Limit Theorems and Statistical Inference for Bessel and Dunkl Processes

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

Der Fakultät für Mathematik
der Technischen Universität Dortmund
vorgelegt von

M.Sc.-Math. Nicole Hufnagel

im November 2022

Dissertation

Limit Theorems and Statistical Inference for Bessel and Dunkl Processes

Fakultät für Mathematik

Technische Universität Dortmund

Erstgutachterin: Prof. Dr. Jeannette Woerner

Zweitgutachter: Prof. Dr. Michael Voit

Tag der mündlichen Prüfung: 5. April 2023

"Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth."

Sherlock Holmes

written by Sir Arthur Conan Doyle

Acknowledgments

First and foremost, I am grateful to my supervisor Prof. Dr. Jeannette Woerner for providing me a great environment for this thesis, for insightful discussions, for introducing me to Prof. Dr. Hiroki Masuda and for the opportunity to study abroad.

Prof. Dr. Michael Voit is not only the second reviewer of my thesis but also my second supervisor throughout my PhD. I am grateful for many discussions on questions regarding multivariate Bessel processes and pointing my attention to helpful papers.

I want to express my gratitude to JProf. Dr. Jan Nagel for helpful advice and inspiring discussions not only during our weekly seminar.

Additionally, I thank Prof. Dr. Hiroki Masuda for welcoming me in Japan, for giving me an insight on Gaussian quasi-likelihood methods and for providing me help at any time regardless the time change. It is always a pleasure to work with him.

My gratitude belongs to Prof. Dr. Sergio Andraus for a lot inspiration during long - sometimes frustrating - discussions and for keeping me motivated. We had an intensive working time in Japan, and we maintained the rhythm afterwards.

I was supported by the Deutsche Forschungsgemeinschaft through the RTG 2131 "High-Dimensional Phenomena in Probability - Fluctuations and Discontinuity" providing me an ideal environment for my PhD. This support promoted my mathematical education, gave me the opportunity to participate in conferences and enabled my stay in Japan to work with Prof. Dr. Sergio Andraus and Prof. Dr. Hiroki Masuda. I loved the inspiring discussions with other PhD students who became my friends. Especially, during the pandemic I really appreciated the coincidental meetings and discussions with Max Kämper.

A wise man said once you will never be completely done nor satisfied with your thesis. In this sense: Thank you to my former and present colleagues.

Many of my friends have accompanied me on this long journey. A special thanks goes to my fellow students Dr. Nils Dabrock, Sascha Knüttel, Till Koch and Roland Piontek who made studying mathematics and informatics into one of the most amazing experiences. Moreover, my gratitude belongs to Kathrin Held who has always supported me despite any distance.

Mama and Papa, you encouraged me to study. Thank you for your belief in me. Dennis, you are everything and more than an elder brother can be: my constant support, my friend, a role model. You showed me how much you can achieve with ambition.

Mirko, no words in this world can express what you are and do for me. Your advice and assistance - mathematically and emotionally - means the world to me. I could not have finished this thesis without you. Not today.

I dedicate this thesis to my great grandma. I will never forget your love.

Contents

Introduction

Dunkl processes gained importance in physical mathematics in the last decades, [10, 80]. However, in the field of statistics, these processes remained invisible. Statistical inference for Dunkl processes, that is, the estimation of the involved multiplicities or the dimension of a Dunkl process, has not been investigated until now. The multiplicities of a Dunkl process are of special interest since they determine the jump activity. If at least one of the multiplicities is sufficiently small, the number of jumps on any finite interval is almost surely infinite. As a first step to provide statistical inference, we consider different variations of a Dunkl process and present estimators for these variations.

In the first chapter we start with a brief overview of the essential aspects of the Dunkl theory relevant for this thesis and how a Dunkl process relates to well-known processes. We introduce Dunkl operators, which are generalizations of the directional derivatives, [27]. By using these operators to define a Dunkl process, this jump process can be viewed as a generalization of the Brownian motion generated by the Laplacian, [74, 76]. The continuous part of this process is obtained by a projection. This is the so-called radial Dunkl process, also known as the multivariate Bessel process. The Euclidean norm of a Dunkl process behaves similarly to that of a Brownian motion, since in both cases we receive a classical Bessel process. In this context, we also consider a type of polynomial processes generalizing the classical Bessel process. Both of these processes admit a stationary modification whose appropriate power turns out to be a Cox-Ingersoll-Ross process. We also discuss the relationship between the parameters of the Dunkl process and the resulting transformations.

In Chapters 2 and 3, we focus on estimators for the multiplicities of a Dunkl process and its dimension. We present well-established methods for inference when the likelihood function is unknown or too complicated. Keeping in mind that observations of a Dunkl process can be transformed into observations of other processes introduced in Chapter 1, we can consider transformations of a Dunkl process and still receive estimators for the parameters of the Dunkl process itself.

In Chapter 2, we concentrate on martingale estimators at low frequency data for the index parameter of a Bessel process and, as an extension, polynomial processes. Since these processes are non-ergodic and most results for inference are developed for stationary and ergodic diffusions, we transform them into processes with such properties by adding a mean reverting term. First, we give an overview of the known results, see [11, 59, 78], which we then apply to our modifications. We published and discussed the optimality of the resulting martingale estimators based on one eigenfunction in the case of a modified polynomial process and on up to two eigenfunctions for the modified Bessel process in [44]. In addition, we will here study the martingale estimator for any finite number of eigenfunctions and apply these results to our original Dunkl process.

Taking a closer look at the estimators from Chapter 2, we will recognize that they depend only on a particular transformation of the data which is a realization of a Cox-Ingersoll-Ross process. In Chapter 3 we hence focus on estimators for this process. A few settings of parameter estimation for the Cox-Ingersoll-Ross process were already studied. For when the process is observed continuously, the asymptotics of the maximum-likelihood estimator was analyzed in [4, 5, 71]. Considering low frequency data, that is, the distance between observations is fixed, local asymptotic normality was proved for moment-matching type estimators in [72]. Despite the popularity of the process in applications, parameter estimation for high-frequency observations has not been fully resolved and will be further explored by us. We apply the Gaussian quasi-likelihood method, another alternative to the maximum likelihood estimator when the density is unknown or too complicated. For this purpose, the density is approximated by the Gaussian density. Even if the Cox-Ingersoll-Ross process has a non-central chi-squared density which is far from being Gaussian, this local approximation works well. We introduce a preliminary estimator that is already considered in [72], then prove asymptotic normality for one-step improvements towards the Gaussian quasi-maximum likelihood estimator. We show that all these estimators are asymptotically equivalent to the Gaussian quasi-maximum likelihood estimator and compare them in a simulation study. This chapter is partially incorporated in the paper [21]. In the end, we contrast the estimators introduced in Chapter 3 and the martingale estimators from Chapter 2.

A key difference of these estimators lies in their asymptotic behaviour. The martingale estimators are consistent for every parameter value whereas the estimators introduced in Chapter 3 require the almost sure positivity of the underlying process. Turning to a different subject, we examine in Chapter 4 the times when a multivariate Bessel process hits the boundary of the Weyl chamber. In the one dimensional case, which is a classical Bessel process, the Weyl chamber represents the origin. This conflicts the positivity assumption above. Hence, we concentrate on the critical cases where the estimators of Chapter 3 do not converge, that is, when the classical Bessel process hits the origin almost surely. Since this process is already well studied, we gather known results from the literature on hitting times and infer new formulas. Then, we focus on its return times to the origin. On the one hand, this set of return times has Lebesgue measure zero almost surely, but on the other hand, its cardinality is infinite almost surely. Therefore, Luqin Liu and Yimin Xiao [61] considered the fractal Hausdorff dimension for the times when self similar processes reach the origin. In particular, they cover a classical Bessel process hitting the origin. We present these calculations and proofs pertaining to the classical Bessel process. In the end, we transfer these calculations to the Hausdorff dimension of the times when a multivariate Bessel process hits the Weyl chamber's boundary. We published this extension in the preprint [43].

We provide further introductory remarks on the various topics at the beginning of the corresponding chapter. This includes motivation and illustrations based on the underlying formulas and would therefore go beyond the scope of the introduction.

1 Preliminaries and background

1.1 Dunkl process

We begin the preliminaries with an overview over one of the central objects in this thesis, namely the Dunkl processes and then elaborate on different variations of these processes. In this section our main reference is [75] and the references therein, for example [27, 28, 29, 74]. In the whole thesis, we denote by $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$ the underlying filtered probability space. Unless otherwise specified, we regard the canonical filtration with respect to the corresponding process $(X_t)_{t\geq 0}$: $\mathcal{F}_t := \sigma\big(X_s \,|\, s \leq t\big)$.

The name of the Dunkl processes is taken from the eponymous operator originated from the field of harmonic analysis. Charles F. Dunkl has introduced Dunkl operators in his paper [27] about 30 years ago. These operators are widely used for the generalization of analytic structures. Dunkl operators play a role in various areas of mathematics and mathematical physics, [10, 80]. The Dunkl processes themselves were not treated until almost ten years later when Margit Rösler solved the heat equation with respect to the Dunkl operator, [74]. Through this discovery these operators gained relevance in stochastics by Margit Rösler and Michael Voit introducing the Dunkl processes, [76]. Since these processes find their origin in the harmonic analysis, we start with a short analytical introduction of the Dunkl operators, defining the main elements and explaining the core steps before moving on to the stochastics.

An essential component of Dunkl operators are root systems. In this regard, we consider the Euclidean space $\big(\mathbb{R}^N, \langle .,. \rangle\big)$ with the standard Euclidean inner product $\langle x, y\rangle = x_1y_1 + \cdots + x_Ny_N$ and an orthogonal reflection operator for $\alpha \in \mathbb{R}\backslash\{0\}$ defined by

$$\sigma_\alpha(x) := x - 2\frac{\langle \alpha, x\rangle}{\langle \alpha, \alpha\rangle}\alpha.$$

We occasionally omit the braces of the argument x in favour of a more concise notation. Any finite set $R \subset \mathbb{R}^N\backslash\{0\}$ which is invariant under the orthogonal reflections along its own elements, that is, $\sigma_\alpha(R) = R$ for all $\alpha \in R$, is called *root system*. In particular, $\alpha \in R$ implying $\sigma_\alpha(\alpha) = -\alpha \in R$ justifies the restriction to one half of the root system, denoted by R_+. In formulas, we decompose the root system into $R = R_+ \dot\cup (-R_+)$ separated by a hyperplane $\langle\beta\rangle^\perp$ with $\beta \notin R$ and $\langle\alpha,\beta\rangle \neq 0$ for all $\alpha \in R$. The generated group $G = G(R) := \langle\sigma_\alpha : \alpha \in R\rangle$ is called the *reflection group* associated with R. We use this reflection group to define a parameterized modification of the usual (partial) derivatives with respect to the standard basis vectors, $e_1, \ldots, e_N$.

Definition 1.1: Let $k : R \to \mathbb{C}$ be a *multiplicity function*, that is invariant under the natural action of G on R, then the *Dunkl operator* $T_i^R = T_i^R(k)$ is defined on $\mathcal{C}^1(\mathbb{R}^N)$ by

$$T_i^R f(x) := \frac{\partial}{\partial x_i} f(x) + \sum_{\alpha \in R_+} k(\alpha)\langle\alpha, e_i\rangle \frac{f(x) - f(\sigma_\alpha x)}{\langle\alpha, x\rangle}$$

for $i = 1, \ldots, N$.

Obviously, we could specify the Dunkl operators in terms of partial derivatives with respect to arbitrary basis vectors $x \in \mathbb{R}^N$ instead of the standard basis vectors but we will only consider $e_1, \ldots, e_N$ throughout this thesis. Upon closer inspection, we observe $\sigma_{c\alpha}(x) = \sigma_\alpha(x)$ for $c \in \mathbb{R}\backslash\{0\}$. Accordingly, different combinations (k, R) may result in the same operator. In order to avoid redundancy, the assumption that R is a *reduced* root system, that is, for all $\alpha \in R$ holds $R \cap \mathbb{R}\alpha = \{-\alpha, \alpha\}$, is hence reasonable. The Dunkl operator satisfies many properties that also apply to the derivative, for more details see [75]. Thus, this motivates the generalization of the well-known Laplacian Δ, the *Dunkl Laplacian* on $\mathcal{C}^2(\mathbb{R}^N)$:

$$\Delta_k^R := \sum_{i=1}^N T_i^R \circ T_i^R.$$

The explicit form

$$\Delta_k^R f(x) = \Delta f(x) + \sum_{\alpha \in R_+} k(\alpha) \left[\frac{2\langle\alpha, \nabla f(x)\rangle}{\langle\alpha, x\rangle} + \|\alpha\|^2 \frac{f(\sigma_\alpha x) - f(x)}{\langle\alpha, x\rangle^2}\right] \tag{1.1}$$

for $f \in \mathcal{C}^2(\mathbb{R}^N)$ was proved in [27, 29]. Obviously, $\Delta_0^R = \Delta$ applies. From now on, we assume non-negative real multiplicity functions, in formulas, $k(\alpha) \geq 0$ for every $\alpha \in R_+$. In the upcoming examples of possible root systems, we note that we can still define the Dunkl operators for complex multiplicities $k : R \to \mathbb{C}$, but do not pursue this for the application in stochastics. Furthermore, the corresponding Dunkl operators and Dunkl Laplacian also live on the same spaces of functions as above which means on $\mathcal{C}^1(\mathbb{R}^N)$ and $\mathcal{C}^2(\mathbb{R}^N)$, respectively.

A Brownian motion is generated by $\frac{\Delta}{2}$. This important Feller process takes a key role in financial mathematics, [13], and physics, [31]. Thus, it is evident to regard processes generated by this generalization of the Laplacian. For this, the existence of a unique solution of the heat equation for Dunkl operators, which is

$$\begin{cases} \Delta_k^R u(t,x) &= \frac{\partial}{\partial t} u(t,x), \quad (t,x) \in (0,\infty) \times \mathbb{R}^N \\ u(0,\cdot) &= f \end{cases}$$

with initial value $f \in \mathcal{C}^b(\mathbb{R}^N)$ and solution u with $u(\cdot,x) \in \mathcal{C}^1\big((0,\infty)\big)$ and $u(t,\cdot) \in \mathcal{C}^2\big(\mathbb{R}^N\big)$, is essential. For more details we refer the reader to [74]. The solution generates a contraction semigroup which provides a Markov process according to [76]. Hence, *Dunkl processes* are the càdlàg (i.e., left-limited and right-continuous) Markov processes, whose infinitesimal generator is $\frac{\Delta_k^R}{2}$ on the domain of all twice continuously differentiable functions f with $\langle \alpha, \nabla f(x) \rangle = 0$ if $\langle \alpha, x \rangle = 0$. Due to the non-singularity we require this reflection which means the probability that the generated process is at such points, $\langle \alpha, x \rangle = 0$, is zero. By observing the explicit formula (1.1), the behaviour of the associated process can already be specified. Especially, $\frac{\Delta}{2}$ generates an N dimensional Brownian motion and the expression containing ∇f generates a drift, cf. [32, Chapter 5.3 Stochastic Integral Equations]. Analyzing the last term in (1.1), we contemplate jump processes as in [32, Chapter 4.2 Markov Jump Processes and Feller Processes]. Here, E is a state space with Borel $\sigma-$algebra $\mathcal{B}(E)$. Jump processes are generally provided by an operator of the form

$$Af(x) = \lambda(x) \int\limits_E \big(f(y) - f(x)\big)\mu(x, \mathrm{d}y), \tag{1.2}$$

where μ is a transition function on $E \times \mathcal{B}(E)$ and λ a non negative function on E. For a fixed $x \in E$, $\mu(x,\cdot)$ is a probability measure specifying the possible jumps from x. Further, λ describes a function on E, which determines the probability of a jump from the state x. In the case of a Dunkl process

$$\mu(x,\cdot) = \sum_{\alpha \in R_+} c_\alpha \delta_{\sigma_\alpha x}$$

is a sum over point measures, meaning we randomly achieve an orthogonal reflection along the roots. Moreover, the jump activity depends on the multiplicity function. This relationship has already been analyzed in more detail: If $k(\alpha) > \frac{1}{2}$ applies for all $\alpha \in R_+$, the number of jumps on any finite interval is almost surely finite, cf. [22, Proposition 3.5].

For the introduction of the transition density of the Dunkl process we first need the *Dunkl kernel* $E_{k,R}(x,y)$ for $x, y \in \mathbb{R}^N$, which is the unique solution $f := E_k(\cdot, y)$ of the eigenfunction problem

$$\begin{cases} T_i^R f &= y_i f, \quad \text{for } i = 1, \dots, N, \\ f(0) &= 1. \end{cases}$$

The existence was proved in [75, 2.27 Theorem], which in turn goes back to [70, Proposition 6.7]. Using this kernel, we specify the transition density

$$p_{k,R}(t,x,y) := \frac{1}{t^{\kappa + \frac{N}{2}} c_k} e^{-\frac{\|x\|^2 + \|y\|^2}{2t}} E_{k,R}\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right) w_{k,R}(y),$$

describing the probability of going from x to $y \in \mathbb{R}^N$ after a time $t > 0$, with the sums of multiplicities

$$\kappa = \kappa(k, R) := \sum_{\alpha \in R_+} k(\alpha),$$

the weight function

$$w_R(x) = w_{k,R}(x) := \prod_{\alpha \in R} |\langle \alpha, x \rangle|^{k(\alpha)}$$

and the normalizing constant

$$c_k = c_k(R) := \int_{\mathbb{R}^N} e^{-\frac{\|x\|^2}{2}} w_{k,R}(x) \, \mathrm{d}x.$$

In this thesis, we first focus on the one dimensional case and proceed to the potentially multivariate Dunkl processes of type A_{N-1} and B_N. In the one dimensional case the root system is $R = \{-1, 1\}$ and the corresponding Dunkl operator with multiplicity parameter $k \geq 0$ is given by

$$Tf(x) = f'(x) + k \frac{f(x) - f(-x)}{x}.$$

Hence, we obtain the Dunkl Laplacian

$$\Delta_k f(x) = f''(x) + 2k\frac{f'(x)}{x} + k\frac{f(-x) - f(x)}{x^2}.$$

Accordingly, in (1.2) $\mu(x, \cdot) = \delta_{-x}$ is the point measure in $-x$ and $\lambda(x) = \frac{k}{2x^2}$. In other words, the only jumps of a one-dimensional Dunkl process are reflections and owing to the behaviour of λ, the probability for a reflection of the values close to zero is large and decreases with increasing distance to zero. Many explicit results are known for the one-dimensional Dunkl process. Here, the transition density is

$$p_k(t, x, y) = \frac{1}{t^{k+\frac{1}{2}}\Gamma(k + \frac{1}{2})} e^{-\frac{\|x\|^2+\|y\|^2}{2t}} E_k\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right) |y|^{2k}.$$

The Dunkl kernel has among others the integral representation

$$E_k(x, y) = \frac{\Gamma(k + \frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(k)} \int_{-1}^{1} e^{txy}(1 - t)^{k-1}(1 + t)^k \,\mathrm{d}t.$$

Since we can define E_k for $k \in \mathbb{C}$ as well, we remark that the condition $k \geq 0$ is essential for this expression to be valid. Furthermore, we contemplate the root system

$$A_{N-1} := \left\{ \pm (e_i - e_j) \,|\, 1 \leq i < j \leq N \right\}$$

with the standard basis vectors $e_1, \ldots, e_N$ on $\mathbb{R}^N$ and the corresponding reflections $\sigma_{ij} := \sigma_{e_i - e_j}$ exchanging the components x_i and x_j of a vector $x \in \mathbb{R}^N$. Since the transpositions (ij) generating the symmetric group S_N act like σ_{ij} on $\mathbb{R}^N$, we derive $G(A_{N-1}) = S_N$. For a fixed multiplicity parameter $k \geq 0$ the Dunkl operators are

$$T_i^{A_{N-1}} f(x) = \frac{\partial}{\partial x_i} f(x) + k \sum_{1 \leq i < j \leq N} \frac{f(x) - f(\sigma_{ij} x)}{x_i - x_j}$$

and therefore the Dunkl Laplacian is

$$\Delta_k^{A_{N-1}} f(x) = \Delta f(x) + 2k \sum_{1 \leq i < j \leq N} \left[\frac{1}{x_i - x_j} \left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j} \right) f(x) + \frac{f(\sigma_{ij} x) - f(x)}{(x_i - x_j)^2} \right].$$

In particular, the jumps of a Dunkl process in the A_{N-1} case are exchanges of two components. The transition density is given via

$$p_{k,A_{N-1}}(t,x,y) = \frac{1}{t^{k\frac{N(N-1)}{2}+\frac{N}{2}}c_k} e^{-\frac{\|x\|^2+\|y\|^2}{2t}} E_{k,A_{N-1}}\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right) w_{A_{N-1}}(y)$$

with

$$w_{A_{N-1}}(y) = \prod_{1\le i<j\le N} |y_i - y_j|^{2k}.$$

The root system B_N incorporates additionally the sign changes $\sigma_i := \sigma_{e_i}$ of the i-th component. Thus, to maintain a root system, the composition $\tau_{ij} := \sigma_{e_i+e_j} = \sigma_{ij}\sigma_i\sigma_j$ needs to be taken into account, which exchanges the ith and jth component and changes their signs. The resulting root system is

$$B_N := \big\{ \pm e_i \,|\, 1 \le i \le N \big\} \cup \big\{ \pm (e_i \pm e_j) \,|\, 1 \le i < j \le N \big\}.$$

The associated group is generated by σ_{ij} (as in the case A_{N-1}) and the sign changes σ_i, which explains why we obtain $|G(B_N)| = 2^N N!$. We regard a fixed multiplicity parameter of the form $k = (k_1, k_2) \in [0,\infty)^2$, where k_1 and k_2 correspond to the sign changes and the σ_{ij} or τ_{ij}, respectively, and obtain the associated Dunkl operators

$$T_i^{B_N} f(x) = \frac{\partial}{\partial x_i} f(x) + k_1 \frac{f(x) - f(\sigma_i x)}{x_i} + k_2 \sum_{1\le i<j\le N} \left[\frac{f(x) - f(\sigma_{ij}x)}{x_i - x_j} + \frac{f(x) - f(\tau_{ij}x)}{x_i + x_j} \right].$$

Consequently, we yield

$$\begin{aligned}\Delta_k^{B_N} f(x) = \Delta f(x) &+ 2k_1 \sum_{i=1}^{N} \frac{1}{x_i}\frac{\partial}{\partial x_i} f(x) + 2k_2 \sum_{\substack{i,j=1\\ j\neq i}}^{N} \left(\frac{1}{x_i - x_j} + \frac{1}{x_i + x_j} \right) \frac{\partial}{\partial x_i} f(x) \\ &+ k_1 \sum_{i=1}^{N} \frac{f(\sigma_i x) - f(x)}{x_i^2} + 2k_2 \sum_{1\le i<j\le N} \left[\frac{f(\sigma_{ij}x) - f(x)}{(x_i - x_j)^2} + \frac{f(\tau_{ij}x) - f(x)}{(x_i + x_j)^2} \right].\end{aligned}$$

Here, the possible jumps are determined by σ_i, σ_{ij} and τ_{ij}. The transition density is specified by

$$p_{k,B_N}(t,x,y) = \frac{1}{t^{k_1 N + k_2 N(N-1) + \frac{N}{2}} c_k} e^{-\frac{\|x\|^2+\|y\|^2}{2t}} E_{k,B_N}\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right) w_{B_N}(y)$$

with

$$\begin{aligned} w_{B_N}(y) &= \prod_{i=1}^{N} |y_i|^{2k_1} \prod_{1\le i<j\le N} |y_i - y_j|^{2k_2} |y_i + y_j|^{2k_2} \\ &= \prod_{i=1}^{N} |y_i|^{2k_1} \prod_{1\le i<j\le N} |y_i^2 - y_j^2|^{2k_2}. \end{aligned}$$

For later reference, we emphasize that the used sums of multiplicities are

$$\kappa(k,R) = \begin{cases} \kappa(k, A_{N-1}) &= \sum\limits_{1\le i<j\le N} k = \frac{kN(N-1)}{2}, \\ \kappa(k, B_N) &= \sum\limits_{i=1}^{N} k_1 + 2 \sum\limits_{1\le i<j\le N} k_2 = k_1 N + k_2 N(N-1) \end{cases}$$

in the special cases of the root systems A_{N-1} and B_N.

1.2 Multivariate Bessel process

With the introduction of Dunkl processes the *G-radial part of the Dunkl process*, which is also called *multivariate Bessel process*, immediately appears as well, [76]. This process is a continuous version of the Dunkl process and is obtained by a projection of the Dunkl process onto a proper quotient space called Weyl chamber. To explain this in more detail we need a few definitions. We assume a reduced root system R and define a *Weyl chamber* W_R to be a fixed connected component of

$$\mathbb{R}^N \setminus \bigcup_{\alpha\in R_+} \langle\alpha\rangle^\perp = \big\{x \in \mathbb{R}^N \,|\, \forall \alpha \in R_+ : \langle \alpha, x\rangle \neq 0\big\}.$$

The Weyl chamber is obviously not unique and the number of possible Weyl chambers depends on the cardinality of R since the root system is reduced. For this purpose, we briefly look at some root systems considered in this thesis. In Figure 1.1 we illustrate the hyperplanes $\langle\alpha\rangle^\perp$ associated with the roots α in the case

$$A_1 := \left\{ \pm \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

(left image) and

$$A_2 := \left\{ \pm \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \pm \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \pm \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

(right image). Each Weyl chamber here corresponds to an ordering of the components of

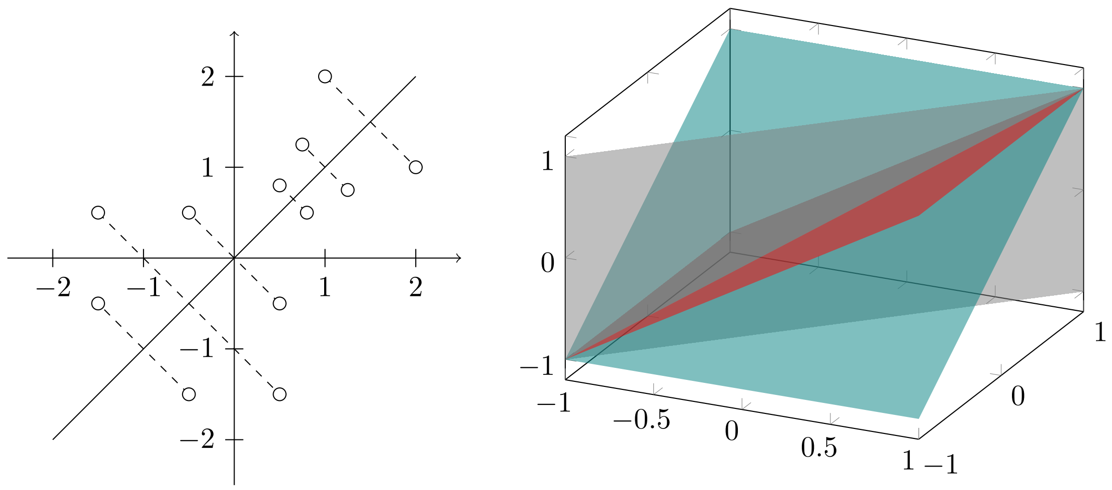

Figure 1.1: Hyperplanes belonging to the roots of A_1 (left) and A_2 (right) and examples of reflections along the roots.

vectors in $\mathbb{R}^N$. Since we incorporate the sign changes in the B_N case as opposed to the A_{N-1} case, already for $N = 2$ there exist $8 = |B_2|$ possible Weyl chambers in the B_N case in contrast to $2 = |A_1|$ in the A_{N-1} case. In general, since the root system R is reduced the number of Weyl chambers is equal to $|R|$. In the B_N case (see Figure 1.2) a Weyl chamber corresponds to an ordering of the norm of the components as well as one specific combination of their signs. Based on the reflections along the roots visualized in the figures above, we consider on $\mathbb{R}^N$ the following equivalence relation:

$$x \sim_R y :\Leftrightarrow \exists \alpha \in G(R) : \sigma_\alpha(x) = y.$$

The properties of an equivalence relation can be easily seen by using the group properties of $G(R)$. In Figures 1.1 and 1.2, we either connect points in the same equivalence class, which we denote by $[\cdot]_R$, with dashed lines or use the same symbol in the case B_1 which equals the one dimensional case. In the case A_{N-1}, the equivalence class of $x \in \mathbb{R}^N$

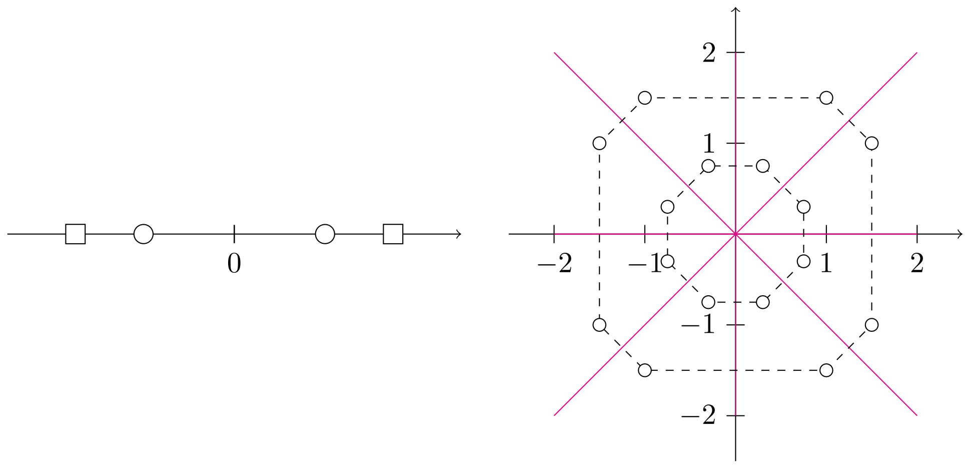

Figure 1.2: Hyperplanes belonging to the roots of B_1 (left) and B_2 (right) and examples of reflections along the roots.

contains arbitrary permutations of its components:

$$[x]_{A_{N-1}} := \{y \in \mathbb{R}^N \mid x \sim_{A_{N-1}} y\} = \Big\{y \in \mathbb{R}^N \mid \exists \tau \in S_N : \tau(x) = y\Big\}.$$

In particular, for A_1 this means that we receive two elements in each of the equivalence classes, see Figure 1.1. When $N = 1$, there exists only one reduced root system, except for multiples of the roots, $B_1 = \{-1, 1\}$. Hence, we get exactly two Weyl chambers, which are separated by the origin, and so we have $[x]_1 = \{-x, x\}$ for any $x > 0$, left-hand side in Figure 1.2. In the case B_2 we already have more roots and hence hyperplanes along which we reflect, right-hand side in Figure 1.2.

Ultimately, we immediately recognize that for the corresponding quotient space the isomorphism $\mathbb{R}^N/G(R) \cong \overline{W}_R$ holds. Now, starting with a Dunkl process $\big(\Xi_t\big)_{t\geq 0}$ associated with the reduced root system R with non-negative real multiplicity function k on R_+ the *G-radial part of the Dunkl process* or *multivariate Bessel process* is $\big(Y_t^R\big)_{t\geq 0} := \big(\pi(\Xi_t)\big)_{t\geq 0}$, the canonical projection on $\mathbb{R}^N/G(R)$. In particular, we fix one Weyl chamber W_R and choose the unique element $\pi(x) \in \overline{W}_R \cap [x]_R$. By the insights of the previous section, jumps of the Dunkl process are orthogonal reflections along its roots, so by fixing the Weyl chamber we eliminate the process's jumps and therefore its discontinuity points. Indeed, this projection of a Dunkl process is almost surely continuous and its generator is

$$\mathcal{L}_R f(x) = \frac{1}{2}\Delta f(x) + \sum_{\alpha\in R_+} k(\alpha)\frac{\langle \alpha, \nabla f(x)\rangle}{\langle \alpha, x\rangle}$$

with $f \in \mathcal{C}^2(W_R) \cap \mathcal{C}^0(\overline{W}_R)$ such that $\langle \alpha, \nabla f(x)\rangle = 0$ if $\langle \alpha, x\rangle = 0$, see [76, Theorem 4.10]. The last condition guarantees that the probability the generated process is at the boundary equals zero. Due to the non-singular boundary, we require such a reflecting boundary. Moreover, this is a Feller processes and as such it satisfies the strong Markov property. This fact follows from [76, Proposition 4.5] and [74, Theorem 4.8]. We can also specify the associated stochastic differential equation

$$\begin{cases} \mathrm{d}Y_t^R &= \mathrm{d}B_t + \sum_{\alpha\in R_+} k(\alpha)\frac{\alpha}{\langle \alpha, Y_t^R\rangle}\,\mathrm{d}t, \\ Y_0^R &= y_0 \in \overline{W}_R, \end{cases}$$

where $(B_t)_{t\geq 0}$ is a standard multivariate Brownian motion. The singularity in the drift, provided that at least one multiplicity $k(\alpha)$ is different from zero, raises the question whether this stochastic differential equation has a unique strong solution. Therefore, it is essential to rephrase this to

$$\mathrm{d}Y_t^R = \mathrm{d}B_t + \sum_{\alpha\in R_+} k(\alpha)\nabla \log\big(\langle \alpha, Y_t^R\rangle\big)\,\mathrm{d}t = \mathrm{d}B_t - \nabla\Phi\big(Y_t^R\big)\,\mathrm{d}t$$

with

$$\Phi(x) := -\sum_{\alpha\in R_+} k(\alpha)\nabla \log\big(\langle \alpha, x\rangle\big).$$

According to [26, Theorem 1] by using [24] there exists indeed a unique strong solution. In particular, this proof is an extension of [19, Theorem 3.1], where we find the proof of the case A_{N-1}, to arbitrary root systems.

Furthermore, the transition probability density

$$q_{k,R}(t,x,y) := \frac{e^{-\frac{\|x\|^2+\|y\|^2}{2t}}}{c_k t^{\frac{N}{2}}} J_{k,R}\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right) w_{k,R}\left(\frac{y}{\sqrt{t}}\right),$$

of starting in x and ending in y in the corresponding Weyl chamber after a time $t > 0$, [25, Eq. (2)], is specialized for $R \in \{A_{N-1}, B_N\}$ by choosing the corresponding forms for

$w_{k,R}$, the normalization constant c_k introduced in Section 1.1 and the *Dunkl-type Bessel function* associated with R and k defined by

$$J_{k,R}(x,y) := \frac{1}{|G(R)|} \sum_{w \in G(R)} E_k(wx, y)$$

for $x, y \in \mathbb{R}^N$. $Q_{k,R}$ denotes the corresponding distribution. For a neater notation, we omit R if the underlying root system is evident. A useful inequality

$$e^{-\|x\| \cdot \|y\|} \leq J_{k,R}(x,y) \leq e^{\|x\| \cdot \|y\|}$$

arises immediately from [75, 2.31 Proposition].

Analogous to the Dunkl process, we gain a characterization when all multiplicities are greater than $\frac{1}{2}$. In this case, the process almost surely never hits $\partial W_R := \big\{x \in \overline{W_R} \,|\, \exists \alpha \in R : \langle x, \alpha \rangle = 0\big\}$ in finite time, [19, Proposition 4.1], whereas the first time the process fulfills $\langle Y^R, \alpha \rangle = 0$ is almost surely finite for every $\alpha \in R_+$ with $k(\alpha) < \frac{1}{2}$, [26, Proposition 1].

