementioned spindle shaped body are depicted in Figure 4. It is worth noticing that the variance of the trapezoidal estimator under the Matérn model seem to satisfy the strong bound of Theorem A.2.8, that is (A.2.7). Increasing the number of intersecting planes to $2/t \leq 100$ the result is more clear, as we find approximate decrease rates of $\hat{\alpha} = 4.94$ and $\hat{\alpha} = 6.07$ for the trapezoidal estimator and Simpson's estimator, respectively.



(a) Cumulative sampling **(b)** Matérn sampling

Figure 4: Empirical variance for the volume estimation of a spindle shaped body in $\mathbb{R}^3$ based on sampling with cumulative errors with $\omega_i \sim \text{Unif}((1-c, 1+c))$ and sampling with a Matérn hard core process of type II with intensity 1 and a hard core distance of 0.4. The value of c is chosen such that the coefficient of error of the increment is 5%.

A.9 Conclusions and future work

Estimating integrals based on known randomized sampling points with unequal increments, we have shown that higher order Newton-Cotes quadratures are to be preferred over naïve Riemann sums, as they are unbiased and have a faster decrease in variance for decreasing average point-increment. In particular, if the measurement function is exactly $(n,1)$-piecewise smooth, applying nth order Newton-Cotes estimation yields an upper bound of the variance decreasing at the same rate as the variance based on equidistant sampling, that is, a rate of $2n+2$. Applying nth order estimation to a function with smoothness of order, say, $m > n$, the variance has an upper bound with a rate of decrease of $2n+2$ in the general case, whereas the bound decreases at the rate $2n+3$ if the set of sampling points is *strongly n-admissible*. We have shown that point processes from the perturbed and cumulative models are strongly admissible and thus the strong bound holds in these cases. Based on a simulation study of the trapezoidal estimator it appears that also sampling from Matérn's hard core model of the second kind satisfies the strong bound. From a practical point of view the trapezoidal estimator is very interesting as the unbiasedness does not require any integrability conditions of the underlying sampling model. Applying this estimator to perturbed and cumulative sampling we have found asymptotic variance expressions, with an overall trend arbitrarily close to the trend of the equidistant case if the perturbations are small and the increments are close to 1, respectively. This asymptotic trend can be calculated if only the derivatives of the covariogram of the measurement function is known at 0, and if moments of the perturbations and increments, respectively, can be computed. This observation allows in principle to estimate the extension term of the variance from measurements of sampling positions and sampled areas in analogy to established methods in the classical, equidistant case. We intend to carry out this program in a future study.

It is an open question if the variance bounds in Theorem A.2.8 are optimal in all cases. As the rate of decrease in (A.2.6) is optimal if the model is weakly admissible or the order of the estimator exceeds the order of smoothness of the measurement function, we expect that the rate in (A.2.6) is optimal for any stationary point process satisfying the assumptions of the theorem. Similarly we know that the bound presented in (A.2.7) yields the optimal decay-rate when $n = 1$ under the perturbed model (assuming non-degenerate perturbations), and thus it is of interest to investigate whether this is the case for all n in perturbed sampling and in general for any admissible point process with unequal increments.

Acknowledgments

This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

Supplementary material

A.I Integrability properties

As mentioned in Section A.4 the weight $\beta_j^{(n)}(x)$ is particularly simple.

Lemma A.I.1. *For all $n \in \mathbb{N}$, $x \in X$ and $j = 0,\dots,n$, the weight $\beta_j^{(n)}(x)$ is a rational function of point-increments,*

$$\beta_j^{(n)}(x) = \frac{q_j^{(n)}(h_1(x),\dots,h_n(x))}{p_j^{(n)}(h_1(x),\dots,h_n(x))}$$

where $q_j^{(n)} : (0,\infty)^n \to \mathbb{R}$ is a homogeneous polynomial of degree $n+1$, and $p_j^{(n)} : (0,\infty)^n \to \mathbb{R}$ is a non-vanishing homogeneous polynomial of degree n with non-negative coefficients.

Proof. Fix $x \in X$, $n \in \mathbb{N}$ and $j \in \{0,\dots,n\}$, and consider $\beta_j^{(n)}(x)$ as defined by (A.3.3). Recall that points in X are distinct and therefore all point-increments are strictly positive. At first we note that the denominator of the integrand in (A.3.3) is constant with each term in the product satisfying

$$s_j(x) - s_k(x) = \begin{cases} \sum_{\ell=k+1}^{j} h_\ell(x) & \text{for } j > k, \\ -\sum_{\ell=j+1}^{k} h_\ell(x) & \text{for } j < k, \end{cases}$$

and hence

$$\prod_{\substack{k=0 \\ k\neq j}}^{n} \Big(s_j(x) - s_k(x)\Big) = (-1)^{n-j} p_j^{(n)}(h_1(x),\dots,h_n(x)),$$

where $p_j^{(n)} : (0,\infty)^n \to \mathbb{R}$ is the polynomial defined by

$$p_j^{(n)}(y_1,\dots,y_n) = \Big(\prod_{k=0}^{j-1} \sum_{\ell=k+1}^{j} y_\ell\Big)\Big(\prod_{k=j+1}^{n} \sum_{\ell=j+1}^{k} y_\ell\Big).$$

The definition of $p_j^{(n)}$ implies that it is non-vanishing with non-negative coefficients and that $p_j^{(n)}(\lambda y_1,\dots,\lambda y_n) = \lambda^n p_j^{(n)}(y_1,\dots,y_n)$ for any $\lambda \in (0,\infty)$. That is, it is homogeneous of degree n.

With the abbreviation $\tilde{s}_k(x) = s_k(x) - x = \sum_{\ell=1}^{k} h_\ell(x)$, a substitution yields

$$\int_x^{s_n(x)} \prod_{\substack{k=0 \\ k\neq j}}^{n} (y - s_k(x))\,\mathrm{d}y = \int_0^{\tilde{s}_n(x)} \prod_{\substack{k=0 \\ k\neq j}}^{n} (y - \tilde{s}_k(x))\,\mathrm{d}y,$$

for $k \geq 0$. The right side of this equation is a polynomial of degree at most $n+1$ in $(\tilde{s}_0(x),\dots,\tilde{s}_n(x))$, as all its derivatives of order $n+2$ vanish. We therefore can define the polynomial $q_j^{(n)} : (0,\infty)^n \to \mathbb{R}$ by

$$q_j^{(n)}(h_1(x),\dots,h_n(x)) = (-1)^{n-j} \int_0^{\tilde{s}_n(x)} \prod_{\substack{k=0 \\ k\neq j}}^{n} (y - \tilde{s}_k(x))\,\mathrm{d}y.$$

A substitution argument shows that the right side is homogeneous of degree $n+1$ as a function of $(\tilde{s}_0(x),\dots,\tilde{s}_n(x))$ and thus also as a function of $(h_1(x),\dots,h_n(x))$. This shows the assertion. □

Assuming either Assumption A.2.1 or Assumption A.4.1, this representation ensures the Palm integrability of $\alpha(0)$, which is used in the proof of Theorem A.2.2.

Lemma A.I.2. *Fix $n \in \mathbb{N}$. If X is a stationary point process such that* (A.4.1) *is satisfied, then*

$$\mathbb{E}^0|\beta_j^{(n)}(s_{-j})| < \infty \tag{A.I.1}$$

for all $j = 0,\dots,n$, and consequently $\mathbb{E}^0|\alpha(0)| < \infty$.

Proof. From Lemma A.I.1 we find real constants $\{c_{\mathbf{m}}^{(n,j)}\}$ and non-negative constants $\{a_{\mathbf{m}'}^{(n,j)}\}$ such that

$$|\beta_j^{(n)}(s_{-j})| = \frac{\left|\sum_{\substack{\mathbf{m}\in\mathbb{N}_0^n\\ |\mathbf{m}|=n+1}} c_{\mathbf{m}}^{(n,j)}\,\mathbf{h}(s_{-j})^{\mathbf{m}}\right|}{\sum_{\substack{\mathbf{m}'\in\mathbb{N}_0^n\\ |\mathbf{m}'|=n}} a_{\mathbf{m}'}^{(n,j)}\,\mathbf{h}(s_{-j})^{\mathbf{m}'}} \le \sum_{\substack{\mathbf{m}\in\mathbb{N}_0^n\\ |\mathbf{m}|=n+1}} \frac{|c_{\mathbf{m}}^{(n,j)}|\,\mathbf{h}(s_{-j})^{\mathbf{m}}}{a_{\mathbf{m}_0'}^{(n,j)}\,\mathbf{h}(s_{-j})^{\mathbf{m}_0'}},$$

where $\mathbf{m}_0'$ is a multi-index such that $a_{\mathbf{m}_0'}^{(n,j)} > 0$ which exists by Lemma A.I.1. By linearity (A.I.1) is satisfied whenever

$$\mathbb{E}^0\left[\frac{\mathbf{h}(s_{-j})^{\mathbf{m}}}{\mathbf{h}(s_{-j})^{\mathbf{m}'}}\right] = \mathbb{E}^0\left[\frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{h}^{\mathbf{m}'}}\right] < \infty, \tag{A.I.2}$$

for all multi-index $\mathbf{m}$ and $\mathbf{m}'$ in $\mathbb{N}_0^n$ with $|\mathbf{m}| = n+1$ and $|\mathbf{m}'| = n$, where the equality is a consequence of the fact that the Palm distribution is invariant under bijective point shifts; see [6, Eq. (13)]. The right side of (A.I.2) is finite by Assumption A.4.1. □

A.II Peano kernels, Bernoulli functions and variance in perturbed sampling

In this section we consider the relation between the Peano kernels K_m and the Bernoulli functions P_m when we sample $X^u = \{U + E_k + k\}_{k\in\mathbb{Z}}$ from the perturbed model (recall that X^u is scaled to have unit-intensity). Note that the unit-intensity equidistant model is obtained with degenerate perturbations concentrated at 0. As in Section A.7 we work with the shifted kernel, K_m^*, defined by

$$K_m^*(r) = K_m(r+U) = K_m(r;X^*),$$

where $X^* = X^u - U = \{E_k + k\}_{k\in\mathbb{Z}}$ is the shifted process. From (A.7.3) the shifted kernel is periodic in law with period 1. Recall that the 1st Bernoulli function is given by $P_1(r) = \tilde{P}_1(r - \lfloor r \rfloor)$, with $\tilde{P}_1(r) = r - \frac{1}{2}$ for $r \in \mathbb{R}$.

Lemma A.II.1. *Let $n \in \mathbb{N}$ be given and let X^u be a unit-intensity process from the perturbed model such that* (A.2.2) *is satisfied. Let $X^* = \{x_k\}_{k\in\mathbb{Z}}$, with $x_k = E_k + k$, be its shifted process. For all $r \in \mathbb{R}$, K_0^* satisfies*

$$\mathbb{E}K_0^*(r) = -\mathbb{E}P_1(E_0 - r) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_0)j\Big] - \frac{n}{2} + Q(r), \tag{A.II.1}$$

where

$$Q(r) = \begin{cases} \mathbb{E}\mathbf{1}_{E_0\geq r}[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1] & \text{for} \quad r - \lfloor r\rfloor < \frac{1}{2}, \\ \mathbb{E}\mathbf{1}_{E_0\geq r-1}[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1] & \text{for} \quad r - \lfloor r\rfloor \geq \frac{1}{2}. \end{cases}$$

Furthermore, if $\mathbb{E}K_0^*(r) = -\mathbb{E}P_1(E_0 - r)$ for all $r \in \mathbb{R}$, then

$$H_m^u(r) = (-1)^m\mathbb{E}[P_{2m+2}(r + E_1 - E_0)] \tag{A.II.2}$$

for $m \leq n$ and all $|r| \geq 2n+2$. If the perturbations are degenerate, that is X^u is the unit-intensity equidistant model, (A.II.2) *is true for all $r \in \mathbb{R}$.*

Proof. By (A.7.3) it is enough to consider $r \in [0,1)$. Let $n \in \mathbb{N}$ and $r \in [0,1)$ be given. Recall that

$$nK_0^*(r) = \sum_{i\in\mathbb{Z}}\mathbf{1}_{x_i<r\leq x_{i+1}}\sum_{\ell=1-n}^{0}R_{x_{i+\ell}}^{(n)}((\cdot - r)_+^0).$$

Only the summands with $i = -1, 0, 1$ can be non-zero, and thus

$$nK_0^*(r) = \mathbf{1}_{E_0\geq r}A_{-1}(r) + \mathbf{1}_{E_0<r}\mathbf{1}_{E_1\geq r-1}A_0(r) + \mathbf{1}_{E_1<r-1}A_1(r),$$

where, for $i = -1, 0, 1$,

$$A_i(r) = \sum_{\ell=1-n}^{0}R_{x_{i+\ell}}^{(n)}((\cdot - r)_+^0) = \sum_{\ell=1-n}^{0}\sum_{j=0}^{n}\beta_j^{(n)}(x_{i+\ell})\mathbf{1}_{\ell+j\geq 1} - \sum_{\ell=1}^{n}(x_{i+\ell} - r).$$

We let q_0 and q_1 be the i.i.d. variables defined by $q_0 = (E_0 - r) - \lfloor E_0 - r\rfloor$ and $q_1 = (E_1 - r) - \lfloor E_1 - r\rfloor$. We will consider the cases $r < \frac{1}{2}$ and $r \geq \frac{1}{2}$ separately.

Let $r < \frac{1}{2}$ be given. As $E_1 \geq r - 1$ the kernel simplifies as

$$nK_0^*(r) = \mathbf{1}_{E_0\geq r}A_{-1}(r) + \mathbf{1}_{E_0<r}A_0(r).$$

Note that $q_0 = E_0 - r$ when $E_0 \geq r$, and $q_0 = E_0 - r + 1$ when $E_0 < r$. Using the independence of the perturbations, $\mathbb{E}E_i = 0$, and the representation of the second power sum, we find

$$\mathbb{E}\mathbf{1}_{E_0\geq r}A_{-1}(r) = \mathbb{E}\mathbf{1}_{E_0\geq r}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_{\ell-1}) - n\tilde{P}_1(q_0) + (n-1)E_0 - \frac{n^2}{2}\Big),$$

$$\mathbb{E}\mathbf{1}_{E_0<r}A_0(r) = \mathbb{E}\mathbf{1}_{E_0<r}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_\ell) - n\tilde{P}_1(q_0) + nE_0 - \frac{n^2}{2}\Big).$$

Constant functions are approximated exactly, and hence $\sum_{j=0}^{n} \beta_j^{(n)}(x_0) = x_n - x_0 = E_n - E_0 + n$. An index-shift in the former term above then implies

$$\begin{aligned}\mathbb{E}K_0^*(r) &= -\mathbb{E}\tilde{P}_1(q_0) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_\ell)\Big]\\ &\quad -\frac{n}{2} + \mathbb{E}\mathbf{1}_{E_0\geq r}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1\Big]\\ &= -\mathbb{E}P_1(E_0 - r) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_0)j\Big]\\ &\quad -\frac{n}{2} + \mathbb{E}\mathbf{1}_{E_0\geq r}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1\Big],\end{aligned}$$

where the last equality follows as $\beta_j^{(n)}(x_\ell)$ equals $\beta_j^{(n)}(x_0)$ in law, as they are rational functions of identically distributed increments.

Now let $r \geq \frac{1}{2}$ be given. Then $E_0 < r$ and the kernel simplifies as

$$nK_0^*(r) = \mathbf{1}_{E_1\geq r-1}A_0(r) + \mathbf{1}_{E_1<r-1}A_1(r).$$

Note that $q_1 = E_1 - r + 1$ when $E_1 \geq r - 1$, and $q_1 = E_0 - r + 2$ when $E_1 < r - 1$. With similar arguments as above we find that

$$\begin{aligned}\mathbb{E}\mathbf{1}_{E_1\geq r-1}A_0(r) &= \mathbb{E}\mathbf{1}_{E_1\geq r-1}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_\ell) - n\tilde{P}_1(q_1) + (n-1)E_1 - \frac{n^2}{2}\Big),\\ \mathbb{E}\mathbf{1}_{E_1<r-1}A_1(r) &= \mathbb{E}\mathbf{1}_{E_1<r-1}\Big(\sum_{j=0}^{n}\sum_{\ell=1-j}^{0}\beta_j^{(n)}(x_{\ell+1}) - n\tilde{P}_1(q_1) + nE_1 - \frac{n^2}{2}\Big).\end{aligned}$$

By the i.i.d. property of the perturbations and the exact arguments as above we conclude that

$$\begin{aligned}\mathbb{E}K_0^*(r) &= -\mathbb{E}P_1(E_0 - r) + \mathbb{E}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_0)j\Big]\\ &\quad -\frac{n}{2} + \mathbb{E}\mathbf{1}_{E_0\geq r-1}\Big[\frac{1}{n}\sum_{j=0}^{n}\beta_j^{(n)}(x_{-j}) - 1\Big]\end{aligned}$$

when $r \geq \frac{1}{2}$. This proves the first part of the lemma.

To show (A.II.2), we note that

$$\mathbb{E}K_m^*(r) - \mathbb{E}_u K_m(0) = -\mathbb{E}P_{m+1}(E_0 - r) \tag{A.II.3}$$

for all $r \in \mathbb{R}$. This is seen by induction using Fubini's theorem, the relations $P_m' = P_{m-1}$ and $K_m' = -K_{m-1}$, the fact that $\mathbb{E}_u K_m(0) = 0$ for all $m < n$ (see Lemma A.4.2), and the continuity properties of the kernels and polynomials. For $|r| \geq 2n+2$, the perturbations in $K_m(r; X^u) = K_m^*(r - U)$ and $K_m(0; X^u) = K_m^*(-U)$ are independent. With $\mathbb{E}_U$, $\mathbb{E}_{X^*}$ and $\mathbb{E}_{E_0,E_1}$ denoting the expectations with respect to the given variables, (A.II.3) and independence then implies

$$\begin{aligned}H_m^u(r) &= \mathbb{E}_U \mathbb{E}_{X^*}[K_m^*(r-U) - \mathbb{E}_u K_m(0)] \mathbb{E}_{X^*}[K_m^*(-U) - \mathbb{E}_u K_m(0)] \\ &= \mathbb{E}_{E_0,E_1} \mathbb{E}_U [P_{m+1}(U+E_0-r) P_{m+1}(U+E_1)] \\ &= (-1)^m \mathbb{E}[P_{2m+2}(r+E_1-E_0)],\end{aligned} \tag{A.II.4}$$

where the last equality is shown in the proof of [7, Prop. 5.2]. This shows (A.II.2). If the model has degenerate perturbations concentrated at 0, (A.II.4) is true for all $r \in \mathbb{R}$ with $X^* = \mathbb{Z}$ deterministic. This concludes the proof. □

Corollary A.II.2. *Let $n \in \mathbb{N}$ be given. If $X^u = U + \mathbb{Z}$ is the unit-intensity equidistant model, then*

$$H_m^u(r) = (-1)^m P_{2m+2}(r)$$

for $m \leq n$ and all $r \in \mathbb{R}$.

Proof. Fix $n \in \mathbb{N}$. Note that $X^* = \mathbb{Z}$ and therefore it is deterministic. We have from Lemma A.II.1 that it suffices to show $K_0^*(r) = -P_1(-r)$ for $r \in [0,1)$. Also, the weights $\beta_j^{(n)}(x)$ do not depend on $x \in X^*$, and we therefore denote the common weights by $\beta_j^{(n)}$. As polynomials of degree 1 are approximated exactly, we find that

$$\frac{1}{n}\sum_{j=0}^{n} \beta_j^{(n)} = 1 \qquad \text{and} \qquad \frac{1}{n}\sum_{j=0}^{n} \beta_j^{(n)} j = \frac{n}{2}.$$

Returning to (A.II.1), we conclude that $K_0^*(r) = -P_1(-r)$. □

Corollary A.II.3. *Let $n = 1$. If X^u is from the unit-intensity perturbed model, then*

$$H_m^u(r) = (-1)^m \mathbb{E}[P_{2m+2}(r + E_1 - E_0)]$$

for $m \in \{0,1\}$ and all $|r| \geq 4$.

Proof. Since $\beta_0^{(1)}(x) = \beta_1^{(1)}(x) = \frac{1}{2} h_1(x)$, $x \in X^u$, holds for all point processes X^u, it is easily seen that $\mathbb{E}\beta_1^{(n)}(E_0) = \frac{1}{2}$ and $Q(r) = 0$. The result follows from (A.II.1). □

References

[1] Asmussen, S. (2003). *Applied Probability and Queues*. Applications of Mathematics: Stochastic Modelling and Applied Probability. Springer, New York.

[2] Baddeley, A.J., K.-A. Dorph-Petersen and E.B.V. Jensen (2006). A note on the stereological implications of irregular spacing of sections. *J. Microsc.* 222(3), 177–181.

[3] Baddeley, A.J. and E.B.V. Jensen (2004). *Stereology for Statisticians*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press, Boca Raton.

[4] De Barra, G. (1981). *Measure Theory and Integration*. New Age International, New Delhi.

[5] Doukhan, P. (1994). *Mixing: Properties and Examples*. Lecture Notes in Statistics. Springer, New York.

[6] Kiderlen, M. and K.-A. Dorph-Petersen (2017). The Cavalieri estimator with unequal section spacing revisited. English. *Image Anal. Stereol.* 36(2), 135–141.

[7] Kiêu, K. (1997). *Three Lectures on Systematic Geometric Sampling*. Technical Report, Aarhus University.

[8] Knopp, K. (1947). *Theorie und Anwendung der unendlichen Reihen*. Die Grundlehren der mathematischen Wissenschaften. Springer, Berlin.

[9] Matheron, G. (1965). *Les Variables régionalisées et leur Estimation*. Masson et Cie, Paris.

[10] Schneider, R. and W. Weil (2008). *Stochastic and Integral Geometry*. Springer, Heidelberg.

[11] Stoer, J. and R. Bulirsch (1993). *Introduction to Numerical Analysis*. Trans. by R. Bartels, W. Gautschi and C. Witzgall. Texts in Applied Mathematics. Springer, New York.

[12] Ziegel, J., A.J. Baddeley, K.-A. Dorph-Petersen and E.B.V. Jensen (2010). Systematic sampling with errors in sample locations. *Biometrika* 97, 1–13.

A note to Paper A

This appendix is devoted to the proof of equation (A.2.11) of (Theorem A.2.9, Paper A); a proof, which was not included in the original paper due to its length. It turns out that (A.2.11) holds true under weaker assumptions than given in Theorem A.2.9, and we therefore restate the theorem in full after which we prove the expression for the remainder under the weaker assumptions. This appendix should be read in the context of Paper A, and we thus refer to Paper A for definitions, descriptions of notation, formulas, etc.

AA.1 Proof of (A.2.11) under weaker assumptions

In the formulation of Theorem A.2.9 concerning the variance of $\hat{V}_1$ under the perturbed model, the expression (A.2.11) (or (AA.1.2) below) for the remainder holds true under the additional assumption that the perturbations have a density with finitely many finite jumps. This is due to the fact that we initially used a relation obtained in [3], where, under the mentioned assumption, the difference between $\sum_{k\in\mathbb{Z}} g(tk + t(E_k - E_0))$ and the integral $\int g\,\mathrm{d}x$ were given in terms of the periodic Bernoulli functions P_m; see in particular [3, Eq. (A2)]. As seen below, this can be obtained without the extra assumption, and Theorem A.2.9 may thus be formulated as follows.

Theorem AA.1.1 (Theorem A.2.9 restated). *Let X be from the perturbed model with intensity $1/t$, and let μ_k be the kth moment of the perturbations E_i. Assume that the measurement function f is exactly $(m,1)$-piecewise smooth with covariogram $g = f * \check{f}$. Then, for $t \downarrow 0$,*

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = -t^2 g'(0^+)(\mu_2 + \tfrac{1}{6}) + Z_0(t) + o(t^2), && \textit{for } m = 0,\\
&\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^4 g^{(3)}(0^+)\tfrac{1}{12}(2\mu_2 + 2\mu_4 + \tfrac{1}{30}) + Z_1(t) + o(t^4), && \textit{for } m = 1,\\
&\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 g^{(4)}(0)\tfrac{1}{8}(2\mu_4 + \mu_2\mu_4 - \mu_2^3 - \mu_3^2) + o(t^5), && \textit{for } m \geq 2,
\end{aligned} \tag{AA.1.1}$$

where the Zitterbewegung $Z_m(t)$ is given by (A.7.7). *It is of order t^{2m+2}, and it is a finite sum of terms oscillating around 0. Moreover, if $m \geq 2$, the remainder $o(t^5)$ is explicitly given by*

$$\begin{aligned}
t^6 g^{(5)}(0^+)\tfrac{1}{720}\Big(&-34\mu_2 - 90\mu_2^2 + 110\mu_4 + 180\mu_2\mu_4\\
&- 180\mu_2^3 - 170\mu_3^2 + 8\mu_6 - \tfrac{1}{21}\Big) + Z_2(t) + o(t^6).
\end{aligned} \tag{AA.1.2}$$

Proof of the remainder (AA.1.2). To ease notation, we will when possible consider X under its Palm-distribution, where we recall that the Palm-version X^0 of X is given by $X^0 = \{t(k + E_k - E_0)\}_{k\in\mathbb{Z}}$. When it is necessary or beneficial to use the exact representation of X^0 we will do so.

From Theorem A.2.2 we have that the trapezoidal estimator $\hat{V}_1(f) = \sum_{x\in X} \alpha(x)f(x)$, with $\alpha(x) = (h_1(x) + h_0(x))/2$, is unbiased for the integral of f, and consequently

$$\left(\mathbb{E}\hat{V}_1(f)\right)^2 = \int_{\mathbb{R}} g(x)\mathrm{d}x.$$

By the refined Campbell Theorem [2, Theorem 3.5.3] we therefore find that

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) &= \mathbb{E}\Big(\sum_{x\in X}\alpha(x)f(x)\Big)^2 - \int_{\mathbb{R}} g(x)\mathrm{d}x \\ &= \tfrac{1}{t}\sum_{k\in\mathbb{Z}}\mathbb{E}^0[\alpha(0)\alpha(s_k)g(s_k)] - \int_{\mathbb{R}} g(x)\mathrm{d}x.\end{aligned} \tag{AA.1.3}$$

First we investigate the difference

$$t\sum_{k\in\mathbb{Z}}\mathbb{E}^0 g(s_k) - \int_{\mathbb{R}} g(x)\mathrm{d}x \tag{AA.1.4}$$

be means of a refined Euler-MacLaurin formula. For all $k \neq 0$ we may write $\mathbb{E}^0 g(s_k) = \mathbb{E}g(tk + t(E_1 - E_0))$, and, noting that $\int_{\mathbb{R}} g(x)\mathrm{d}x = \mathbb{E}\int_{\mathbb{R}} g(x + t(E_1 - E_0))\mathrm{d}x$, equation (AA.1.4) reads

$$\begin{aligned}\mathbb{E}\Big[t\sum_{k\in\mathbb{Z}} g(tk + t(E_1 - E_0)) - \int_{\mathbb{R}} g(x + t(E_1 - E_0))\mathrm{d}x\Big] \\ + tg(0) - t\mathbb{E}[g(t(E_1 - E_0))].\end{aligned} \tag{AA.1.5}$$

Since g is $(2m+1,1)$-piecewise smooth by [1, Corollary 5.8], also the function $x \mapsto \tilde{g}(x) = g(x + t(E_1 - E_0))$ is $(2m+1,1)$-piecewise smooth. Since $\tilde{g}$ clearly depends on t, the refined Euler-MacLaurin formula [1, Proposition 4.2] cannot be used directly on the first term in (AA.1.5), however, following its proof we conclude that

$$\begin{aligned}t\sum_{k\in\mathbb{Z}}\mathbb{E}^0 g(s_k) - \int_{\mathbb{R}} g(x)\mathrm{d}x &= tg(0) - t\mathbb{E}[g(t(E_1 - E_0))] \\ &\quad - t^{2m+2}\sum_{a\in D_{g^{(2m+1)}}} J_{g^{(2m+1)}}(a)\mathbb{E}[P_{2m+2}(\tfrac{a}{t} + E_1 - E_0)] \\ &\quad - t^{2m+2}\int_{\mathbb{R}} g^{(2m+2)}(s)\mathbb{E}[P_{2m+2}(\tfrac{s}{t} + E_1 - E_0)]\mathrm{d}s.\end{aligned} \tag{AA.1.6}$$

By dominated convergence and arguments as in the proof of [1, Proposition 4.2], the last term above is seen to be of order $o(t^{2m+2})$ as $t \downarrow 0$.

Now we turn to the sum

$$\tfrac{1}{t}\sum_{k\in\mathbb{Z}}\mathbb{E}^0[(\alpha(0)\alpha(s_k) - t^2)g(s_k)] = t\sum_{k\in\mathbb{Z}}\mathbb{E}^0_u[(\alpha(0)\alpha(s_k) - 1)g(ts_k)],$$

where $\mathbb{E}^0_u$ denotes the expectation with respect to the Palm distribution of X/t, i.e. the expectation with respect to X^0/t. We see that $\alpha(0)$, $\alpha(s_k)$ and s_k are independent for $|k| \geq 3$, and since $\mathbb{E}^0_u\alpha(s_k) = 1$ for all k (seen directly or shown in the proof of Theorem A.2.2) the sum simplifies as

$$t\sum_{k=-2}^{2}\mathbb{E}^0_u[(\alpha(0)\alpha(s_k) - 1)g(ts_k)].$$

Since g is $(2m+1,1)$-piecewise smooth, we find for all k and sufficiently small t that

$$g(ts_k) = \sum_{j=0}^{m} \frac{g^{(2j)}(0)}{(2j)!}(ts_k)^{2j} + \frac{g^{(2m+1)}(\xi_{k,t})}{(2m+1)!}(ts_k)^{2m+1},$$

where $\xi_{k,t}$ is in the open interval with endpoints 0 and ts_k. Here it has been used that $g^{(j)}$ is odd and continuous for all odd $j \le 2m$, and hence $g^{(j)}(0) = 0$. Moreover, as $g^{(2m+1)}$ is odd and has a jump in 0 by [1, Corollary 5.8], we obtain $g^{(2m+1)}(0^+) = -g^{(2m+1)}(0^-)$. Hence, as $|s_k| = |k + E_k - E_0| \le |k| + 1$,

$$g(ts_k) = \sum_{j=0}^{m} \frac{g^{(2j)}(0)}{(2j)!} t^{2j} s_k^{2j} + \frac{g^{(2m+1)}(0^+)}{(2m+1)!} t^{2m+1} |s_k|^{2m+1} + o(t^{2m+1})$$

as $t \downarrow$ for all $|k| \le 2$. Similarly we see that

$$\begin{aligned} g(0) - \mathbb{E}g(t(E_1 - E_0)) = &-\sum_{j=0}^{m} \frac{g^{(2j)}(0)}{(2j)!} t^{2j} \mathbb{E}(E_1 - E_0)^{2j} \\ &- \frac{g^{(2m+1)}(0^+)}{(2m+1)!} t^{2m+1} \mathbb{E}|E_1 - E_0|^{2m+1} + o(t^{2m+1}) \end{aligned}$$

by dominated convergence. Let the Zitterbewegung be given by

$$Z_m(t) = -t^{2m+2} \sum_{a \in D_{g^{(2m+1)}} \setminus \{0\}} J_{g^{(2m+1)}}(a) \mathbb{E}[P_{2m+2}(\tfrac{a}{t} + E_1 - E_0)].$$

Combining (AA.1.3) and (AA.1.6), and using the fact that $J_{g^{(2m+1)}}(0) = 2g^{(2m+1)}(0^+)$, we now conclude by dominated convergence that

$$\begin{aligned} \mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = &\sum_{j=0}^{m} t^{2j+1} \frac{g^{(2j)}(0)}{(2j)!} \Big(\sum_{k=-2}^{2} \mathbb{E}_u^0[s_k^{2j}(\alpha(0)\alpha(s_k) - 1)] - \mathbf{1}_{j \ge 1} \mathbb{E}[(E_1 - E_0)^{2j}] \Big) \\ &+ t^{2m+2} \frac{g^{(2m+1)}(0^+)}{(2m+1)!} \Big(\sum_{k=-2}^{2} \mathbb{E}_u^0[|s_k|^{2m+1}(\alpha(0)\alpha(s_k) - 1)] \\ &- \mathbb{E}|E_1 - E_0|^{2m+1} - 2(2m+1)!\, \mathbb{E}[P_{2m+2}(E_1 - E_0)] \Big) + Z_m(t) + o(t^{2m+2}) \end{aligned} \tag{AA.1.7}$$

as $t \downarrow 0$. Using the fact that s_k can be represented as $k + E_k - E_0$, lengthy but otherwise elementary calculations show that

$$\sum_{k=-2}^{2} \mathbb{E}_u^0[s_k^{2j}(\alpha(0)\alpha(s_k) - 1)] - \mathbf{1}_{j \ge 1} \mathbb{E}[(E_1 - E_0)^{2j}] = 0$$

for $j \in \{0, 1\}$.

Now let $m = 2$. It can be seen that the 6th periodic Bernoulli function is given by

$$P_6(x) = \tfrac{1}{6!}\Big(x^6 - 3|x|^5 + \tfrac{5}{2}x^4 - \tfrac{1}{2}x^2 + \tfrac{1}{42}\Big)$$

for all $x \in [-1,1]$. Since $|E_i| < 1/2$ almost surely, we conclude that the variance is given by

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\hat{V}_1(f)) = t^5 \frac{g^{(4)}(0)}{4!}\Big(\sum_{k=-2}^{2} \mathbb{E}_u^0[s_k^4(\alpha(0)\alpha(s_k)-1)] - \mathbb{E}[(E_1-E_0)^4]\Big)\\ + t^6 \frac{g^{(5)}(0^+)}{5!}\Big(\sum_{k=-2}^{2} \mathbb{E}_u^0[|s_k|^5(\alpha(0)\alpha(s_k)-1)]\\ - \tfrac{1}{3!}\mathbb{E}\Big[(E_1-E_0)^6 + \tfrac{5}{2}(E_1-E_0)^4 - \tfrac{1}{2}(E_1-E_0)^2 + \tfrac{1}{42}\Big]\Big) + Z_2(t) + o(t^{2m+2})\end{aligned} \tag{AA.1.8}$$

as $t \downarrow 0$. Using the representation of s_k again, we see by lengthy but straightforward calculations that (AA.1.1) and (AA.1.2) follow from (AA.1.8).

The case $m > 2$ follows from (AA.1.7) and by the realization that $g^{(5)}$ is continuous and odd, hence $g^{(5)}(0^+) = 0$ and $Z_2(t) = 0$ for all t, and thus the remainder (AA.1.2) simply reads $o(t^6)$. □

References

[1] Kiêu, K. (1997). *Three Lectures on Systematic Geometric Sampling*. Technical Report, Aarhus University.

[2] Schneider, R. and W. Weil (2008). *Stochastic and Integral Geometry*. Springer, Heidelberg.

[3] Ziegel, J., A.J. Baddeley, K.-A. Dorph-Petersen and E.B.V. Jensen (2010). Systematic sampling with errors in sample locations. *Biometrika* 97, 1–13.

PAPER B

Improving the Cavalieri Estimator Under Non-Equidistant Sampling and Dropouts

Mads Stehr and Markus Kiderlen

Submitted to *Image Analysis and Stereology*

Abstract

Motivated by the stereological problem of volume estimation from parallel section profiles, the so-called Newton-Cotes integral estimators based on random sampling nodes are analyzed. These estimators generalize the classical Cavalieri estimator and its variant for non-equidistant sampling nodes, the generalized Cavalieri estimator, and have typically a substantially smaller variance than the latter. The present paper focuses on the following points in relation to Newton-Cotes estimators: the treatment of dropouts, the construction of variance estimators, and, finally, their application in volume estimation of convex bodies.

Dropouts are eliminated points in the initial stationary point process of sampling nodes, modeled by independent thinning. Among other things, exact representations of the variance are given in terms of the thinning probability and increments of the initial points under two practically relevant sampling models.

The paper presents a general estimation procedure for the variance of Newton-Cotes estimators based on the sampling nodes in a bounded interval. Finally, the findings are illustrated in an application of volume estimation for three-dimensional convex bodies with sufficiently smooth boundaries.

Keywords: Newton-Cotes quadrature; numerical integration with random nodes; stationary point process; variance estimation,; dropouts; Cavalieri estimator; weakly (m,p)*-piecewise smooth function*

B.1 Introduction

In the present paper we consider variance behavior and variance approximations of the so-called Newton-Cotes estimators defined in [6] and investigated further in Paper A. The estimators can be seen as generalizations of the well-known classical Cavalieri volume estimator, and as an improvement of the generalized Cavalieri estimator defined initially in [1] and described further in [12, 13]. We will give a short introduction to all estimators but refer to (Sections A.1 and A.2, Paper A) for a more thorough description and comparison. As all the above-mentioned estimators can be seen as instances of Monte Carlo integration schemes for a function on the real axis, we adopt this more general setting in the following.

Throughout this paper we assume that a compactly supported and integrable function f is known at all points of a stationary point process $X \subseteq \mathbb{R}$ (random locally finite collection of points in $\mathbb{R}$ with a translation-invariant distribution). The aim is to estimate the integral $\int_{\mathbb{R}} f(x)\mathrm{d}x$ from this data. From now on we refer to f as the *measurement function*.

In the classical setting the sampling points $X = \{x_k\}_{k\in\mathbb{Z}}$ are equidistant and satisfy $x_k = t(U + k)$, $k \in \mathbb{Z}$, where $t > 0$ and U is a uniform random variable on the interval $(0, 1)$, turning X into a stationary point process. Its intensity (expected number of points per unit interval) is $1/t$. The classical Cavalieri estimator is then defined by

$$\hat{V}(f) = t \sum_{x\in X} f(x), \tag{B.1.1}$$

as described in [1]. The estimator (B.1.1) can also be used if the points of the stationary point process X with intensity $1/t$ are not equidistant, in which case it is referred to as the *generalized Cavalieri estimator*. Both estimators are unbiased for the integral $\int_{\mathbb{R}} f(x)\mathrm{d}x$ [1, Theorem 1], however, as indicated in [1] and quantified in [12], the variance of the generalized Cavalieri estimator may be substantially higher than in the equidistant case. [12, 13] also consider the practically relevant situation where the f-values at some of the points of X are unavailable (*dropouts*). This situation is modeled as independent p-thinning, where independently for each point its corresponding function value is discarded with probability $p \in [0, 1)$ before (B.1.1) is applied to the remaining observations. [13] suggests a better alternative to (B.1.1) under independent p-thinning where discarded function values are approximated by the average of the closest two neighboring known measurements before (B.1.1) is applied. In either case, the variance is higher than in the equidistant case.

Given that the increments of X are available, [6] suggest using Newton-Cotes quadrature rules to avoid the above-described variance inflation. For fixed $n \in \mathbb{N}$, on the interval from $x_0 \in X$ to its nth neighbor $x_n \in X$, the function f is approximated by a piecewise polynomial of degree at most n passing through the points $\{x_j, f(x_j)\}_{j=0}^{n}$, where $x_1 < \cdots < x_{n-1}$ are the ordered points in $X \cap (x_0, x_n)$. The *nth (order) Newton-Cotes estimator* $\hat{V}_n(f)$ is then obtained as the sum of integrals of such function approximations averaged with respect to the starting point. By this construction, $\hat{V}_n(f)$ is given as the weighted sum

$$\hat{V}_n(f) = \sum_{x\in X} \alpha_n(x) f(x), \tag{B.1.2}$$

where $\alpha_n(x) = \alpha_n(x;X)$ is a rational function of n point increments to the left and right of $x \in X$; see (Eq.'s (A.3.3) and (A.3.4), A). In particular, $\alpha_1(x) = (h_1(x)+h_0(x))/2$, where $h_0(x)$ and $h_1(x)$ denotes the increment to the left and right of x, respectively. This gives the *trapezoidal estimator*

$$\hat{V}_1(f) = \sum_{x \in X} \frac{h_1(x)+h_0(x)}{2} f(x),$$

which will be of particular interest throughout this paper. As for the classical and generalized Cavalieri estimators, the nth Newton-Cotes estimator (B.1.2) is unbiased for the integral $\int_{\mathbb{R}} f(x)\mathrm{d}x$ as long as the typical increments satisfy certain integrability conditions; see (B.2.1) below and (Theorem A.2.2, Paper A). Interestingly, Newton-Cotes estimators of any order coincide with the classical Cavalieri estimator when the points in X are equidistant, and moreover, the trapezoidal estimator applied to the equidistant process combined with independent p-thinning coincides with the correction method of [13]. The latter claim will be shown in the section on independent p-thinning. One great advantage of Newton-Cotes estimators is that they can be applied to any stationary point process satisfying (B.2.1) with a variance-order independent of the underlying point process. This will be clarified in the next paragraph. In particular, the variances of estimators applied to a stationary point process or a thinning hereof are of the same order; see the section on independent p-thinning.

The order of the variance depends not only on the order n of $\hat{V}_n$ but also on the smoothness of the measurement function f. The smoothness concepts described below are given in terms of jumps of a function, which we define as follows: The function $h : \mathbb{R} \to \mathbb{R}$ jumps at the point $a \in \mathbb{R}$ if the limits and the difference in

$$J_h(a) = h(a^+) - h(a^-) = \lim_{x \downarrow a} h(x) - \lim_{x \uparrow a} h(x)$$

are defined on $\mathbb{R} \cup \{-\infty, \infty\}$ and $J_h(a) \neq 0$. We refer to J_h as the jump-function of h and we let D_h denote the set of jump-points of h. The classical smoothness concept is that of $(m,1)$-piecewise smoothness, and this is used first and foremost in [7] and subsequently in Paper A. For a given $m \in \mathbb{N}_0$, a compactly supported function is said to be $(m,1)$-piecewise smooth if it is $(m-1)$-times continuously differentiable and if the mth and $(m+1)$st derivatives exist and are continuous except in at most finitely many points, where they may have finite jumps. However, the condition on the $(m+1)$st derivative turns out to be rather restrictive from a practical point of view: In the section on stereological volume estimation below, we give an example of a practically relevant measurement function whose second derivative has infinite jumps at the boundary of its support. It turns out that this less restrictive smoothness property actually suffices for the variance results presented in this paper to hold; in particular, the entire Paper A could have been formulated under this milder smoothness property. We refer to the appendix for a justification. For this reason, from now on we consider what we call *weakly* $(m,1)$-piecewise smooth functions, which differ from $(m,1)$-piecewise smooth functions by the less restrictive property that the $(m+1)$st derivative is allowed to have finitely many (possibly infinite) jumps. As the variance of $\hat{V}_n$ is essentially given by the order of the first non-continuous derivative of the measurement function, we assume throughout the following that

m is chosen largest possible for a given f, that is, we assume that the mth derivative has at least one jump. It is not difficult to see that the $(m+1)$st derivative of a weakly $(m,1)$-piecewise smooth function is integrable.

It should be mentioned that a third smoothness concept for measurement functions has been introduced in the literature by [3]. In that paper $(q,1)$-piecewise smoothness has been extended to allow for arbitrary real $q \geq 0$, and functions with this property are called q-smooth (for $p = 1$). Their theory is based on fractional derivatives. For $m \in \mathbb{N}_0$ an $(m,1)$-piecewise smooth function is also m-smooth, but the converse is false. The function $f(x) = 1_{[0,1]}x\log(x)$ is weakly $(0,1)$-piecewise smooth, but neither $(0,1)$-piecewise smooth nor 0-smooth, showing that our new notion covers previously intractable measurement functions. On the other hand, even when applying the theory to the problem of estimating the volume of a compact convex object, one cannot avoid fractional smoothness; see [3]. Although smoothness assumptions – like the one we will impose in the section on volume estimation – can avoid fractional smoothness of the measurement function, a unifying approach is still missing but beyond the scope of the present paper.

The variance of $\hat{V}_n(f)$ is given in terms of the jumps of the $(2m+1)$st derivative of the covariogram $g(z) = \int_{\mathbb{R}} f(x)f(x+z)\mathrm{d}x$, $z \in \mathbb{R}$, associated to f. Here we make use of the fact that g is weakly $(2m+1,1)$-piecewise smooth by Corollary B.6.3 in the appendix (an adaption of [7, Corollary 5.8] to weakly piecewise smooth functions) and, since the mth derivative has non-zero jumps, the fact that $(2m+1)$ is the order of the first discontinuous derivative of g. If the measurement function is weakly $(m,1)$-piecewise smooth, a decomposition of the variance of the classical Cavalieri estimator (B.1.1), known for $(m,1)$-piecewise smooth functions, holds:

$$\mathbb{V}\mathrm{ar}(\hat{V}(f)) = \mathbb{V}\mathrm{ar}_E(\hat{V}(f)) + Z(t) + r(t). \tag{B.1.3}$$

Here, the *extension term* $\mathbb{V}\mathrm{ar}_E(\hat{V}(f))$ explains the general behavior of the variance and is given by $\mathbb{V}\mathrm{ar}_E(\hat{V}(f)) = t^{2m+2}g^{(2m+1)}(0^+)c_m$ for a non-zero constant c_m independent of t and the measurement function f. For instance, $c_0 = -1/6$ and $c_1 = 1/360$. The *Zitterbewegung* $Z(t)$, which is also of order t^{2m+2}, is a finite sum of terms oscillating around 0 with the oscillation given in terms of the jumps $a \in D_{g^{(2m+1)}}$, $a \neq 0$, of $g^{(2m+1)}$ away from the origin, and $r(t)$ is a lower-order remainder. The main message here is that the variance decreases at a rate of t^{2m+2} when f is weakly $(m,1)$-piecewise smooth and this rate of decrease is not achievable with the generalized Cavalieri estimator in the non-equidistant case; see e.g. [12, Proposition 1]. However, as shown in (Theorem A.2.8, Paper A), if X is stationary with intensity $1/t$, applying a Newton-Cotes estimator of order m also yields a decrease rate of t^{2m+2} for the variance. Moreover, under certain assumptions on the covariance structure of X, the variances of Newton-Cotes estimators decompose similar to (B.1.3) but with the Zitterbewegung not necessarily being oscillating. This is the result presented in Lemma B.2.4 below.

Two point models will receive particular interest in this paper as they did in [12, 13] and Paper A, namely the perturbed and cumulative models. Note that the perturbed model is formulated in terms of perturbations from the equidistant model, and hence it includes the equidistant model if the perturbations have a degenerate distribution concentrated at 0.

Definition B.1.1 (Perturbed model). *A stationary point process $X = \{x_k\}_{k\in\mathbb{Z}}$ with intensity $1/t$ is from the perturbed model if $x_k = t(U + k + E_k)$, where U is a uniform random variable on $(0,1)$ and the perturbations $\{E_k\}_{k\in\mathbb{Z}}$ are independent and identically distributed with $\mathbb{E}E_k = 0$ and $|E_k| < 1/2$ almost surely. Moreover, U and $\{E_k\}$ are assumed to be stochastically independent.*

Definition B.1.2 (Cumulative model). *A stationary point process X with intensity $1/t$ is from the cumulative model if it has independent and identically distributed increments $\{\omega_k\}_{k\in\mathbb{Z}}$, where ω_k (necessarily) has expectation t. Furthermore, the increments are assumed to have a continuous distribution and a moment-generating function $\eta \mapsto \mathbb{E}e^{\eta\omega_k}$ which is finite in a neighborhood of 0.*

Note that the cumulative model is defined slightly differently than the *model with cumulative errors* in Paper A, hence the different naming. In contrast to that paper, we always require the condition on the moment-generation function. In fact, all variance results in Paper A for the model with cumulative errors also require the latter condition.

The purpose of this paper is to extend the results of Paper A on Newton-Cotes estimation. The first goal is to allow for independently p-thinned point processes; in particular, we will state explicit variance expressions in terms of the thinning probability p for the trapezoidal estimator when sampling from the perturbed or cumulative models. By letting $p = 0$ the exact expressions (Theorems A.2.9 & A.2.10, Paper A) for the original models are recovered. Secondly, we derive and present statistical estimates for the variance of Newton-Cotes estimators applied to any point process. In the special cases of sampling from the perturbed or cumulative model with potential thinning alternative estimates are suggested which appear to be particularly robust. Finally, we substantiate our findings by application to the stereological problem of estimating the volume of a compact convex set $Y \subset \mathbb{R}^3$ from parallel section profiles. We will state variance relations for $\hat{V}_n$ in terms of principal curvatures of the boundary of Y. We will also discuss the case where the joint direction of the section planes is isotropically randomized and show among other things that the variance of the Newton-Cotes estimator then is essentially proportional to the surface area of the object. This is well-known for the classical Cavalieri estimator based on equidistant points (see [2, 9]) but exact conditions on Y for this statement to hold have not been specified before, not even in the classical case.

The paper is organized as follows. First we introduce relevant notation and rephrase two important results on the variance of Newton-Cotes estimators in general and in particular of the trapezoidal estimator. We introduce independent p-thinning to model dropouts, show that the results for Newton-Cotes estimators also apply to a thinned process and give explicit variance expressions for $\hat{V}_n$ when sampling from the perturbed or cumulative model with potential thinning. Next, we give an overview of techniques to estimate the variance of Newton-Cotes estimators, concluding with a section on stereological volume estimation.

B.2 Variance of Newton-Cotes estimators

In this section we introduce the notation used throughout the paper, and we give a brief overview of the results on Newton-Cotes estimators presented in Paper A.

For a function $f : \mathbb{R} \to \mathbb{R}$, we let $f^{(k)}$ denote its kth derivative whenever it is defined. At times, we will also use the notation f' and f'' for the first and second derivative, respectively.

For any finite-intensity stationary point process X on $\mathbb{R}$ we define points and increments in X relative to $x \in X$ as follows: For any $j \in \mathbb{Z}$, $s_j(x) = s_j(x;X)$ denotes the jth successor (predecessor for $j < 0$) of x in X with $s_0(x) = x$ by definition, and $h_j(x) = h_j(x;X) = s_j(x) - s_{j-1}(x)$ denotes the distance from the jth successor (predecessor) of x to its left neighbor in X.

Throughout the paper, a subscript u on a point process indicates a scaling to unit intensity, i.e. $X_u = X/t$ has intensity 1 when X has intensity $1/t$, $t > 0$.

We let $\mathbb{P}_X^0$ denote the Palm distribution of X, that is, the conditional distribution of X given that it contains the origin (see e.g. [11, Section 3.3]), and we let $\mathbb{E}_X^0$ denote the expectation with respect to the Palm measure. When considering a point process under its Palm distribution, we suppress the dependence on the origin 0, writing for instance $s_j(0) = s_j$ and $h_j(0) = h_j$, $j \in \mathbb{Z}$.

The results in Paper A are formulated under the overall assumption of finite positive and negative moments of the typical point increment, and for completeness we include it here. Note that (Assumption A.2.1, Paper A) is slightly stronger than Assumption B.2.1 below, but the arguments and comments in that paper show that (B.2.1) is actually sufficient for all results to hold.

Assumption B.2.1. *For fixed Newton-Cotes order $n \in \mathbb{N}$, we have*

$$\mathbb{E}_X^0 h_1^j < \infty \qquad \textit{for all} \begin{cases} j \in \mathbb{N} & \textit{if} \quad n = 1, \\ j \in \mathbb{Z} & \textit{if} \quad n \geq 2. \end{cases} \tag{B.2.1}$$

Before proceeding to the results on Newton-Cotes estimators, we introduce a stochastic process $K_m(\cdot) = K_m(\cdot;X)$ on $\mathbb{R}$, $m \in \mathbb{N}_0$. It is commonly referred to as the *(mth) Peano kernel*, as it appears as an integration kernel in a Peano-type error representation of the Newton-Cotes estimator for a weakly $(m,1)$-piecewise smooth function f (adaption of (Theorem A.2.3, Paper A)):

$$\hat{V}_n(f) - \int_{\mathbb{R}} f(x)\mathrm{d}x = \int_{\mathbb{R}} f^{(m+1)}(r)K_m(r)\mathrm{d}r + \sum_{a \in D_{f^{(m)}}} J_{f^{(m)}}(a)K_m(a) \tag{B.2.2}$$

for $m \leq n$. For details on K_m and the summarized properties below, see (Sections A.3 & A.4, Paper A). The Peano kernel is a piecewise polynomial of order at most $(m+1)$ with coefficient given in terms of the underlying point process X and the order n of the Newton-Cotes estimator. More specifically, $K_m(r)$ depends on the n points in X to the left and right of $r \in \mathbb{R}$. It is translation covariant, i.e. $K_m(r+s;X+s) = K_m(r;X)$ for all $s, r \in \mathbb{R}$. When X is a stationary point process satisfying (B.2.1), K_m is a stationary stochastic process with finite absolute moments of any positive order, i.e. $\mathbb{E}|K_m(r)|^j = \mathbb{E}|K_m(0)|^j < \infty$ for all $j \in \mathbb{N}$ and $r \in \mathbb{R}$. We let $H_m(\cdot) = H_m(\cdot;X)$ denote the covariance function of K_m associated to X, that is, $H_m(s) = \mathrm{Cov}(K_m(r), K_m(r+s))$. It is independent of $r \in \mathbb{R}$ and, by Cauchy-Schwarz inequality, $H_m(s) \le H_m(0)$ for all $s \in \mathbb{R}$ where equality cannot hold on non-degenerate intervals; see Lemma B.2.2 below. Moreover, K_m is co-homogeneous of degree $m+1$, and, thus

$$H_m(tr;tX) = t^{2m+2}H_m(r;X) = t^{2m+2}H_m(r) \tag{B.2.3}$$

for all $r \in \mathbb{R}$ and $0 < t < \infty$.

Lemma B.2.2. *Let $n \in \mathbb{N}$, $m \le n$ and a compact interval $I \subset \mathbb{R}$ be given. If $H_m(s) = H_m(0)$ for all $s \in I$, then I is a singleton.*

Proof. For contradiction, assume that $H_m(s) = H_m(0)$ for all $s \in I$, where I is a compact interval of positive length $r > 0$. Without loss of generality we may assume that I has rational endpoints. For $j \in \mathbb{Z}$ and $s \in \mathbb{Q} \cap I$, we have $K_m(jr+s) = K_m(jr)$ for all $s \in \mathbb{Q} \cap I$, almost surely. Here we used the assumed equality, stationarity and the fact that $\mathbb{Q}$ is countable. As the intervals in $\{jr + I \,:\, j \in \mathbb{Z}\}$ cover $\mathbb{R}$, and any two neighboring intervals have rational endpoints in common, we conclude that

$$K_m(s) = K_m(0) \text{ for all } s \in \mathbb{Q}, \tag{B.2.4}$$

almost surely. As K_0 is linear on each connected component of $\mathbb{R} \setminus X$ by (Eq. (A.3.7), Paper A), Equation (B.2.4) implies $K_m \equiv K_m(0)$ on $\mathbb{R} \setminus X$ almost surely when $m = 0$. For $m \ge 1$, the process $K_m(\cdot)$ has continuous paths by (Lemma A.3.2, Paper A), so (B.2.4) implies $K_m \equiv K_m(0)$ on $\mathbb{R}$ almost surely. In either case (B.2.2) yields $\hat{V}_n(f) = \int_{\mathbb{R}} f(x)\mathrm{d}x$ for almost all realizations of X and all weakly $(m,1)$-piecewise smooth functions f. Taking two such functions f_1, f_2 coinciding on all points of X but with $\int_{\mathbb{R}} f_1(x)\mathrm{d}x \neq \int_{\mathbb{R}} f_2(x)\mathrm{d}x$ we obtain a contradiction. □

An important property introduced in Paper A is that of admissibility of a point process. Strong admissibility improves the order of the variance when the degree of smoothness of the measurement function exceeds the order of the estimator (Theorem A.2.8, Paper A), and weak admissibility ensures a variance decomposition similar to (B.1.3) as seen in Lemma B.2.4 below. Note that the class of admissible point processes is closed under scaling.

Definition B.2.3. *Let $n \in \mathbb{N}$ be given and let X be a stationary point process satisfying* (B.2.1). *X is called* strongly n-admissible *if $\int_0^z H_n(s)\mathrm{d}s$ is uniformly bounded in $z \ge 0$. X is called* weakly n-admissible *if $\lim_{z\to\infty} \frac{1}{z}\int_0^z H_n(s)\mathrm{d}s = 0$.*

Lemma B.2.4. *Let $n \in \mathbb{N}$ be given and assume that X is a stationary point process satisfying* (B.2.1) *and with intensity $1/t > 0$. If f is weakly $(m,1)$-piecewise smooth with $m \le n$ and covariogram g, then*

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = \mathbb{V}\mathrm{ar}_E(\hat{V}_n(f)) + Z_m(t) + r(t), \tag{B.2.5}$$

where the extension term $\mathbb{V}\mathrm{ar}_E(\hat{V}_n(f))$ is of order t^{2m+2} and is given by

$$\mathbb{V}\mathrm{ar}_E(\hat{V}_n(f)) = (-1)^{m+1}2g^{(2m+1)}(0^+)H_m(0), \tag{B.2.6}$$

the Zitterbewegung $Z_m(t)$ is of order $O(t^{2m+2})$ and satisfies

$$Z_m(t) = (-1)^{m+1}\sum_{a\in D_{g^{(2m+1)}}\setminus\{0\}} J_{g^{(2m+1)}}(a)H_m(a), \tag{B.2.7}$$

and the remainder $r(t)$ is of order $O(t^{2m+2})$. If $m < n$ or X is weakly n-admissible, $r(t)$ is of order $o(t^{2m+2})$.

Proof. Applying (B.2.3), we may write

$$H_m(r) = H_m(r;X) = t^{2m+2} H_m(r/t;X_u) \tag{B.2.8}$$

for all $r \in \mathbb{R}$, where $X_u = X/t$ is a unit-intensity point process. By properties of the Peano kernel, $H_m(0;X_u)$ is independent of t and $H_m(r/t;X_u)$ is uniformly bounded in t. The result now follows from an adaption of (Proposition A.6.1, Paper A) to weakly piecewise smooth functions. □

Both the perturbed and cumulative models are strongly n-admissible for all $n \in \mathbb{N}$ and, in particular, the remainder $r(t)$ in Lemma B.2.4 is of order $o(t^{2m+2})$. Even stronger, if X is from the perturbed model then $r \mapsto H_m(r;X_u)$ is periodic with period 1 for all $m \in \mathbb{N}_0$ and sufficiently large $|r|$, and $\int_r^{r+1} H_m(s;X_u)\mathrm{d}s = 0$ (Lemma A.7.1, Paper A). Hence, if X has intensity $1/t$, the Zitterbewegung (B.2.7) is of order t^{2m+2} and it is a finite sum of terms each oscillating around 0. If X is from the cumulative model there is $\epsilon > 0$ such that $H_m(r/t;X_u) = O(\mathrm{e}^{-\epsilon r/t})$ for all $r \in \mathbb{R}$ as $t \downarrow 0$, due to (Lemma A.7.3, Paper A). Hence, by (B.2.8), it follows that the Zitterbewegung $Z(t)$ is $o(t^{2m+2})$ and thus (B.2.5) reads

$$\mathbb{V}\mathrm{ar}(\hat{V}_n(f)) = \mathbb{V}\mathrm{ar}_E(\hat{V}_n(f)) + o(t^{2m+2}) \tag{B.2.9}$$

as $t \downarrow 0$ in this case.

In (Eq.'s (A.7.8) and (A.7.9), Paper A) $H_m(0;X_u)$ is simplified for $n = 1$ and $m \in \{0,1\}$. Using these representations, the following corollary to Lemma B.2.4 is a consequence of the fact that $\mathbb{E}_X^0 h_1^j = t^j \mathbb{E}_{X_u}^0 h_1^j$ for all $j \in \mathbb{N}$.

Corollary B.2.5. *Let X be a stationary point process satisfying* (B.2.1) *and with intensity* $1/t > 0$, *and let f be weakly* $(m,1)$*-piecewise smooth with covariogram g. For $m = 0$ we have*

$$\begin{aligned}\mathbb{V}\mathrm{ar}_E(\hat{V}_1(f)) &= -g^{(1)}(0^+)\tfrac{1}{6}\tfrac{1}{t}\mathbb{E}_X^0 h_1^3 \\ &= -t^2 g^{(1)}(0^+)\tfrac{1}{6}\mathbb{E}_{X_u}^0 h_1^3,\end{aligned} \tag{B.2.10}$$

and for $m = 1$ we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}_E(\hat{V}_1(f)) &= g^{(3)}(0^+)\tfrac{1}{360}\Big(6\tfrac{1}{t}\mathbb{E}_X^0 h_1^5 - 5(\tfrac{1}{t}\mathbb{E}_X^0 h_1^3)^2\Big) \\ &= t^4 g^{(3)}(0^+)\tfrac{1}{360}\Big(6\mathbb{E}_{X_u}^0 h_1^5 - 5(\mathbb{E}_{X_u}^0 h_1^3)^2\Big).\end{aligned}$$

B.3 Independent p-thinning

In this section we consider the point process X to be obtained by so-called independent p-thinning, where $0 \le p < 1$ denotes the probability of an initial point being removed. We let $\tilde{X} = \{\tilde{x}_k\}_{k\in\mathbb{Z}}$ denote the underlying stationary point process on $\mathbb{R}$ with finite intensity $1/\tilde{t} > 0$, and we define X by $X = \{\tilde{x}_k \in \tilde{X} : U_k > p\}$, where $\{U_k\}_{k\in\mathbb{Z}}$ denotes an i.i.d. sequence of uniform $(0,1)$ variables independent of $\tilde{X}$. Hence, U_k is the thinning variable associated to the point $\tilde{x}_k \in \tilde{X}$, and therefore $\tilde{x}_k \in X$ with probability $\mathbb{P}(U_k > p) = 1-p$. We refer to $\tilde{X}$ as the *initial* process and X as the *observed* or *thinned* process. If $p = 0$ no dropouts occur and naturally $\tilde{X} = X$. If X has intensity $1/t$, it is not difficult to see that $t = \tilde{t}/(1-p)$. As in the previous section we let $X_u = X/t$ denote the scaled unit-intensity process of X.

As mentioned previously, the correction method introduced in [13, Section 4] coincides with the trapezoidal estimator applied to the thinned process X if the initial process $\tilde{X}$ is equidistant. To see this, recall that $\alpha_1(x) = (h_1(x) + h_0(x))/2$ for all $x \in X$, where, for some $k \in \mathbb{Z}$, $x = \tilde{x}_k$. If $\tilde{X}$ has intensity $1/\tilde{t}$ and $x \in X$ we have $h_1(x) = \tilde{t}\min\{j \geq 1 \,:\, U_{k+j} > p\}$ and $h_0(x) = \tilde{t}\min\{j \geq 1 \,:\, U_{k-j} > p\}$. Consequently, the trapezoidal estimator reads

$$\hat{V}_1(f) = \tilde{t}\sum_{k\in\mathbb{Z}} \psi_k f(\tilde{x}_k),$$

with the weight

$$\psi_k = \tfrac{1}{2}\mathbf{1}_{(U_k>p)}\Big(\min\{j \geq 1 \,:\, U_{k+j} > p\} + \min\{j \geq 1 \,:\, U_{k-j} > p\}\Big).$$

This is exactly the estimator introduced in [13, Section 4]. Note that the estimators only coincide when the underlying point process is equidistant, and if this is not the case, the trapezoidal estimator is superior. For instance, the correction method of [13] is of order $\tilde{t}^3$ when the underlying process is from the perturbed model and the measurement function is weakly $(1,1)$-piecewise smooth, whereas the trapezoidal estimator is always of order t^4 (and hence also of order $\tilde{t}^4$) for weakly $(1,1)$-piecewise smooth functions; see [13, Proposition 4] and Lemma B.2.4 above.

We now aim for variance expressions of Newton-Cotes estimators when applied to the observed process X. As the results in the previous section rest on Assumption B.2.1, we have to assure that this condition is satisfied by the observed process X. Lemma B.3.1 below shows that this is in fact the case if the initial process satisfies Assumption B.2.1. Subsequently, we find explicit variance expressions for the trapezoidal estimator applied to the observed process X, when the initial process $\tilde{X}$ is from the perturbed or cumulative model. We show that X is strongly n-admissible in either case, and hence it satisfies (B.2.5) with the remainder $r(t)$ of order $o(t^{2m+2})$. $\mathbb{V}\mathrm{ar}_E(\hat{V}_n(f))$ is given by (B.2.6) for arbitrary n and explicitly in Corollary B.2.5 for the trapezoidal estimator.

Lemma B.3.1. *If the initial process $\tilde{X}$ satisfies* (B.2.1), *then so does the thinned process X.*

Proof. Between two consecutive points in the thinned process X, the number of removed points from the initial process $\tilde{X}$ follows a geometric distribution on $\mathbb{N}_0$ with success probability $(1-p)$. By independence, this implies

$$\mathbb{E}^0_X h_1^j = \sum_{k=1}^{\infty}(1-p)p^{k-1}\,\mathbb{E}^0_{\tilde{X}} s_k^j. \tag{B.3.1}$$

As $\mathbb{E}^0_X h_1^j \leq \mathbb{E}^0_{\tilde{X}} h_1^j < \infty$ for all $j < 0$ we may assume $j \geq 0$. By an application of the multinomial theorem and Hölder's inequality, $\mathbb{E}^0_{\tilde{X}} s_k^j$ has the upper bound

$$\mathbb{E}^0_{\tilde{X}}\big(h_1 + \cdots + h_k\big)^j \leq \sum_{j_1+\cdots+j_k=j}\binom{j}{j_1,\ldots,j_k}\mathbb{E}^0_{\tilde{X}}\big[h_1^j\big] = k^j\,\mathbb{E}^0_{\tilde{X}}\big[h_1^j\big],$$

and we conclude the proof as $\mathbb{E}^0_X h_1^j \leq \mathbb{E}^0_{\tilde{X}}\big[h_1^j\big](1-p)\sum_{k=1}^{\infty} k^j p^{k-1} < \infty$, due to the assumption on $\tilde{X}$. □

As mentioned, a point process from the perturbed model is strongly n-admissible for any $n \in \mathbb{N}$, and this is also the case under independent thinning.

Lemma B.3.2. *Let the initial process $\tilde{X}$ be from the perturbed model and let X be the observed process. For all $n, j \in \mathbb{N}$ and $m \in \mathbb{N}_0$, X satisfies*

(i) $H_m(r; X_u) = H_m(r+1; X_u) + o(r^{-j})$,

(ii) $\int_r^{r+1} H_m(s; X_u)\mathrm{d}s = o(r^{-j})$,

as $r \to \pm\infty$. Furthermore,

(iii) *X is strongly n-admissible .*

Proof. As $r \mapsto H_m(r)$ is an even function due to stationarity, it is enough to show the claims (i) and (ii) for positive $r \to \infty$.

Before proving (i)–(iii) we introduce some notation which will be used throughout the proof. First, to simplify notation, we will write $K_m(r) = K_m(r; X_u)$ and $H_m(r) = H_m(r; X_u)$, as we do not consider the observed process but only the associated unit-scaled process. For $s \in \mathbb{R}$ and $r \geq 0$ let

$$A_-(s) = A_-(s; X_u) = \left\{\#\left(X_u \cap [s - \tfrac{r}{3}, s)\right) \geq n\right\}, \quad \text{and}$$
$$A_+(s) = A_+(s; X_u) = \left\{\#\left(X_u \cap [s, s + \tfrac{r}{3})\right) \geq n\right\}$$

be the events that X_u has at least n points to the left and right of s, respectively, within a distance of at most $r/3$. Furthermore define $I_\pm(s) = I_\pm(s; X_u) = \mathbf{1}_{A_\pm(s; X_u)}$ as the associated indicator functions, and note that they are vanishing for small r. We also define the shifted process $X^* = X_u - U = \{k + E_k \,:\, U_k > p,\ k \in \mathbb{Z}\}$ and correspondingly $K_m^*(s) = K_m(s; X^*)$ and $I_\pm^*(s) = I_\pm(s; X^*)$. Due to the i.i.d. property of perturbations and thinning variables, the distributional equivalence

$$K_m^*(s) I_\pm^*(s) \stackrel{\mathcal{D}}{=} K_m^*(s+1) I_\pm^*(s+1) \tag{B.3.2}$$

holds for all $s \in \mathbb{R}$, where the translation covariance of $K_m(s; X)$ and $I_\pm(s; X)$ was used. Furthermore, by construction, $K_m^*(s) I_-^*(s)$ and $K_m^*(s) I_+^*(s)$ only depend on the variables

$$\{(E_i, U_i) \,:\, i + E_i \geq s - r/3\} \subseteq \{(E_i, U_i) \,:\, i \geq s - r/3 - 1/2\},$$

and

$$\{(E_i, U_i) \,:\, i + E_i < s + r/3\} \subseteq \{(E_i, U_i) \,:\, i < s + r/3 + 1/2\},$$

respectively. In particular, $K_m^*(s) I_-^*(s)$ and $K_m^*(s') I_+^*(s')$ are stochastically independent for $s - s' \geq 2r/3 + 1$.

Let $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$ be fixed and consider $r \geq 0$. To prove (i) we observe by stationarity that $\mathbb{P}((A_-(s) \cap A_+(0))^c) \leq 2\mathbb{P}(A_+(0)^c) = 2\mathbb{P}(\#(X_u \cap [0, r/3)) < n)$, and thus

$$\mathbb{P}\left((A_-(s) \cap A_+(0))^c\right) = o(r^{-j}) \tag{B.3.3}$$

for all $j \in \mathbb{N}$ and $s \in \mathbb{R}$, as $r \to \infty$. This is seen as follows: With $\lfloor\cdot\rfloor$ denoting integer part, there are at least $N_r = \lfloor r/3 \rfloor - 2$ points in $\tilde{X}_u \cap [0, r/3)$, and the number B of these points also in X_u is a binomial random variable with N_r trials and success probability $1-p$. As $p<1$ it is not difficult to see that $\mathbb{P}(B<n)$ is of order $o(N_r^{-j})$ for any $j \in \mathbb{N}$, and hence also of order $o(r^{-j})$ as $r \to \infty$. Equation (B.3.3) now follows as $\mathbb{P}(\#(X_u \cap [0, \frac{r}{3})) < n) \le \mathbb{P}(B<n)$.

By the Cauchy-Schwarz inequality and the fact that $\mathbb{E}K_m^4(r) = \mathbb{E}K_m^4(0) < \infty$, (B.3.3) implies that

$$\mathbb{E}[K_m(r)K_m(0)] = S(r) + o(r^{-j}) \tag{B.3.4}$$

for all $j \in \mathbb{N}$, where

$$S(r) = \mathbb{E}[K_m(r)I_-(r)K_m(0)I_+(0)]$$

is defined for notational convenience. Conditioning on U and using the translation covariance of K_m and $I_\pm$, $S(r)$ reads

$$\begin{aligned} S(r) &= \int_0^1 \mathbb{E}[K_m^*(r-u)I_-^*(r-u)K_m^*(-u)I_+^*(-u)]\mathrm{d}u \\ &= \int_0^1 \mathbb{E}[K_m^*(r-u)I_-^*(r-u)]\mathbb{E}[K_m^*(-u)I_+^*(-u)]\mathrm{d}u \end{aligned}$$

for $r \ge 3$, where it has been used that $K_m^*(r-u)I_-^*(r-u)$ and $K_m^*(-u)I_+^*(-u)$ are independent for such r. Equation (B.3.2) now implies that

$$S(r) = S(r+1)$$

for all $j \in \mathbb{N}$ and $r \ge 3$. Using this and (B.3.4) with r and $r+1$ yields (i).

Defining $R_\pm^*(s) = \mathbb{E}[K_m^*(s)I_\pm^*(s)]$, we note that

$$\int_0^1 R_\pm^*(s-u)\mathrm{d}u = \mathbb{E}[K_m(s)I_\pm(s)] = \mathbb{E}[K_m(0)I_\pm(0)] = \mathbb{E}[K_m(0)] + o(r^{-j}) \tag{B.3.5}$$

for all $s \in \mathbb{R}$, as $r \to \infty$. The first equality is seen by conditioning on U, the second equality is due to stationarity, and the third equality is obtained by Cauchy-Schwarz' inequality using the fact that $\mathbb{P}(A_\pm(0)^c) = o(r^{-j})$ and $\mathbb{E}K^2(0) < \infty$. Equations (B.3.4)–(B.3.5), Fubini's theorem and a substitution now yield

$$\begin{aligned} \int_r^{r+1} \mathbb{E}[K_m(0)K_m(s)]\mathrm{d}s &= \int_0^1 R_+^*(-u)\int_0^1 R_-^*(r+1-u-s)\mathrm{d}s\mathrm{d}u + o(r^{-j}) \\ &= (\mathbb{E}[K_m(0)])^2 + o(r^{-j}), \end{aligned}$$

which proves (ii).

That X_u (and consequently X by (B.2.3)) is strongly n-admissible is easily seen by (ii) as

$$\begin{aligned} \left|\int_0^z H_n(r)\mathrm{d}r\right| &= \left|\sum_{i=0}^{\lfloor z\rfloor -1}\int_i^{i+1} H_n(s)\mathrm{d}s + \int_{\lfloor z\rfloor}^z H_n(s)\mathrm{d}s\right| \\ &\le \sum_{i=0}^{\infty}\left|\int_i^{i+1} H_n(s)\mathrm{d}s\right| + \mathbb{V}\mathrm{ar}(K_n(0)) < \infty. \end{aligned}$$

This is exactly assertion (iii). □

Theorem B.3.3. *Let the initial process $\tilde{X}$ be from the perturbed model with intensity $1/\tilde{t}$ and let X be the observed process with intensity $1/t$. Let $\theta_k = \mathbb{E}(E_1 - E_0)^k$ be the kth moment of the difference in perturbations. If f is weakly $(m,1)$-piecewise smooth with covariogram g, then $\mathbb{V}\mathrm{ar}_E(\hat{V}_1(f))$ coincides with*

$$-\tilde{t}^2 g^{(1)}(0^+)\frac{1}{6}\Big(1 + 3\theta_2 + \frac{6p}{(1-p)^2}\Big), \tag{B.3.6}$$

for $m = 0$, and

$$\tilde{t}^4 g^{(3)}(0^+)\frac{1}{360}\Big(1 + \frac{120p^3 + 300p^2 + 120p}{(1-p)^4} + 30\theta_2\frac{p^2+4p+1}{(1-p)^2} - 45\theta_2^2 + 30\theta_4\Big), \tag{B.3.7}$$

for $m = 1$. Moreover, the Zitterbewegung $Z_m(t)$ is of order t^{2m+2} (hence $\tilde{t}^{2m+2}$) and it is a finite sum of terms each oscillating around 0.

Before proving this result, we make a few comments on the extension term. As mentioned, the trapezoidal estimator and the correction method of [13] coincide when the initial point process is equidistant. In particular, when $\theta_k = 0$ for all $k \in \mathbb{N}$, the result above should match [13, Proposition 3]. However, according to [13], the second summand in (B.3.7) is erroneously claimed to be $(120p^3 + 390p^2 + 120p)/(1-p)^4$.

Considering perturbed sampling without thinning ($p = 0$), how much do we underestimate the extension term if we use the classical variance, that is, if we disregard the perturbations by setting $\theta_k = 0$ above? In the rather extreme case where E_1 is uniform on $(-1/2, 1/2)$ the extension term equals 3/2 and 27/4 times the classical extension term for $m = 0, 1$, respectively. By comparison, the extension term of the generalized Cavalieri estimator equals 2 times the classical extension term when $m = 0$; see [13, p. 190]. In the worst case where the errors E_i are arbitrarily close to $\pm 1/2$, the extension term equals 5/2 and 79/4 times the classical extension term for $m = 0, 1$, respectively.

Proof. As processes from the perturbed model satisfy (B.2.1), also X satisfies (B.2.1) due to Lemma B.3.1. Recall that $\theta_k = \mathbb{E}(E_j - E_0)^k = \mathbb{E}(E_1 - E_0)^k$ for all $j \neq 0$, with $\theta_k = 0$ for all odd k. It is not difficult to see that $\mathbb{E}^0_{\tilde{X}} s_j^5 = \tilde{t}^5 \mathbb{E}(j + E_j - E_0)^5 = \tilde{t}^5(j^5 + 10\theta_2 j^3 + 5\theta_4 j)$ for all $j \in \mathbb{N}$, and consequently, using (B.3.1),

$$\begin{aligned}
\mathbb{E}^0_X h_1^5 &= \sum_{j=1}^{\infty} p^{j-1}(1-p)\mathbb{E}^0_{\tilde{X}} s_j^5 \\
&= \tilde{t}^5 \sum_{j=1}^{\infty} p^{j-1}(1-p)(j^5 + 10\theta_2 j^3 + 5\theta_4 j) \\
&= \tilde{t}^5\Big(\frac{p^4 + 26p^3 + 66p^2 + 26p + 1}{(1-p)^5} + 10\theta_2\frac{p^2+4p+1}{(1-p)^3} + 5\theta_4\frac{1}{1-p}\Big).
\end{aligned}$$

By similar arguments it is shown that $\mathbb{E}^0_X h_1^3 = \tilde{t}^3((1+3\theta_2)/(1-p) + 6p/(1-p)^3)$. Since $t = \tilde{t}/(1-p)$, (B.3.6) and (B.3.7) follow by applying Corollary B.2.5 to X. The claim on the Zitterbewegung is a consequence of Lemma B.3.2 and (B.2.7). □

Next we turn to the cumulative model from Definition B.1.2. Interestingly, this model class is closed under independent thinning. In the following, $\{\tilde{\omega}_k\}$ and $\{\omega_k\}$ denotes the increments of the initial process $\tilde{X}$ and the thinned process X, respectively.

Lemma B.3.4. *If the initial process $\tilde{X}$ is from the cumulative model, then so is the thinned process X. In this case, for all $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$, there is some $\epsilon > 0$ such that $H_m(s) = O(\mathrm{e}^{-\epsilon s})$ as $s \to \infty$. In particular, X is strongly n-admissible.*

Proof. Let $\tilde{F}$ be the continuous distribution function of the initial increments. The independence of these increments and the thinning variables show that X has independent and identically distributed increments with continuous distribution function F given by $F(x) = \sum_{k=1}^{\infty}(1-p)p^{k-1}\tilde{F}^{*k}(x)$. Here $\tilde{F}^{*k}$ denotes the k-fold convolution of $\tilde{F}$, that is, the distribution of $\tilde{\omega}_1+\cdots+\tilde{\omega}_k$. Furthermore, $\mathbb{E}\mathrm{e}^{\eta\omega_1} = \sum_{k=1}^{\infty}(1-p)p^{k-1}(\mathbb{E}\mathrm{e}^{\eta\tilde{\omega}_1})^k$, which is finite if $\mathbb{E}\mathrm{e}^{\eta\tilde{\omega}_1} < 1/p$. Since there is $\tilde{\eta} > 0$ such that $\mathbb{E}\mathrm{e}^{\tilde{\eta}\tilde{\omega}_k} < \infty$, a dominated convergence argument shows that $\mathbb{E}\mathrm{e}^{\eta\tilde{\omega}_1} < 1/p$ if we choose $\eta > 0$ small enough. The claim on the convergence rate of H_m is now a consequence of (Lemma A.7.3, Paper A). □

Theorem B.3.5. *Let the initial process $\tilde{X}$ be from the cumulative model with intensity $1/\tilde{t}$ and let X be the observed process with intensity $1/t$. Define $\tilde{\nu}_k = \mathbb{E}\tilde{\omega}_1^k/\tilde{t}^k$ as the kth moment of the increments in the initial unit-intensity scaled process $\tilde{X}_u$. If f is weakly $(m,1)$-piecewise smooth with covariogram g, then $\mathbb{V}\mathrm{ar}_E(\hat{V}_1(f))$ coincides with*

$$-\tilde{t}^2 g^{(1)}(0^+)\frac{1}{6}\Big(\tilde{\nu}_3 + \tilde{\nu}_2\frac{6p}{1-p} + \frac{6p^2}{(1-p)^2}\Big),$$

for $m = 0$, and

$$\tilde{t}^4 g^{(3)}(0^+)\frac{1}{360}\Big(6\tilde{\nu}_5 - 5\tilde{\nu}_3^2 + \frac{60p(\tilde{\nu}_4+\tilde{\nu}_3\tilde{\nu}_2)}{1-p} + \frac{60p^2(5\tilde{\nu}_3+6\tilde{\nu}_2^2)}{(1-p)^2} + \frac{\tilde{\nu}_2 1080p^3}{(1-p)^3} + \frac{540p^4}{(1-p)^4}\Big),$$

for $m = 1$. Moreover, $Z_m(t)$ is of order $o(t^{2m+2})$, that is $o(\tilde{t}^{2m+2})$, for $m = 0,1$.

Proof. By definition, the cumulative model has finite exponential expectation and in particular (B.2.1) is satisfied for both the initial and the observed model. By the multinomial theorem, the independence and identical law of the increments, and by the fact that $\tilde{\nu}_1 = 1$,

$$\begin{aligned}
\mathbb{E}(\tilde{\omega}_1+\cdots+\tilde{\omega}_j)^5/\tilde{t}^5 &= \sum_{k_1+\cdots+k_j=5}\binom{5}{k_1,\ldots,k_j}\prod_{i=1}^{j}\tilde{\nu}_{k_j}\\
&= \binom{j}{1}\tilde{\nu}_5 + \binom{j}{2}\binom{2}{1}(5\tilde{\nu}_4+10\tilde{\nu}_3\tilde{\nu}_2) + \binom{j}{3}\binom{3}{2}(20\tilde{\nu}_3+30\tilde{\nu}_2^2) + \binom{j}{4}\binom{4}{3}60\tilde{\nu}_2 + \binom{j}{5}120\\
&= j\tilde{\nu}_5 + j(j-1)(5\tilde{\nu}_4+10\tilde{\nu}_3\tilde{\nu}_2) + j(j-1)(j-2)(10\tilde{\nu}_3+15\tilde{\nu}_2^2)\\
&\quad + j(j-1)(j-2)(j-3)10\tilde{\nu}_2 + j(j-1)(j-2)(j-3)(j-4)
\end{aligned}$$

for all $j \in \mathbb{N}$. Consequently,

$$\begin{aligned}
\mathbb{E}_X^0 h_1^5 &= \sum_{j=1}^{\infty}p^{j-1}(1-p)\mathbb{E}_{\tilde{X}}^0 s_j^5 = \sum_{j=1}^{\infty}p^{j-1}(1-p)\mathbb{E}(\tilde{\omega}_1+\cdots+\tilde{\omega}_j)^5\\
&= \tilde{t}^5\Big(\frac{\tilde{\nu}_5}{1-p} + \frac{(10\tilde{\nu}_4+20\tilde{\nu}_3\tilde{\nu}_2)p}{(1-p)^2} + \frac{(60\tilde{\nu}_3+90\tilde{\nu}_2^2)p^2}{(1-p)^3} + \frac{240\tilde{\nu}_2 p^3}{(1-p)^4} + \frac{120p^4}{(1-p)^5}\Big).
\end{aligned}$$

Similarly, it is seen that $\mathbb{E}(\tilde{\omega}_1 + \cdots + \tilde{\omega}_j)^3/\tilde{t}^3 = j\tilde{\nu}_3 + 3j(j-1)\tilde{\nu}_2 + j(j-1)(j-2)$ for all $j \in \mathbb{N}$ and hence $\mathbb{E}_X^0 h_1^3 = \tilde{t}^3\Big(\tilde{\nu}_3/(1-p) + \tilde{\nu}_2 6p/(1-p)^2 + 6p^2/(1-p)^3\Big)$. Since the relation $t = \tilde{t}/(1-p)$ holds, the lemma is concluded by Corollary B.2.5 (applied to X), Lemma B.3.4 and (B.2.9). □

B.4 Variance estimation

Traditionally, the variance of the classical Cavalieri estimator is approximated by its extension term, and hence an estimation of the extension term serves as an estimation of the variance as a whole. It is debatable if this is appropriate, as a particular sample may result in a large Zitterbewegung. Hence, using the extension term only, one risks actually underestimating the variance. However, at least in the stereological application that is discussed at the end of this paper, we see that the Zitterbewegung never can exceed the extension term, see (B.5.6), so $2\mathbb{V}\mathrm{ar}_E(\hat{V}_n)$ is an upper bound for the variance, only neglecting the remainder term in this application. The extension term is (relatively) easy to estimate and is traditionally used as an approximation for the unknown variance of the classical Cavalieri estimator (B.1.1). In view of the variance decomposition (B.2.5) we can follow these lines also for the new Newton-Cotes estimators and focus therefore on the estimation of the extension term. The estimation naturally depends on the available information on the sampling points. In this section we discuss a general estimation approach based on the observed process, and, if we sample from the perturbed or cumulative models with information on the initial process available, we mention another approach for the trapezoidal estimator exploiting the exact representations from Theorems B.3.3 and B.3.5.

Throughout this section we assume that the measurement function f is weakly $(m,1)$-piecewise smooth with known $m \in \mathbb{N}_0$, and we estimate the variance based on the points $X \cap L$ of the observed process X falling inside a bounded interval $L \subseteq \mathbb{R}$. We assume that f is known at the points $X \cap L$ and that L contains the support of f. We let the observed process X have intensity $1/t$.

Before proceeding, we introduce an estimation procedure for the Palm expectation of a function of increments as this will be of great relevance below. More precisely, for the point process X and a fixed $n \in \mathbb{N}$, we aim to estimate $\Theta = t^{-1}\mathbb{E}_X^0 F(h_1,\ldots,h_n)$, where $F : \mathbb{R}^n \to \mathbb{R}$ is an integrable function. For the n largest points in $X \cap L$, some of the subsequent n point increments are not accessible from the information of X in L. To correct for this censoring close to the boundary of L, we use the following Hanisch estimator, where we only consider points x in $X \cap L$ for which all of the subsequent n increments $h_1(x),\ldots,h_n(x)$ are observed:

$$\hat{\Theta} = \sum_{\substack{x \in X \cap L \\ s_n(x) \in L}} \frac{F\Big(h_1(x),\ldots,h_n(x)\Big)}{\mathcal{H}^1\Big(L \cap (L - s_n(x) + x)\Big)}. \tag{B.4.1}$$

Here $\mathcal{H}^1$ denotes the 1-dimensional Hausdorff measure, that is, the length measure. By the refined Campbell theorem [11, Theorem 3.5.3] it is seen that $\hat{\Theta}$ is unbiased for Θ. To estimate the intensity $1/t$ we simply use the unbiased estimator $\#(X \cap L)/\mathcal{H}^1(L)$.

As the extension term (B.2.6) factorizes similar to the classical case, we can estimate the contribution from the point process (through H_m) and from the measurement function (through g) separately. We first estimate the variance $H_m(0)$ depending on the point process X only. More specifically, if $n \in \mathbb{N}$ is the order of the Newton-Cotes estimator, using (Eq.'s (A.3.6) and (A.3.7), Paper A) and the refined Campbell theorem [11, Theorem 3.5.3], it is not difficult to see that there are (known) rational functions $p_m, q_m : \mathbb{R}^n \to \mathbb{R}$ such that

$$\begin{aligned} H_m(0) &= \mathbb{E}K_m^2(0) - (\mathbb{E}K_m(0))^2 \\ &= \tfrac{1}{t}\mathbb{E}_X^0 p_m(h_1,\dots,h_n) - (\tfrac{1}{t}\mathbb{E}_X^0 q_m(h_1,\dots,h_n))^2, \end{aligned}$$

where the latter term vanishes for all $m < n$; see (Lemma A.4.2, Paper A). For $n = 1$, comparing with Corollary B.2.5, we see that $p_0(h_1) = h_1^3/12$, $p_1(h_1) = h_1^5/120$ and $q_1(h_1) = h_1^3/12$. Applying (B.4.1) with F substituted by p_m and q_m, $H_m(0)$ can be estimated by the information of X in L.

Secondly, we estimate the derivative $g^{(2m+1)}(0^+)$. To do so, we apply similar techniques as in [7, Section 6.2], where we explicitly use the fact that the covariogram g is weakly $(2m+1,1)$-piecewise smooth. An induction argument using Lemma B.6.1 in the appendix (an adaption of the refined partial integration formula [7, Lemma 4.1] to weakly piecewise smooth functions) yields

$$g(y) = \sum_{j=0}^{2m+1} \frac{g^{(j)}(0^+)}{j!} y^j + R_{2m+1}(y;g) \tag{B.4.2}$$

for all $y > 0$, where the remainder R_{2m+1} satisfies

$$R_{2m+1}(y;g) = \sum_{a \in D_{g^{(2m+1)}} \cap (0,y)} \frac{(y-a)^{2m+1}}{(2m+1)!} J_{g^{(2m+1)}}(a) + \frac{1}{(2m+1)!}\int_0^y g^{(2m+2)}(t)(y-t)^{2m+1}\,\mathrm{d}t.$$

For $k \in \mathbb{N}_0$, we estimate the covariogram g using

$$\hat{g}(k,t) = t \sum_{x \in X \cap L} f(x) f(s_k(x)), \tag{B.4.3}$$

which, by the refined Campbell theorem [11, Theorem 3.5.3] and the fact that L contains the support of f, satisfies $\mathbb{E}\hat{g}(k,t) = \mathbb{E}_X^0 g(s_k)$. Let β_k be given by $\beta_k = g^{(k)}(0^+)/k!$. As $g^{(j)}$ is odd and continuous for all odd $j \le 2m$, we find using (B.4.2) that

$$\hat{g}(k,t) = \sum_{j=0}^{m} \mathbb{E}_X^0[s_k^{2j}]\beta_{2j} + \mathbb{E}_X^0[s_k^{2m+1}]\beta_{2m+1} + \epsilon_{k,t}, \tag{B.4.4}$$

where the error $\epsilon_{k,t}$ arises from a Taylor expansion and the covariogram estimation. Consequently, $\mathbb{E}\epsilon_{k,t} = \mathbb{E}_X^0 R_{2m+1}(s_k;g)$, and hence

$$\mathbb{E}\epsilon_{k,t} = \mathbb{E}_X^0\Big[\sum_{\substack{a \in D_{g^{(2m+1)}} \\ a>0}} \frac{(s_k-a)_+^{2m+1}}{(2m+1)!} J_{g^{(2m+1)}}(a) + \frac{1}{(2m+1)!}\int_0^\infty g^{(2m+2)}(r)(s_k-r)_+^{2m+1}\,\mathrm{d}r\Big],$$

where $(x-r)_+^k = \mathbf{1}_{x>r}(x-r)^k$. By definition, $s_k^j(0) = t^j s_k^j(0;X_u)$, and as $\mathbb{E}_{X_u}^0 s_k^j < \infty$ for all $k, j \in \mathbb{N}_0$ and g^{2m+2} is integrable by assumption, we conclude by dominated convergence that $\mathbb{E}\epsilon_{k,t} = o(t^{2m+1})$ as $t \downarrow 0$. Applying this convergence rate, the proposition below follows from (B.4.4).

Proposition B.4.1. *Let X be the observed process with intensity $1/t$ and let f be weakly $(m,1)$-piecewise smooth with covariogram g. For $N \geq m+1$ define the $(N+1)\times(m+2)$ matrix $Q_X = (\mathbb{E}_X^0 s_i^j)_{i\in\{0,\dots,N\},j\in\{0,2,\dots,2m,2m+1\}}$. If the linear system $Q_X\eta = (0,\dots,0,1)^\top$ has a solution $\eta = \{\eta_k\}_{k=0}^N$, then, with $\hat{g}(k,t)$ given by* (B.4.3),

$$\hat{\beta}_{2m+1} = \sum_{k=0}^{N} \hat{g}(k,t)\eta_k \tag{B.4.5}$$

estimates $g^{(2m+1)}(0^+)/(2m+1)!$ with a bias of order $o(1)$ as $t \downarrow 0$.

In the corollary below we present estimators for the extension term of the trapezoidal estimator when $m \in \{0,1\}$. To use Proposition B.4.1 for $m=1$, we require the denominator in (B.4.7) to be non-zero. This ensures that the linear system has a solution, and it is for instance always satisfied for the perturbed model. Note that for $m=0$ we choose $N=2$ larger than necessary since this is traditionally used in equidistant sampling; see e.g. [7, Formula (6.9)] and recall (B.2.10). If we instead choose $N=1$ we estimate $g'(0^+)$ by $(-\hat{g}(0,t)+\hat{g}(1,t))/t$.

Corollary B.4.2. *Let X be the observed process with intensity $1/t$ and let f be weakly $(m,1)$-piecewise smooth. Define $\tilde{g}(k,t) = \hat{g}(k,t)/t$, with $\hat{g}(k,t)$ given by* (B.4.3), *and $\gamma_{i,j} = \frac{1}{t}\mathbb{E}_X^0 s_i^j$. The extension term $\mathbb{V}\mathrm{ar}_E(\hat{V}_1(f))$ of the trapezoidal estimator is estimated with a bias of order $o(1)$ by:*

$$\Big(\tfrac{1}{4}\tilde{g}(0,t) - \tfrac{1}{3}\tilde{g}(1,t) + \tfrac{1}{12}\tilde{g}(2,t)\Big)\gamma_{1,3} \tag{B.4.6}$$

for $m=0$, and

$$\frac{\tilde{g}(0,t)(\gamma_{2,2}-\gamma_{1,2}) - \tilde{g}(1,t)\gamma_{2,2} + \tilde{g}(2,t)\gamma_{1,2}}{\gamma_{1,2}\gamma_{2,3} - \gamma_{2,2}\gamma_{1,3}}\Big(\tfrac{1}{10}\gamma_{1,5} - \tfrac{1}{12}(\gamma_{1,3})^2\Big) \tag{B.4.7}$$

for $m=1$, if the denominator in (B.4.7) *is non-zero. All $\gamma_{i,j}$ can be unbiasedly estimated using $X \cap L$ by the Hanish estimator* (B.4.1).

We have described a general estimation approach where the only requirement on the observed process X (other than stationarity) is the existence of a solution to the linear system in Proposition (B.4.1). However, if the model of X is known this can also be used to construct other estimates for $H_m(0)$ and the weights $\{\eta_k\}$ in (B.4.5). In particular, if both the observed process X and its underlying initial process $\tilde{X}$ are accessible, a decomposition of the extension term similar to Theorems B.3.3 and B.3.5 can be constructed, representing the Palm expectations in terms of the thinning probability p and increments of the initial process. Preliminary simulations for the trapezoidal estimator under perturbed and cumulative sampling indicate that such a model-specific estimation has a smaller variance than the general approach, and furthermore it appears to be much more robust for varying intensity, especially when thinning is present. This will be investigated further in a future paper.

B.5 An application to volume estimation

We now specialize the above results to stereological volume estimation. This had been the original starting point of our research and extends the settings of the well-established Cavalieri estimator and its generalizations. The target is the volume of a compact set $Y \subset \mathbb{R}^d$. For a unit vector $\omega \in S^{d-1}$ let $\omega^\perp$ be the hyperplane with normal ω and let $\mathcal{H}^k$ be the k-dimensional Hausdorff-measure in $\mathbb{R}^d$. We assume that the measurement function

$$f(x) = \mathcal{H}^{d-1}\Big(Y \cap (x\omega + \omega^\perp)\Big) \tag{B.5.1}$$

is available at all points x of a stationary point process X in $\mathbb{R}$. In order to apply the theory of Newton-Cotes estimators we assume throughout the following that Assumption B.2.1 is satisfied for a given $n \in \mathbb{N}$. Then, the Newton-Cotes estimator of order n is unbiased for $\int_{\mathbb{R}} f(x)\mathrm{d}x$, which is equal to $\mathcal{H}^d(Y)$ by Fubini's theorem. In the special case where the points in X are equidistant, Newton-Cotes estimators of any order coincide with the classical Cavalieri volume-estimator.

We recall a number of fundamental notions from convex geometry; see e.g. [10]. A set $Y \subset \mathbb{R}^d$ is called a *convex body*, if it is non-empty, compact and convex. We say that a hyperplane H in $\mathbb{R}^d$ *supports* the convex body Y, if $H \cap Y \neq \emptyset$ and one of the two open half-spaces generated by H does not contain any points of Y. If $H = x\omega + \omega^\perp$ supports Y and ω points in the direction of the open half-space that is disjoint from Y, the unit vector ω is called an *outer normal vector* of Y at the *support set* $H \cap Y$. We say that Y is *strictly convex* if its boundary does not contain any non-degenerate line-segments. A convex body is *of class* C^2_+ if its boundary is a regular submanifold of $\mathbb{R}^d$ that is twice continuously differentiable (in the sense of differential geometry) and all principal curvatures are positive at all boundary points. In particular, if Y is of class C^2_+, it is strictly convex.

If the convex body $Y \subset \mathbb{R}^d$ and a unit vector $\omega \in S^{d-1}$ are given, the support of the measurement function (B.5.1) is a compact interval $[x_-, x_+]$. The hyperplanes $H_- = x_-\omega + \omega^\perp$ and $H_+ = x_+\omega + \omega^\perp$ support Y at support sets Y_- and Y_+ with outer unit normals $-\omega$ and ω, respectively. The number $d(\omega) = x_+ - x_-$ is called the *width* of Y in direction ω. The function $\omega \mapsto d(\omega)$ is continuous on S^{d-1}. If the width $d(\omega)$ does not depend on ω, the set Y is called a *body of constant width*. Clearly, any ball is of constant width with $d = d(\omega)$ being its diameter, but there are other convex bodies of constant width in $\mathbb{R}^d$. However, among all *point symmetric bodies* Y (meaning that there is a point $z \in \mathbb{R}^d$ with $Y - z = \{-x : x \in Y - z\}$) balls can be characterized by having constant width. These and further results on bodies of constant width can be found in [5, p. 3.2]; see in particular Theorem 3.2.7 of this monograph.