In the following, we choose the Weyl chamber

$$W_R := \big\{x \in \mathbb{R}^N \,|\, \forall \alpha \in R_+ : \langle \alpha, x \rangle > 0\big\}.$$

Now, we take a closer look at the special cases. The one-dimensional case is the well-known classical Bessel process which justifies perceiving the G-radial part of the Dunkl process as their multivariate extension. More details on this process will be given in the next section.

As mentioned above, the unique strong solution of a multivariate Bessel process in the case A_{N-1} was proved before introducing Dunkl processes and specifically the multivariate Bessel processes for arbitrary root systems R. This particle system is furthermore equivalent to the Dyson model, which is widely studied in mathematical physics, [30]. In particular, the multivariate Bessel process of type A_{N-1} is described by its generator

$$\begin{aligned} \mathcal{L}_{A_{N-1}} f(x) &= \frac{\Delta}{2} f(x) + k \sum_{1 \leq i < j \leq N} \frac{1}{x_i - x_j} \left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j} \right) f(x) \\ &= \frac{\Delta}{2} f(x) + k \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{1}{x_i - x_j} \frac{\partial}{\partial x_i} f(x). \end{aligned}$$

We can also state the stochastic differential equation given via

$$\begin{cases} \mathrm{d}Y_{t,i}^A &= \mathrm{d}B_{t,i} + k \sum\limits_{\substack{j=1 \\ j\neq i}}^N \frac{1}{Y_{t,i}^A - Y_{t,j}^A}\,\mathrm{d}t, \\ Y_0^A &= y \in W_{A_{N-1}} \end{cases}$$

for $i = 1, \ldots, N$, where $(B_t)_{t\geq 0}$ is a standard multivariate Brownian motion. The process lives on the closure of the Weyl chamber $W_{A_{N-1}} := \{x \in \mathbb{R}^N \,|\, x_1 < \cdots < x_N\}$, which means $Y_t^A \in \overline{W}_{A_{N-1}}$ for all t.

The case B_1 obviously equals the one-dimensional case, that is, a classical Bessel process. Generally, the B_N case is component-wise equivalent to the square root of the Wishart-Laguerre process studied in mathematical physics, [17, 60]. Similarly, the multivariate Bessel process of type B_N lives on the closure of $W_{B_N} := \{x \in \mathbb{R}^N \,|\, 0 < x_1 < \cdots < x_N\}$. This process is defined by its generator

$$\mathcal{L}_{B_N} f(x) = \frac{\Delta}{2} f(x) + k_1 \sum_{i=1}^N \frac{1}{x_i} \frac{\partial}{\partial x_i} f(x) + k_2 \sum_{\substack{i,j=1 \\ j\neq i}}^N \left(\frac{1}{x_i - x_j} + \frac{1}{x_i + x_j} \right) \frac{\partial}{\partial x_i} f(x),$$

or equivalently via the stochastic differential equation

$$\begin{cases} \mathrm{d}Y_{t,i}^B &= \mathrm{d}B_{t,i} + k_1 \frac{1}{Y_{t,i}^B}\,\mathrm{d}t + k_2 \sum\limits_{\substack{j=1 \\ j\neq i}}^N \left(\frac{1}{Y_{t,i}^B - Y_{t,j}^B} + \frac{1}{Y_{t,i}^B + Y_{t,j}^B} \right) \mathrm{d}t, \\ Y_0^B &= y \in W_{B_N} \end{cases}$$

for $i = 1, \ldots, N$ with a standard multivariate Brownian motion $(B_t)_{t\geq 0}$.

1.3 Bessel process and polynomial processes

The origin of Brownian motion is in the field of botany, when Robert Brown observed the irregular movement of pollen in water nearly two centuries ago, [16]. The relevance of Brownian motion in mathematics began with the existence proof of Norbert Wiener almost one century later, [84, 85]. In this context, it was Henry P. McKean jr. who introduced the Bessel process almost 40 years later as the Euclidean norm of this process, [66]. The origin of the name comes from the eponymous functions that appear in the transition density and were named after Friedrich W. Bessel to acknowledge the groundwork, [77]. Recalling that the Brownian motion is generated by $\frac{\Delta}{2}$ and for our Dunkl Laplacian $\Delta_0 = \Delta$ holds, we could regard Dunkl processes as a generalization of the Brownian motion. The question then arises naturally whether there exists a similar result for the Euclidean norm of an arbitrary Dunkl process which was answered by Margit Rösler and Michael Voit in [76].

The *classical Bessel process of index* $\vartheta > -\frac{1}{2}$ is the Markov process with generator

$$\mathcal{L}_\vartheta f(x) = \frac{1}{2} f''(x) + \left(\vartheta + \frac{1}{2}\right) \frac{1}{x} f'(x)$$

for $f \in \mathcal{C}^2([0,\infty))$ with $f(0) = 0$. We can also specify the associated stochastic differential equation, that is,

$$\begin{cases} \mathrm{d}Y_t &= \mathrm{d}B_t + \left(\vartheta + \frac{1}{2}\right) \frac{1}{Y_t} \,\mathrm{d}t, \\ Y_0 &= y_0 > 0 \end{cases} \tag{1.3}$$

for a one dimensional Brownian motion $(B_t)_{t\geq 0}$. If $(\Xi_t)_{t\geq 0}$ is a Dunkl process, then $(Y_t)_{t\geq 0} := (\|\Xi_t\|)_{t\geq 0}$ represents a classical Bessel process of index $\vartheta = \kappa + \frac{N}{2} - 1 > -\frac{1}{2}$, cf. [76, Theorem 4.11]. Sometimes the term *classical Bessel process of dimension* d is used because the Euclidean norm of a d dimensional Brownian motion, case $\kappa = 0$ and $N = d$, is a Bessel process with index $\vartheta = \frac{d}{2} - 1 > -\frac{1}{2}$. We further observe that the projection used to obtain a multivariate Bessel process is norm-preserving, therefore $\|\pi(\Xi_t)\| = \|\Xi_t\|$ holds for every $t \geq 0$ and hence the Euclidean norm of the multivariate Bessel process is also a Bessel process with index $\vartheta = \kappa + \frac{N}{2} - 1 > -\frac{1}{2}$. There exists a unique strong solution of (1.3) based on [48, Example 8.3]. However, since in the one dimensional case the discontinuities of a Dunkl process are reflections, its π projection is identical to the Euclidean norm of the process and hence it is a classical Bessel process. Accordingly, we have already explained the existence of a unique strong solution in the previous section.

The density with respect to the Lebesgue measure of the classical Bessel process is provided by

$$q_\vartheta(t,x,y) = \frac{2}{(2t)^\vartheta \Gamma(\vartheta+1)} j_\vartheta\left(\frac{ixy}{t}\right) e^{-\frac{x^2+y^2}{2t}} y^{2\vartheta+1} \mathbb{1}_{(0,\infty)}$$

for every $x, y, t > 0$ where

$$j_\vartheta(z) := \frac{\Gamma(\vartheta+1)}{\Gamma(\vartheta+\frac{1}{2})\Gamma(\frac{1}{2})} \int_{-1}^{1} e^{isz}(1-s^2)^{\vartheta-\frac{1}{2}} \,\mathrm{d}s$$

is the spherical Bessel function of the first kind with index ϑ (see for instance [49]). Q_ϑ denotes the corresponding distribution. We notice that $(Y_t)_{t\geq 0}$ is a Feller process [73, p. 252] and hence has the strong Markov property [73, p. 102 Theorem 3.1]. Thus, $(Y_t)_{t\geq 0}$ is a time homogeneous strong Markov process on $[0,\infty)$, which is a well-known fact for the Bessel process.

By supplementing another parameter $p < 1$ we obtain an intuitive generalization of the Bessel process to a polynomial process given by the stochastic differential equation

$$\begin{cases} \mathrm{d}Y_{t,p} &= Y_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}B_t + \left(\vartheta + \frac{1}{2}\right) Y_{t,p}^p\,\mathrm{d}t, \\ Y_{0,p} &= y_0 > 0 \end{cases} \tag{1.4}$$

or equivalently through the generator

$$\mathcal{L}_{\vartheta,p} f(x) = \frac{1}{2} x^{p+1} f''(x) + \left(\vartheta + \frac{1}{2}\right) x^p f'(x)$$

for $f \in \mathcal{C}^2([0,\infty))$ with $f(0) = 0$. For $p = -1$ we receive the Bessel process. There exist other types of polynomial processes, but for our purposes we will just concentrate on these. In order to discuss the existence of a unique strong solution we reformulate the stochastic differential equation to

$$\begin{aligned} \frac{2}{1-p}\,\mathrm{d}Y_{t,p}^{\frac{1-p}{2}} &= Y_{t,p}^{-\frac{p+1}{2}}\,\mathrm{d}Y_{t,p} - \frac{1}{2}\cdot\frac{1+p}{2} Y_{t,p}^{-\frac{p+3}{2}}\,\mathrm{d}[Y_{t,p}] \\ &\overset{(1.4)}{=} \mathrm{d}B_t + \left(\vartheta + \frac{1}{2}\right) Y_{t,p}^{\frac{p-1}{2}}\,\mathrm{d}t - \frac{1+p}{4} Y_{t,p}^{-\frac{p+3}{2}} Y_{t,p}^{p+1}\,\mathrm{d}t \\ &= \mathrm{d}B_t + \left(\vartheta + \frac{1-p}{4}\right) Y_{t,p}^{\frac{p-1}{2}}\,\mathrm{d}t. \end{aligned}$$

Hence, the process $\big(\widetilde{Y}_{t,p}\big)_{t\geq 0} := \left(\frac{2}{1-p} Y_{t,p}^{\frac{1-p}{2}}\right)_{t\geq 0}$ fulfills the stochastic differential equation

$$\begin{aligned} \mathrm{d}\widetilde{Y}_{t,p} &= \mathrm{d}B_t + \left(\vartheta + \frac{1-p}{4}\right) \frac{2}{1-p} \frac{\mathrm{d}t}{\widetilde{Y}_{t,p}} \\ &= \mathrm{d}B_t + \left(\frac{2\vartheta}{1-p} + \frac{1}{2}\right) \frac{\mathrm{d}t}{\widetilde{Y}_{t,p}}, \end{aligned}$$

which is as well a classical Bessel process with index $\frac{2\vartheta}{1-p} > -\frac{1}{2}$ for every $\vartheta > \frac{p-1}{4}$. Thus, the existence of a unique solution and the strong Markov property for a Feller process transfers from the classical Bessel process to the polynomial process $(Y_{t,p})_{t\geq 0}$. This polynomial process can thus be seen as a natural generalization of the classical Bessel process, but also as a simple transformation of it.

1.4 Modified polynomial process

Since polynomial processes and in particular a classical Bessel process are non-ergodic and most results on parameter estimation for diffusions are developed for ergodic diffusions, we introduce a modification of a polynomial process which is ergodic. We consider the generator

$$\mathcal{L}_{\vartheta,\alpha,p}f(x) = \frac{1}{2}x^{p+1}f''(x) + \left[\left(\vartheta + \frac{1}{2}\right)x^p - \alpha x\right]f'(x)$$

for $f \in \mathcal{C}^2([0,\infty))$ with $f(0) = 0$ and some fixed $\alpha > 0$, $p < 1$ and the parameter $\vartheta > -\frac{1}{2}$. We can also state the stochastic differential equation

$$\begin{cases} \mathrm{d}X_{t,p} &= X_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}B_t + \left[\left(\vartheta + \frac{1}{2}\right)X_{t,p}^p - \alpha X_{t,p}\right]\,\mathrm{d}t, \\ X_{0,p} &= x_0 > 0 \end{cases} \tag{1.5}$$

where again $(B_t)_{t\geq 0}$ is a Brownian motion. The equation (1.5) is similar to the equation defining a polynomial process except for the drift term $-\alpha X_{t,p}\,\mathrm{d}t$, which we add to ensure ergodicity and stationarity. Further details on ergodicity and stationarity are given below. In particular, $(X_{t,0})_{t\geq 0}$ describes a Cox-Ingersoll-Ross process, more details on which will be given in the next section. In this paragraph we examine the essential properties of the process $(X_{t,p})_{t\geq 0}$ that we need for the analysis of our martingale estimators in Chapter 2. We keep in mind that we derive all properties also for the special case $p = -1$, the stationary version of the Bessel process. In order to accurately characterize the relationship between polynomial processes and their modifications, we consider a space time transformation

$$Y_{t,p} := f(t)X_{g(t),p} \tag{1.6}$$

with suitable functions $f, g \in \mathcal{C}^1([0,\infty))$ such that $(Y_{t,p})_{t\geq 0}$ is a polynomial process. In particular, g shall increase monotonically with $g(0) = 0$. For reconstruction of $(X_t)_{t\geq 0}$ from $(Y_t)_{t\geq 0}$ it is sufficient to claim $f(t) \neq 0$ for every $t \geq 0$. Our aim is to choose f and

g such that the stochastic differential equation

$$\mathrm{d}Y_{t,p} = Y_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}W_t + \left(\vartheta + \frac{1}{2}\right) Y_{t,p}^{p}\,\mathrm{d}t$$

holds for some Brownian motion $(W_t)_{t\geq 0}$. Using first product rule for Itô integrals leads to

$$\begin{aligned}
\mathrm{d}Y_{t,p} &\overset{(1.6)}{=} \mathrm{d}(f(t)X_{g(t),p}) = f(t)\,\mathrm{d}X_{g(t),p} + X_{g(t),p}\,\mathrm{d}f(t) + \underbrace{[f(\cdot), X_{g(\cdot),p}]_t}_{=0} \\
&\overset{(1.5)}{=} f(t)\left[X_{g(t),p}^{\frac{p+1}{2}}\,\mathrm{d}B_{g(t)} + \left(\vartheta + \frac{1}{2}\right) X_{g(t),p}^{p}\,\mathrm{d}g(t) - \alpha X_{g(t),p}\,\mathrm{d}g(t)\right] + X_{g(t)} f'(t)\,\mathrm{d}t \\
&= f(t)\sqrt{g'(t)}X_{g(t),p}^{\frac{p+1}{2}}\,\mathrm{d}W_t + \left(\vartheta + \frac{1}{2}\right) f(t)g'(t)X_{g(t),p}^{p}\,\mathrm{d}t \\
&\qquad + X_{g(t),p}\big(-\alpha f(t)g'(t) + f'(t)\big)\,\mathrm{d}t \\
&\overset{(1.6)}{=} f(t)^{1-\frac{p+1}{2}}\sqrt{g'(t)}Y_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}W_t + \left(\vartheta + \frac{1}{2}\right) f(t)^{1+p}g'(t)Y_{t,p}^{p}\,\mathrm{d}t \\
&\qquad + X_{g(t),p}\big(-\alpha f(t)g'(t) + f'(t)\big)\,\mathrm{d}t \\
&= f(t)^{\frac{1-p}{2}}\sqrt{g'(t)}Y_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}W_t + \left(\vartheta + \frac{1}{2}\right) f(t)^{1+p}g'(t)Y_{t,p}^{p}\,\mathrm{d}t \\
&\qquad + X_{g(t),p}\big(-\alpha f(t)g'(t) + f'(t)\big)\,\mathrm{d}t,
\end{aligned}$$

where $[f(\cdot), X_{g(\cdot),p}]_t$ is the covariation process of $f(\cdot)$ and $X_{g(\cdot),p}$ at time t. The validity of the third equality was proved in [68, Theorem 8.5.7 (Time change formula for Itô integrals)] whereas the covariation process $[f(\cdot), X_{g(\cdot),p}]_t$ is zero since f is of finite variation as a continuously differentiable function. To this end, we use the mean value theorem to immediately observe a zero quadratic variation

$$\begin{aligned}
\sum_{k=1}^{m} |f(t_k) - f(t_{k-1})|^2 &\leq \max_{k=1,\dots,m} \big\{|f(t_k) - f(t_{k-1})|\big\} \sum_{k=1}^{m} |f(t_k) - f(t_{k-1})| \\
&= \max_{k=1,\dots,m} \big\{|f'(\xi_k)|(t_k - t_{k-1})\big\} \sum_{k=1}^{m} \frac{|f(t_k) - f(t_{k-1})|}{t_k - t_{k-1}}(t_k - t_{k-1}) \\
&\leq \max_{s\in[0,t]} |f'(s)| \max_{k=1,\dots,m} \big\{(t_k - t_{k-1})\big\} \sum_{k=1}^{m} \frac{|f(t_k) - f(t_{k-1})|}{t_k - t_{k-1}}(t_k - t_{k-1}) \\
&\to 0 \cdot \int_0^t |f'(s)|\,\mathrm{d}s = 0
\end{aligned}$$

for some partition of the interval $[0,t]$ with mesh tending to zero and apply the inequality $[f(\cdot), X_{g(\cdot),p}]_t^2 \leq [f(\cdot)]_t \cdot [X_{g(\cdot),p}]_t$, [54, 1.5.7 Problem (iii)] afterwards. Comparing the differential equation with our preliminary considerations, we receive the following conditions:

$$\begin{cases} f^{1-p}(t)g'(t) & = & 1, \\ -\alpha f(t)g'(t) + f'(t) & = & 0, \\ g(0) & = & 0. \end{cases}$$

With a few easy arithmetic steps we solve this differential equation by

$$\begin{cases} g(t) & = & \frac{\log((1-p)\alpha t+1)}{(1-p)\alpha}, \\ f(t) & = & \sqrt[1-p]{(1-p)\alpha t + 1}. \end{cases}$$

By (1.6) we obtain

$$\begin{aligned} X_{t,p} &= \frac{1}{f(g^{-1}(t))} Y_{g^{-1}(t),p} \\ &= \exp(-\alpha t) Y_{\frac{\exp((1-p)\alpha t)-1}{(1-p)\alpha},p}. \end{aligned} \tag{1.7}$$

In the following, we check that $(X_{t,p})_{t\geq 0}$ is indeed stationary and ergodic and determine the invariant measure. We notice that, due to the singularity in the fraction of the drift divided by the squared volatility, we initially have to consider some positive interior point ξ.

Proposition 1.2: The density of the invariant probability measure with respect to the Lebesgue measure on $(0,\infty)$ is provided by

$$\mu_{\vartheta,p}(x) = \frac{1-p}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)} \left(\frac{2\alpha}{1-p}\right)^{\frac{2\vartheta+2}{1-p}} x^{2\vartheta+1} e^{-\frac{2\alpha}{1-p}x^{1-p}}$$

for $p < 1$. Therefore, $(X_{t,p})_{t\geq 0}$ is *stationary* which means if $X_{0,p} \sim \mu_{\vartheta,p}$ then $X_{t,p} \sim \mu_{\vartheta,p}$ for all $t \geq 0$.

Proof: We use [42, 9.13 Proposition] for the proof. First, we calculate the measure for a fixed $\xi \in (0,\infty)$:

$$s(x) := \exp\left(- \int_{\xi}^{x} \frac{\left(\vartheta + \frac{1}{2}\right) y^p - 2\alpha y}{y^{p+1}} \, \mathrm{d}y \right)$$

$$= \exp\left(-(2\vartheta+1)\log\left(\frac{x}{\xi}\right) + \frac{2\alpha}{1-p}\left(x^{1-p} - \xi^{1-p}\right)\right)$$
$$= \left(\frac{x}{\xi}\right)^{-(2\vartheta+1)} \exp\left(\frac{2\alpha}{1-p}\left(x^{1-p} - \xi^{1-p}\right)\right).$$

We receive the invariant probability measure by normalizing the inverse of this measure s with

$$\begin{aligned}
\int_0^\infty \frac{1}{s(x)}\,\mathrm{d}x &= \xi^{-(2\vartheta+1)} e^{\frac{2\alpha}{1-p}\xi^2} \int_0^\infty x^{2\vartheta+1} \exp\left(-\frac{2\alpha}{1-p}x^{1-p}\right)\mathrm{d}x \\
&= \xi^{-(2\vartheta+1)} e^{\frac{2\alpha}{1-p}\xi^2} \int_0^\infty \left(\frac{1-p}{2\alpha}y\right)^{\frac{2\vartheta+1}{1-p}+\frac{p}{1-p}} e^{-y}\,\frac{\mathrm{d}y}{2\alpha} \\
&= \frac{\xi^{-(2\vartheta+1)} e^{\frac{2\alpha}{1-p}\xi^2}}{1-p} \left(\frac{1-p}{2\alpha}\right)^{\frac{2\vartheta+p+1}{1-p}+1} \Gamma\left(\frac{2\vartheta+p+1}{1-p}+1\right) \\
&= \frac{\xi^{-(2\vartheta+1)} e^{\frac{2\alpha}{1-p}\xi^2}}{1-p} \left(\frac{1-p}{2\alpha}\right)^{\frac{2\vartheta+2}{1-p}} \Gamma\left(\frac{2\vartheta+2}{1-p}\right).
\end{aligned}$$

□

In the case of a polynomial process, in formulas $\alpha = 0$, the inverse of the scale measure is not normable. In particular, there cannot exist an invariant probability measure and accordingly the process is not stationary.

Definition 1.3: A stationary process $(X_t)_{t\geq 0}$ is called *ergodic* if $\frac{1}{T}\int_0^T X_t\,\mathrm{d}t$ converges to $\mathbb{E}(X_t)$ in squared mean as $T \to \infty$.

Corollary 1.4: The process $(X_{t,p})_{t\geq 0}$ is ergodic.

Proof: It suffices to show

$$\int_0^\xi s(x)\,\mathrm{d}x = \infty, \quad \int_\xi^\infty s(x)\,\mathrm{d}x = \infty \quad \text{and} \quad \int_0^\infty \frac{1}{s(x)}\,\mathrm{d}x < \infty,$$

(cf. [78, Condition 3.1] and the remarks afterwards), which is partly already proved in Proposition 1.2. We deduce

$$\begin{aligned}\int_0^\xi s(x)\,\mathrm{d}x &= \int_0^\xi \left(\frac{x}{\xi}\right)^{-(2\vartheta+1)} e^{-\frac{2\alpha}{1-p}x^{1-p}}\,\mathrm{d}x\\ &\geq e^{-\frac{2\alpha}{1-p}\xi^2}\int_0^\xi \left(\frac{x}{\xi}\right)^{-(2\vartheta+1)}\,\mathrm{d}x\\ &= \frac{e^{-\frac{2\alpha}{1-p}\xi^2}\xi^{2\vartheta+1}}{-2\vartheta}x^{-2\vartheta}\Bigg|_{x=0}^{\xi} = \infty,\\ \int_\xi^\infty s(x)\,\mathrm{d}x &= \xi^{2\vartheta+1}\int_\xi^\infty x^{-(2\vartheta+1)}e^{\frac{2\alpha}{1-p}(x^2-\xi^2)}\,\mathrm{d}x = \infty.\end{aligned}$$

□

Corollary 1.5: The $(1-p)\eta$th moment of the invariant measure is

$$\int_0^\infty x^{(1-p)\eta}\mu_{\vartheta,p}(x)\,\mathrm{d}x = \frac{\Gamma\left(\eta+\frac{2\vartheta+2}{1-p}\right)}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)}\left(\frac{1-p}{2\alpha}\right)^\eta; \quad \eta\in\mathbb{N}.$$

Proof: By a short calculation, we obtain

$$\begin{aligned}\int_0^\infty x^{(1-p)\eta}\mu_\vartheta(x)\,\mathrm{d}x &\overset{1.2}{=} \frac{1-p}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)}\left(\frac{2\alpha}{1-p}\right)^{\frac{2\vartheta+2}{1-p}}\int_0^\infty x^{(1-p)\eta+2\vartheta+1}e^{-\frac{2\alpha}{1-p}x^{1-p}}\,\mathrm{d}x\\ &= \frac{1-p}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)}\left(\frac{2\alpha}{1-p}\right)^{\frac{2\vartheta+2}{1-p}}\int_0^\infty\left(\frac{1-p}{2\alpha}y\right)^{\eta+\frac{2\vartheta+1}{1-p}+\frac{p}{1-p}}e^{-y}\frac{\mathrm{d}y}{2\alpha}\\ &= \frac{\Gamma\left(\eta+\frac{2\vartheta+1+p}{1-p}+1\right)}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)}\left(\frac{1-p}{2\alpha}\right)^{\eta+\frac{2\vartheta+1+p}{1-p}+1-\frac{2\vartheta+2}{1-p}}\\ &= \frac{\Gamma\left(\eta+\frac{2\vartheta+2}{1-p}\right)}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)}\left(\frac{1-p}{2\alpha}\right)^\eta.\end{aligned}$$

□

In particular, for a Bessel process, in formulas $p = -1$, Corollary 1.5 implies the even moments are easily computable. In this case we additionally use the space-time transformation to derive the distribution of $(X_t)_{t\geq 0} = (X_{t,-1})_{t\geq 0}$. By using the well-known distribution of the Bessel process $(Y_t)_{t\geq 0}$ we yield

$$\begin{aligned}
\mathbb{P}(X_t \leq z \,|\, X_0 = x) &\overset{(1.7)}{=} \mathbb{P}\big(Y_{\frac{\exp(2\alpha t)-1}{2\alpha}} \leq \exp(\alpha t) z \,|\, Y_0 = x\big) \\
&= \frac{2}{(2\frac{\exp(2\alpha t)-1}{2\alpha})^{\vartheta}\Gamma(\vartheta+1)} \int\limits_0^{\exp(\alpha t)z} j_\vartheta\left(\frac{ixy}{\frac{\exp(2\alpha t)-1}{2\alpha}}\right) \cdot \\
&\qquad \exp\left(-\frac{x^2+y^2}{2\frac{\exp(2\alpha t)-1}{2\alpha}}\right) y^{2\vartheta+1}\,\mathrm{d}y \\
&= \frac{2\alpha^{\vartheta}(\exp(2\alpha t))^{\vartheta+1}}{\Gamma(\vartheta+1)(\exp(2\alpha t)-1)^{\vartheta}} \int\limits_0^{z} j_\vartheta\left(ixy\frac{2\alpha\exp(\alpha t)}{\exp(2\alpha t)-1}\right) \cdot \\
&\qquad \exp\left(-\alpha\frac{x^2+y^2\exp(2\alpha t)}{\exp(2\alpha t)-1}\right) y^{2\vartheta+1}\,\mathrm{d}y \\
&= C_{\vartheta,\alpha,t}\int\limits_0^{z} j_\vartheta\left(ixy\frac{2\alpha\exp(\alpha t)}{\exp(2\alpha t)-1}\right)\exp\left(-\alpha\frac{x^2+y^2\exp(2\alpha t)}{\exp(2\alpha t)-1}\right) y^{2\vartheta+1}\,\mathrm{d}y
\end{aligned}$$

with

$$C_{\vartheta,\alpha,t} := \frac{2\alpha^{\vartheta}(\exp(2\alpha t))^{\vartheta+1}}{\Gamma(\vartheta+1)(\exp(2\alpha t)-1)^{\vartheta}}.$$

We denote the density of X_t with starting point x by $p_\vartheta(x,\cdot,t)$ and the distribution of X_t by P_ϑ. For simplicity we omit the index $p = -1$ in the case of a Bessel process.

1.5 Cox-Ingersoll-Ross process

Within financial mathematics, a Cox-Ingersoll-Ross process is commonly used to describe interest rates. This model was introduced by the mathematicians John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross in 1985, cf. [23]. In the last section we have already mentioned that the stationary version of the polynomial process $(X_{t,p})_{t\geq 0}$ is a Cox-Ingersoll-Ross process for $p = 0$. It will turn out that this process to the power $1-p$ becomes a Cox-Ingersoll-Ross process for arbitrary p. In particular, in the case of a classical Bessel process the square of the corresponding stationary version is such a process.

We assume a three dimensional parameter $\theta := (\alpha, \beta, \gamma) \subset (0, \infty)^3$, then the Cox-Ingersoll-Ross process $(Z_t)_{t\geq 0} = \left(Z_t^\theta\right)_{t\geq 0}$ is the Markov process with generator on $\mathcal{C}^2(\mathbb{R})$

$$\mathcal{L}_\theta f(x) = \frac{\gamma x}{2} f''(x) + (\alpha - \beta x) f'(x)$$

or equivalently is the solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}Z_t &= (\alpha - \beta Z_t)\,\mathrm{d}t + \sqrt{\gamma Z_t}\,\mathrm{d}B_t, \\ Z_0 &= z_0 > 0, \end{cases} \tag{1.8}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion, see among others [23] and [38]. It is known that (1.8) admits a unique strong solution, cf. [48, Example 8.2]. Due to the comparison theorem for one-dimensional diffusion processes in [73, Chapter IX. Theorem (3.7)] for $\alpha > 0$ the Cox-Ingersoll-Ross process is non-negative at any time t. Additionally, if $2\alpha > \gamma$ holds, $(Z_t)_{t\geq 0}$ stays positive almost surely, see [23]. Furthermore, the transition density with respect to the Lebesgue measure on $(0, \infty)$ is well-known, cf. [23] and [6, Eq. (1.5)], which is a noncentral chi-squared density given by

$$q_\theta^{\mathrm{CIR}}(t, x, y) = \frac{2\beta}{\gamma(1 - e^{-\beta t})} \left(\frac{y}{x e^{-\beta t}}\right)^{\frac{\nu}{2}} \exp\left(- \frac{2\beta}{\gamma} \cdot \frac{x + e^{\beta t} y}{e^{\beta t} - 1}\right) I_\nu \left(\frac{2\beta\sqrt{xy}}{\gamma \sinh\left(\frac{\beta t}{2}\right)}\right) \tag{1.9}$$

for every $x, y, t > 0$ and $\nu = \frac{2\alpha}{\gamma} - 1$ with

$$I_\nu(x) := \sum_{n=0}^{\infty} \frac{1}{n!\Gamma(\nu + n + 1)} \left(\frac{x}{2}\right)^{2n+\nu},$$

that is, the modified Bessel function of the first kind. We denote the distribution by Q_θ^{CIR}. By an analogous calculation as in the previous section in Proposition 1.2, we receive the density of the invariant probability measure with respect to the Lebesgue measure on $(0, \infty)$,

$$\pi_\theta(x) = \frac{1}{\Gamma\left(\frac{2\alpha}{\gamma}\right)} \left(\frac{2\beta}{\gamma}\right)^{\frac{2\alpha}{\gamma}} x^{\frac{2\alpha}{\gamma} - 1} e^{-\frac{2\beta}{\gamma} x},$$

see also [55, p. 22], and hence we verify the stationarity and ergodicity similar to Corollary 1.4, for which we assume $Z_0 \sim \pi_\theta$.

Corollary 1.6: The qth moment of the invariant measure is

$$\int_0^\infty x^q \pi_\theta(x)\,\mathrm{d}x = \frac{\Gamma\big(q+\frac{2\alpha}{\gamma}\big)}{\Gamma\big(\frac{2\alpha}{\gamma}\big)} \left(\frac{\gamma}{2\beta}\right)^q .$$

Proof: We can easily verify

$$\begin{aligned}
\int_0^\infty x^q \pi_\theta(x)\,\mathrm{d}x &= \frac{1}{\Gamma\big(\frac{2\alpha}{\gamma}\big)} \left(\frac{2\beta}{\gamma}\right)^{\frac{2\alpha}{\gamma}} \int_0^\infty x^{q+\frac{2\alpha}{\gamma}-1} e^{-\frac{2\beta}{\gamma}x}\,\mathrm{d}x \\
&= \frac{1}{\Gamma\big(\frac{2\alpha}{\gamma}\big)} \left(\frac{2\beta}{\gamma}\right)^{\frac{2\alpha}{\gamma}} \left(\frac{\gamma}{2\beta}\right)^{q+\frac{2\alpha}{\gamma}-1+1} \int_0^\infty x^{q+\frac{2\alpha}{\gamma}-1} e^{-x}\,\mathrm{d}x \\
&= \frac{\Gamma\big(q+\frac{2\alpha}{\gamma}\big)}{\Gamma\big(\frac{2\alpha}{\gamma}\big)} \left(\frac{\gamma}{2\beta}\right)^q .
\end{aligned}$$

□

In particular, the qth moment is finite if $q > -\frac{2\alpha}{\gamma}$.

Lemma 1.7: The conditional mean and conditional variance are

$$\begin{aligned}
\mathbb{E}\left(Z_t \mid Z_{t_0}\right) &= e^{-\beta(t-t_0)} Z_{t_0} + \frac{\alpha}{\beta}\big(1-e^{-\beta(t-t_0)}\big), \\
\mathbb{V}\mathrm{ar}\left(Z_t \mid Z_{t_0}\right) &= \frac{\gamma}{\beta}\left(1-e^{-\beta(t-t_0)}\right)\left[e^{-\beta(t-t_0)} Z_{t_0} + \frac{\alpha}{2\beta}\left(1-e^{-\beta(t-t_0)}\right)\right].
\end{aligned}$$

Proof: Applying Itô's formula

$$\begin{aligned}
\mathrm{d}(e^{\beta t} Z_t) &= e^{\beta t}\,\mathrm{d}Z_t + Z_t\,\mathrm{d}e^{\beta t} + \underbrace{\left[e^{\beta\cdot}, Z_\cdot\right]_t}_{=0} \\
&\overset{(1.8)}{=} e^{\beta t}\left[(\alpha-\beta Z_t)\,\mathrm{d}t + \sqrt{\gamma Z_t}\,\mathrm{d}B_t\right] + \beta e^{\beta t} Z_t\,\mathrm{d}t \\
&= \alpha e^{\beta t}\,\mathrm{d}t + e^{\beta s}\sqrt{\gamma Z_t}\,\mathrm{d}B_t
\end{aligned}$$

yields

$$e^{\beta t} Z_t - e^{\beta t_0} Z_{t_0} = \alpha \int_{t_0}^t e^{\beta s}\,\mathrm{d}s + \int_{t_0}^t \sqrt{\gamma Z_s}\,\mathrm{d}B_s$$

$$= \frac{\alpha}{\beta}\left(e^{\beta t} - e^{\beta t_0}\right) + \int_{t_0}^{t} e^{\beta s}\sqrt{\gamma Z_s}\,\mathrm{d}B_s. \tag{1.10}$$

Hence, we deduce

$$\begin{aligned}\mathbb{E}\left(Z_t \,|\, Z_{t_0}\right) &= e^{-\beta(t-t_0)} Z_{t_0} + \frac{\alpha}{\beta}\left(1 - e^{-\beta(t-t_0)}\right) + \sqrt{\gamma} e^{-\beta t}\,\mathbb{E}\left(\int_{t_0}^{t} e^{\beta s}\sqrt{Z_s}\,\mathrm{d}B_s \,\Big|\, Z_{t_0}\right) \\ &= e^{-\beta(t-t_0)} Z_{t_0} + \frac{\alpha}{\beta}\left(1 - e^{-\beta(t-t_0)}\right) \end{aligned} \tag{1.11}$$

almost surely since

$$\begin{aligned}\mathbb{E}\left(\int_{t_0}^{t} \left(e^{\beta s}\sqrt{Z_s}\right)^2 \mathrm{d}s\right) &= \int_{t_0}^{t} e^{2\beta s}\,\mathbb{E}\left(Z_s\right)\mathrm{d}s \\ &= \int_{t_0}^{t} e^{2\beta s}\,\mathrm{d}s \int_{0}^{\infty} x\pi_\theta(x)\,\mathrm{d}x \\ &= \frac{\left(\frac{2\beta}{\gamma}\right)^{\frac{2\alpha}{\gamma}}}{\Gamma\left(\frac{2\alpha}{\gamma}\right)} \cdot \frac{e^{2\beta t} - e^{2\beta t_0}}{2\beta} \int_{0}^{\infty} x^{\frac{2\alpha}{\gamma}} e^{-\frac{2\beta}{\gamma}x}\,\mathrm{d}x \\ &= \frac{\left(\frac{2\beta}{\gamma}\right)^{\frac{2\alpha}{\gamma}}}{\Gamma\left(\frac{2\alpha}{\gamma}\right)} \cdot \frac{e^{2\beta t} - e^{2\beta t_0}}{2\beta} \left(\frac{2\beta}{\gamma}\right)^{-\frac{2\alpha}{\gamma}-1} \Gamma\left(\frac{2\alpha}{\gamma} + 1\right) \\ &= \frac{\alpha\left(e^{2\beta t} - e^{2\beta t_0}\right)}{2\beta^2} < \infty, \end{aligned}$$

cf. [68, Theorem 3.2.1 (iii)]. Furthermore, applying Itô isometry yields

$$\begin{aligned}\mathbb{E}\left(\left(Z_t - \mathbb{E}\left[Z_t \,|\, Z_{t_0}\right]\right)^2 \,|\, Z_{t_0}\right) &\overset{(1.10)}{\underset{(1.11)}{=}} \gamma e^{-2\beta t}\,\mathbb{E}\left(\left(\int_{t_0}^{t} e^{\beta s}\sqrt{Z_s}\,\mathrm{d}B_s\right)^2 \,\Big|\, Z_{t_0}\right) \\ &= \gamma e^{-2\beta t}\,\mathbb{E}\left(\int_{t_0}^{t} e^{2\beta s} Z_s\,\mathrm{d}s \,\Big|\, Z_{t_0}\right) \\ &= \gamma e^{-2\beta t} \int_{t_0}^{t} e^{2\beta s}\,\mathbb{E}\left(Z_s \,|\, Z_{t_0}\right)\mathrm{d}s \end{aligned}$$

$$
\begin{aligned}
&\overset{(1.11)}{=} \gamma e^{-2\beta t} \int_{t_0}^{t} e^{2\beta s} \left(e^{-\beta(s-t_0)} Z_{t_0} + \frac{\alpha}{\beta} \left(1 - e^{-\beta(s-t_0)} \right) \right) \mathrm{d}s \\
&= \gamma e^{-2\beta t} \int_{t_0}^{t} e^{\beta(s+t_0)} Z_{t_0} + \frac{\alpha}{\beta} \left(e^{2\beta s} - e^{\beta(s+t_0)} \right) \mathrm{d}s \\
&= \frac{\gamma}{\beta} e^{-2\beta t} \Big[\left(e^{\beta(t+t_0)} - e^{2\beta t_0} \right) Z_{t_0} \\
&\qquad\qquad + \frac{\alpha}{2\beta} \left(e^{2\beta t} - e^{2\beta t_0} - 2e^{\beta(t+t_0)} + 2e^{2\beta t_0} \right) \Big] \\
&= \frac{\gamma}{\beta} \left[\left(e^{-\beta(t-t_0)} - e^{-2\beta(t-t_0)} \right) Z_{t_0} + \frac{\alpha}{2\beta} \left(1 - e^{-\beta(t-t_0)} \right)^2 \right] \\
&= \frac{\gamma}{\beta} \left(1 - e^{-\beta(t-t_0)} \right) \left[e^{-\beta(t-t_0)} Z_{t_0} + \frac{\alpha}{2\beta} \left(1 - e^{-\beta(t-t_0)} \right) \right].
\end{aligned}
$$

□

For later reference, we state a few basic tools for the Cox-Ingersoll-Ross process.

Lemma 1.8: We assume $2\alpha > \gamma$, which means that the origin is non-attracting, and consider equidistant times $t_j = jh_n$ for $j = 0, \ldots, n$ with $h_n > 0$, $h_n \to 0$ and $nh_n \to \infty$ as $n \to \infty$.

(i) For every $p < \frac{2\alpha}{\gamma}$, we get

$$\sup_{t\in[0,\infty)} \mathbb{E}\left(Z_t^{-p}\right) < \infty.$$

(ii) For every $p \geq 1$, there exists a constant $C_p > 0$ such that

$$\sup_{s,t\in[0,\infty):\, 0<|t-s|<1} \mathbb{E}\left(|Z_t - Z_s|^p\right) \leq C_p |t-s|^{\frac{p}{2}}.$$

(iii) For every $p < \frac{1}{2}\left(\frac{2\alpha}{\gamma} - 1\right)$, we get

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} Z_{t_{j-1}}^{-p} = \int_0^\infty z^{-p} \pi_\theta(z)\, \mathrm{d}z = \left(\frac{2\beta}{\gamma}\right)^p \frac{\Gamma\left(\frac{2\alpha}{\gamma} - p\right)}{\Gamma\left(\frac{2\alpha}{\gamma}\right)}$$

in probability.

(iv) For every $p < \frac{2\alpha}{\gamma}$ and any $\varepsilon_n > 0$ with $\lim_{n\to\infty} \varepsilon_n = 0$, we receive

$$\lim_{n\to\infty} \frac{\varepsilon_n}{n} \sum_{j=1}^{n} Z_{t_{j-1}}^{-p} = 0$$

in probability.

(v) We have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f\big(Z_{t_{j-1}}\big) = \int_0^\infty f(z)\pi_{\theta_0}(z)\,\mathrm{d}z$$

in probability for each $\mathcal{C}^1([0,\infty))$-function with f and f' being of at most polynomial growth for $x \to \infty$.

In [5] we find statement (i) as Proposition 3 and (ii) as a combination of Propositions 4 and 5. Statements (iii) and (iv) can be found in [6, Lemmas 3.1 and 3.2]. The last statement is then a conclusion in conjunction with the ergodic theorem since $(Z_t)_{t\geq 0}$ is (exponentially) strong-mixing by [35, Corollary 2.1].

Example 1.9: If we go back to the stationary modification of a polynomial process as in Section 1.4, which was introduced as the solution of

$$\mathrm{d}X_{t,p} = X_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}B_t + \left[\left(\vartheta + \frac{1}{2}\right) X_{t,p}^{p} - \alpha X_{t,p}\right] \mathrm{d}t,$$

we can easily show that $\big(X_{t,p}^{1-p}\big)_{t\geq 0}$ is a Cox-Ingersoll-Ross process. By using Itô's formula, we derive

$$\begin{aligned}
\mathrm{d}X_{t,p}^{1-p} &= (1-p)X_{t,p}^{-p}\,\mathrm{d}X_{t,p} + \frac{(1-p)(-p)}{2} X_{t,p}^{-p-1}\,\mathrm{d}[X_{\cdot,p}]_t \\
&= (1-p)X_{t,p}^{-p}\Big[\Big(\vartheta + \frac{1}{2}\Big)X_{t,p}^{p} - \alpha X_{t,p}\Big]\,\mathrm{d}t + (1-p)X_{t,p}^{\frac{p+1}{2}-p}\,\mathrm{d}B_t \\
&\qquad + \frac{(1-p)(-p)}{2} X_{t,p}^{-p-1} X_{t,p}^{p+1}\,\mathrm{d}t \\
&= \Big[\frac{1-p}{2}(2\vartheta + 1 - p) - \alpha(1-p)X_{t,p}^{1-p}\Big]\,\mathrm{d}t + (1-p)X_{t,p}^{\frac{1-p}{2}}\,\mathrm{d}B_t.
\end{aligned}$$

Therefore, $\big(X_{t,p}^{1-p}\big)_{t\geq 0}$ is a Cox-Ingersoll-Ross process with parameter

$$\theta = \left(\frac{1-p}{2}(2\vartheta + 1 - p), \alpha(1-p), (1-p)^2\right).$$

In particular, we can apply Lemma 1.7 to achieve

$$\begin{aligned}\mathbb{E}\left(X_{t,p}^{1-p}\,|\,X_{t_0,p}\right) &= e^{-\alpha(1-p)(t-t_0)}X_{t_0,p}^{1-p} + \frac{\frac{1-p}{2}(2\vartheta+1-p)}{\alpha(1-p)}\left(1-e^{-\alpha(1-p)(t-t_0)}\right)\\ &= e^{-\alpha(1-p)(t-t_0)}X_{t_0,p}^{1-p} + \frac{2\vartheta+1-p}{2\alpha}\left(1-e^{-\alpha(1-p)(t-t_0)}\right)\end{aligned}$$

and

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(X_{t,p}^{1-p}\,|\,X_{t_0,p}\right) = &\frac{(1-p)^2}{\alpha(1-p)}\left(1-e^{-\alpha(1-p)(t-t_0)}\right)\cdot\\ &\left[e^{-\alpha(1-p)(t-t_0)}X_{t_0,p}^{1-p} + \frac{\frac{1-p}{2}(2\vartheta+1-p)}{2\alpha(1-p)}\left(1-e^{-\alpha(1-p)(t-t_0)}\right)\right]\\ = &\frac{(1-p)}{\alpha}\left(1-e^{-\alpha(1-p)(t-t_0)}\right)\cdot\\ &\left[e^{-\alpha(1-p)(t-t_0)}X_{t_0,p}^{1-p} + \frac{2\vartheta+1-p}{4\alpha}\left(1-e^{-\alpha(1-p)(t-t_0)}\right)\right].\end{aligned}$$

In the case of the stationary version of a Bessel process $p=-1$, these formulas simplify to

$$\begin{aligned}\mathbb{E}\left(X_t^2\,|\,X_{t_0}\right) &= e^{-2\alpha(t-t_0)}X_{t_0}^2 + \frac{\vartheta+1}{\alpha}\left(1-e^{-2\alpha(t-t_0)}\right),\\ \mathbb{V}\mathrm{ar}\left(X_t^2\,|\,X_{t_0}\right) &= \frac{2}{\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\left[e^{-2\alpha(t-t_0)}X_{t_0}^2 + \frac{\vartheta+1}{2\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\right].\end{aligned}$$

Since we require the formula of the conditional expected value for $X_{t_0}^4$ later in Chapter 2 as well, we calculate

$$\begin{aligned}\mathbb{E}\left(X_t^4\,|\,X_{t_0}\right) &= \mathbb{V}\mathrm{ar}\left(X_t^2\,|\,X_{t_0}\right) + \mathbb{E}\left(X_t^2\,|\,X_{t_0}\right)^2\\ &= \frac{2}{\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\left[e^{-2\alpha(t-t_0)}X_{t_0}^2 + \frac{\vartheta+1}{2\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\right]\\ &\qquad + \left(e^{-2\alpha(t-t_0)}X_{t_0}^2 + \frac{\vartheta+1}{\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\right)^2\\ &= e^{-4\alpha(t-t_0)}X_{t_0}^4 + \frac{1}{\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\left[(2\vartheta+4)e^{-2\alpha(t-t_0)}X_{t_0}^2\right.\\ &\qquad\left. + \frac{(\vartheta+1)(\vartheta+2)}{2\alpha}\left(1-e^{-2\alpha(t-t_0)}\right)\right].\end{aligned}$$

2 Martingale estimation functions for the Bessel process

The content of this chapter is partially incorporated in the article

Martingale estimation functions for Bessel processes
Statistical Inference for Stochastic Processes **25**, 337–353 (2022).
Nicole Hufnagel, Jeannette H. C. Woerner.

2.1 Existing results on ergodic diffusions

In this section, we introduce the theory of martingale estimation functions studied by Michael Sørensen [78]. An *estimation function*

$$G_n(\vartheta) = G_n(\vartheta, X_{t_0}, \dots, X_{t_n})$$

depends on the parameter of interest ϑ and observations of a process $(X_t)_{t\geq 0}$ at discrete time points $t_0, \dots, t_n$. If the function G_n is additionally a martingale, we speak of a *martingale estimation function*. Our estimator itself is a solution of the equation

$$G_n(\vartheta) = 0.$$

We will discuss the existence and uniqueness of this solution later. Martingale estimation functions provide a well-established method for inference in discretely observed diffusion processes, when the likelihood function is unknown or too complicated. The idea behind martingale estimation functions is to provide a simple approximation of the true likelihood, which forms a martingale and hence under suitable regularity assumptions leads to consistent and asymptotically normal estimators. We will focus on two examples approximating the likelihood function. One way is by Taylor expansion leading to linear and quadratic martingale estimation functions, cf. Bo M. Bibby and Michael Sørensen [11]. Another possibility is to use the eigenfunctions of the associated diffusion operator, cf. Mathieu Kessler and Michael Sørensen [59].

The general setting, which we present below, is based on [78, 1.3 Martingale estimating functions]. We consider low frequency data, so in particular equidistant time observations $X_\Delta, \dots, X_{n\Delta}$ with a fixed $\Delta > 0$, of a one dimensional diffusion process defined through the stochastic differential equation

$$\mathrm{d}X_t = a(X_t, \vartheta)\,\mathrm{d}t + b(X_t, \vartheta)\,\mathrm{d}B_t, \tag{2.1}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion and $\vartheta \in \Theta \subset \mathbb{R}$. The functions a and b are given such that a weak solution exists and are assumed to be smooth enough sucht that its distribution P_ϑ is unique. The state space is $I \subset \mathbb{R}$, a not necessarily finite interval. We assume a positive transition density on I with respect to the Lebesgue measure denoted by $p_\vartheta(t, x, \cdot)$ for the density after time t conditioned on the starting point x. Furthermore, we presume $X_0 \sim \mu_\vartheta$ so that the process is ergodic with invariant probability density μ_ϑ with respect to the Lebesgue measure and hence the distribution of two consecutive observations is determined by

$$Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) = \mu_\vartheta(x) p_\vartheta(\Delta, x, y)\,\mathrm{d}x\,\mathrm{d}y. \tag{2.2}$$

In both cases that we examine, the martingale estimation function can be written in the following form:

$$G_n(\vartheta) = \sum_{i=1}^{n} g(X_{(i-1)\Delta}, X_{i\Delta}, \vartheta), \tag{2.3}$$

with a suitable function g, which will be specified later. Let the true value ϑ_0 be in the interior of Θ. For the function g we require a few conditions, which we combine from [78, Condition 1.1 and Eq. (1.17)] adapted to our setting and receive the following properties.

Condition 2.1: (i) The function $g(\cdot, \cdot, \vartheta)$ is integrable with respect to Q_Δ^ϑ for all $\vartheta \in \Theta$ and its integral with respect to Q_Δ^ϑ is equal to zero.

(ii) The function $g(x, y, \cdot)$ is continuously differentiable on Θ for all $x, y \in I$.

(iii) The function $\left|\frac{\partial}{\partial\vartheta} g(\cdot,\cdot,\vartheta)\right|$ is dominated for all $\vartheta \in \Theta$ by a function which is integrable with respect to Q_Δ^ϑ.

(iv) The integral $f(\vartheta_0) := \int\limits_I \int\limits_I \frac{\partial}{\partial\vartheta_0} g(x,y,\vartheta_0) Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y)$ has a value different from zero.

(v) The integral $v(\vartheta_0) := \int\limits_I \int\limits_I g^2(x,y,\vartheta_0) Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y)$ is finite.

In particular, due to the Markov property the first condition ensures that G_n is a martingale, while the last guarantees the applicability of the law of large numbers. These conditions enable us to state the existence of an estimator which is consistent and asymptotically normal. For more details and the proof of the following theorem, see [78, Theorem 1.5].

Theorem 2.2: Under Condition 2.1 there exists a solution of

$$G_n(\widehat{\vartheta}_n) = 0$$

with a probability tending to one as $n \to \infty$ under P_{ϑ_0} such that

(i) $\lim\limits_{n\to\infty} \widehat{\vartheta}_n = \vartheta_0$ in probability,

(ii) $\lim\limits_{n\to\infty} \sqrt{n}(\widehat{\vartheta}_n - \vartheta_0) = \mathcal{N}\left(0, \frac{v(\vartheta_0)}{f^2(\vartheta_0)}\right)$ in distribution

under P_{ϑ_0}.

As mentioned, martingale estimation functions grant simple approximations of the true likelihood function. We take a closer look at the maximum likelihood estimator itself, see [78, 1.3.2 Likelihood inference]. We assume that the transition density is differentiable with respect to ϑ. In our case, the diffusion process $(X_t)_{t\geq 0}$ is a Markov process and therefore the corresponding *log likelihood function* is denoted as

$$l_n(\vartheta) := \sum_{i=1}^n \log p_\vartheta(\Delta, X_{(i-1)\Delta}, X_{i\Delta})$$

and the estimation function, in this case called *score function*, is

$$s_n(\vartheta) := \frac{\partial}{\partial\vartheta} l_n(\vartheta)$$

$$= \sum_{i=1}^{n} \frac{\frac{\partial}{\partial\vartheta} p_\vartheta(\Delta, X_{(i-1)\Delta}, X_{i\Delta})}{p_\vartheta(\Delta, X_{(i-1)\Delta}, X_{i\Delta})}.$$

The *maximum likelihood estimator* is a solution of $s_n(\vartheta) = 0$. Now, we want to establish a connection to martingale estimators and for this we calculate

$$\mathbb{E}_\vartheta \left(\frac{\frac{\partial}{\partial\vartheta} p_\vartheta(\Delta, X_{(i-1)\Delta}, X_{i\Delta})}{p_\vartheta(\Delta, X_{(i-1)\Delta}, X_{i\Delta})} \mid \mathcal{F}_{(i-1)\Delta} \right) = \int_I \frac{\frac{\partial}{\partial\vartheta} p_\vartheta(\Delta, X_{(i-1)\Delta}, y)}{p_\vartheta(\Delta, X_{(i-1)\Delta}, y)} p_\vartheta(\Delta, X_{(i-1)\Delta}, y)\,\mathrm{d}y$$
$$= \frac{\partial}{\partial\vartheta} \int_I p_\vartheta(\Delta, X_{(i-1)\Delta}, y)\,\mathrm{d}y = 0.$$

$\mathbb{E}_\vartheta$ indicates the expectation with respect to P_ϑ and $\mathcal{F}_{(i-1)\Delta} := \sigma(X_\Delta, \ldots, X_{(i-1)\Delta})$ indicates the canonical sigma algebra of the process. In this chapter we write the index ϑ attached to the expectation to emphasize the dependence on the parameter. This is especially significant when we calculate derivatives with respect to ϑ.

Hence, the score function s_n is itself a martingale with respect to $(\mathcal{F}_{i\Delta})_{i\in\mathbb{N}}$ when the integral and derivative are interchangeable. Furthermore, the maximum likelihood estimator as a solution of $s_n(\vartheta) = 0$ is in fact already a martingale estimator and fulfills the statement in Theorem 2.2. Nevertheless, we intend to simplify the estimator by now using approximations, especially as we consider Bessel processes in Sections 2.2 and 2.4 and thus a Bessel function is contained in the density.

In the following, we introduce an estimation function based on [11] and a concise summary of the derivation. For this, we discretize the continuous-time score function and optimize it. First, we derive the likelihood function by finding an equivalent probability measure. We apply Girsanov's theorem and for this purpose we define

$$Z_\tau := \exp\left(-\int_0^\tau \frac{a(X_s, \vartheta)}{b(X_s, \vartheta)}\,\mathrm{d}B_s - \frac{1}{2}\int_0^\tau \frac{a^2(X_s, \vartheta)}{b^2(X_s, \vartheta)}\,\mathrm{d}s \right)$$

on the interval $\tau \in [0, t]$ with a given constant $0 < t \leq \infty$ and

$$\mathrm{d}Q_t := Z_t\,\mathrm{d}P_\vartheta$$

on $\mathcal{F}_t$. Assuming $(Z_\tau)_{\tau\in[0,t]}$ is a martingale with respect to $(\mathcal{F}_\tau)_{\tau\in[0,t]}$ and P_ϑ, then, according to Girsanov's theorem [68, Theorem 8.6.4], Q_t is a probability measure on $\mathcal{F}_t$. Hereafter, the likelihood function is obtained with respect to the probability measure Q_t, which enables us to derive:

$$\begin{aligned}
L_t(\vartheta) :=\ & \frac{\mathrm{d}P_\vartheta}{\mathrm{d}Q_t} = Z_t^{-1} \\
=\ & \exp\left(\int_0^t \frac{a(X_s,\vartheta)}{b(X_s,\vartheta)}\,\mathrm{d}B_s + \frac{1}{2}\int_0^t \frac{a^2(X_s,\vartheta)}{b^2(X_s,\vartheta)}\,\mathrm{d}s \right) \\
\overset{(2.1)}{=}\ & \exp\left(\int_0^t \frac{a(X_s,\vartheta)}{b(X_s,\vartheta)} \left[\frac{\mathrm{d}X_s}{b(X_s,\vartheta)} - \frac{a(X_s,\vartheta)}{b(X_s,\vartheta)}\,\mathrm{d}s \right] + \frac{1}{2}\int_0^t \frac{a^2(X_s,\vartheta)}{b^2(X_s,\vartheta)}\,\mathrm{d}s \right) \\
=\ & \exp\left(\int_0^t \frac{a(X_s,\vartheta)}{b^2(X_s,\vartheta)}\,\mathrm{d}X_s - \frac{1}{2}\int_0^t \frac{a^2(X_s,\vartheta)}{b^2(X_s,\vartheta)}\,\mathrm{d}s \right).
\end{aligned}$$

From here on, we use the same simplifications and approximations as in [11] to obtain an intuitive linear martingale estimator. First, using an Itô and Riemann sum we approximate

$$\widetilde{L}_n(\vartheta) = \exp\left(\sum_{i=1}^n \frac{a(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta},\vartheta)}(X_{i\Delta} - X_{(i-1)\Delta}) - \frac{\Delta}{2}\sum_{i=1}^n \frac{a^2(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta},\vartheta)} \right).$$

In particular, assuming b to be independent of ϑ facilitates the derivative of the approximated log likelihood function, the approximated score function:

$$\begin{aligned}
\frac{\partial}{\partial\vartheta}\log(\widetilde{L}_n(\vartheta)) = & \sum_{i=1}^n \frac{\frac{\partial}{\partial\vartheta}a(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta})}(X_{i\Delta} - X_{(i-1)\Delta}) \\
& - \Delta\sum_{i=1}^n \frac{a(X_{(i-1)\Delta},\vartheta)\frac{\partial}{\partial\vartheta}a(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta})}.
\end{aligned}$$

For simplicity, we use the same estimation function in the case of dependence on ϑ,

$$S_n(\vartheta) := \sum_{i=1}^n \frac{\frac{\partial}{\partial\vartheta}a(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta},\vartheta)}\big(X_{i\Delta} - X_{(i-1)\Delta}\big) - \Delta\sum_{i=1}^n \frac{a(X_{(i-1)\Delta},\vartheta)\frac{\partial}{\partial\vartheta}a(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta},\vartheta)}.$$

To ensure a martingale with respect to $\big(\mathcal{F}_{n\Delta}\big)_{n\in\mathbb{N}_0}$, we subtract the compensator, that is,

$$\sum_{i=1}^n \mathbb{E}_\vartheta\big(S_i(\vartheta) - S_{i-1}(\vartheta)\,|\,\mathcal{F}_{(i-1)\Delta}\big) = \sum_{i=1}^n \frac{\frac{\partial}{\partial\vartheta}a(X_{(i-1)\Delta},\vartheta)}{b^2(X_{(i-1)\Delta},\vartheta)}\big(\mathbb{E}_\vartheta(X_{i\Delta}\,|\,\mathcal{F}_{(i-1)\Delta}) - X_{(i-1)\Delta}\big)$$

$$- \Delta \sum_{i=1}^{n} \frac{a(X_{(i-1)\Delta}, \vartheta) \frac{\partial}{\partial \vartheta} a(X_{(i-1)\Delta}, \vartheta)}{b^2(X_{(i-1)\Delta}, \vartheta)}$$

and maintain

$$\begin{aligned} S_n - \sum_{i=1}^{n} \mathbb{E}_{\vartheta}\big(S_i(\vartheta) - S_{i-1}(\vartheta) \,|\, \mathcal{F}_{(i-1)\Delta}\big) \\ = \sum_{i=1}^{n} \frac{\frac{\partial}{\partial \vartheta} a(X_{(i-1)\Delta}, \vartheta)}{b^2(X_{(i-1)\Delta}, \vartheta)} \big(\mathbb{E}_{\vartheta}(X_{i\Delta} \,|\, X_{(i-1)\Delta}) - X_{(i-1)\Delta}\big). \end{aligned}$$

In order to avoid being limited by this specific weight, we consider

$$\Sigma_n(\vartheta) = \sum_{i=1}^{n} \omega(X_{(i-1)\Delta}, \vartheta) \big(\mathbb{E}_{\vartheta}(X_{i\Delta} \,|\, X_{(i-1)\Delta}) - X_{(i-1)\Delta}\big)$$

with an $\mathcal{F}_{i-1}$−measurable arbitrary weight $\omega(X_{(i-1)\Delta}, \cdot)$ which is continuously differentiable with respect to ϑ. By construction, the function Σ_n is a martingale estimation function. The optimal estimation function of this form in the sense of Vidyadhar P. Godambe and Christopher C. Heyde [37, 41], which means having the smallest asymptotic confidence interval around the true value ϑ_0 and an estimator with the smallest asymptotic dispersion, is given by

$$\omega(X_{(i-1)\Delta}, \vartheta) := \frac{\frac{\partial}{\partial \vartheta} \mathbb{E}_{\vartheta}(X_{i\Delta} \,|\, X_{(i-1)\Delta})}{\mathbb{V}\mathrm{ar}_{\vartheta}(X_{i\Delta} \,|\, X_{(i-1)\Delta})}, \tag{2.4}$$

see [11, Eq. (2.10)]. Similar to the mean, $\mathbb{V}\mathrm{ar}_{\vartheta}$ represents the variance with respect to P_{ϑ}. For small Δ the fraction in (2.4) is an approximation of $\frac{\frac{\partial}{\partial \vartheta} a(X_{(i-1)\Delta}, \vartheta)}{b^2(X_{(i-1)\Delta}, \vartheta)}$, see for more details [11, Eq. (2.11) and (2.12)].

In the case of the estimation function Σ_n, we can simplify Condition 2.1. As Σ_n is a martingale, (i) is already satisfied. The differentiability in (ii) can be inferred from ω and the conditional expectation $\mathbb{E}_{\vartheta}\big(X_{\Delta} \,|\, X_0 = x\big)$. Furthermore, by short calculations,

$$\begin{aligned} f(\vartheta) &= \int_I \int_I \frac{\partial}{\partial \vartheta}\Big(\omega(x, \vartheta)\big(y - \mathbb{E}_{\vartheta}(X_{\Delta} \,|\, X_0 = x)\big)\Big) Q_{\Delta}^{\vartheta_0}(\,\mathrm{d}x, \,\mathrm{d}y) \\ &= \int_I \int_I \Big(\frac{\partial}{\partial \vartheta}\omega(x, \vartheta)\big(y - \mathbb{E}_{\vartheta}(X_{\Delta} \,|\, X_0 = x)\big) \\ &\qquad - \omega(x, \vartheta) \frac{\partial}{\partial \vartheta} \mathbb{E}_{\vartheta}(X_{\Delta} \,|\, X_0 = x)\Big) Q_{\Delta}^{\vartheta_0}(\,\mathrm{d}x, \,\mathrm{d}y) \end{aligned}$$

$$
\begin{aligned}
&\overset{(2.2)}{=} \int_I \frac{\partial}{\partial\vartheta}\omega(x,\vartheta) \underbrace{\int_I \big(y - \mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\big) p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=0}\, \mu_\vartheta(x)\,\mathrm{d}x \\
&\quad - \int_I \underbrace{\int_I p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=1}\, \omega(x,\vartheta)\frac{\partial}{\partial\vartheta}\mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x \\
&= -\int_I \omega(x,\vartheta)\frac{\partial}{\partial\vartheta}\mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x
\end{aligned}
$$

and

$$
\begin{aligned}
v(\vartheta) &= \int_I \int_I \omega^2(x,\vartheta)\big(y - \mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\big)^2 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{(2.2)}{=} \int_I \omega^2(x,\vartheta) \int_I \big(y - \mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\big)^2 p_\vartheta(\Delta,x,y)\,\mathrm{d}y\,\mu_\vartheta(x)\,\mathrm{d}x \\
&= \int_I \omega^2(x,\vartheta)\,\mathbb{V}\mathrm{ar}_\vartheta(X_\Delta \mid X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x,
\end{aligned}
$$

we receive the desired properties such that Theorem 2.2 holds.[1] This specific case can be found in [11, Theorem 3.2].

Condition 2.3: (i) For all $x \in I$ and $\vartheta \in \Theta$, $\omega(x,\vartheta)$ and $\mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)$ are continuously differentiable with respect to ϑ.

(ii) The function

$$
\frac{\partial}{\partial\vartheta}\omega(x,\vartheta)\big(y - \mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\big) - \omega(x,\vartheta)\frac{\partial}{\partial\vartheta}\mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)
$$

is dominated in $x, y \in I$ for all $\vartheta \in \Theta$ by a function which is integrable with respect to $Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)$.

(iii) The integral

$$
f(\vartheta) = -\int_I \omega(x,\vartheta)\frac{\partial}{\partial\vartheta}\mathbb{E}_\vartheta(X_\Delta \mid X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x
$$

does not vanish at ϑ_0.

[1] The derivative always refers only to the following function otherwise the term to which the derivative refers is enclosed in brackets.

(iv) The integral

$$v(\vartheta) = \int_I \omega^2(x,\vartheta)\,\mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x$$

is finite at ϑ_0.

In the case of the discussed optimal weight (2.4), the formulas

$$\begin{aligned} f(\vartheta) &= -\int_I \omega(x,\vartheta)\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x \\ &= -\int_I \frac{\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)}{\mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)}\cdot\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x \\ &= -\int_I \frac{\left(\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)\right)^2}{\mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)}\mu_\vartheta(x)\,\mathrm{d}x \end{aligned}$$

and

$$\begin{aligned} v(\vartheta) &= \int_I \omega^2(x,\vartheta)\,\mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)\mu_{\vartheta_0}(x)\,\mathrm{d}x \\ &= \int_I \left(\frac{\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)}{\mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)}\right)^2 \mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)\mu_{\vartheta_0}(x)\,\mathrm{d}x \\ &= -f(\vartheta) \end{aligned}$$

simplify the reciprocal of the asymptotic variance

$$\begin{aligned} \frac{1}{\sigma^2(\vartheta)} = \frac{f^2(\vartheta)}{v(\vartheta)} &= \int_I \frac{\left(\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)\right)^2}{\mathbb{V}\mathrm{ar}_{\vartheta}(X_\Delta \,|\, X_0 = x)}\mu_\vartheta(x)\,\mathrm{d}x \\ &= \int_I \omega(x,\vartheta)\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\vartheta}(X_\Delta \,|\, X_0 = x)\mu_\vartheta(x)\,\mathrm{d}x. \end{aligned} \tag{2.5}$$

Now, we present an estimator based on eigenfunctions, cf. [59]. For this, we consider the corresponding generator of the diffusion

$$\mathcal{L}_{a,b}f(x) = a(x,\vartheta)f'(x) + \frac{1}{2}b^2(x,\vartheta)f''(x) \tag{2.6}$$

and search for eigenfunctions $\varphi_\eta(\cdot,\vartheta) \in \mathcal{C}^2(\mathbb{R})$ of $\mathcal{L}_{a,b}$ with eigenvalues[2] $\lambda_\eta(\vartheta)$ given via

$$\mathcal{L}_{a,b}\varphi_\eta(x,\vartheta) = -\lambda_\eta(\vartheta)\varphi_\eta(x,\vartheta).$$

For the construction of a martingale estimator, the following lemma is fundamental, cf. [59, Chapter 5. Eigenfunctions and martingales] and [78, Theorem 1.16].

Lemma 2.4: If the integral

$$\int_{\mathbb{R}} \left(\frac{\partial}{\partial x}\varphi_\eta(x,\vartheta)\right)^2 b^2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$

is finite, then

$$e^{\lambda_\eta t}\varphi_\eta(X_t,\vartheta)$$

is a martingale or, equivalently, in formulas

$$\mathbb{E}_\vartheta\big(\varphi_\eta(X_t,\vartheta)\,|\,X_s\big) = e^{-\lambda_\eta(t-s)}\varphi_\eta(X_s,\vartheta).$$

Proof: Using Itô's formula leads to

$$\begin{aligned}
\mathrm{d}\left(e^{\lambda_\eta t}\varphi_\eta(X_t,\vartheta)\right) &= \lambda_\eta e^{\lambda t}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}t + e^{\lambda_\eta t}\frac{\partial}{\partial x}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}X_t + \frac{1}{2}e^{\lambda_\eta t}\frac{\partial^2}{\partial x^2}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}[X]_t \\
&\overset{(2.6)}{=} e^{\lambda_\eta t}\Big(\lambda_\eta\varphi_\eta(X_t,\vartheta) + a(X_t,\vartheta)\frac{\partial}{\partial x}\varphi_\eta(X_t,\vartheta)\Big)\,\mathrm{d}t \\
&\quad + e^{\lambda_\eta t}b(X_t,\vartheta)\frac{\partial}{\partial x}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}B_t + \frac{1}{2}e^{\lambda_\eta t}b^2(X_t,\vartheta)\frac{\partial^2}{\partial x^2}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}t \\
&= e^{\lambda_\eta t}\underbrace{\big(\lambda\varphi_\eta(X_t,\vartheta) + \mathcal{L}_{a,b}\varphi_\eta(X_t,\vartheta)\big)}_{=0}\,\mathrm{d}t + e^{\lambda_\eta t}b(X_t,\vartheta)\frac{\partial}{\partial x}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}B_t \\
&= e^{\lambda_\eta t}b(X_t,\vartheta)\frac{\partial}{\partial x}\varphi_\eta(X_t,\vartheta)\,\mathrm{d}B_t.
\end{aligned}$$

Hence, this process is a local martingale. Then, the conclusion that this is a true martingale follows immediately by first using Burkholder inequality and Hölder inequality afterwards

$$\left[\mathbb{E}\left(\sup_{s\in[0,t]}\left|\int_0^s e^{\lambda_\eta \xi}b(X_\xi,\vartheta)\frac{\partial}{\partial x}\varphi_\eta(X_\xi,\vartheta)\,\mathrm{d}B_\xi\right|\right)\right]^2$$

[2] Technically, the eigenvalue is $-\lambda_\eta(\vartheta)$, but we omit the sign to simplify the following calculations.

$$
\begin{aligned}
&\leq C\Bigg[\mathbb{E}\Bigg(\int_0^t e^{2\lambda_\eta\xi}b^2(X_\xi,\vartheta)\Big(\frac{\partial}{\partial x}\varphi_\eta(X_\xi,\vartheta)\Big)^2 \mathrm{d}\xi\Bigg)^{\frac{1}{2}}\Bigg]^2\\
&\leq C\,\mathbb{E}\Bigg(\int_0^t e^{2\lambda_\eta\xi}b^2(X_\xi,\vartheta)\Big(\frac{\partial}{\partial x}\varphi_\eta(X_\xi,\vartheta)\Big)^2 \mathrm{d}\xi\Bigg)\\
&= C\int_0^t e^{2\lambda_\eta\xi}\,\mathrm{d}\xi\int_{\mathbb{R}} b^2(x,\vartheta)\Big(\frac{\partial}{\partial x}\varphi_\eta(x,\vartheta)\Big)^2\mu_\vartheta(x)\,\mathrm{d}x\\
&= C\frac{e^{2\lambda_\eta t}}{2\lambda_\eta}\int_{\mathbb{R}} b^2(x,\vartheta)\Big(\frac{\partial}{\partial x}\varphi_\eta(x,\vartheta)\Big)^2\mu_\vartheta(x)\,\mathrm{d}x.
\end{aligned}
$$

Since this expectation is finite due to the assumption, the martingale property holds by the dominated convergence theorem. □

For the construction of the estimation function, we assume m eigenfunctions φ_j ordered with respect to increasing eigenvalues λ_j satisfying Lemma 2.4. Therefore,

$$
\begin{aligned}
g(x,y,\vartheta) &:= \sum_{j=1}^m g_j(x,y,\vartheta)\\
&:= \sum_{j=1}^m \omega_j(x,\vartheta)\Big(\varphi_j(y,\vartheta) - \mathbb{E}_\vartheta\big(\varphi_j(X_\Delta,\vartheta)\,|\,X_0 = x\big)\Big)\\
&= \sum_{j=1}^m \omega_j(x,\vartheta)\big(\varphi_j(y,\vartheta) - e^{-\lambda_j(\vartheta)\Delta}\varphi_j(x,\vartheta)\big)
\end{aligned}
$$

leads to a martingale estimation function of the form (2.3) with arbitrary weight functions ω_j. We choose an increasing order since most information is gathered through the eigenfunctions belonging to the smallest eigenvalues. Under the assumption of Lemma 2.4 g has the desired martingale property (i) in Condition 2.1 and hence the function f simplifies to

$$
\begin{aligned}
f(\vartheta) &:= \int_I\int_I \frac{\partial}{\partial\vartheta} g(x,y,\vartheta) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)\\
&= \sum_{j=1}^m \int_I\int_I \frac{\partial}{\partial\vartheta} g_j(x,y,\vartheta) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=1}^{m} \int_I \int_I \frac{\partial}{\partial\vartheta} \Big(\omega_j(x,\vartheta)\big(\varphi_j(y,\vartheta)\big) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big)\, Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{(2.2)}{=} \sum_{j=1}^{m} \Bigg(\int_I \frac{\partial}{\partial\vartheta}\omega_j(x,\vartheta) \underbrace{\int_I \big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big) p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=0}\, \mu_\vartheta(x)\,\mathrm{d}x \\
&\qquad\qquad + \int_I \int_I \omega_j(x,\vartheta) \frac{\partial}{\partial\vartheta}\big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)\Bigg) \\
&= \sum_{j=1}^{m} \int_I \int_I \omega_j(x,\vartheta) \frac{\partial}{\partial\vartheta}\big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y).
\end{aligned}
$$

Additionally, we compute

$$
\begin{aligned}
v(\vartheta) &= \int_I \int_I g^2(x,y,\vartheta) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&= \int_I \int_I \left(\sum_{j=1}^{m} \omega_j(x,\vartheta) \Big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big) \right)^2 Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} \int_I \int_I \omega_i(x,\vartheta) \Big(\varphi_i(y,\vartheta) - e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\Big) \cdot \\
&\qquad\qquad \omega_j(x,\vartheta) \Big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{(2.2)}{=} \sum_{i=1}^{m} \sum_{j=1}^{m} \int_I \omega_j(x,\vartheta)\omega_i(x,\vartheta) \cdot \\
&\quad \underbrace{\int_I \Big(\varphi_i(y,\vartheta) - e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\Big) \Big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big) p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=:\omega_{ij}(x,\vartheta)}\, \mu_\vartheta(x)\,\mathrm{d}x \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} \int_I \omega_j(x,\vartheta)\omega_i(x,\vartheta)\omega_{ij}(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x
\end{aligned}
$$

and summarize the conditions such that Theorem 2.2 is valid for this estimator besides the requirement from Lemma 2.4. We collect the corresponding conditions, which can be found in [59, Condition 4.2], see also [59, Theorem 4.3] for the respective theorem in the case of eigenfunctions.