The asymptotic variance behavior depends on the smoothness of f, which in turn reflects properties of the set Y. To illustrate how basic geometric regularity of a set Y yields smoothness properties of f, we restrict considerations to convex objects in three-dimensional space ($d = 3$), but note that generalizations to sets of higher dimension and with a boundary being a smooth manifold are also possible. As a consequence of the Brunn-Minkowski inequality [4, p. 361], the measurement function $f^{1/2}$ for a given $Y \subset \mathbb{R}^3$ is concave on its support $[x_-, x_+]$. If the support set Y_+ (Y_-) is a singleton or a line-segment, the function f is continuous in x_+ (x_-). In particular, f is continuous when Y is strictly convex.

If Y is of class C^2_+ the principal radii $r_1(\omega)$, $r_2(\omega)$ of curvature exist at the support point of Y with outer normal vector ω. The second normalized symmetric function $s_2(\omega) = r_1(\omega)r_2(\omega)$ of the principal radii is continuous and has integral

$$\int_{S^2} s_2(\omega)\mathcal{H}^2(\mathrm{d}\omega) = \mathcal{H}^2(\partial Y), \tag{B.5.2}$$

where the right side is the surface area of the boundary ∂Y of Y; see [10, (4.26) and (4.32)]. For later use, we also remark that the function s_2 determines the convex set Y of class C^2_+ uniquely up to translations; see [10, (4.26) and Theorem 8.1.1].

Proposition B.5.1. *Let $Y \subset \mathbb{R}^3$ be a convex body of class C^2_+. Then the measurement function f in* (B.5.1) *is twice continuously differentiable on $\mathbb{R} \setminus \{x_-, x_+\}$, and f' jumps exactly at the endpoints $x_- < x_+$ of the support of f. The jump at x_+ is $2\pi\sqrt{s_2(\omega)}$ and the jump at x_- is $2\pi\sqrt{s_2(-\omega)}$.*

If f'' has a right sided limit at x_- and a left sided limit at x_+ then f is weakly $(1,1)$*-piecewise smooth.*

Proof. We first set out to prove that f is twice continuously differentiable on $\mathbb{R} \setminus \{x_-, x_+\}$, and can restrict attention to the interior of its support. Translating Y appropriately, it is enough to show that f is twice continuously differentiable in a neighborhood of $x = 0$ when the origin is an interior point of Y. Assuming $0 \in \operatorname{int} Y$, we show first that the radial function $\rho_Y(u) = \max\{t \geq 0 : tu \in Y\}$, $u \in S^2$, is a twice continuously differentiable function on the sphere S^2. It is enough to show this claim in the neighborhood of one unit vector, which we may assume to coincide with the last standard basis vector e_3. As Y is of class C^2_+ and 0 is an interior point of Y, there is an open ball U in $e_3^\perp$ centered at 0 (and with a radius strictly smaller than 1) and a local C^2-parametrization $h : U \to \mathbb{R}$ such that $(x, h(x))$, $x \in U$, parametrizes a patch of the boundary of Y close to $(0, h(0))$. Any point u in a neighborhood of e_3 in S^2 can be written as $u = (x, \sqrt{1 - \|x\|^2})$, $x \in U$, and thus tu is a boundary point of Y if and only if $t\sqrt{1 - \|x\|^2} = h(tx)$. The implicit definition of $\rho_Y(u) = t$ through

$$F(x,t) = h(tx) - t\sqrt{1 - \|x\|^2} = 0$$

shows that ρ_Y is C^2 in a neighborhood of e_3 in S^2 by the implicit function theorem. As the origin is an interior point of Y, this implies that the Minkowski functional ρ_Y^{-1} of Y is in $C^2(S^2)$ and Y is therefore 2-smooth in the sense of [8]. Like in [8, Lemma 2.4] one now shows that f is twice continuously differentiable in a neighborhood of $x = 0$, noting that the origin-symmetry required in the statement of the lemma is not needed for the proof.

We have already seen that strict convexity implies continuity of f. We now show that the first derivative of f has finite jumps at the endpoints x_- and x_+ of its support. Without loss of generality, we may now assume $\omega = e_3$, that Y is rescaled and that the origin $0 \in Y$ is chosen such that e_3 is the boundary point of Y where the support plane at position $x_+ = 1$ meets Y. As we have already seen, there is a local C^2-parametrization $h : U \to \mathbb{R}$ with U as above, such that $(x, h(x))$, $x \in U$, parametrizes a patch of the boundary of Y close to $(0, h(0)) = e_3$. A first order Taylor expansion with a second order remainder term, using the fact that the gradient of h must be zero at $x = 0$, shows

$$h(x) = 1 - \tfrac{1}{2} x^\top A(\xi_x) x \tag{B.5.3}$$

in U, where $A(\cdot)$ is the Hessian matrix of $-h$ and ξ_x is a point on the line segment with endpoints 0 and x.

The eigenvalues λ_1 and λ_2 of $A(0)$ are the principal curvatures of Y at e_3 and thus coincide with $1/r_1(e_3)$ and $1/r_2(e_3)$ up to permutation [10, Section 2.5] and are positive by assumption. As $A(\cdot)$ is continuous, there is a compact neighborhood V of 0 in U such that $x^\top A(\xi_x)x \geq \frac{\min\{\lambda_1,\lambda_2\}}{2}\|x\|^2$ for all $x \in V$. This and the convexity of Y imply the existence of a constant ε_0 such that $M_\varepsilon = \{x \in U : h(x) \geq 1-\varepsilon\} \subset V$ for all $0 < \varepsilon < \varepsilon_0$. We claim that for such ε, the orthogonal projection of $Y_\varepsilon = Y \cap \Big((1-\varepsilon)e_3 + e_3^\perp\Big)$ onto $e_3^\perp$ coincides with M_ε, so

$$f(1-\varepsilon) = \mathcal{H}^2\Big(M_\varepsilon\Big). \tag{B.5.4}$$

In fact, if this was not the case, there would be a point $(x', t) \in Y_\varepsilon$ with $x' \notin U$. The convexity of Y implies that all points $x \in U$ of the line segment s between x' and 0 are in M_ε. But $s \cap (U \setminus V) \neq \emptyset$, so M_ε contains points outside V, a contradiction.

In view of (B.5.4) and (B.5.3) we have

$$\begin{aligned}\frac{1}{\varepsilon} f(1-\varepsilon) &= \mathcal{H}^2\Big(\frac{1}{\sqrt{\varepsilon}} M_\varepsilon\Big) \\ &= \mathcal{H}^2\Big(\{x \in \tfrac{1}{\sqrt{\varepsilon}} U : x^\top A(\xi_{\sqrt{\varepsilon}x})x \leq 2\}\Big) \\ &\to \mathcal{H}^2\Big(\{x \in e_3^\perp : x^\top A(0)x \leq 2\}\Big),\end{aligned}$$

as $\varepsilon \downarrow 0$, where the continuity of $A(\cdot)$ was used again. The limit set in the last displayed formula is an ellipse in $e_3^\perp$ with half axes $\sqrt{2/\lambda_i}$, $i = 1, 2$ and area $2\pi/\sqrt{\lambda_1\lambda_2}$. This implies

$$f'(1-) = -\frac{2\pi}{\sqrt{\lambda_1\lambda_2}} = -2\pi\sqrt{s_2(e_3)},$$

which is the negative of the jump of f' at $x_+ = 1$, as $f'(x) = 0$ for $x > x_+$. Replacing ω by $-\omega$, the above arguments show a corresponding equality for the jump at x_-.

As f'' exists and is continuous on $\mathbb{R} \setminus \{x_-, x_+\}$, this function can at most have two discontinuities. Under the assumption in the last claim of the theorem, these discontinuities are points where f'' has jumps, so f is weakly $(1,1)$-piecewise smooth. This concludes the proof. □

As mentioned in the introduction, there are practically relevant examples of measurement functions with unbounded second derivative. In particular, there are convex bodies of class C_+^2 with a measurement function f which is not $(1,1)$-piecewise smooth but only weakly $(1,1)$-piecewise smooth. An example can be constructed by modifying the measurement function $\pi(1-(1-x)^2) = \pi(x(2-x))$, $0 \leq x \leq 2$, of the unit ball centered at the point ω. We consider the measurement function

$$f(x) = \pi(x + x^{3/2})(2-x) = \pi(2x + 2x^{3/2} - x^2 - x^{5/2}),$$

$0 \leq x \leq 2$ (and zero elsewhere), of a body of revolution Y with axis ω. Because $f''(0^+) = \infty$, the function f is not $(1,1)$-piecewise smooth, but only weakly $(1,1)$-piecewise smooth. For the corresponding surface of revolution, the principal curvatures at a boundary point with level x are known to be

$$\begin{aligned}\kappa_1(x) &= \frac{1}{\rho(x)\sqrt{1+\rho'(x)^2}} \\ &= \frac{2}{\sqrt{(2\rho)^2(x) + (2 + 3x^{1/2} - 2x - \frac{5}{2}x^{3/2})^2}},\end{aligned}$$

and

$$\kappa_2(x) = -\frac{\rho''(x)}{\sqrt{1+\rho'(x)^2}^3} = -\rho^3(x)\rho''(x)\kappa_1^3(x),$$

where $\rho(x) = \sqrt{f(x)/\pi}$ is the radius of the section disk of Y at level x. This can be used to show that Y is a convex body of class C_+^2.

In the last statement of the proposition, we assumed that all discontinuity points of f'' are jump points. It is an open problem if this condition can be replaced by a more geometric assumption on Y. But it is well-known that the measurement function of any ellipsoid in $\mathbb{R}^3$ has a second derivative which satisfies this property for any direction ω. In addition, f'' is uniformly bounded in ω in this case; see, for instance, [7, Appendix B.1].

Corollary B.5.2. *For $\omega \in S^2$ assume that $Y \subset \mathbb{R}^3$ is a convex body of class C_+^2 with measurement function f supported by $[x_-, x_+]$, and such that f'' has one-sided limits at x_- and x_+. Then the covariogram g of the measurement function* (B.5.1) *is weakly* $(3,1)$*-piecewise smooth and $g^{(3)}$ has three jumps. These jumps have positions $-d(\omega)$,* 0 *and $d(\omega) = x_+ - x_-$ and their heights are*

$$J_{g^{(3)}}(0) = 4\pi^2(s_2(\omega) + s_2(-\omega)), \qquad \textit{and} \qquad J_{g^{(3)}}\Big(\pm d(\omega)\Big) = 4\pi^2\sqrt{s_2(\omega)s_2(-\omega)}.$$

Proof. That g is weakly $(3,1)$-piecewise smooth follows from Corollary B.6.3 of the appendix. The relation

$$J_{g^{(3)}}(c) = \sum_{b-a=c} J_{f'}(a)J_{f'}(b), \quad c \in \mathbb{R},$$

which is a special case of (B.6.3), in combination with Proposition B.5.1 yields the remaining claims. □

Note that $2g^{(3)}(0^+) = J_{g^{(3)}}(0)$ as $g^{(3)}$ is odd. Hence, $g^{(3)}(0^+) = 2\pi^2(s_2(\omega) + s_2(-\omega))$. When Y is point symmetric, we have $s_2(\omega) = s_2(-\omega)$ for all $\omega \in S^2$ and thus

$$J_{g^{(3)}}(0) = 2J_{g^{(3)}}\Big(\pm d(\omega)\Big) = 8\pi^2 s_2(\omega).$$

Theorem B.5.3. *Let $n \in \mathbb{N}$ be given. Assume that X_u is a unit-intensity stationary point process satisfying Assumption B.2.1, and that $Y \subset \mathbb{R}^3$ is a convex body of class C_+^2 whose measurement function supported by $[x_-, x_+]$ in direction $\omega \in S^2$ has a second derivative with one-sided limits at x_- and x_+. If $V_n(Y) = V_{n,\omega}(Y)$ is the nth Newton-Cotes estimator based on intersections of Y with the hyperplanes $\{x\omega + \omega^\perp : x \in tX_u\}$ with $t > 0$, then*

$$\mathbb{V}\mathrm{ar}(V_{n,\omega}(Y)) = \mathbb{V}\mathrm{ar}_E(V_{n,\omega}(Y)) + Z_\omega(t) + r_\omega(t), \tag{B.5.5}$$

where the extension term is given by

$$\mathbb{V}\mathrm{ar}_E(V_{n,\omega}(Y)) = 4\pi^2\Big(s_2(\omega) + s_2(-\omega)\Big)H_1(0; X_u)t^4,$$

and the Zitterbewegung

$$Z_\omega(t) = 8\pi^2\sqrt{s_2(\omega)s_2(-\omega)}\,H_1\Big(\tfrac{d(\omega)}{t}; X_u\Big)t^4$$

satisfies

$$Z_\omega(t) \le \mathbb{V}\mathrm{ar}_E(V_{n,\omega}(Y)). \tag{B.5.6}$$

The remainder $r_\omega(t)$ is of order $O(t^4)$. If X_u is weakly 1*-admissible, $r_\omega(t)$ is of order $o(t^4)$.*

We remark that (B.5.6) also holds for more general objects Y, as long as its measurement function f is weakly $(1,1)$-piecewise smooth and f' has exactly two jumps.

Proof. The explicit forms of $\mathbb{V}\mathrm{ar}_E(V_{n,\omega}(Y))$ and $Z_\omega(t)$ are obtained by inserting the jumps of the third derivative of the covariogram $g^{(3)}$ given in Corollary B.5.2 into the corresponding equations of Lemma B.2.4, taking (B.2.3) into account. The asymptotic behavior of $r_\omega(t)$ has also been established in Lemma B.2.4. The bound (B.5.6) follows from $H_1(\cdot;X_u) \le H_1(0;X_u)$ and the inequality of arithmetic and geometric means. □

The explicit expression for the extension term in Theorem B.5.3 also shows which directions $\omega \in S^2$ are best possible in terms of asymptotic variance when the Zitterbewegung is neglected. If possible, one should choose ω in the set of all unit vectors for which the average $(s_2 + \check{s}_2)/2$ attains its minimum, where $\check{s}_2$ denotes the reflection of s_2. When Y is point symmetric, this is the set where the second normalized symmetric curvature function s_2 is minimal. Of course, s_2 is not available in applications, but with this in mind, one might want to choose ω such that the corresponding support point has large curvature (that is, the vicinity is 'peaked').

A common strategy to determine the orientation of the hyperplane stack in applications is to randomize the direction ω in an isotropic way. We will write $\mathbb{E}_\omega$ for the expectation with respect to ω in this case. The assumptions of the following corollary are for instance satisfied if Y is an ellipsoid; see the comment right before Corollary B.5.2.

Corollary B.5.4. *Let the assumptions of Theorem B.5.3 be satisfied for all $\omega \in S^2$. Assume in addition that the second derivative of the measurement function is uniformly integrable in $\omega \in S^2$. If ω is a uniform random unit vector which is independent of X_u, then* (B.2.5) *with $m = 1$ holds with*

$$\mathbb{V}\mathrm{ar}_E(V_n(Y)) = 2\pi\mathcal{H}^2(\partial Y)H_1(0;X_u)t^4. \tag{B.5.7}$$

The term corresponding to the Zitterbewegung is

$$\overline{Z}(t) = 8\pi^2\,\mathbb{E}_\omega\Big[\sqrt{s_2(\omega)s_2(-\omega)}\,H_1\big(\tfrac{d(\omega)}{t};X_u\big)\Big]t^4,$$

and satisfies

$$\overline{Z}(t) \le \mathbb{V}\mathrm{ar}_E(V_n(Y)) \tag{B.5.8}$$

with equality if and only if Y is a ball with diameter d_0, say, and $H_1(d_0/t;X_u) = H_1(0;X_u)$. The remainder is of order $O(t^4)$, and even of order $o(t^4)$ when X_u is weakly 1-admissible.

Proof. Due to the law of total variance and the fact that $V_{n,\omega}(f)$ is unbiased for all ω, we have $\mathbb{V}\mathrm{ar}(V_n(Y)) = \mathbb{E}_\omega\mathbb{V}\mathrm{ar}\Big[V_{n,\omega}(Y) \mid |\omega\Big]$, where the latter variance has been described in Theorem B.5.3. Hence, taking expectations in (B.5.5) shows

$$\mathbb{V}\mathrm{ar}(V_n(Y)) = \mathbb{E}_\omega\mathbb{V}\mathrm{ar}_E(V_{n,\omega}(Y)) + \mathbb{E}_\omega Z_\omega(t) + \mathbb{E}_\omega r_\omega(t).$$

The first term on the right equals (B.5.7) due to (B.5.2) and the second term is obviously $\overline{Z}(t)$.

Now let g_ω denote the covariogram of the measurement function $f = f_\omega$ with respect to the direction $\omega \in S^2$. By (Proposition A.6.1, Paper A)

$$r_\omega(t) = t^4\int_{\mathbb{R}} g_\omega^{(4)}(s)H_1(\tfrac{s}{t};X_u)\mathrm{d}s,$$

and, for all ω, it is of order $o(t^4)$ if X_u is weakly 1-admissible. As $H_1(\cdot;X_u)$ is uniformly bounded and independent of ω, the asserted lower order properties of $r(t) = \mathbb{E}_\omega r_\omega(t)$ follow by dominated convergence if

$$\mathbb{E}_\omega \int_{\mathbb{R}} |g_\omega^{(4)}(s)|\mathrm{d}s < \infty. \tag{B.5.9}$$

The function $g_\omega^{(4)}$ is uniformly integrable in $\omega \in S^2$ by (B.6.1) and the assumption that f_ω'' is uniformly integrable in ω. Thus (B.5.9) follows.

We now fix $t > 0$. Inequality (B.5.8) clearly holds due to (B.5.6) and (B.5.2). If the former holds with equality for a given t, we must have

$$H_1\Big(\tfrac{d(\cdot)}{t};X_u\Big) = H_1(0;X_u) \tag{B.5.10}$$

almost surely, and $s_2 = \check{s}_2$, almost surely. The function s_2 is continuous by assumption, so this implies $s_2 = \check{s}_2$ on S^2, which shows that Y must be point symmetric, as Y is determined by s_2 up to translation. As the width $d(\cdot)$ is continuous and positive on S^2, its range is a compact interval in $(0,\infty)$, which, by the almost sure equality (B.5.10) and Lemma B.2.2, implies that the range is degenerate and hence Y has constant width. This means that $d(\omega)$ does not depend on ω, so there is a constant $d_0 > 0$ such that $d = d_0$ on S^2. But the only origin symmetric convex bodies with constant width d_0 are balls with diameter d_0. This and (B.5.10), which now reads $H_1(d_0/t;X_u) = H_1(0;X_u)$, shows one of the implications of the characterization of equality in (B.5.8). The other is trivially satisfied. □

Under the assumptions of Corollary B.5.4 one can also show $\overline{Z}(t) \geq -\mathbb{V}\mathrm{ar}_E(V_n(Y))$. However, this lower bound is not sharp even in the equidistant case.

B.6 Appendix

In this appendix we list generalizations of important results from [7] such that they now apply not only to functions with finitely many finite jumps, but also to integrable functions with finitely many, possibly infinite jumps. With these generalizations in mind, it is easily seen that all the results of Paper A involving $(m,1)$-piecewise smooth functions also apply for *weakly* $(m,1)$-piecewise smooth functions. One simply has to note that integrability of $f^{(m+1)}$ and $g^{(2m+2)}$ actually suffice where their boundedness was used in previous papers (here f denotes the measurement function and g its associated covariogram).

To state the results below, we follow the notation in [7], and let C_K be the set of all compactly supported piecewise continuous functions with finitely many finite jumps, and we let C_b be the set of all piecewise continuous bounded functions with locally finitely many and finite jumps. Furthermore, we define $\overline{C}_K$ to be the set of compactly supported piecewise continuous and integrable functions with finitely many (possibly infinite) jumps. In particular, if $h \in C_K$ such that h' is continuous in all but finitely many points with possibly infinite jumps, then h' is integrable by the first fundamental theorem of calculus, and thus $h' \in \overline{C}_K$.

The following generalization of [7, Lemma 4.1] follows simply by noticing that the integrability of h' is enough to guarantee the integrability of $h'\phi$ in its proof.

Lemma B.6.1 (Generalization of [7, Lemma 4.1]). *Let $h \in C_K$ and $\phi \in C_b$ such that $h' \in \overline{C}_K$ and $\phi' \in C_b$. Then*

$$\int_{\mathbb{R}} h(x)\phi'(x)\mathrm{d}x + \int_{\mathbb{R}} h'(x)\phi(x)\mathrm{d}x = -\sum_{a \in D_{h\phi}} J_{h\phi}(a).$$

In the proposition below, $\check{f}$ means the reflection of f, i.e. $\check{f}(x) = f(-x)$, and $*$ denotes convolution. Similarly, $\check{f}^{(k)}$ is the kth derivative of $\check{f}$, and $\check{J}_{f^{(k)}}$ is the reflection of the jump-function associated to $f^{(k)}$.

Proposition B.6.2 (Generalization of [7, Proposition 5.7]). *Let f be a function with all derivatives up to order m in C_K and derivative of order $m+1$ in $\overline{C}_K$. Then all derivatives up to order $2m+1$ of its covariogram g are in C_K, and the derivative of order $2m+2$ of g is in $\overline{C}_K$ with*

$$g^{(2k)} = f^{(k)} * \check{f}^{(k)} + \sum_{0 \le \ell < k} J_{f^{(\ell)}} * \check{f}^{(2k-\ell-1)} + \sum_{0 \le \ell < k} (-1)^{\ell+1} \check{J}_{f^{(\ell)}} * f^{(2k-\ell-1)}, \tag{B.6.1}$$

$$g^{(2k+1)} = f^{(k)} * \check{f}^{(k+1)} + \sum_{0 \le \ell < k} J_{f^{(\ell)}} * \check{f}^{(2k-\ell-1)} + \sum_{0 \le \ell \le k} (-1)^{\ell+1} \check{J}_{f^{(\ell)}} * f^{(2k-\ell-1)} \tag{B.6.2}$$

for all $2k$ and $2k+1$ less than or equal to $2m+2$.

Proof. The claim for $2k$ and $2k+1$ less than or equal to $2m$ is exactly [7, Proposition 5.7]. The expression (B.6.2) for $k = m$ and the claim that $g^{(2m+1)} \in C_K$ follow by differentiation from [7, Proposition 5.6] realizing that this proposition holds true even under the weaker assumption $f_1, f_2 \in C_K$ and $f_2' \in \overline{C}_K$, where the notation of this proposition has been adopted. One simply has to realize that $f_1 * f_2'$ is continuous also for $f_2' \in \overline{C}_K$.

The expression (B.6.1) for $k = m+1$ follows by differentiation using another generalization of [7, Proposition 5.6]: Let $f_1 \in \overline{C}_K$ and $f_2 \in C_K$ such that $f_2' \in \overline{C}_K$. Then $f_1 * f_2 \in C_K$ is continuous, $f_1 * f_2' \in \overline{C}_K$ and $(f_1 * f_2)' \in \overline{C}_K$ coincides with $f_1 * f_2' + J_{f_2} * f_1$ outside $D_{(f_1 * f_2)'}$. Moreover, $J_{f_2} * f_1 = J_{f_2} * J_{f_1}$. The proof follows along the same lines as that of [7, Proposition 5.6] using Lemma B.6.1 and the appropriate generalizations of [7, Lemmas 5.4 and 5.5]. □

The following corollary generalizes the claims of [7, Corollary 5.8] relevant for this paper and for Paper A.

Corollary B.6.3 (Generalization of [7, Corollary 5.8]). *If the function f is weakly $(m,1)$-piecewise smooth then its covariogram g is weakly $(2m+1,1)$-piecewise smooth with*

$$J_{g^{(2m+1)}} = (-1)^{m+1} \check{J}_{f^{(m)}} * J_{f^{(m)}}. \tag{B.6.3}$$

Proof. The fact that $g^{(k)}$ is continuous for $k \le 2m$ is shown in the original corollary. The fact that g is then weakly $(2m+1,1)$-piecewise smooth with the given jumps of $g^{(2m+1)}$ follows easily from Proposition B.6.2. □

Acknowledgments

This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

References

[1] Baddeley, A.J., K.-A. Dorph-Petersen and E.B.V. Jensen (2006). A note on the stereological implications of irregular spacing of sections. *J. Microsc.* 222(3), 177–181.

[2] Cruz-Orive, L.M (1989). On the precision of systematic sampling: a review of Matheron's transitive methods. *J. Microsc.* 153, 315–333.

[3] García-Fiñana, M. and L.M Cruz-Orive (2004). Improved variance prediction for systematic sampling on $\mathbb{R}$. *Statistics* 38, 243–272.

[4] Gardner, R.J. (2002). The Brunn–Minkowski inequality. *Bull. Amer. Math. Soc.* 39, 355–405.

[5] Gardner, R.J. (2006). *Geometric Tomography*. Second edition. Cambridge University Press, Cambridge.

[6] Kiderlen, M. and K.-A. Dorph-Petersen (2017). The Cavalieri estimator with unequal section spacing revisited. English. *Image Anal. Stereol.* 36(2), 135–141.

[7] Kiêu, K. (1997). *Three Lectures on Systematic Geometric Sampling*. Technical Report, Aarhus University.

[8] Koldobsky, A. (2005). *Fourier Analysis in Convex Geometry*. Vol. 116. Mathematical Surveys and Monographs. Am. Math. Soc., Providence RI.

[9] Matheron, G. (1965). *Les Variables régionalisées et leur Estimation*. Masson et Cie, Paris.

[10] Schneider, R. (2014). *Convex Bodies: The Brunn–Minkowski Theory*. Second edition. Cambridge University Press, Cambridge.

[11] Schneider, R. and W. Weil (2008). *Stochastic and Integral Geometry*. Springer, Heidelberg.

[12] Ziegel, J., A.J. Baddeley, K.-A. Dorph-Petersen and E.B.V. Jensen (2010). Systematic sampling with errors in sample locations. *Biometrika* 97, 1–13.

[13] Ziegel, J., E.B.V. Jensen and K.-A. Dorph-Petersen (2011). Variance estimation for generalized Cavalieri estimators. *Biometrika* 98, 187–198.

Paper C

Tail Asymptotics of an Infinitely Divisible Space-Time Model with Convolution Equivalent Lévy Measure

Mads Stehr and Anders Rønn-Nielsen

Accepted for publication by the Applied Probability Trust
(http://www.appliedprobability.org) in *Journal of Applied Probability* 58.1 (March 2021)

Abstract

We consider a space-time random field on $\mathbb{R}^d \times \mathbb{R}$ given as an integral of a kernel function with respect to a Lévy basis with a convolution equivalent Lévy measure. The field obeys causality in time and is thereby not continuous along the time-axis. For a large class of such random fields we study the tail behavior of certain functionals of the field. It turns out that the tail is asymptotically equivalent to the right tail of the underlying Lévy measure. Particular examples are the asymptotic probability that there is a time-point and a rotation of a spatial object with fixed radius, in which the field exceeds the level x, and that there is a time-interval and a rotation of a spatial object with fixed radius, in which the average of the field exceeds the level x.

MSC: 60G60; 60G17; 60E07; 60D05

Keywords: Convolution equivalence; infinite divisibility; Lévy-based modeling; asymptotic equivalence; sample paths for random fields

C.1 Introduction

In the present paper we investigate the extremal behavior of a space-time random field $(X_{v,t})_{(v,t)\in B\times[0,T]}$ defined by

$$X_{v,t} = \int_{\mathbb{R}^d\times(-\infty,t]} f(|v-u|, t-s)M(\mathrm{d}u, \mathrm{d}s), \tag{C.1.1}$$

where M is an infinitely divisible, independently scattered random measure on $\mathbb{R}^{d+1}$, $d \in \mathbb{N}$, f is some kernel function, and B and $[0,T]$ are compact index sets. We think of v and t as the position in space and time, respectively. Similarly, the first d coordinates of M refers to the spatial position, while the last coordinate is interpreted as time. The random field defined in (C.1.1) is a causal model in the sense that $X_{v,t}$ only depends on the noise, accounted for by M, up to time t, i.e. the restriction of M to $\mathbb{R}^d \times (-\infty, t]$. We shall make continuity assumptions on f ensuring that X is continuous in the space-direction. Discontinuities in the time-direction will however be possible, and we therefore have to pay particular attention to the assumptions on f to obtain sample paths that are both continuous in space and càdlàg in time; see Definition C.2.1, Assumption C.2.3, and Theorem C.5.7 below.

Lévy-driven moving average models, where a kernel function is integrated with respect to a Lévy basis, provide a flexible and tractable modeling framework and have been used for a variety of modeling purposes. Recent applications that, similarly to (C.1.1), include both time and space are modeling of turbulent flows ([5]) and growth processes ([13]). Spatial models without an additional time axis have e.g. been applied to define Cox point processes ([11]) and have served as a modeling framework for brain imaging data ([12, 22]). Lévy-based models for a stochastic process in time have gained recent popularity in finance. A simple example is a Lévy-driven Ornstein-Uhlenbeck process, with $f(t) = e^{-\lambda t}$, that has e.g. been used as a model for option pricing as illustrated in [6]. In [21] estimators for the mean and variogram in Lévy-driven moving average models are proposed, and central limit theorems for these estimators are derived.

In this paper, we will assume that the Lévy measure ρ of the random measure M has a convolution equivalent right tail ([8, 9, 16]) with index $\beta > 0$, with the notation $\mathcal{S}_\beta$ for this class of probability measures; see (C.2.2) and (C.2.3) below. Measures with a convolution equivalent tail cover the important cases of an inverse Gaussian and a normal inverse Gaussian (NIG) basis, respectively; see [19] and in particular examples 2.1 and 2.2. We derive that certain functionals of the field will have a right tail that is equivalent to the tail of the underlying Lévy measure. More precisely, we show that for a functional Ψ satisfying Assumptions C.3.1, C.3.5 and C.3.8 given below, there exist known constants C and c such that

$$\mathbb{P}(\Psi(X) > x) \sim C\rho((x/c, \infty)) \qquad \text{as } x \to \infty.$$

We give three important examples of the functional Ψ to illustrate the generality of the setting. The simplest is $\Psi(X) = \sup_{v,t} X_{v,t}$, where it is concluded that under appropriate assumptions on f it holds that $\sup_{v,t} X_{v,t}$ asymptotically has the same right tail as ρ.

A second example, see Example C.3.3, involves the spatial excursion set at level x and time t

$$A_{x,t} = \{v \in B : X_{v,t} > x\}. \tag{C.1.2}$$

Under some further regularity conditions we show that the asymptotic probability that there exists a t for which the excursion set at level x contains some rotation of an object D with a fixed radius has a tail that is equivalent with the tail of ρ. An asymptotic result for this probability will give information about the size of the excursion sets in the asymptotic scenario where $x \to \infty$. Studying the asymptotic behavior of excursion sets has previously appeared in the literature in various contexts. In [26] an overview is given of central limit theorem results concerning the volume of the excursion set for a broad class of stationary processes. Note however that statements on the volume are different from statements about excursion sets given here. As explained in the paragraph on convolution equivalence below, the literature offers some, although different, asymptotic results for excursion sets in the Gaussian and the subexponential case.

In concrete applications with spatial data, possibly observed over time, it is often of interest to detect locations where the observations are significantly large. A result for the probability defined above will make it possible to detect whether a cluster of neighboring locations, all with large observations, experienced at a certain time point within some period jointly constitutes an extreme observation. A specific application could be observations of temperatures in time and space. See e.g. [25] for an application of excursion sets to climate data.

In the last example, see Example C.3.4, we show a similar result for the probability that there is a time-interval and a translation and rotation of some fixed spatial object D such that the field in average, over both the time-interval and the resulting spatial object, exceeds the level x. While Example C.3.3 studied the probability of the excursion set being large enough to contain (a rotation and translation of) the set D, the present example has a slightly different scope. Here we are concerned with the probability that the field in some area (in time or space) in average is large. In a climate application this could be the existence of a 24–hour–average or a land area of a certain size with an average above a certain level.

In [7] sub–additive functionals of similar random fields, also with convolution equivalent tails, are studied. Here it is shown that under appropriate regularity conditions there exists constants $C_1 < C_2$ and a constant c such that

$$C_1\rho((x/c,\infty)) \le \mathbb{P}(\Psi(X) > x) \le C_2\rho((x/c,\infty)).$$

Note that the functional Ψ in the present paper is not necessarily required to be sub–additive. In particular, the functional corresponding to the excursion set framework is indeed not sub–additive.

In [19] and [20] the extremal behavior of spatial random fields of the form

$$X_v = \int_{\mathbb{R}^d} f(|v-u|)M(\mathrm{d}u) \qquad (v \in B)$$

is studied, when M is assumed to have a convolution equivalent Lévy measure. Here assumptions are imposed on the kernel function f to ensure that $v \mapsto X_v$ is continuous. Under some further regularity conditions it is shown in [19] that $\sup_{v\in B} X_v$ has a tail that is asymptotically equivalent with the tail of the underlying Lévy measure. In [20] this result is extended to the asymptotic probability that there exists a rotation of a fixed spatial object that is contained in the excursion set $A_x = \{v \in B : X_v > x\}$. The present paper extends the results of [19, 20] leading towards the generality of [7], as it includes a time-dependence in X and furthermore shows the asymptotic tail behavior for a certain class of functionals Ψ acting on X, including the ones considered in [19, 20]. By including the time-dependence, time-discontinuity and asymmetry are imposed on the kernel function, which necessitates particular care in the arguments. However, a few proofs are of structure similar to those of [19, 20] and are therefore found in the supplementary material.

In [10], results for a moving average process on $\mathbb{R}$, obtained as an integral with respect to a Lévy process with convolution equivalent tail, are derived. Here the process $(X_t)_{t\in[0,T]}$ is given by

$$X_t = \int_{-\infty}^{t} f(t-s)M(\mathrm{d}s),$$

where, again, M has a convolution equivalent Lévy measure. In agreement with the similar but more general result of the present paper for the field defined in (C.1.1) it is derived in [10] that $\sup_t X_t$ has a tail that asymptotically is equivalent with this.

Note that convolution equivalent distributions, as studied in the present paper, have heavier tails than Gaussian distributions and lighter tails than those of subexponential distributions (including regularly varying), that is, distributions in $\mathcal{S}_0$. For Gaussian random fields it is known that the distribution of the supremum of the field can be approximated by the expected Euler characteristic of an excursion set (see [4] and references therein). The extremal behavior of a non-Gaussian random field given by integrals with respect to an infinitely divisible random measure with a subexponential Lévy measure has already been studied in the literature. Results for the asymptotic distribution of the supremum are found in [23], and these results are refined to results on asymptotics for more high level geometric properties in [2] and [3]. The proofs rely heavily on the assumption of subexponential tails, where, asymptotically, the tail of a sum of independent variables is completely given in terms of that of one variable. For distributions in $\mathcal{S}_\beta$ with $\beta > 0$ this is not the case and the proofs can therefore not be applied in this context.

The paper is organised as follows. In Section C.2 we formally define the random field (C.1.1) and introduce some necessary assumptions for the field to be well defined and to have sample paths that are continuous in space and càdlàg in time. In Section C.3 we state and prove the main result for a general functional Ψ and introduce two specific examples of the functional. Some of the proofs in this section will apply many of the same techniques as in [19] and [20] and are therefore deferred to the supplementary material. In Section C.4 we state conditions for each of the two examples under which we afterwards show that the main result can be obtained. Section C.5 is devoted to showing that under appropriate regularity conditions, the field defined in (C.1.1) is continuous in space and càdlàg in time.

C.2 Preliminaries and initial assumptions

We define a Lévy basis to be an infinitely divisible and independently scattered random measure. Then the random measure M on $\mathbb{R}^{d+1}$ is independently scattered, such that for all disjoint Borel sets $(A_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^{d+1}$, the random variables $(M(A_n))_{n\in\mathbb{N}}$ are independent and furthermore satisfy $M(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} M(A_n)$. Furthermore, $M(A)$ is infinitely divisible for all Borel sets $A \subseteq \mathbb{R}^{d+1}$.

Moreover, in this paper we assume M to be a stationary and isotropic Lévy basis on $\mathbb{R}^{d+1}$. With $m(\cdot)$ denoting the Lebesgue measure, and $C(\lambda \dagger Y) = \log \mathbb{E}e^{i\lambda Y}$ the cumulant function for a random variable Y, this means that the random variable $M(A)$ has Lévy-Khintchine representation

$$C(\lambda \dagger M(A)) = i\lambda a m(A) - \frac{1}{2}\lambda^2 \theta m(A) + \int_{A\times\mathbb{R}} \left(e^{i\lambda x} - 1 - i\lambda x \mathbf{1}_{[-1,1]}(x)\right) F(\mathrm{d}u, \mathrm{d}x), \quad \text{(C.2.1)}$$

where $a \in \mathbb{R}$, $\theta \geq 0$ and F is the product measure $m \otimes \rho$ of the Lebesgue measure and a Lévy measure ρ. The notion of the so-called *spot variable* M' will be useful. It is a random variable equivalent in distribution to $M(A)$ when $m(A) = 1$.

We assume that the Lévy basis M has a convolution equivalent Lévy measure ρ with index $\beta > 0$, by which we formally mean that the probability measure ρ_1, the normalized restriction of ρ to $(1, \infty)$, is in $\mathcal{S}_\beta$. This means that $\rho_1 \in \mathcal{L}_\beta$, the class of probability measures with an exponential right tail β, i.e.

$$\frac{\rho_1((x-y,\infty))}{\rho_1((x,\infty))} \to e^{\beta y} \qquad \text{as } x \to \infty, \quad \text{(C.2.2)}$$

for all $y \in \mathbb{R}$, and that it furthermore satisfies the convolution property

$$\frac{(\rho_1 * \rho_1)((x,\infty))}{\rho_1((x,\infty))} \to 2\int_{\mathbb{R}} e^{\beta y} \rho_1(\mathrm{d}y) < \infty \qquad \text{as } x \to \infty, \quad \text{(C.2.3)}$$

where $*$ denotes convolution. To ease notation, we write $\rho \in \mathcal{S}_\beta$ when $\rho_1 \in \mathcal{S}_\beta$.

For later reference, we list the mentioned properties as part of Assumption C.2.2 below.

We write the tail of ρ as $\rho((x,\infty)) = L(x)\exp(-\beta x)$, so for all $y \in \mathbb{R}$, (C.2.2) implies that

$$\frac{L(x-y)}{L(x)} \to 1 \qquad \text{as } x \to \infty. \quad \text{(C.2.4)}$$

Equation (C.2.4) implies that the mapping $x \mapsto L(\log(x))$ is slowly varying. A consequence is (see formula (3.6) in [20]) that for all $\gamma > 0$ there exist $x_0 > 0$ and $C_0 > 0$ such that

$$\frac{L(\alpha x)}{L(x)} \leq C_0 \exp((\alpha - 1)\gamma x) \qquad \text{for all } x \geq x_0,\ \alpha \geq 1. \quad \text{(C.2.5)}$$

Before we define the kernel function f and consequently the field X, we introduce a continuity property called *t-càdlàg* which is of importance in this paper. Under assumptions on the basis M and the integration kernel f (appearing below), the entire field $X = (X_{v,t})$ exhibits this regularity, thus ensuring that the supremum of such fields on bounded sets behaves nicely; see e.g. the proof of Lemma C.3.12 in the supplementary material. Moreover, this continuity property will explicitly be used in the proofs of Section C.4.

Definition C.2.1 (t-càdlàg). *A field $(y_{v,t})_{(v,t)}$ is t-càdlàg if it for all (v,t) satisfies*

$$\lim_{(u,s)\to(v,t^-)} y_{u,s} \text{ exists in } \mathbb{R}, \qquad \text{and} \qquad \lim_{(u,s)\to(v,t^+)} y_{u,s} = y_{v,t}. \tag{C.2.6}$$

In defining the field $X = (X_{v,t})_{(v,t)\in B'\times T'}$ below, we make the following assumptions on the Lévy basis M and the integration kernel f. The assumptions are stronger than needed to ensure the existence of the integral (C.2.10) defining $X_{v,t}$: By [18, Theorem 2.7] the integral is well-defined if only the stationary and isotropic basis has a Lévy measure ρ satisfying $\int_{|y|>1}|y|\rho(\mathrm{d}y)<\infty$, and if the bounded integration kernel f is integrable in the sense of (C.2.8). However, we make the stronger assumptions below as these also guarantee the existence of a t-càdlàg version of X; see Theorem C.5.7.

Assumption C.2.2. *The Lévy basis M on $\mathbb{R}^{d+1}$ is stationary and isotropic with a Lévy measure $\rho \in \mathcal{S}_\beta$, $\beta > 0$. Moreover, ρ satisfies*

$$\int_{|y|>1} |y|^k \rho(\mathrm{d}y) < \infty \qquad \forall k \in \mathbb{N}. \tag{C.2.7}$$

Note that the integrability along the right tail is already given from the exponential tail property, and since ρ is a Lévy measure it also satisfies $\int_{[-1,1]} y^2\rho(\mathrm{d}y) < \infty$. Also, by [24, Theorem 25.3], (C.2.7) is equivalent to finite moments $\mathbb{E}|M'|^k < \infty$ of the spot variable. This is explicitly used when showing that there is a t-càdlàg version of X. Here we use a result from [1] which requires finite moments of a certain high order. In Sections C.2 to C.4, it is assumed that Assumption C.2.2 is satisfied.

Assumption C.2.3. *The kernel $f:[0,\infty)\times\mathbb{R}\to[0,\infty)$ is bounded, it satisfies $f(x,y)=0$ for all $x\in[0,\infty)$ and $y<0$, it is integrable in the sense that*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}} f(|u|,s)\mathrm{d}s\mathrm{d}u < \infty, \tag{C.2.8}$$

and it is Lipschitz continuous on $[0,\infty)\times[0,\infty)$, that is, there is $C_L \in (0,\infty)$ such that

$$|f(x_1,y_1) - f(x_2,y_2)| \le C_L |(x_1,y_1)-(x_2,y_2)| \tag{C.2.9}$$

for all $(x_1,y_1),(x_2,y_2) \in [0,\infty)\times[0,\infty)$.

Let $B \subseteq \mathbb{R}^d$ be a compact set with strictly positive Lebesgue measure, and consider $[0,T]$ for deterministic $0<T<\infty$. For $r,\ell \ge 0$ fixed, define the expanded sets $B' = B\oplus C_r(0) = \{x+y : x\in B, |y|\le r\}$ and $T' = [0,T+\ell]$. Here $C_r(u)\subseteq\mathbb{R}^d$ is the d-dimensional closed ball with radius r and center in $u\in\mathbb{R}^d$. Under Assumptions C.2.2 and C.2.3 we define the random field $X = (X_{v,t})_{(v,t)\in B'\times T'}$ by

$$X_{v,t} = \int_{\mathbb{R}^d\times\mathbb{R}} f(|v-u|, t-s) M(\mathrm{d}u,\mathrm{d}s). \tag{C.2.10}$$

Note that alternatively we can write

$$X_{v,t} = \int_{\mathbb{R}^d\times(-\infty,t]} f(|v-u|, t-s) M(\mathrm{d}u,\mathrm{d}s)$$

due to the assumptions on f. Thus X has a causal structure in the time direction in the sense that $X_{v,t}$ only depends on M restricted to the subset $\mathbb{R}^d \times (-\infty, t]$.

We are ultimately interested in extremal probabilities of the form

$$\mathbb{P}\Big(\Psi((X_{v,t})_{(v,t)\in B'\times T'}) > x\Big), \tag{C.2.11}$$

where $\Psi : \mathbb{R}^{B'\times T'} \to \mathbb{R}$ is a functional satisfying some assumptions that will be given in Section C.3. For notational convenience, we usually write $\Psi(y_{v,t})$, when applying Ψ to a field $(y_{v,t})_{(v,t)\in B'\times T'}$, however, when it is necessary to clarify the indices of the field, we write it fully. For the type of functionals Ψ we shall consider, it will be convenient to make some further assumptions on the kernel. The following Assumption C.2.4 clearly implies Assumption C.2.3 above. In Sections C.2 to C.4, Assumption C.2.4 is assumed satisfied.

Assumption C.2.4. *The kernel $f : [0,\infty)\times\mathbb{R} \to [0,\infty)$ satisfies $f(0,0) = 1$ and $f(x,y) = 0$ for all $x \in [0,\infty)$ and $y < 0$. Moreover,*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}} \sup_{v\in B'}\sup_{t\in T'} f(|v-u|, t-s)\mathrm{d}s\mathrm{d}u < \infty, \tag{C.2.12}$$

and f is Lipschitz on $[0,\infty)\times[0,\infty)$, i.e. it satisfies (C.2.9).

It turns out that the infinite divisibility of M is inherited to the field X. We shall spend the remainder of this section establishing this property and use it to obtain a useful representation of the field as an independent sum of a compound Poisson term and a term with a lighter tail than exponentials. The procedure is inspired by a similar technique used in [19], [20] and [23]. Here, we present the procedure fully to introduce all relevant notation.

The cumulant function of $X_{v,t}$ takes the form, cf. [18, Theorem 2.7],

$$\begin{aligned} C(\lambda \dagger X_{v,t}) &= i\lambda a \int_{\mathbb{R}^d}\int_{\mathbb{R}} f(|v-u|, t-s)\mathrm{d}s\mathrm{d}u - \frac{1}{2}\theta\lambda^2 \int_{\mathbb{R}^d}\int_{\mathbb{R}} f(|v-u|, t-s)^2\mathrm{d}s\mathrm{d}u \\ &+ \int_{\mathbb{R}^d}\int_{\mathbb{R}}\int_{\mathbb{R}}\Big(e^{if(|v-u|,t-s)\lambda z} - 1 - if(|v-u|, t-s)\lambda z 1_{[-1,1]}(z)\Big)\rho(\mathrm{d}z)\mathrm{d}s\mathrm{d}u. \end{aligned}$$

A similar expression can be obtained for any finite linear combination of $X_{v,t}$'s by substitution f with a relevant linear combination of f's. Thus, all finite-dimensional distributions of $(X_{v,t})_{(v,t)\in B'\times T'}$ are infinitely divisible, and consequently any countably indexed field $(X_{v,t})$ is infinitely divisible. Define the countable set $\mathbb{K} = (B'\times T')\cap\mathbb{Q}^{d+1}$, and let $\nu = (m\otimes m\otimes\rho)\circ H^{-1}$ be the measure on $(\mathbb{R}^{\mathbb{K}}, \mathcal{B}(\mathbb{R}^{\mathbb{K}}))$ defined as the image-measure of H on $m\otimes m\otimes\rho$, where $H : \mathbb{R}^d\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}^{\mathbb{K}}$ is given by

$$H(u,s,z) = (zf(|v-u|, t-s))_{(v,t)\in\mathbb{K}}.$$

Then direct manipulations show that ν is the Lévy measure of $(X_{v,t})_{(v,t)\in\mathbb{K}}$, and furthermore the Lévy-Khintchine representation is

$$\begin{aligned} C(\beta \dagger (X_{v,t})_{(v,t)\in\mathbb{K}}) &= i\sum_{(v,t)}\beta_{v,t}a_{v,t} - \frac{1}{2}\theta\int_{\mathbb{R}^d}\int_{\mathbb{R}}\Big(\sum_{(v,t)}\beta_{v,t}f(|v-u|, t-s)\Big)^2\mathrm{d}s\mathrm{d}u \\ &+ \int_{\mathbb{R}^{\mathbb{K}}}\Big(\mathrm{e}^{i\sum_{(v,t)}\beta_{v,t}z_{v,t}} - 1 - i\sum_{(v,t)}\beta_{v,t}z_{v,t}1_{[-1,1]^{\mathbb{K}}}(z)\Big)\nu(\mathrm{d}z) \end{aligned}$$

for suitable $(a_{v,t})_{(v,t)\in\mathbb{K}} \in \mathbb{R}^{\mathbb{K}}$. Here $\beta \in \mathbb{R}^{\mathbb{K}}$ with $\beta_{v,t} \neq 0$ for at most finitely many $(v,t) \in \mathbb{K}$. From the infinite divisibility, $(X_{v,t})_{(v,t)\in\mathbb{K}}$ can be represented as the independent sum

$$X_{v,t} = X^1_{v,t} + X^2_{v,t}.$$

The field $(X^1_{v,t})_{(v,t)\in\mathbb{K}}$ is a compound Poisson sum

$$X^1_{v,t} = \sum_{n=1}^{N} V^n_{v,t},$$

where N is Poisson distributed with intensity $\nu(A) < \infty$ and

$$A = \Big\{ z \in \mathbb{R}^{\mathbb{K}} : \sup_{(v,t)\in\mathbb{K}} z_{v,t} > 1 \Big\}.$$

The finiteness of $\nu(A)$ follows from arguments similar to those of [19, Lemma A.1] using (C.2.12). The fields $((V^n_{v,t})_{(v,t)\in\mathbb{K}})_{n\in\mathbb{N}}$ are i.i.d. with common distribution $\nu_1 = \nu_A/\nu(A)$, that is, the normalization of the restriction of ν to A. Also $(X^2_{v,t})_{(v,t)\in\mathbb{K}}$ is infinitely divisible and has Lévy measure ν_{A^c}, the restriction of ν to A^c.

It will be essential that there exist extensions of the fields $(X^1_{v,t})$ and $(X^2_{v,t})$ to $B' \times T'$ with t-càdlàg sample paths. In law, each of the fields $(V^n_{v,t})$ can be represented by $(Zf(|v-U|, t-S))_{(v,t)\in\mathbb{K}}$, where $(U,S,Z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}$ has distribution F_1, the normalized restriction of F to the set

$$H^{-1}(A) = \Big\{ (u,s,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} : \sup_{(v,t)\in\mathbb{K}} zf(|v-u|, t-s) > 1 \Big\}.$$

Hence, clearly a t-càdlàg extension $(V_{v,t})_{(v,t)\in B'\times T'}$ exists, and it is represented by $(Zf(|v-U|, t-S))_{(v,t)\in B'\times T'}$. As X^1 is a finite sum of such fields it also has an extension to $B' \times T'$ which is t-càdlàg. As mentioned above and shown in Theorem C.5.7, the entire field $(X_{v,t})_{(v,t)\in B'\times T'}$ has a version with t-càdlàg sample paths, and hence also X^2 has en extension with such paths.

C.3 Functional assumptions and main theorem

In this section we introduce assumptions on Ψ and related functionals, and we derive the main theorem on the asymptotic behavior of the extremal probability $\mathbb{P}(\Psi(X_{v,t}) > x)$ as $x \to \infty$. As the proofs of some of the results follow the same ideas as in [19] and [20], we refer to the supplementary material for these.

Throughout this section we shall assume the following.

Assumption C.3.1. *The functional $\Psi : \mathbb{R}^{B'\times T'} \to \mathbb{R}$ satisfies*

(i) *For all deterministic fields $(y_{v,t})_{(v,t)\in B'\times T'}$ and all $a \geq 0$ and $b \in \mathbb{R}$ it holds that*

$$\Psi(a y_{v,t} + b) = a\Psi(y_{v,t}) + b.$$

(ii) *Ψ is increasing, i.e.*

$$\Psi(y_{v,t} + z_{v,t}) \geq \Psi(y_{v,t})$$

whenever the field $(z_{v,t})_{(v,t)\in B'\times T'}$ satisfies that $z_{v,t} \geq 0$ for all $(v,t) \in B' \times T'$.

(iii) *For all $x > 0$, $u \in \mathbb{R}^d$ and $s \in \mathbb{R}$, there is a functional $\psi_{x,u,s,f} : \mathbb{R}^{B' \times T'} \to \mathbb{R}$ such that*

$$\Psi\Big(af(|v-u|, t-s) + y_{v,t}\Big) > x \qquad \text{if and only if} \qquad \psi_{x,u,s,f}(y_{v,t}) < a$$

for all $a \geq 0$ and all fields $(y_{v,t})$.

Proposition C.3.2. *The functionals Ψ and $\psi_{x,u,s,f}$ satisfy*

(i) *$\psi_{x,u,s,f}$ is decreasing, that is, for all $x > 0$, $u \in \mathbb{R}^d$ and $s \in \mathbb{R}$ and all fields $(y_{v,t})$*

$$\psi_{x,u,s,f}(y_{v,t}) \geq \psi_{x,u,s,f}(y_{v,t} + z_{v,t})$$

if $z_{v,t} \geq 0$ for all $(v,t) \in B' \times T'$.

(ii) *For all fields $(y_{v,t})$ and any constant $y \in \mathbb{R}$,*

$$\psi_{x,u,s,f}(y_{v,t} + y) = \psi_{x-y,u,s,f}(y_{v,t}).$$

(iii) *For all $x > 0$, $u \in \mathbb{R}^d$ and $s \in \mathbb{R}$ and all fields $(y_{v,t})$,*

$$\psi_{x,u,s,f}(y_{v,t}) \geq \psi_{x,u,s,f}(y^*) = \frac{x - y^*}{\Psi((f(|v-u|, t-s))_{(v,t)})},$$

where $y^* = \sup_{(v,t) \in B' \times T'} y_{v,t}$.

Proof. Statement (i) is seen as follows: Let x, u, s be fixed, and assume for contradiction the existence of $\epsilon > 0$ such that $\psi_{x,u,s,f}(y_{v,t}) + \epsilon = \psi_{x,u,s,f}(y_{v,t} + z_{v,t})$. Now choose a such that $a - \epsilon \leq \psi_{x,u,s,f}(y_{v,t}) < a$, and therefore $\psi_{x,u,s,f}(y_{v,t} + z_{v,t}) \geq a$. However, appealing to Assumption C.3.1(ii) and Assumption C.3.1(iii) we also conclude that

$$x < \Psi(af(|v-u|, t-s) + y_{v,t}) \leq \Psi(af(|v-u|, t-s) + y_{v,t} + z_{v,t}),$$

so also $\psi_{x,u,s,f}(y_{v,t} + z_{v,t}) < a$; a contradiction.

Part (ii) and (iii) are seen using Assumption C.3.1(i) and Assumption C.3.1(iii). □

Before giving two examples of functionals easily seen to satisfy Assumption C.3.1, we introduce some notation. Let $D \subseteq C_r(0) \subseteq \mathbb{R}^d$ be a fixed spatial object, and for all rotations $R \in SO(d)$ and translations $v \in \mathbb{R}^d$, define $D^R(v) = RD + v$. Similarly, let $D(v) = D + v$. Furthermore, let $I(t) = [t, t + \ell]$ for all $t \geq 0$. In Example C.3.3 below we assume that the set D in fact has radius $r/2 \geq 0$, by which we mean there is $\alpha \in \mathbb{S}^{d-1}$ such that $\{-\alpha r/2, \alpha r/2\} \subseteq D \subseteq C_{r/2}(0)$.

Example C.3.3. Suppose we are interested in the probability that there exist a time-point t, a translation v_0 and a rotation R of a given set D such that the field exceeds the level x on the entire set $\{t\} \times D^R(v_0)$. More formally, we assume that $D \subseteq C_{r/2}(0) \subseteq \mathbb{R}^d$ has radius $r/2$ and study the probability

$$\mathbb{P}\Big(\text{there exist } t \in [0,T], v_0 \in B, R \in SO(d) : X_{v,t} > x \text{ for all } v \in D^R(v_0)\Big).$$

To put this within the more general framework introduced in (C.2.11), we define Ψ by

$$\Psi(y_{v,t}) = \sup_{t \in [0,T]} \sup_{v_0 \in B} \sup_{R \in SO(d)} \inf_{v \in D^R(v_0)} y_{v,t}.$$

Consequently (obtained by straightforward manipulations),

$$\psi_{x,u,s,f}(y_{v,t}) = \inf_{t\in[0,T]}\inf_{v_0\in B}\inf_{R\in SO(d)}\sup_{v\in D^R(v_0)}\frac{x-y_{v,t}}{f(|v-u|,t-s)}.$$

Note that the probability above can be reformulated in terms of the excursion set $A_{x,t}$, defined in (C.1.2), as

$$\mathbb{P}\Big(\text{there exist } t\in[0,T], v_0\in B, R\in SO(d) : D^R(v_0)\subseteq A_{x,t}\Big).$$

Example C.3.4. Suppose we are interested in the probability that there is a time-interval and a location and rotation of the fixed spatial object D, in which the average of the field exceeds the level x. For this, let $D\subseteq C_r(0)\subseteq\mathbb{R}^d$ be given and consider the probability

$$\mathbb{P}\Big(\text{there exist } t_0\in[0,T], v_0\in B, R\in SO(d) : \tfrac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)}X_{v,t}\mathrm{d}t\mathrm{d}v > x\Big),$$

where $K=\int_D\int_0^\ell 1\,\mathrm{d}t\mathrm{d}v$. The set D can both be of full dimension in $\mathbb{R}^d$ and a subset of some lower dimensional subspace. In either case, $\mathrm{d}v$ refers to the relevant version of the Lebesgue measure. The special cases of

$$\mathbb{P}\Big(\text{there exist } t\in[0,T], v_0\in B, R\in SO(d) : \tfrac{1}{K}\int_{D^R(v_0)}X_{v,t}\mathrm{d}v > x\Big),$$

with a time-point instead of an interval (and K defined appropriately), and

$$\mathbb{P}\Big(\text{there exist } t_0\in[0,T], v\in B : \tfrac{1}{K}\int_{I(t_0)}X_{v,t}\mathrm{d}t > x\Big),$$

with a single spatial point, will be covered by the general formulation of the example, simply be defining $\int_{I(t_0)}X_{v,t}\mathrm{d}t = X_{t_0,v}$ when $\ell=0$, and $\int_{D^R(v_0)}X_{v,t}\mathrm{d}v = X_{t,v_0}$ when $D=\{0\}$ and hence $D^R(v_0)=\{v_0\}$. In the same spirit, the special case of

$$\mathbb{P}\Big(\text{there exist } t\in[0,T], v\in B : X_{v,t}>x\Big)$$

corresponds to letting $\ell=0$ and $D=\{0\}$. Note that this probability alternatively could be formulated as

$$\mathbb{P}\Big(\sup_{t\in[0,T]}\sup_{v\in B}X_{v,t}>x\Big).$$

To put this example in the framework of functionals, we define

$$\Psi(y_{v,t}) = \sup_{t_0\in[0,T]}\sup_{v_0\in B}\sup_{R\in SO(d)}\tfrac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)}y_{v,t}\,\mathrm{d}t\mathrm{d}v,$$

leading to

$$\psi_{x,u,s,f}(y_{v,t}) = \inf_{t_0\in[0,T]}\inf_{v_0\in B}\inf_{R\in SO(d)}\frac{x-\frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)}y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)}f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v}.$$

For the further arguments to hold it will be important that $\psi_{x,u,s,f}$ converges in a particular way as $x \to \infty$. The following assumption is satisfied under further case specific assumptions on the kernel f in each of the Examples C.3.3 and C.3.4 as illustrated in Section C.4.

Assumption C.3.5. *With the functionals Ψ and $\psi_{x,u,s,f}$ as in Assumption C.3.1, there exists c such that*

$$c = \Psi((f(|v-u|, t-s))_{(v,t)}) \tag{C.3.1}$$

for all $(u,s) \in B \times [0,T]$, and $\Psi((f(|v-u|, t-s))_{(v,t)}) < c$ for all $(u,s) \notin B \times [0,T]$. Furthermore, for all $(u,s) \in B \times [0,T]$ there is a functional $\lambda_{u,s} : \mathbb{R}^{B' \times T'} \to \mathbb{R}$, such that

$$\psi_{x,u,s,f}(y_{v,t}) - \frac{x}{c} + \lambda_{u,s}(y_{v,t}) \to 0 \tag{C.3.2}$$

as $x \to \infty$, holds for all t-càdlàg fields $(y_{v,t})_{(v,t)\in B' \times T'}$.

The following proposition is easily seen from Assumption C.3.1 and Proposition C.3.2.

Proposition C.3.6. *With c and $\lambda_{u,s}$ as in Assumption C.3.5 it holds that*

(i) *If the field $(y_{v,t})_{(v,t)\in B' \times T'}$ is constantly equal to $y \in \mathbb{R}$, then*

$$\lambda_{u,s}(y_{v,t}) = \lambda_{u,s}(y) = \frac{y}{c}.$$

(ii) *For all constants $y \in \mathbb{R}$ and fields $(y_{v,t})$,*

$$\lambda_{u,s}(y_{v,t} + y) = \lambda_{u,s}(y_{v,t}) + \frac{y}{c}.$$

(iii) *$\lambda_{u,s}$ is increasing.*

In the remainder of this section, it as assumed that also Assumption C.3.5 is satisfied.

The first step in proving the asymptotic behavior of the extremal probability $\mathbb{P}(\Psi(X_{v,t}) > x)$ is to consider the asymptotic behavior of extremal sets of a single jump-field $V = (V_{v,t})_{(v,t)\in B' \times T'}$ with distribution ν_1.

Theorem C.3.7. *Let $(V_{v,t})_{(v,t)\in B' \times T'}$ have distribution ν_1 and let $(y_{v,t})_{(v,t)\in B' \times T'}$ be t-càdlàg. As $x \to \infty$, it holds that*

$$\frac{\mathbb{P}(\Psi(V_{v,t} + y_{v,t}) > x)}{L(x/c)\exp(-\beta x/c)} \to \frac{1}{\nu(A)} \int_B \int_0^T \exp(\beta \lambda_{u,s}(y_{v,t}))\, \mathrm{d}s \mathrm{d}u. \tag{C.3.3}$$

Proof. For sufficiently large $x > 0$ we find

$$\begin{aligned} &\nu(A)\,\mathbb{P}(\Psi(V_{v,t} + y_{v,t}) > x) \\ &\quad = F\Big(\Big\{(u,s,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}_+ : \Psi\Big(zf(|v-u|, t-s) + y_{v,t}\Big) > x\Big\}\Big) \\ &\quad = F\Big(\Big\{(u,s,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}_+ : \psi_{x,u,s,f}(y_{v,t}) < z\Big\}\Big) \\ &\quad = \int_{B\times[0,T]} L\Big(\psi_{x,u,s,f}(y_{v,t})\Big) \exp\Big(-\beta \psi_{x,u,s,f}(y_{v,t})\Big)\, m(\mathrm{d}u, \mathrm{d}s) \\ &\qquad + \int_{(B\times[0,T])^c} L\Big(\psi_{x,u,s,f}(y_{v,t})\Big) \exp\Big(-\beta \psi_{x,u,s,f}(y_{v,t})\Big)\, m(\mathrm{d}u, \mathrm{d}s). \end{aligned} \tag{C.3.4}$$

First we show that the latter integral is of order $o(L(x/c)\exp(-\beta x/c))$ as $x \to \infty$. Let $y^* = \sup_{(v,t)\in B'\times T'} y_{v,t}$. Using Proposition C.3.2(iii) and that $x \mapsto L(x)\exp(-\beta x)$ is decreasing, we obtain that the second integral in (C.3.4) is bounded from above by

$$\int_{(B\times[0,T])^c} L\left(\frac{x-y^*}{\Psi(f(|v-u|,t-s))}\right)\exp\left(-\beta\frac{x-y^*}{\Psi(f(|v-u|,t-s))}\right)m(\mathrm{d}u,\mathrm{d}s).$$

Let $h(u,s;x)$ denote the integrand. For all $(u,s)\in(B\times[0,T])^c$ we have $\Psi(f(|v-u|,t-s)) < c$. In combination with (C.2.4) and (C.2.5), this implies the existence of $\gamma > 0$ and $C > 0$ such that

$$\frac{h(u,s;x)}{L(x/c)\exp(-\beta x/c)} \le C\exp(-\gamma x)$$

for sufficiently large x. Thus, $h(u,s;x)$ is of order $o(L(x/c)\exp(-\beta x/c))$ at infinity. By dominated convergence, also the integral is of order $o(L(x/c)\exp(-\beta x/c))$ if we can find an integrable function $g : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ such that

$$\frac{h(u,s;x)}{L(x/c)\exp(-\beta x/c)} \le g(u,s)$$

for all $(u,s)\in\mathbb{R}^d\times\mathbb{R}$. Returning to (C.2.5) we see that for all $0<\gamma<\beta/c$ there is $C>0$ and $x_0 > y^*$ such that

$$\frac{h(u,s;x)}{L(x/c)\exp(-\beta x/c)} \le C\exp\left(-(x_0-y^*)(\beta-\gamma c)\left(\frac{1}{\Psi(f(|v-u|,t-s))}-\frac{1}{c}\right)\right) \qquad \text{(C.3.5)}$$

for all $x \ge x_0$. Independent of (u,s) there is a constant $\tilde{C}$ such that the right hand side of (C.3.5) is bounded by

$$\tilde{C}\,\Psi(f(|v-u|,t-s)) \le \tilde{C}\sup_{(v,t)\in B'\times T'} f(|v-u|,t-s),$$

where we used Assumption C.3.1(i) and (ii). By Assumption C.2.4, this is integrable.

It remains to show that the first integral in (C.3.4) has the desired mode of convergence. For this, we have from (C.3.2), the representation of L, and the fact that ρ has an exponential tail, that for any $(u,s)\in B\times[0,T]$,

$$\frac{L\big(\psi_{x,u,s,f}(y_{v,t})\big)\exp\big(-\beta\psi_{x,u,s,f}(y_{v,t})\big)}{L\big(\frac{x}{c}\big)\exp\big(-\beta\frac{x}{c}\big)} \to \exp\left(\beta\lambda_{u,s}(y_{v,t})\right)$$

as $x \to \infty$. Since $x \mapsto L(x)\exp(-\beta x)$ is decreasing, we find using Proposition C.3.2(iii) that for sufficiently large x,

$$\frac{L\big(\psi_{x,u,s,f}(y_{v,t})\big)\exp\big(-\beta\psi_{x,u,s,f}(y_{v,t})\big)}{L\big(\frac{x}{c}\big)\exp\big(-\beta\frac{x}{c}\big)} \le \frac{L\big(\frac{x-y^*}{c}\big)\exp\big(-\beta\frac{x-y^*}{c}\big)}{L\big(\frac{x}{c}\big)\exp\big(-\beta\frac{x}{c}\big)} \le C\exp(\beta y^*/c),$$

for any $(u,s)\in B\times[0,T]$, where, according to (C.2.4), C is such that $L\big(\frac{x-y^*}{c}\big)/L\big(\frac{x}{c}\big) \le C$. As $B\times[0,T]$ is compact, the upper bound is integrable over $B\times[0,T]$ and (C.3.3) then follows by dominated convergence. □

The next step is to extend the relation (C.3.3) to an asymptotic result for $\mathbb{P}(\Psi(V_{v,t}^1 + \cdots + V_{v,t}^n + y_{v,t}) > x)$, where, for $i = 1,\ldots,n$, V^i are independent and identically distributed with common distribution ν_1. Here it will be useful to recall that each V^i can be represented by $(Z^i f(|v-U^i|, t-S^i))_{(v,t)\in B'\times T'}$, where (U^i, S^i, Z^i) has distribution F_1. Before being able to extend (C.3.3), we need a final assumption on the existence of a function ϕ ensuring sufficient integrability properties.

For the assumption we need some notation representing a deterministic version of the sum $V_{v,t}^1 + \cdots + V_{v,t}^n$. Thus let for each $i = 1,\ldots,n$ the field $(y_{v,t}^i)_{(v,t)\in B'\times T'}$ be given by

$$y_{v,t}^i = z^i f(|v-u^i|, t-s^i),$$

where all $z^i \geq 0$, $u^i \in \mathbb{R}^d$ and $s^i \in \mathbb{R}$.

Assumption C.3.8. *There exists a Lebesgue integrable function $\phi : \mathbb{R}^d \times \mathbb{R} \to [0,\infty)$ such that*

$$\phi(u,s)\begin{cases} = c & \text{for} \quad (u,s) \in B' \times T' \\ < c & \text{for} \quad (u,s) \notin B' \times T', \end{cases} \tag{C.3.6}$$

where $c > 0$ is the constant defined in (C.3.1)*.*

The function ϕ satisfies

$$\Psi\Big(\sum_{i=1}^n y_{v,t}^i\Big) \leq \sum_{i=1}^n z^i \phi(u^i, s^i), \tag{C.3.7}$$

and

$$\sup_{s\in[0,T]} \sup_{u\in B} \lambda_{u,s}\Big(\sum_{i=1}^n y_{v,t}^i\Big) \leq \frac{1}{c}\sum_{i=1}^n z^i \phi(u^i, s^i). \tag{C.3.8}$$

The definition of ϕ ensures that the tail of $Z\phi(U,S)$ is asymptotically equivalent to $\rho((x/c,\infty))$ and hence, $Z\phi(U,S)$ is convolution equivalent with index β/c; see Lemma C.3.9 below. Equation (C.3.7) then provides a convolution equivalent upper bound of the extremal probability for the functional Ψ of a sum of jump-fields. Finally, finiteness of relevant exponential moments of $\lambda_{u,s}$ applied to jump-fields is ensured by (C.3.8). This result is seen in Theorem C.3.10 below.

When showing the convolution equivalence of $Z\phi(U,S)$, we use the integrability of ϕ, although the weaker assumption that $\exp(-\gamma/\phi(u,s))$ is integrable for some $\gamma > 0$ is sufficient; see the proof of Lemma C.3.9 in the supplementary material. However, as seen in Section C.4, in practice $\phi(u,s)$ is often bounded by $\sup_{(v,t)} f(|v-u|, t-s)$ and hence its integrability follows from that of $\sup_{(v,t)} f(|v-u|, t-s)$.

In the remainder of this section it is also assumed that Assumption C.3.8 is satisfied. The proof of Lemma C.3.9 below follows by similar arguments as the proof of Theorem C.3.7 above, however, for completeness the proof can be found in the supplementary material.

Lemma C.3.9. *Let (U,S,Z) have distribution F_1. Then, as $x \to \infty$,*

$$\frac{\mathbb{P}(Z\phi(U,S) > x)}{L(x/c)\exp(-\beta x/c)} \to \frac{1}{\nu(A)} m(B' \times T'). \tag{C.3.9}$$

In particular, the distribution of $Z\phi(U,S)$ is convolution equivalent with index β/c and

$$\mathbb{E}\Big[\exp\Big(\frac{\beta}{c} Z\phi(U,S)\Big)\Big] < \infty. \tag{C.3.10}$$

As mentioned, the convolution equivalence of $Z\phi(U,S)$ is translated into a convolution equivalent upper bound for the extremal probability of a sum of jump-fields. Here (C.3.7) is applied together with the relation $\overline{F^{*n}}(x) \sim n\overline{F}(x)\Big(\int e^{\beta y}F(\mathrm{d}y)\Big)^{n-1}$, $x \to \infty$, when F is a convolution equivalent distribution with index β, F^{*n} is its n-fold convolution, and $\overline{F}$ is its tail. For this relation see e.g. [9, Corollary 2.11].

In Theorem C.3.10 below, a similar convolution equivalence for the sum of jump-fields is obtained.

Theorem C.3.10. *Let $V^1, V^2, \dots$ be i.i.d. fields with common distribution ν_1, and assume that $(y_{v,t})_{(v,t)\in B'\times T'}$ is t-càdlàg. For all $n \in \mathbb{N}$ it holds that*

$$\frac{\mathbb{P}(\Psi(V^1_{v,t} + \cdots + V^n_{v,t} + y_{v,t}) > x)}{\mathbb{P}(\Psi(V^1_{v,t}) > x)} \to \frac{n}{m(B\times[0,T])}\int_B\int_0^T \mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\Big(V^1_{v,t} + \cdots + V^{n-1}_{v,t} + y_{v,t}\Big)\Big)\Big]\mathrm{d}s\mathrm{d}u$$

as $x \to \infty$.

Recall that the field X^1 is defined as the compound Poisson sum with i.i.d. jump-fields $V^1, V^2, \dots$ and an independent Poisson distributed variable N with intensity $\nu(A) < \infty$. The following result on the extremal behavior of X^1 follows from Theorem C.3.10 by conditioning on the value of N.

Theorem C.3.11. *For each $(u,s) \in B\times[0,T]$ it holds that $\mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\Big(X^1_{v,t}\Big)\Big)\Big] < \infty$. For a field $(y_{v,t})_{(v,t)\in B'\times T'}$ satisfying* (C.2.6),

$$\frac{\mathbb{P}(\Psi(X^1_{v,t} + y_{v,t}) > x)}{L(x/c)\exp(-\beta x/c)} \to \int_B\int_0^T \mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\Big(X^1_{v,t} + y_{v,t}\Big)\Big)\Big]\mathrm{d}s\mathrm{d}u$$

as $x \to \infty$.

Now recall that we write the field $X = (X_{v,t})_{(v,t)\in B'\times T'}$ defined in (C.2.10) as the independent sum $X = X^1 + X^2$, where X^1 is the compound Poisson sum of fields with distribution ν_1. Also, the fields in the decomposition can be assumed to be t-càdlàg. One can show that the tail of $\sup_{(v,t)} X^2_{v,t}$ is lighter than that of $\Psi(X^1_{v,t})$, which is equivalent to the tail of ρ by Theorem C.3.11. Combining this fact with [16, Lemma 2.1], an argument based on independence and dominated convergence can be used to conclude Theorem C.3.13 below from Theorem C.3.11.

Lemma C.3.12. *For all $(u,s) \in B\times[0,T]$ it holds that $\mathbb{E}[\exp(\beta\lambda_{u,s}(X_{v,t}))] < \infty$.*

Theorem C.3.13. *Let the field X be given by* (C.2.10), *where the Lévy basis M satisfies Assumption C.2.2 and the kernel function f satisfies Assumption C.2.4. Let the functionals Ψ and $\lambda_{u,s}$ satisfy Assumptions C.3.1 and C.3.5, respectively. Then*

$$\lim_{x\to\infty}\frac{\mathbb{P}(\Psi(X_{v,t}) > x)}{\rho((x/c,\infty))} = \int_B\int_0^T \mathbb{E}[\exp(\beta\lambda_{u,s}(X_{v,t}))]\mathrm{d}s\mathrm{d}u.$$

C.4 Example results

In this section we return to Examples C.3.3 and C.3.4 to show versions of Theorem C.3.13 when Ψ is specifically given as in the examples. We make further assumptions on the kernel f that guarantee Assumptions C.3.5 and C.3.8.

In the setting of Example C.3.3 we assume the following.

Assumption C.4.1. *The kernel $f : [0,\infty)\times\mathbb{R} \to [0,\infty)$ is decreasing in both coordinates on $[0,\infty)\times[0,\infty)$, and it is strictly decreasing in the point $(r/2,0)$ in the sense that*

$$f(x,y) < f(r/2,0) \qquad \text{for all } (x,y) \in \Big([r/2,\infty)\times[0,\infty)\Big)\setminus\Big\{(r/2,0)\Big\}. \tag{C.4.1}$$

Moreover, the derivative $f_1(x) = \frac{\partial f}{\partial x}(x,0)$ exists for all $x \geq 0$, and there is a function g such that

$$g(x) = f_1(r/2)(x-r/2) + f(r/2,0) \tag{C.4.2}$$

for all $x \in [0,r]$, where also $f(x,0) \leq g(x)$ for all $x \in [0,r]$.

Such a g exists in particular when f is concave on $[0,r]$. The following lemma shows that Assumption C.3.5 is satisfied when the kernel satisfies Assumption C.4.1.

Lemma C.4.2. *If Ψ and $\psi_{x,u,s,f}$ are given as in Example C.3.3 and f satisfies Assumption C.4.1, then Assumption C.3.5 is satisfied with $c = f(r/2,0)$. Furthermore, for a t-càdlàg field $y = (y_{v,t})_{(v,t)\in B'\times[0,T]}$, the functional $\lambda_{u,s}$ takes the form*

$$\lambda_{u,s}\Big((y_{v,t})_{(v,t)\in B'\times[0,T]}\Big) = \lambda_u\Big((y_{v,s})_{v\in B'}\Big)$$

for a functional $\lambda_u : \mathbb{R}^{B'} \to \mathbb{R}$.

Proof. From [20, Lemma 3.1] we have, for fixed $s \in [0,T]$ and for all $u \in B$, a functional λ_u such that

$$\inf_{v_0\in B}\inf_{R\in SO(d)}\sup_{v\in D^R(v_0)}\frac{x-y_{v,s}}{f(|v-u|,0)} - \frac{x}{f(r/2,0)} + \lambda_u((y_{v,s})_{v\in B'}) \to 0 \tag{C.4.3}$$

as $x \to \infty$. With $\lambda_{u,s}$ defined by $\lambda_{u,s}\Big((y_{v,t})_{(v,t)\in B'\times[0,T]}\Big) = \lambda_u((y_{v,s})_{v\in B'})$, we claim that Assumption C.3.5 is satisfied. For notational convenience, we write $C = -\lambda_u(y_{v,s})$.

For all sufficiently large x, we can choose $t_x \in [s,T]$, $v_x \in B$ and $R_x \in SO(d)$ such that

$$\sup_{v\in D^{R_x}(v_x)}\frac{x-y_{v,t_x}}{f(|v-u|,t_x-s)} = \inf_{t\in[0,T]}\inf_{v_0\in B}\inf_{R\in SO(d)}\sup_{v\in D^R(v_0)}\frac{x-y_{v,t}}{f(|v-u|,t-s)}.$$

With $y^* = \sup_{(v,t)\in B'\times[0,T]} y_{v,t}$ and $y_* = \inf_{(v,t)\in B'\times[0,T]} y_{v,t}$, we then find

$$\begin{aligned}\frac{x-y^*}{\inf_{v\in D^{R_x}(v_x)} f(|v-u|,t_x-s)} &\leq \sup_{v\in D^{R_x}(v_x)}\frac{x-y_{v,t_x}}{f(|v-u|,t_x-s)}\\ &\leq \sup_{v\in D^{R_x}(u)}\frac{x-y_{v,s}}{f(|v-u|,0)} \leq \frac{x-y_*}{f(r/2,0)}.\end{aligned}$$

Going to the limit $x\to\infty$ and using that $\inf_{v\in D^R(v_0)} f(|v-u|,t-s) \leq f(r/2,0)$, shows $\inf_{v\in D^{R_x}(v_x)} f(|v-u|,t_x-s) \to f(r/2,0)$ as $x\to\infty$. For all $(v_0,t) \neq (u,s)$ and $R \in SO(d)$

we in fact have that $\inf_{v\in D^R(v_0)} f(|v-u|, t-s) < f(r/2, 0)$, and thus the convergence implies that also $v_x \to u$ and $t_x \to s$. We will show the desired convergence

$$\sup_{v\in D^{R_x}(v_x)} \frac{x - y_{v,t_x}}{f(|v-u|, t_x - s)} - \frac{x}{f(r/2,0)} \to C \qquad (x\to\infty)$$

by contradiction. Since

$$\sup_{v\in D^{R_x}(v_x)} \frac{x - y_{v,t_x}}{f(|v-u|, t_x - s)} \leq \inf_{v_0\in B} \inf_{R\in SO(d)} \sup_{v\in D^R(v_0)} \frac{x - y_{v,s}}{f(|v-u|, 0)},$$

we assume the existence of $\epsilon > 0$ and a sequence (x_n), $x_n \to \infty$, such that

$$\sup_{v\in D^{R_n}(v_n)} \frac{x_n - y_{v,t_n}}{f(|v-u|, t_n - s)} - \frac{x_n}{f(r/2,0)} \leq C - \epsilon \tag{C.4.4}$$

for all n, where $t_n = t_{x_n}$, $v_n = v_{x_n}$ and $R_n = R_{x_n}$. By t-càdlàg properties of the y-field, we can find n_0 such that

$$\sup_{v\in B'} \left| \frac{y_{v,t_n} - y_{v,s}}{f(|v-u|, 0)} \right| \leq \frac{\epsilon}{2}$$

for all $n \geq n_0$. Consequently and using that f is decreasing and (C.4.4)

$$\begin{aligned} \sup_{v\in D^{R_n}(v_n)} \frac{x_n - y_{v,s}}{f(|v-u|, 0)} - \frac{x_n}{f(r/2,0)} &\leq \sup_{v\in D^{R_n}(v_n)} \frac{x_n - y_{v,t_n}}{f(|v-u|, 0)} - \frac{x_n}{f(r/2,0)} + \frac{\epsilon}{2} \\ &\leq \sup_{v\in D^{R_n}(v_n)} \frac{x_n - y_{v,t_n}}{f(|v-u|, t_n - s)} - \frac{x_n}{f(r/2,0)} + \frac{\epsilon}{2} \leq C - \frac{\epsilon}{2}, \end{aligned}$$

which contradicts the limit relation (C.4.3). □

Theorem C.4.3. *Let the field X be given by* (C.2.10)*, where the Lévy basis M satisfies Assumption C.2.2 and the kernel function f satisfies Assumptions C.2.4 and C.4.1. Let $D \subseteq C_{r/2}(0)$ have radius $r/2 > 0$ and let Ψ be defined by*

$$\Psi(y_{v,t}) = \sup_{t\in[0,T]} \sup_{v_0\in B} \sup_{R\in SO(d)} \inf_{v\in D^R(v_0)} y_{v,t}.$$

Furthermore, let $\lambda_{u,s}$ be the functional given in Lemma C.4.2 and write $c = f(r/2,0)$. Then

$$\lim_{x\to\infty} \frac{\mathbb{P}(\Psi(X_{v,t}) > x)}{\rho((x/c,\infty))} = \int_B \int_0^T \mathbb{E}\left[\exp(\beta\lambda_{u,s}(X_{v,t}))\right] \mathrm{d}s\mathrm{d}u.$$

Proof. The result follows from Theorem C.3.13 and Lemma C.4.2 once we show the existence of a function ϕ satisfying Assumption C.3.8. Now define ϕ as

$$\phi(u,s) = f(r/2,0)\mathbf{1}_{B'\times[0,T]}(u,s) + \sup_{t\in[0,T]} \sup_{v\in B\oplus C_{r/2}} f(|v-u|, t-s)\mathbf{1}_{(B'\times[0,T])^c}(u,s),$$

which is integrable by (C.2.12) and satisfies (C.3.6) by (C.4.1). By a combination of Lemma C.4.2 and [20, Lemma 3.2] we find that

$$\lambda_{u,s}(y_{v,t}) = \frac{1}{2f(r/2,0)} \sup_{\alpha\in\mathbb{S}^{d-1}} (y_{u+\alpha r/2,s} + y_{u-\alpha r/2,s}) \tag{C.4.5}$$

for all $(u,s) \in B \times [0,T]$, if $D = \{-\alpha r/2, \alpha r/2\}$ for some $\alpha \in \mathbb{S}^{d-1}$. Adapting the proof of [20, Lemma 3.3] to this time-dependent setting, it is seen using (C.4.5) that (C.3.7) and (C.3.8) follow when it is shown that

$$\frac{1}{2}\Big(y^i_{u+\alpha r/2,s} + y^i_{u-\alpha r/2,s}\Big) \leq z^i \phi(u^i, s^i) \qquad \forall (u,s) \in B \times [0,T]\,, \alpha \in \mathbb{S}^{d-1}, \tag{C.4.6}$$

with $y^i_{v,t}$ defined as just before Assumption C.3.8. Since f is decreasing in both coordinates,

$$\frac{1}{2}\Big(y^i_{u+\alpha r/2,s} + y^i_{u-\alpha r/2,s}\Big) \leq \frac{z^i}{2}\Big(f(|u+\alpha r/2 - u^i|, 0) + f(|u-\alpha r/2 - u^i|, 0)\Big).$$

Using the upper bound g assumed by (C.4.2), arguments as in [20, Lemma 3.3] show that (C.4.6) is satisfied when $(u^i, s^i) \in B' \times [0,T]$. When $(u^i, s^i) \in (B' \times [0,T])^c$ it is immediately seen that

$$\frac{1}{2}\Big(y^i_{u+\alpha r/2,s} + y^i_{u-\alpha r/2,s}\Big) \leq z^i \sup_{t\in[0,T]} \sup_{v \in B \oplus C_{r/2}} f(|v-u^i|, t-s^i) = z^i \phi(u^i, s^i).$$

This concludes the proof. □

In the setting of Example C.3.4 the following is assumed.

Assumption C.4.4. *For the set $D \subseteq C_r(0) \subseteq \mathbb{R}^d$ the kernel function f satisfies*

$$\int_{D(v_0)} \int_{I(t_0)} f(|v-u|, t-s)\mathrm{d}t\mathrm{d}v < \int_D \int_0^\ell f(|v|, t)\mathrm{d}t\mathrm{d}v \tag{C.4.7}$$

for all $(v_0, t_0) \neq (u,s) \in \mathbb{R}^d \times \mathbb{R}$.

Lemma C.4.5. *Let $y = (y_{v,t})_{(v,t)\in B'\times T'}$ be a t-càdlàg field. For all $(u,s) \in B \times [0,T]$ it holds that*

$$\begin{aligned} \inf_{t_0\in[0,T]} \inf_{v_0 \in B} \inf_{R\in SO(d)} & \frac{x - \frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} f(|v-u|, t-s)\,\mathrm{d}t\mathrm{d}v} \\ & - \frac{x}{c} + \sup_{R\in SO(d)} \frac{1}{c}\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\mathrm{d}t\mathrm{d}v \to 0 \end{aligned} \tag{C.4.8}$$

as $x \to \infty$, where $c = \frac{1}{K}\int_D \int_0^\ell f(|v|, t)\mathrm{d}t\mathrm{d}v$. That is, with Ψ and $\psi_{x,u,s,f}$ as in Example C.3.4, and with $\lambda_{u,s}(y_{v,t}) = \sup_R \frac{1}{c}\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\mathrm{d}t\mathrm{d}v$, Assumption C.3.5 is satisfied.

Proof. For all sufficiently large $x > 0$, choose $t_x \in [s,T]$, $v_x \in B$ and $R_x \in SO(d)$ with

$$\inf_{t_0, v_0, R} \frac{x - \frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} f(|v-u|, t-s)\,\mathrm{d}t\mathrm{d}v} = \frac{x - \frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} f(|v-u|, t-s)\,\mathrm{d}t\mathrm{d}v}.$$

By definition of t_x and v_x we find that

$$
\begin{aligned}
\frac{x-y^*}{\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v} &\le \frac{x-\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v} \\
&\le \inf_{R\in SO(d)} \frac{x-\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D(u)}\int_{I(s)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v} \\
&= \frac{x-\sup_{R\in SO(d)}\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{c},
\end{aligned}
$$

where $y^* = \sup y_{v,t}$. Rearranging and noting that $\frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v < c$ for all $(v_0,t_0)\neq(u,s)$ and any $R\in SO(d)$, we conclude that $\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v \to c$ and consequently $v_x\to u$ and $t_x\to s$ as $x\to\infty$. Since the field $(y_{v,t})$ is t-càdlàg, we furthermore find that, as $x\to\infty$,

$$
\sup_{R\in SO(d)}\int_{D^R(v_x)}\int_{I(t_x)} y_{v,t}\,\mathrm{d}t\mathrm{d}v \to \sup_{R\in SO(d)}\int_{D^R(u)}\int_{I(s)} y_{v,t}\,\mathrm{d}t\mathrm{d}v.
$$

Recalling that $\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v \le c$ for all x, and turning to the inequalities above, we conclude (C.4.8) by

$$
\begin{aligned}
0 &\le \frac{x-\sup_R\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{c} - \frac{x-\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} f(|v-u|,t-s)\,\mathrm{d}t\mathrm{d}v} \\
&\le \frac{\frac{1}{K}\int_{D^{R_x}(v_x)}\int_{I(t_x)} y_{v,t}\,\mathrm{d}t\mathrm{d}v - \sup_R\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{c} \\
&\le \frac{\sup_R\frac{1}{K}\int_{D^R(v_x)}\int_{I(t_x)} y_{v,t}\,\mathrm{d}t\mathrm{d}v - \sup_R\frac{1}{K}\int_{D^R(u)}\int_{I(s)} y_{v,t}\,\mathrm{d}t\mathrm{d}v}{c} \to 0
\end{aligned}
$$

as $x\to\infty$. □

Theorem C.4.6. *Let the field X be given by* (C.2.10), *where the Lévy basis M satisfies Assumption C.2.2 and the kernel function f satisfies Assumptions C.2.4 and C.4.4. Let $D\subseteq C_r(0)\subseteq\mathbb{R}^d$ for $r\ge 0$ be given, and let Ψ be defined by*

$$
\Psi(y_{v,t}) = \sup_{t_0\in[0,T]}\sup_{v_0\in B}\sup_{R\in SO(d)}\frac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} y_{v,t}\,\mathrm{d}t\mathrm{d}v,
$$

where $K=\int_D\int_0^\ell 1\,\mathrm{d}t\mathrm{d}v$. Furthermore, let $c=\frac{1}{K}\int_D\int_0^\ell f(|v|,t)\mathrm{d}t\mathrm{d}v$. Then

$$
\lim_{x\to\infty}\frac{\mathbb{P}(\Psi(X_{v,t})>x)}{\rho((x/c,\infty))} = m(B\times[0,T])\mathbb{E}\Big[\exp\Big(\beta\sup_{R\in SO(d)}\frac{1}{c}\frac{1}{K}\int_{D^R(u)}\int_{I(s)} X_{v,t}\mathrm{d}t\mathrm{d}v\Big)\Big],
$$

where $(u,s)\in B\times[0,T]$ is chosen arbitrarily.

Proof. The result follows from Theorem C.3.13 and Lemma C.4.5 once we show the existence of a function ϕ satisfying Assumption C.3.8. Note that the integrand in the limit in Theorem C.3.13 is constant due to the stationarity of X and $\lambda_{u,s}$. Define

$$
\phi(u,s) = c\,\mathbf{1}_{B'\times T'}(u,s) + \sup_{t_0\in[0,T]}\sup_{v_0\in B}\frac{1}{K}\int_{D(v_0)}\int_{I(t_0)} f(|v-u|,t-s)\mathrm{d}t\mathrm{d}v\,\mathbf{1}_{(B'\times T')^c}(u,s),
$$

which is integrable by (C.2.12) and satisfies (C.3.6) by (C.4.7). Now let $n \in \mathbb{N}$ be fixed, and let $(y^i_{v,t})_{(v,t)\in B'\times T'}$ for $i = 1,\dots,n$ be t-càdlàg fields. Then

$$\begin{aligned}\Psi\Big(\sum_{i=1}^n y^i_{v,t}\Big) &= \sup_{t_0\in[0,T]}\sup_{v_0\in B}\sup_{R\in SO(d)} \tfrac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)}\sum_{i=1}^n y^i_{v,t}\,\mathrm{d}t\mathrm{d}v\\ &\le \sum_{i=1}^n \sup_{t_0\in[0,T]}\sup_{v_0\in B}\sup_{R\in SO(d)} \tfrac{1}{K}\int_{D^R(v_0)}\int_{I(t_0)} y^i_{v,t}\,\mathrm{d}t\mathrm{d}v = \sum_{i=1}^n \Psi(y^i_{v,t}).\end{aligned}$$

Furthermore, if $y^i_{v,t} = z^i f(|v-u^i|, t-s^i)$, it is easily seen that $\Psi(y^i_{v,t}) \le z^i\phi(u^i,s^i)$, and hence, (C.3.7) is satisfied. Since

$$\sup_{s\in[0,T]}\sup_{u\in B}\lambda_{u,s}\big(y^i_{v,t}\big) = \frac{1}{c}\Psi(y^i_{v,t}),$$

(C.3.8) is also satisfied, which concludes the proof. □

As mentioned in Example C.3.4, the case of $\mathbb{P}(\sup_{t\in[0,T]}\sup_{v\in B} X_{v,t} > x)$ follows from Theorem C.4.6 by letting $\ell = 0$ and $D = \{0\}$. In this case, the constant $c = f(0,0) = 1$ and (C.4.7) translates into $f(|v_0-u|, t_0-s) < f(0,0)$ for all $(v_0,t_0) \ne (u,s)$, or equivalently $f(x,y) < f(0,0)$ for all $(x,y) \ne (0,0)$.

Theorem C.4.7. *Let the field X be given by* (C.2.10), *where the Lévy basis M satisfies Assumption C.2.2 and the kernel function f satisfies Assumption C.2.4 and $f(x,y) < f(0,0)$ for all $(x,y) \ne (0,0)$. Then*

$$\lim_{x\to\infty}\frac{\mathbb{P}(\sup_{t\in[0,T]}\sup_{v\in B} X_{v,t} > x)}{\rho((x,\infty))} = m(B\times[0,T])\mathbb{E}\left[\exp\left(\beta X_{u,s}\right)\right],$$

where $(u,s) \in B\times[0,T]$ is chosen arbitrarily.

C.5 Continuity properties

The main purpose of this section is to show that the field defined in (C.2.10) has a version with t-càdlàg sample paths. This result will be obtained in Theorem C.5.7 below. However, the proof involves showing two other results on continuity properties of related random fields of independent value. Therefore these results are formulated as separate theorems; see Theorems C.5.2 and C.5.4 below. Only the main results, Theorems C.5.2, C.5.4 and C.5.7, are stated fully with all assumptions included in the statement. The rest are to be understood in relation to the context. As stated in Section C.2, Assumption C.2.2 on the Lévy basis is partly used to guarantee that the field is t-càdlàg. However, if the aim is solely to obtain the t-càdlàg property, we can relax the assumption. In this section we therefore consider Assumption C.5.1 below. It will both be referred to with the dimension of the Lévy basis being d and $d+1$. Thus, both the assumption and the subsequent Theorem C.5.2 will be formulated with $m \in \mathbb{N}$ indicating the dimension.

Assumption C.5.1. *The Lévy basis M on $\mathbb{R}^m$ is stationary and isotropic satisfying* (C.2.1). *Moreover, the Lévy measure, denoted ρ, satisfies*

$$\int_{|y|>1}|y|^k\rho(\mathrm{d}y) < \infty \qquad \forall k \in \mathbb{N}. \tag{C.5.1}$$

For the first result in this section, consider a compact set $K \subseteq \mathbb{R}^m$ and define the random field $Y = (Y_v)_{v\in K}$ by

$$Y_v = \int_{\mathbb{R}^m} h(|v-u|)M(\mathrm{d}u), \tag{C.5.2}$$

where M is a Lévy basis on $\mathbb{R}^m$ satisfying Assumption C.5.1. It is shown in [19, Theorem A.1] that such a field has a continuous version when $h : [0,\infty) \to \mathbb{R}$ satisfies certain properties including being differentiable. Under much less restrictive assumptions on the kernel function h, we show that this is still the case. We only assume that h is bounded and integrable

$$\int_{\mathbb{R}^m} h(|u|)\mathrm{d}u < \infty, \tag{C.5.3}$$

and that h is Lipschitz continuous. That is, there exist $C_L > 0$ such that

$$|h(x) - h(y)| \le C_L |x-y| \tag{C.5.4}$$

for all $x, y \ge 0$. Having Assumption C.5.1 satisfied for the basis M and (C.5.3) and (C.5.4) satisfied for the bounded kernel function ensures in particular that the integral (C.5.2) exists; see [18, Theorem 2.7].

To show continuity, we appeal to a result in [1], in which finite moments and cumulants of the spot variable M' of the basis M are needed. As already mentioned, (C.5.1) is equivalent to saying that M' has finite moments and thus cumulants of any order; see [17, Corollary 3.2.2] for the relation between moments and cumulants.