Condition 2.5: (i) The function $g_j(x,y,\cdot)$ is continuously differentiable on Θ for every

$x, y \in I$.

(ii) The function $\left|\frac{\partial}{\partial\vartheta} g(\cdot,\cdot,\vartheta)\right|$ is dominated for all $\vartheta \in \Theta$ by a function which is integrable with respect to Q_Δ^ϑ.

(iii) The integral

$$f(\vartheta) = \sum_{j=1}^{m} \int_I \int_I \omega_j(x,\vartheta) \frac{\partial}{\partial\vartheta} \big(\varphi_j(y,\vartheta) - e^{-\lambda_j \Delta} \varphi_j(x,\vartheta)\big) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)$$

does not vanish at ϑ_0.

(iv) The integral

$$v(\vartheta) = \sum_{i=1}^{m} \sum_{j=1}^{m} \int_I \omega_j(x,\vartheta)\omega_i(x,\vartheta)\omega_{ij}(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$

with

$$\omega_{ij}(x,\vartheta) := \int_I \Big(\varphi_i(y,\vartheta) - e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\Big)\Big(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big)\, p_\vartheta(\Delta,x,y)\,\mathrm{d}y$$

is finite at ϑ_0.

In the case of a single eigenfunction the notation

$$v(\vartheta) = \int_0^\infty \int_0^\infty g^2(x,y,\vartheta) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)$$

is more concise. We intend to discuss optimality in this case likewise. Toward this goal, we examine weights that depend on the previous observation and the true parameter,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \omega_j^\star(X_{(i-1)\Delta},\vartheta)\big(\varphi_j(X_{i\Delta},\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big).$$

The optimal weights $\omega_j^\star$ in the sense of Vidyadhar P. Godambe and Christopher C. Heyde

[37] are given in [59, p. 305], which are specified by the equation

$$\begin{pmatrix} u_{11} & \cdots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{1m} & \cdots & u_{mm} \end{pmatrix} \begin{pmatrix} \omega_1^\star \\ \vdots \\ \omega_m^\star \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \tag{2.7}$$

involving

$$u_{ij}(x,\vartheta) := \mathbb{E}_\vartheta\Big(\big(\varphi_i(X_\Delta,\vartheta) - e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\big)\big(\varphi_j(X_\Delta,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big)\,\big|\, X_0 = x\Big)$$

for $1 \leq i < j \leq m$ and

$$v_j(x,\vartheta) := -\,\mathbb{E}_\vartheta\Big(\frac{\partial}{\partial\vartheta}\big(\varphi_j(X_\Delta,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big)\,\big|\, X_0 = x\Big)$$

for $j = 1, \ldots, m$.

2.2 Martingale estimating functions based on eigenfunctions

We present in this section a light entry-level example of a non-ergodic process on which we apply the results from the previous section. Our aim is to estimate the dimensionality or index parameter $\vartheta \in \Theta \subset (-\frac{1}{2}, \infty)$ of a classical Bessel process specified in Section 1.3 via the stochastic differential equation

$$\begin{cases} \mathrm{d}Y_t &= \mathrm{d}B_t + \big(\vartheta + \frac{1}{2}\big)\,\frac{1}{Y_t}\,\mathrm{d}t, \\ Y_0 &= y_0 > 0, \end{cases}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion. We proceed similarly to [11] and [59], introduced in Section 2.1, to construct martingale estimation functions for our parameter of interest ϑ. Since a Bessel process is non-ergodic, we transform it into a stationary and ergodic process by adding a mean reverting term with speed of mean reversion $\alpha > 0$ in the drift, that is,

$$\begin{cases} \mathrm{d}X_t &= \mathrm{d}B_t + \Big[\big(\vartheta + \frac{1}{2}\big)\,\frac{1}{X_t} - \alpha X_t\Big]\,\mathrm{d}t, \\ X_0 &= y_0 > 0. \end{cases}$$

We call this process a modified Bessel process, for more details see Section 1.4. We assume $X_0 \sim \mu_\vartheta$ and observe the modified Bessel process at discrete times, that is, $X_\Delta, \ldots, X_{n\Delta}$.

Lemma 2.6: The eigenfunctions of the generator

$$\mathcal{L}_{\vartheta,\alpha} f(x) = \frac{1}{2} f''(x) + \left[\left(\vartheta + \frac{1}{2} \right) \frac{1}{x} - \alpha x \right] f'(x),$$

which are the solutions of $\mathcal{L}_{\vartheta,\alpha}\varphi_\eta = -\lambda_\eta \varphi_\eta$ are given by

$$\lambda_\eta = 2\alpha\eta, \quad \varphi_\eta(x,\vartheta) = \sum_{l=0}^{\eta} \frac{(-\eta)_l}{(\vartheta+1)_l l!} (\alpha x^2)^l, \quad \eta \in \mathbb{N},$$

with the Pochhammer symbols $(x)_0 := 1$ and $(x)_l := \frac{\Gamma(x+l)}{\Gamma(x)} = x(x+1)\dots(x+l-1)$ for $l \in \mathbb{N}$.

Proof: First, we calculate the image of the monomials

$$\begin{aligned}
\mathcal{L}_{\vartheta,\alpha} 1 &= 0, \\
\mathcal{L}_{\vartheta,\alpha} x &= \left(\vartheta + \frac{1}{2} \right) \frac{1}{x} - \alpha x, \\
\mathcal{L}_{\vartheta,\alpha} x^l &= \frac{1}{2} l(l-1) x^{l-2} + l \left[\left(\vartheta + \frac{1}{2} \right) \frac{1}{x} - \alpha x \right] x^{l-1} \\
&= \frac{l(l+2\vartheta)}{2} x^{l-2} - \alpha l x^l
\end{aligned}$$

for $l \geq 2$. Due to the last line, we require the eigenfunctions to be the sum of at least two monomials with difference two in the degree. Furthermore, we observe in the second line that the coefficient belonging to x shall be zero. In combination with the previous condition, the eigenfunctions thus consist of the sum of even monomials. We assume

$$\varphi_\eta(x) := \sum_{l=0}^{\eta} c_l x^{2l}$$

with $c_0 = 1$ and compute

$$\begin{aligned}
\mathcal{L}_{\vartheta,\alpha} \varphi_\eta(x) &= \sum_{l=1}^{\eta} c_l \left(\frac{2l(2l+2\eta)}{2} x^{2l-2} - 2\alpha l x^{2l} \right) \\
&= \sum_{l=0}^{\eta} 2(l+1)(l+1+\vartheta) c_{l+1} x^{2l} - \sum_{l=1}^{\eta} 2\alpha l c_l x^{2l} \\
&= c_1(2\vartheta+2) + \sum_{l=1}^{\eta} \big(2(l+1)(l+1+\vartheta) c_{l+1} - 2\alpha l c_l \big) x^{2l} - 2\alpha\eta c_\eta x^{2\eta}
\end{aligned}$$

$$\stackrel{!}{=} -\lambda_\eta \sum_{l=0}^{\eta} c_l x^{2l}.$$

Equating the coefficients of x^0 and $x^{2\eta}$ directly yields

$$c_1(2\vartheta + 2) = -\lambda_\eta,$$
$$2\alpha\eta = \lambda_\eta,$$

or $c_1 = -\frac{\alpha\eta}{\vartheta+1}$, respectively, while for the coefficient of x^{2l} ensues

$$\begin{aligned} 2(l+1)(l+1+\vartheta)c_{l+1} - 2\alpha l c_l &= -\lambda_\eta c_l \\ &= -2\alpha\eta c_l \end{aligned}$$

for all $l = 1, \dots, \eta - 1$. The result follows from solving the equation by c_{l+1} and successive substitution

$$c_{l+1} = c_l \frac{\alpha(-\eta + l)}{(l+1)(\vartheta + 1 + l)} = \cdots = \frac{\alpha^{l+1}(-\eta)_{l+1}}{(l+1)!(\vartheta+1)_{l+1}}.$$

□

Remark: The eigenfunctions of this Lemma coincide with the even eigenfunctions of a modified Dunkl process of dimension one. We consider the following eigenfunction problem

$$\widetilde{\Delta}_k \psi_\nu := \left(\frac{\Delta_k}{2} - \sum_{i=1}^{d} x_i \frac{\partial}{\partial x_i} \right) \psi_\nu = -\lambda_\nu \psi_\nu.$$

We call the process generated by $\widetilde{\Delta}_k$ a modified Dunkl process. This equation is solved by

$$\lambda_{\nu,k} = |\nu|, \quad \psi_\nu = H_\nu^k,$$

cf. [75, 2.58 Corollary (i)], for $\nu \in \mathbb{N}^d$ and the generalized Hermite polynomials H_ν^k. For more detailed information on these polynomials the reader is referred to [81]. The extent to which these two results are related can be seen if we examine the one-dimensional

Dunkl process. In this case, there exist simple explicit formulas

$$\begin{cases} H^k_{2\eta}(x) &= (-1)^\eta 2^{2\eta}\eta! L^{k-\frac{1}{2}}_\eta(x^2), \\ H^k_{2\eta+1}(x) &= (-1)^\eta 2^{2\eta+1}\eta! L^{k+\frac{1}{2}}_\eta(x^2) \end{cases}$$

for $\eta \in \mathbb{N}$ with the generalized Laguerre polynomials

$$L^p_\eta(x) = \frac{1}{\eta!} x^{-p} e^x \frac{\partial^\eta}{\partial x^\eta}\left(x^{\eta+p}e^{-x}\right)$$

for $p > 0$. From the Laguerre polynomials

$$\begin{aligned} L^p_0(x) &= 1, \\ L^p_1(x) &= -x + p + 1, \\ L^p_2(x) &= \frac{1}{2}\left[x^2 - 2(p+2)x + (p+1)(p+2)\right], \end{aligned}$$

we thus derive the first Hermite polynomials

$$\begin{aligned} H^k_0(x) &= L^{k-\frac{1}{2}}_0(x^2) = 1, \\ H^k_1(x) &= 2xL^{k+\frac{1}{2}}_0(x^2) = 2x, \\ H^k_2(x) &= -1 \cdot 2^2 \cdot 1! L^{k-\frac{1}{2}}_1(x^2) = -4\left[-x^2 + k + \frac{1}{2}\right] \\ &= 4\left[x^2 - k - \frac{1}{2}\right], \\ H^k_3(x) &= -1 \cdot 2^3 \cdot 2! x L^{k+\frac{1}{2}}_1(x^2) = 8\left[x^3 - \left(k + \frac{3}{2}\right)x\right], \\ H^k_4(x) &= 2^4 \cdot 2! L^{k-\frac{1}{2}}_2(x^2) = \frac{32}{2}\left[x^4 - 2\left(k+\frac{3}{2}\right)x^2 + \left(k+\frac{1}{2}\right)\left(k+\frac{3}{2}\right)\right]. \end{aligned}$$

We compare now this result with the eigenfunctions of the modified Bessel process by setting $\alpha = 1$,

$$\begin{aligned} \varphi_1(x,\vartheta) &= 1 - \frac{x^2}{\vartheta+1} = -\frac{1}{\vartheta+1}\left[x^2 - (\vartheta+1)\right] \\ &= -\frac{1}{4(\vartheta+1)} H^{\vartheta+\frac{1}{2}}_2(x), \\ \varphi_2(x,\vartheta) &= 1 - 2\frac{x^2}{\vartheta+1} + \frac{x^4}{(\vartheta+1)(\vartheta+2)}(x) \end{aligned}$$

$$= \frac{1}{16(\vartheta+1)(\vartheta+2)} \left[x^4 - (2\vartheta+4)x^2 + (\vartheta+1)(\vartheta+2)\right]$$
$$= \frac{1}{(\vartheta+1)(\vartheta+2)} H_4^{\vartheta+\frac{1}{2}}(x)$$

and the corresponding eigenvalues $\lambda_1 = \lambda_{2,\vartheta+\frac{1}{2}} = 2$ and $\lambda_2 = \lambda_{4,\vartheta+\frac{1}{2}} = 4$. This connection is not surprising, since the Euclidean norm of a Dunkl process equals a classical Bessel process. In the case of a modified Dunkl process, we preserve as eigenfunctions not only the sum of even monomials. Omitting the jumps produces a loss of information, thus we presumably can obtain a better estimator for the Dunkl process by examining the Hermite polynomials instead of φ_η.

We return to the modified Bessel process. According to Lemma 2.4, the property

$$\int_0^\infty \left(\frac{\partial}{\partial x}\varphi_\eta(x,\vartheta)\right)^2 \mu_\vartheta(\,\mathrm{d}x) = \frac{2\alpha^{\vartheta+1}}{\Gamma(\vartheta+1)} \int_0^\infty \left(\frac{\partial}{\partial x}\varphi_\eta(x,\vartheta)\right)^2 x^{2\vartheta+1} e^{-\alpha x^2}\,\mathrm{d}x < \infty$$

for the polynomials φ_η is sufficient to deduce

$$\mathrm{E}_\vartheta\big(\varphi_\eta(X_{i\Delta},\vartheta)|X_{(i-1)\Delta}\big) = e^{-\lambda_\eta\Delta}\varphi_\eta(X_{(i-1)\Delta},\vartheta).$$

Consequently, we may use the general theory on estimators based on eigenfunctions introduced in Section 2.1. However, in our case we will calculate the involved quantities and obtain explicit results. For the first eigenfunction $\varphi_1(x,\vartheta) = 1 - \frac{\alpha x^2}{\vartheta+1}$ we consider the estimator based on the martingale estimation function

$$\begin{aligned} G_n(\vartheta) &= \sum_{i=1}^n \big(\varphi_1(X_{i\Delta},\vartheta) - e^{-\lambda_1\Delta}\varphi_1(X_{(i-1)\Delta},\vartheta)\big) \\ &= \sum_{i=1}^n \left(1 - \frac{\alpha X_{i\Delta}^2}{\vartheta+1} - e^{-2\alpha\Delta}\left(1 - \frac{\alpha X_{(i-1)\Delta}^2}{\vartheta+1}\right)\right) \\ &= n(1-e^{-2\alpha\Delta}) - \frac{\alpha}{\vartheta+1}\sum_{i=1}^n \left(X_{i\Delta}^2 - e^{-2\alpha\Delta}X_{(i-1)\Delta}^2\right). \end{aligned}$$

The unique solution of $G_n\big(\widehat{\vartheta}_n\big) = 0$ is

$$\widehat{\vartheta}_n = \frac{\alpha \sum_{i=1}^n \left(X_{i\Delta}^2 - e^{-2\alpha\Delta}X_{(i-1)\Delta}^2\right)}{n(1-e^{-2\alpha\Delta})} - 1. \tag{2.8}$$

Now, we may deduce consistency and asymptotic normality along the same lines as for general martingale estimation functions.

Theorem 2.7: For every true value $\vartheta_0 \in \Theta \subset (-\frac{1}{2}, \infty)$, we have

(i) $\lim\limits_{n\to\infty} \widehat{\vartheta}_n = \vartheta_0$ in probability and

(ii) $\lim\limits_{n\to\infty} \sqrt{n}\big(\widehat{\vartheta}_n - \vartheta_0\big) = \mathcal{N}\big(0, \sigma^2(\vartheta_0)\big)$ in distribution

under P_{ϑ_0} with $\sigma^2(\vartheta_0) := (\vartheta_0 + 1)\frac{1+e^{-2\alpha\Delta}}{1-e^{-2\alpha\Delta}}$.

Proof: The proof is straightforward. We first validate the assumptions in Condition 2.5. For the calculation of the asymptotic variance we will need the symmetric distribution Q_Δ^ϑ of two consecutive observations $X_{(i-1)\Delta}$ and $X_{i\Delta}$ on $(0,\infty)^2$. It is given by

$$\begin{aligned} Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) &= \mu_\vartheta(x) p_\vartheta(\Delta, x, y)\,\mathrm{d}x\,\mathrm{d}y \\ &= C_\vartheta j_\vartheta\left(ixy\frac{2\alpha\exp(\alpha\Delta)}{\exp(2\alpha\Delta)-1}\right)\exp\left(-\frac{\alpha\exp(2\alpha\Delta)}{\exp(2\alpha\Delta)-1}(x^2+y^2)\right)(xy)^{2\vartheta+1}\,\mathrm{d}y\,\mathrm{d}x \end{aligned}$$

with

$$C_\vartheta := \frac{4\alpha^{2\vartheta}(\exp(2\alpha\Delta))^{\vartheta+1}}{\Gamma(\vartheta+1)^2(\exp(2\alpha\Delta)-1)^\vartheta}.$$

We define

$$\begin{aligned} g(x,y,\vartheta) &:= \varphi_1(y,\vartheta) - e^{-\lambda_1\Delta}\varphi_1(x,\vartheta) \\ &= 1 - \frac{\alpha y^2}{\vartheta+1} - e^{-2\alpha\Delta}\left(1 - \frac{\alpha x^2}{\vartheta+1}\right) \end{aligned}$$

a continuously differentiable function with respect to ϑ. The absolute value of the derivative

$$\frac{\partial}{\partial\vartheta} g(x,y,\vartheta) = \frac{\alpha}{(\vartheta+1)^2}(y^2 - e^{-2\alpha\Delta}x^2)$$

is dominated by $4\alpha(y^2 + e^{-2\alpha\Delta}x^2)$, which is independent of ϑ and square integrable with respect to $Q_\Delta^{\vartheta_0}$. Moreover, the symmetry in x and y of the density of $Q_\Delta^{\vartheta_0}$ implies

$$f(\vartheta_0) := \int_0^\infty\int_0^\infty \frac{\partial}{\partial\vartheta} g(x,y,\vartheta_0) Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y)$$

$$= \underbrace{\frac{\alpha}{(\vartheta_0+1)^2}(1-e^{-2\alpha\Delta})}_{>0} \underbrace{\int_0^\infty \int_0^\infty x^2 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y)}_{>0} \neq 0.$$

Owing to the exponential function in Q_Δ^ϑ and $g(\cdot,\cdot,\vartheta)$ being a polynomial, the integral

$$v(\vartheta_0) := \int_0^\infty \int_0^\infty g^2(x,y,\vartheta_0) Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y)$$

is finite, which completes the proof of (i) and (ii) according to Theorem 2.2 and Condition 2.5. Therefore, we are left with the task of computing $\sigma^2(\vartheta_0)$. Due to Theorem 2.2, the asymptotic variance is given by $\sigma^2(\vartheta_0) = \frac{v(\vartheta_0)}{f^2(\vartheta_0)}$. Because of

$$\begin{aligned}
g^2(x,y,\vartheta) &= \left(\left(1-e^{-2\alpha\Delta}\right) - \frac{\alpha}{\vartheta+1}y^2 + \frac{\alpha}{\vartheta+1}e^{-2\alpha\Delta}x^2 \right)^2 \\
&= (1-e^{-2\alpha\Delta})^2 + \frac{\alpha^2}{(\vartheta+1)^2}y^4 + \frac{\alpha^2 e^{-4\alpha\Delta}}{(\vartheta+1)^2}x^4 - (1-e^{-2\alpha\Delta})\frac{2\alpha}{\vartheta+1}y^2 \\
&\qquad + (1-e^{-2\alpha\Delta})\frac{2\alpha e^{-2\alpha\Delta}}{\vartheta+1}x^2 - \frac{2\alpha^2 e^{-2\alpha\Delta}}{(\vartheta+1)^2}x^2y^2 \\
&= (1-e^{-2\alpha\Delta})^2 - (1-e^{-2\alpha\Delta})\frac{2\alpha}{\vartheta+1}\left(y^2 - e^{-2\alpha\Delta}x^2\right) \\
&\qquad + \frac{\alpha^2}{(\vartheta+1)^2}\left(y^4 + e^{-4\alpha\Delta}x^4\right) - \frac{2\alpha^2 e^{-2\alpha\Delta}}{(\vartheta+1)^2}x^2y^2
\end{aligned}$$

and the symmetry of $Q_\Delta^{\vartheta_0}$, we get

$$\begin{aligned}
v(\vartheta_0) = (1-e^{-2\alpha\Delta})^2 &\left(1 - \frac{2\alpha}{\vartheta_0+1}\int_0^\infty \int_0^\infty x^2 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y)\right) \\
&+ \frac{\alpha^2(1+e^{-4\alpha\Delta})}{(\vartheta_0+1)^2}\int_0^\infty \int_0^\infty x^4 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y) \\
&- \frac{2\alpha^2 e^{-2\alpha\Delta}}{(\vartheta_0+1)^2}\int_0^\infty \int_0^\infty x^2y^2 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y).
\end{aligned}$$

Furthermore, we derived in Corollary 1.5

$$\begin{aligned}
\int_0^\infty \int_0^\infty x^{2\eta} Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y) &= \int_0^\infty \int_0^\infty x^{2\eta} \mu_{\vartheta_0}(x) p_\vartheta(\Delta, x, y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty x^{2\eta} \underbrace{\int_0^\infty p_\vartheta(\Delta, x, y)\,\mathrm{d}y}_{=1}\, \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&= \int_0^\infty x^{2\eta} \mu_{\vartheta_0}(x)\,\mathrm{d}x \overset{1.5}{=} \frac{\Gamma(\eta + \vartheta_0 + 1)}{\alpha^\eta \Gamma(\vartheta_0 + 1)} \\
&= \frac{(\vartheta_0 + 1)_\eta}{\alpha^\eta}.
\end{aligned}$$

Additionally using the explicit formula of $\mathbb{E}_{\vartheta_0}(X_{i\Delta}^2 \,|\, X_{(i-1)\Delta} = x)$ in Example 1.9, we conclude

$$\begin{aligned}
\int_0^\infty \int_0^\infty x^2 y^2 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y) &= \int_0^\infty \int_0^\infty x^2 y^2 \mu_{\vartheta_0}(x) p_\vartheta(\Delta, x, y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_0^\infty x^2 \int_0^\infty y^2 p_\vartheta(\Delta, x, y)\,\mathrm{d}y\; \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&= \int_0^\infty x^2\, \mathbb{E}_{\vartheta_0}(X_{i\Delta}^2 \,|\, X_{(i-1)\Delta} = x) \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&\overset{1.9}{=} \int_0^\infty x^2 \left(x^2 e^{-2\alpha\Delta} - \frac{\vartheta_0 + 1}{\alpha}(e^{-2\alpha\Delta} - 1) \right) \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&= \frac{(\vartheta_0 + 1)(\vartheta_0 + 2)}{\alpha^2} e^{-2\alpha\Delta} - \frac{(\vartheta_0 + 1)^2}{\alpha^2}(e^{-2\alpha\Delta} - 1) \\
&= \frac{(\vartheta_0 + 1)^2}{\alpha^2} + e^{-2\alpha\Delta} \frac{\vartheta_0 + 1}{\alpha^2}.
\end{aligned}$$

Applying these formulas, we establish

$$\begin{aligned}
f(\vartheta_0) &= \frac{\alpha}{(\vartheta_0 + 1)^2}(1 - e^{-2\alpha\Delta}) \int_0^\infty \int_0^\infty x^2 Q_\Delta^{\vartheta_0}(\,\mathrm{d}x,\,\mathrm{d}y) \\
&= \frac{1 - e^{-2\alpha\Delta}}{\vartheta_0 + 1}
\end{aligned}$$

and

$$\begin{aligned} v(\vartheta_0) &= \left(1-e^{-2\alpha\Delta}\right)^2\left(1-\frac{2\alpha}{\vartheta_0+1}\cdot\frac{\vartheta_0+1}{\alpha}\right)+\frac{\alpha^2(1+e^{-4\alpha\Delta})}{(\vartheta_0+1)^2}\cdot\frac{(\vartheta_0+1)(\vartheta_0+2)}{\alpha^2} \\ &\qquad -\frac{2\alpha^2e^{-2\alpha\Delta}}{(\vartheta_0+1)^2}\left(\frac{(\vartheta_0+1)^2}{\alpha^2}+e^{-2\alpha\Delta}\frac{\vartheta_0+1}{\alpha^2}\right) \\ &= -\left(1-e^{-2\alpha\Delta}\right)^2+\frac{\vartheta_0+2}{\vartheta_0+1}\left(1+e^{-4\alpha\Delta}\right)-2e^{-2\alpha\Delta}-e^{-4\alpha\Delta}\frac{2}{\vartheta_0+1} \\ &= \left(-1+\cancel{2e^{-2\alpha\Delta}}-e^{-4\alpha\Delta}\right)+\frac{\vartheta_0+2}{\vartheta_0+1}+\frac{\vartheta_0}{\vartheta_0+1}e^{-4\alpha\Delta}-\cancel{2e^{-2\alpha\Delta}} \\ &= \frac{1-e^{-4\alpha\Delta}}{\vartheta_0+1}=\frac{(1-e^{-2\alpha\Delta})(1+e^{-2\alpha\Delta})}{\vartheta_0+1}. \end{aligned}$$

Hence, we infer

$$\sigma^2(\vartheta_0)=\frac{v(\vartheta_0)}{f^2(\vartheta_0)}=(\vartheta_0+1)\frac{1+e^{-2\alpha\Delta}}{1-e^{-2\alpha\Delta}}.$$

□

Let us discuss the results. Looking at the asymptotic variance, we see that it decreases when $\alpha\Delta$ is increasing. This seems surprising at first glance, since it implies that the asymptotic variance decreases when the distance between observations increases, as we keep the mean reverting parameter α fixed. Notice, that we have the observation scheme $X_\Delta, \cdots, X_{n\Delta}$, hence $n \to \infty$ and $\Delta \to 0$ such that $n\Delta \to \infty$ would correspond to continuous observations. However, keeping in mind that equidistant observations for the stationary version of the Bessel process $(X_t)_{t\geq 0}$ mean that the distance between two observations of the underlying Bessel process $(Y_t)_{t\geq 0}$ is exponentially growing

$$X_t = e^{-\alpha t} Y_{\frac{e^{2\alpha t}-1}{2\alpha}},$$

which leads to a fast growing observation interval. This might capture the non-stationary behaviour of the original Bessel process. Furthermore, we see that the asymptotic variance tends to infinity as the mean reverting parameter tends to zero.

Having a closer look at the estimator, we see that it only depends on the square of the observations, hence we could reformulate our problem and consider the squared process $Z_t := X_t^2$. As explained in Example 1.9, this results in

$$\mathrm{d}Z_t = 2\sqrt{Z_t}\,\mathrm{d}B_t + (2\vartheta + 2 - 2\alpha Z_t)\,\mathrm{d}t,$$

an equation describing a Cox-Ingersoll-Ross process. We consider now the canonical linear martingale estimation function

$$\begin{aligned}\Sigma_n(\vartheta) &:= \sum_{i=1}^{n} \big(Z_{i\Delta} - \mathrm{E}(Z_{i\Delta}|Z_{(i-1)\Delta})\big)\\ &= \sum_{i=1}^{n} \Big(Z_{i\Delta} - Z_{(i-1)\Delta}e^{-2\alpha\Delta} + \frac{\vartheta+1}{\alpha}(e^{-2\alpha\Delta}-1)\Big)\\ &= -\frac{\vartheta+1}{\alpha}G_n(\vartheta).\end{aligned}$$

For $\vartheta > -\frac{1}{2}$, the unique solution of $\Sigma_n\big(\widehat{\vartheta}_n\big) = 0$ is again

$$\widehat{\vartheta}_n = \frac{\alpha \sum\limits_{i=1}^{n} \Big(X_{i\Delta}^2 - X_{(i-1)\Delta}^2 e^{-2\alpha\Delta}\Big)}{n(1-e^{-2\alpha\Delta})} - 1.$$

Hence, we see that the two estimators coincide. In Theorem 2.7 we have already established the consistency and asymptotic normality of $\widehat{\vartheta}_n$.

The next step is to increase the flexibility of Σ_n by adding the weight ω depending on the parameter of interest and the previous observation

$$\sum_{i=1}^{n} \omega(\vartheta, X_{(i-1)\Delta}) \left(X_{i\Delta}^2 - X_{(i-1)\Delta}^2 e^{-2\alpha\Delta} + \frac{\vartheta+1}{\alpha}\big(e^{-2\alpha\Delta}-1\big)\right),$$

where $\omega(\cdot, X_{(i-1)\Delta})$ is $\sigma(X_\Delta, \ldots, X_{(i-1)\Delta})-$measurable to keep the martingale property and continuously differentiable to apply our method. Using the same technique, we search for the optimal estimator with the smallest asymptotic variance. Considering this second approach via linear martingale estimation functions for the squared process allows us to easily determine this optimal estimator. By (2.4) the optimal weight is given by

$$\begin{aligned}\omega(\vartheta, X_{(i-1)\Delta}) &:= \frac{\frac{\partial}{\partial\vartheta}\,\mathbb{E}_\vartheta(X_{i\Delta}^2 \,|\, X_{(i-1)\Delta})}{\mathbb{V}\mathrm{ar}_\vartheta(X_{i\Delta}^2 \,|\, X_{(i-1)\Delta})}\\ &\overset{1.9}{=} \frac{\frac{1-e^{-2\alpha\Delta}}{\alpha}}{\frac{2}{\alpha}\,(1-e^{-2\alpha\Delta})\,\big(X_{i\Delta}^2 + \frac{\vartheta+1}{2\alpha}\,(1-e^{-2\alpha\Delta})\big)}\\ &= \frac{1}{2X_{(i-1)\Delta}^2 e^{-2\alpha\Delta} + \frac{\vartheta+1}{\alpha}(1-e^{-2\alpha\Delta})}.\end{aligned}$$

Unfortunately, the equation defining the optimal estimator

$$\sum_{i=1}^{n} \frac{1}{2X_{(i-1)\Delta}^2 e^{-2\alpha\Delta} + \frac{\vartheta+1}{\alpha}(1-e^{-2\alpha\Delta})} \cdot \left(X_{i\Delta}^2 - X_{(i-1)\Delta}^2 e^{-2\alpha\Delta} + \frac{\vartheta+1}{\alpha}(e^{-2\alpha\Delta}-1) \right) = 0$$

is not explicitly solvable with respect to ϑ. However, we can nevertheless determine the improvement in the asymptotic variance. Therefore, we have to establish the finiteness of

$$\begin{aligned}
&\int_0^\infty \omega(\vartheta_0, X_{(i-1)\Delta}) \frac{\partial}{\partial\vartheta_0} \mathbb{E}_{\vartheta_0}(X_{i\Delta}^2 \,|\, X_{(i-1)\Delta}) \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&= \int_0^\infty \frac{1}{2X_{(i-1)\Delta}^2 e^{-2\alpha\Delta} + \frac{\vartheta_0+1}{\alpha}(1-e^{-2\alpha\Delta})} \cdot \frac{1-e^{-2\alpha\Delta}}{\alpha} \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&= \int_0^\infty \frac{1}{\frac{2\alpha e^{-2\alpha\Delta}}{1-e^{-2\alpha\Delta}} X_{(i-1)\Delta}^2 + \vartheta_0 + 1} \mu_{\vartheta_0}(x)\,\mathrm{d}x \\
&< \int_0^\infty \frac{1}{\vartheta_0+1} \mu_{\vartheta_0}(x)\,\mathrm{d}x = \frac{1}{\vartheta_0+1},
\end{aligned}$$

the reciprocal of the asymptotic variance, that is, the asymptotic information, cf. (2.5). Consequently, we can deduce that a lower bound of the optimal variance is given by ϑ_0+1.

Figure 2.1 shows the asymptotic information of the 10.000 simulated optimal estimator (triangles) and $\widehat{\vartheta}_n$ (dots) for $n = 1.000$. The solid line corresponds to the calculated asymptotic information of $\widehat{\vartheta}_n$ in Theorem 2.7. The dotted line represents our computed bound above. As the lines nearly touch around $\Delta = 3$, the improvement of the optimal estimator quickly tends to zero. Starting from the value $\Delta = 1$, the simulated asymptotic information is almost the same for both estimators. Beforehand, the improvement is clearly visible but we do not want to maintain such a high variance as we can choose the value of $\alpha\Delta$ so that the asymptotic variance is close to the lower bound.

We take a closer look at the asymptotic variance of $\widehat{\vartheta}_n$ from Theorem 2.7, which decreases

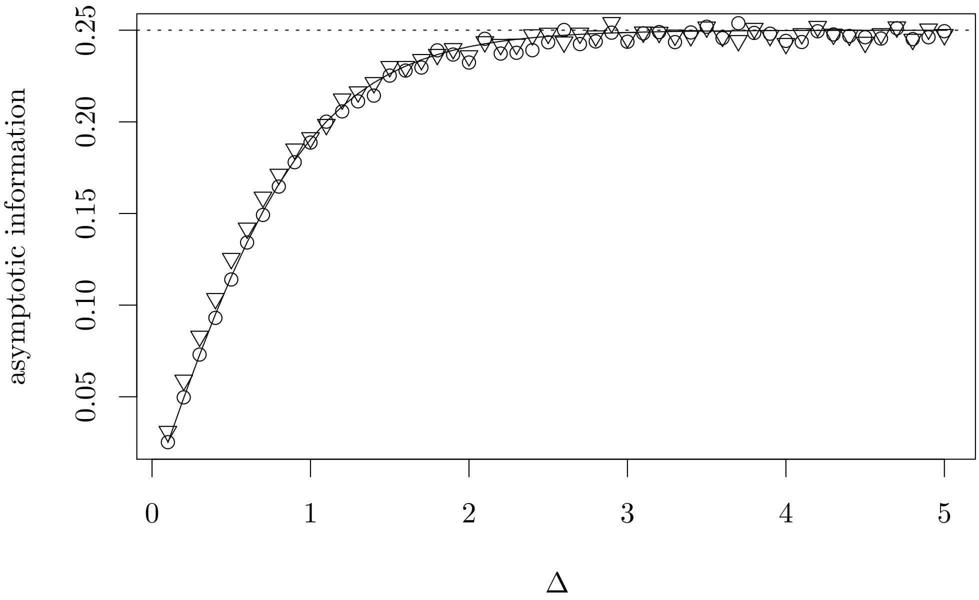

Figure 2.1: The asymptotic behaviour for $\alpha = 1, x_0 = 0.1, \vartheta_0 = 3$.

monotonously in $\alpha\Delta$:

$$\lim_{\alpha\Delta\to\infty} (\vartheta_0 + 1)\frac{1 + e^{-2\alpha\Delta}}{1 - e^{-2\alpha\Delta}} = \vartheta_0 + 1.$$

Due to the fast convergence to the lower bound $\vartheta_0 + 1$, we can for practical purposes restrict ourselves to the estimator $\widehat{\vartheta}_n$ and hence have an explicit estimator.

We can transfer this result to the Dunkl process. Previously, we established that the eigenfunctions φ_i coincide with those of a modified version of the one-dimensional Dunkl process. Now, we consider a potentially multidimensional Dunkl process $(\Xi_t)_{t\geq 0}$. Since its Euclidean norm is a classical Bessel process of index $\vartheta = \kappa + \frac{N}{2} - 1$ we can give an estimator of $\kappa \in K \subset [0, \infty)$, that is, the sum of the multiplicities, assuming the dimension N is known. For this reason, we know that $\left(e^{-\alpha t}\left\|\Xi_{\frac{\exp(2\alpha t)}{2\alpha}}\right\|\right)_{t\geq 0}$ is the corresponding stationary version of the classical Bessel process, see (1.7). We define

$$\begin{aligned}\widehat{\kappa}_n &:= \frac{\alpha \sum_{i=1}^{n} \left(e^{-2\alpha i \Delta} \left\| \Xi_{\frac{\exp(2\alpha i\Delta)}{2\alpha}} \right\|^2 - e^{-2\alpha(i-1)\Delta} \left\| \Xi_{\frac{\exp(2\alpha(i-1)\Delta)}{2\alpha}} \right\|^2 e^{-2\alpha\Delta} \right)}{n(1-e^{-2\alpha\Delta})} - \frac{N}{2} \\ &= \frac{\alpha \sum_{i=1}^{n} \left(e^{-2\alpha i \Delta} \left\| \Xi_{\frac{\exp(2\alpha i\Delta)}{2\alpha}} \right\|^2 - e^{-2\alpha(i+1)\Delta} \left\| \Xi_{\frac{\exp(2\alpha(i-1)\Delta)}{2\alpha}} \right\|^2 \right)}{n(1-e^{-2\alpha\Delta})} - \frac{N}{2}\end{aligned}$$

analogously to (2.8) and obtain the following corollary from Theorem 2.7.

Corollary 2.8: For every true value $\kappa_0 \in K \subset [0,\infty)$, we have

(i) $\lim_{n\to\infty} \widehat{\kappa}_n = \kappa_0$ in probability and

(ii) $\lim_{n\to\infty} \sqrt{n}\big(\widehat{\kappa}_n - \kappa_0\big) = \mathcal{N}\big(0, \sigma^2(\kappa_0)\big)$ in distribution

under $P_{k_0,R}$ with $\sigma^2(\kappa_0) := \big(\kappa_0 + \frac{N}{2}\big)\frac{1+e^{-2\alpha\Delta}}{1-e^{-2\alpha\Delta}}$ where k_0 denotes the corresponding true multiplicity function.

The asymptotic variance increases with the dimension of the Dunkl process so that the estimate becomes less accurate. We analyze this corollary for specific root systems. First of all, in the one dimensional case κ is equal to the multiplicity k, which we then estimate directly. In the A_{N-1} case, $\kappa(k, A_{N-1}) = \frac{kN(N-1)}{2}$ holds and hence $\frac{2\widehat{\kappa}_n}{N(N-1)}$ estimates the multiplicity.

In the B_N case we have two multiplicities which means we cannot say anything about the specific values since $\kappa(k, B_N) = k_1 N + k_2 N(N-1)$ holds.

Conversely, if we assume κ to be known, we can estimate N by considering the estimator:

$$\widehat{N}_n := \frac{2\alpha \sum_{i=1}^{n} \left(e^{-2\alpha i \Delta} \left\| \Xi_{\frac{\exp(2\alpha i\Delta)}{2\alpha}} \right\|^2 - e^{-2\alpha(i+1)\Delta} \left\| \Xi_{\frac{\exp(2\alpha(i-1)\Delta)}{2\alpha}} \right\|^2 \right)}{n(1-e^{-2\alpha\Delta})} - 2\kappa.$$

As before, the parameter N is obtained via an affine linear transformation of ϑ and hence the asymptotic behaviour of $\widehat{N}_n$ is inherited from $\widehat{\vartheta}_n$.

Corollary 2.9: For every true value $N_0 \in \mathbb{N}$, we have

(i) $\lim_{n\to\infty} \widehat{N}_n = N_0$ in probability and

(ii) $\lim_{n\to\infty} \sqrt{n}\big(\widehat{N}_n - N_0\big) = \mathcal{N}\big(0, \sigma^2(N_0)\big)$ in distribution

under $P_{k,R}$ with $\sigma^2(N_0) := 4\big(\kappa + \frac{N_0}{2}\big)\frac{1+e^{-2\alpha\Delta}}{1-e^{-2\alpha\Delta}}$.

2.3 An extension to some polynomial processes

In the previous section, we introduced an idea of performing a martingale estimating function while using an ergodic transformation of a non-ergodic process, the Bessel process. In particular, this seems like a straightforward method to deal with non-ergodic processes. The question arises whether this can find applications for other processes. We aim to extend the developed technique to some larger class of processes. We consider some non-ergodic polynomial processes solving the stochastic differential equation

$$\begin{cases} \mathrm{d}Y_{t,p} &= Y_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}B_t + \big(\vartheta + \frac{1}{2}\big)\,Y_{t,p}^{p}\,\mathrm{d}t, \\ Y_{0,p} &= x_0 > 0 \end{cases} \tag{2.9}$$

for a Brownian motion $(B_t)_{t\geq 0}$, the parameter of interest $\vartheta \in \Theta \subset (-\frac{1}{2}, \infty)$ and the additional parameter $p < 1$. Note that for $p = -1$, we get the Bessel process back. We briefly analyze a martingale estimator based on the first eigenfunction with the same technique as before. Using the space-time transformation

$$X_{t,p} := e^{-\alpha t}Y_{\frac{e^{(1-p)\alpha t}-1}{(1-p)\alpha},p}$$

for some $\alpha > 0$, we receive by Itô's formula an ergodic and stationary version

$$\begin{cases} \mathrm{d}X_{t,p} &= X_{t,p}^{\frac{p+1}{2}}\,\mathrm{d}B_t + \big[\big(\vartheta + \frac{1}{2}\big)\,X_{t,p}^{p} - \alpha X_{t,p}\big]\,\mathrm{d}t, \\ X_{0,p} &= x_0 > 0, \end{cases} \tag{2.10}$$

for more details see Section 1.4. The corresponding generator can be stated as

$$\mathcal{L}_{\vartheta,\alpha,p}f(x) = \frac{1}{2}x^{p+1}f''(x) + \left[\left(\vartheta + \frac{1}{2}\right)x^p - \alpha x\right]f'(x).$$

The first eigenfunction arises after a brief calculation. We suppose a monomial of degree at least two

$$\mathcal{L}_{\vartheta,\alpha,p}x^\eta = \frac{\eta(\eta-1)}{2}x^{p+1+\eta-2} + \left[\left(\vartheta + \frac{1}{2}\right)x^p - \alpha x\right]nx^{\eta-1}$$

$$= \left(\frac{\eta(\eta-1)}{2} + \vartheta + \frac{1}{2}\right) x^{\eta+p-1} - \alpha\eta x^{\eta}$$

and immediately see by choosing $\eta = 1-p$ that the formula

$$\begin{aligned}\mathcal{L}_{\vartheta,\alpha,p} x^{1-p} &= \frac{(1-p)(-p)}{2} + \vartheta + \frac{1}{2} - \alpha(1-p)x^{1-p} \\ &= -\alpha(1-p)\left(x^{1-p} - \frac{2\vartheta+1-p}{2\alpha}\right)\end{aligned}$$

can be used to read the first eigenfunction

$$\varphi_{1,p} = x^{1-p} - \frac{2\vartheta+1-p}{2\alpha}$$

with eigenvalue $\lambda_{1,p} = (1-p)\alpha$. We assume $(X_{t,p})_{t\geq 0}$ to be stationary, that is, $X_{0,p} \sim \mu_{\vartheta,p}$, and assume $X_{\Delta,p}, \ldots, X_{n\Delta,p}$ to be discrete observations of (2.10). We consider the estimator based on the martingale estimation function

$$\begin{aligned}G_{n,p}(\vartheta) &= \sum_{i=1}^{n} \left(\varphi_{1,p}(X_{i\Delta,p}, \vartheta) - e^{-\lambda_{1,p}\Delta}\varphi_{1,p}(X_{(i-1)\Delta,p}, \vartheta)\right) \\ &= \sum_{i=1}^{n} \left(X_{i\Delta}^{1-p} - e^{-(1-p)\alpha} X_{(i-1)\Delta}^{1-p}\right) - \frac{2\vartheta+1-p}{2\alpha} n\big(1 - e^{-(1-p)\alpha\Delta}\big).\end{aligned}$$

The unique solution of $G_{n,p}(\widehat{\vartheta}_{n,p}) = 0$ is

$$\widehat{\vartheta}_{n,p} = \frac{\alpha \sum_{i=1}^{n} \left(X_{i\Delta,p}^{1-p} - X_{(i-1)\Delta,p}^{1-p} e^{-(1-p)\alpha\Delta}\right)}{n\big(1 - e^{-(1-p)\alpha\Delta}\big)} - \frac{1-p}{2}. \tag{2.11}$$

Next, we review how this process is related to a linear martingale estimation function. In particular, $\big(X_{t,p}^{1-p}\big)_{t\geq 0}$ is again a Cox-Ingersoll-Ross process due to Example 1.9, where we as well calculated the conditional mean

$$\mathbb{E}_{\vartheta}\big(X_{i\Delta,p}^{1-p} \,|\, X_{(i-1)\Delta,p}\big) = e^{-\alpha(1-p)\Delta} X_{(i-1)\Delta,p}^{1-p} + \frac{2\vartheta+1-p}{2\alpha}\big(1 - e^{-\alpha(1-p)\Delta}\big).$$

Thus, we receive the linear martingale estimation function

$$\begin{aligned}\Sigma_{n,p}(\vartheta) &:= \sum_{i=1}^{n} \left(X_{i\Delta,p}^{1-p} - \mathbb{E}_{\vartheta}\big(X_{i\Delta,p}^{1-p} | X_{(i-1)\Delta,p}\big)\right) \\ &= \sum_{i=1}^{n} \left(X_{i\Delta,p}^{1-p} - X_{(i-1)\Delta,p}^{1-p} e^{-(1-p)\alpha\Delta} + \frac{2\vartheta+1-p}{2\alpha}(e^{-(1-p)\alpha\Delta} - 1)\right)\end{aligned}$$

$$= G_{n,p}(\vartheta)$$

and see that the unique solution of $\Sigma_{n,p}(\vartheta_{n,p}) = 0$ is again (2.11).

Theorem 2.10: For every true value $\vartheta_0 \in \Theta \subset (-\frac{1}{2}, \infty)$, we have

(i) $\widehat{\vartheta}_{n,p} \to \vartheta_0$ in probability and

(ii) $\sqrt{n}(\widehat{\vartheta}_{n,p} - \vartheta_0) \to \mathcal{N}\big(0, \sigma^2(\vartheta_0)\big)$ in distribution

under P_{ϑ_0} with $\sigma^2(\vartheta_0) := \frac{(1-p)(\vartheta_0+1)e^{-(1-p)\alpha\Delta}}{1-e^{-(1-p)\alpha\Delta}} + \frac{(2\vartheta_0+1-p)(1-p)}{4}$.

Proof: Obviously, (1) and (2) from Condition 2.3 are satisfied. As $\sigma^2(\vartheta_0) \in (0, \infty)$ applies, the convergences (i) and (ii) are given if the equation

$$\sigma^2(\vartheta_0) = \frac{v(\vartheta_0)}{f(\vartheta_0)^2}$$

holds, where

$$\begin{aligned} f(\vartheta) &:= -\int_0^\infty \frac{\partial}{\partial\vartheta} \mathbb{E}_\vartheta\big(X_{\Delta,p}^{1-p} \,|\, X_{0,p} = x\big)\mu_{\vartheta,p}(x)\,\mathrm{d}x \\ &= -\int_0^\infty \frac{1}{\alpha}\left(1 - e^{-(1-p)\alpha\Delta}\right)\mu_{\vartheta,p}(x)\,\mathrm{d}x \\ &= \frac{e^{-(1-p)\alpha\Delta} - 1}{\alpha}, \\ v(\vartheta) &:= \int_0^\infty \mathbb{V}\mathrm{ar}_\vartheta(X_{\Delta,p}^{1-p} \,|\, X_{0,p} = x)\mu_{\vartheta,p}(x)\,\mathrm{d}x. \end{aligned}$$

We have already performed the main calculation in Chapter 1. The conditional variance of $X_{\Delta,p}^{1-p}$ given $X_{0,p}$ is determined by

$$\begin{aligned} \mathbb{V}\mathrm{ar}_\vartheta(X_{t,p}^{1-p} \,|\, X_{t_0,p}) = \frac{(1-p)}{\alpha}&\left(1 - e^{-\alpha(1-p)(t-t_0)}\right) \cdot \\ &\left[e^{-(1-p)\alpha(t-t_0)}X_{t_0,p}^{1-p} + \frac{2\vartheta+1-p}{4\alpha}\left(1 - e^{-(1-p)\alpha(t-t_0)}\right)\right], \end{aligned}$$

see Example 1.9. Combined with

$$\int_0^\infty x^{1-p}\mu_{\vartheta,p}(x)\,\mathrm{d}x \overset{1.5}{=} \frac{\Gamma\left(1+\frac{2\vartheta+2}{1-p}\right)}{\Gamma\left(\frac{2\vartheta+2}{1-p}\right)}\cdot\frac{1-p}{2\alpha} = \frac{2\vartheta+2}{1-p}\cdot\frac{1-p}{2\alpha} = \frac{\vartheta+1}{\alpha}$$

we establish

$$\begin{aligned} v(\vartheta) &= \int_0^\infty \mathbb{V}\mathrm{ar}_{\,\vartheta}(X_{\Delta,p}^{1-p}\,|\,X_{0,p}=x)\mu_{\vartheta,p}(x)\,\mathrm{d}x \\ &= \frac{(1-p)}{\alpha}\left(1-e^{-\alpha(1-p)\Delta}\right)\cdot \\ &\qquad \int_0^\infty \left[e^{-(1-p)\alpha\Delta}x^{1-p}+\frac{2\vartheta+1-p}{4\alpha}\left(1-e^{-(1-p)\alpha\Delta}\right)\right]\mu_{\vartheta,p}(x)\,\mathrm{d}x \\ &= \frac{(1-p)}{\alpha}\left(1-e^{-\alpha(1-p)\Delta}\right)\cdot\left[e^{-(1-p)\alpha\Delta}\frac{\vartheta+1}{\alpha}+\frac{2\vartheta+1-p}{4\alpha}\left(1-e^{-(1-p)\alpha\Delta}\right)\right] \end{aligned}$$

and hence the equation $\sigma^2(\vartheta_0) = \frac{v(\vartheta_0)}{f(\vartheta_0)^2}$ is valid. □

We want to increase the flexibility of $\Sigma_{n,p}$ using the same scheme as for $\Sigma_n = \Sigma_{n,-1}$ and once more obtain the optimal weight

$$\begin{aligned} \omega(\vartheta, X_{(i-1)\Delta,p}) &:= \frac{\frac{\partial}{\partial\vartheta}\,\mathbb{E}_{\,\vartheta}\big(X_{i\Delta,p}^{1-p}\,|\,X_{(i-1)\Delta,p}\big)}{\mathbb{V}\mathrm{ar}_{\,\vartheta}\big(X_{i\Delta,p}\,|\,X_{(i-1)\Delta,p}\big)} \\ &= \frac{1}{\frac{(2\vartheta+1-p)(1-p)}{4\alpha}(1-e^{-(1-p)\alpha\Delta})+(1-p)X_{(i-1)\Delta,p}^{1-p}e^{-(1-p)\alpha\Delta}} \end{aligned}$$

for the estimation function

$$\sum_{i=1}^n \omega(\vartheta, X_{(i-1)\Delta,p})\left(X_{i\Delta,p}^{1-p} - X_{(i-1)\Delta,p}^{1-p}e^{-(1-p)\alpha\Delta} + \frac{2\vartheta+1-p}{2\alpha}(e^{-(1-p)\alpha\Delta}-1)\right),$$

cf. (2.4). As before, we cannot explicitly derive the estimator as a solution of

$$\sum_{i=1}^n \frac{1}{\frac{(2\vartheta+1-p)(1-p)}{4\alpha}(1-e^{-(1-p)\alpha\Delta})+(1-p)X_{(i-1)\Delta,p}^{1-p}e^{-(1-p)\alpha\Delta}}.$$

$$\left(X_{i\Delta,p}^{1-p} - X_{(i-1)\Delta,p}^{1-p} e^{-(1-p)\alpha\Delta} + \frac{2\vartheta + 1 - p}{2\alpha}(e^{-(1-p)\alpha\Delta} - 1) \right) = 0,$$

but we can analyze the improvement with respect to the estimator $\widehat{\vartheta}_{n,p}$. Using the formula (2.4), we have to establish the finiteness of

$$\begin{aligned} &\int_0^\infty \left(\omega(\vartheta_0, X_{(i-1)\Delta,p}) \frac{\partial}{\partial\vartheta} \mathbb{E}_\vartheta \big(X_{i\Delta,p}^{1-p} \,|\, X_{(i-1)\Delta,p} \big) \right) \mu_{\vartheta_0}(x)\, \mathrm{d}x \\ =& \int_0^\infty \frac{1}{\frac{(2\vartheta_0+1-p)(1-p)}{4} + \frac{(1-p)e^{-(1-p)\alpha\Delta}}{\alpha(1-e^{-(1-p)\alpha\Delta})} X_{(i-1)\Delta,p}^{1-p}} \mu_{\vartheta_0}(x)\, \mathrm{d}x \\ <& \frac{4}{(2\vartheta_0 + 1 - p)(1-p)}, \end{aligned}$$

the reciprocal of the asymptotic variance, to achieve consistency and asymptotic normality. Comparing this result to the limit

$$\begin{aligned} \lim_{\alpha\Delta\to\infty} \sigma^2(\vartheta_0) &= \lim_{\alpha\Delta\to\infty} \frac{(1-p)(\vartheta_0+1)e^{-(1-p)\alpha\Delta}}{1 - e^{-(1-p)\alpha\Delta}} + \frac{(2\vartheta_0 + 1 - p)(1-p)}{4} \\ &= \frac{(2\vartheta_0 + 1 - p)(1-p)}{4}, \end{aligned}$$

we recognize a fast convergence to the asymptotic variance's lower bound of the optimal estimator. This result resembling the case of the Bessel process justifies the restriction to the explicit estimator $\widehat{\vartheta}_{n,p}$ from a practical point of view.

2.4 Estimator based on two and more eigenfunctions

Now, we turn back to the modified Bessel process and try to improve the asymptotic variance further by considering martingale estimation functions based on two eigenfunctions. Yet, this approach suffers from the drawback that we do not get explicit results for the estimators anymore, but we do for the asymptotic variance at least for weights depending only on the unknown parameter.

As in the previous sections we start with a class of martingale estimation functions with weight depending on the unknown parameter only. We consider

$$G_{n,2}(\vartheta) := \sum_{i=1}^{n}\sum_{j=1}^{2}\omega_j(\vartheta)\left(\varphi_j(X_{i\Delta},\vartheta) - e^{-\lambda_j(\vartheta)\Delta}\varphi_j(X_{(i-1)\Delta},\vartheta)\right),$$

where ω_1 and ω_2 are continuously differentiable functions only depending on ϑ.[3] Under suitable conditions on the interaction between the weights ω_i and the eigenfunctions, we can easily achieve a consistent and asymptotically normal estimator.

Theorem 2.11: If for every $\vartheta \in \Theta$

$$f(\omega_1,\omega_2,\vartheta) := \omega_1(\vartheta)\frac{1-e^{-2\alpha\Delta}}{\vartheta+1} + \omega_2(\vartheta)\frac{1-e^{-4\alpha\Delta}}{(\vartheta+1)(\vartheta+2)} \neq 0$$

is satisfied, then there exists a solution of $G_{n,2}(\widehat{\vartheta}_{n,2}) = 0$ with a probability tending to one as $n \to \infty$ under P_{ϑ_0}. Furthermore, for every true value $\vartheta_0 \in \Theta \subset (-\frac{1}{2}, \infty)$ we have

(i) $\lim\limits_{n\to\infty} \widehat{\vartheta}_{n,2} = \vartheta_0$ in probability and

(ii) $\lim\limits_{n\to\infty} \sqrt{n}(\widehat{\vartheta}_{n,2} - \vartheta_0) = \mathcal{N}\left(0, \frac{v(\omega_1,\omega_2,\vartheta_0)}{f^2(\omega_1,\omega_2,\vartheta_0)}\right)$ in distribution

under P_{ϑ_0} with

$$v(\omega_1,\omega_2,\vartheta_0) := \omega_1^2(\vartheta_0)\frac{1-e^{-4\alpha\Delta}}{\vartheta_0+1} + \omega_2^2(\vartheta_0)\frac{2-2e^{-8\alpha\Delta}}{(\vartheta_0+1)(\vartheta_0+2)}.$$

Proof: As by the assumption $f(\cdot,\cdot,\vartheta) \neq 0$ for every $\vartheta \in \Theta$, we conclude $\omega_1(\vartheta) \neq 0$ or $\omega_2(\vartheta) \neq 0$ and consequently $v(\cdot,\cdot,\vartheta) \neq 0$ for every $\vartheta \in \Theta$. Using again Theorem 2.2, we only have to establish the formulas of f and v given in Condition 2.5. In our calculations below we need the following straightforward properties

(a) Q_Δ^ϑ symmetric,

(b) $\int_0^\infty \varphi_1(x,\vartheta)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x = 0$,

(c) $\int_0^\infty \varphi_j(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x = 0$ for $j = 1, 2$,

[3] In the last section, we considered $G_{n,p}$. Here, $G_{n,2}$ is not meant to be the special case $p = 2$. To distinguish these two estimators, we could alternatively write $G_{n,1,p}$ to emphasize that one eigenfunction is considered and accordingly name the estimator $\widehat{\vartheta}_{n,1,p}$. Since we are examining these estimators in different sections, we leave it this way to simplify the notations.

(d) $\int_0^\infty x^{2\eta}\mu_\vartheta(x)\,\mathrm{d}x = \frac{(\vartheta+1)_\eta}{\alpha^\eta}$ for $\eta \in \mathbb{N}$.

Statement (d) was proved in Corollary 1.5. Furthermore, we verify

$$\begin{aligned}
\int_0^\infty \varphi_1(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x &= \int_0^\infty \left(1 - \frac{\alpha x^2}{\vartheta+1}\right)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(d)}}{=} 1 - \frac{\alpha}{\vartheta+1}\cdot\frac{\vartheta+1}{\alpha} = 0, \\
\int_0^\infty \varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x &= \int_0^\infty \left(1 - 2\frac{\alpha x^2}{\vartheta+1} + \frac{\alpha^2 x^4}{(\vartheta+1)(\vartheta+2)}\right)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(d)}}{=} 1 - 2 + 1 = 0
\end{aligned}$$

and

$$\begin{aligned}
\int_0^\infty \varphi_1(x,\vartheta)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x &= \int_0^\infty \varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x - \int_0^\infty \frac{\alpha x^2}{\vartheta+1}\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(c)}}{=} \int_0^\infty \left(-\frac{\alpha x^2}{\vartheta+1} + 2\frac{\alpha^2x^4}{(\vartheta+1)^2} - \frac{\alpha^3x^6}{(\vartheta+1)^2(\vartheta+2)}\right)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(d)}}{=} -1 + \frac{2(\vartheta+2)}{\vartheta+1} - \frac{\vartheta+3}{\vartheta+1} = 0.
\end{aligned}$$

Step 1: We separate the proof into two steps. Looking at the definition

$$f(\omega_1,\omega_2,\vartheta) := \sum_{i=1}^{2}\int_0^\infty\int_0^\infty \omega_i(\vartheta)\frac{\partial}{\partial\vartheta}\left(\varphi_i(x,\vartheta) - e^{-2\alpha\Delta}\varphi_i(y,\vartheta)\right)Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)$$

in Condition 2.5, the first step is to obtain the explicit expression given in Theorem 2.11. We can easily calculate the two summands

$$\begin{aligned}
&\int_0^\infty\int_0^\infty \omega_1(\vartheta)\frac{\partial}{\partial\vartheta}\left(\varphi_1(x,\vartheta) - e^{-2\alpha\Delta}\varphi_1(y,\vartheta)\right)Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{\text{(a)}}{=} \omega_1(\vartheta)\big(1-e^{-2\alpha\Delta}\big)\int_0^\infty\int_0^\infty \frac{\partial}{\partial\vartheta}\varphi_1(x,\vartheta)Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{(2.2)}{=} \omega_1(\vartheta)\big(1-e^{-2\alpha\Delta}\big)\int_0^\infty\int_0^\infty \frac{\alpha x^2}{(\vartheta+1)^2}p_\vartheta(\Delta,x,y)\mu_\vartheta(x)\,\mathrm{d}x\,\mathrm{d}y
\end{aligned}$$

$$\begin{aligned}
&= \omega_1(\vartheta)\big(1-e^{-2\alpha\Delta}\big)\int_0^\infty \underbrace{\int_0^\infty p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=1}\,\frac{\alpha x^2}{(\vartheta+1)^2}\mu_\vartheta(x)\,\mathrm{d}x \\
&= \omega_1(\vartheta)\big(1-e^{-2\alpha\Delta}\big)\int_0^\infty \frac{\alpha x^2}{(\vartheta+1)^2}\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{(d)}{=} \omega_1(\vartheta)\frac{1-e^{-2\alpha\Delta}}{\vartheta+1}
\end{aligned}$$

and similarly

$$\begin{aligned}
&\int_0^\infty\int_0^\infty \omega_2(\vartheta)\frac{\partial}{\partial\vartheta}\big(\varphi_2(x,\vartheta)-e^{-4\alpha\Delta}\varphi_2(y,\vartheta)\big)\,Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{(a)}{=} \omega_2(\vartheta)\big(1-e^{-4\alpha\Delta}\big)\int_0^\infty \frac{\partial}{\partial\vartheta}\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&= \omega_2(\vartheta)\big(1-e^{-4\alpha\Delta}\big)\int_0^\infty \frac{\partial}{\partial\vartheta}\left(1-2\frac{\alpha x^2}{\vartheta+1}+\frac{\alpha^2x^4}{(\vartheta+1)(\vartheta+2)}\right)\mu_\vartheta(x)\,\mathrm{d}x \\
&= \omega_2(\vartheta)\big(1-e^{-4\alpha\Delta}\big)\int_0^\infty \left(\frac{2\alpha}{(\vartheta+1)^2}x^2-\frac{(2\vartheta+3)\alpha^2}{(\vartheta+1)^2(\vartheta+2)^2}x^4\right)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{(d)}{=} \omega_2(\vartheta)\big(1-e^{-4\alpha\Delta}\big)\left(\frac{2}{\vartheta+1}-\frac{2\vartheta+3}{(\vartheta+1)(\vartheta+2)}\right) \\
&= \omega_2(\vartheta)\frac{1-e^{-4\alpha\Delta}}{(\vartheta+1)(\vartheta+2)}.
\end{aligned}$$

Step 2: According to Condition 2.5, we receive

$$v(\vartheta)=\sum_{i,j=1}^{2}\omega_i(\vartheta)\omega_j(\vartheta)\omega_{ij}(\vartheta)$$

with

$$\omega_{ij}(\vartheta):=\int_0^\infty\int_0^\infty\Big(\varphi_i(y,\vartheta)-e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\Big)\cdot\Big(\varphi_j(y,\vartheta)-e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big)\,Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y).$$

In the following, we explicitly compute these integrals, starting with ω_{11}. If we take a look at the proof of Theorem 2.7, we recognize the already calculated value

$$\omega_{11}(\vartheta) = \int_0^\infty \int_0^\infty \left(1 - \frac{\alpha y^2}{\vartheta+1} - e^{-2\alpha\Delta}\left(1 - \frac{\alpha x^2}{\vartheta+1}\right)\right)^2 Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y) = \frac{1-e^{-4\alpha\Delta}}{\vartheta+1}.$$

For the next term $\omega_{12}(\vartheta) = \omega_{21}(\vartheta)$, it holds

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \left(\varphi_1(y,\vartheta) - e^{-2\alpha\Delta}\varphi_1(x,\vartheta)\right) \cdot \left(\varphi_2(y,\vartheta) - e^{-4\alpha\Delta}\varphi_2(x,\vartheta)\right) Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{\text{(a)}}{=} \left(1 + e^{-6\alpha\Delta}\right) \int_0^\infty \varphi_1(x,\vartheta)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\qquad\qquad - (e^{-2\alpha\Delta} + e^{-4\alpha\Delta}) \int_0^\infty \int_0^\infty \varphi_1(y,\vartheta)\varphi_2(x,\vartheta) Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{\text{(b)}}{=} -(e^{-2\alpha\Delta} + e^{-4\alpha\Delta}) \int_0^\infty \int_0^\infty \varphi_1(y,\vartheta)\varphi_2(x,\vartheta) Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&= -(e^{-2\alpha\Delta} + e^{-4\alpha\Delta}) \int_0^\infty \int_0^\infty \left(1 - \frac{\alpha y^2}{\vartheta+1}\right) p_\vartheta(\Delta,x,y)\,\mathrm{d}y\, \varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(c)}}{=} (e^{-2\alpha\Delta} + e^{-4\alpha\Delta}) \int_0^\infty \frac{\alpha}{\vartheta+1}\, \mathbb{E}_\vartheta(X_\Delta^2 \mid X_0 = x)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{1.9}{=} (e^{-2\alpha\Delta} + e^{-4\alpha\Delta}) \int_0^\infty \left(\frac{\alpha}{\vartheta+1} x^2 e^{-2\alpha\Delta} + 1 - e^{-2\alpha\Delta}\right) \varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(c)}}{=} \frac{\alpha(e^{-4\alpha\Delta} + e^{-6\alpha\Delta})}{\vartheta+1} \int_0^\infty x^2 \varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&= \frac{\alpha(e^{-4\alpha\Delta} + e^{-6\alpha\Delta})}{\vartheta+1} \int_0^\infty \left(x^2 - \frac{2\alpha x^4}{\vartheta+1} + \frac{\alpha^2 x^6}{(\vartheta+1)(\vartheta+2)}\right) \mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{(d)}}{=} \frac{e^{-4\alpha\Delta} + e^{-6\alpha\Delta}}{\vartheta+1} \left[\vartheta + 1 - 2(\vartheta+2) + \vartheta + 3\right] \\
&= 0
\end{aligned}$$

and similarly we obtain for $\omega_{22}(\vartheta)$

$$\int_0^\infty\int_0^\infty \left(\varphi_2(y,\vartheta) - e^{-4\alpha\Delta}\varphi_2(x,\vartheta)\right)^2 Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y)$$
$$\overset{(a)}{=} (1+e^{-8\alpha\Delta})\int_0^\infty \varphi_2^2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x - 2e^{-4\alpha\Delta}\int_0^\infty\int_0^\infty \varphi_2(x,\vartheta)\varphi_2(y,\vartheta)Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y)$$
$$\overset{!}{=} \frac{2-2e^{-8\alpha\Delta}}{(\vartheta+1)(\vartheta+2)}.$$

For the last equation we verify the two integrals separately

$$\int_0^\infty\int_0^\infty \varphi_2(x,\vartheta)\varphi_2(y,\vartheta)Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y)$$
$$= \int_0^\infty \left(1-\frac{2\alpha}{\vartheta+1}x^2+\frac{\alpha^2}{(\vartheta+1)(\vartheta+2)}x^4\right)\varphi_2(x,\vartheta)p_\vartheta(\Delta,x,y)\mu_\vartheta(x)\,\mathrm{d}y\,\mathrm{d}x$$
$$\overset{(c)}{=} -\frac{2\alpha}{\vartheta+1}\int_0^\infty \mathbb{E}_\vartheta(X_\Delta^2\,|\,X_0=x)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$
$$+\frac{\alpha^2}{(\vartheta+1)(\vartheta+2)}\int_0^\infty \mathbb{E}_\vartheta(X_\Delta^4\,|\,X_0=x)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x. \tag{2.12}$$

We contemplate

$$\int_0^\infty \mathbb{E}_\vartheta(X_\Delta^2\,|\,X_0=x)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$
$$\overset{1.9}{=} \int_0^\infty \left(x^2e^{-2\alpha\Delta}-\frac{\vartheta+1}{\alpha}(e^{-2\alpha\Delta}-1)\right)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$
$$\overset{(c)}{=} e^{-2\alpha\Delta}\int_0^\infty x^2\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \overset{\text{s.a.}}{=} 0$$

and if we directly omit in the integral (2.12) the summand of

$$\mathbb{E}_\vartheta\big(X_\Delta^4\,|\,X_0=x\big) = e^{-4\alpha\Delta}x^4+\frac{1}{\alpha}\left(1-e^{-2\alpha\Delta}\right)\Big[(2\vartheta+4)e^{-2\alpha\Delta}x^2$$

$$+ \frac{(\vartheta+1)(\vartheta+2)}{2\alpha}\Big(1-e^{-2\alpha(t-t_0)}\Big)\Bigg],$$

which is independent of x, cf. Example 1.9, we obtain

$$\begin{aligned}
&\int_0^\infty \mathbb{E}_\vartheta(X_\Delta^4 \,|\, X_0 = x)\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{(c)}{=} \frac{2\vartheta+4}{\alpha}\left(e^{-2\alpha\Delta}-e^{-4\alpha\Delta}\right)\int_0^\infty x^2\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x + e^{-4\alpha\Delta}\int_0^\infty x^4\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{\text{s.a.}}{=} e^{-4\alpha\Delta}\int_0^\infty x^4\varphi_2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x \\
&= e^{-4\alpha\Delta}\int_0^\infty x^4\left(1-\frac{2\alpha x^2}{\vartheta+1}+\frac{\alpha^2x^4}{(\vartheta+1)(\vartheta+2)}\right)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{(d)}{=} \frac{e^{-4\alpha\Delta}}{\alpha^2}\big[(\vartheta+1)(\vartheta+2)-2(\vartheta+2)(\vartheta+3)+(\vartheta+3)(\vartheta+4)\big] \\
&= \frac{e^{-4\alpha\Delta}}{\alpha^2}\big[\vartheta^2+3\vartheta+2-\vartheta(\vartheta+3)\big] \\
&= \frac{2e^{-4\alpha\Delta}}{\alpha^2}.
\end{aligned}$$