Theorem C.5.2. *If the field Y is given by* (C.5.2) *with the Lévy basis M on $\mathbb{R}^m$ satisfying Assumption C.5.1, and if the kernel is bounded and satisfies* (C.5.3) *and* (C.5.4), *then the field has a continuous version.*

Proof. For $r \in \mathbb{R}$ and $n \in \mathbb{N}$ we shall consider moments of the form $\mathbb{E}[(Y_{v+r} - Y_v)^n]$. Note that only indices in K are relevant, so in particular, $0 \le |r| \le \mathrm{diam}(K)$. By (C.5.4) and the triangle inequality, there is a finite C such that

$$\Big|h(|v+r-u|) - h(|v-u|)\Big| \le C|r|.$$

Now let $\kappa_n[\,\cdot\,]$ denote the nth cumulant of a random variable; for a brief overview of the relation between cumulants and moments we refer to [17, Chapter 3] and in particular [17, Corollary 3.2.2]. The cumulants κ_n of the difference $Y_{v+r} - Y_v$ satisfy $\kappa_1[Y_{v+r} - Y_v] = 0$ and, for $n > 1$,

$$\begin{aligned}|\kappa_n[Y_{v+r} - Y_v]| &\le \Big|\kappa_n[M']\Big| \int_{\mathbb{R}^m} \Big|h(|v+r-u|) - h(|v-u|)\Big|^n \mathrm{d}u \\ &\le \Big|\kappa_n[M']\Big| C^{n-1}|r|^{n-1} \int_{\mathbb{R}^m} \Big|h(|v+r-u|) - h(|v-u|)\Big| \mathrm{d}u \le C_n |r|^{n-1},\end{aligned}$$

where $C_n \ge 0$ is a finite constant, chosen independently of r and $v \in K$ by

$$C_n = \Big|\kappa_n[M']\Big| C^{n-1} 2 \int_{\mathbb{R}^m} h(|u|)\mathrm{d}u < \infty,$$

see e.g. [21, Appendix A] for the cumulant formulas. Consequently, for all $n \in \mathbb{N}$, there exist finite constants C_n' and natural numbers $n' \geq n/2$ such that

$$\mathbb{E}[(Y_{v+r} - Y_v)^n] \leq C_n' |r|^{n'}$$

with the equality $n' = n/2$ whenever n is even; see [17, Corollary 3.2.2]. Using the fact that $|r| \leq \operatorname{diam}(K)$, we find finite $C' \geq 0$ and $\eta > 4(m+1)$ such that

$$\mathbb{E}|Y_{v+r} - Y_v|^{4(m+1)} \leq C'_{4(m+1)} |r|^{2m} |r|^2 \leq \frac{C'|r|^{2m}}{|\log|r||^{1+\eta}}$$

for all $v \in K$. From a corollary to [1, Theorem 3.2.5] we conclude that $(Y_v)_{v \in K}$ has a continuous version on K. □

Next, we consider a field indexed by $\mathbb{R}^d \times \mathbb{R}$ allowing for discontinuities in time, and we show that it has a t-càdlàg version. For compact sets $K \subseteq \mathbb{R}^d$ and $[0,S]$, $S > 0$, we let the random field $Z = (Z_{v,t})_{(v,t) \in K \times [0,S]}$ be given by

$$Z_{v,t} = \int_{\mathbb{R}^d} \int_{[0,t]} g(|v-u|) M(\mathrm{d}s, \mathrm{d}u), \tag{C.5.5}$$

where M is a Lévy basis satisfying Assumption C.5.1 with $m = d+1$, and the integration kernel $g : [0,\infty) \to \mathbb{R}$ is assumed to be bounded, integrable and Lipschitz continuous, i.e. it satisfies (C.5.3) and (C.5.4) with $m = d$.

Choose $0 = t_0 < \cdots < t_n$ in $[0,S]$ and $v \in K$. Arguing as in Section C.2, the cumulant function for $(Z_{v,t_1}, Z_{v,t_2} - Z_{v,t_1}, \ldots, Z_{v,t_n} - Z_{v,t_{n-1}})$ can be found to be

$$\begin{aligned}
&C(\boldsymbol{\lambda} \dagger (Z_{v,t_1}, Z_{v,t_2} - Z_{v,t_1}, \ldots, Z_{v,t_n} - Z_{v,t_{n-1}})) \\
&\quad = \sum_{j=1}^{n} (t_j - t_{j-1}) \Big(i\lambda_j a \int_{\mathbb{R}^d} g(|v-u|) \mathrm{d}u - \frac{1}{2}\theta \lambda_j^2 \int_{\mathbb{R}^d} g(|v-u|)^2 \mathrm{d}u \\
&\qquad\qquad + \int_{\mathbb{R}^d} \int_{\mathbb{R}} \mathrm{e}^{ig(|v-u|)\lambda_j z} - 1 - ig(|v-u|)\lambda_j z 1_{[-1,1]}(z) \rho(\mathrm{d}z) \mathrm{d}u \Big)
\end{aligned}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$. By a change of measure, we see for fixed $v \in \mathbb{K}$ that $(Z_{v,t})_{t \in [0,S]}$ is a one-dimensional Lévy process in law. In the following we shall extend this to a result concerning the process of random fields indexed by time.

In this section we will often consider the field as being a collection of real-valued functions defined on space K or $\tilde{K} = K \cap \mathbb{Q}^d$, with the functions indexed by time in $[0,S]$ or $\tilde{S} = [0,S] \cap \mathbb{Q}$. As such we introduce the notation $\boldsymbol{Z}_t = (Z_{v,t})_{v \in K}$, with the entire field denoted by $\boldsymbol{Z} = (\boldsymbol{Z}_t)_{t \in [0,S]}$ when considered as a collection of random functions. We use the same notation when space and time are indexed by $\tilde{K}$ and $\tilde{S}$, respectively, although when it is unclear which is meant and it is necessary to distinguish the cases, we explicitly state it.

Let $t \in [0,S]$ be fixed and choose $v_1,\dots,v_n \in K$. Then $(Z_{v_1,t},\dots,Z_{v_n,t})$ has cumulant function given by

$$\begin{aligned}
C(\boldsymbol{\lambda} \dagger (Z_{v_1,t},\dots,Z_{v_n,t})) &= tia \int_{\mathbb{R}^d} \sum_{j=1}^n \lambda_j g(|v_j - u|)\mathrm{d}u \\
&- t\frac{1}{2}\theta \int_{\mathbb{R}^d} \Big(\sum_{j=1}^n \lambda_j g(|v_j - u|)\Big)^2 \mathrm{d}u \qquad \text{(C.5.6)} \\
&+ t \int_{\mathbb{R}^d} \int_{\mathbb{R}} \mathrm{e}^{i\sum_{j=1}^n \lambda_j g(|v_j-u|)z} - 1 - i \sum_{j=1}^n \lambda_j g(|v_j - u|) z 1_{[-1,1]}(z)\rho(\mathrm{d}z)\mathrm{d}u.
\end{aligned}$$

Replacing (a,θ,ρ) by $(ta,t\theta,t\rho)$, we see from (C.5.6) that $\boldsymbol{Z}_t$ is the type of field defined in (C.5.2). Thus, by Theorem C.5.2, $(Z_{v,t})_{v\in\tilde{K}}$ is almost surely uniformly continuous. This holds jointly for all rational time points $t \in \tilde{S}$, and therefore a version of $(\boldsymbol{Z}_t)_{t\in\tilde{S}}$ can be chosen with $\boldsymbol{Z}_t$ being continuous for all $t \in \tilde{S}$, i.e. it has values in the space of real-valued functions on the compact set K. It will be useful in the following that this space equipped with the uniform norm, here denoted $(\mathcal{C}(K,\mathbb{R}),\|\cdot\|_\infty)$, is a separable Banach space; see [14, Theorem 4.19]. The following lemma concerns this specific version of $(\boldsymbol{Z})_{t\in\tilde{S}}$ taking its values in $(\mathcal{C}(K,\mathbb{R}),\|\cdot\|_\infty)$.

Lemma C.5.3. *The process $(\boldsymbol{Z}_t)_{t\in\tilde{S}}$ is a Lévy process in law, i.e. each $\boldsymbol{Z}_t$ has an infinitely divisible distribution, the process has stationary and independent increments, and it is stochastically continuous with respect to the uniform norm.*

Proof. As $\boldsymbol{Z}_t$ is a version of the field studied in (C.5.6) for each $t \in \tilde{S}$, also the cumulant function for $(Z_{v_1,t},\dots,Z_{v_n,t})$ will be as in (C.5.6). With similar considerations, but heavier notation, it can be realised that for $v_1,\dots,v_n \in K$ and $0 = t_0 < t_1 < \dots < t_m \in \tilde{S}$, and defining $Z^n_{t_j} = (Z_{v_1,t_j},\dots,Z_{v_n,t_j})$ for $j = 1,\dots,m$, it holds that

$$\begin{aligned}
&C(\boldsymbol{\lambda} \dagger (Z^n_{t_1}, Z^n_{t_2} - Z^n_{t_1},\dots,Z^n_{t_m} - Z^n_{t_{m-1}})) \\
&\quad = \sum_{j=1}^m C(\lambda_j \dagger Z^n_{t_j} - Z^n_{t_{j-1}}) = \sum_{j=1}^m C(\lambda_j \dagger Z^n_{t_j - t_{j-1}}),
\end{aligned}$$

where $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_n)$ and each $\lambda_j \in \mathbb{R}^n$, and the natural convention $Z^n_0 = (0,\dots,0)$ is applied. This shows that $(\boldsymbol{Z}_t)_{t\in\tilde{S}}$ has stationary and independent increments.

To show stochastic continuity it suffices to show that

$$\lim_{n\to\infty} \mathbb{P}(\|\boldsymbol{Z}_{t_n}\|_\infty \geq \epsilon) = 0$$

for any rational sequence (t_n) satisfying $t_n \downarrow 0$. As $(\mathcal{C}(K,\mathbb{R}),\|\cdot\|_\infty)$ is a separable Banach space, this is equivalent to showing that $\boldsymbol{Z}_{t_n}$ converges to $\delta_{\boldsymbol{0}}$ in law in the uniform norm, where $\delta_{\boldsymbol{0}}$ is the degenerate probability measure concentrated at $\boldsymbol{0}$. For $t \in \tilde{S}$, let ν_t denote the distribution of $\boldsymbol{Z}_t$ and let $\hat{\nu}_{t_n}$ be its characteristic function defined on the dual space of $\mathcal{C}(K,\mathbb{R})$, see [15, Section 1.7]. Since $(\mathcal{C}(K,\mathbb{R}),\|\cdot\|_\infty)$ is separable, ν_t is a Radon measure [15, Proposition 1.1.3] and the results in [15, Chapters 2 & 5] apply. Due to the infinite divisibility, $\nu_1 = \nu_{t_n} * \nu_{1-t_n}$ for any $n \in \mathbb{N}$ (assuming $t_n \leq 1$), and we conclude that $\{\nu_{t_n}\}$ is relatively shift compact [15, Theorem 2.3.1]. Following the proofs of [15, Propositions 5.1.4 & 5.1.5] we obtain that $\lim_{n\to\infty}\hat{\nu}_{t_n} \to 1$ uniformly on bounded sets of the dual space. Combining [15, Propositions 2.3.9 & 1.8.2] shows that $\boldsymbol{Z}_{t_n}$ converges in law to $\delta_{\mathbf{0}}$ as claimed. □

The next theorem states that the field Z defined in (C.5.5) indeed has a t-càdlàg version.

Theorem C.5.4. *Let the field Z be given by* (C.5.5) *such that the Lévy basis M on $\mathbb{R}^{d+1}$ satisfies Assumption C.5.1 with $m = d+1$, and the bounded kernel g satisfies* (C.5.3) *and* (C.5.4)*, with $m = d$. There is a field $Z' = (Z'_{v,t})_{(v,t)\in K\times[0,S]}$ that is a version of Z, i.e. $\mathbb{P}(Z'_{v,t} = Z_{v,t}) = 1$ for all $(v,t)\in K\times[0,S]$, and such that $\lim_{s\downarrow t}\boldsymbol{Z}'_s(\omega) = \boldsymbol{Z}'_t(\omega)$ and $\lim_{s\uparrow t}\boldsymbol{Z}'_s(\omega)$ exists with respect to $\|\cdot\|_\infty$ for all ω. Furthermore, the map $v\mapsto Z'_{v,t}$ from K into $\mathbb{R}$ is continuous for all $t\in[0,S]$. In particular, Z' has t-càdlàg sample paths.*

The desired t-càdlàg version will be an extension of the field $(\boldsymbol{Z}_t)_{t\in\tilde{S}}$ studied in Lemma C.5.3. Thus, $(\boldsymbol{Z}_t)_{t\in\tilde{S}}$ will still be a version chosen such that each $\boldsymbol{Z}_t$ is a continuous random field. The result relies on a sequence of lemmas that are shown using an adaption of the ideas of [24, Theorems 11.1 & 11.5] and [24, Lemmas 11.2-11.4] for Lévy processes on $\mathbb{R}$. Lemmas C.5.5 and C.5.6 are shown using similar techniques for the Lévy process $(\boldsymbol{Z}_t)_{t\in\tilde{S}}$, and therefore we omit the proofs here, and refer to the supplementary material for completeness.

For the statement and proof of these lemmas, the following notation will be useful. We say that $\boldsymbol{Z}(\omega)$ has ϵ-oscillation n times in a set $M\subseteq\mathbb{Q}\cap[0,\infty)$ if there exist $t_0<t_1<\cdots<t_n\in M$ such that

$$\|\boldsymbol{Z}_{t_j}(\omega)-\boldsymbol{Z}_{t_{j-1}}(\omega)\|_\infty = \sup_{v\in K}|Z_{v,t_j}(\omega)-Z_{v,t_{j-1}}(\omega)| > \epsilon$$

for all $j=1,\dots,n$. We say that $\boldsymbol{Z}(\omega)$ has ϵ-oscillation infinitely often in M if it has ϵ-oscillation n times in M for any $n\in\mathbb{N}$. Consider Ω_1 given by

$$\begin{aligned}\Omega_1 = \{\omega\in\Omega \mid & \lim_{s\in\mathbb{Q},s\downarrow t}\boldsymbol{Z}_s(\omega) \text{ exists with respect to } \|\cdot\|_\infty \text{ for all } t\in[0,S] \text{ and} \\ & \lim_{s\in\mathbb{Q},s\uparrow t}\boldsymbol{Z}_s(\omega) \text{ exists with respect to } \|\cdot\|_\infty \text{ for all } t\in[0,S]\}.\end{aligned}$$

Furthermore define the sets

$$A_k = \{\omega\in\Omega \mid \boldsymbol{Z}(\omega) \text{ does not have } \tfrac{1}{k}\text{-oscillation infinitely often in } \tilde{S}\},$$

and from these define $\Omega'_1 = \cap_{k\in\mathbb{N}}A_k$. Each A_k is measurable as each $\boldsymbol{Z}_t$ is continuous on K for $t\in\tilde{S}$, such that $\|\cdot\|_\infty = \sup_{v\in K}|\cdot| = \sup_{v\in\tilde{K}}|\cdot|$.

Lemma C.5.5. $\Omega'_1\subseteq\Omega_1$.

Lemma C.5.6. $\mathbb{P}(\Omega'_1) = 1$.

Having established that $\lim_{s\in\mathbb{Q},s\downarrow t}\boldsymbol{Z}_s$ and $\lim_{s\in\mathbb{Q},s\uparrow t}\boldsymbol{Z}_s$ exist almost surely, we now prove the main result Theorem C.5.4 on the existence of a t-càdlàg version of Z.

Proof of Theorem C.5.4. We have $\mathbb{P}(\Omega_1') = 1$ by Lemma C.5.6. For all $t \in [0,S]$, define $\boldsymbol{Z}_t'(\omega) = \mathbf{1}_{\Omega_1'}(\omega)(\lim_{s\in\mathbb{Q},s\downarrow t} \boldsymbol{Z}_s(\omega))$, where the limit is with respect to $\|\cdot\|_\infty$, and exists according to Lemma C.5.5. The càdlàg-assertion is trivially true for $\omega \notin \Omega_1'$. Now consider $\omega \in \Omega_1'$ but suppress ω in ease of notation. By definition of $\boldsymbol{Z}_t'$,

$$\forall \epsilon > 0\ \exists N\ \forall s \in (t, t+\tfrac{1}{N}) \cap \mathbb{Q}:\ \|\boldsymbol{Z}_t' - \boldsymbol{Z}_s\|_\infty < \epsilon. \tag{C.5.7}$$

Let (t_n) be any sequence satisfying $t_n \downarrow t$. Fix $\epsilon > 0$, and let $N \in \mathbb{N}$ satisfy (C.5.7) with the bound $\frac{\epsilon}{2}$. There is $n_0 \in \mathbb{N}$ such that $|t_n - t| < \frac{1}{N}$ for all $n \geq n_0$. Now fix such n. By another application of (C.5.7) there exist N_n such that $t_n + \frac{1}{N_n} \leq t + \frac{1}{N}$ and $\|\boldsymbol{Z}_{t_n}' - \boldsymbol{Z}_s\|_\infty < \frac{\epsilon}{2}$ for all $s \in (t_n, t_n + \frac{1}{N_n}) \cap \mathbb{Q}$. For any of those s we in particular find that

$$\|\boldsymbol{Z}_{t_n}' - \boldsymbol{Z}_t'\|_\infty \leq \|\boldsymbol{Z}_{t_n}' - \boldsymbol{Z}_s\|_\infty + \|\boldsymbol{Z}_t' - \boldsymbol{Z}_s\|_\infty < \epsilon.$$

As this is true for all $n \geq n_0$ we conclude that $\boldsymbol{Z}' = (\boldsymbol{Z}_t')_{t\in[0,S]}$ is right-continuous with respect to $\|\cdot\|_\infty$. Similar arguments show that $\boldsymbol{Z}'$ has limits from the left and that the limits are unique. The mapping $v \mapsto Z_{v,t}'$ is continuous because the space $(\mathcal{C}(K,\mathbb{R}), \|\cdot\|_\infty)$ is complete, and $\boldsymbol{Z}_t'$ is defined as the limit of such functions.

We now argue that $\boldsymbol{Z}'$ is indeed a version $\boldsymbol{Z}$. If $(t_n) \subset \tilde{S}$ with $t_n \downarrow t$ then $Z_{v,t_n} \xrightarrow{\mathbb{P}} Z_{v,t}$ for all $v \in K$ as $(Z_{v,t})_{t\in[0,S]}$ is a Lévy process in law and thus especially stochastically continuous. Since $\mathbb{P}(\Omega_1') = 1$ we have $Z_{v,t_n} \to Z_{v,t}'$ almost surely, and by uniqueness of limits we conclude that $\mathbb{P}(Z_{v,t}' = Z_{v,t}) = 1$ for all $(v,t) \in K \times [0,S]$.

It remains to show that Z' is t-càdlàg. Since we for given $(v,t) \in K \times [0,S]$ can write

$$|Z_{v,t}' - Z_{u,s}'| \leq |Z_{v,t}' - Z_{u,t}'| + \|\boldsymbol{Z}_t' - \boldsymbol{Z}_s'\|_\infty$$

for any choice of $(u,s) \in K \times [0,S]$, we conclude that $\lim_{(u,s)\to(v,t^+)} Z_{u,s}' = Z_{v,t}'$, from the continuity of $\boldsymbol{Z}_t'$ and the uniform càdlàg property of $(\boldsymbol{Z}_t')_{t\in[0,S]}$. Similar arguments give that the limit $\lim_{(u,s)\to(v,t^-)} Z_{u,s}'$ exists in $\mathbb{R}$ and that it is unique. □

Theorem C.5.7 below is stated under Assumption C.2.2. However, in order to establish t-càdlàg sample paths, the milder Assumption C.5.1 would have been sufficient.

Theorem C.5.7. *Let the field $X = (X_{v,t})_{(v,t)\in B'\times T'}$ be given by* (C.2.10)*, where the Lévy basis M satisfies Assumption C.2.2 and the kernel function f satisfies Assumption C.2.3. Then X has a version with t-càdlàg sample paths.*

Proof. We decompose the field $(X_{v,t})$ as

$$\begin{aligned} X_{v,t} &= \int_{\mathbb{R}^d\times[0,t]} f(|v-u|, t-s) - f(|v-u|, 0) M(\mathrm{d}s, \mathrm{d}u) \\ &\quad + \int_{\mathbb{R}^d\times(-\infty,0)} f(|v-u|, t-s) M(\mathrm{d}s, \mathrm{d}u) + \int_{\mathbb{R}^d\times[0,t]} f(|v-u|, 0) M(\mathrm{d}s, \mathrm{d}u). \end{aligned}$$

By Theorem C.5.4, choosing $g(\cdot) = f(\cdot, 0)$, the third term has a t-càdlàg version. Due to continuity of the integrands, the first and second terms have continuous versions by arguments similar to those in the proof of Theorem C.5.2: Defining the continuous function $\phi : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ by

$$\phi(x,y) = \mathbf{1}_{[0,\infty)}(y)\big(f(x,y) - f(x,0)\big),$$

the first term above reads

$$Y'_{v,t} = \int_{\mathbb{R}^d \times [0,\infty)} \phi(|v-u|, t-s) M(\mathrm{d}u, \mathrm{d}s) = Y_{v,t} + y_{v,t},$$

where $y_{v,t} = \mathbb{E}Y'_{v,t}$ and $Y_{v,t} = Y'_{v,t} - y_{v,t}$. The field $(Y_{v,t})$ is continuous by previous arguments replacing the assumptions (C.5.3) and (C.5.4) by the conditions

$$\int_{\mathbb{R}^d} \int_0^\infty \Big|\phi(|u|, T + \ell - s)\Big| \mathrm{d}s \mathrm{d}u < \infty \tag{C.5.8}$$

and the Lipschitz continuity of ϕ

$$\Big|\phi(|u_1|, t_1 - s) - \phi(|u_2|, t_2 - s)\Big| \leq C\Big|(u_1 - u_2, t_1 - t_2)\Big|$$

for all $u_1, u_2 \in \mathbb{R}^d$ and $t_1, t_2 \in T'$. These conditions are easily seen to be satisfied under Assumption C.2.3. As

$$y_{v,t} = y_{0,t} = \mathbb{E}[M'] \int_{\mathbb{R}^d} \int_0^\infty \phi(|u|, t-s) \mathrm{d}s \mathrm{d}u < \infty$$

the deterministic field $(y_{v,t})$ is continuous by a dominated convergence argument using (C.5.8). The continuity of the second term follows similarly. □

Acknowledgments

This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

Supplementary material

C.I Proofs of Section C.3

Proof of Lemma C.3.9. For sufficiently large x we find that

$$\begin{aligned}\mathbb{P}(Z\phi(U,S)>x) &= \frac{1}{\nu(A)}F\Big(\Big\{(u,s,z)\in\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}_+ : z\phi(u,s)>x\Big\}\Big)\\ &= \frac{1}{\nu(A)}\int_{B'\times T'} L\Big(\frac{x}{c}\Big)\exp\Big(-\beta\frac{x}{c}\Big)m(\mathrm{d}u,\mathrm{d}s)\\ &\quad+\frac{1}{\nu(A)}\int_{(B'\times T')^c} L\Big(\frac{x}{\phi(u,s)}\Big)\exp\Big(-\beta\frac{x}{\phi(u,s)}\Big)m(\mathrm{d}u,\mathrm{d}s),\end{aligned}$$

where the first term equals $L(x/c)\exp(-\beta x/c)$ times the desired limit. The result follows when the latter integral is shown to be of order $o(L(x/c)\exp(-\beta x/c))$, as $x\to\infty$. Let $h(u,s;x)$ denote the integrand. For all $(u,s)\notin B'\times T'$ we have $\phi(u,s)<c$. Combined with (C.2.4), this implies the existence of $\gamma>0$ and $C>0$ such that

$$\frac{h(u,s;x)}{L(x/c)\exp(-\beta x/c)}\le C\exp(-\gamma x)$$

for sufficiently large x. Thus, the integrand $h(u,s;x)$ is $o(L(x/c)\exp(-\beta x/c))$ at infinity. By dominated convergence, the integral is of order $o(L(x/c)\exp(-\beta x/c))$ if we can find an integrable function $g:\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$ such that

$$\frac{h(u,s;x)}{L(x/c)\exp(-\beta x/c)}\le g(u,s)$$

for all $(u,s)\in\mathbb{R}^d\times\mathbb{R}$. Returning to (C.2.5) we see that for all $0<\gamma<\beta/c$ there is $C>0$ and x_0 such that

$$\frac{h(u,s;x)}{L(x/c)\exp(-\beta x/c)}\le C\exp\Big(-x_0(\beta-\gamma c)\Big(\frac{1}{\phi(u,s)}-\frac{1}{c}\Big)\Big) \tag{C.I.1}$$

for all $x\ge x_0$. Independent of (u,s) we can find a finite constant $\tilde{C}$ such that the right hand side of (C.I.1) is bounded by $\tilde{C}\phi(u,s)$, which is integrable by assumption. This shows the desired order of convergence.

From [16, Lemma 2.4(i)] the distribution of $Z\phi(U,S)$ is convolution equivalent with index β/c. The integrability result follows from [16, Corollary 2.1(ii)]. □

Corollary C.I.1. *If $V^1,V^2,\dots$ are i.i.d. fields with distribution ν_1, then*

$$\mathbb{E}\Big[\exp\Big(\beta\sup_{u\in B}\sup_{s\in[0,T]}\lambda_{u,s}\big((V^1_{v,t}+\cdots+V^n_{v,t})_{(v,t)}\big)\Big)\Big]<\infty$$

for all $n\in\mathbb{N}$.

Proof. Because each V^i can be represented by $(Z^i f(|v-U^i|,t-S^i)_{(v,t)\in B'\times T'}$, the result follows from (C.3.8) and (C.3.10). □

Proof of Theorem C.3.10. We will show the claim by induction over n: We note that the case $n=1$ follows easily from Theorem C.3.7. Now assume that the result holds true for some $n \in \mathbb{N}$ and let for convenience $V^{*n} = V^1 + \cdots + V^n$. Also, let $y^* = \sup_{(v,t)\in B'\times T'} y_{v,t}$. Using (C.3.7) and the representation $V^i = Z^i f(|v-U^i|, t-S^i)$, we find

$$\begin{aligned}
&\mathbb{P}(\Psi(V_{v,t}^{*n} + V_{v,t}^{n+1} + y_{v,t}) > x) \\
&\quad \leq \mathbb{P}\Big(\sum_{i=1}^{n} Z^i \phi(U^i, S^i) > \frac{x-y^*}{2}, Z^{n+1}\phi(U^{n+1}, S^{n+1}) > \frac{x-y^*}{2}, \\
&\qquad\quad \Psi(V_{v,t}^{*n} + V_{v,t}^{n+1} + y_{v,t}) > x\Big) \\
&\qquad + \mathbb{P}\Big(\sum_{i=1}^{n} Z^i \phi(U^i, S^i) \leq \frac{x-y^*}{2}, \Psi(V_{v,t}^{*n} + V_{v,t}^{n+1} + y_{v,t}) > x\Big) \\
&\qquad + \mathbb{P}\Big(Z^{n+1}\phi(U^{n+1}, S^{n+1}) \leq \frac{x-y^*}{2}, \Psi(V_{v,t}^{*n} + V_{v,t}^{n+1} + y_{v,t}) > x\Big)
\end{aligned} \tag{C.I.2}$$

The first term in (C.I.2) is bounded from above by

$$\mathbb{P}\Big(\sum_{i=1}^{n} Z^i \phi(U^i, S^i) > \frac{x-y^*}{2}\Big)\mathbb{P}\Big(Z^{n+1}\phi(U^{n+1}, S^{n+1}) > \frac{x-y^*}{2}\Big).$$

In Lemma C.3.9 we showed that the distribution of $Z^i\phi(U^i, S^i)$ is convolution equivalent with index β/c, and hence, from [9, Corollary 2.11] and (C.3.9), both factors are asymptotically equivalent to $\rho_1((x/(2c), \infty))$ as $x \to \infty$. Following the proof of [8, Lemma 2] we see that the product is $o((\rho_1 * \rho_1)((x/c, \infty)))$, and as such the first term in (C.I.2) is $o(\rho_1((x/c, \infty)))$ due to the convolution equivalence of ρ_1. By Theorem C.3.7 it is of order $o(\mathbb{P}(\Psi(V_{v,t}^1) > x))$ as $x \to \infty$.

By independence, the two remaining terms in (C.I.2) divided by $\mathbb{P}(\Psi(V_{v,t}^1) > x)$ are

$$\begin{aligned}
&\int_{C_x} \frac{\mathbb{P}(\Psi(\sum_{i=1}^{n} z^i f(|v-u^i|, t-s^i) + V_{v,t}^{n+1} + y_{v,t}) > x)}{\mathbb{P}(\Psi(V_{v,t}^1) > x)} F_1^{\otimes n}(\mathrm{d}(u^1, s^1, z^1; \ldots; u^n, s^n, z^n)) \\
&\qquad + \int_{\tilde{C}_x} \frac{\mathbb{P}(\Psi(V_{v,t}^{*n} + z^1 f(|v-u^1|, t-s^1) + y_{v,t}) > x)}{\mathbb{P}(\Psi(V_{v,t}^1) > x)} F_1(\mathrm{d}(u^1, s^1, z^1)),
\end{aligned} \tag{C.I.3}$$

where $F_1^{\otimes n}$ is the n-fold product measure of F_1 and

$$\begin{aligned}
C_x &= \Big\{(u^1, s^1, z^1; \ldots; u^n, s^n, z^n) : \sum_{i=1}^{n} z^i \phi(u^i, s^i) \leq \frac{x-y^*}{2}\Big\}, \\
\tilde{C}_x &= \Big\{(u^1, s^1, z^1) : z^1 \phi(u^1, s^1) \leq \frac{x-y^*}{2}\Big\}.
\end{aligned}$$

Above we used the representation $V^i = Z^i f(|v-U^i|, t-S^i)$ again. By Theorem C.3.7 and the induction assumption, the integrands of (C.I.3) have the following limits as

$x \to \infty$,

$$f_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n) = \frac{\int_B \int_0^T \exp\left(\beta\lambda_{u,s}\left(\sum_{i=1}^n z^i f(|v-u^i|,t-s^i) + y_{v,t}\right)\right)\mathrm{d}s\mathrm{d}u}{m(B\times[0,T])},$$

$$f_2(u^1,s^1,z^1) = \frac{n\int_B \int_0^T \mathbb{E}\left[\exp\left(\beta\lambda_{u,s}\left(V_{v,t}^1+\cdots+V_{v,t}^{n-1} + z^1 f(|v-u^1|,t-s^1) + y_{v,t}\right)\right)\right]\mathrm{d}s\mathrm{d}u}{m(B\times[0,T])},$$

respectively. When integrated with respect to the relevant measures we find

$$\begin{aligned}&\int f_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n)F_1^{\otimes n}(\mathrm{d}(u^1,s^1,z^1;\dots;u^n,s^n,z^n))\\ &\quad+\int f_2(u^1,s^1,z^1)F_1(\mathrm{d}(u^1,s^1,z^1))\\ =&\frac{n+1}{m(B\times[0,T])}\int_B\int_0^T \mathbb{E}\left[\exp\left(\beta\lambda_{u,s}\left(V_{v,t}^1+\cdots+V_{v,t}^n+y_{v,t}\right)\right)\right]\mathrm{d}s\mathrm{d}u,\end{aligned}$$

which is the desired expression. To show convergence of the integrals in (C.I.3), using Fatou's lemma, it suffices to find integrable functions $g_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n;x)$ and $g_2(u^1,s^1,z^1;x)$ that are upper bounds of the integrands such that their limits exist when $x\to\infty$ and such that

$$\begin{aligned}&\int g_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n;x)F_1^{\otimes n}(\mathrm{d}(u^1,s^1,z^1;\dots;u^n,s^n,z^n))\\ &\quad+\int g_2(u^1,s^1,z^1;x)F_1(\mathrm{d}(u^1,s^1,z^1))\\ \to&\int \lim_{x\to\infty} g_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n;x)F_1^{\otimes n}(\mathrm{d}(u^1,s^1,z^1;\dots;u^n,s^n,z^n))\\ &\quad+\int \lim_{x\to\infty} g_2(u^1,s^1,z^1;x)F_1(\mathrm{d}(u^1,s^1,z^1))\end{aligned}$$

as $x\to\infty$. Using (C.3.7) and properties of Ψ, we can choose the functions

$$g_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n;x) = \mathbf{1}_{C_x}\frac{\mathbb{P}(Z^1\phi(U^1,Z^1) > x - y^* - \sum_{i=1}^n z^i\phi(u^i,s^i))}{\mathbb{P}(\Psi(V_{v,t}^1) > x)}$$

and

$$g_2(u^1,s^1,z^1;x) = \mathbf{1}_{\tilde{C}_x}\frac{\mathbb{P}(\sum_{i=1}^n Z^i\phi(U^i,Z^i) > x - y^* - z^1\phi(u^1,s^1))}{\mathbb{P}(\Psi(V_{v,t}^1) > x)}.$$

From Theorem C.3.7 and (C.3.9) we find that

$$\mathbb{P}(Z^1\phi(U^1,S^1) > x) \sim \frac{m(B'\times T')}{m(B\times[0,T])}\mathbb{P}(\Psi(V_{v,t}^1) > x) \tag{C.I.4}$$

as $x\to\infty$. The fact that the distribution of $Z^1\phi(U^1,S^1)$ is convolution equivalent and in particular has an exponential tail implies

$$g_1(u^1,s^1,z^1;\dots;u^n,s^n,z^n;x) \to \frac{m(B'\times T')}{m(B\times[0,T])}\exp\left(\frac{\beta}{c}\left(y^* + \sum_{i=1}^n z^i\phi(u^i,s^i)\right)\right)$$

as $x \to \infty$. Similarly, (C.I.4) and an application of [9, Corollary 2.11] gives

$$\begin{aligned} &g_2(u^1, s^1, z^1; x) \\ &\quad \to \frac{m(B' \times T')}{m(B \times [0,T])} n \exp\Big(\frac{\beta}{c}(y^* + z^1 \phi(u^1, s^1))\Big)\Big(\mathbb{E}\exp\Big(\frac{\beta}{c} Z^1 \phi(U^1, S^1)\Big)\Big)^{n-1} \end{aligned}$$

as $x \to \infty$, and we conclude that

$$\begin{aligned} &\int \lim_{x\to\infty} g_1(u^1, s^1, z^1; \dots; u^n, s^n, z^n; x) F_1^{\otimes n}(\mathrm{d}(u^1, s^1, z^1; \dots; u^n, s^n, z^n)) \\ &\qquad + \int \lim_{x\to\infty} g_2(u^1, s^1, z^1; x) F_1(\mathrm{d}(u^1, s^1, z^1)) \\ &= \frac{m(B' \times T')}{m(B \times [0,T])}(n+1)\exp(\beta y^*/c)\Big(\mathbb{E}\exp\Big(\frac{\beta}{c} Z^1 \phi(U^1, S^1)\Big)\Big)^n. \end{aligned} \tag{C.I.5}$$

For notational convenience, we let μ denote the distribution of $Z^i\phi(U^i, S^i)$. Then, again by [9, Corollary 2.11] and (C.I.4), (C.I.5) equals

$$\lim_{x\to\infty} \frac{m(B' \times T')}{m(B \times [0,T])} \frac{\mu^{*(n+1)}((x - y^*, \infty))}{\mu((x, \infty))} = \lim_{x\to\infty} \frac{\mu^{*(n+1)}((x - y^*, \infty))}{\mathbb{P}(\Psi(V_{v,t}^1) > x)}. \tag{C.I.6}$$

Furthermore, we see

$$\begin{aligned} &\mathbb{P}(\Psi(V_{v,t}^1) > x)\Big(\int g_1(z^1; \dots; z^n; x)\mu^{\otimes n}(\mathrm{d}(z^1; \dots; z^n)) + \int g_2(z; x)\mu(\mathrm{d}z)\Big) \\ &\quad = \int_0^{(x-y^*)/2} \mu((x - y^* - z, \infty))\mu^{*n}(\mathrm{d}z) + \int_0^{(x-y^*)/2} \mu^{*n}((x - y^* - z, \infty))\mu(\mathrm{d}z). \end{aligned}$$

Since, in particular, the tails of μ and μ^{*n} are exponential with index β/c, we see from [8, Lemma 2] that the sum of integrals is asymptotically equivalent to $\mu^{*(n+1)}((x - y^*, \infty))$. Returning to (C.I.6) concludes the proof. □

Before proving the theorem on the extremal behavior of X^1, we need the following lemma for a dominated convergence argument.

Lemma C.I.2. *Let $V^1, V^2, \dots$ be i.i.d. fields with distribution ν_1, and let (U, S, Z) be distributed according to F_1. There exist a constant K such that*

$$\mathbb{P}(\Psi(V_{v,t}^1 + \dots + V_{v,t}^n) > x) \le K^n \mathbb{P}(Z\phi(U, S) > x)$$

for all $n \in \mathbb{N}$ and all $x \ge 0$.

Proof. By Lemma C.3.9 the distribution of $Z\phi(U, S)$ is convolution equivalent, and it follows from [9, Lemma 2.8] that there is a constant K such that

$$\mathbb{P}\Big(\sum_{i=1}^n Z^i \phi(U^i, S^i) > x\Big) \le K^n \mathbb{P}(Z\phi(U, S) > x),$$

for i.i.d. variables $(U^1, S^1, Z^1), (U^2, S^2, Z^2), \dots$ with distribution F_1. The result follows directly from (C.3.7). □

Proof of Theorem C.3.11. From (C.3.8) and the representation $V^i = (Z^i f(|v-U^i|, t-S^i))_{(v,t)}$, we see that

$$\mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\big(X^1_{v,t}\big)\Big)\Big] \le \exp\Big(\nu(A)\Big(\mathbb{E}[\exp(\tfrac{\beta}{c}Z\phi(U,S)] - 1\Big)\Big).$$

The first claim now follows from (C.3.10).

For the limit result, we find by independence and Lemma C.I.2,

$$\begin{aligned}
&\mathbb{P}(\Psi(X^1_{v,t}+y_{v,t})>x)\\
&\quad = \mathrm{e}^{-\nu(A)}\sum_{n=1}^{\infty}\frac{\nu(A)^n}{n!}\mathbb{P}(\Psi(V^1_{v,t}+\cdots+V^n_{v,t}+y_{v,t})>x)\\
&\quad \le \mathbb{P}(Z\phi(U,S)>x-y^*)\mathrm{e}^{-\nu(A)}\sum_{n=1}^{\infty}\frac{\nu(A)^nK^n}{n!},
\end{aligned}$$

where $y^* = \sup_{(v,t)} y_{v,t}$ and $\mathrm{e}^{-\nu(A)}\sum_{n=1}^{\infty}\frac{\nu(A)^nK^n}{n!} < \infty$. With the convention that $V^1_{v,t}+\cdots+V^{n-1}_{v,t} = 0$ for $n=1$, by dominated convergence, Theorems C.3.7 and C.3.10 and Lemma C.3.9 yield

$$\begin{aligned}
&\lim_{x\to\infty}\frac{\mathbb{P}(\Psi(X^1_{v,t}+y_{v,t})>x)}{L(x/c)\exp(-\beta x/c)}\\
&\quad = \frac{n}{\nu(A)}\mathrm{e}^{-\nu(A)}\sum_{n=1}^{\infty}\frac{\nu(A)^n}{n!}\int_B\int_0^T\mathbb{E}\Big[\mathrm{e}^{\beta\lambda_{u,s}\left(V^1_{v,t}+\cdots+V^{n-1}_{v,t}+y_{v,t}\right)}\Big]\mathrm{d}s\mathrm{d}u\\
&\quad = \mathrm{e}^{-\nu(A)}\sum_{n=0}^{\infty}\frac{\nu(A)^n}{n!}\int_B\int_0^T\mathbb{E}\Big[\mathrm{e}^{\beta\lambda_{u,s}\left(V^1_{v,t}+\cdots+V^{n}_{v,t}+y_{v,t}\right)}\Big]\mathrm{d}s\mathrm{d}u\\
&\quad = \int_B\int_0^T\mathbb{E}\Big[\mathrm{e}^{\beta\lambda_{u,s}\left(X^1_{v,t}+y_{v,t}\right)}\Big]\mathrm{d}s\mathrm{d}u.
\end{aligned}$$

This concludes the proof. □

Proof of Lemma C.3.12. First we show that

$$\mathbb{E}\exp(\gamma\sup_{(v,t)\in B'\times T'}X^2_{v,t})<\infty \tag{C.I.7}$$

for all $\gamma>0$. Applying arguments as in Section C.2, we write X^2 as the independent sum $X^2_{v,t} = Y^1_{v,t}+Y^2_{v,t}$. Here Y^1 is a compound Poisson sum

$$Y^1_{v,t} = \sum_{k=1}^{M}J^k_{v,t}$$

with finite intensity $\nu(A^c\cap D)<\infty$ and jump distribution $\nu_2 = \nu_{A^c\cap D}/\nu(A^c\cap D)$, where $D = \{z\in\mathbb{R}^{\mathbb{K}} : \inf_{(v,t)\in\mathbb{K}} z_{v,t} < -1\}$. Furthermore, Y^2 is infinitely divisible with Lévy measure $\nu_{A^c\cap D^c}$, the restriction of ν to the set $A^c\cap D^c = \{z\in\mathbb{R}^{\mathbb{K}} : \sup_{(v,t)\in\mathbb{K}}|z_{v,t}|\le 1\}$. By arguments as before, both fields have t-càdlàg extensions to $B'\times T'$. For each k, $J^k_{v,t}\le 0$ for all $(v,t)\in B'\times T'$ almost surely, and in particular $\mathbb{E}\exp(\gamma\sup_{(v,t)\in B'\times T'}Y^1_{v,t})<\infty$ for all $\gamma>0$. As $(Y^2_{v,t})_{(v,t)\in B'\times T'}$ is t-càdlàg on the compact set $B'\times T'$, we find that

$\mathbb{P}(\sup_{(v,t)\in B'\times T'}|Y^2_{v,t}|<\infty)=1$. Since also $\nu_{A^c\cap D^c}(\{z\in\mathbb{R}^{\mathbb{K}}:\sup_{(v,t)\in\mathbb{K}}|z_{v,t}|>1\})=0$, we obtain from [7, Lemma 2.1] that $\mathbb{E}\exp(\gamma\sup_{(v,t)\in B'\times T'}|Y^2_{v,t}|)<\infty$ for all $\gamma>0$, which yields the claim (C.I.7).

Appealing to properties of $\lambda_{u,s}$ we find that

$$\lambda_{u,s}(X_{v,t})\le\lambda_{u,s}\Big(X^1_{v,t}+\sup_{(v,t)\in B'\times T'}X^2_{v,t}\Big)=\lambda_{u,s}\left(X^1_{v,t}\right)+\frac{\sup_{(v,t)\in B'\times T'}X^2_{v,t}}{c}.$$

The assertion now follows from (C.I.7) and the first claim of Theorem C.3.11. $\square$

Proof of Theorem C.3.13. Let π be the distribution of $(X^2_{v,t})_{(v,t)\in B'\times T'}$. Conditioning on $(X^2_{v,t})_{(v,t)}=(y_{v,t})_{(v,t)}$ we find by independence that

$$\frac{\mathbb{P}(\Psi(X_{v,t})>x)}{\mathbb{P}(\Psi(X^1_{v,t})>x)}=\int\frac{\mathbb{P}(\Psi(X^1_{v,t}+y_{v,t})>x)}{\mathbb{P}(\Psi(X^1_{v,t})>x)}\pi(\mathrm{d}y)=\int f(y;x)\pi(\mathrm{d}y)$$

with $f(y;x)=\mathbb{P}(\Psi(X^1_{v,t}+y_{v,t})>x)/\mathbb{P}(\Psi(X^1_{v,t})>x)$, which, according to Theorem C.3.11, satisfies

$$f(y;x)\to f(y)=\frac{\int_B\int_0^T\mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\Big(X^1_{v,t}+y_{v,t}\Big)\Big)\Big]\mathrm{d}s\mathrm{d}u}{\int_B\int_0^T\mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\Big(X^1_{v,t}\Big)\Big)\Big]\mathrm{d}s\mathrm{d}u}$$

as $x\to\infty$. By another application of Theorem C.3.11 and since

$$\int f(y)\pi(\mathrm{d}y)=\frac{\int_B\int_0^T\mathbb{E}\left[\exp\left(\beta\lambda_{u,s}(X_{v,t})\right)\right]\mathrm{d}s\mathrm{d}u}{\int_B\int_0^T\mathbb{E}\Big[\exp\Big(\beta\lambda_{u,s}\Big(X^1_{v,t}\Big)\Big)\Big]\mathrm{d}s\mathrm{d}u},$$

the proof is completed if we can find non-negative and integrable functions $g(y;x)$ and $g(y)=\lim_{x\to\infty}g(y;x)$ such that $f(y;x)\le g(y;x)$ and such that

$$\int g(y;x)\pi(\mathrm{d}y)\to\int g(y)\pi(\mathrm{d}y)$$

as $x\to\infty$. With $y^*=\sup_{(v,t)\in B'\times T'}y_{v,t}$ we use the function

$$g(y;x)=\mathbb{P}(\Psi(X^1_{v,t})+y^*>x)/\mathbb{P}(\Psi(X^1_{v,t})>x)$$

which, according to properties of $\lambda_{u,s}$ and Theorem C.3.11, satisfies

$$g(y;x)\to g(y)=\exp(\beta y^*/c)$$

as $x\to\infty$. From [16, Lemma 2.4(i)] and Theorem C.3.11 the distribution of $\Psi(X^1_{v,t})$ is convolution equivalent with index β/c. Now let G and H denote the distributions of $\Psi(X^1_{v,t})$ and $\sup_{(v,t)\in B'\times T'}X^2_{v,t}$, respectively. If $\overline{H}(x)=o(\overline{G}(x))$, $x\to\infty$, it follows from the integrability statement (C.I.7) and [16, Lemma 2.1] that

$$\begin{aligned}\int g(y;x)\pi(\mathrm{d}y)&=\frac{\mathbb{P}(\Psi(X^1_{v,t})+\sup_{(v,t)\in B'\times T'}X^2_{v,t}>x)}{\mathbb{P}(\Psi(X^1_{v,t})>x)}\\&\to\mathbb{E}\exp\Big(\frac{\beta}{c}\sup_{(v,t)\in B'\times T'}X^2_{v,t}\Big)=\int g(y)\pi(\mathrm{d}y)\end{aligned}$$

as $x\to\infty$. From (C.I.7) we find that $\lim_{x\to\infty}\mathrm{e}^{\gamma x}\mathbb{P}(\sup_{(v,t)\in B'\times T'}X^2_{v,t}>x)=0$ for all $\gamma>0$. Combined with the convolution equivalence of the distribution of $\Psi(X^1_{v,t})$, this yields $\overline{H}(x)=o(\overline{G}(x))$ and the claim follows. $\square$

C.II Proofs of Section C.5

Proof of Lemma C.5.5. Let $\omega \in \Omega_1'$ and $(s_n) \subset \tilde{S}$ such that $s_n \downarrow t \in [0,S]$. For all $k \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that

$$\|\boldsymbol{Z}_{s_n}(\omega) - \boldsymbol{Z}_{s_N}(\omega)\|_\infty \le \tfrac{1}{k} \quad \text{for all} \quad n \ge N. \tag{C.II.1}$$

This is seen by contradiction as follows: Assume that for any $N \in \mathbb{N}$ there exists $n \ge N$ such that

$$\|\boldsymbol{Z}_{s_n}(\omega) - \boldsymbol{Z}_{s_N}(\omega)\|_\infty > \tfrac{1}{k}.$$

Now fix $p \in \mathbb{N}$. By this there exist $n_0 < n_1 < n_2 < \cdots < n_p$ such that

$$\|\boldsymbol{Z}_{s_{n_j}}(\omega) - \boldsymbol{Z}_{s_{n_{j-1}}}(\omega)\|_\infty > \tfrac{1}{k} \quad \text{for} \quad j = 1,\dots,p$$

and we conclude that $\boldsymbol{Z}(\omega)$ has $\frac{1}{k}$-oscillation p times in $\tilde{S}$ for any p. Hence $\omega \in A_k^c$, which is a contradiction. From (C.II.1) and the fact that the metric space $(\mathcal{C}(K,\mathbb{R}), \|\cdot\|_\infty)$ is complete, we know that $\lim_{n\to\infty} \boldsymbol{Z}_{s_n}(\omega)$ exists with respect to $\|\cdot\|_\infty$ as a continuous function on K. To show uniqueness of the limit, let $(t_n) \subset \tilde{S}$ be another sequence such that $t_n \downarrow t$. Then $\lim_{n\to\infty} \boldsymbol{Z}_{s_n}(\omega) = \lim_{n\to\infty} \boldsymbol{Z}_{t_n}(\omega)$: Let (r_n) be the union of (s_n) and (t_n) ordered such that $r_n \downarrow t$. Then similarly for any $\epsilon > 0$ there is N' such that

$$\|\boldsymbol{Z}_{r_n}(\omega) - \boldsymbol{Z}_{r_{N'}}(\omega)\|_\infty < \tfrac{\epsilon}{2} \quad \text{for} \quad n \ge N'.$$

Also there is $N \in \mathbb{N}$ such that $(s_n)_{n\ge N}, (t_n)_{n \ge N} \subseteq (r_n)_{n\ge N'}$, and hence

$$\|\boldsymbol{Z}_{s_n}(\omega) - \boldsymbol{Z}_{t_n}(\omega)\|_\infty \le \|\boldsymbol{Z}_{s_n}(\omega) - \boldsymbol{Z}_{r_{N'}}(\omega)\|_\infty + \|\boldsymbol{Z}_{t_n}(\omega) - \boldsymbol{Z}_{r_{N'}}(\omega)\|_\infty < \epsilon$$

for all $n \ge N$. Thus, the limit $\lim_{s\in\mathbb{Q}, s\downarrow t} \boldsymbol{Z}_s(\omega)$ exists uniquely with respect to $\|\cdot\|_\infty$. Similarly for $\lim_{s\in\mathbb{Q}, s\uparrow t} \boldsymbol{Z}_s(\omega)$. □

We let

$$B(p,\epsilon,D) = \{\omega \in \Omega \mid \boldsymbol{Z}(\omega) \text{ has } \epsilon\text{-oscillation } p \text{ times in } D\},$$

with $D \subseteq \mathbb{Q} \cap [0,\infty)$, and

$$\alpha_\epsilon(r) = \sup\{\mathbb{P}(\|\boldsymbol{Z}_t\|_\infty \ge \epsilon) \mid t \in [0,r] \cap \mathbb{Q}\}.$$

Note that a direct consequence of the stochastic continuity from Lemma C.5.3 is that $\alpha_\epsilon(r) \to 0$ as $r \to 0$ for all $\epsilon > 0$.

Lemma C.II.1. *Let p be a positive integer, $D = \{t_1,\dots,t_n\} \subseteq \mathbb{Q} \cap [0,\infty)$ and $u, r \in \mathbb{Q}$ such that $0 \le u \le t_1 < \cdots < t_n \le r$. Then $\mathbb{P}(B(p, 4\epsilon, D)) \le (2\alpha_\epsilon(r-u))^p$.*

Proof. We will show the statement by induction in p. For this, define

$$\begin{aligned} C_k &= \{\|\boldsymbol{Z}_{t_j} - \boldsymbol{Z}_u\|_\infty \le 2\epsilon,\ j = 1,\dots,k-1,\ \|\boldsymbol{Z}_{t_k} - \boldsymbol{Z}_u\|_\infty > 2\epsilon\}\,, \\ D_k &= \{\|\boldsymbol{Z}_{t_k} - \boldsymbol{Z}_r\|_\infty > \epsilon\} \end{aligned}$$

and note that $C_1, \dots, C_n$ are disjoint and

$$\begin{aligned} B(1, 4\epsilon, D) &\subseteq \bigcup_{k=1}^{n} \{\|\boldsymbol{Z}_{t_k} - \boldsymbol{Z}_u\|_\infty > 2\epsilon\} = \bigcup_{k=1}^{n} C_k \\ &= \bigcup_{k=1}^{n} (C_k \cap D_k^c) \cup (C_k \cap D_k) \subseteq \{\|\boldsymbol{Z}_r - \boldsymbol{Z}_u\|_\infty \geq \epsilon\} \cup \bigcup_{k=1}^{n} (C_k \cap D_k). \end{aligned}$$

By the Lévy properties in Lemma C.5.3 we have $\mathbb{P}(\|\boldsymbol{Z}_r - \boldsymbol{Z}_u\|_\infty \geq \epsilon) \leq \alpha_\epsilon(r-u)$ and furthermore that $\mathbb{P}(C_k \cap D_k) = \mathbb{P}(C_k)\mathbb{P}(D_k) \leq \mathbb{P}(C_k)\alpha_\epsilon(r-u)$. The fact that $C_1, \dots, C_n$ are disjoint then implies

$$\mathbb{P}(B(1, 4\epsilon, D)) \leq \mathbb{P}(\|\boldsymbol{Z}_r - \boldsymbol{Z}_u\|_\infty \geq \epsilon) + \sum_{k=1}^{n} \mathbb{P}(C_k \cap D_k) \leq 2\alpha_\epsilon(r-u),$$

which is the desired expression for $p = 1$. We now assume the result to be true for arbitrary $p \in \mathbb{N}$. We define the sets

$$\begin{aligned} F_k &= \{\omega : \boldsymbol{Z}(\omega) \text{ has } 4\epsilon\text{-oscillation } p \text{ times in } \{t_1, \dots, t_k\}, \\ &\qquad \text{but does not have } 4\epsilon\text{-oscillation } p \text{ times in } \{t_1, \dots, t_{k-1}\}\}\,, \\ G_k &= \{\omega : \boldsymbol{Z}(\omega) \text{ has } 4\epsilon\text{-oscillation one time in } \{t_k, \dots, t_n\}\}. \end{aligned}$$

Then $F_1, \dots, F_n$ are disjoint, and $\mathbb{P}(F_k \cap G_k) = \mathbb{P}(F_k)\mathbb{P}(G_k)$ for all $k = 1, \dots, n$ due to the Lévy properties. Also $B(p, 4\epsilon, D) = \cup_{k=1}^{n} F_k$, and furthermore

$$B(p+1, 4\epsilon, D) = \bigcup_{k=1}^{n} (F_k \cap G_k)$$

with the inclusion $\subseteq$ seen as follows: If $\omega \in B(p+1, 4\epsilon, D)$ then $\boldsymbol{Z}(\omega)$ has 4ϵ-oscillation $p+1$ times in some $\{t_{n_0}, \dots, t_{n_{p+1}}\} \subseteq D$ with $n_0 < n_1 < \dots < n_{p+1}$. Hence there is $k \leq n_p$ such that $\omega \in F_k$. Also $\|\boldsymbol{Z}_{t_{n_{p+1}}}(\omega) - \boldsymbol{Z}_{t_{n_p}}(\omega)\|_\infty > 4\epsilon$ and as such also $\omega \in G_k$. From the induction assumption, the case $p = 1$ and the fact that $F_1, \dots, F_n$ are disjoint we find that

$$\begin{aligned} \mathbb{P}(B(p+1, 4\epsilon, D)) &= \sum_{k=1}^{n} \mathbb{P}(G_k)\mathbb{P}(F_k) \leq 2\alpha_\epsilon(r-u)\mathbb{P}\Big(\bigcup_{k=1}^{n} F_k\Big) \\ &= 2\alpha_\epsilon(r-u)\mathbb{P}(B(p, 4\epsilon, M)) \leq (2\alpha_\epsilon(r-u))^{p+1}\,. \end{aligned}$$

□

Proof of Lemma C.5.6. To show that $\mathbb{P}(\Omega_1') = 1$ it suffices to prove $\mathbb{P}(A_k^c) = 0$ for any fixed $k \in \mathbb{N}$. Since $\alpha_\epsilon(r) \to 0$ as $r \downarrow 0$ for any $\epsilon > 0$, we can choose $\ell \in \mathbb{N}$ such that $2\alpha_{1/(4k)}(S/\ell) < 1$. Then by continuity of $\mathbb{P}$ we get

$$\begin{aligned} \mathbb{P}(A_k^c) &\leq \mathbb{P}(\boldsymbol{Z} \text{ has } \tfrac{1}{k}\text{-oscillation infinitely often in } \tilde{S}) \\ &\leq \sum_{j=1}^{\ell} \mathbb{P}(\boldsymbol{Z} \text{ has } \tfrac{1}{k}\text{-oscillation infinitely often in } [\tfrac{j-1}{\ell}S, \tfrac{j}{\ell}S] \cap \mathbb{Q}) \\ &= \sum_{j=1}^{\ell} \lim_{p \to \infty} \mathbb{P}(B(p, \tfrac{1}{k}, [\tfrac{j-1}{\ell}S, \tfrac{j}{\ell}S] \cap \mathbb{Q})). \end{aligned}$$

Now fix $j = 1,\dots,\ell$, and enumerate the elements of $[\frac{j-1}{\ell}S, \frac{j}{\ell}S] \cap \mathbb{Q}$ by $(t_m)_{m\in\mathbb{N}}$. From Lemma C.II.1 we know that

$$\mathbb{P}(B(p, \tfrac{1}{k}, \{t_1,\dots,t_n\})) \leq (2\alpha_{1/(4k)}(\tfrac{S}{\ell}))^p$$

for any $n \in \mathbb{N}$. By continuity of $\mathbb{P}$ we see that

$$\mathbb{P}(B(p, \tfrac{1}{k}, [\tfrac{j-1}{\ell}S, \tfrac{j}{\ell}S] \cap \mathbb{Q})) = \lim_{n\to\infty} \mathbb{P}(B(p, \tfrac{1}{k}, \{t_1,\dots,t_n\})) \leq (2\alpha_{1/(4k)}(\tfrac{S}{\ell}))^p$$

which tends to 0 as $p \to \infty$ since ℓ is chosen such that $2\alpha_{1/(4k)}(S/\ell) < 1$. As this holds for all $j = 1,\dots,\ell$ we conclude that $\mathbb{P}(A_k^c) = 0$. □

References

[1] Adler, R.J. (1981). *The Geometry of Random Fields*. John Wiley & Sons, New York.

[2] Adler, R.J., G. Samorodnitsky and J.E. Taylor (2010). Excursion sets of three classes of stable random fields. *Adv. Appl. Probab.* 42, 293–318.

[3] Adler, R.J., G. Samorodnitsky and J.E. Taylor (2013). High level excursion set geometry for non–Gaussian infinitely divisible random fields. *Ann. Probab.* 41(1), 134–169.

[4] Adler, R.J. and J.E. Taylor (2007). *Random Fields and Geometry*. Springer, New York.

[5] Barndorff-Nielsen, O.E. and J. Schmiegel (2004). Lévy-based Tempo-Spatial Modelling with Applications to Turbulence. *Uspekhi Mat. Nauk* 159, 63–90.

[6] Barndorff-Nielsen, O.E. and N. Shephard (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B* 63(2), 167–241.

[7] Braverman, M. and G. Samorodnitsky (1995). Functionals of infinitely divisible stochastic processes with exponential tails. *Stoch. Process. Their Appl.* 56(2), 207–231.

[8] Cline, D.B.H. (1986). Convolution tails, product tails and domains of attraction. *Probab. Theory Relat. Fields* 72(4), 529–557.

[9] Cline, D.B.H. (1987). Convolutions of Distributions With Exponential and Subexponential Tails. *J. Aust. Math. Soc. Ser. A* 43(3), 347–365.

[10] Fasen, V. (2008). Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. *Extremes* 12(3), 265–296.

[11] Hellmund, G., M. Prokešová and E.B.V. Jensen (Sept. 2008). Lévy-based Cox point processes. *Adv. Appl. Probab.* 40, 603–629.

[12] Jónsdóttir, K.Ý., A. Rønn-Nielsen, K. Mouridsen and E.B.V. Jensen (2013). Lévy-based Modelling in Brain Imaging. *Scand. Stat. Theory Appl.* 40(3), 511–529.

[13] Jónsdóttir, K.Ý., J. Schmiegel and E.B.V. Jensen (2008). Lévy-based growth models. *Bernoulli* 14(1), 62–90.

[14] Kechris, A. (2012). *Classical Descriptive Set Theory*. Graduate Texts in Mathematics. Springer, New York.

[15] Linde, W. (1983). *Infinitely Divisible and Stable Measures on Banach Spaces*. Teubner-Texte zur Mathematik. B.G. Teubner Verlagsgesellschaft.

[16] Pakes, A.G. (2004). Convolution equivalence and infinite divisibility. *J. Appl. Probab.* 41, 407–424.

[17] Peccati, G. and M.S. Taqqu (2011). *Combinatorial Expressions of Cumulants and Moments*. Springer, Milan.

[18] Rajput, B.S. and J. Rosinski (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Relat. Fields* 82, 451–487.

[19] Rønn-Nielsen, A. and E.B.V. Jensen (2016). Tail asymptotics for the supremum of an infinitely divisible field with convolution equivalent Lévy measure. *J. Appl. Probab.* 53(1), 244–261.

[20] Rønn-Nielsen, A. and E.B.V. Jensen (2017). Excursion sets of infinitely divisible random fields with convolution equivalent Lévy measure. *J. Appl. Probab.* 54(3), 833–851.

[21] Rønn-Nielsen, A. and E.B.V. Jensen (2019). Central limit theorem for mean and variogram estimators in Lévy-based models. *J. Appl. Probab.* 56(1), 209–222.

[22] Rønn-Nielsen, A., J. Sporring and E.B.V. Jensen (2017). Estimation of sample spacing in stochastic processes. *Image Anal. Stereol.* 36(1), 43–49.

[23] Rosinski, J. and G. Samorodnitsky (1993). Distributions of Subadditive Functionals of Sample Paths of Infinitely Divisible Processes. *Ann. Probab.* 21(2), 996–1014.

[24] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.

[25] Sommerfeld, M., S. Sain and A. Schwartzman (2018). Confidence regions for spatial excursion sets from repeated random field observations, with an application to climate. *J. Am. Stat. Assoc.* 113(523), 1327–1340.

[26] Spodarev, E. (2014). "Limit theorems for excursion sets of stationary random fields". *Modern Stochastics and Applications*. Ed. by V. Korolyuk, N. Limnios, Y. Mishura, L. Sakhno and G. Shevchenko. Switzerland: Springer, New York, 221–241.

Paper D

Extreme Value Theory for Spatial Random Fields – With Application to a Lévy-Driven Field

Mads Stehr and Anders Rønn-Nielsen

Submitted to *Extremes*

Abstract

First, we consider a stationary random field indexed by an increasing sequence of subsets of $\mathbb{Z}^d$. Under certain mixing and anti–clustering conditions combined with a very broad assumption on how the sequence of spatial index sets increases, we obtain an extremal result that relates a normalized version of the distribution of the maximum of the field over the index sets to the tail distribution of the individual variables. Furthermore, we identify the limiting distribution as an extreme value distribution.

Secondly, we consider a continuous, infinitely divisible random field indexed by $\mathbb{R}^d$ given as an integral of a kernel function with respect to a Lévy basis with convolution equivalent Lévy measure. When observing the supremum of this field over an increasing sequence of (continuous) index sets, we obtain an extreme value theorem for the distribution of this supremum. The proof relies on discretization and a conditional version of the technique applied in the first part of the paper, as we condition on the high activity and light–tailed part of the field.

MSC: 60G70; 60G60; 60E07; 60D05; 52A39

Keywords: Extreme value theory; spatial models; Lévy–based modeling; geometric probability; intrinsic volumes; convolution equivalence; random fields

D.1 Introduction

In classical extreme value theory the aim is to describe the asymptotic distributional properties of

$$M_n = \max\{\xi_1, \dots, \xi_n\}$$

as $n \to \infty$, where $\xi_1, \xi_2, \dots$ are independent and identically distributed. A main result is that if there exist sequences (a_n) and (b_n) such that $\mathbb{P}(a_n(M_n - b_n) \le x) \to G(x)$ for all x, where G is a non–degenerate function, then G is in fact the distribution function of a distribution belonging to one of three specific types called *extreme value distributions*; see e.g the monographs [5, 11, 15] for thorough treatments of both the classical extreme value theory and many important extensions and applications.

A very useful result, when relating the limiting extremal distribution with the distribution of the individual ξ–variables, is the following theorem, cf. [11, Theorem 1.5.1].

Theorem D.1.1. *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed real random variables and let $0 \le \tau \le \infty$. Then for a real sequence $(x_n)_{n\in\mathbb{N}}$,*

$$n\mathbb{P}(\xi_1 \le x_n) \to \tau \quad \text{as } n \to \infty$$

if and only if

$$\mathbb{P}(M_n \le x_n) \to e^{-\tau} \quad \text{as } n \to \infty.$$

A relevant extension, studied in the literature, of the independent case is assuming that $(\xi_n)_{n\in\mathbb{N}}$ is a stationary sequence of random variables that are not necessarily independent. Obtaining results on the extremal behavior of the maximum of dependent variables will now be a question of controlling this dependence. A key property in this framework is an adapted version of Theorem D.1.1: The result of the theorem is still true under additional mixing and anti–clustering conditions ensuring that the sequence on a large scale behaves like an independent sequence; see [11, Chapter 3] for a detailed exposition. A generalization to stationary stochastic processes in continuous time can be found in [11, Chapter 13].

In the present paper we extend the index set from a one–dimensional time axis to a spatial setting with indices in $\mathbb{Z}^d$ and $\mathbb{R}^d$. The contribution of the paper will be two–fold. First we consider a stationary random field $(\xi_v)_{v\in\mathbb{Z}^d}$ and a sequence of finite, increasing index sets $(D_n)_{n\in\mathbb{N}}$ with $D_n \subseteq \mathbb{Z}^d$. A main result will be a new version of Theorem D.1.1 in this setting, where we relate the convergence of $\mathbb{P}(\xi_v > x_n)$ to that of $\mathbb{P}(M(D_n) \le x_n)$ with the notation $M(D_n) = \max_{v\in D_n} \xi_v$, which will be used throughout. This is formulated as Theorem D.3.5 in Section D.3 below. To obtain this result, we need to impose conditions on the spatial dependence structure in the same spirit as the mixing and anti–clustering conditions needed in the one–dimensional case, cf. the conditions $\mathcal{D}(x_n; K_n)$ and $\mathcal{D}'(x_n)$ defined below.

While the index set in the one–dimensional case always has the form $D_n = \{1, \dots, m_n\}$, most often with $m_n = n$, a challenging part of the generalization to a spatial setting is to obtain the desired result under realistic and useful assumptions on the increasing behavior of the sequence of index sets (D_n). It is needed that D_n expands in a particularly nice way such that it can be approximated by a certain class of expanding cubes. The sufficient assumption, Assumption D.2.4, on (D_n) will be given in Section D.2 together with the introduction of relevant geometrical notation. The assumption is formulated in terms of the so–called intrinsic volumes of the (continuous version of) the sets.

The geometric assumption on (D_n) is formulated in a large generality, but as illustrated in Example D.2.6 it includes the simple, but useful, situation, where a fixed set is scaled up by an increasing sequence (r_n),

$$D_n = (r_n C) \cap \mathbb{Z}^d.$$

Here C is a union of finitely many bounded, convex and full–dimensional sets.

Extreme value theory formulated in a spatial setting is, to the best of the authors' knowledge, rare in the literature. However, in [12] a coordinate–wise spatial mixing condition is formulated. Furthermore it is obtained that under this condition the limiting distribution of a normalization of $M(D_n)$ is an extreme value distribution in the asymptotic scenario, where the index sets D_n constitutes an increasing sequence of boxes.

The second part of the paper concerns a random field $(X_v)_{v\in\mathbb{R}^d}$ defined by

$$X_v = \int_{\mathbb{R}^d} f(|v-u|)M(\mathrm{d}u), \tag{D.1.1}$$

where M is an infinitely divisible, independently scattered random measure on $\mathbb{R}^d$ and f is an appropriate kernel function. We furthermore assume that the Lévy measure of the random measure M has a convolution equivalent right tail; see [3, 4] for details about convolution equivalent distributions and [13] for the relation between convolution equivalence and infinite divisibility. Examples of convolution equivalent distributions that are also infinitely divisible counts the important cases of the inverse Gaussian and the normal inverse Gaussian distribution; see examples 2.1 and 2.2 in [16].

Lévy-driven moving average models as defined in (D.1.1) form a very flexible framework that recently have been used for multiple modeling purposes. This includes modeling of turbulent flows ([1]), growth processes ([9]), Cox point processes ([7]), and brain imaging data ([8, 19])

For a sequence $(C_n)_{n\in\mathbb{N}}$ of index sets satisfying a continuous version of the assumption imposed on discrete sets $(D_n)_{n\in\mathbb{N}}$ above, and under mild restrictions on the kernel function f, we obtain the main result that for all $x \in \mathbb{R}$,

$$\mathbb{P}\Big(a_n^{-1}\Big(\sup_{v\in C_n} X_v - b_n\Big) \le x\Big) \to \exp\Big(-\mathrm{e}^{-x}\mathbb{E}\mathrm{e}^{\beta X_u}\rho((1,\infty))\Big) \tag{D.1.2}$$

as $n \to \infty$, where a_n, b_n are norming constants chosen according to the extremal behavior of the Lévy measure ρ relative to the volume $|C_n|$ of the index set, such that $\lim_n |C_n|\rho((a_n x + b_n, \infty)) = \mathrm{e}^{-x}\rho((1,\infty))$ for all $x \in \mathbb{R}$.

The proof of this result relies on a discretization, writing the supremum as a maximum of suprema over unit cubes. Here a result from [16] becomes rather beneficial: The tail distribution of e.g. $\sup_{v\in[0,1]^d} X_v$ is asymptotically equivalent with that of the underlying Lévy measure ρ. Obtaining the result will, however, not be a direct application of the result for stationary, discretely indexed random fields from the first part of the paper. Showing directly that the mixing and anti–clustering conditions

$\mathcal{D}(x_n;K_n)$ and $\mathcal{D}'(x_n)$ are satisfied is rather challenging, while it is easier to apply an independent decomposition $X = Z + Y$, where Z is a compound Poisson random field with relatively heavy tails, and Y represents the part of X with infinite activity but lighter tails. The proof strategy will be to condition on $Y = y$ and then establish a conditional result on the extremal behavior of the field $Z + y$. The final result is obtained by applying ergodic properties of the conditioning Y–field. This proof strategy has the additional advantage that an extension of (D.1.2) follows directly with X replaced by $X + \tilde{Y}$, where $\tilde{Y}$ is a sufficiently light–tailed, stationary and ergodic random field that is independent of X, cf. Theorem D.4.11 below.

In [6] a stochastic process $(X_t)_{t\geq 0}$ on the form

$$X_t = \int_{\mathbb{R}_+\times\mathbb{R}} f(r,t-s)M(\mathrm{d}r,\mathrm{d}s)$$

is considered. Note that the index set here is one–dimensional. Furthermore, the kernel function f is assumed to satisfy $f(r,s) = 0$ for $s < 0$. Under very similar conditions on the random measure M, in particular it is assumed that ρ has a convolution equivalent right tail, an asymptotic result, similar to (D.1.2), is obtained for $\sup_{t\in[0,T]} X_t$ as $T \to \infty$.

An asymptotic result for a discretely observed field on the form (D.1.1) in a scenario with an increasing spatial index set can be found in [18]. The paper proposes estimators for the mean and covariance function of the field and furthermore provides central limit theorems for these estimators. The spatial asymptotic scenario of the index sets is, however, less restrictive compared to the requirements in the present paper. There it is only needed that the discrete index set D_n increases in a way, such that the surface of D_n is asymptotically inferior to the volume of D_n.

The paper is organized as follows. In Section D.2 we introduce a few geometrical concepts and formulate the assumption on the increasing sequence of index sets both in the discrete and continuous setting. In Section D.3 we state and prove a spatial version of Theorem D.1.1 under the assumption on (D_n) from Section D.2 and additional mixing and anti–clustering conditions. We furthermore show that the limiting distribution of a normalized version of $M(D_n)$ is an extreme value distribution.

Section D.4 is devoted to the introduction of the Lévy–based model (D.1.1) and the proofs of the main extremal results. The proofs that relates to the conditioning (Y_v)–field – in particular results on its ergodic behavior – are found in Section D.5, while proofs of some further technical lemmas, used in the proof of the main theorem, are deferred to Section D.6.

D.2 Geometrical assumption and result

The main purpose of this section will be to formulate the sufficient assumption on the sequence of index sets. It turns out that in both the discrete and the continuous setting a relevant property should be formulated through conditions on sets in $\mathbb{R}^d$. That means that for the sequence (D_n) of sets in $\mathbb{Z}^d$ used for the discrete result we will assume the existence of a sequence (C_n) of sets in $\mathbb{R}^d$ satisfying certain conditions and that $D_n = C_n \cap \mathbb{Z}^d$.

The relevant condition on the sequence of sets can be found in Assumption D.2.4 below. However, the formulation of this assumption requires a few concepts and results from the classical theory of geometry for convex sets. This is summarized below in Theorem D.2.1 and the properties immediately thereafter; see e.g. [21, Chapter 4] for a detailed exposition of this topic.

Throughout the paper, the following notation will be used. For a subset A of either $\mathbb{Z}^d$ or $\mathbb{R}^d$, we let $|A|$ denote the size of the set A. Hence, $|A|$ is the number of points in $A \subseteq \mathbb{Z}^d$, and $|A|$ is the Lebesgue measure of $A \subseteq \mathbb{R}^d$. However, we will only consider the Lebesgue measure of unions of convex and full–dimensional sets in $\mathbb{R}^d$, so $|A|$ will in fact also be the Hausdorff measure of A. For a set $A \subseteq \mathbb{R}^d$ and a constant $\lambda > 0$ we define $\lambda A = \{\lambda x \mid x \in A\}$. For two sets $A, B \subseteq \mathbb{R}^d$, we define the Minkowski sum by $A \oplus B = \{a + b \mid a \in A, b \in B\}$. We let $0 \in \mathbb{R}^d$ denote the origin of $\mathbb{R}^d$, and we define $B(r)$ as the closed ball in $\mathbb{R}^d$ centered in 0 and with radius $r \geq 0$.

A compact convex set with non–empty interior will in the following be called a convex body. The following theorem, called Steiner's Theorem, constitutes a classical result from convex geometry.

Theorem D.2.1. *Assume that C is a convex body and let $r > 0$. Then the volume $|C \oplus B(r)|$ is a polynomial in r, i.e.*

$$|C \oplus B(r)| = \sum_{j=0}^{d} \omega_{d-j} V_j(C) r^{d-j},$$

where $V_j(C)$ are coefficients that only depend on C, and ω_j is the volume of the j–dimensional unit ball.

The coefficients $V_j(C)$ are called intrinsic volumes of the convex body C and satisfy

1. $V_0(C) = 1$,
2. $V_1(C) = \frac{d\omega_d}{2\omega_{d-1}} b(C)$, where $b(C)$ is the mean width of C,
3. $V_{d-1}(C) = F(C)/2$, where $F(C)$ denotes the surface area of C, i.e. the $d-1$ dimensional Hausdorff measure of the surface ∂C,
4. $V_d(C) = |C|$.

Furthermore, the functionals $V_j : \mathcal{K} \to \mathbb{R}$, where $\mathcal{K}$ denotes the collection of all convex bodies, satisfy some important properties, among which we mention

(i) They are non–negative, i.e. $V_j(C) \geq 0$ for all $C \in \mathcal{K}$.

(ii) They are motion invariant, i.e. $V_j(gC) = V_j(C)$ for any euclidean motion g.

(iii) They are homogeneous, i.e. $V_j(\gamma C) = \gamma^j V_j(C)$ for all $\gamma > 0$.

(iv) They are monotone, i.e. $C \subseteq D$ implies $V_j(C) \leq V_j(D)$.

Corollary D.2.2. *Let $C \subseteq \mathbb{R}^d$ be a convex body and $r > 0$. Then*

$$\sum_{j=0}^{d-1} \omega_{d-j} V_j(C) r^{d-j} \leq |\partial C \oplus B(r)| \leq 2 \sum_{j=0}^{d-1} \omega_{d-j} V_j(C) r^{d-j}.$$

Proof. For the first inequality, we apply Theorem D.2.1 to find

$$|\partial C \oplus B(r)| \geq |\big(C \oplus B(r)\big) \setminus C| = |C \oplus B(r)| - |C| = \sum_{j=0}^{d-1} \omega_{d-j} V_j(C) r^{d-j}.$$

For the second inequality, we define $C_{-r} = C \setminus \big(\partial C \oplus B(r)\big)$. Then C_{-r} is a convex set: Let $x, y \in C_{-r}$ and choose a point z on the line segment from x to y. If we can show $B(r) + z \subseteq C$, we will have $z \in C_{-r}$ and thus convexity is obtained. For this, choose $z' \in z + B(r)$, i.e. there is $b \in B(r)$ such that $z' = z + b$. Then z' is on the line segment from $x + b$ to $y + b$, and furthermore both $x + b$ and $y + b$ are in C. The convexity of C gives the desired result.

Arguing as above, using monotonicity of the intrinsic volumes and the fact that $C = C_{-r} \oplus B(r)$, we find

$$|C \setminus C_{-r}| = |C_{-r} \oplus B(r)| - |C_{-r}| \leq \sum_{i=0}^{d-1} \omega_{d-j} V_j(C) r^{d-j}$$

Since $\partial C \oplus B(r) = \big((C \oplus B(r)) \setminus C\big) \cup \big(C \setminus C_{-r}\big)$, we can deduce the desired inequality. □

Definition D.2.3. *A set $C \subseteq \mathbb{R}^d$ is said to be p–convex, if it is connected and has the form*

$$C = \bigcup_{i=1}^{p} C_i,$$

where $C_1, \ldots, C_p$ are convex bodies in $\mathbb{R}^d$.

In the following we give the assumption on the index sets (C_n) used in the paper. Due to stationarity of all fields involved, we can without loss of generality in later results assume that $0 \in C_n$ for all $n \in \mathbb{N}$. Although not formulated in the assumption, this will be assumed throughout the paper.

Assumption D.2.4. *The sequence $(C_n)_{n\in\mathbb{N}}$ consists of p–convex bodies, where*

$$C_n = \bigcup_{i=1}^{p} C_{n,i}$$

and $|C_n| \to \infty$ as $n \to \infty$. Furthermore,

$$\frac{\sum_{i=1}^{p} V_j(C_{n,i})}{|C_n|^{j/d}} \quad \textit{is bounded for each } j = 1, \ldots, d-1. \tag{D.2.1}$$

Concerning the sequence $(D_n)_{n\in\mathbb{N}}$ of discrete sets in $\mathbb{Z}^d$, we say that $(D_n)_{n\in\mathbb{N}}$ satisfies Assumption D.2.4 if there exists a sequence $(C_n)_{n\in\mathbb{N}}$ of sets in $\mathbb{R}^d$ satisfying the assumption, such that $D_n = C_n \cap \mathbb{Z}^d$.

For a sequence $(C_n)_{n\in\mathbb{N}}$ satisfying Assumption D.2.4 we will for each $k \in \mathbb{N}$ and a fixed λ close to 1, define $t_{n,k,\lambda} = \lfloor \sqrt[d]{|C_n|/(\lambda k)} \rfloor$. Most often we will choose $\lambda = 1$, but in Theorems D.2.5 and D.3.4 and Lemmas D.3.2 and D.3.3 we will use the full generality. Subsequent to Theorem D.3.4, we set $\lambda = 1$ and λ will be suppressed from all notation. For $z = (z_1, \ldots, z_d) \in \mathbb{Z}^d$ and n large enough relative to k we furthermore define the cube $I_z^{n,k,\lambda}$ as

$$I_z^{n,k,\lambda} = \bigtimes_{i=1}^{d} \Big[z_i t_{n,k,\lambda}, (z_i+1) t_{n,k,\lambda} \Big).$$

Let $P_{n,k,\lambda}$ be the set of indices z for which $I_z^{n,k,\lambda}$ is contained in C_n and $Q_{n,k,\lambda}$ be the set of indices z for which $I_z^{n,k,\lambda}$ is intersected by C_n, respectively. I.e.

$$P_{n,k,\lambda} = \{z \in \mathbb{Z}^d \,:\, I_z^{n,k,\lambda} \subseteq C_n\}, \quad \text{and} \quad Q_{n,k,\lambda} = \{z \in \mathbb{Z}^d \,:\, I_z^{n,k,\lambda} \cap C_n \neq \emptyset\}.$$

Let furthermore

$$p_{n,k,\lambda} = |P_{n,k,\lambda}| \qquad \text{and} \qquad q_{n,k,\lambda} = |Q_{n,k,\lambda}|.$$

Note that by construction, $p_{n,k,\lambda} \leq \lambda k$ and $q_{n,k,\lambda} \geq \lambda k$ for values of n large enough relative to k. Let $J_z^{n,k,\lambda} = I_z^{n,k,\lambda} \cap \mathbb{Z}^d$ be the integer numbers in $I_z^{n,k,\lambda}$. Define

$$D_{n,k,\lambda}^{-} = \bigcup_{z \in P_{n,k,\lambda}} J_z^{n,k,\lambda} \qquad \text{and} \qquad D_{n,k,\lambda}^{+} = \bigcup_{z \in Q_{n,k,\lambda}} J_z^{n,k,\lambda}.$$

With $D_n = C_n \cap \mathbb{Z}^d$ we have $J_z^{n,k,\lambda} \subseteq D_n$ for all $z \in P_{n,k,\lambda}$, and that $z \in Q_{n,k,\lambda}$ for all $J_z^{n,k,\lambda}$ with $J_z^{n,k,\lambda} \cap D_n \neq \emptyset$. That gives

$$D_{n,k,\lambda}^{-} \subseteq D_n \subseteq D_{n,k,\lambda}^{+}. \tag{D.2.2}$$

Theorem D.2.5. *Let $(C_n)_{n \in \mathbb{N}}$ satisfy Assumption D.2.4, and let $D_n = C_n \cap \mathbb{Z}^n$. Then*

(i) *$|D_n| \sim |C_n|$ as $n \to \infty$,*

(ii) *for all λ the sequences $p_{n,k,\lambda}$ and $q_{n,k,\lambda}$, defined above, satisfy that*

$$\liminf_{n \to \infty} p_{n,k,\lambda} \sim \lambda k \qquad \textit{and} \qquad \limsup_{n \to \infty} q_{n,k,\lambda} \sim \lambda k$$

as $k \to \infty$,

(iii) *for each k, λ and n with n large enough relative to k, it holds that $D_{n,k,\lambda}^{+} \subseteq K_{n,k,\lambda}$, where $K_{n,k,\lambda}$ is the cube*

$$K_{n,k,\lambda} = \bigcup_{z \in N_{k,\lambda}} J_z^{n,k,\lambda},$$

and $N_{k,\lambda}$ is on the form $N_{k,\lambda} = [-c_{k,\lambda}\,,\, c_{k,\lambda}]^d \cap \mathbb{Z}^d$ for some $c_{k,\lambda} < \infty$,

(iv) *there exists $c < \infty$ such that for all $n \in \mathbb{N}$*

$$D_{n,k,\lambda}^{+} \subseteq K_n = \Big[-c \cdot |C_n|^{1/d} \,,\, c \cdot |C_n|^{1/d} \Big]^d \cap \mathbb{Z}^d.$$

In (iii) above the important property is that $N_{k,\lambda}$ does not depend on n. That means in particular that for k and λ fixed, all D_n are included in the same finite collection of (increasing) cubes $J_z^{n,k,\lambda}$.

Proof. We start by demonstrating statement (ii). Defining $\tilde{t}_{n,k,\lambda} = \left(\frac{|C_n|}{\lambda k}\right)^{1/d}$ for each $k, n \in \mathbb{N}$ and using Corollary D.2.2, we find

$$|\partial C_n \oplus B(\tilde{t}_{n,k,\lambda})| \leq \sum_{i=1}^{p} |\partial C_{n,i} \oplus B(\tilde{t}_{n,k,\lambda})| \leq 2\sum_{j=0}^{d-1} \omega_{d-j}\left(\sum_{i=1}^{p} V_j(C_{n,i})\right)\tilde{t}_{n,k,\lambda}^{d-j}.$$

From straightforward calculations,

$$\frac{1}{\lambda k}\frac{1}{\tilde{t}_{n,k,\lambda}^{d}}|\partial C_n \oplus B(\tilde{t}_{n,k,\lambda})| \leq 2\sum_{j=0}^{d-1} \omega_{d-j}\frac{\sum_{i=1}^{p} V_j(C_{n,i})}{|C_n|^{j/d}}\left(\frac{1}{\lambda k}\right)^{\frac{d-j}{d}},$$

such that (D.2.1) gives

$$\limsup_{n\to\infty} \frac{1}{\lambda k}\frac{1}{\tilde{t}_{n,k,\lambda}^{d}}|\partial C_n \oplus B(\tilde{t}_{n,k,\lambda})| \to 0$$

as $k \to \infty$. Using that $\tilde{t}_{n,k,\lambda/c^d} = c\tilde{t}_{n,k,\lambda}$ and replacing λ by λ/c^d in the limit above, we find that in fact

$$\limsup_{n\to\infty} \frac{1}{\lambda k}\frac{1}{\tilde{t}_{n,k,\lambda}^{d}}|\partial C_n \oplus B(c\tilde{t}_{n,k,\lambda})| \to 0$$

as $k \to \infty$ for all $c > 0$. Since $t_{n,k,\lambda} \leq \tilde{t}_{n,k,\lambda}$, and $t_{n,k,\lambda} \sim \tilde{t}_{n,k,\lambda}$ as $n \to \infty$, we also have

$$\limsup_{n\to\infty} \frac{1}{\lambda k}\frac{1}{t_{n,k,\lambda}^{d}}|\partial C_n \oplus B(ct_{n,k,\lambda})| \to 0 \tag{D.2.3}$$

as $k \to \infty$ for all $c > 0$. Recall that the side length of $I_z^{n,k,\lambda}$ is $t_{n,k,\lambda}$, so if $I_z^{n,k,\lambda} \cap \partial C_n \neq \emptyset$ then clearly $I_z^{n,k,\lambda} \subseteq C_n \oplus B(2\sqrt{d}\, t_{n,k\lambda})$. Using that $|I_z^{n,k,\lambda}| = t_{n,k,\lambda}^d$, we then find

$$\frac{q_{n,k,\lambda} - p_{n,k,\lambda}}{\lambda k} \leq \frac{1}{\lambda k}\frac{1}{t_{n,k,\lambda}^{d}}\Bigg|\bigcup_{I_z^{n,k,\lambda}\cap\partial C_n\neq\emptyset} I_z^{n,k,\lambda}\Bigg| \leq \frac{1}{\lambda k}\frac{1}{t_{n,k,\lambda}^{d}}|C_n \oplus B(2\sqrt{d}\, t_{n,k,\lambda})|.$$

Together with (D.2.3) and the fact that $p_{n,k,\lambda} \leq \lambda k \leq q_{n,k,\lambda}$ for n chosen large enough relative to k, this gives the statement in (ii).

For statement (i) we now use that $|D_{n,k,\lambda}^-| = p_{n,k,\lambda} \cdot t_{n,k,\lambda}^d$ and $|D_{n,k,\lambda}^+| = q_{n,k,\lambda} \cdot t_{n,k,\lambda}^d$ together with (D.2.2) to find

$$\frac{p_{n,k,\lambda} \cdot t_{n,k,\lambda}^d}{|C_n|} \leq \frac{|D_n|}{|C_n|} \leq \frac{q_{n,k,\lambda} \cdot t_{n,k,\lambda}^d}{|C_n|}.$$

Letting $n \to \infty$ and subsequently $k \to \infty$ gives the desired result.

To see statement (iii), we have for each k and λ that the sequence $(q_{n,k,\lambda})_n$ is bounded by a constant $c_{k,\lambda}$ for n large enough. Since $0 \in C_n$, we have $0 \in Q_{n,k,\lambda}$. Furthermore, C_n is connected, so $Q_{n,k,\lambda}$ consists of at most $c_{k,\lambda}$ points that are pairwise neighbors. Therefore,

$$Q_{n,k,\lambda} \subseteq [-c_{k,\lambda}\,,\, c_{k,\lambda}]^d \cap \mathbb{Z}^d$$

which, again, implies the desired result. Finally, (iv) follows easily from (iii). □

Example D.2.6. Let $C = \cup_{i=1}^{p} C_i$ be a p–convex set and define the sequence $(C_n)_{n\in\mathbb{N}}$ by

$$C_n = r_n C = \bigcup_{i=1}^{p} r_n C_i,$$

where $r_n \uparrow \infty$ as $n \to \infty$. Then $V_j(r_n C) = r_n^j V_j(C)$ for $j = 0,\dots,d$. In particular,

$$\frac{\sum_{i=1}^{d} V_j(r_n C_i)}{|r_n C|^{j/d}} = \frac{\sum_{i=1}^{d} V_j(C_i)}{|C|^{j/d}}$$

will be constant for $j = 1,\dots,d-1$. Thus the sequence $(C_n)_{n\in\mathbb{N}}$ satisfies Assumption D.2.4.

D.3 Extreme value theorem for a spatial stationary field on $\mathbb{Z}^d$

In this section we consider a stationary random field $\xi = (\xi_v)_{v\in\mathbb{Z}^d}$ and a sequence of index sets $(D_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}^d$ satisfying Assumption D.2.4. Below we formulate the sufficient mixing condition for ξ in terms of the behavior on the sufficient sets (K_n) associated to (D_n) by Theorem D.2.5, however, before doing so we introduce some further notation which will be used in this and the remaining sections of the paper.

We say that two subsets A, B of $\mathbb{Z}^d$ (or $\mathbb{R}^d$) are r-separated if $\operatorname{dist}(A,B) = \inf\{|a-b| : a \in A,\, b \in B\} \ge r$ and there are two disjoint sets $A', B' \subseteq \mathbb{R}^d$, which are both connected, such that $A \subseteq A'$ and $B \subseteq B'$. Moreover, when talking about a d–dimensional cube with side–length $s > 0$, we mean a box with all side–lengths equal to s. If the side–lengths of a box are not identical, we list them all.

Condition ($\mathcal{D}(x_n;K_n)$). *The condition $\mathcal{D}(x_n;K_n)$ is satisfied for the stationary random field $(\xi_v)_{v\in\mathbb{Z}^d}$ if there exists a sequence $\gamma_n = o(\sqrt[d]{|D_n|})$ such that for all $n \in \mathbb{N}$ and all γ_n-separated sets $A, B \subseteq K_n$ it holds that*

$$\Big|\mathbb{P}(M(A\cup B) \le x_n) - \mathbb{P}(M(A) \le x_n)\mathbb{P}(M(B) \le x_n)\Big| \le \alpha_n, \tag{D.3.1}$$

where $\alpha_n \to 0$ as $n \to \infty$.

We need an "approximate independence"similar to (D.3.1) for more than two separated sets. The lemma below follows easily by induction (and possibly a reordering), using the fact that $\cup_{i=1}^{r-1} A_i$ and A_r are γ_n-separated if all A_i, $i = 1,\dots,r$, are pairwise γ_n-separated for all $i \ne j$.

Lemma D.3.1. *Let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary field satisfying $\mathcal{D}(x_n;K_n)$, and let for $r \in \mathbb{N}$ the sets $A_1,\dots,A_r$ be pairwise γ_n-separated. Then*

$$\left|\mathbb{P}\Big(\bigcap_{i=1}^{r}\{M(A_i) \le x_n\}\Big) - \prod_{i=1}^{r}\mathbb{P}(M(A_i) \le x_n)\right| \le (r-1)\alpha_n.$$

Assuming $\mathcal{D}(x_n;K_n)$ we have $\gamma_n < t_{n,k,\lambda}$ for n sufficiently large relative to a fixed k. For $z \in N_{k,\lambda}$ we divide each $J_z = J_z^{n,k,\lambda}$ into two disjoint subsets, H_z and H_z^*, where

$$H_z = \Big\{u \in \mathbb{Z}^d \;:\; z_j t_{n,k,\lambda} \le u_j \le (z_j+1)t_{n,k,\lambda} - 1 - \gamma_n, \text{ for all } j = 1,\dots,d\Big\}.$$

We note that H_z^* is the union of the overlapping boxes $L_{z,1}^*,\dots,L_{z,d}^*$ given by

$$L_{z,j}^* = \Big\{u \in J_z \,:\, (z_j+1)t_{n,k,\lambda} - \gamma_n \le u_j \le (z_j+1)t_{n,k,\lambda} - 1\Big\}$$

for all $j = 1,\dots,d$.

Lemma D.3.2. *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of sets satisfying Assumption D.2.4, and let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary field satisfying $\mathcal{D}(x_n;K_n)$. Then*

$$\begin{aligned}&\Big|\mathbb{P}(M(D_{n,k,\lambda}^-) \le x_n) - \mathbb{P}^{p_{n,k,\lambda}}(M(J_0) \le x_n)\Big| \\ &\le 2p_{n,k,\lambda}\mathbb{P}(M(H_0) \le x_n < M(H_0^*)) + (p_{n,k,\lambda}-1)\alpha_n,\end{aligned} \tag{D.3.2}$$

and similarly

$$\begin{aligned}&\Big|\mathbb{P}(M(D_{n,k,\lambda}^+) \le x_n) - \mathbb{P}^{q_{n,k,\lambda}}(M(J_0) \le x_n)\Big| \\ &\le 2q_{n,k,\lambda}\mathbb{P}(M(H_0) \le x_n < M(H_0^*)) + (q_{n,k,\lambda}-1)\alpha_n.\end{aligned} \tag{D.3.3}$$

Proof. Since H_z is a subset of J_z, it is easily seen that

$$\begin{aligned}0 &\le \mathbb{P}\Big(\bigcap_{z\in P_{n,k,\lambda}} \{M(H_z) \le x_n\}\Big) - \mathbb{P}(M(D_{n,k,\lambda}^-) \le x_n) \\ &\le \mathbb{P}\Big(\bigcup_{z\in P_{n,k,\lambda}} \{M(H_z) \le x_n < M(H_z^*)\}\Big) \\ &\le p_{n,k,\lambda}\mathbb{P}(M(H_0) \le x_n < M(H_0^*))\end{aligned}$$

by stationarity. Turning to Lemma D.3.1 and using stationarity again show that

$$\Big|\mathbb{P}\Big(\bigcap_{z\in P_{n,k,\lambda}} \{M(H_z) \le x_n\}\Big) - \mathbb{P}^{p_{n,k,\lambda}}(M(H_0) \le x_n)\Big| \le (p_{n,k,\lambda}-1)\alpha_n,$$

and realizing that

$$\begin{aligned}0 &\le \mathbb{P}^{p_{n,k,\lambda}}(M(H_0) \le x_n) - \mathbb{P}^{p_{n,k,\lambda}}(M(J_0) \le x_n) \\ &\le p_{n,k,\lambda}\Big(\mathbb{P}(M(H_0) \le x_n) - \mathbb{P}(M(J_0) \le x_n)\Big) \\ &= p_{n,k,\lambda}\mathbb{P}(M(H_0) \le x_n < M(H_0^*))\end{aligned}$$

concludes (D.3.2). The claim (D.3.3) follows similarly. □

Lemma D.3.3. *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of sets satisfying Assumption D.2.4, and let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary field satisfying $\mathcal{D}(x_n;K_n)$. Then*

$$\Big|\mathbb{P}(M(D_n) \le x_n) - \mathbb{P}^{\lambda k}(M(J_0^{n,k,\lambda}) \le x_n)\Big| \le R_{n,k,\lambda}, \tag{D.3.4}$$

where $R_{n,k,\lambda}$ satisfies

$$\limsup_{n\to\infty} R_{n,k,\lambda} = o(1) \tag{D.3.5}$$

as $k \to \infty$.

Proof. Let $R^p_{n,k,\lambda} \geq 0$ and $R^q_{n,k,\lambda} \geq 0$ denote the upper bounds in (D.3.2) and (D.3.3), respectively. First we show that

$$\mathbb{P}(M(H_0) \leq x_n < M(H_0^*)) \to 0 \tag{D.3.6}$$

as $n \to \infty$. Fix an integer r such that $t_{n,k,\lambda} \geq (2r+1)\gamma_n$. Then, for all $j = 1,\dots,d$, we can write $L^*_{0,j}$ as the union of $2(d-1)$ overlapping boxes each with side-lengths $\gamma_n, (t_{n,k,\lambda} - \gamma_n), \dots, (t_{n,k,\lambda} - \gamma_n)$, where we need one box (interval) if $d = 1$. Now let E_j^* be such a box. Since $t_{n,k,\lambda} \geq (2r+1)\gamma_n$, we can construct boxes $E_1, \dots, E_r$ contained in H_0 which are simply translations of E_j^*, such that they are separated from E_j^* and each other by γ_n. Hence, by Lemma D.3.1,

$$\begin{aligned}
&\mathbb{P}(M(H_0) \leq x_n < M(E_j^*)) \\
&\quad \leq \mathbb{P}\Big(\bigcap_{s=1}^r \{M(E_s) \leq x_n < M(E_j^*)\}\Big) \\
&\quad = \mathbb{P}\Big(\bigcap_{s=1}^r \{M(E_s) \leq x_n\}\Big) - \mathbb{P}\Big(\bigcap_{s=1}^r \{M(E_s) \leq x_n\} \cap \{M(E_j^*) \leq x_n\}\Big) \\
&\quad \leq x^r - x^{r+1} + 2r\alpha_n,
\end{aligned}$$

where $x = \mathbb{P}(M(E_s) \leq x_n) = \mathbb{P}(M(E_j^*) \leq x_n)$ for all s by stationarity. For $x \in [0,1]$ the mapping $x \mapsto x^r - x^{r+1}$ is bounded by $1/r$, and we conclude that

$$\mathbb{P}(M(H_0) \leq x_n < M(E_j^*)) \leq \frac{1}{r} + 2r\alpha_n.$$

This can be done for all the $2(d-1)$ sub-boxes of $L^*_{0,j}$ for all $j = 1,\dots,d$. Since $H_0^* = \cup_{j=1}^d L^*_{1,j}$, we therefore find that

$$\mathbb{P}(M(H_0) \leq x_n < M(H_0^*)) \leq 2d(d-1)\Big(\frac{1}{r} + 2r\alpha_n\Big).$$

Letting $n \to \infty$ and then $r \to \infty$ show that (D.3.6) is satisfied, which implies that also

$$\lim_{n\to\infty} R^p_{n,k,\lambda} = \lim_{n\to\infty} R^q_{n,k,\lambda} = 0. \tag{D.3.7}$$

Now define $\tilde{R}^p_{n,k,\lambda} \geq 0$ and $\tilde{R}^q_{n,k,\lambda} \geq 0$ by

$$\begin{aligned}
\tilde{R}^p_{n,k,\lambda} &= R^p_{n,k,\lambda} + \mathbb{P}^{p_{n,k,\lambda}}(M(J_0) \leq x_n) - \mathbb{P}^{\lambda k}(M(J_0) \leq x_n), \qquad \text{and} \\
\tilde{R}^q_{n,k,\lambda} &= R^q_{n,k,\lambda} + \mathbb{P}^{\lambda k}(M(J_0) \leq x_n) - \mathbb{P}^{q_{n,k,\lambda}}(M(J_0) \leq x_n).
\end{aligned}$$

Since $D^-_{n,k,\lambda} \subseteq D_n \subseteq D^+_{n,k,\lambda}$, it is seen that

$$\begin{aligned}
\mathbb{P}^{\lambda k}(M(J_0) \leq x_n) - \tilde{R}^q_{n,k,\lambda} &\leq \mathbb{P}(M(D^+_{n,k,\lambda}) \leq x_n) \\
&\leq \mathbb{P}(M(D_n) \leq x_n) \\
&\leq \mathbb{P}(M(D^-_{n,k,\lambda}) \leq x_n) \leq \mathbb{P}^{\lambda k}(M(J_0) \leq x_n) + \tilde{R}^p_{n,k,\lambda},
\end{aligned}$$

and defining $R_{n,k,\lambda} = \max\{\tilde{R}^p_{n,k,\lambda}, \tilde{R}^q_{n,k,\lambda}\}$ then shows (D.3.4).