Consequently, we have

$$\begin{aligned}
\int_0^\infty\int_0^\infty \varphi_2(x,\vartheta)\varphi_2(y,\vartheta)Q_\Delta(\,\mathrm{d}x,\,\mathrm{d}y) &= -\frac{2\alpha}{\vartheta+1}\cdot 0 + \frac{\alpha^2}{(\vartheta+1)(\vartheta+2)}\cdot\frac{2e^{-4\alpha\Delta}}{\alpha^2} \\
&= \frac{2e^{-4\alpha\Delta}}{(\vartheta+1)(\vartheta+2)}.
\end{aligned}$$

Therefore, we are left with the task of calculating

$$\begin{aligned}
\int_0^\infty \varphi_2^2(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x &= \int_0^\infty \varphi_2(x)\big(2\varphi_1(x)-2\varphi_1(x)+\varphi_2(x)\big)\mu_\vartheta(x)\,\mathrm{d}x \\
&\overset{(b)}{=} \int_0^\infty \varphi_2(x)\big(-2\varphi_1(x)+\varphi_2(x)\big)\mu_\vartheta(x)\,\mathrm{d}x \\
&= \int_0^\infty \varphi_2(x)\left(-1+\frac{\alpha^2x^4}{(\vartheta+1)(\vartheta+2)}\right)\mu_\vartheta(x)\,\mathrm{d}x
\end{aligned}$$

$$\overset{(c)}{=} \frac{\alpha^2}{(\vartheta+1)(\vartheta+2)} \int_0^\infty x^4 \varphi_2(x) \mu_\vartheta(x)\,\mathrm{d}x$$
$$\overset{\text{s.a.}}{=} \frac{\alpha^2}{(\vartheta+1)(\vartheta+2)} \cdot \frac{2}{\alpha^2} = \frac{2}{(\vartheta+1)(\vartheta+2)}.$$

Ultimately, we receive

$$\begin{aligned}\omega_{22}(\vartheta) &= \left(1+e^{-8\alpha\Delta}\right)\frac{2}{(\vartheta+1)(\vartheta+2)} - 2e^{-4\alpha\Delta}\frac{2e^{-4\alpha\Delta}}{(\vartheta+1)(\vartheta+2)} \\ &= \frac{1-2e^{-8\alpha\Delta}}{(\vartheta+1)(\vartheta+2)}.\end{aligned}$$

□

Our aim is now to find the ω_i's which lead to the smallest asymptotic variance as $\alpha\Delta \to \infty$. Therefore, we define for fixed $\vartheta \in \Theta$ the approximating functions

$$\begin{aligned}\widetilde{v}(\omega_1,\omega_2) &:= \frac{\omega_1^2(\vartheta)}{\vartheta+1} + \frac{2\omega_2^2(\vartheta)}{(\vartheta+1)(\vartheta+2)}, \\ \widetilde{f}(\omega_1,\omega_2) &:= \frac{\omega_1(\vartheta)}{\vartheta+1} + \frac{\omega_2(\vartheta)}{(\vartheta+1)(\vartheta+2)},\end{aligned}$$

for which

$$\lim_{\alpha\Delta\to\infty} \left| \frac{v(\vartheta)}{f^2(\vartheta)} - \frac{\widetilde{v}(\vartheta)}{\widetilde{f}^2(\vartheta)} \right| = 0$$

holds. This property justifies the search for the global minimum of

$$(\omega_1,\omega_2) \mapsto \frac{\widetilde{v}(\omega_1,\omega_2)}{\widetilde{f}^2(\omega_1,\omega_2)}.$$

To establish the minimum we first simplify the function

$$\begin{aligned}\frac{\widetilde{v}(\omega_1,\omega_2)}{\widetilde{f}^2(\omega_1,\omega_2)} &= \frac{\frac{\omega_1^2}{\vartheta+1} + \frac{2\omega_2^2}{(\vartheta+1)(\vartheta+2)}}{\left(\frac{\omega_1}{\vartheta+1} + \frac{\omega_2}{(\vartheta+1)(\vartheta+2)}\right)^2} \\ &= (\vartheta+1)(\vartheta+2)\frac{(\vartheta+2)\omega_1^2 + 2\omega_2^2}{\left((\vartheta+2)\omega_1 + \omega_2\right)^2} \\ &=: (\vartheta+1)(\vartheta+2)\frac{\bar{v}(\omega_1,\omega_2)}{\bar{f}^2(\omega_1,\omega_2)}\end{aligned}$$

and determine the first derivatives

$$\begin{aligned}
\frac{\partial}{\partial \omega_1} \frac{\widetilde{v}(\omega_1,\omega_2)}{\widetilde{f}^2(\omega_1,\omega_2)} &= (\vartheta+1)(\vartheta+2)\cdot \frac{\left(\frac{\partial}{\partial \omega_1}\bar{v}(\omega_1,\omega_2)\right)\bar{f}(\omega_1,\omega_2) - 2\bar{v}(\omega_1,\omega_2)\left(\frac{\partial}{\partial \omega_1}\bar{f}(\omega_1,\omega_2)\right)}{\bar{f}^3(\omega_1,\omega_2)} \\
&= (\vartheta+1)(\vartheta+2)\cdot \\
&\quad \frac{2(\vartheta+2)\omega_1\cdot\big(\cancel{(\vartheta+2)\omega_1}+\omega_2\big) - \big(\cancel{(\vartheta+2)\omega_1^2}+2\omega_2^2\big)\cdot 2(\vartheta+2)}{\big((\vartheta+2)\omega_1+\omega_2\big)^3} \\
&= 2(\vartheta+1)(\vartheta+2)^2\frac{\omega_1\omega_2-2\omega_2^2}{\big((\vartheta+2)\omega_1+\omega_2\big)^3}, \\
\frac{\partial}{\partial \omega_2} \frac{\widetilde{v}(\omega_1,\omega_2)}{\widetilde{f}^2(\omega_1,\omega_2)} &= (\vartheta+1)(\vartheta+2)\frac{4\omega_2\big((\vartheta+2)\omega_1+\cancel{\omega_2}\big) - 2\big((\vartheta+2)\omega_1^2+\cancel{2\omega_2^2}\big)\cdot 1}{\big((\vartheta+2)\omega_1+\omega_2\big)^3} \\
&= 2(\vartheta+1)(\vartheta+2)^2\frac{2\omega_1\omega_2-\omega_1^2}{\big((\vartheta+2)\omega_1+\omega_2\big)^3}.
\end{aligned}$$

Taking into account the properties of the ω_i's in Theorem 2.11, we get as possible minima $\omega_1 = 2\omega_2 \neq 0$ with value

$$\frac{\widetilde{v}(2\omega_2,\omega_2)}{\widetilde{f}^2(2\omega_2,\omega_2)} = \frac{2(\vartheta+1)(\vartheta+2)}{2\vartheta+5}.$$

In order to check if we indeed have minima, we consider $\omega_1 \neq 2\omega_2$ and see

$$\begin{aligned}
\frac{\widetilde{v}(\omega_1,\omega_2)}{\widetilde{f}^2(\omega_1,\omega_2)} - \frac{2(\vartheta+1)(\vartheta+2)}{2\vartheta+5} &= (\vartheta+1)(\vartheta+2)\left(\frac{(2\vartheta+5)(\vartheta+2)\omega_1^2+2(2\vartheta+5)\omega_2^2}{(2\vartheta+5)\big((\vartheta+2)\omega_1+\omega_2\big)^2}\right. \\
&\qquad \left. - \frac{2(\vartheta+2)^2\omega_1^2+4(\vartheta+2)\omega_1\omega_2+2\omega_2^2}{(2\vartheta+5)\big((\vartheta+2)\omega_1+\omega_2\big)^2}\right) \\
&= (\vartheta+1)(\vartheta+2)\frac{(\vartheta+2)\omega_1^2-4(\vartheta+2)\omega_1\omega_2+2(2\vartheta+4)\omega_2^2}{(2\vartheta+5)\big((\vartheta+2)\omega_1+\omega_2\big)^2} \\
&= (\vartheta+1)(\vartheta+2)^2\frac{\omega_1^2-4\omega_1\omega_2+4\omega_2^2}{(2\vartheta+5)\big((\vartheta+2)\omega_1+\omega_2\big)^2} \\
&= (\vartheta+1)(\vartheta+2)^2\frac{(\omega_1-2\omega_2)^2}{(2\vartheta+5)\big((\vartheta+2)\omega_1+\omega_2\big)^2} > 0.
\end{aligned}$$

Hence, these critical points are global minima. Finally, we may specify the improvement of the asymptotic variance

$$\vartheta+1-\frac{2(\vartheta+1)(\vartheta+2)}{2\vartheta+5} = (\vartheta+1)\frac{2\vartheta+5-2(\vartheta+2)}{2\vartheta+5} = \frac{\vartheta+1}{2\vartheta+5} > 0$$

if we consider the asymptotic behaviour $\alpha\Delta \to \infty$. Hence, we see that the relative improvement in comparison to $\vartheta + 1$, the bound of the asymptotic variance in the case of only one eigenfunction, is $\frac{1}{2\vartheta+5}$ and decreases as ϑ increases. However, for the boundary case $\vartheta = -\frac{1}{2}$, we get an improvement of 25%. Recalling that the Euclidean norm of a Dunkl process equals a Bessel process, we still preserve an improvement of 20% in case $\vartheta = 0$, which for a Dunkl process separates between finite and infinite jump activity. In Figure 2.2, the asymptotic information, the reciprocal of the asymptotic variance, is visualized for various ϑ. We can distinctly see that especially for $\alpha\Delta \geq 1$ not only the absolute, but also the relative improvement decreases significantly with increasing ϑ.

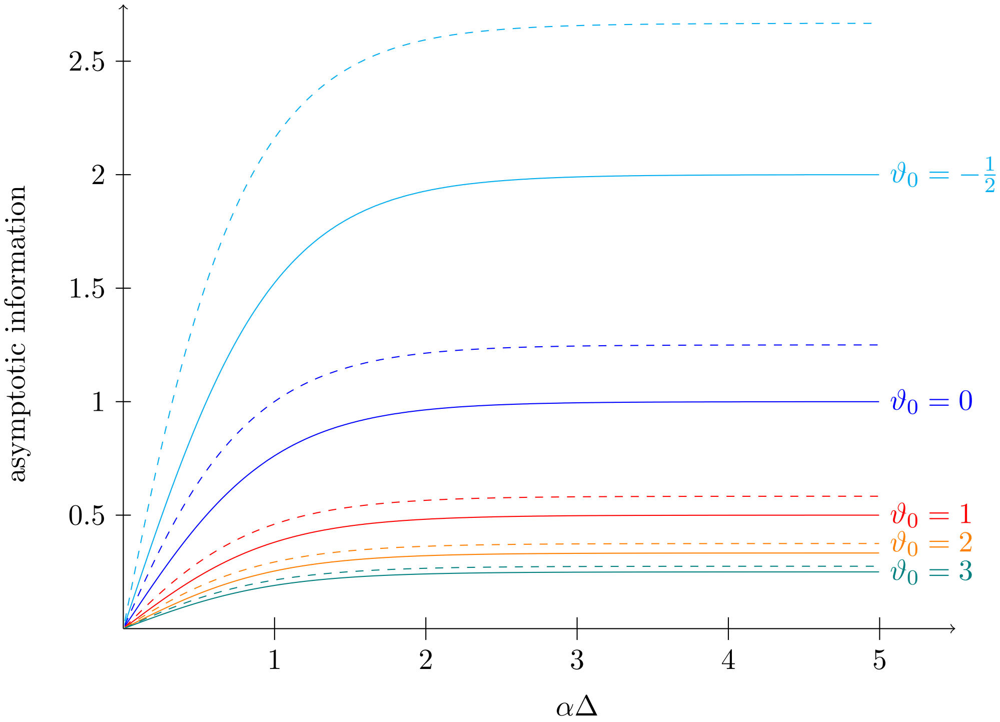

Figure 2.2: Comparison of the asymptotic information from Theorem 2.7 (solid line) to the one from Theorem 2.11 (dashed line) for $\omega_1 = 2$ and $\omega_2 = 1$.

As a second step we may consider weights which also depend on the observations. Note that though we may determine the optimal weights as solutions to a system of linear equations with coefficients depending on higher order conditional moments, which is theoretically feasible, we cannot provide an explicit result for the optimal asymptotic variance. Hence, we are not able to quantify the improvement compared to the simpler weights before.

If we take into account weights $\omega_j^\star$ that additionally depend on the trajectories, that is if we consider estimation functions

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\omega_j^\star(X_{(i-1)\Delta},\vartheta)\big(\phi_j(X_{i\Delta},\vartheta)-e^{-\lambda_j\Delta}\phi_j(x,\vartheta)\big),$$

the optimal weights are specified in (2.7) by an equation

$$\begin{pmatrix} u_{11} & u_{12}\\ u_{12} & u_{22}\end{pmatrix}\begin{pmatrix}\omega_1^\star\\ \omega_2^\star\end{pmatrix}=\begin{pmatrix}v_1\\ v_2\end{pmatrix}.$$

Now, we evaluate the elements of this matrix. By means of the calculations from Section 2.2, we recognize

$$\varphi_1(y,\vartheta)-e^{-2\alpha\Delta}\varphi_1(x,\vartheta)=-\frac{\alpha}{\vartheta+1}\Big(y^2-\mathbb{E}_\vartheta\big(X_\Delta^2\,|\,X_0=x\big)\Big),$$

which implies

$$\begin{aligned}
u_{11}(x,\vartheta) &= \mathbb{E}_\vartheta\Big(\big(\varphi_1(X_\Delta,\vartheta)-e^{-2\alpha\Delta}\varphi_1(x,\vartheta)\big)^2\,|\,X_0=x\Big)\\
&= \frac{\alpha^2}{(\vartheta+1)^2}\,\mathbb{E}_\vartheta\Big(\big(X_\Delta^2-\mathbb{E}_\vartheta(X_\Delta^2\,|\,x=x)\big)^2\,|\,X_0=x\Big)\\
&= \frac{\alpha^2}{(\vartheta+1)^2}\,\mathbb{V}\mathrm{ar}_\vartheta\big(X_\Delta^2\,|\,X_0=x\big)\\
&\overset{1.9}{=} \frac{(1-e^{-2\alpha\Delta})^2}{\vartheta+1}+\frac{2\alpha x^2}{(\vartheta+1)^2}\left(e^{-2\alpha\Delta}-e^{-4\alpha\Delta}\right).
\end{aligned}$$

Using Lemma 2.4, we evaluate

$$\begin{aligned}
u_{22}(x,\vartheta) &= \mathbb{E}_\vartheta\Big(\big(\varphi_2(X_\Delta,\vartheta)-e^{-4\alpha\Delta}\varphi_2(x,\vartheta)\big)^2\,|\,X_0=x\Big)\\
&\overset{2.4}{=} \mathbb{E}_\vartheta\Big(\big(\varphi_2(X_\Delta,\vartheta)-\mathbb{E}_\vartheta(\varphi_2(X_\Delta,\vartheta)\,|\,X_0=x)\big)^2\,|\,X_0=x\Big)\\
&= \mathbb{E}_\vartheta\Bigg(\bigg[\frac{\alpha^2}{(\vartheta+1)(\vartheta+2)}\big(X_\Delta^4-\mathbb{E}_\vartheta(X_\Delta^4\,|\,X_0=x)\big)\\
&\qquad\qquad -\frac{2\alpha}{\vartheta+1}\big(X_\Delta^2-\mathbb{E}_\vartheta(X_\Delta^2\,|\,X_0=x)\big)\bigg]^2\,\bigg|\,X_0=x\Bigg)\\
&= \frac{\alpha^4}{(\vartheta+1)^2(\vartheta+2)^2}\,\mathbb{V}\mathrm{ar}_\vartheta\big(X_\Delta^4\,|\,X_0=x\big)+\frac{4\alpha^2}{(\vartheta+1)^2}\,\mathbb{V}\mathrm{ar}_\vartheta\big(X_\Delta^2\,|\,X_0=x\big)\\
&\quad -\frac{4\alpha^3}{(\vartheta+1)^2(\vartheta+2)}\Big(\mathbb{E}_\vartheta\big(X_\Delta^6\,|\,X_0=x\big)-\mathbb{E}_\vartheta\big(X_\Delta^4\,|\,X_0=x\big)\,\mathbb{E}_\vartheta\big(X_\Delta^2\,|\,X_0=x\big)\Big).
\end{aligned}$$

and similarly

$$
\begin{aligned}
u_{12}(x,\vartheta) &= \mathbb{E}_\vartheta\Big(\big(\varphi_1(X_\Delta,\vartheta) - e^{-2\alpha\Delta}\varphi_1(x,\vartheta)\big)\big(\varphi_2(X_\Delta,\vartheta) - e^{-4\alpha\Delta}\varphi_2(x,\vartheta)\big) \mid X_0 = x\Big) \\
&\overset{2.4}{=} -\frac{\alpha^3}{(\vartheta+1)^2(\vartheta+2)}\,\mathbb{E}_\vartheta\Big(\big(X_\Delta^2 - \mathbb{E}_\vartheta(X_\Delta^2 \mid X_0 = x)\big)\big(X_\Delta^4 - \mathbb{E}_\vartheta(X_\Delta^4 \mid X_0 = x)\big) \mid X_0 = x\Big) \\
&\qquad + \frac{2\alpha^2}{(\vartheta+1)^2}\,\mathbb{E}_\vartheta\Big(\big(X_\Delta^2 - \mathbb{E}_\vartheta(X_\Delta^2 \mid X_0 = x)\big)^2 \mid X_0 = x\Big) \\
&= -\frac{\alpha^3}{(\vartheta+1)^2(\vartheta+2)}\Big(\mathbb{E}_\vartheta(X_\Delta^6 \mid X_0 = x) - \mathbb{E}_\vartheta(X_\Delta^2 \mid X_0 = x)\,\mathbb{E}_\vartheta(X_\Delta^4 \mid X_0 = x)\Big) \\
&\qquad + \frac{2\alpha^2}{(\vartheta+1)^2}\,\mathbb{V}\mathrm{ar}_\vartheta(X_\Delta^2 \mid X_0 = x).
\end{aligned}
$$

All these conditional expectations within the matrix can be calculated but this notation is more concise. We can easily work out

$$
\begin{aligned}
v_1(x,\vartheta) &= -\,\mathbb{E}_\vartheta\Big(\frac{\partial}{\partial\vartheta}\Big[\varphi_1(X_\Delta,\vartheta) - e^{-2\alpha\Delta}\varphi_1(x,\vartheta)\Big] \Big| X_0 = x\Big) \\
&= -\,\mathbb{E}_\vartheta\Big(\frac{\partial}{\partial\vartheta}\Big[1 - \frac{\alpha}{\vartheta+1}X_\Delta^2 - e^{-2\alpha\Delta}\Big(1 - \frac{\alpha}{\vartheta+1}x^2\Big)\Big] \Big| X_0 = x\Big) \\
&= -\frac{\alpha}{(\vartheta+1)}\,\mathbb{E}_\vartheta\Big(X_\Delta^2 - e^{-2\alpha\Delta}x^2 \mid X_0 = x\Big) \\
&= -\frac{\alpha}{(\vartheta+1)^2}\Big(\mathbb{E}_\vartheta\big(X_\Delta^2 \mid X_0 = x\big) - e^{-2\alpha\Delta}x^2\Big) \\
&\overset{1.9}{=} -\frac{\alpha}{\vartheta+1}\Big(x^2 e^{-2\alpha\Delta} - \frac{\vartheta+1}{\alpha}\big(e^{-2\alpha\Delta} - 1\big) - e^{-2\alpha\Delta}x^2\Big) \\
&= -\frac{1 - e^{-2\alpha\Delta}}{\vartheta+1}
\end{aligned}
$$

and

$$
\begin{aligned}
v_2(x,\vartheta) &= -\,\mathbb{E}_\vartheta\Big(\frac{\partial}{\partial\vartheta}\Big[\varphi_2(X_\Delta,\vartheta) - e^{-4\alpha\Delta}\varphi_2(x,\vartheta)\Big] \Big| X_0 = x\Big) \\
&= -\,\mathbb{E}_\vartheta\Big(\frac{\partial}{\partial\vartheta}\Big[1 - \frac{2\alpha}{\vartheta+1}X_\Delta^2 + \frac{1}{(\vartheta+1)(\vartheta+2)}X_\Delta^4 \\
&\qquad - e^{-4\alpha\Delta}\Big(1 - \frac{2\alpha}{\vartheta+1}x^2 + \frac{\alpha^2}{(\vartheta+1)(\vartheta+2)}x^4\Big)\Big] \Big| X_0 = x\Big) \\
&= -\frac{2\alpha}{(\vartheta+1)^2}\Big(\mathbb{E}_\vartheta(X_\Delta^2 \mid X_0 = x) - e^{-4\alpha\Delta}x^2\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{(2\vartheta+3)\alpha^2}{(\vartheta+1)^2(\vartheta+2)^2}\Big(\mathbb{E}_\vartheta(X_\Delta^4 \mid X_0=x)-e^{-4\alpha\Delta}x^4\Big)\\
&\overset{1.9}{=}-\frac{2\alpha}{(\vartheta+1)^2}\Big(e^{-2\alpha\Delta}x^2-e^{-4\alpha\Delta}x^2+\frac{\vartheta+1}{\alpha}\big(1-e^{-2\alpha\Delta}\big)\Big)\\
&+\frac{(2\vartheta+3)\alpha^2}{(\vartheta+1)^2(\vartheta+2)^2}\cdot\frac{\big(1-e^{-2\alpha\Delta}\big)}{\alpha}\Big(2(\vartheta+2)e^{-2\alpha\Delta}x^2+\frac{(\vartheta+1)(\vartheta+2)}{2\alpha}\big(1-e^{-2\alpha\Delta}\big)\Big)\\
&=\frac{2\alpha}{(\vartheta+1)^2}\big(e^{-2\alpha\Delta}-e^{-4\alpha\Delta}\big)x^2\Big(-1+\frac{2\vartheta+3}{\vartheta+2}\Big)\\
&\qquad\qquad-\frac{1-e^{-2\alpha\Delta}}{\vartheta+1}\Big(2-\frac{2\vartheta+3}{2(\vartheta+2)}\big(1-e^{-2\alpha\Delta}\big)\Big)\\
&=\frac{e^{-2\alpha\Delta}-e^{-4\alpha\Delta}}{(\vartheta+1)(\vartheta+2)}\Big(2\alpha x^2-\frac{2\vartheta+3}{2}\Big)-\frac{1-e^{-2\alpha\Delta}}{2(\vartheta+1)(\vartheta+2)}\big(2\vartheta+5\big).
\end{aligned}
$$

From these formulas, we can determine the optimal weights, which we omit for the sake of simplicity.

Finally, we examine the martingale estimation function that incorporates m eigenfunctions φ_j and weights ω_j depending on the parameter of interest:

$$G_{n,m}(\vartheta) := \sum_{i=1}^{n}\sum_{j=1}^{m}\omega_j(\vartheta)\Big(\varphi_j(X_{i\Delta},\vartheta)-e^{-\lambda_j(\vartheta)\Delta}\varphi_j(X_{(i-1)\Delta},\vartheta)\Big).$$

With the consideration of up to two eigenfunctions, we could simply determine the values of the asymptotic variance when the weights ω_j are independent of the trajectories. In the following, we notice that we can still calculate the associated asymptotic variance for all m, but this requires computing $\mathbb{E}\big(X_\Delta^{2\eta} \mid X_0=x\big)$ for all $\eta=1,...,m$.

Theorem 2.12: Assuming that the weights ω_j are bounded and continuously differentiable with bounded derivative for every $j=1,\dots,m$ such that

$$f_m(\omega_1,\dots,\omega_m,\vartheta) := \sum_{l=1}^{m}\Big(\sum_{j=1}^{l}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)(-j)_l\Big)\frac{\frac{\partial}{\partial\vartheta}(\vartheta+1)_l}{l!(\vartheta+1)_l}$$

does not vanish at ϑ_0, there exists a solution of

$$G_{n,m}(\widehat{\vartheta}_{n,m})=0$$

with a probability tending to one as $n\to\infty$ under P_{ϑ_0} such that

(i) $\lim\limits_{n\to\infty}\widehat{\vartheta}_{n,m}=\vartheta_0$ in probability,

(ii) $\lim_{n\to\infty} \sqrt{n}(\widehat{\vartheta}_{n,m} - \vartheta_0) = \mathcal{N}\left(0, \frac{v_m(\omega_1,\dots,\omega_m,\vartheta_0)}{f_m^2(\omega_1,\dots,\omega_m,\vartheta_0)}\right)$ in distribution

under P_{ϑ_0} with

$$v_m(\omega_1,\dots,\omega_m,\vartheta) := \sum_{i=1}^{m}\sum_{j=1}^{m} \omega_j(\vartheta)\omega_i(\vartheta)\nu_{ij}(\vartheta)$$

and

$$\begin{aligned}\nu_{jj}(\vartheta) :=& \big(1 + e^{-2\lambda_j\Delta}\big) \int_0^\infty \big(\varphi_j(x,\vartheta)\big)^2 \mu_\vartheta(x)\,\mathrm{d}x \\ &- 2e^{-\lambda_j\Delta} \sum_{l=0}^{j} \frac{(-j)_l \alpha^l}{(\vartheta+1)_l l!} \int_0^\infty \varphi_j(x,\vartheta)\,\mathbb{E}_\vartheta\big(X_\Delta^{2l} \mid X_0 = x\big)\mu_\vartheta(x)\,\mathrm{d}x, \\ \nu_{ij}(\vartheta) :=& -\big(e^{-\lambda_i\Delta} + e^{-\lambda_j\Delta}\big) \sum_{l=0}^{j} \frac{(-j)_l \alpha^l}{(\vartheta+1)_l l!} \int_0^\infty \varphi_i(x,\vartheta)\,\mathbb{E}_\vartheta\big(X_\Delta^{2l} \mid X_0 = x\big)\mu_\vartheta(x)\,\mathrm{d}x\end{aligned}$$

for $1 \le i \ne j \le m$.

Remark: Due to the symmetry $\nu_{ij} = \nu_{ji}$ it is enough to calculate ν_{ij} for $i < j$.

Proof: We again verify (i) to (iv) of Condition 2.5. For the notation we recall Section 2.1:

$$\begin{aligned} g(x,y,\vartheta) &:= \sum_{j=1}^{m} g_j(x,y,\vartheta) \\ &:= \sum_{j=1}^{m} \omega_j(\vartheta)\big(\varphi_j(y,\vartheta) - e^{-\lambda_j(\vartheta)\Delta}\varphi_j(x,\vartheta)\big).\end{aligned}$$

First, $g_j(x,y,\cdot)$ is obviously continuously differentiable on Θ for every $x,y \ge 0$ as the sum of products of a polynomial and a continuously differentiable function which ensures Condition 2.5 (i). Next, we need the boundedness with respect to ϑ of $\frac{\partial}{\partial\vartheta} g(x,y,\vartheta)$. For this purpose, we take a closer look at the eigenfunctions

$$\varphi_j(x,\vartheta) = \sum_{l=0}^{j} \frac{(-j)_l}{(\vartheta+1)_l l!}(\alpha x^2)^l.$$

The parameter ϑ appears here only in the term $\frac{1}{(\vartheta+1)_l}$. Owing to the condition on the weights, we confine the boundedness of $\frac{\partial}{\partial\vartheta} g(x,y,\vartheta)$ to showing that $\frac{\partial}{\partial\vartheta}\frac{1}{(\vartheta+1)_l}$ is bounded, which we prove by induction.

The initial case $l = 1$

$$\frac{\partial}{\partial\vartheta}\frac{1}{\vartheta+1} = \frac{1}{(\vartheta+1)^2} < \frac{1}{\left(\frac{1}{2}\right)^2} = 4$$

is trivial and thus we can assume there exists an $l \in \mathbb{N}$ such that

$$\left|\frac{\partial}{\partial\vartheta}\frac{1}{(\vartheta+1)_l}\right| \leq C_l$$

holds. The induction step follows immediately:

$$\begin{aligned}
\left|\frac{\partial}{\partial\vartheta}\frac{1}{(\vartheta+1)_{l+1}}\right| &= \left|\frac{\partial}{\partial\vartheta}\frac{1}{(\vartheta+1)_l(\vartheta+l+1)}\right| \\
&= \left|\left(\frac{\partial}{\partial\vartheta}\frac{1}{(\vartheta+1)_l}\right)\cdot\frac{1}{\vartheta+l+1} + \frac{1}{(\vartheta+1)_l}\cdot\frac{\partial}{\partial\vartheta}\frac{1}{\vartheta+l+1}\right| \\
&= \left|\left(\frac{\partial}{\partial\vartheta}\frac{1}{(\vartheta+1)_l}\right)\cdot\frac{1}{\vartheta+l+1} - \frac{1}{(\vartheta+1)_l(\vartheta+l+1)^2}\right| \\
&\leq \frac{C_l}{l+\frac{1}{2}} + \frac{1}{\left(\frac{1}{2}\right)_l\left(l+\frac{1}{2}\right)^2} =: C_{l+1}.
\end{aligned}$$

For the above inequality, note that $(\vartheta+1)_l$ is a polynomial in $\vartheta+1 > \frac{1}{2}$ with positive coefficients and thus monotonically increasing which means

$$0 < \frac{1}{(\vartheta+1)_l} < \frac{1}{\left(\frac{1}{2}\right)_l}$$

for every $\vartheta > -\frac{1}{2}$ and $l \in \mathbb{N}$.

For the verification of Condition 2.5 (iii) we check

$$f_m(\omega_1,\ldots,\omega_m,\vartheta) \stackrel{!}{=} \sum_{j=1}^{m}\int_0^\infty\int_0^\infty \omega_j(\vartheta)\frac{\partial}{\partial\vartheta}\left(\varphi_j(y,\vartheta) - e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\right)Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y).$$

By using the symmetry of Q_Δ^ϑ and straight forward calculations we derive

$$\begin{aligned}
&\sum_{j=1}^{m}\int_0^\infty\int_0^\infty \omega_j(\vartheta)\frac{\partial}{\partial\vartheta}\big(\varphi_j(y,\vartheta)-e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\big)Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)\\
&= \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\int_0^\infty\int_0^\infty\left(\frac{\partial}{\partial\vartheta}\varphi_j(x,\vartheta)\right)Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)\\
&\overset{(2.2)}{=} \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\int_0^\infty\underbrace{\int_0^\infty p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=1}\left(\frac{\partial}{\partial\vartheta}\varphi_j(x,\vartheta)\right)\mu_\vartheta(x)\,\mathrm{d}x\\
&= \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\int_0^\infty\left(\frac{\partial}{\partial\vartheta}\varphi_j(x,\vartheta)\right)\mu_\vartheta(x)\,\mathrm{d}x\\
&\overset{2.6}{=} \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\int_0^\infty\sum_{l=0}^{j}\left(\frac{\partial}{\partial\vartheta}\frac{(-j)_l}{(\vartheta+1)_l l!}\right)(\alpha x^2)^l\mu_\vartheta(x)\,\mathrm{d}x\\
&= \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\sum_{l=0}^{j}\left(\frac{\partial}{\partial\vartheta}\frac{(-j)_l}{(\vartheta+1)_l l!}\right)\int_0^\infty(\alpha x^2)^l\mu_\vartheta(x)\,\mathrm{d}x\\
&\overset{1.5}{=} \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\sum_{l=0}^{j}\left(\frac{\partial}{\partial\vartheta}\frac{(-j)_l}{(\vartheta+1)_l l!}\right)(\vartheta+1)_l\\
&= \sum_{j=1}^{m}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)\sum_{l=0}^{j}(-j)_l\frac{\frac{\partial}{\partial\vartheta}(\vartheta+1)_l}{(\vartheta+1)_l l!}\\
&= \sum_{l=1}^{m}\left(\sum_{j=1}^{l}\omega_j(\vartheta)\big(1-e^{-\lambda_j\Delta}\big)(-j)_l\right)\frac{\frac{\partial}{\partial\vartheta}(\vartheta+1)_l}{l!(\vartheta+1)_l} = f_m(\omega_1,\dots,\omega_m,\vartheta).
\end{aligned}$$

Finally, Condition 2.5 (iv) is obviously true, that is, the finiteness of $v_m(\omega_1,\dots,\omega_m,\vartheta_0)$. We are left with the task of verifying the formula

$$\nu_{ij}(\vartheta)\overset{!}{=}\int_0^\infty\int_0^\infty\Big(\varphi_i(y,\vartheta)-e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\Big)\Big(\varphi_j(y,\vartheta)-e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big)\,Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y).$$

Using the symmetry of Q_Δ^ϑ and that $p_\vartheta(\Delta,x,\cdot)$ is a probability measure leads to

$$\int_0^\infty\int_0^\infty\Big(\varphi_i(y,\vartheta)-e^{-\lambda_i\Delta}\varphi_i(x,\vartheta)\Big)\Big(\varphi_j(y,\vartheta)-e^{-\lambda_j\Delta}\varphi_j(x,\vartheta)\Big)\,Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y)$$

$$= \left(1 + e^{-(\lambda_i+\lambda_j)\Delta}\right) \int_0^\infty \varphi_i(x,\vartheta)\varphi_j(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$
$$- \left(e^{-\lambda_i\Delta} + e^{-\lambda_j\Delta}\right) \int_0^\infty \int_0^\infty \varphi_i(x,\vartheta)\varphi_j(y,\vartheta) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y).$$

The integral

$$\int_0^\infty \varphi_i(x,\vartheta)\varphi_j(x,\vartheta)\mu_\vartheta(x)\,\mathrm{d}x$$

vanishes for $i \neq j$, cf. [55, 15.13 The Spectral Representation of the Transition Density for a Diffusion. Eq. (13.9)]. Next, we derive

$$\begin{aligned}
\int_0^\infty \int_0^\infty \varphi_i(x,\vartheta)\varphi_j(y,\vartheta) Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) &= \int_0^\infty \int_0^\infty \varphi_i(x,\vartheta) \sum_{l=0}^{j} \frac{(-j)_l}{(\vartheta+1)_l l!} (\alpha y^2)^l Q_\Delta^\vartheta(\,\mathrm{d}x,\,\mathrm{d}y) \\
&\overset{(2.2)}{=} \sum_{l=0}^{j} \frac{(-j)_l \alpha^l}{(\vartheta+1)_l l!} \int_0^\infty \varphi_i(x,\vartheta) \underbrace{\int_0^\infty y^{2l} p_\vartheta(\Delta,x,y)\,\mathrm{d}y}_{=\mathbb{E}_\vartheta(X_\Delta^{2l}\,|\,X_0=x)} \mu_\vartheta(x)\,\mathrm{d}x \\
&= \sum_{l=0}^{j} \frac{(-j)_l \alpha^l}{(\vartheta+1)_l l!} \int_0^\infty \varphi_i(x,\vartheta)\,\mathbb{E}_\vartheta\big(X_\Delta^{2l}\,|\,X_0 = x\big)\mu_\vartheta(x)\,\mathrm{d}x,
\end{aligned}$$

which completes the proof. □

Considering a Dunkl process $(\Xi_t)_{t\geq 0}$, we can transfer this result as well to estimate the sum of multiplicities κ. For this purpose we set $(X_t)_{t\geq 0} := \big(e^{-\alpha t}\big\|\Xi_{\frac{\exp(2\alpha t)}{2\alpha}}\big\|\big)_{t\geq 0}$ to receive a modified Bessel process with index $\vartheta = \kappa + \frac{N}{2} - 1$. Under the conditions of Theorem 2.12 and for the corresponding observations of the Dunkl process there exists the estimator $\widehat{\vartheta}_{n,m}$ with a probability tending to one as $n \to \infty$ under P_{k_0}. We can define

$$\widehat{\kappa}_{n,m} := \widehat{\vartheta}_{n,m} - \frac{N}{2} + 1$$

for which we have

(i) $\lim\limits_{n\to\infty} \widehat{\kappa}_{n,m} = \kappa_0$ in probability,

(ii) $\lim_{n\to\infty} \sqrt{n}(\widehat{\kappa}_{n,m} - \kappa_0) = \mathcal{N}\left(0, \frac{v_m\left(\omega_1,\ldots,\omega_m,\kappa_0+\frac{N}{2}-1\right)}{f_m^2\left(\omega_1,\ldots,\omega_m,\kappa_0+\frac{N}{2}-1\right)}\right)$ in distribution

under $P_{k_0,R}$, with the functions and conditions specified as in Theorem 2.12. Analogous to Section 2.2, similar results hold for

$$\widehat{N}_{n,m} := 2\widehat{\vartheta}_{n,m} - 2\kappa + 2.$$

In each of these results concerning the Dunkl process, the process can be replaced by the corresponding multivariate Bessel process since these processes have the same Euclidean norm.

3 Estimation of the Cox-Ingersoll-Ross process under high-frequency sampling

In Chapter 2 we investigated martingale estimation functions for a stationary version of the classical Bessel process. These estimators depend only on the square of this process, a Cox-Ingersoll-Ross process with parameter $\theta = (2\vartheta + 2, 2\alpha, 4)$, see Example 1.9. We already estimated the first parameter at low frequency data. Furthermore, we presented a martingale estimator for a modified polynomial process as well, whose $(1-p)$th power is a Cox-Ingersoll-Ross process. Accordingly, in the specific case $p = 0$, the modified polynomial process is itself a Cox-Ingersoll-Ross process. We now focus immediately on the Cox-Ingersoll-Ross process and estimate all three parameters at high-frequency data.

The content of this chapter is partially incorporated in the paper

Estimation of ergodic square-root diffusion under high-frequency sampling
Econometrics and Statistics Article Number: 346 (2022)
Yuzhong Cheng, Nicole Hufnagel, Hiroki Masuda.

Yuzhong Cheng wrote his master's thesis based on this cooperation, [20]. The simulations are available through the yuima package in R, cf. [45, 46].

3.1 Gaussian quasi-likelihood function and existing results for the Cox-Ingersoll-Ross process

In this section we present a classical method from statistics applied to stochastic processes: the Gaussian quasi-likelihood method. This method is another alternative to the maximum likelihood estimator, such as in the previous chapter, when the density is unknown or too complicated.

We consider observations $Z_{t_0}, \dots, Z_{t_n}$ of a stochastic process $(Z_t)_{t\ge 0}$ depending on a parameter of interest $\theta \in \Theta \subset \mathbb{R}^d$. For a Gaussian approximation, we assume that the conditional mean as well as the conditional variance are known, that is,

$$\begin{aligned} \mu_{j-1}(\theta) &:= \mathbb{E}\left(Z_{t_j} \,|\, Z_{t_{j-1}}\right), \\ \sigma^2_{j-1}(\theta) &:= \mathbb{V}\mathrm{ar}\left(Z_{t_j} \,|\, Z_{t_{j-1}}\right) := \mathbb{E}\left((Z_{t_j} - \mu_{j-1}(\theta))^2 \,|\, Z_{t_{j-1}}\right). \end{aligned} \tag{3.1}$$

The *Gaussian quasi-likelihood function (GQLF)* is then defined by

$$\mathbb{H}_n(\theta) := \sum_{j=1}^{n} \log\Big(\phi\left(Z_{t_j}, \mu_{j-1}(\theta), \sigma^2_{j-1}(\theta)\right)\Big).$$

This estimation function looks similar to the log likelihood function. The only difference is that we approximate the density function by the Gaussian density $\phi(\cdot, \mu, \Sigma)$ with mean vector μ and covariance matrix Σ. The *Gaussian quasi-maximum likelihood estimator* (GQMLE) is any

$$\widehat{\theta}_n \in \underset{\overline{\Theta}}{\operatorname{argmax}}\, \mathbb{H}_n.$$

The existence and properties will be discussed later in our specific setting. It is well-known that the GQLF works effectively for uniformly elliptic diffusions with coefficients smooth enough, see [58, 88]. Even when the diffusion coefficients are not uniformly elliptic, it is quite often assumed that the inverse of the diffusion coefficients can be bounded above by a constant multiple of the function $1 + |x|^C$ for some $C > 0$. In the following, we will deal with a process that does not satisfy any of these properties but nevertheless receive asymptotic efficiency using the GQMLE.

We now recall the diffusion introduced in Section 1.5, that is, the Cox-Ingersoll-Ross process satisfying the stochastic differential equation

$$\left\{ \begin{array}{rcl} \mathrm{d}Z_t &=& (\alpha - \beta Z_t)\,\mathrm{d}t + \sqrt{\gamma Z_t}\,\mathrm{d}B_t, \\ Z_0 &=& z_0 > 0, \end{array} \right. \tag{3.2}$$

where $(B_t)_{t\ge 0}$ is a Brownian motion. The parameter of interest is

$$\theta := (\alpha, \beta, \gamma) \in \Theta \subset (0, \infty)^3.$$

Through this chapter, we assume the process to be stationary, in formulas $Z_0 \sim \pi_\theta$, where π_θ is the invariant measure specified in Section 1.5. For simplicity, we regard only equidistant times $t_j = jh$ and consider high frequency data, in formulas $h = h_n \to 0$ and $T_n := nh \to \infty$ when $n \to \infty$. As mentioned in Section 1.5, the origin is non-attracting for $2\alpha > \gamma$. Here, we assume an even stronger property that the parameter space Θ is a bounded convex domain satisfying

$$\overline{\Theta} \subset \left\{(\alpha, \beta, \gamma) \in (0, \infty)^3 : 2\alpha > 5\gamma\right\}. \tag{3.3}$$

This choice of the parameters is important for the application of Lemma 1.8 within our proofs later. We then have $2 < \frac{1}{2}\left(\frac{2\alpha}{\gamma} - 1\right)$, hence in particular Lemma 1.8 (iii) holds for $p = 1, 2$, respectively Lemma 1.8 (iv) works for every $p \leq 5$.

Recalling the parameters of the modified polynomial process in Chapter 2, we recognize that its $(1 - p)$th power is a Cox-Ingersoll-Ross process with parameter

$$\theta = \left(\frac{1-p}{2}(2\vartheta + 1 - p), \alpha(1-p), (1-p)^2\right),$$

where the condition on the right-hand side of (3.3) is fulfilled if additionally $\vartheta > 2(1 - p)$ holds. Especially, we can apply the following results in this chapter on the modified polynomial process and in particular on the modified Bessel process, case $p = -1$.

Next, we specify the GQLF for the Cox-Ingersoll-Ross model. Therefore, we recall the conditional mean and the conditional variance from Lemma 1.7:

$$\mu_{j-1}(\alpha, \beta) = e^{-\beta h} Z_{t_{j-1}} + \frac{\alpha}{\beta}(1 - e^{-\beta h}), \tag{3.4}$$

$$\sigma^2_{j-1}(\theta) = \frac{\gamma}{\beta}\left(1 - e^{-\beta h}\right)\left[e^{-\beta h} Z_{t_{j-1}} + \frac{\alpha}{2\beta}\left(1 - e^{-\beta h}\right)\right]. \tag{3.5}$$

In particular, the conditional mean is independent of the diffusion parameter γ. Using these expressions we receive as GQLF:

$$\begin{aligned}
\mathbb{H}_n(\theta) &= \sum_{j=1}^{n} \log\left(\phi\left(Z_{t_j}, \mu_{j-1}(\alpha, \beta), \sigma^2_{j-1}(\theta)\right)\right) \\
&= -n \log\left(\sqrt{2\pi}\right) - \frac{1}{2}\sum_{j=1}^{n}\left(\log \sigma^2_{j-1}(\theta) + \frac{1}{\sigma^2_{j-1}(\theta)}\left(Z_{t_j} - \mu_{j-1}(\alpha, \beta)\right)^2\right) \\
&= -n \log\left(\sqrt{2\pi}\right) - \frac{1}{2}\sum_{j=1}^{n} \log\left(\frac{\gamma}{\beta}\left(1 - e^{-\beta h}\right)\left[e^{-\beta h} Z_{t_{j-1}} + \frac{\alpha}{2\beta}\left(1 - e^{-\beta h}\right)\right]\right)
\end{aligned}$$

$$-\frac{1}{2}\sum_{j=1}^{n}\frac{\left(Z_{t_j}-e^{-\beta h}Z_{t_{j-1}}-\frac{\alpha}{\beta}\left(1-e^{-\beta h}\right)\right)^2}{\frac{\gamma}{\beta}(1-e^{-\beta h})\left[e^{-\beta h}Z_{t_{j-1}}+\frac{\alpha}{2\beta}\left(1-e^{-\beta h}\right)\right]}. \tag{3.6}$$

Then, the GQMLE

$$\widehat{\theta}_n=\left(\widehat{\alpha}_n,\widehat{\beta}_n,\widehat{\gamma}_n\right)\in\underset{\overline{\Theta}}{\operatorname{argmax}}\,\mathbb{H}_n \tag{3.7}$$

exists obviously since $\overline{\Theta}$ is compact and $\mathbb{H}_n$ is continuous. As we lack the uniform ellipticity condition for the present model (3.2), we need to take care of the irregularity of the diffusion coefficient when proving moment bounds and basic limit theorems. The GQMLE $\widehat{\theta}_n$ cannot be given in a closed form because of its nonlinearity in the parameters. It is well-known, see Section 1.5, that the Cox-Ingersoll-Ross process has a noncentral chi-squared transition density and hence is far from being Gaussian. Nevertheless, this local approximation seems natural since the driving noise is Gaussian.

However, as in [72, Section 3], we will first introduce the explicit initial estimator $\widehat{\theta}_{0,n}=\left(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n},\widehat{\gamma}_{0,n}\right)$, defined below, since it is easily derivable. Its asymptotics are rather similar [72], even though we need to take care of some moment bounds in the present high-frequency setup.

First, we look at the drift parameter (α,β). We introduce the conditional least-squares estimator $(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n})$ defined to be a maximizer of

$$\begin{aligned}\mathbb{H}_{1,n}(\alpha,\beta)&:=-\sum_{j=1}^{n}\left(Z_{t_j}-\mu_{j-1}(\alpha,\beta)\right)^2\\&=-\sum_{j=1}^{n}\left(Z_{t_j}-e^{-\beta h}Z_{t_{j-1}}-\frac{\alpha}{\beta}\left(1-e^{-\beta h}\right)\right)^2,\end{aligned}$$

which uses only the conditional mean μ_{j-1} so that we receive an estimation function independent of γ.

Lemma 3.1: The least-squares estimator is given by

$$\begin{aligned}\widehat{\alpha}_{0,n}&:=\frac{\overline{Z}_n-e^{-\widehat{\beta}_{0,n}h}\overline{Z}'_n}{1-e^{-\widehat{\beta}_{0,n}h}}\widehat{\beta}_{0,n},\\\widehat{\beta}_{0,n}&:=-\frac{1}{h}\log\left(\frac{\sum_{j=1}^{n}\left(Z_{t_{j-1}}-\overline{Z}'_n\right)\left(Z_{t_j}-\overline{Z}_n\right)}{\sum_{j=1}^{n}\left(Z_{t_{j-1}}-\overline{Z}'_n\right)^2}\right)\end{aligned}$$

with $\overline{Z}_n := \frac{1}{n}\sum_{j=1}^{n} Z_{t_j}$ and $\overline{Z}'_n := \frac{1}{n}\sum_{j=1}^{n} Z_{t_{j-1}}$.

Proof: Initially, we identify all critical points by calculating

$$\frac{\partial}{\partial\alpha}\mathbb{H}_{1,n}(\alpha,\beta) = \frac{2\big(1-e^{-\beta h}\big)}{\beta}\sum_{j=1}^{n}\Big(Z_{t_j} - e^{-\beta h}Z_{t_{j-1}} - \frac{\alpha}{\beta}\big(1-e^{-\beta h}\big)\Big)$$

which is zero only for

$$\begin{aligned}\alpha_0(\beta) &= \frac{\beta}{1-e^{-\beta h}}\cdot\frac{1}{n}\sum_{j=1}^{n}\Big(Z_{t_j} - e^{-\beta h}Z_{t_{j-1}}\Big)\\ &= \frac{\overline{Z}_n - e^{-\beta h}\overline{Z}'_n}{1-e^{-\beta h}}\beta.\end{aligned} \tag{3.8}$$

Due to the definitions we directly see

$$\begin{aligned}\sum_{j=1}^{n}\big(Z_{t_j} - \overline{Z}_n\big) &= 0,\\ \sum_{j=1}^{n}\big(Z_{t_{j-1}} - \overline{Z}'_n\big) &= 0,\end{aligned} \tag{3.9}$$

which we will often use in the following to specify the value of β such that $\nabla\mathbb{H}_{1,n}\big(\alpha_0(\beta),\beta\big) = \mathbf{0}$ holds.[4] Therefore, we derive

$$\begin{aligned}\frac{\partial}{\partial\beta}\mathbb{H}_{1,n}(\alpha,\beta)\Big|_{\alpha=\alpha_0(\beta)} &= -2\sum_{j=1}^{n}\Big(Z_{t_j} - e^{-\beta h}Z_{t_{j-1}} - \frac{\alpha_0(\beta)}{\beta}\big(1-e^{-\beta h}\big)\Big)\\ &\qquad\cdot\left(he^{-\beta h}Z_{t_{j-1}} - \alpha_0(\beta)\frac{e^{-\beta h}\big(\beta h + 1 - e^{-\beta h}\big)}{\beta^2}\right)\\ &\overset{(3.8)}{=} -2\sum_{j=1}^{n}\Big(Z_{t_j} - e^{-\beta h}Z_{t_{j-1}} - \big(\overline{Z}_n - e^{-\beta h}\overline{Z}'_n\big)\Big)\\ &\qquad\cdot\left(he^{-\beta h}Z_{t_{j-1}} - \alpha_0(\beta)\frac{e^{-\beta h}\big(\beta h + 1 - e^{-\beta h}\big)}{\beta^2}\right)\\ &\overset{(3.9)}{=} -2he^{-\beta h}\sum_{j=1}^{n}\Big(Z_{t_j} - \overline{Z}_n - e^{-\beta h}\big(Z_{t_{j-1}} - \overline{Z}'_n\big)\Big)Z_{t_{j-1}}.\end{aligned}$$

[4] In this chapter we use the notation **0** for the multidimensional origin.

Since the factor $2he^{-\beta h}$ is positive, the term is zero only if

$$e^{-\beta h} = \frac{\sum_{j=1}^{n} Z_{t_{j-1}}\big(Z_{t_j} - \overline{Z}_n\big)}{\sum_{j=1}^{n} Z_{t_{j-1}}\big(Z_{t_{j-1}} - \overline{Z}'_n\big)} \overset{(3.9)}{=} \frac{\sum_{j=1}^{n} \big(Z_{t_{j-1}} - \overline{Z}'_n\big)\big(Z_{t_j} - \overline{Z}_n\big)}{\sum_{j=1}^{n} \big(Z_{t_{j-1}} - \overline{Z}'_n\big)^2}$$

leading to the closed forms of $\widehat{\alpha}_{0,n}$ and $\widehat{\beta}_{0,n}$ specified in the statement. Now, we verify that the Hessian matrix is negative-definite:

$$\frac{\partial^2}{\partial\alpha^2}\mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) = -2n\frac{\big(1 - e^{-\widehat{\beta}_{0,n}h}\big)^2}{\widehat{\beta}_{0,n}^2} < 0.$$

Furthermore, we require

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha\partial\beta}\mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) &= \frac{\partial}{\partial\beta}\left(\frac{\partial}{\partial\alpha}\mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big)\right) \\
&= \frac{\partial}{\partial\beta}\left(\frac{2\big(1 - e^{-\beta h}\big)}{\beta}\sum_{j=1}^{n}\left(Z_{t_j} - e^{-\beta h}Z_{t_{j-1}} - \frac{\alpha}{\beta}\big(1 - e^{-\beta h}\big)\right)\right)\Bigg|_{(\alpha,\beta)=(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n})} \\
&= \left(\frac{\partial}{\partial\beta}\frac{1 - e^{-\beta h}}{\beta}\right)\Bigg|_{\beta=\widehat{\beta}_{0,n}} \cdot \frac{\widehat{\beta}_{0,n}}{1 - e^{-\widehat{\beta}_{0,n}h}}\underbrace{\frac{\partial}{\partial\alpha}\mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big)}_{=0} \\
&\quad + \frac{2\big(1 - e^{-\widehat{\beta}_{0,n}h}\big)}{\widehat{\beta}_{0,n}}\sum_{j=1}^{n}\left(he^{-\widehat{\beta}_{0,n}h}Z_{t_{j-1}} - \widehat{\alpha}_{0,n}\frac{e^{-\widehat{\beta}_{0,n}h}\big(\widehat{\beta}_{0,n}h + 1 - e^{-\widehat{\beta}_{0,n}h}\big)}{\widehat{\beta}_{0,n}^2}\right) \\
&= \frac{2n\big(1 - e^{-\widehat{\beta}_{0,n}h}\big)}{\widehat{\beta}_{0,n}}e^{-\widehat{\beta}_{0,n}h}\left(h\overline{Z}'_n - \frac{\widehat{\alpha}_{0,n}\big(\widehat{\beta}_{0,n}h + 1 - e^{-\widehat{\beta}_{0,n}h}\big)}{\widehat{\beta}_{0,n}^2}\right)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial^2}{\partial\beta^2}\mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) \\
&= -2\frac{\partial}{\partial\beta}\left(\sum_{j=1}^{n}\left(Z_{t_j} - e^{-\beta h}Z_{t_{j-1}} - \frac{\widehat{\alpha}_{0,n}}{\beta}\big(1 - e^{-\beta h}\big)\right)\right)
\end{aligned}$$

$$\cdot \left(he^{-\beta h} Z_{t_{j-1}} - \widehat{\alpha}_{0,n} \frac{e^{-\beta h}\big(\beta h + 1 - e^{-\beta h}\big)}{\beta^2} \right)\Bigg)\Bigg|_{(\alpha,\beta)=(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n})}$$

$$= -2 \sum_{j=1}^{n} \left(he^{-\widehat{\beta}_{0,n} h} Z_{t_{j-1}} - \widehat{\alpha}_{0,n} \frac{e^{-\widehat{\beta}_{0,n} h}\big(\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}\big)}{\widehat{\beta}_{0,n}^2} \right)^2$$

$$- 2 \sum_{j=1}^{n} \left(Z_{t_j} - e^{-\widehat{\beta}_{0,n} h} Z_{t_{j-1}} - \frac{\widehat{\alpha}_{0,n}}{\widehat{\beta}_{0,n}}\big(1 - e^{-\widehat{\beta}_{0,n} h}\big) \right)^2$$

$$\cdot \left(- h^2 e^{-\widehat{\beta}_{0,n} h} Z_{t_{j-1}} - \widehat{\alpha}_{0,n} \frac{e^{-\widehat{\beta}_{0,n} h}\big(- \widehat{\beta}_{0,n} h^2 - 2\widehat{\beta}_{0,n} h + 2e^{\widehat{\beta}_{0,n} h} - 2\big)}{\widehat{\beta}_{0,n}^3} \right)$$

$$= -2e^{-2\widehat{\beta}_{0,n} h} \sum_{j=1}^{n} \left(hZ_{t_{j-1}} - \widehat{\alpha}_{0,n} \frac{\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}}{\widehat{\beta}_{0,n}^2} \right)^2$$

$$= -2e^{-2\widehat{\beta}_{0,n} h} \left(\sum_{j=1}^{n} h^2 Z_{t_{j-1}}^2 - 2n\widehat{\alpha}_{0,n} \frac{\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}}{\widehat{\beta}_{0,n}^2} h\overline{Z}'_n + n\widehat{\alpha}_{0,n}^2 \frac{\big(\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}\big)^2}{\widehat{\beta}_{0,n}^4} \right).$$

The second-last equality is valid since $\nabla \mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) = \mathbf{0}$. To obtain the determinant of the Hessian matrix we calculate

$$\left(\frac{\partial^2}{\partial\alpha\partial\beta} \mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) \right)^2$$

$$= \frac{4n^2\big(1 - e^{-\widehat{\beta}_{0,n} h}\big)^2}{\widehat{\beta}_{0,n}^2} e^{-2\widehat{\beta}_{0,n} h} \left(h\overline{Z}'_n - \frac{\widehat{\alpha}_{0,n}\big(\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}\big)}{\widehat{\beta}_{0,n}^2} \right)^2$$

$$= \frac{4n^2\big(1 - e^{-\widehat{\beta}_{0,n} h}\big)^2}{\widehat{\beta}_{0,n}^2} e^{-2\widehat{\beta}_{0,n} h} \left(h^2 \big(\overline{Z}'_n\big)^2 - 2\widehat{\alpha}_{0,n} \frac{\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}}{\widehat{\beta}_{0,n}^2} h\overline{X}'_n + \widehat{\alpha}_{0,n}^2 \frac{\big(\widehat{\beta}_{0,n} h + 1 - e^{-\widehat{\beta}_{0,n} h}\big)^2}{\widehat{\beta}_{0,n}^4} \right)$$

and conclude with the Cauchy-Schwarz inequality:

$$\left(\frac{\partial^2}{\partial\alpha^2} \mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) \right)\left(\frac{\partial^2}{\partial\beta^2} \mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) \right) - \left(\frac{\partial^2}{\partial\alpha\partial\beta} \mathbb{H}_{1,n}\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) \right)^2$$

$$= 4h^2 \frac{\big(1 - e^{-\widehat{\beta}_{0,n} h}\big)^2}{\widehat{\beta}_{0,n}^2} e^{-2\widehat{\beta}_{0,n} h} \left(n \sum_{j=1}^{n} Z_{t_{j-1}}^2 - \big(\overline{Z}'_n\big)^2 \right)$$

$$= 4h^2 \frac{\left(1 - e^{-\widehat{\beta}_{0,n}h}\right)^2}{\widehat{\beta}_{0,n}^2} e^{-2\widehat{\beta}_{0,n}h} \left[n \sum_{j=1}^n Z_{t_{j-1}}^2 - \left(\sum_{j=1}^n Z_{t_{j-1}} \right)^2 \right]$$
$$\geq 4h^2 \frac{\left(1 - e^{-\widehat{\beta}_{0,n}h}\right)^2}{\widehat{\beta}_{0,n}^2} e^{-2\widehat{\beta}_{0,n}h} (n-1) \sum_{j=1}^n Z_{t_{j-1}}^2 > 0.$$

Hence, the Hessian matrix is negative definite and we indeed have a local maximum. Since the function $\mathbb{H}_n$ is continuously differentiable on $\mathbb{R}^2$ and we have just this critical point, it is indeed a global maximum. Otherwise, if there exists another point where the function has a greater value we would get a saddle point using mountain pass theorem, see [7] and [9, Theorem 5]. This would be a contradiction since we do not have a saddle point. □

For the estimation of the diffusion parameter γ, we substitute $(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n})$ into (α, β) in the GQLF (3.6) and denote the resulting function by $\mathbb{H}_{2,n}(\gamma)$:

$$\begin{aligned}
\mathbb{H}_{2,n}(\gamma) &:= \mathbb{H}_n\left(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}, \gamma\right) \\
&= \sum_{j=1}^n \log \phi \left(Z_{t_j}, \mu_{j-1}(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}), \sigma_{j-1}^2(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}, \gamma) \right) \\
&= -n \log\left(\sqrt{2\pi}\right) - \frac{1}{2} \sum_{j=1}^n \Bigg(\log \sigma_{j-1}^2(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}, \gamma) \\
&\qquad + \frac{1}{\sigma_{j-1}^2(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}, \gamma)} \left(Z_{t_j} - \mu_{j-1}(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}) \right)^2 \Bigg).
\end{aligned}$$

We are looking for a maximizer here as well. Via the representation of σ_{j-1}^2, see (3.5), we immediately recognize

$$\frac{\frac{\partial}{\partial \gamma} \sigma_{j-1}^2(\alpha, \beta, \gamma)}{\sigma_{j-1}^2(\alpha, \beta, \gamma)} = \frac{1}{\gamma}$$

and deduce

$$\frac{\partial}{\partial \gamma} \mathbb{H}_{2,n}(\gamma) = -\frac{n}{2\gamma} + \frac{1}{2\gamma} \sum_{j=1}^n \frac{1}{\sigma_{j-1}^2(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}, \gamma)} \left(Z_{t_j} - \mu_{j-1}(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}) \right)^2,$$

which is zero only for

$$\widehat{\gamma}_{0,n} := \frac{1}{n} \sum_{j=1}^n \frac{\left(Z_{t_j} - \mu_{j-1}(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}) \right)^2}{\frac{1}{\widehat{\beta}_{0,n}} \left(1 - e^{-\widehat{\beta}_{0,n}h}\right) \left(e^{-\widehat{\beta}_{0,n}h} Z_{t_{j-1}} + \frac{\widehat{\alpha}_{0,n}}{2\widehat{\beta}_{0,n}} \left(1 - e^{-\widehat{\beta}_{0,n}h}\right) \right)}. \tag{3.10}$$

Moreover, we derive

$$\begin{aligned}
\frac{\partial^2}{\partial\gamma^2}\mathbb{H}_{2,n}(\widehat{\gamma}_{0,n}) &= \frac{n}{2\widehat{\gamma}_{0,n}^2} - \frac{1}{\widehat{\gamma}_{0,n}^2}\sum_{j=1}^{n}\frac{1}{\sigma_{j-1}^2(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n},\widehat{\gamma}_{0,n})}\big(Z_{t_j}-\mu_{j-1}(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n})\big)^2 \\
&= \frac{n}{2\widehat{\gamma}_{0,n}^2} - \frac{2}{\widehat{\gamma}_{0,n}}\left(\frac{\partial}{\partial\gamma}\mathbb{H}_{2,n}(\widehat{\gamma}_{0,n}) + \frac{n}{2\widehat{\gamma}_{0,n}}\right) \\
&= \frac{n}{2\widehat{\gamma}_{0,n}^2} - \frac{n}{\widehat{\gamma}_{0,n}^2} = -\frac{n}{2\widehat{\gamma}_{0,n}^2} < 0,
\end{aligned}$$

so that (3.10) is a local maximum of $\mathbb{H}_{2,n}$. Additionally, $\lim_{\gamma\to\infty}\mathbb{H}_{2,n}(\gamma) = -\infty$ due to $\lim_{\gamma\to\infty}\sigma_{j-1}^2(\alpha,\beta,\gamma) = \infty$ and similarly $\lim_{\gamma\to 0}\mathbb{H}_{2,n}(\gamma) = -\infty$ leads to a global maximum of $\mathbb{H}_{2,n}$.

Raymond J. Carroll and David Ruppert call this a pseudo-likelihood method, see [18, 3.2 Pseudo-likelihood estimation of variance functions]. Next, we examine the asymptotic behaviour of the estimator $\widehat{\theta}_{0,n} = (\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n},\widehat{\gamma}_{0,n})$. In [72], although this estimator is introduced, asymptotic normality is shown only for the preliminary estimator $(\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n})$. We do not show asymptotic normality below since we do not need this for our results, which overall leads to a different proof than one finds therein. For the asymptotic proof we will apply the delta method, [86, p. 73].

Theorem 3.2: Given random variables $(X_n)_{n\in\mathbb{N}}$ and X with values in $\mathbb{R}^2$, a function $g:\mathbb{R}^2\to\mathbb{R}^2$ differentiable in a and a real sequence $(c_n)_{n\in\mathbb{N}}$ satisfying $\lim_{n\to\infty} c_n = \infty$, then from

$$\lim_{n\to\infty} c_n(X_n - a) = X$$

in distribution follows

$$\lim_{n\to\infty} c_n\big(g(X_n)-g(a)\big) = \big(\mathrm{D}g(a)\big)^\top X$$

in distribution, where D denotes the total differential.

For the following proof, we introduce some Landau symbols for a random variable ζ_n. We write $\zeta_n = O_p(r_n)$ for some positive sequence $r_n > 0$ if $\limsup_{n\to\infty}|r_n^{-1}\zeta_n| < \infty$ in distribution and $\zeta_n = o_p(r_n)$ if $\lim_{n\to\infty} r_n^{-1}\zeta_n(\theta) = 0$ in distribution, respectively. In case of a vector or matrix the notation should be understood component-wise. Furthermore, we define the rate of convergence

$$\begin{aligned}\mathbb{D}_n :&= \operatorname{diag}\big(\sqrt{T_n},\, \sqrt{T_n},\, \sqrt{n}\big)\\ &= \operatorname{diag}\big(\sqrt{nh},\, \sqrt{nh},\, \sqrt{n}\big).\end{aligned} \tag{3.11}$$

Lemma 3.3: For every true parameter $\theta_0 \in \Theta$ the asymptotic

$$\mathbb{D}_n\big(\widehat{\theta}_{0,n} - \theta_0\big) = O_p(1)$$

holds.

Proof: Step 1: We initially verify $\sqrt{nh}\big(\widehat{\beta}_{0,n} - \beta_0, \widehat{\alpha}_{0,n} - \alpha_0\big) = O_p(1)$. We specify $\rho := (\alpha, \beta) \in \Theta_\rho \subset (0,\infty)^2$ with Θ_ρ denoting the parameter space of ρ and $\widehat{\rho}_{0,n} := (\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n})$. First, we want to rewrite the function $\mathbb{H}_{1,n}$ in a compact way. Therefore, we define a few objects:

$$V(\rho) = V(\alpha,\beta) := \begin{pmatrix} e^{-\beta h} \\ \frac{\alpha}{\beta}\big(1 - e^{-\beta h}\big)\end{pmatrix}, \quad z_{j-1} := \Big(Z_{t_{j-1}}, 1\Big)$$

such that

$$z_{j-1}V(\rho) = \mu_{j-1}(\rho)$$

holds and by setting $\boldsymbol{z}_n := (z_0, \dots, z_{n-1})^\top$ and $\boldsymbol{y}_n := (Z_{t_1}, \dots, Z_{t_n})^\top$ we receive

$$\mathbb{H}_{1,n}(\rho) = -\|\boldsymbol{y}_n - \boldsymbol{z}_n V(\rho)\|^2.$$

The corresponding estimating equation $\nabla \mathbb{H}_{1,n} = \mathbf{0}$ is equivalent to

$$\begin{aligned}\mathbf{0} = \nabla_V \mathbb{H}_{1,n} &= \nabla_V\Big(-\|\boldsymbol{y}_n\|^2 + 2\langle \boldsymbol{y}_n, \boldsymbol{z}_n V\rangle - \|\boldsymbol{z}_n V\|^2\Big)\\ &= \nabla_V\Big(-\|\boldsymbol{y}_n\|^2 + 2\big\langle \boldsymbol{z}_n^\top \boldsymbol{y}_n, V\big\rangle - V^\top \boldsymbol{z}_n^\top \boldsymbol{z}_n V\Big)\\ &= 2\big(\boldsymbol{z}_n^\top \boldsymbol{y}_n\big)^\top - 2V^\top \boldsymbol{z}_n^\top \boldsymbol{z}_n\\ &= 2\big(\boldsymbol{z}_n^\top \boldsymbol{y}_n\big)^\top - 2\big(\boldsymbol{z}_n^\top \boldsymbol{z}_n V\big)^\top.\end{aligned}$$

The known solution $\widehat{V}_n := V(\widehat{\rho}_{0,n})$, see Lemma 3.1, is due to this new expression also given by $V(\widehat{\rho}_{0,n}) = (\boldsymbol{z}_n^\top \boldsymbol{z}_n)^{-1}\boldsymbol{z}_n^\top \boldsymbol{y}_n$. Using Lemma 1.8 (iii) we notice

$$\lim_{n\to\infty} \frac{1}{n}\boldsymbol{z}_n^\top \boldsymbol{z}_n = \lim_{n\to\infty} \begin{pmatrix} \frac{1}{n}\sum\limits_{j=1}^n Z_{t_j}^2 & \frac{1}{n}\sum\limits_{j=1}^n Z_{t_j} \\ \frac{1}{n}\sum\limits_{j=1}^n Z_{t_j} & 1 \end{pmatrix} = \begin{pmatrix} \int\limits_0^\infty z^2\pi_{\theta_0}(z)\,\mathrm{d}z & \int\limits_0^\infty z\pi_{\theta_0}(z)\,\mathrm{d}z \\ \int\limits_0^\infty z\pi_{\theta_0}(z)\,\mathrm{d}z & 1 \end{pmatrix}$$

$$\overset{1.6}{=} \begin{pmatrix} \frac{\alpha_0(2\alpha_0+\gamma_0)}{2\beta_0^2} & \frac{\alpha_0}{\beta_0} \\ \frac{\alpha_0}{\beta_0} & 1 \end{pmatrix}. \tag{3.12}$$

In particular, the determinant of this matrix is

$$\frac{\alpha_0(2\alpha_0+\gamma_0)}{2\beta_0^2} - \frac{\alpha_0^2}{\beta_0^2} = \frac{\alpha_0\gamma_0}{2\beta_0^2} > 0.$$