It remains to show that (D.3.5) is satisfied, which follows if both $\limsup_{n\to\infty} \tilde{R}^p_{n,k,\lambda}$ and $\limsup_{n\to\infty} \tilde{R}^q_{n,k,\lambda}$ are of order $o(1)$ as $k \to \infty$. We only show this for $\tilde{R}^p_{n,k,\lambda}$ as the behavior for $\tilde{R}^q_{n,k,\lambda}$ follows similarly. First, since $p_{n,k,\lambda} \le \lambda k$, the mapping $x \mapsto x^{p_{n,k,\lambda}} - x^{\lambda k}$ is bounded by $(1 - \frac{p_{n,k,\lambda}}{\lambda k})$ for all $x \in [0,1]$. Hence

$$\mathbb{P}^{p_{n,k,\lambda}}(M(J_0) \le x_n) - \mathbb{P}^{\lambda k}(M(J_0) \le x_n) \le 1 - \frac{p_{n,k,\lambda}}{\lambda k},$$

and, using (D.3.7) and (ii) in Theorem D.2.5, we conclude that

$$0 \le \limsup_{n\to\infty} \tilde{R}^p_{n,k,\lambda} \le 1 - \frac{\liminf_{n\to\infty} p_{n,k,\lambda}}{\lambda k} \to 0$$

as $k \to \infty$. Equation (D.3.5) is thus satisfied. □

The next theorem shows that under Assumption D.2.4 and $\mathcal{D}(x_n; K_n)$, and when normalized correctly, the limiting distribution of $M(D_n)$ is necessarily an extreme value distribution; see [11, Chapter 1] for a thorough exposition of extreme value distributions and their connection to max–stable distributions.

Theorem D.3.4. *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of sets satisfying Assumption D.2.4, and let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary field. Assume that there are sequences $(a_n)_{n\in\mathbb{N}}$ with $a_n > 0$ and $(b_n)_{n\in\mathbb{N}}$ such that $\mathcal{D}(a_n x + b_n; K_n)$ is satisfied for all $x \in \mathbb{R}$. Assume furthermore that there exists a constant $a > 0$ such that*

$$|D_n| \sim a \cdot n. \tag{D.3.8}$$

If there exists a non–degenerate distribution function G such that for all $x \in \mathbb{R}$

$$P(M(D_n) \le a_n x + b_n) \to G(x), \tag{D.3.9}$$

then G is the distribution function of an extreme value distribution.

Note that (D.3.8) is in particular satisfied if $C_n = r_n C$ for a p–convex set C and a sequence (r_n) with $r_n \sim c\sqrt[d]{n}$, cf. Example D.2.6 and (i) in Theorem D.2.5.

Proof. According to [11, Theorem 1.3.1] it suffices to show that there exists a sequence of distribution functions $(F_n)_{n\in\mathbb{N}}$ such that for all $x \in \mathbb{R}$ and all $k \in \mathbb{N}$

$$F_n(a_{nk} x + b_{nk}) \to G^{1/k}(x) \tag{D.3.10}$$

as $n \to \infty$. To obtain this, we first relate the limiting distribution G to the distribution of $M(\tilde{D}^{k_0}_n)$, where $(\tilde{D}^{k_0}_n)_{n\in\mathbb{N}}$ is a sequence of discrete cubes defined as follows: Choose $k_0 \in \mathbb{N}$ such that $2/\sqrt[d]{\lambda k_0} \le c$ for all $\lambda \in (1-\epsilon, 1+\epsilon)$, with $\epsilon > 0$ and c defined in Theorem D.2.5, and define

$$\tilde{C}^{k_0,\lambda}_n = \left[0, \sqrt[d]{|C_n|/(\lambda k_0)}\right]^d \qquad \text{and} \qquad \tilde{D}^{k_0,\lambda}_n = \tilde{C}^{k_0,\lambda}_n \cap \mathbb{Z}^d,$$

which are easily seen to satisfy Assumption D.2.4. Note in addition that $2\tilde{D}^{k_0,\lambda}_n \subseteq K_n$ and $|\tilde{C}^{k_0,\lambda}_n| = \frac{|C_n|}{\lambda k_0}$ for all n. Now, with a notation very similar to the notation from Section D.2, we define for each $k \in \mathbb{N}$

$$\tilde{P}^{k_0,\lambda}_{n,k} = \{z \in \mathbb{Z}^d \,:\, I^{n,k_0\cdot k,\lambda}_z \subseteq \tilde{C}^{k_0,\lambda}_n\} \qquad \text{and}$$
$$\tilde{Q}^{k_0,\lambda}_{n,k} = \{z \in \mathbb{Z}^d \,:\, I^{n,k_0\cdot k,\lambda}_z \cap \tilde{C}^{k_0,\lambda}_n \ne \emptyset\}.$$

With the same arguments as before and in the proof of Theorem D.2.5 it is easily seen that

$$|\tilde{P}_{n,k}^{k_0,\lambda}| \le k \le |\tilde{Q}_{n,k}^{k_0,\lambda}|$$

and

$$\liminf_{n\to\infty} \left|\tilde{P}_{n,k}^{k_0,\lambda}\right| \sim k \qquad \text{resp.} \qquad \limsup_{n\to\infty} \left|\tilde{Q}_{n,k}^{k_0,\lambda}\right| \sim k$$

as $k \to \infty$. Using that $\tilde{D}_n^{k_0,\lambda}$ is approximated by roughly k of the cubes $J_z^{n,k_0\cdot k,\lambda}$ in exactly the same way as D_n can be approximated by roughly k of the cubes $J_z^{n,k,1}$, we find from Lemma D.3.3 that

$$\left|\mathbb{P}\Big(M(\tilde{D}_n^{k_0,\lambda}) \le a_n x + b_n\Big) - \mathbb{P}^k\Big(M(J_0^{n,k_0\cdot k,\lambda}) \le a_n x + b_n\Big)\right| \le \tilde{R}_{n,k,\lambda},$$

where $\tilde{R}_{n,k,\lambda}$ satisfies $\limsup_{n\to\infty} \tilde{R}_{n,k,\lambda} = o(1)$ as $k \to \infty$. Using that $x \mapsto x^{\lambda k_0}$ is Lipschitz continuous when restricted to $[0,1]$, we find that also

$$\left|\mathbb{P}^{\lambda k_0}\Big(M(\tilde{D}_n^{k_0,\lambda}) \le a_n x + b_n\Big) - \mathbb{P}^{\lambda k_0\cdot k}\Big(M(J_0^{n,k_0\cdot k,\lambda}) \le a_n x + b_n\Big)\right| \le \tilde{R}'_{n,k,\lambda},$$

where $\limsup_{n\to\infty} \tilde{R}'_{n,k,\lambda} = o(1)$ as $k \to \infty$. Since we already have from Lemma D.3.3 that

$$\left|\mathbb{P}\Big(M(D_n) \le a_n x + b_n\Big) - \mathbb{P}^{\lambda k_0\cdot k}\Big(M(J_0^{n,k_0\cdot k,\lambda}) \le a_n x + b_n\Big)\right| \le R_{n,k,\lambda},$$

we can, when combining with (D.3.9), conclude that for all $x \in \mathbb{R}$

$$\mathbb{P}^{\lambda k_0}\Big(M(\tilde{D}_n^{k_0,\lambda}) \le a_n x + b_n\Big) \to G(x)$$

as $n \to \infty$, and thereby also

$$\mathbb{P}^{k_0}\Big(M(\tilde{D}_n^{k_0,\lambda}) \le a_n x + b_n\Big) \to G^{1/\lambda}(x).$$

Now define

$$\overline{C}_n^{k_0} = \left[0, \sqrt[d]{a\cdot n/k_0}\right]^d \qquad \text{and} \qquad \overline{D}_n^{k_0} = \overline{C}_n^{k_0} \cap \mathbb{Z}^d.$$

By (D.3.8) and (i) in Theorem D.2.5 we find for all $\lambda^- < 1 < \lambda^+$ that

$$\begin{aligned}\mathbb{P}\Big(M(\tilde{D}_n^{k_0,\lambda^-}) \le a_n x + b_n\Big) &\le \mathbb{P}\Big(M(\overline{D}_n^{k_0}) \le a_n x + b_n\Big)\\ &\le \mathbb{P}\Big(M(\tilde{D}_n^{k_0,\lambda^+}) \le a_n x + b_n\Big)\end{aligned}$$

for n large enough. Letting $\lambda^- \uparrow 1$ and $\lambda^+ \downarrow 1$ this gives

$$\mathbb{P}^{k_0}\Big(M(\overline{D}_n^{k_0}) \le a_n x + b_n\Big) \to G(x) \tag{D.3.11}$$

as $n \to \infty$. Let $k \in \mathbb{N}$ and replace k_0 by $k_0\cdot k$ and n by $n\cdot k$ in (D.3.11). Since furthermore

$$\overline{D}_{n\cdot k}^{k_0\cdot k} = \overline{D}_n^{k_0}$$

by construction, we find that for all $k \in \mathbb{N}$ and $x \in \mathbb{R}$ the convergence

$$\mathbb{P}^{k_0}\Big(M(\overline{D}_n^{k_0}) \le a_{nk} x + b_{nk}\Big) \to G^{1/k}(x)$$

as $n \to \infty$ is satisfied. This shows (D.3.10) as desired. □

For the remainder of this paper we let $\lambda = 1$ and suppress it from the notation. So far, assuming $\mathcal{D}(x_n;K_n)$ has made it possible to relate the limiting distribution of $M(D_n)$ to the distribution of e.g. $M(J_0^{n,k})$. However, to be able to relate this to the distribution of the individual ξ–variables, it is necessary to make the following further assumption.

Condition ($\mathcal{D}'(x_n)$). *The condition $\mathcal{D}'(x_n)$ is satisfied for the stationary field $(\xi_v)_{v\in\mathbb{Z}^d}$ if*

$$S_{n,k} = \limsup_{n\to\infty} t_{n,k}^d \sum_{0\neq v\in J_0^{n,k}} \mathbb{P}(\xi_0 > x_n, \xi_v > x_n) = o(k^{-1})$$

as $k \to \infty$.

Theorem D.3.5. *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of sets satisfying Assumption D.2.4, and let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary field satisfying $\mathcal{D}(x_n;K_n)$ and $\mathcal{D}'(x_n)$. Let $0 \le \tau < \infty$. Then, for any $v \in \mathbb{Z}^d$,*

$$|D_n|\mathbb{P}(\xi_v > x_n) \to \tau \qquad \text{as } n \to \infty, \tag{D.3.12}$$

if and only if

$$\mathbb{P}(\max_{v\in D_n} \xi_v \le x_n) \to \mathrm{e}^{-\tau} \qquad \text{as } n \to \infty. \tag{D.3.13}$$

Proof. Let F_ξ denote the common distribution function of $(\xi_v)_{v\in\mathbb{Z}^d}$, i.e.

$$F_\xi(x) = 1 - \overline{F}_\xi(x) = \mathbb{P}(\xi_v \le x), \qquad \text{for all} \quad v \in \mathbb{Z}^d.$$

Writing $J_0 = J_0^{n,k}$, it is not difficult to see that

$$\begin{aligned}\sum_{v\in J_0} \mathbb{P}(\xi_v > x_n) - \sum_{v<v'\in J_0} &\mathbb{P}(\xi_v > x_n, \xi_{v'} > x_n) \\ &\le \mathbb{P}(M(J_0) > x_n) \\ &\le \sum_{v\in J_0} \mathbb{P}(\xi_v > x_n),\end{aligned}$$

where the summation over $\{v < v' \in J_0\}$ indicates the double sum of points in $v \in J_0$ and subsequent points $v' \in J_0$ falling strictly after v under some underlying enumeration. This notation will be used in the following sections as well. By stationarity of (ξ_v) the above implies that

$$t_{n,k}^d \overline{F}_\xi(x_n) - S_{n,k} \le P(M(J_0) > x_n) \le t_{n,k}^d \overline{F}_\xi(x_n),$$

where $S_{n,k}$ satisfies $\limsup_n S_{n,k} = o(k^{-1})$ as $k \to \infty$ according to $\mathcal{D}'(x_n)$. Appealing to Lemma D.3.3 this implies that

$$\begin{aligned}(1 - t_{n,k}^d &\overline{F}_\xi(x_n))^k - R_{n,k} \\ &\le P(M(D_n) \le x_n) \\ &\le (1 - t_{n,k}^d \overline{F}_\xi(x_n) + S_{n,k})^k + R_{n,k}.\end{aligned} \tag{D.3.14}$$

Now assume that (D.3.12) is satisfied, that is, $|D_n|\overline{F}_\xi(x_n) \to \tau$ and equivalently $t_{n,k}^d \overline{F}_\xi(x_n) \to \tau/k$. Then (D.3.5) and (D.3.14) imply that

$$\begin{aligned}\left(1-\frac{\tau}{k}\right)^k + o(1) &\le \liminf_{n\to\infty} \mathbb{P}(M(D_n) \le x_n) \\ &\le \limsup_{n\to\infty} \mathbb{P}(M(D_n) \le x_n) \\ &\le \left(1-\frac{\tau}{k}+o(k^{-1})\right)^k + o(1).\end{aligned}$$

Taking the limit $k \to \infty$ and using the equivalence $\log(1-y) \sim -y$ $(y \to 0)$ show that $\lim_{n\to\infty} \mathbb{P}(M(D_n) \le x_n) = \mathrm{e}^{-\tau}$.

Now assume that (D.3.13) is satisfied, i.e. $\mathbb{P}(M(D_n) \le x_n) \to \mathrm{e}^{-\tau}$. Using this assumption, (D.3.5) and (D.3.14) imply that

$$\begin{aligned}1-\left(\mathrm{e}^{-\tau}+o(1)\right)^{1/k} &\le k^{-1} \liminf_{n\to\infty} |D_n|\overline{F}_\xi(x_n) \\ &\le k^{-1} \limsup_{n\to\infty} |D_n|\overline{F}_\xi(x_n) \\ &\le 1-\left(\mathrm{e}^{-\tau}+o(1)\right)^{1/k} + o(k^{-1}).\end{aligned}$$

Multiplying by k and taking the limit $k \to \infty$ show that $\lim_{n\to\infty}|D_n|\overline{F}_\xi(x_n) = \tau$. □

The following corollary follows exactly as [11, Corollary 3.4.2].

Corollary D.3.6. *For $\tau = \infty$, the conclusions of Theorem D.3.5 hold if the conditions $\mathcal{D}(x_n;K_n)$ and $\mathcal{D}'(x_n)$ are replaced by the following: For all $\tau' < \infty$ there is a real sequence $(s_n)_{n\in\mathbb{N}}$ such that $|D_n|\mathbb{P}(\xi_v > s_n) \to \tau'$ and such that $\mathcal{D}(s_n;K_n)$ and $\mathcal{D}'(s_n)$ are satisfied.*

Below we provide an example to which the results can be used, namely that of a stationary Gaussian field, i.e. a field such that all finite dimensional distributions are multivariate Gaussian.

Corollary D.3.7. *Let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary Gaussian field with correlation function r_v, $v \in \mathbb{Z}^d$, satisfying*

$$\log(m) \sup_{|v|>m} |r_v| \to 0 \tag{D.3.15}$$

as $m \to \infty$. Furthermore, let $(D_n)_n$ be a sequence of sets satisfying Assumption D.2.4. Then, for all $0 \le \tau \le \infty$,

$$|D_n|\mathbb{P}(\xi_v > x_n) \to \tau \qquad \text{if and only if} \qquad \mathbb{P}(\max_{v\in D_n} \xi_v \le x_n) \to \mathrm{e}^{-\tau}$$

as $n \to \infty$.

Proof. Without loss of generality we assume that (ξ_v) is a standard Gaussian field. First we note that

$$\sup_{v\neq 0} |r_v| < 1, \tag{D.3.16}$$

which is seen by the following considerations: From the assumption (D.3.15) we have

$$\sup_{|v|>m} |r_v| \to 0 \tag{D.3.17}$$

as $m \to \infty$. Now assume that $|r_v| = 1$ for some $0 \neq v \in \mathbb{Z}^d$. Then, by the Cauchy-Schwarz inequality, $\xi_0 = \pm\xi_{kv}$ almost surely for all $k \in \mathbb{N}$, and consequently $|r_{kv}| = 1$ for all $k \in \mathbb{N}$ contradicting (D.3.17). Hence, $|r_v| < 1$ for all $0 \neq v \in \mathbb{Z}^d$, which, by (D.3.17), implies (D.3.16). For all $x \in \mathbb{R}$ and subsets $A \subseteq K_n$, a trivial generalization of [11, Corollary 4.2.4] now gives

$$\Big|\mathbb{P}\Big(\max_{v\in A} \xi_v \leq x\Big) - \Phi(x)^{|A|}\Big| \leq K|K_n| \sum_{0\neq v\in K_n} |r_v| \exp\Big(-\frac{x^2}{1+|r_v|}\Big), \tag{D.3.18}$$

where K is an appropriate constant, and Φ denotes the standard normal distribution function, i.e. the distribution of ξ_v. If furthermore $|K_n|(1-\Phi(x_n))$ is bounded, the right-hand side of (D.3.18) tends to 0 as $n \to \infty$: Let $\delta = \sup_{v\neq 0}|r_v|$ and choose $0 < \alpha < (1-\delta)/(1+\delta)$. Splitting the sum into the two parts for which $|v| \leq |K_n|^{\alpha/d}$ and $|v| > |K_n|^{\alpha/d}$, respectively, the result follows by the same arguments as in the proof of [11, Lemma 4.3.2], realizing that

$$\log|K_n| \sup_{|v|>|K_n|^{\alpha/d}} |r_v| = \tfrac{d}{\alpha} \log|K_n|^{\alpha/d} \sup_{|v|>|K_n|^{\alpha/d}} |r_v| \to 0$$

as $n \to \infty$ by (D.3.15). It is now not difficult to see that $\mathcal{D}(x_n; K_n)$ and $\mathcal{D}'(x_n)$ are satisfied.

Suppose that $|K_n|(1-\Phi(x_n))$ is bounded, and hence $\mathcal{D}(x_n; K_n)$ and $\mathcal{D}'(x_n)$ are satisfied. Then the claim follows for all $0 \leq \tau < \infty$ from Theorem D.3.5. Now suppose that $|K_n|(1-\Phi(x_n))$ is unbounded. Let $\tau' < \infty$ and define the sequence (s_n) such that $|K_n|(1-\Phi(s_n)) = \tau'$. Then, from the considerations above, the conditions $\mathcal{D}(s_n; K_n)$ and $\mathcal{D}'(s_n)$ are satisfied, and the claim follows for $\tau = \infty$ from Corollary D.3.6. □

Theorem D.3.8. *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of sets satisfying Assumption D.2.4, and let $(\xi_v)_{v\in\mathbb{Z}^d}$ be a stationary field. Assume that there are sequences $(a_n)_{n\in\mathbb{N}}$ with $a_n > 0$ and $(b_n)_{n\in\mathbb{N}}$ such that $\mathcal{D}(a_n x + b_n; K_n)$ and $\mathcal{D}'(a_n x + b_n)$ are satisfied for all $x \in \mathbb{R}$. Assume furthermore that $|D_{n+1}|/|D_n| \to 1$ as $n \to \infty$. If there exists a non–degenerate distribution function G such that for all $x \in \mathbb{R}$*

$$\mathbb{P}\Big(M(D_n) \leq a_n x + b_n\Big) \to G(x),$$

then G is the distribution function of an extreme value distribution.

The assumption $|D_{n+1}|/|D_n| \to 1$ ensures that the growth of the (D_n) is not too explosive. It will in particular be satisfied under the assumption given by (D.3.8). It will also be fulfilled if e.g. $C_n = r_n C$ for a p–convex set C and a sequence (r_n), where $r_{n+1}/r_n \to 1$.

Proof. Let again F_ξ denote the common distribution function of $(\xi_v)_{v\in\mathbb{Z}^d}$. By Theorem D.3.5 we have

$$|D_n|\overline{F}_\xi(a_n x + b_n) \to -\log G(x) \tag{D.3.19}$$

for all $x \in \mathbb{R}$. Now define the two sequences $(k_m)_{m\in\mathbb{N}}$ and $(\ell_m)_{m\in\mathbb{N}}$ by

$$k_m = \max\{|D_n| : |D_n| < m\} \qquad \text{and} \qquad \ell_m = \min\{|D_n| : |D_n| \geq m\}.$$

Then $k_m < m \le \ell_m$ and by assumption $k_m/\ell_m \to 1$ as $m \to \infty$. We find

$$\frac{k_m}{\ell_m}\ell_m \overline{F}_\xi(a_{\ell_m}x + b_{\ell_m}) < m\overline{F}_\xi(a_{\ell_m}x + b_{\ell_m}) \le \ell_m \overline{F}_\xi(a_{\ell_m}x + b_{\ell_m}).$$

Define $a'_m = a_{\ell_m}$ and $b'_m = b_{\ell_m}$ and let $m \to \infty$. Using the limit (D.3.19) then gives

$$m\overline{F}_\xi(a'_m x + b'_m) \to -\log G(x)$$

for all $x \in \mathbb{R}$. Let $(\xi'_m)_{m\in\mathbb{N}}$ be an independent and identically distributed sequence with common distribution function F_ξ and define $M_m = \max\{\xi'_1, \ldots, \xi'_m\}$. Then, by [11, Theorem 1.5.1] (which is in fact Theorem D.1.1 from the introduction), we find

$$\mathbb{P}\Big(M_m \le a'_m x + b'_m\Big) \to G(x)$$

for all $x \in \mathbb{R}$, showing by the classical Extremal Types Theorem [11, Theorem 1.4.2] that G is indeed an extreme value distribution. □

D.4 Extreme theorem for stationary Lévy-driven random fields

In this section we consider a stationary random field $(X_v)_{v\in\mathbb{R}^d}$, given as an integral of a kernel function with respect to a Lévy basis, and we wish to characterize the tail behavior of $\sup_{v\in C_n} X_v$, where $(C_n)_{n\in\mathbb{N}}$ is a sequence of index sets in $\mathbb{R}^d$ satisfying Assumption D.2.4.

We define a Lévy basis to be an infinitely divisible and independently scattered random measure. The random measure M on $\mathbb{R}^d$ is independently scattered if for all disjoint Borelsets $(A_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ the random variables $(M(A_n))_{n\in\mathbb{N}}$ are independent and furthermore satisfy $M(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} M(A_n)$. The random measure M is infinitely divisible if $M(A)$ is infinitely divisible for all Borelsets $A \subseteq \mathbb{R}^d$.

Moreover, in this paper we assume M to be a stationary and isotropic Lévy basis on $\mathbb{R}^d$. With $C(\lambda \dagger Y) = \log \mathbb{E}e^{i\lambda Y}$ denoting the cumulant function for a random variable Y, this means that the random variable $M(A)$ has Lévy-Khintchine representation

$$C(\lambda \dagger M(A)) = i\lambda a|A| - \frac{1}{2}\lambda^2\theta|A| + \int_{A\times\mathbb{R}} \Big(e^{i\lambda x} - 1 - i\lambda x \mathbf{1}_{[-1,1]}(x)\Big) F(\mathrm{d}u, \mathrm{d}x).$$

Here $a \in \mathbb{R}$, $\theta \ge 0$ and F is the product measure $m \otimes \rho$ of the Lebesgue measure m and a Lévy measure ρ.

We assume that the Lévy measure ρ is convolution equivalent with index $\beta > 0$, and we write $\rho \in \mathcal{S}_\beta$. Convolution equivalence is a property of the right tail of a finite measure, and thus we may equivalently define it through the restriction ρ_1 of ρ to the set $(1,\infty)$. Note that $\mathcal{S}_\beta$ usually denotes the class of convolution equivalent distributions, however, we say that ρ is in $\mathcal{S}_\beta$ if its $(1,\infty)$-restriction $\rho_1 \in \mathcal{L}_\beta$, the class of finite measures with an exponential right tail with index β, i.e.

$$\frac{\rho_1((x-y,\infty))}{\rho_1((x,\infty))} \to e^{\beta y} \qquad \text{as } x \to \infty, \tag{D.4.1}$$

for all $y \in \mathbb{R}$, and if it furthermore satisfies the convolution property

$$\frac{(\tilde{\rho}_1 * \tilde{\rho}_1)((x,\infty))}{\tilde{\rho}_1((x,\infty))} \to 2\int_{\mathbb{R}} e^{\beta y}\tilde{\rho}_1(\mathrm{d}y) < \infty \qquad \text{as } x \to \infty,$$

where $*$ denotes convolution and $\tilde{\rho}_1$ is the normalization of ρ_1. Moreover, we note that ρ_1 (equivalently ρ) lies in the *maximum domain of attraction* MDA(Λ) of the Gumbel distribution, by which we mean that there are norming constants $\tilde{a}_n > 0$ and $\tilde{b}_n \in \mathbb{R}$ such that

$$n\tilde{\rho}_1((\tilde{a}_n x + \tilde{b}_n, \infty)) \to \mathrm{e}^{-x}, \qquad (n \to \infty),$$

for all $x \in \mathbb{R}$. This is seen by [5, Theorem 3.3.27] and (D.4.1) choosing the function $\tilde{a}(\,\cdot\,)$ (as in the formulation of the theorem) constantly equal to $1/\beta$. The norming constants can be seen to satisfy

$$\tilde{a}_n \to \beta^{-1}, \qquad \text{and} \qquad \tilde{b}_n \to \infty$$

as $n \to \infty$, where we refer to [5, Chapter 3] for a description on the extreme value distributions, their maximum domains of attraction, and the associated norming constants. For convenience, we collect the assumptions on M and ρ in the following. Assumption D.4.1 is assumed satisfied in the remainder of this paper.

Assumption D.4.1. *The Lévy basis M on $\mathbb{R}^d$ is stationary and isotropic with a Lévy measure $\rho \in \mathcal{S}_\beta$ where $\beta > 0$. Moreover, ρ satisfies*

$$\int_{|y|>1} |y|^k \rho(\mathrm{d}y) < \infty \qquad \forall k \in \mathbb{N}. \tag{D.4.2}$$

The integrability of ρ along its right tail is already given from the fact that $\rho \in \mathcal{S}_\beta$, and, since ρ is a Lévy measure, it also satisfies $\int_{[-1,1]} y^2 \rho(\mathrm{d}y) < \infty$. Moreover, by [20, Theorem 25.3], the integrability $\mathbb{E}|M(A)|^k < \infty$ with $|A| > 0$ is equivalent to (D.4.2).

We consider the Lévy-driven field $X = (X_v)_{v\in\mathbb{R}^d}$ defined by

$$X_v = \int_{\mathbb{R}^d} f(|v-u|)M(\mathrm{d}u), \tag{D.4.3}$$

which, by [14, Theorem 2.7], is well-defined if only the Lévy measure satisfies (D.4.2) for $k = 1$, and if the integration kernel $f : [0,\infty) \to [0,\infty)$ is bounded and satisfies $\int_{\mathbb{R}^d} f(|u|)\mathrm{d}u < \infty$. However, below we make a set of stronger assumptions on f which are assumed satisfied throughout this paper. Combined with Assumption D.4.1, these guarantee the existence of a continuous version of $(X_v)_{v\in\mathbb{R}^d}$, and furthermore they give a sufficient mixing structure of the field.

Assumption D.4.2. *The integration kernel $f : [0,\infty) \to [0,\infty)$ satisfying*

$$f(0) = 1, \qquad f(x) < 1 \text{ for } x > 0, \tag{D.4.4}$$

is bounded from above by a decreasing function g such that

$$\int_{\mathbb{R}^d} g(|u|)\mathrm{d}u < \infty. \tag{D.4.5}$$

Moreover, the kernel f is Lipschitz continuous, that is, there is a constant C_L such that

$$|f(x_1) - f(x_2)| \le C_L |x_1 - x_2|$$

for all $x_1, x_2 \ge 0$.

Furthermore, if $d \geq 2$, there is a sequence $\gamma_n = o(\sqrt[d]{|C_n|})$ such that, as $n \to \infty$,

$$|C_n| g(\gamma_n/2) \to 0 \qquad \textit{and} \qquad |C_n|^{(d-1)/d} \int_{\gamma_n/2}^{\infty} g(x)\mathrm{d}x \to 0 \tag{D.4.6}$$

as $n \to \infty$.

We remark that the integrability of g and the fact that it is decreasing in particular implies that

$$\int_{\mathbb{R}^d} \sup_{v \in [0,1]^d} g(|v-u|)\mathrm{d}u < \infty.$$

This will be used when referring to the results of [16] below.

Example D.4.3. If f is a Lipschitz continuous integration kernel satisfying (D.4.4), and if there exist $c \in \mathbb{R}$ and $\epsilon > 0$ such that

$$f(x) \leq g(x) = \frac{c}{(1+x)^{d+\epsilon}} \qquad \text{for all } x \in [0,\infty),$$

then (D.4.5) is satisfied. Defining $\gamma_n = \sqrt[d+\epsilon/2]{|C_n|}$ it can be seen that also (D.4.6) is satisfied, and thus Assumption D.4.2 holds.

We write $M = M_1 + M_2$ as the independent sum of two Lévy bases with Lévy–Khintchine representations

$$C(\lambda \dagger M_1(A)) = \int_{A\times\mathbb{R}} \left(\mathrm{e}^{i\lambda x} - 1\right) F_1(\mathrm{d}u, \mathrm{d}x)$$
$$C(\lambda \dagger M_2(A)), = i\lambda a|A| - \frac{1}{2}\lambda^2\theta|A| + \int_{A\times\mathbb{R}} \left(\mathrm{e}^{i\lambda x} - 1 - i\lambda x \mathbf{1}_{[-1,1]}(x)\right) F_2(\mathrm{d}u, \mathrm{d}x),$$

respectively. Here, $F_i = m \otimes \rho_i$, where ρ_1 and ρ_2 are the restrictions of ρ to $(1,\infty)$ and $(-\infty, 1]$, respectively, and m is the Lebesgue measure. Similarly, we decompose the field $(X_v)_v$ into a sum of two independent random fields $X_v = Z_v + Y_v$ for all $v \in \mathbb{R}^d$, where

$$Z_v = \int_{\mathbb{R}^d} f(|v-u|) M_1(\mathrm{d}u)$$

and

$$Y_v = \int_{\mathbb{R}^d} f(|v-u|) M_2(\mathrm{d}u) \tag{D.4.7}$$

are both stationary. By (Theorem C.5.2, Paper C) all three fields have continuous versions on compact sets.

In [16, Theorem 4.2] an equivalence between the tail of the supremum of a field defined as (D.4.3) and its associated convolution equivalent Lévy measure is given. The result relies on a set of assumptions on the Lévy basis and the kernel function, which differ slightly from Assumptions D.4.1 and D.4.2 given here. However, by (Theorem C.5.2, Paper C) and Assumptions D.4.1 and D.4.2, the field $(X_v)_{v\in\mathbb{R}^d}$ is continuous on compact sets, and it can easily be seen that the results of [16] are valid. In particular, if $C \subseteq [0,1]^d$ (or a translation hereof) then

$$\frac{\mathbb{P}(\sup_{v\in C} X_v > x)}{\rho((x,\infty))} \to |C| \mathbb{E}\mathrm{e}^{\beta X_u}, \qquad (x \to \infty), \tag{D.4.8}$$

and

$$\frac{\mathbb{P}(\sup_{v\in C} Z_v > x)}{\rho((x,\infty))} \to |C|\mathbb{E}e^{\beta Z_u}, \qquad (x\to\infty), \tag{D.4.9}$$

where $u \in \mathbb{R}^d$ is arbitrarily chosen.

In our arguments below we need to find a lower bound of Z_v, which (as a field) is approximately independent for large lag, and which has essentially the same limiting behavior as Z_v. Since $f \geq 0$ and $M_1 \geq 0$, the field $Z_v^{(t)}$ defined by

$$Z_v^{(t)} = \int_{\{|v-u|\leq t\}} f(|v-u|)M_1(\mathrm{d}u)$$

therefore bounds Z_v from below for all $t > 0$. Also, $Z_v^{(t)}$ and $Z_{v'}^{(t)}$ are independent for all $v, v' \in \mathbb{R}^d$ satisfying $|v - v'| > 2t$. Lastly, it can be seen by the arguments in [16] that (D.4.9) also holds for the field $(Z_v^{(t)})_v$ if t is large enough relative to the size of $C \subseteq [0,1]^d$, i.e.

$$\frac{\mathbb{P}(\sup_{v\in C} Z_v^{(t)} > x)}{\rho((x,\infty))} \to |C|\mathbb{E}e^{\beta Z_u^{(t)}}, \qquad (x\to\infty), \tag{D.4.10}$$

where $u \in \mathbb{R}^d$ is arbitrarily chosen.

For the remainder of the paper, we assume that $(C_n)_{n\in\mathbb{N}}$ satisfies Assumption D.2.4.

As mentioned, the conditions $\mathcal{D}$ and $\mathcal{D}'$ do not show easily when C_n is properly discretized, and we therefore continue by studying the extremal behavior of the semi-deterministic field $(Z_v + y_v)_{v\in\mathbb{R}^d}$, where $(y_v)_v$ is seen as a realization of the field $(Y_v)_v$. Having characterized the extremal behaviour of $(Z_v + y_v)_v$, we conclude the behaviour of $(X_v)_v$ by an independence argument. To discretize, we need the notion of a unit-cube $C(z)$ in $\mathbb{R}^d$: For a point $z = (z_1,\dots,z_d) \in \mathbb{Z}^d$, we let $C(z) \subseteq \mathbb{R}^d$ denote the closed unit-cube given by

$$C(z) = \bigtimes_{j=1}^{d} [z_j, z_j + 1],$$

and we say that the *corner* $z \in \mathbb{Z}^d$ has (associated) *unit-cube* $C(z)$, and vice versa. We note that

$$\bigcup_{z\in D_{n,k}^-} C(z) \subseteq C_n \subseteq \bigcup_{z\in D_{n,k}^+} C(z)$$

Thus,

$$\max_{z\in D_{n,k}^-} \sup_{u\in C(z)} (Z_u + y_u) \leq \sup_{v\in C_n} (Z_v + y_v) \leq \max_{z\in D_{n,k}^+} \sup_{u\in C(z)} (Z_u + y_u). \tag{D.4.11}$$

Before proceeding, we introduce some notation which will be convenient in the formulation and proof of the results of this section. These should be read in the context described above, however, the main result, Theorem D.4.10, is self-contained. For any discrete set $A \subseteq \mathbb{Z}^d$, we let $M_y(A)$ and $M_y^{(t)}(A)$ be the supremas over the union of (continuous) unit-cubes with corners in A,

$$M_y(A) = \max_{z\in A} \sup_{u\in C(z)} (Z_u + y_u), \qquad \text{and} \qquad M_y^{(t)}(A) = \max_{z\in A} \sup_{u\in C(z)} (Z_u^{(t)} + y_u).$$

Hence, with this notation, (D.4.11) translates to

$$M_y(D_{n,k}^-) \leq \sup_{v\in C_n} (Z_v + y_v) \leq M_y(D_{n,k}^+). \tag{D.4.12}$$

From now on, we let $(x_n)_{n\in\mathbb{N}}$ be a real sequence given by

$$x_n = a_n x + b_n, \qquad x \in \mathbb{R},$$

where a_n, b_n denotes the norming constants relative to $|C_n|$, i.e

$$|C_n|\tilde{\rho}_1((a_n x + b_n, \infty)) \to \mathrm{e}^{-x}$$

for all $x \in \mathbb{R}$ as $n \to \infty$. Then $x_n \to \infty$ as $n \to \infty$, and consequently we conclude from (D.4.8) to (D.4.10) that

$$|C_n|\mathbb{P}(\sup_{u\in C(v)} X_u > x_n) \to \mathrm{e}^{-x}\rho((1,\infty))\mathbb{E}\mathrm{e}^{\beta X_0}$$

$$|C_n|\mathbb{P}(\sup_{u\in C(v)} Z_u > x_n) \to \mathrm{e}^{-x}\rho((1,\infty))\mathbb{E}\mathrm{e}^{\beta Z_0}, \tag{D.4.13}$$

$$|C_n|\mathbb{P}(\sup_{u\in C(v)} Z_u^{(t)} > x_n) \to \mathrm{e}^{-x}\rho((1,\infty))\mathbb{E}\mathrm{e}^{\beta Z_0^{(t)}} \tag{D.4.14}$$

as $n \to \infty$. For each fixed $x \in \mathbb{R}$ we let for notational convenience τ and $\tau^{(t)}$ be defined by

$$\tau = \mathrm{e}^{-x}\rho((1,\infty))\mathbb{E}\mathrm{e}^{\beta X_0}, \qquad \text{and} \qquad \tau^{(t)} = \mathrm{e}^{-x}\rho((1,\infty))\mathbb{E}\mathrm{e}^{\beta(Z_0^{(t)}+Y_0)},$$

where $t > 0$. Note that $\tau^{(t)} \to \tau$ as $t \to \infty$ by monotone convergence.

For the results below, it is important that the tails of $(Z_v + y_v)_v$ and $(Z_v^{(t)} + y_v)_v$ behave essentially like those of the stationary fields $(X_v)_v$ and $(Z_v^{(t)} + Y_v)_v$, respectively. The following lemma will be shown in Section D.6.

Lemma D.4.4. *Let $(Z_v)_v$, $(Z_v^{(t)})_v$ and $(Y_v)_v$ be given as above. Then, for almost all realizations $(y_v)_v$ of $(Y_v)_v$, it holds for all $z \in N_k$ that*

$$\frac{|C_n|}{|J_z|}\sum_{v\in J_z}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u + y_u) > x_n\Big) \to \tau \tag{D.4.15}$$

and

$$\frac{|C_n|}{|J_z|}\sum_{v\in J_z}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u^{(t)} + y_u) > x_n\Big) \to \tau^{(t)}. \tag{D.4.16}$$

The result also holds true if J_z is replaced with a subset of J_z in the shape of a box, which increases in size asymptotically as J_z.

The following lemma, which will be proved in Section D.6, gives that a conditional version of $\mathcal{D}(x_n; K_n)$ is satisfied for the field.

Lemma D.4.5. *Let $(Z_v)_v$, $(Z_v^{(t)})_v$ and $(Y_v)_v$ be given as above. There is a sequence $\gamma_n = o(\sqrt[d]{|C_n|})$ such that for all γ_n-separated sets $A, B \subseteq K_n$ it holds that*

$$\Big|\mathbb{P}(M_y(A\cup B) \le x_n) - \mathbb{P}(M_y(A) \le x_n)\mathbb{P}(M_y(B) \le x_n)\Big| \le \alpha_{y,n},$$

where $\alpha_{y,n} \to 0$ as $n \to \infty$ for almost all realizations $(y_v)_v$ of $(Y_v)_v$.

As in the stationary case, the following generalization follows by induction.

Lemma D.4.6. *Let $(Z_v)_v$, $(Z_v^{(t)})_v$ and $(Y_v)_v$ be given as above, and let $(y_v)_v$ be a realization of $(Y_v)_v$. Let for $r \in \mathbb{N}$ the sets $A_1, \dots, A_r$ be pairwise γ_n-separated. Then*

$$\Big|\mathbb{P}\Big(\bigcap_{i=1}^{r}\{M_y(A_i) \le x_n\}\Big) - \prod_{i=1}^{r}\mathbb{P}(M_y(A_i) \le x_n)\Big| \le (r-1)\alpha_{y,n}.$$

We state the following lemma without proof, as it follows by the exact arguments as Lemma D.3.2 taking the lack of stationarity of the discretely indexed field $(\sup_{u\in C(v)}(Z_u + y_u))_{v\in\mathbb{Z}^d}$ into account.

Lemma D.4.7. *Let $(Z_v)_v$ and $(Y_v)_v$ be given as above, and let $(y_v)_v$ be a realization of $(Y_v)_v$. Then it holds that*

$$\begin{aligned}&\Big|\mathbb{P}(M_y(D^-_{n,k}) \le x_n) - \prod_{z\in P_{n,k}}\mathbb{P}(M_y(J_z) \le x_n)\Big| \\ &\le 2\sum_{z\in P_{n,k}}\mathbb{P}(M_y(H_z) \le x_n < M_y(H_z^*)) + (p_{n,k}-1)\alpha_{y,n},\end{aligned} \tag{D.4.17}$$

and similarly

$$\begin{aligned}&\Big|\mathbb{P}(M_y(D^+_{n,k}) \le x_n) - \prod_{z\in Q_{n,k}}\mathbb{P}(M_y(J_z) \le x_n)\Big| \\ &\le 2\sum_{z\in Q_{n,k}}\mathbb{P}(M_y(H_z) \le x_n < M_y(H_z^*)) + (q_{n,k}-1)\alpha_{y,n}.\end{aligned} \tag{D.4.18}$$

Lemma D.4.8. *Let $(Z_v)_v$ and $(Y_v)_v$ be given as above. Then, for almost all realizations $(y_v)_v$ of $(Y_v)_v$, the following is satisfied*

$$\begin{aligned}\Big(\liminf_{n\to\infty}\min_{z\in N_k}\mathbb{P}(M_y(J_z) \le x_n)\Big)^{\tilde{q}_k} &\le \liminf_{n\to\infty}\mathbb{P}\Big(\sup_{v\in C_n}(Z_v + y_v) \le x_n\Big) \\ &\le \limsup_{n\to\infty}\mathbb{P}\Big(\sup_{v\in C_n}(Z_v + y_v) \le x_n\Big) \\ &\le \Big(\limsup_{n\to\infty}\max_{z\in N_k}\mathbb{P}(M_y(J_z) \le x_n)\Big)^{\tilde{p}_k},\end{aligned} \tag{D.4.19}$$

where $\tilde{p}_k = \liminf_n p_{n,k}$ and $\tilde{q}_k = \limsup_n q_{n,k}$.

Proof. Let $R^p_{n,k} \ge 0$ and $R^q_{n,k} \ge 0$ denote the upper bounds in (D.4.17) and (D.4.18), respectively. For all $z \in N_k$, we have

$$\begin{aligned}\mathbb{P}(M_y(H_z) \le x_n < M_y(H_z^*)) &\le \mathbb{P}(M_y(H_z^*) > x_n) \\ &\le \sum_{j=1}^{d}\mathbb{P}(M_y(L^*_{z,j}) > x_n) \\ &\le \sum_{j=1}^{d}\sum_{v\in L^*_{z,j}}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u + y_u) > x_n\Big),\end{aligned}$$

where we recall that $|L^*_{z,j}| = t_{n,k}^{d-1}\gamma_n$. We then find that $|L^*_{z,j}| = o(|J_z|)$ as $n \to \infty$ for all $j = 1,\dots,d$, and in particular $J_z \setminus L^*_{z,j}$ is a box, which increases in size asymptotically as J_z. Since the limit in (D.4.15) is finite and $|J_z|, |J_z \setminus L^*_{z,j}|$ and $|C_n|$ are asymptotically of the same order, we conclude by Lemma D.4.4 that

$$\begin{aligned}\mathbb{P}(M_y(H_z) \le x_n < M_y(H_z^*)) &\le \sum_{j=1}^{d}\sum_{v\in J_z}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u + y_u) > x_n\Big)\\ &\quad - \sum_{j=1}^{d}\sum_{v\in J_z\setminus L^*_{z,j}}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u + y_u) > x_n\Big)\\ &\to 0\end{aligned}$$

almost surely for all $z \in N_k$. By Lemma D.4.5 it now follows that

$$\lim_{n\to\infty} R^p_{n,k} = \lim_{n\to\infty} R^q_{n,k} = 0$$

almost surely. Turning to (D.4.12) and using Lemma D.4.7 show that

$$\begin{aligned}\liminf_n \prod_{z\in Q_{n,k}} \mathbb{P}(M_y(J_z) \le x_n) &\le \liminf_n \mathbb{P}\Big(\sup_{v\in C_n}(Z_v + y_v) \le x_n\Big)\\ &\le \limsup_n \mathbb{P}\Big(\sup_{v\in C_n}(Z_v + y_v) \le x_n\Big)\\ &\le \limsup_n \prod_{z\in P_{n,k}} \mathbb{P}(M_y(J_z) \le x_n).\end{aligned}$$

Since the involved factors are probabilities and thus lie in the interval $[0,1]$, we easily obtain (D.4.19) as desired. □

The following lemma shows that a conditional version of the anti–clustering condition $\mathcal{D}'(x_n)$ is satisfied. The proof is deferred to Section D.6.

Lemma D.4.9. *Let $(Z_v^{(t)})_v$ and $(Y_v)_v$ be given as above. Then there is a function g of order $g(k) = o(k^{-1})$ as $k \to \infty$ such that*

$$\limsup_{n\to\infty} \sum_{v<v'\in J_z} \mathbb{P}\Big(\sup_{u\in C(v)}(Z_u^{(t)} + y_u) > x_n,\ \sup_{u\in C(v')}(Z_u^{(t)} + y_u) > x_n\Big) \le g(k) \tag{D.4.20}$$

for almost all realizations $(y_v)_v$ of $(Y_v)_v$ and all t large enough.

The following theorem is the main result of the section and in the formulation we explicitly state the assumptions under which the limit holds.

Theorem D.4.10. *Let $(X_v)_{v\in\mathbb{R}^d}$ be a Lévy-driven stationary field given by* (D.4.3) *where the Lévy basis M satisfies Assumption D.4.1 and the kernel function f satisfies Assumption D.4.2. Let $(C_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathbb{R}^d$ satisfying Assumption D.2.4, and let a_n, b_n be the norming constants of the Lévy measure ρ relative to $|C_n|$, i.e. $\lim_n |C_n|\rho((a_n x + b_n, \infty)) = \mathrm{e}^{-x}\rho((1,\infty))$ for all $x \in \mathbb{R}$. Then, as $n \to \infty$,*

$$\mathbb{P}\Big(a_n^{-1}\Big(\sup_{v\in C_n} X_v - b_n\Big) \le x\Big) \to \exp\Big(-\mathrm{e}^{-x}\mathbb{E}\mathrm{e}^{\beta X_u}\rho((1,\infty))\Big) \tag{D.4.21}$$

for all $x \in \mathbb{R}$, where $u \in \mathbb{R}^d$ is arbitrarily chosen.

Proof. We use the same notation already used throughout this section. In particular, $x_n = a_n x + b_n$, $\tau = \mathrm{e}^{-x}\mathbb{E}\mathrm{e}^{\beta X_u}\rho((1,\infty))$ and $\tau^{(t)} = \mathrm{e}^{-x}\mathbb{E}\mathrm{e}^{\beta(Z_u^{(t)}+Y_u)}\rho((1,\infty))$, where $u \in \mathbb{R}^d$ is arbitrarily chosen due to stationarity. Moreover, $\tau^{(t)} \to \tau$ as $t \to \infty$.

Similarly as in the previous section, we can find upper and lower bounds to the probability $\mathbb{P}(M_y(J_z) \le x_n)$, now taking the lack of stationarity into account. Using these in combination with Lemma D.4.8 and the fact that $\sup_{C(v)}(Z_u^{(t)} + y_u) \le \sup_{C(v)}(Z_u + y_u)$ imply that

$$\begin{aligned}
\Big(\liminf_n \min_{z\in N_k}\Big(1-\sum_{v\in J_z}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u+y_u)>x_n\Big)\Big)\Big)^{\tilde{q}_k}&\\
\le \liminf_n \mathbb{P}\Big(\sup_{v\in C_n}(Z_v+y_v)\le x_n\Big)&\\
\le \limsup_n \mathbb{P}\Big(\sup_{v\in C_n}(Z_v+y_v)\le x_n\Big)&\\
\le \Big(\limsup_n \max_{z\in N_k}\Big(1-\sum_{v\in J_z}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u^{(t)}+y_u)>x_n\Big)+S_{n,k}(z)\Big)\Big)^{\tilde{p}_k}&,
\end{aligned} \tag{D.4.22}$$

where

$$S_{n,k}(z) = \sum_{v<v'\in J_z}\mathbb{P}\Big(\sup_{u\in C(v)}(Z_u^{(t)}+y_u)>x_n,\ \sup_{u\in C(v')}(Z_u^{(t)}+y_u)>x_n\Big),$$

and $\tilde{q}_k = \limsup q_{n,k}$ and $\tilde{p}_k = \liminf_n p_{n,k}$. By Lemma D.4.9, $\limsup_n S_{n,k}(z) = o(k^{-1})$ as $k \to \infty$ uniformly in z. Since $t_{n,k}^d \sim |C_n|/k$, we find by Lemma D.4.4 and (D.4.22) that

$$\begin{aligned}
\Big(1-\frac{\tau}{k}\Big)^{\tilde{q}_k} &\le \liminf_n \mathbb{P}\Big(\sup_{v\in C_n}(Z_v+y_v)\le x_n\Big)\\
&\le \limsup_n \mathbb{P}\Big(\sup_{v\in C_n}(Z_v+y_v)\le x_n\Big)\\
&\le \Big(1-\frac{\tau^{(t)}}{k}+o(k^{-1})\Big)^{\tilde{p}_k}
\end{aligned}$$

almost surely. First taking the limit $k \to \infty$ combined with the fact that $\tilde{p}_k \sim \tilde{q}_k \sim k$, and secondly taking the limit $t \to \infty$ show

$$\lim_{n\to\infty}\mathbb{P}\Big(\sup_{v\in C_n}(Z_v+y_v)\le x_n\Big) = \tau$$

almost surely. Let π denote the distribution of the field $(Y_v)_v$. Then, by independence and dominated convergence,

$$\mathbb{P}\Big(\sup_{v\in C_n}X_v\le x_n\Big) = \int \mathbb{P}\Big(\sup_{v\in C_n}(Z_v+y_v)\le x_n\Big)\pi(\mathrm{d}y) \to \tau$$

as $n \to \infty$. This is exactly (D.4.21). □

Going through the arguments leading up to Theorem D.4.10 it is seen that the Lévy based form behind the field $(Y_v)_v$ is only used to obtain that it is independent of $(Z_v)_v$ and furthermore stationary, ergodic and satisfying the integrability result of Lemma D.5.2. Therefore, Theorem D.4.10 is immediately extended to

Theorem D.4.11. *Let $(X_v)_{v\in\mathbb{R}^d}$ be a Lévy-driven stationary field given by* (D.4.3) *where the Lévy basis M satisfies Assumption D.4.1 and the kernel function f satisfies Assumption D.4.2. Let $(C_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathbb{R}^d$ satisfying Assumption D.2.4, and let a_n, b_n be the norming constants of the Lévy measure ρ relative to $|C_n|$, i.e. $\lim_n |C_n|\rho((a_n x + b_n, \infty)) = \mathrm{e}^{-x}\rho((1,\infty))$ for all $x \in \mathbb{R}$. Let furthermore $(\tilde{Y}_v)_v$ be a stationary and ergodic field independent of $(X_v)_v$ which satisfies*

$$\mathbb{E}\exp\Big(\gamma \sup_{v\in C(0)} \tilde{Y}_v\Big) < \infty$$

for some $\gamma > 2\beta$, where β is introduced in Assumption D.4.1. Then, as $n \to \infty$,

$$\mathbb{P}\Big(a_n^{-1}\Big(\sup_{v\in C_n} X_v + \tilde{Y}_v - b_n\Big) \le x\Big) \to \exp\Big(-\mathrm{e}^{-x}\mathbb{E}\mathrm{e}^{\beta(X_u+\tilde{Y}_u)}\rho((1,\infty))\Big)$$

for all $x \in \mathbb{R}$, where $u \in \mathbb{R}^d$ is arbitrarily chosen.

D.5 Proofs related to the $(\boldsymbol{Y_v})$–field

We start this section by considering the tail of a distribution in $\mathcal{L}_\beta$, and we give a bound which will be useful in the proofs below. In the formulation of the result, y^+ is defined as $y^+ = \max\{y, 0\}$ for all $y \in \mathbb{R}$.

Lemma D.5.1. *Let $G \in \mathcal{L}_\beta$ be a distribution with exponential right tail with index $\beta \ge 0$. Let $\overline{G}$ be its tail. For all $\gamma > \beta$ there is $x_0 \in \mathbb{R}$ and $\tilde{C} \in \mathbb{R}$ such that*

$$\overline{G}(x-y) \le \overline{G}(x)\tilde{C}\exp(\gamma y^+) \tag{D.5.1}$$

for all $x \ge x_0$ and $y \in \mathbb{R}$.

Proof. Since $G \in \mathcal{L}_\beta$ is follows that

$$\overline{G}(x) = a(x)\exp\Big(-\int_0^x \beta(y)\mathrm{d}y\Big),$$

where $a(x) \to a > 0$ and $\beta(x) \to \beta$ as $x \to \infty$. This is due to Karamata's representation theorem (see e.g. [5, Theorem A3.3]) and the fact that $\overline{G}\circ\log$ is a regularly varying function of index $-\beta$. Fix $\gamma > \beta$ and find x_0 such that

$$\beta(x) < \gamma, \qquad \text{and} \qquad |a(x) - a| \le a/3 \tag{D.5.2}$$

for all $x \ge x_0$.

Now consider only $x \ge x_0$ and $y \ge 0$. If $y \le x - x_0$ and hence $x - y \ge x_0$, we find from (D.5.2) above that

$$\begin{aligned}\overline{G}(x-y) &= \overline{G}(x)\frac{a(x-y)}{a(x)}\exp\Big(\int_{x-y}^x \beta(y)\mathrm{d}y\Big)\\ &\le \overline{G}(x)2\exp(\gamma y).\end{aligned}$$

If on the other hand $y > x - x_0$, we see that

$$\begin{aligned}\overline{G}(x-y) &\le \frac{\overline{G}(x)}{\overline{G}(x)} = \overline{G}(x)\Big[a(x)\exp\Big(-\int_0^{x_0}\beta(y)\mathrm{d}y\Big)\Big]^{-1}\exp\Big(\int_{x_0}^x \beta(y)\mathrm{d}y\Big)\\ &\le \overline{G}(x)\tilde{C}_0\exp(\gamma y),\end{aligned}$$

where the constant $\tilde{C}_0$ can be chosen as

$$\tilde{C}_0 = \Big[\frac{2a}{3}\exp\Big(-\int_0^{x_0}\beta(y)\mathrm{d}y\Big)\Big]^{-1}.$$

Choosing $\tilde{C} = \max\{2, \tilde{C}_0\}$ shows (D.5.1) for $y \geq 0$.

If $y < 0$ the claim (D.5.1) reads $\overline{G}(x-y) \leq \overline{G}(x)\tilde{C}$, which is clearly true since $x \mapsto \overline{G}(x)$ is decreasing. □

The following result will be used repeatedly in the subsequent proofs and helps ensuring that the field $(Y_v)_v$ has minor importance when determining the extremal behaviour of $(X_v)_v$.

Lemma D.5.2. *Let the field* $(Y_v)_{v\in\mathbb{R}^d}$ *be given by* (D.4.7). *Then*

$$\mathbb{E}\exp\Big(\gamma \sup_{v\in C(0)} Y_v\Big) < \infty \tag{D.5.3}$$

for all $\gamma > 0$.

Note that the result in Lemma D.5.2 is equivalent with $\mathbb{E}\exp\Big(\gamma(\sup_{C(0)} Y_v)^+\Big) < \infty$ for all $\gamma > 0$, which will be used specifically in Section D.6.

Proof. By considerations as in [16], [17] and Paper C, the countable field $(Y_v)_{v\in\mathbb{Q}^d}$ is infinitely divisible with characteristic function as in [2, Eq. (1.1)], where its Lévy measure ν on $\mathbb{R}^{\mathbb{Q}^d}$ is given as follows: Define $H : \mathbb{R}^d \times (-\infty, 1] \to \mathbb{R}^{\mathbb{Q}^d}$ as

$$H(u,x) = (xf(|v-u|))_{v\in\mathbb{Q}^d}.$$

With m denoting the Lebesgue measure, we let $\nu = (m\otimes\rho_2)\circ H^{-1}$ be the image-measure of $m\otimes\rho_2$. Since $v \mapsto Y_v$ is continuous on compact sets,

$$\mathbb{P}(\sup_{v\in C(0)} Y_v < \infty) = \mathbb{P}(\sup_{v\in C(0)\cap\mathbb{Q}^d} Y_v < \infty) = 1.$$

Moreover, $\nu(\{z \in \mathbb{R}^{\mathbb{Q}^d} : \sup_{C(0)\cap\mathbb{Q}^d} z_v > 1\}) = 0$, and the claim now follows from [2, Lemma 2.1]. □

In the remainder of this section we establish some useful ergodic properties of the field $(Y_v)_v$. First, we recast some notation and a result from [10]. Let $(S, \mathcal{A}, \mu)$ be a probability space, and assume that $T_i : S \to S$ is a measurable map for $i = 1, \dots, d$ such that $T_1, \dots, T_d$ commute, i.e. $T_i \circ T_j = T_j \circ T_i$ for all i, j. Furthermore, assume for all $i = 1, \dots, d$ that

$$T_i(\mu) = \mu.$$

Let $\mathbb{V} = \mathbb{Z}_+^d$, and define for each $v = (v_1, \dots, v_d) \in \mathbb{V}$ the map $T_v : S \to S$ by

$$T_v = T_1^{v_1} T_2^{v_2} \cdots T_d^{v_d},$$

where e.g. $T_1^{v_1}$ means the composition of T_1 with itself v_1 times. Note that μ is also T_v–invariant for all $v \in \mathbb{V}$.

We define a subset $I \subseteq \mathbb{V}$ to be a box if it has the form

$$I = \Big(\bigtimes_{i=1}^{d} [u_i, v_i[\Big) \cap \mathbb{Z}^d,$$

for $u = (u_1, \ldots, u_d), v = (v_1, \ldots, v_d) \in \mathbb{V}$. The set of all such boxes in $\mathbb{V}$ will be denoted $\mathcal{I}$.

Definition D.5.3. *A sequence $I_1, I_2, \cdots \subseteq \mathbb{V}$ is said to be regular if there exists an increasing sequence $I_1' \subseteq I_2' \subseteq \cdots \in \mathcal{I}$ and $c < \infty$ such that $I_i \subseteq I_i'$ and $|I_i'| \leq c|I_i|$ for each i.*

Now we can formulate an ergodic theorem, which can be found as Theorem 6.2.8 in [10]. The theorem is followed by a few definitions and theorems also know from the classical ergodic theory.

Theorem D.5.4. *Assume that $(S, \mathcal{A}, \mu)$ and $(T_i)_{i=1,\ldots,d}$ satisfies the above. Let furthermore $g : S \to \mathbb{R}$ be measurable and μ–integrable, and assume that $I_1, I_2, \cdots \in \mathcal{I}$ is a regular sequence. Then*

$$\frac{1}{|I_n|} \sum_{v \in I_n} g \circ T_v \to \mathbb{E}[g \mid \mathbb{I}]$$

μ–almost everywhere as $n \to \infty$, where $\mathbb{I}$ is the invariant σ–algebra, i.e. the σ–algebra consisting of all sets in $\mathcal{A}$ invariant to T_i for $i = 1, \ldots, d$.

If the invariant σ–algebra $\mathbb{I}$ is trivial, we say that the family $(T_v)_{v \in \mathbb{V}}$ is ergodic.

Definition D.5.5. *The family $(T_v)_{v \in \mathbb{V}}$ is mixing if*

$$\mu(F \cap T^{-v_n}(G)) \to \mu(F)\mu(G) \tag{D.5.4}$$

for all $F, G \in \mathcal{A}$ and $(v_n)_{n \in \mathbb{N}} \subset \mathbb{V}$ with $|v_n| \to \infty$.

Theorem D.5.6. *If $(T_v)_{v \in \mathbb{V}}$ is mixing, then it is ergodic.*

Proof. Let $F \in \mathbb{I}$ and choose any sequence $(v_n) \subset \mathbb{V}$ with $|v_n| \to \infty$. Then $T^{-v_n}(F) = F$, so

$$\mu(F) = \mu(F \cap T^{-v_n}(F)) \to \mu(F)^2$$

leading to F being a trivial set. □

The following theorem is obtained by a standard extension argument.

Theorem D.5.7. *For $(T_v)_{v \in \mathbb{V}}$ being mixing, it is sufficient that* (D.5.4) *is satisfied for all F and G in an intersection stable generating system for $\mathcal{A}$.*

We will apply the ergodic theory to the random field $Y = (Y_v)_{\mathbb{R}}$ defined in (D.4.7). Thus we let $S = C(\mathbb{R}^d)$ be the set of all continuous functions on $\mathbb{R}^d$, and let $\mathcal{A}$ be the corresponding σ–algebra generated by all coordinate projections. Finally, we let $\mu = Y(\mathbb{P})$. Each map $T_i : S \to S$ is defined as

$$T_i\Big((x_{t_1,\ldots,t_d})_{(t_1,\ldots,t_d) \in \mathbb{R}^d}\Big) = \Big((x_{t_1,\ldots,t_{i-1},t_i+1,t_{i+1},\ldots,t_d})_{(t_1,\ldots,t_d) \in \mathbb{R}^d}\Big),$$

which obviously commutes with T_j, i.e. $T_i \circ T_j = T_j \circ T_i$ for all $i, j = 1, \ldots, d$. Moreover, by stationarity of $(Y_v)_v$, the maps satisfy $T_i(\mu) = \mu$ for all i.

Lemma D.5.8. *Let $u_1,\dots,u_p,v_1,\dots,v_q \in \mathbb{R}^d$ and $(z_n)_{n\in\mathbb{N}} \subseteq \mathbb{V}$ with $|z_n| \to \infty$ be given. Then*

$$(Y_{u_1},\dots,Y_{u_p},Y_{v_1+z_n},\dots,Y_{v_q+z_n}) \to (Y_{u_1},\dots,Y_{u_p})(\mathbb{P}) \otimes (Y_{v_1},\dots,Y_{v_q})(\mathbb{P})$$

in distribution.

Proof. For $\lambda_1,\dots,\lambda_p,\beta_1,\dots,\beta_q \in \mathbb{R}$ we show that

$$\begin{aligned}\log \mathbb{E}\Big[\mathrm{e}^{i(\lambda_1 Y_{u_1}+\dots+\lambda_p Y_{u_p}+\beta_1 Y_{v_1+z_n}+\dots+\beta_q Y_{v_q+z_n})}\Big]&\\ \to \log \mathbb{E}\Big[\mathrm{e}^{i(\lambda_1 Y_{u_1}+\dots+\lambda_p Y_{u_p})}\Big] + \log \mathbb{E}\Big[(\mathrm{e}^{i(\beta_1 Y_{v_1}+\dots+\beta_q Y_{v_q})}\Big]&\end{aligned}$$

as $n \to \infty$. Defining $g_1(w) = \sum_{i=1}^p \lambda_i f(|u_i - w|)$ and $g_2(w) = \sum_{i=1}^q f(|v_i - w|)$, and utilizing that $\int_{\mathbb{R}} y^2 \rho(\mathrm{d}y) < \infty$, we can write

$$\begin{aligned}&\log \mathbb{E}\Big[\mathrm{e}^{i(\lambda_1 Y_{u_1}+\dots+\lambda_p Y_{u_p}+\beta_1 Y_{v_1+z_n}+\dots+\beta_q Y_{v_q+z_n})}\Big]\\ &\quad= ia' \int_{\mathbb{R}^d} g_1(w) + g_2(w-z_n)\mathrm{d}w - \frac{1}{2}\theta^2 \int_{\mathbb{R}^d} \Big(g_1(w) + g_2(w-z_n)\Big)^2 \mathrm{d}w\\ &\qquad+ \int_{\mathbb{R}^d}\int_{\mathbb{R}} \mathrm{e}^{i(g_1(w)+g_2(w-z_n))x} - 1 - i(g_1(w)+g_2(w-z_n))x\,\rho(\mathrm{d}x)\,\mathrm{d}w\end{aligned} \tag{D.5.5}$$

for an appropriate constant a'. The proof is complete, when it is shown that the limit of (D.5.5) is

$$\sum_{j=1}^2 \Big[ia' \int_{\mathbb{R}^d} g_j(w)\mathrm{d}w - \frac{1}{2}\theta^2 \int_{\mathbb{R}^d} g_j(w)^2 \mathrm{d}w + \int_{\mathbb{R}^d}\int_{\mathbb{R}} \mathrm{e}^{g_j(w)x} - 1 - ig_j(w)x\,\rho(\mathrm{d}x)\,\mathrm{d}w\Big].$$

For the first term in (D.5.5) we have the equality

$$\int_{\mathbb{R}^d} g_1(w) + g_2(w-z_n)\mathrm{d}w = \int_{\mathbb{R}^d} g_1(w)\mathrm{d}w + \int_{\mathbb{R}^d} g_2(w)\mathrm{d}w.$$

The integral in the second term of (D.5.5) equals

$$\int_{\mathbb{R}^d} g_1(w)^2 \mathrm{d}w + \int_{\mathbb{R}^d} g_2(w)^2 \mathrm{d}w + 2\int_{\mathbb{R}^d} g_1(w) g_2(w-z_n)\mathrm{d}w,$$

where the last term converges 0 due to dominated convergence, since g_1 is integrable, and $g_2(w - z_n)$ is bounded and has pointwise limit 0 by assumptions on the kernel f.

Finally, for the convergence of the third term in (D.5.5) we let $\epsilon > 0$ be given and choose $D > 0$ such that for $i = 1,2$,

$$\int_{B(D)^c} g_i(w)^2 \mathrm{d}w \cdot C < \epsilon/4, \tag{D.5.6}$$

where $C = \int_{\mathbb{R}} x^2 \rho(\mathrm{d}x)$. With $h(w,x)$ denoting the integrand in the third term of (D.5.5), the integral, denoted $\mathcal{J}_n$, can for large n be rewritten as

$$\begin{aligned}\mathcal{J}_n = &\int_{B(D)}\int_{\mathbb{R}} h(w,x)\rho(\mathrm{d}x)\mathrm{d}w + \int_{B(D)}\int_{\mathbb{R}} h(w+z_n,x)\rho(\mathrm{d}x)\mathrm{d}w\\ &+ \int_{(B(D)\cup(B(D)+z_n))^c}\int_{\mathbb{R}} h(w,x)\rho(\mathrm{d}x)\mathrm{d}w.\end{aligned} \tag{D.5.7}$$

Using $|e^{ix} - 1 - ix| \leq x^2$ it is seen that the third term in (D.5.7) is bounded from above by

$$C \cdot \Big(\int_{B(D)^c} g_1(w)^2 \mathrm{d}w + \int_{(B(D)+z_n)^c} g_2(w-z_n)^2 \mathrm{d}w + 2 \int_{\mathbb{R}^d} g_1(w) g_2(w-z_n) \mathrm{d}w \Big)$$
$$\leq \epsilon/2 + 2C \int_{\mathbb{R}^d} g_1(w) g_2(w-z_n) \mathrm{d}w,$$

where the integral has limit 0 with an argument similarly as above.

The limit of the sum of the first two terms in (D.5.7) is

$$\int_{B(D)} \int_{\mathbb{R}} \mathrm{e}^{ig_1(w)x} - 1 - ig_1(w)x\, \rho(\mathrm{d}x)\, \mathrm{d}w$$
$$+ \int_{B(D)} \int_{\mathbb{R}} \mathrm{e}^{ig_2(w)x} - 1 - ig_2(w)x\, \rho(\mathrm{d}x)\, \mathrm{d}w$$

due to dominated convergence. Collecting the limit results for the three terms of (D.5.7) and referring to (D.5.6) again, we find

$$\limsup_{n\to\infty} \Big| \mathcal{J}_n - \int_{\mathbb{R}^d} \int_{\mathbb{R}} \mathrm{e}^{ig_1(w)x} - 1 - ig_1(w)x\, \rho(dx)\, \mathrm{d}w$$
$$- \int_{\mathbb{R}^d} \int_{\mathbb{R}} \mathrm{e}^{ig_2(w)x} - 1 - ig_2(w)x\, \rho(dx)\, \mathrm{d}w \Big| \leq \epsilon,$$

from which the desired conclusion is obtained, since ϵ was chosen arbitrarily. □

Theorem D.5.9. *Let $(Y_v)_v$ be defined as above and let $(I_n)_n$ be a regular sequence of boxes in $\mathbb{Z}_+^d$. If $g : S \to \mathbb{R}$ satisfies $\mathbb{E}|g((Y_u)_u)| < \infty$ then*

$$\frac{1}{|I_n|} \sum_{v \in I_n} g((Y_{u+v})_u) \to \mathbb{E} g((Y_u)_u)$$

almost surely as $n \to \infty$.

Proof. Let A be the set of continuity points for the distribution of, say, Y_0, which due to the stationarity is the set of continuity points for all Y_v. Note that A is dense in $\mathbb{R}$. Lemma D.5.8 implies that for all $a_1, \ldots, a_p$, $b_1, \ldots, b_q \in A$,

$$\mathbb{P}(Y_{u_1} \leq a_1, \ldots, Y_{u_p} \leq a_p, Y_{v_1+z_n} \leq b_1, \ldots, Y_{v_q+z_n} \leq b_p)$$
$$\to \mathbb{P}(Y_{u_1} \leq a_1, \ldots, Y_{u_p} \leq a_p) \mathbb{P}(Y_{v_1} \leq b_1, \ldots, Y_{v_q} \leq b_p).$$

Since sets on the form

$$\{y \in S \,:\, y_{u_1} \leq a_1, \ldots, y_{u_p} \leq a_p\},$$

where $p \in \mathbb{N}_0$, $u_1, \ldots, u_p$, $a_1, \ldots, a_p \in A$ constitutes an intersection stable generating system for $\mathcal{A}$, we have from Theorems D.5.4, D.5.6 and D.5.7 that for any μ–integrable map $g : S \to \mathbb{R}$,

$$1 = \mu\Big(\frac{1}{|I_n|} \sum_{v \in I_n} g \circ T_v \to \int g \mathrm{d}\mu \Big) = \mathbb{P}\Big(\frac{1}{|I_n|} \sum_{v \in I_n} g(T_v(Y)) \to \mathbb{E}(g(Y)) \Big),$$

This concludes the proof. □

The following corollary adapts Theorem D.5.9 into the specific setting that will be useful in the further arguments.

Corollary D.5.10. *Let the field $(Y_v)_{v\in\mathbb{R}^d}$ be given by* (D.4.7), *and let g be a function satisfying $\mathbb{E}|g((Y_u)_{u\in C(0)})| < \infty$. For all $z \in N_k$ it then holds that*

$$\frac{1}{|J_z|}\sum_{v\in J_z} g((Y_{u+v})_{u\in C(0)}) \to \mathbb{E}g((Y_u)_{u\in C(0)})$$

almost surely as $n\to\infty$. The result also holds true if J_z is replaced with a subset of J_z in the shape of a box, which increases in size asymptotically as J_z.

Proof. In principle, we can only apply Theorem D.5.9 to the collection $J_z^{n,k}$ of sets contained in $\mathbb{Z}_+^d$. But by symmetry, the same result can be obtained for $J_z^{n,k}$ sets in all the 2^d other placements relative to 0.

Let $z \in N_k$ be fixed and consider $J_z = J_z^{n,k}$. By construction, $(K_{n,k})_n$ is an increasing sequence of cubes such that $J_z^{n,k} \subseteq K_{n,k}$ and $|K_{n,k}| = |N_k|\cdot|J_z^{n,k}|$. In particular, $(J_z^{n,k})_n$ is a regular sequence of boxes in $\mathbb{Z}^d$, and the first part of the result thus follows from Theorem D.5.9.

Now let $L_z^n \subseteq J_z^n$ be a box such that $|J_z^n|/|L_z^n| \to c \in [1,\infty)$ as $n\to\infty$. Then, for sufficiently large n, the relation $|K_{n,k}| \le 2c|N_k|\cdot|L_z^n|$ holds, and $(L_z^n)_n$ is thus a regular sequence of boxes in $\mathbb{Z}^d$. The proof is completed by yet another application of Theorem D.5.9. □

D.6 Remaining proofs

In what follows, $C_r(v)$ is an r-cube with corner $v\in\mathbb{R}^d$, that is, a box in $\mathbb{R}^d$ with side-length equal to $r>0$. Moreover, as up until now, we let $C(v) = C_1(v)$ denote the unit-cube with corner v.

Proof of Lemma D.4.4. We only show the convergence (D.4.15) as (D.4.16) and the expressions for J_z replaced by an asymptotically size-equivalent box follow identically.

Let $L\in\mathbb{N}$ be fixed. For all $v\in\mathbb{Z}^d$ define $A_L(v)$ as the set of corners in a grid with distance $1/L$ for which the associated $1/L$-cubes are contained in $C(v)$, i.e.

$$A_L(v) = \{u \in (L^{-1}\mathbb{Z})^d \,:\, C_{1/L}(u)\subseteq C(v)\}.$$

With this construction it follows that

$$C(v) = \bigcup_{u\in A_L(v)} C_{1/L}(u).$$

For $v\in\mathbb{Z}^d$, define $y^*(v) = \sup_{u\in C(v)} y_u$. Similarly, for all $u\in(L^{-1}\mathbb{Z})^d$, define $y_L^*(u) = \sup_{s\in C_{1/L}(u)} y_s$, and $\overline{y}_L(u) = \inf_{s\in C_{1/L}(u)} y_s$. We now let F_L denote the distribution of $\sup_{s\in C_{1/L}(u)} Z_s$, which, by stationarity, is independent of $v\in\mathbb{Z}^d$ and $u\in A_L(v)$. Then, from (D.4.13),

$$|C_n|\overline{F}_L(x_n) \to L^{-d}\mathrm{e}^{-x}\rho((1,\infty))\mathbb{E}\mathrm{e}^{\beta Z_0} \tag{D.6.1}$$

as $n \to \infty$, where $\overline{F}_L$ is the tail of F_L. Since $\rho \in \mathcal{L}_\beta$ the equivalence (D.4.9) implies that also $\overline{F}_L \in \mathcal{L}_\beta$. From Lemma D.5.1 we conclude for any $\gamma > \beta$ the existence of a finite constant $\tilde{C}_L$ such that

$$\overline{F}_L(x_n - y) \le \overline{F}_L(x_n)\tilde{C}_L \exp(\gamma y^+), \qquad \text{for all} \quad y \in \mathbb{R}, \tag{D.6.2}$$

for n sufficiently large.

Writing the supremum over $C(v)$ as a maximum of supremas over $C_{1/L}(u)$ for $u \in A_L(v)$, it is not difficult to see that

$$\begin{aligned} \sum_{v \in J_z} \sum_{u \in A_L(v)} \overline{F}_L(x_n - \overline{y}_L(u)) - \sum_{v \in J_z} S_L(v) & \\ \le \sum_{v \in J_z} \mathbb{P}\Big(\sup_{u \in C(v)} (Z_u + y_u) > x_n \Big) & \\ \le \sum_{v \in J_z} \sum_{u \in A_L(v)} \overline{F}_L(x_n - y_L^*(u)), & \end{aligned} \tag{D.6.3}$$

where

$$S_L(v) = \sum_{u < u' \in A_L(v)} \mathbb{P}\Big(\sup_{s \in C_{1/L}(u)} Z_s > x_n - y^*(v), \sup_{s \in C_{1/L}(u')} Z_s > x_n - y^*(v) \Big).$$

First, we consider the upper bound in (D.6.3). Since $F_L \in \mathcal{L}_\beta$, we find that the convergence $\overline{F}_L(x_n - y)/\overline{F}_L(x_n) \to \exp(\beta y)$, $n \to \infty$, is uniform for $y \le K$ for all $K \in \mathbb{N}$; see e.g [13, Definition 2.1]. Using Corollary D.5.10 and this uniform convergence, (D.6.1) and (D.6.2) now imply for all fixed $K \in \mathbb{N}$ that

$$\begin{aligned} &\frac{|C_n|}{|J_z|} \sum_{v \in J_z} \sum_{u \in A_L(v)} \overline{F}_L(x_n - y_L^*(u)) \\ &\le |C_n| \overline{F}_L(x_n) \frac{1}{|J_z|} \sum_{v \in J_z} \sum_{u \in A_L(v)} \Big(\frac{\overline{F}_L(x_n - y_L^*(u))}{\overline{F}_L(x_n)} 1_{y^*(v) \le K} \\ &\qquad\qquad + \tilde{C}_L \exp(\gamma (y_L^*(u))^+) 1_{y^*(v) > K} \Big) \\ &\to \mathrm{e}^{-x} \rho((1,\infty)) \mathbb{E}\mathrm{e}^{\beta Z_0} \mathbb{E}\Big[\exp\Big(\beta \sup_{v \in C_{1/L}(0)} Y_v\Big) 1_{\sup_{v \in C_1(0)} Y_v \le K} \\ &\qquad\qquad + \tilde{C}_L \exp\Big(\gamma (\sup_{v \in C_{1/L}(0)} Y_v)^+\Big) 1_{\sup_{v \in C_1(0)} Y_v > K} \Big] \end{aligned}$$

almost surely as $n \to \infty$, where the stationarity of $(Y_v)_v$ has also been used. Since (D.5.3) holds for all $\gamma > 0$ and $v \mapsto Y_v$ is continuous, letting $K \to \infty$ and then $L \to \infty$ show by dominated convergence that

$$\begin{aligned} \limsup_{n \to \infty} \frac{|C_n|}{|J_z|} \sum_{v \in J_z} \mathbb{P}\Big(\sup_{u \in C(v)} (Z_u + y_u) > x_n \Big) &\le \mathrm{e}^{-x} \rho((1,\infty)) \mathbb{E}\mathrm{e}^{\beta (Z_0 + Y_0)} \\ &= \mathrm{e}^{-x} \rho((1,\infty)) \mathbb{E}\mathrm{e}^{\beta X_0} = \tau \end{aligned}$$

almost surely. Concerning the lower bound of (D.6.3), we find for all fixed $u \ne u' \in (L^{-1}\mathbb{Z})^d$ that

$$\frac{1}{\overline{F}_L(x_n)} \mathbb{P}\Big(\sup_{s \in C_{1/L}(u)} Z_s > x_n - y, \sup_{s \in C_{1/L}(u')} Z_s > x_n - y \Big) \to 0 \tag{D.6.4}$$

uniformly for $y \le K$, for all $K \in \mathbb{N}$. This is easily seen from (D.4.9) and the inclusion-exclusion principle, with the convergence being uniform due to $\overline{F}_L \in \mathcal{L}_\beta$. Turning to (D.6.2) and repeating the arguments above, we conclude for all $L \in \mathbb{N}$ that

$$\frac{|C_n|}{|J_z|} \sum_{v \in J_z} S_L(v) \to 0$$

almost surely as $n \to \infty$. For the former term of the lower bound, arguing as for the upper bound shows that

$$\liminf_{n\to\infty} \frac{|C_n|}{|J_z|} \sum_{v \in J_z} \mathbb{P}\Big(\sup_{u \in C(v)} (Z_u + y_u) > x_n \Big) \ge \tau$$

almost surely, which concludes the proof. □

Proof of Lemma D.4.5. We show the claim for $d \ge 2$ first. Let the sets A and B be given as in the lemma and define

$$\mathcal{A} = \bigcup_{v \in A} C(v) \qquad \text{and} \qquad \mathcal{B} = \bigcup_{v \in B} C(v).$$

Recall that $B(r)$ denotes the closed ball in $\mathbb{R}^d$ of radius $r \ge 0$ with center in $0 \in \mathbb{R}^d$, and define $\mathcal{A}_n = \mathcal{A} \oplus B(\gamma_n/2)$ and similarly $\mathcal{B}_n = \mathcal{B} \oplus B(\gamma_n/2)$, where γ_n is given by Assumption D.4.2. For all $v \in \mathcal{A}$ let

$$Z_v^A = \int_{\mathcal{A}_n} f(|v-u|) M_1(\mathrm{d}u), \qquad \text{and} \qquad \overline{Z}_v^A = \int_{\mathcal{A}_n^c} f(|v-u|) M_1(\mathrm{d}u).$$

Similarly, for all $v \in \mathcal{B}$,

$$Z_v^B = \int_{\mathcal{B}_n} f(|v-u|) M_1(\mathrm{d}u), \qquad \text{and} \qquad \overline{Z}_v^B = \int_{\mathcal{B}_n^c} f(|v-u|) M_1(\mathrm{d}u).$$

Since M_1 is a positive measure all Z_v^A, $\overline{Z}_v^A$, Z_v^B and $\overline{Z}_v^B$ are non–negative, and we have

$$\sup_{v \in \mathcal{A}} \overline{Z}_v^A \le \int_{\mathcal{A}_n^c} \sup_{v \in \mathcal{A}} f(|v-u|) M_1(\mathrm{d}u) \tag{D.6.5}$$

and

$$\sup_{v \in \mathcal{B}} \overline{Z}_v^B \le \int_{\mathcal{B}_n^c} \sup_{v \in \mathcal{B}} f(|v-u|) M_1(\mathrm{d}u).$$

Let $r_n \sim c\sqrt[d]{|C_n|}$ for some $c < \infty$, such that $K_n \subseteq B(r_n)$. By Assumption D.4.2, equations (D.4.5) and (D.4.6), we may choose a sequence $(\epsilon_n)_{n\in\mathbb{N}}$ such that $\epsilon_n \downarrow 0$ and

$$\frac{1}{\epsilon_n}|C_n| g(\gamma_n/2) \to 0 \qquad \text{and} \qquad \frac{1}{\epsilon_n}\int_{\gamma_n/2}^{\infty} g(x)(x+r_n)^{d-1}\mathrm{d}x \to 0 \tag{D.6.6}$$

as $n \to \infty$, where we recall that g is a decreasing upper bound to f. Define the events S_n^A and S_n^B by

$$S_n^A = \Big(\sup_{v \in \mathcal{A}} \overline{Z}_v^A \le \epsilon_n \Big) \qquad \text{and} \qquad S_n^B = \Big(\sup_{v \in \mathcal{B}} \overline{Z}_v^B \le \epsilon_n \Big).$$

From Markov's inequality and (D.6.5) we find

$$\begin{aligned}\mathbb{P}\Big((S_n^A)^c\Big) &\le \frac{1}{\epsilon_n}\mathbb{E}(M_1')\int_{\mathcal{A}_n^c}\sup_{v\in\mathcal{A}} f(|v-u|)\mathrm{d}u \\ &\le \frac{1}{\epsilon_n}\mathbb{E}(M_1')\Big[\Big|B(r_n+\gamma_n/2)\Big|g(\gamma_n/2)+\int_{B(r_n+\gamma_n/2)^c} g(|u|-r_n)\mathrm{d}u\Big],\end{aligned} \tag{D.6.7}$$

with an identical upper bound for the probability $\mathbb{P}\Big((S_n^B)^c\Big)$. Similarly to the notation introduced in Section D.4, we define

$$\begin{aligned} M_y^A(A) &= \max_{v\in A}\sup_{u\in C(v)}(Z_u^A+y_u) = \sup_{v\in\mathcal{A}}(Z_v^A+y_v), \qquad \text{and} \\ M_y^B(B) &= \max_{v\in B}\sup_{u\in C(v)}(Z_u^B+y_u) = \sup_{v\in\mathcal{B}}(Z_v^B+y_v). \end{aligned}$$

Utilizing that M_y^A and M_y^B are independent, it can be seen from straightforward calculations that

$$\begin{aligned} &\Big|\mathbb{P}(M_y(A\cup B)\le x_n)-\mathbb{P}(M_y(A)\le x_n)\mathbb{P}(M_y(B)\le x_n)\Big| \\ &\quad\le \mathbb{P}(M_y^A(A)\le x_n, S_n^A)\mathbb{P}(M_y^B(B)\le x_n, S_n^B) \\ &\qquad - \mathbb{P}(M_y^A(A)\le x_n-\epsilon_n, S_n^A)\mathbb{P}(M_y^B(B)\le x_n-\epsilon_n, S_n^B) \\ &\qquad + 2\Big(\mathbb{P}\Big((S_n^A)^c\Big)+\mathbb{P}\Big((S_n^B)^c\Big)\Big) \\ &\quad\le \mathbb{P}(M_y^A(A)\le x_n, S_n^A)-\mathbb{P}(M_y^A(A)\le x_n-\epsilon_n, S_n^A) \\ &\qquad + \mathbb{P}(M_y^B(B)\le x_n, S_n^B)-\mathbb{P}(M_y^B(B)\le x_n-\epsilon_n, S_n^B) \\ &\qquad + 2\Big(\mathbb{P}\Big((S_n^A)^c\Big)+\mathbb{P}\Big((S_n^B)^c\Big)\Big). \end{aligned} \tag{D.6.8}$$

To obtain the desired conclusion it suffices to show that all three terms of (D.6.8) have upper bounds which are independent of A and B and which tend to 0 as $n\to\infty$. Concerning the first term we see that

$$\begin{aligned} 0 &\le \mathbb{P}(M_y^A(A)\le x_n, S_n^A)-\mathbb{P}(M_y^A(A)\le x_n-\epsilon_n, S_n^A) \\ &\le \mathbb{P}(\exists v\in A : x_n-\epsilon_n < M_y^A(\{v\})\le x_n, S_n^A) \\ &\le \mathbb{P}(\exists v\in A : x_n-\epsilon_n < M_y(\{v\})\le x_n+\epsilon_n, S_n^A) \\ &\le \mathbb{P}(\exists v\in K_n : x_n-\epsilon_n < M_y(\{v\})\le x_n+\epsilon_n) \\ &\le \sum_{v\in K_n}\mathbb{P}(M_y(\{v\})\le x_n+\epsilon_n)-\sum_{v\in K_n}\mathbb{P}(M_y(\{v\})\le x_n-\epsilon_n). \end{aligned}$$

Since $\epsilon_n\to 0$, the considerations that led to (D.4.15) also show that the two sums above have the same limit as $n\to\infty$ for almost all realizations $(y_v)_v$ of $(Y_v)_v$. As neither sum depend on the choice of A and B, we conclude that the first term of (D.6.8) satisfy the desired convergence. The convergence of the second term follows identically. Concerning the convergence of the third term, we see from (D.6.7) above that $\mathbb{P}\Big((S_n^A)^c\Big)$ and $\mathbb{P}\Big((S_n^B)^c\Big)$ have a common upper bound independent of A and B. Realizing that $\Big|B(r_n+\gamma_n/2)\Big|$ asymptotically equals $|C_n|$, we obtain from (D.6.6) and a coordinate change that the upper bound in (D.6.7) tends to 0 as $n\to\infty$. This completes the proof when $d\ge 2$.

Now let $d = 1$ and let γ_n be any sequence such that $\gamma_n = o(\sqrt[d]{|C_n|})$. Note that we must have a z_0 such that

$$\mathcal{A} \subseteq (-\infty, z_0 - \gamma_n/2] \qquad \text{and} \qquad \mathcal{B} \subseteq (z_0 + \gamma_n/2, \infty),$$

where $\mathcal{A}$ and $\mathcal{B}$ again denote the continuous extensions of A and B, respectively. Hence, with the notation as before, we may consider $\mathcal{A}_n$ and $\mathcal{B}_n$ as given by

$$\mathcal{A}_n = (-\infty, z_0] \qquad \text{and} \qquad \mathcal{B}_n = (z_0, \infty).$$

Following the exact same arguments as above it suffices to show that

$$\int_{\mathcal{A}_n^c} \sup_{v \in \mathcal{A}} f(|v-u|)\mathrm{d}u \leq \int_{z_0}^{\infty} \sup_{v \leq z_0 - \gamma_n/2} f(|v-u|)\mathrm{d}u \to 0$$

as $n \to \infty$ (with similar arguments for $\mathcal{B}$). As g is an upper bound of f and furthermore decreasing, we find that the integral is bounded by

$$\begin{aligned}\int_{z_0}^{\infty} \sup_{v \leq z_0 - \gamma_n/2} g(|v-u|)\mathrm{d}u &\leq \int_{z_0}^{\infty} g(u - z_0 + \gamma_n/2)\mathrm{d}u \\ &\leq \int_{\gamma_n/2}^{\infty} g(u)\mathrm{d}u \to 0\end{aligned}$$

as $n \to \infty$, which proves the claim for $d = 1$. □

Proof of Lemma D.4.9. With a slight change of notation (as compared with F_L defined in the proof of Lemma D.4.4), we now let $F_t \in \mathcal{L}_\beta$ denote the distribution of $\sup_{u \in C(v)} Z_u^{(t)}$ for $t > 0$. Hence, by (D.4.14),

$$|C_n|\overline{F}_t(x_n) \to \tau^{(t)} \tag{D.6.9}$$

as $n \to \infty$, for t large enough.

For all $z \in N_k$ and all $v, v' \in J_z$ with $|v - v'| > 2t$ we have by construction that $Z_v^{(t)}$ and $Z_{v'}^{(t)}$ are independent. Writing $y^*(v) = \sup_{u \in C(v)} y_u$ for all $v \in \mathbb{Z}^d$, turning to Lemma D.5.1, we find for all $\gamma > \beta$ that there is a constant $\tilde{C}$ such that

$$\begin{aligned}&\sum_{\substack{v < v' \in J_z \\ |v-v'|>2t}} \mathbb{P}\Big(\sup_{u \in C(v)} (Z_u^{(t)} + y_u) > x_n, \sup_{u \in C(v')} (Z_u^{(t)} + y_u) > x_n\Big) \\ &\qquad \leq \sum_{v < v' \in J_i} \overline{F}_t(x_n - y^*(v))\overline{F}_t(x_n - y^*(v')) \\ &\qquad \leq \tilde{C}\Big(\sum_{v \in J_i} \overline{F}_t(x_n)\exp(\gamma(y^*(v))^+)\Big)^2\end{aligned}$$

for sufficiently large n and all $z \in N_k$. Since the field $(Y_v)_v$ satisfies (D.5.3), and since $|C_n|/k \sim |J_z|$ as $n \to \infty$, we conclude by Corollary D.5.10 and (D.6.9) that

$$\limsup_{n \to \infty} \tilde{C}\Big(\sum_{v \in J_z} \overline{F}_t(x_n)\exp(\gamma|y^*(v)|)\Big)^2 = \frac{1}{k^2}\tilde{C}\Big(\tau^{(t)}\mathbb{E}\exp\big(\gamma(\sup_{v \in C(0)} Y_v)^+\big)\Big)^2$$

almost surely. This is independent of z and of order $o(k^{-1})$ as $k \to \infty$. This shows (D.4.20) for the terms in the sum with indices more than $2t$ apart.

Now consider $v, v' \in J_z$ such that $|v - v'| \le 2t$. As in (D.6.4), we have for all fixed $v \ne v' \in \mathbb{Z}^d$ that

$$\frac{1}{\overline{F}_t(x_n)} \mathbb{P}\Big(\sup_{u \in C(v)} Z_u^{(t)} > x_n - y, \sup_{u \in C(v')} Z_u^{(t)} > x_n - y \Big) \to 0$$

uniformly for $y \le K$, for all $K \in \mathbb{N}$. Define $y^*(v, v') = \max\{y^*(v), y^*(v')\}$ and similarly for $(Y_v)_v$ and note that

$$\mathbb{E} \exp(\gamma (Y^*(v, v'))^+) \le \mathbb{E} \exp(2\gamma (\sup_{v \in C(0)} Y_v)^+) < \infty \tag{D.6.10}$$

for all $\gamma > 0$ by the Cauchy-Schwarz inequality and (D.5.3). Arguing as in the proof of Lemma D.4.4, the uniform convergence combined with Corollary D.5.10 and the stationarity of $(Y_v)_v$ then yield

$$\begin{aligned}
&\limsup_{n\to\infty} \sum_{\substack{v < v' \in J_z \\ |v - v'| \le 2t}} \mathbb{P}\Big(\sup_{u \in C(v)} (Z_u^{(t)} + y_u) > x_n, \sup_{u \in C(v')} (Z_u^{(t)} + y_u) > x_n \Big) \\
&\quad \le \limsup_{n\to\infty} \sum_{\substack{v < v' \in J_i \\ |v - v'| \le 2t}} \mathbb{P}\Big(\sup_{u \in C(v)} Z_u^{(t)} > x_n - y^*(v, v'), \sup_{u \in C(v')} Z_u^{(t)} > x_n - y^*(v, v') \Big) \\
&\quad \le \frac{\tilde{C}}{k} \sum_{|v| \le 2t} \mathbb{E}\Big[\exp(\gamma (Y^*(0, v))^+) 1_{Y^*(0,v) > K} \Big]
\end{aligned}$$

for all $\gamma > 0$ and $K \in \mathbb{N}$, where the constant $\tilde{C}$ is chosen according to (D.5.1). Since this is independent of z, and due to the fact that there are only finitely many terms in the sum, we conclude (D.4.20) by letting $K \to \infty$ using a dominated convergence argument, which is justified by (D.6.10). □

Acknowledgments

We are grateful to Professor Gennady Samorodnitsky, Cornell University, for the very fruitful discussions and suggestions concerning Lévy–driven fields in the early phase of the project.

This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

References

[1] Barndorff-Nielsen, O.E. and J. Schmiegel (2004). Lévy-based Tempo-Spatial Modelling with Applications to Turbulence. *Uspekhi Mat. Nauk* 159, 63–90.

[2] Braverman, M. and G. Samorodnitsky (1995). Functionals of infinitely divisible stochastic processes with exponential tails. *Stoch. Process. Their Appl.* 56(2), 207–231.

[3] Cline, D.B.H. (1986). Convolution tails, product tails and domains of attraction. *Probab. Theory Relat. Fields* 72(4), 529–557.

[4] Cline, D.B.H. (1987). Convolutions of Distributions With Exponential and Subexponential Tails. *J. Aust. Math. Soc. Ser. A* 43(3), 347–365.

[5] Embrechts, P., C. Klüppelberg and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Applications of mathematics. Springer, New York.

[6] Fasen, V. (2008). Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. *Extremes* 12(3), 265–296.

[7] Hellmund, G., M. Prokešová and E.B.V. Jensen (Sept. 2008). Lévy-based Cox point processes. *Adv. Appl. Probab.* 40, 603–629.

[8] Jónsdóttir, K.Ý., A. Rønn-Nielsen, K. Mouridsen and E.B.V. Jensen (2013). Lévy-based Modelling in Brain Imaging. *Scand. Stat. Theory Appl.* 40(3), 511–529.

[9] Jónsdóttir, K.Ý., J. Schmiegel and E.B.V. Jensen (2008). Lévy-based growth models. *Bernoulli* 14(1), 62–90.

[10] Krengel, U. (1985). *Ergodic Theorems*. de Gruyter Studies in Mathematics. Walter de Gruyter, Berlin.

[11] Leadbetter, M.R., G. Lindgren and H. Rootzen (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer, New York.

[12] Leadbetter, M.R. and H. Rootzén (1998). "On Extreme Values in Stationary Random Fields". *Karatzas I., Rajput B.S., Taqqu M.S. (eds) Stochastic Processes and Related Topics*. Trends in Mathematics. Birkhäuser, Boston.

[13] Pakes, A.G. (2004). Convolution equivalence and infinite divisibility. *J. Appl. Probab.* 41, 407–424.

[14] Rajput, B.S. and J. Rosinski (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Relat. Fields* 82, 451–487.

[15] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

[16] Rønn-Nielsen, A. and E.B.V. Jensen (2016). Tail asymptotics for the supremum of an infinitely divisible field with convolution equivalent Lévy measure. *J. Appl. Probab.* 53(1), 244–261.

[17] Rønn-Nielsen, A. and E.B.V. Jensen (2017). Excursion sets of infinitely divisible random fields with convolution equivalent Lévy measure. *J. Appl. Probab.* 54(3), 833–851.

[18] Rønn-Nielsen, A. and E.B.V. Jensen (2019). Central limit theorem for mean and variogram estimators in Lévy-based models. *J. Appl. Probab.* 56(1), 209–222.

[19] Rønn-Nielsen, A., J. Sporring and E.B.V. Jensen (2017). Estimation of sample spacing in stochastic processes. *Image Anal. Stereol.* 36(1), 43–49.

[20] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.

[21] Schneider, R. (2014). *Convex Bodies: The Brunn–Minkowski Theory*. Second edition. Cambridge University Press, Cambridge.