By (1.10), we have

$$\boldsymbol{y}_n = \boldsymbol{z}_n V(\rho_0) + M_n, \tag{3.13}$$

where $M_n = \big(M_{n,1},\dots,M_{n,n}\big)^\top$ is defined by

$$M_{n,j} := \gamma_0 \int\limits_{t_{j-1}}^{t_j} e^{-(t_j-s)\beta_0}\sqrt{Z_s}\,\mathrm{d}B_s$$

and fulfills

$$\begin{aligned}
\sup_{n\in\mathbb{N}}\sup_{1\le j\le n} \mathbb{E}\left(\left|h^{-\frac{1}{2}}M_{n,j}\right|^q\right) &\le \sup_{n\in\mathbb{N}}\sup_{1\le j\le n} \mathbb{E}\left(\left|h^{-1}\big[M_{n,j}\big]\right|^{\frac{q}{2}}\right) \\
&= \sup_{n\in\mathbb{N}}\sup_{1\le j\le n} \mathbb{E}\left(\left(h^{-1}\gamma_0^2\int\limits_{t_{j-1}}^{t_j} e^{-2(t_j-s)\beta_0}Z_s\,\mathrm{d}s\right)^{\frac{q}{2}}\right) \\
&= \sup_{n\in\mathbb{N}}\sup_{1\le j\le n} \int\limits_0^\infty \left(h^{-1}\gamma_0^2\int\limits_{t_{j-1}}^{t_j} e^{-2(t_j-s)\beta_0}z\,\mathrm{d}s\right)^{\frac{q}{2}}\pi_{\theta_0}(z)\,\mathrm{d}z \\
&= \sup_{n\in\mathbb{N}}\sup_{1\le j\le n} \left(\gamma_0^2\frac{1-e^{-2h\beta_0}}{2h\beta_0}\right)^{\frac{q}{2}}\int\limits_0^\infty z^{\frac{q}{2}}\pi_{\theta_0}(z)\,\mathrm{d}z \\
&\overset{1.6}{=} \sup_{n\in\mathbb{N}}\sup_{1\le j\le n} \left(\gamma_0^2\frac{1-e^{-2h\beta_0}}{2h\beta_0}\right)^{\frac{q}{2}}\frac{\Gamma\big(\frac{1}{2}+\frac{2\alpha}{\gamma}\big)}{\Gamma\big(\frac{2\alpha}{\gamma}\big)}\sqrt{\frac{\gamma}{2\beta}}
\end{aligned}$$

$$\leq \gamma_0^q \frac{\Gamma\big(\frac{1}{2} + \frac{2\alpha}{\gamma}\big)}{\Gamma\big(\frac{2\alpha}{\gamma}\big)} \sqrt{\frac{\gamma}{2\beta}} < \infty$$

for every $q > 0$ due to Burkholder inequality in the first line and mean value theorem in the last line, which implies $\frac{e^{-x}-e^{-y}}{y-x} \leq 1$ for every $0 \leq x < y$. By an analogous calculation, we again receive the finiteness by replacing $M_{n,j}$ with $M_{n,j}Z_{t_{j-1}}$. Since $\big\{\big(h^{-\frac{1}{2}}M_{n,j}z_{j-1}, \mathcal{F}_{t_j}\big)\big\}_{1\leq j\leq n}$ forms a martingale-difference array satisfying

$$\sup_{n\in\mathbb{N}} \sup_{1\leq j\leq n} \mathbb{E}\left(\left|h^{-\frac{1}{2}}M_{n,j}z_{j-1}\right|^q\right) \leq \sup_{n\in\mathbb{N}} \sup_{1\leq j\leq n} \sum_{k=0}^{\lfloor q\rfloor} \binom{q}{k} \mathbb{E}\left(\left|h^{-\frac{1}{2}}M_{n,j}\right|^q + \left|h^{-\frac{1}{2}}M_{n,j}Z_{t_{j-1}}\right|^q\right) < \infty$$

for every $q > 0$ due to the triangle inequality and the binomial theorem, we conclude

$$\begin{aligned}
\sqrt{\frac{n}{h}}\big(V(\widehat{\rho}_{0,n}) - V(\rho_0)\big) &= \sqrt{\frac{n}{h}}\Big((\boldsymbol{z}_n^\top \boldsymbol{z}_n)^{-1}\boldsymbol{z}_n^\top \boldsymbol{y}_n - V(\rho_0)\Big) \\
&\overset{(3.13)}{=} \sqrt{\frac{n}{h}}\Big((\boldsymbol{z}_n^\top \boldsymbol{z}_n)^{-1}\boldsymbol{z}_n^\top\big(\boldsymbol{z}_n V(\rho_0) + M_n\big) - V(\rho_0)\Big) \\
&= \left(\frac{1}{n}\boldsymbol{z}_n^\top \boldsymbol{z}_n\right)^{-1} \frac{1}{\sqrt{n}} \sum_{j=1}^n h^{-\frac{1}{2}} M_{n,j} z_{j-1} \overset{(3.12)}{=} O_p(1).
\end{aligned}$$

We apply the delta method, Theorem 3.2, on the function $(x, y) \mapsto (\log x, y)$, reminding $V(\rho) = \Big(e^{-\beta h}, \frac{\alpha}{\beta}\big(1 - e^{-\beta h}\big)\Big)^\top$, to obtain

$$\sqrt{\frac{n}{h}}\left((-\widehat{\beta}_{0,n}h) - (-\beta_0 h),\ \frac{\widehat{\alpha}_{0,n}h}{\widehat{\beta}_{0,n}h}(1 - e^{-\widehat{\beta}_{0,n}h}) - \frac{\alpha_0 h}{\beta_0 h}(1 - e^{-\beta_0 h})\right) = O_p(1).$$

Another application of the delta method with the function $(x, y) \mapsto (\frac{xy}{e^x - 1}, -x)$ yields

$$\begin{aligned}
O_p(1) &= \sqrt{\frac{n}{h}}\left(-\widehat{\beta}_{0,n}h\frac{\widehat{\alpha}_{0,n}}{\widehat{\beta}_{0,n}} \cdot \frac{1 - e^{-\widehat{\beta}_{0,n}h}}{e^{-\widehat{\beta}_{0,n}h} - 1} + \beta_0 h \frac{\alpha_0}{\beta_0} \cdot \frac{1 - e^{-\beta_0 h}}{e^{-\beta_0 h} - 1}, \widehat{\beta}_{0,n}h - \beta_0 h\right) \\
&= \sqrt{\frac{n}{h}}\left(\widehat{\alpha}_{0,n}h - \alpha_0 h, \widehat{\beta}_{0,n}h - \beta_0 h\right) \\
&= \sqrt{nh}\left(\widehat{\alpha}_{0,n} - \alpha_0, \widehat{\beta}_{0,n} - \beta_0\right).
\end{aligned}$$

Step 2: Next, we verify $\sqrt{n}(\widehat{\gamma}_{0,n} - \gamma_0) = O_p(1)$. Before proceeding to the proof, we make a few remarks on the notation. To emphasize the independence of γ, we introduce

$$c(z,\rho) = \gamma^{-1}\sigma^2(z,\theta), \tag{3.14}$$

where $\sigma^2(z,\theta)$ is defined by (3.5), replacing $Z_{t_{j-1}}$ by z and $\mu(z,\rho)$ by (3.4) accordingly. In what follows, we use $\mu_{j-1}(\rho) := \mu(Z_{t_{j-1}},\rho)$, $\sigma^2_{j-1}(\rho) := \sigma^2(Z_{t_{j-1}},\theta)$ and $c_{j-1}(\rho) := c(Z_{t_{j-1}},\rho)$ for a shorter notation. We recall

$$\sqrt{n}(\widehat{\gamma}_{0,n} - \gamma_0) \underset{(3.10)}{\overset{(3.5)}{=}} \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{\big(Z_{t_j} - \mu_{j-1}(\widehat{\rho}_{0,n})\big)^2}{c_{j-1}(\widehat{\rho}_{0,n})} - \gamma_0\right)$$

and split this expression in four summands which we will analyze separately:

$$\begin{aligned}
G_{1,n}(\rho) &:= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} 2\frac{\mu_{j-1}(\rho_0) - \mu_{j-1}(\rho)}{c_{j-1}(\rho)}\left(Z_{t_j} - \mu_{j-1}(\rho_0)\right),\\
G_{2,n}(\rho) &:= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} \frac{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2 - \gamma c_{j-1}(\rho_0)}{c_{j-1}(\rho)},\\
G_{3,n} &:= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} \frac{\big(\mu_{j-1}(\rho_0) - \mu_{j-1}(\widehat{\rho}_{0,n})\big)^2}{c_{j-1}(\widehat{\rho}_{0,n})},\\
G_{4,n} &:= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} \frac{\gamma_0\big(c_{j-1}(\rho_0) - c_{j-1}(\widehat{\rho}_{0,n})\big)}{c_{j-1}(\widehat{\rho}_{0,n})}\\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} \gamma_0\left(\frac{c_{j-1}(\rho_0)}{c_{j-1}(\widehat{\rho}_{0,n})} - 1\right).
\end{aligned}$$

We can easily see

$$G_{2,n}(\widehat{\rho}_{0,n}) + G_{4,n} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2}{c_{j-1}(\widehat{\rho}_{0,n})} - \gamma_0\right)$$

and for the sum over the numerators with common denominator $c_{j-1}(\widehat{\rho}_{0,n})$ we observe

$$\begin{aligned}
&\underbrace{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2}_{\text{from } G_{2,n}(\widehat{\rho}_{0,n}) + G_{4,n}} + \underbrace{2\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)\cdot\big(\mu_{j-1}(\rho_0) - \mu_{j-1}(\widehat{\rho}_{0,n})\big)}_{\text{from } G_{1,n}(\widehat{\rho}_{0,n})} + \underbrace{\big(\mu_{j-1}(\rho_0) - \mu_{j-1}(\widehat{\rho}_{0,n})\big)^2}_{\text{from } G_{3,n}}\\
&= \big(Z^2_{t_{j-1}} - \cancel{2\mu_{j-1}(\rho)Z_{t_j}} + \cancel{\mu^2_{j-1}(\rho_0)}\big) + \big(\cancel{2Z_{t_j}\mu_{j-1}(\rho_0)} - 2Z_{t_j}\mu_{j-1}(\widehat{\rho}_{0,n}) - \cancel{2\mu^2_{j-1}(\rho_0)}\\
&\qquad + \cancel{2\mu_{j-1}(\widehat{\rho}_{0,n})\mu_{j-1}(\rho_0)}\big) + \big(\cancel{\mu^2_{j-1}(\rho_0)} - \cancel{2\mu_{j-1}(\rho_0)\mu_{j-1}(\widehat{\rho}_{0,n})} + \mu^2_{j-1}(\widehat{\rho}_{0,n})\big)\\
&= \big(Z_{t_j} - \mu_{j-1}(\widehat{\rho}_{0,n})\big)^2,
\end{aligned}$$

which implies

$$\begin{aligned}\sqrt{n}(\widehat{\gamma}_{0,n}-\gamma_0) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{\big(Z_{t_j}-\mu_{j-1}(\widehat{\rho}_{0,n})\big)^2}{c_{j-1}(\widehat{\rho}_{0,n})}-\gamma_0\right)\\ &= G_{1,n}(\widehat{\rho}_{0,n})+G_{2,n}(\widehat{\rho}_{0,n})+G_{3,n}+G_{4,n}.\end{aligned}$$

In the following, we use the notation $a_n \lesssim b_n$ if $\sup\big\{\frac{a_n}{b_n}\,|\,n\in\mathbb{N}\big\}$ is almost surely bounded by some universal constant. We start to bound $G_{3,n}$ and $G_{4,n}$. For both we notice that

$$\begin{aligned}0\le \frac{h}{c_{j-1}(\widehat{\rho}_{0,n})} &\overset{(3.5)}{=} \frac{\widehat{\beta}_{0,n}h}{\big(1-e^{-\widehat{\beta}_{0,n}h}\big)\big[e^{-\widehat{\beta}_{0,n}h}Z_{t_{j-1}}+\frac{\widehat{\alpha}_{0,n}}{2\widehat{\beta}_{0,n}}\big(1-e^{-\widehat{\beta}_{0,n}h}\big)\big]}\\ &\le \frac{\widehat{\beta}_{0,n}h}{\big(e^{-\widehat{\beta}_{0,n}h}-e^{-2\widehat{\beta}_{0,n}h}\big)Z_{t_{j-1}}}\\ &= \frac{1}{e^{-\xi_{n,h}}Z_{t_{j-1}}}\end{aligned}$$

for some $\xi_{n,h}\in(\widehat{\beta}_{0,n}h, 2\widehat{\beta}_{0,n}h)\subset(0,\infty)$ due to the mean value theorem. In particular, $\xi_{n,h}\to 0$ almost surely for $nh\to\infty$ while $h\to 0$ and hence

$$\sup_{n\in\mathbb{N}}\left|\frac{h}{c_{j-1}(\widehat{\rho}_{0,n})}\right| \lesssim Z_{t_{j-1}}^{-1}. \tag{3.15}$$

Therefore, we conclude

$$\begin{aligned}|G_{3,n}| &\le \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left|\frac{h}{c_{j-1}(\widehat{\rho}_{0,n})}\right|\frac{\big(\mu_{j-1}(\rho_0)-\mu_{j-1}(\widehat{\rho}_{0,n})\big)^2}{h}\\ &\lesssim \frac{2}{\sqrt{n}}\sum_{j=1}^{n}Z_{t_{j-1}}^{-1}\Bigg(Z_{t_{j-1}}^2\frac{\big(e^{-\beta_0 h}-e^{-\widehat{\beta}_{0,n}h}\big)^2}{h}\\ &\qquad +\frac{1}{h}\Big(\frac{\alpha_0}{\beta_0}\big(1-e^{-\beta_0 h}\big)-\frac{\widehat{\alpha}_{0,n}}{\widehat{\beta}_{0,n}}\big(1-e^{-\widehat{\beta}_{0,n}h}\big)\Big)^2\Bigg).\end{aligned}$$

The last inequality is valid since $(a-b)^2\le 2a^2+2b^2$ for every $a,b\in\mathbb{R}$ and by virtue of (3.4). Now, we analyze the two summands separately. By the mean value theorem and the first step, we can easily see

$$\frac{\big(e^{-\beta_0 h}-e^{-\widehat{\beta}_{0,n}h}\big)^2}{h} \le \Big(\sup_{x\in(0,\infty)} e^{-x}\Big)^2\frac{\big(\widehat{\beta}_{0,n}h-\beta_0 h\big)^2}{h}$$

$$
\begin{aligned}
&\le h\big(\widehat{\beta}_{0,n}-\beta_0\big)^2\\
&= n^{-1}\big(\sqrt{nh}(\widehat{\beta}_{0,n}-\beta_0)\big)^2\\
&= n^{-1}O_p(1)
\end{aligned}
$$

and analogously

$$
\begin{aligned}
&\frac{1}{h}\Big(\frac{\alpha_0}{\beta_0}\big(1-e^{-\beta_0 h}\big)-\frac{\widehat{\alpha}_{0,n}}{\widehat{\beta}_{0,n}}\big(1-e^{-\widehat{\beta}_{0,n}h}\big)\Big)^2\\
&\le \frac{2}{h}\Big(\frac{\alpha_0}{\beta_0}\big(1-e^{-\beta_0 h}\big)-\frac{\widehat{\alpha}_{0,n}}{\beta_0}\big(1-e^{-\beta_0 h}\big)\Big)^2+\frac{2}{h}\left(\frac{\widehat{\alpha}_{0,n}}{\beta_0}\big(1-e^{-\beta_0 h}\big)-\frac{\widehat{\alpha}_{0,n}}{\widehat{\beta}_{0,n}}\big(1-e^{-\widehat{\beta}_{0,n}h}\big)\right)^2\\
&= 2h\big(\widehat{\alpha}_{0,n}-\alpha_0\big)^2\left(\frac{1-e^{-\beta_0 h}}{\beta_0 h}\right)^2+2h\widehat{\alpha}_{0,n}^2\left(\frac{1-e^{-\beta_0 h}}{\beta_0 h}-\frac{1-e^{-\widehat{\beta}_{0,n}h}}{\widehat{\beta}_{0,n}}\right)^2\\
&\le 2h\big(\widehat{\alpha}_{0,n}-\alpha_0\big)^2+2h\widehat{\alpha}_{0,n}^2\sup_{x\in(0,\infty)}\left(\frac{\mathrm{d}}{\mathrm{d}x}\frac{1-e^{-x}}{x}\right)^2\big(\beta_0 h-\widehat{\beta}_{0,n}h\big)^2\\
&= 2h\big(\widehat{\alpha}_{0,n}-\alpha_0\big)^2+2h\widehat{\alpha}_{0,n}^2\big(\beta_0 h-\widehat{\beta}_{0,n}h\big)^2\\
&= 2n^{-1}\Big(\sqrt{nh}\big(\widehat{\alpha}_{0,n}-\alpha_0\big)\Big)^2+2h^2n^{-1}\widehat{\alpha}_{0,n}^2\Big(\sqrt{nh}\big(\widehat{\beta}_{0,n}-\beta_0\big)\Big)^2\\
&= n^{-1}O_p(1)+n^{-1}o_p(1)=n^{-1}O_p(1).
\end{aligned}
$$

In particular, the last line in the calculations follows directly by the first step whereas the third and fourth line are again valid due to the mean value theorem. Combining the calculations implies

$$
\big|G_{3,n}\big|\lesssim\frac{1}{\sqrt{n}}\left[\frac{1}{n}\sum_{j=1}^n Z_{t_{j-1}}+\frac{1}{n}\sum_{j=1}^n Z_{t_{j-1}}^{-1}\right]O_p(1)
$$

and the terms in the brackets are $O_p(1)$ by the law of large numbers, Lemma 1.8 (iii). Likewise, by (3.5) and the mean value theorem we derive

$$
\begin{aligned}
|G_{4,n}| &\overset{(3.15)}{\lesssim} \frac{1}{\sqrt{n}}\sum_{j=1}^n Z_{t_{j-1}}^{-1}\gamma_0\Bigg(\left|\frac{e^{-\beta_0 h}-e^{-2\beta_0 h}}{\beta_0 h}-\frac{e^{-\widehat{\beta}_{0,n}h}-e^{-2\widehat{\beta}_{0,n}h}}{\widehat{\beta}_{0,n}h}\right|Z_{t_{j-1}}\\
&\qquad\qquad+\left|\frac{\alpha_0}{2}\cdot\frac{\big(1-e^{-\beta_0 h}\big)^2}{\beta_0^2 h}-\frac{\widehat{\alpha}_{0,n}}{2}\cdot\frac{\big(1-e^{-\widehat{\beta}_{0,n}h}\big)^2}{\widehat{\beta}_{0,n}^2 h}\right|\Bigg)\\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^n Z_{t_{j-1}}^{-1}\gamma_0\Big(\big|f'(\chi_{n,h})\big|\cdot\big|\widehat{\beta}_{0,n}h-\beta_0 h\big|Z_{t_{j-1}}
\end{aligned}
$$

$$+ \left|\left\langle \nabla g\big(\xi_{n,h}, \zeta_{n,h}\big), h\big(\widehat{\rho}_{0,n} - \rho_0\big)\right\rangle\right|\Big)$$

for the functions $f(x) := \frac{e^{-x}-e^{-2x}}{x}$ and $g(x,y) := \frac{x}{2} \cdot \frac{(1-e^{-y})^2}{y^2}$ and some interior points $\chi_{n,h}$ and $\xi_{n,h}$. Since $\big(\widehat{\alpha}_{0,n}, \widehat{\beta}_{0,n}\big) \to \big(\alpha_0, \beta_0\big)$ we immediately observe $\chi_{n,h} \to 0$ and $\big(\xi_{n,h}, \zeta_{n,h}\big) \to \big(0, 0\big)$ for $n \to \infty$. Therefore, the calculations

$$f'(\chi_{n,h}) = \frac{e^{-2\chi_{n,h}}\big(2\chi_{n,h} - e^{\chi_{n,h}}(\chi_{n,h}+1)+1\big)}{\chi_{n,h}^2} \to \frac{3}{2},$$

$$\nabla g\big(\xi_{n,h} \quad \zeta_{n,h}\big) = \begin{pmatrix} \frac{(1-e^{-\zeta_{n,h}})}{2\zeta_{n,h}^2} \\ \xi_{n,h}\frac{e^{-2\zeta_{n,h}}(e^{\zeta_{n,h}}-1)(-\zeta_{n,h}+e^{\zeta_{n,h}}-1)}{\zeta_{n,h}^3} \end{pmatrix} \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

result in

$$\begin{aligned} |G_{4,n}| &\lesssim \frac{1}{\sqrt{n}} \sum_{j=1}^n \Big(h\big(1 + Z_{t_{j-1}}^{-1}\big)\big|\widehat{\beta}_{0,n} - \beta_0\big| + hZ_{t_{j-1}}^{-1}\big|\widehat{\alpha}_{0,n} - \alpha_0\big|\Big) \\ &= \sqrt{h}\Big(1 + \frac{1}{n}\sum_{j=1}^n Z_{t_{j-1}}^{-1}\Big)\sqrt{nh}\big|\widehat{\beta}_{0,n} - \beta_0\big| + \sqrt{h}\Big(\frac{1}{n}\sum_{j=1}^n Z_{t_{j-1}}^{-1}\Big)\sqrt{nh}\big|\widehat{\alpha}_{0,n} - \alpha_0\big|, \end{aligned}$$

which is in $O_p(\sqrt{h})$ due to the first step and the law of large numbers, Lemma 1.8 (iii). It remains to deduce that both $G_{i,n}(\widehat{\rho}_{0,n}) = O_p(1)$ for $i = 1, 2$. We regard $G_{i,n}(\rho)$ as stochastic processes in $\mathcal{C}(\overline{\Theta}_\rho)$. Since Burkholder's inequality ensures that $G_{i,n}(\rho) = O_p(1)$ for each $\rho \in \overline{\Theta}_\rho$, it suffices to verify the tightness of $G_{i,n}$ in $\mathcal{C}(\overline{\Theta}_\rho)$. In view of the Kolmogorov tightness criterion [47], it is in turn sufficient to show the moment bound

$$\exists \delta, q, C > 0 \ \forall \rho, \rho' \in \Theta_\rho : \quad \sup_{n\in\mathbb{N}} \mathbb{E}\left(\big|G_{i,n}(\rho) - G_{i,n}(\rho')\big|^q\right) \le C|\rho - \rho'|^{2+\delta}. \tag{3.16}$$

We perform this calculation only for $i = 1$, since it is similar in the other case. We can easily derive

$$\sup_{\rho\in\Theta_\rho} \Big|\nabla g_{j-1}(\rho)\Big| \lesssim 1 + Z_{t_{j-1}}^{-1} + Z_{t_{j-1}}^{-2} \tag{3.17}$$

for the function

$$g_{j-1}(\rho) = 2\frac{\mu_{j-1}(\rho_0) - \mu_{j-1}(\rho)}{c_{j-1}(\rho)}.$$

Therefore, we simplify

$$\begin{aligned}\frac{\mu_{j-1}(\rho)}{c_{j-1}(\rho)} &= \frac{\beta}{1-e^{-\beta h}} + \frac{\frac{\alpha}{2\beta}\big(1-e^{-\beta h}\big)}{\frac{1-e^{-\beta h}}{\beta}\big[e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{2\beta}(1-e^{-\beta h})\big]}\\ &= \frac{\beta}{1-e^{-\beta h}} + \frac{\alpha}{2e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{\beta}(1-e^{-\beta h})}\end{aligned}$$

and then calculate

$$\begin{aligned}\left|\frac{\partial}{\partial\alpha}\frac{\mu_{j-1}(\rho)}{c_{j-1}(\rho)}\right| &= \left|\frac{1\cdot\big[2e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{\beta}(1-e^{-\beta h})\big] - \alpha\cdot\frac{1}{\beta}(1-e^{-\beta h})}{\big(2e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{\beta}(1-e^{-\beta h})\big)^2}\right|\\ &= \frac{2e^{-\beta h}Z_{t_{j-1}}}{\big(2e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{\beta}(1-e^{-\beta h})\big)^2}\\ &\le \frac{2e^{-\beta h}Z_{t_{j-1}}}{\big(2e^{-\beta h}Z_{t_{j-1}}\big)^2}\\ &\lesssim Z_{t_{j-1}}^{-1}\end{aligned}$$

and

$$\begin{aligned}\left|\frac{\partial}{\partial\beta}\frac{\mu_{j-1}(\rho)}{c_{j-1}(\rho)}\right| &= \left| -\frac{e^{-\beta h}(\beta h+1)-1}{\beta^2} - \frac{\alpha\big[-2he^{-\beta h}Z_{t_{j-1}} + \alpha\cdot\frac{e^{-\beta h}(\beta h+1)-1}{\beta^2}\big]}{\big(2e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{\beta}(1-e^{-\beta h})\big)^2}\right|\\ &\lesssim 1 + \frac{2\alpha he^{-\beta h}Z_{t_{j-1}}}{\big(2e^{-\beta h}Z_{t_{j-1}}\big)^2} + \left|\frac{\alpha e^{-\beta h}(\beta h+1)-1}{\beta^2\big(2e^{-\beta h}Z_{t_{j-1}}\big)^2}\right|\\ &\lesssim 1 + Z_{t_{j-1}}^{-1} + Z_{t_{j-1}}^{-2},\end{aligned}$$

which completes the proof of (3.17). We remark here that the boundedness of Θ_ρ is essential. Furthermore, we derive

$$\begin{aligned}\mathbb{E}\left(|Z_{t_j} - \mu_{j-1}(\rho_0)|^q \mid Z_{t_{j-1}}\right) &\lesssim \mathbb{E}\left(|Z_{t_j}|^q + |\mu_{j-1}(\rho_0)|^q \mid Z_{t_{j-1}}\right)\\ &= \mathbb{E}\left(|Z_{t_j}|^q \mid Z_{t_{j-1}}\right) + |\mu_{j-1}(\rho_0)|^q\\ &\lesssim \Big(\big(1+Z_{t_{j-1}}^q\big) + \big(1+Z_{t_{j-1}}\big)^q\Big)\\ &\lesssim \big(1+Z_{t_{j-1}}\big)^q\end{aligned}$$

for every $q > 1$. In particular, the inequality of the conditional expectation in the second-last line was proved in [6, Lemma 4.2] whereas $1 + x^q \le (1+x)^q$ is valid for every $x \ge 0$ due to the binomial theorem. By means of Burkholder's and Jensen's inequalities, we see that

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \mathbb{E}\left(\left|G_{1,n}(\rho)-G_{1,n}(\rho')\right|^q\right)\\
&=\sup_{n\in\mathbb{N}} \mathbb{E}\left(\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(g_{j-1}(\rho)-g_{j-1}(\rho')\right)\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)\right|^q\right)\\
&\lesssim \sup_{n\in\mathbb{N}} \mathbb{E}\left(\left|\frac{1}{n}\sum_{j=1}^{n}\left(g_{j-1}(\rho)-g_{j-1}(\rho')\right)^2\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^2\right|^{\frac{q}{2}}\right)\\
&\lesssim \sup_{n\in\mathbb{N}} \frac{1}{n^{\frac{q}{2}}}\sum_{j=1}^{n}\mathbb{E}\left(\left|g_{j-1}(\rho)-g_{j-1}(\rho')\right|^q\,\mathbb{E}\left(|Z_{t_j}-\mu_{j-1}(\rho_0)|^q \mid Z_{t_{j-1}}\right)\right)\\
&\lesssim \sup_{n\in\mathbb{N}} \frac{1}{n^{\frac{q}{2}}}\sum_{j=1}^{n}\mathbb{E}\left(\sup_{\rho\in\Theta_\rho}\left|\nabla g_{j-1}(\rho)\right|^q(1+Z_{t_{j-1}})^q\right)|\rho-\rho'|^q\\
&\lesssim \sup_{n\in\mathbb{N}} \frac{1}{n^{\frac{q}{2}}}\sum_{j=1}^{n}\mathbb{E}\left(\left(1+Z_{t_{j-1}}^{-1}+Z_{t_{j-1}}^{-2}\right)\left(1+Z_{t_{j-1}}\right)^q\right)|\rho-\rho'|^q\\
&=\frac{1}{n^{\frac{q}{2}-1}}\int_0^\infty\left(1+x^{-1}+x^{-2}\right)\left(1+x\right)^q\pi_{\theta_0}(x)\,\mathrm{d}x\cdot|\rho-\rho'|^q\\
&\lesssim |\rho-\rho'|^q.
\end{aligned}$$

Similarly, we can deduce (3.16) for $G_{2,n}$ by using

$$\sup_{\rho\in\Theta_\rho}\left|\nabla_\rho\,\frac{1}{c_{j-1}(\rho)}\right|\lesssim Z_{t_{j-1}}^{-1}+Z_{t_{j-1}}^{-2},$$

which completes the proof of $\sqrt{n}(\widehat{\gamma}_{0,n}-\gamma_0)=O_p(1)$. □

We use the preliminary estimator $\widehat{\theta}_{0,n}$ in our following section. There, we show the asymptotic behavior of the GQMLE and consider one step improvements of our preliminary estimator, which are asymptotically equivalent to the GQMLE.

3.2 Gaussian quasi-likelihood for the Cox-Ingersoll-Ross process

3.2.1 Asymptotics for joint Gaussian quasi-maximum likelihood estimator

The objective now is to deduce the asymptotic normality of the GQMLE $\widehat{\theta}_n$, (3.7). In this section we use the notation $\zeta_n = O_p^*(r_n)$ if $\sup_{\theta\in\Theta}|\zeta_n(\theta)| = O_p(r_n)$ and $\zeta_n = o_p^*(r_n)$, respectively, for some positive sequence $r_n > 0$ and a random function ζ_n on Θ. We omit p if ζ_n is non-random.

For the proof of the asymptotic normality, we first turn to the consistency, which can be derived similarly to [58], through applying the argmax theorem twice, see [65, Lemma 6.6]. In particular, we will often use the following argument, [34, Lemma 9].

Lemma 3.4: For a series of $\mathcal{G}_{j,n} := \sigma\big(Z_s : s \le t_j\big)$-measurable random variables $\chi_{j,n}$ for $n \in \mathbb{N}$ and $j = 1, \ldots, n$ with a random variable χ satisfying

$$\lim_{n\to\infty}\sum_{j=1}^{n}\mathbb{E}\left(\chi_{j,n}\,|\,\mathcal{G}_{j-1,n}\right) = \chi,$$
$$\lim_{n\to\infty}\sum_{j=1}^{n}\mathbb{E}\left((\chi_{j,n})^2\,|\,\mathcal{G}_{j-1,n}\right) = 0$$

in probability we have

$$\lim_{n\to\infty}\sum_{j=1}^{n}\chi_{j,n} = \chi$$

in probability.

Lemma 3.5: For every true value $\theta_0 \in \Theta$

$$\lim_{n\to\infty}\widehat{\theta}_n = \theta_0$$

holds in probability.

Proof: At first stage we define

$$\mathbb{Y}_{1,n}(\gamma) := \frac{1}{n}\big(\mathbb{H}_n(\widehat{\rho}_n,\gamma) - \mathbb{H}_n(\widehat{\rho}_n,\gamma_0)\big),$$
$$\mathbb{Y}_{2,n}(\rho) := \frac{1}{T_n}\big(\mathbb{H}_n(\rho,\widehat{\gamma}_n) - \mathbb{H}_n(\rho_0,\widehat{\gamma}_n)\big)$$

to rewrite

$$\mathbb{H}_n(\theta) - \mathbb{H}_n(\theta_0) = n\mathbb{Y}_{1,n}(\gamma) + T_n\mathbb{Y}_{2,n}(\rho).$$

Using the argmax theorem [65, Lemma 6.6], the consistency follows immediately if we can find a suitable non-random function $\mathbb{Y}_{1,0}$ such that

$$\lim_{n\to\infty}\sup_{\gamma\in\Theta_\gamma}\big|\mathbb{Y}_{1,n}(\gamma) - \mathbb{Y}_{1,0}(\gamma)\big| = 0 \quad \text{in probability},$$
$$\operatorname{argmax}\ \mathbb{Y}_{1,0} = \gamma_0$$

and respectively some $\mathbb{Y}_{2,0}$ with

$$\lim_{n\to\infty}\sup_{\rho\in\Theta_\rho}\big|\mathbb{Y}_{2,n}(\rho) - \mathbb{Y}_{2,0}(\rho)\big| = 0 \quad \text{in probability},$$
$$\operatorname{argmax}\ \mathbb{Y}_{2,0} = \rho_0.$$

Again Θ_ρ and Θ_γ denote the parameter spaces of ρ and γ, respectively, so that $\Theta = \Theta_\rho \times \Theta_\gamma$. In the following, we keep the notation of the proof ofLemma 3.3 for $\mu_{j-1}(\rho)$ and $\sigma^2_{j-1}(\theta)$.

Case $\mathbb{Y}_{1,n}$: We start by rewriting:

$$\begin{aligned}
\mathbb{Y}_{1,n}(\gamma) &\overset{(3.6)}{=} -\frac{1}{2n}\sum_{j=1}^{n}\left(\log\left(\frac{\sigma^2_{j-1}(\rho_0,\gamma)}{\sigma^2_{j-1}(\theta_0)}\right) + \frac{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2}{\sigma^2_{j-1}(\rho_0,\gamma) - \sigma^2_{j-1}(\theta_0)}\right)\\
&\overset{(3.5)}{=} -\frac{1}{2n}\sum_{j=1}^{n}\left(\log\left(\frac{\gamma}{\gamma_0}\right) + \left(\frac{1}{\gamma} - \frac{1}{\gamma_0}\right)\frac{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2}{\gamma_0^{-1}\sigma^2_{j-1}(\theta_0)}\right)\\
&= -\frac{1}{2}\log\left(\frac{\gamma}{\gamma_0}\right) - \frac{1}{2}\left(\frac{\gamma_0}{\gamma} - 1\right)\frac{1}{n}\sum_{j=1}^{n}\frac{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2}{\sigma^2_{j-1}(\theta_0)}.
\end{aligned}$$

We can easily compute:

$$
\begin{aligned}
\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left(\frac{\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^2}{\sigma_{j-1}^2(\theta_0)}\,\Big|\,Z_{t_{j-1}}\right) &= 1,\\
0 \ \le \frac{1}{n^2}\sum_{j=1}^{n}\mathbb{E}\left(\frac{\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^4}{\sigma_{j-1}^4(\theta_0)}\,\Big|\,Z_{t_{j-1}}\right)& \\
= \ \frac{1}{n^2}\sum_{j=1}^{n}\frac{\mathbb{E}\left(\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^4\,|\,Z_{t_{j-1}}\right)}{\frac{\gamma_0^2}{\beta_0^2}\left(1-e^{-\beta_0 h}\right)^2\left[e^{-\beta_0 h}Z_{t_{j-1}}+\frac{\alpha_0}{2\beta_0}\left(1-e^{-\beta_0 h}\right)\right]^2}& \\
\le \ \frac{1}{n^2}\sum_{j=1}^{n}\frac{\mathbb{E}\left(\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^4\,|\,Z_{t_{j-1}}\right)}{\frac{\gamma_0^2}{\beta_0^2}\left(1-e^{-\beta_0 h}\right)^2 e^{-2\beta_0 h}Z_{t_{j-1}}^2}.&
\end{aligned}
$$

Since $\mathbb{E}\left(\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^4\,|\,Z_{t_{j-1}}\right)$ is a polynomial[5] in $Z_{t_{j-1}}$, we conclude

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{j=1}^{n}\mathbb{E}\left(\frac{\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^4}{\sigma_{j-1}^4(\theta_0)}\,\Big|\,Z_{t_{j-1}}\right)=0$$

in probability according to Lemma 1.8 (iv). Therefore, using Lemma 3.4 results in

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\frac{\left(Z_{t_j}-\mu_{j-1}(\rho_0)\right)^2}{\sigma_{j-1}^2(\theta_0)}=1$$

in probability. By defining

$$\mathbb{Y}_{1,0}(\gamma):=-\frac{1}{2}\left(\log\frac{\gamma}{\gamma_0}+\frac{\gamma_0}{\gamma}-1\right)$$

we achieve

$$\lim_{n\to\infty}\left|\mathbb{Y}_{1,n}(\gamma)-\mathbb{Y}_{1,0}(\gamma)\right|=0$$

in probability while the function $\mathbb{Y}_{1,0}$ is maximized locally at γ_0 due to

$$
\begin{aligned}
\frac{\partial}{\partial\gamma}\ \mathbb{Y}_{1,0}(\gamma) &= -\frac{1}{2}\Big(\frac{1}{\gamma}-\frac{\gamma_0}{\gamma^2}\Big)\overset{!}{=}0 \qquad \Leftrightarrow \quad \gamma=\gamma_0,\\
\frac{\partial^2}{\partial\gamma^2}\mathbb{Y}_{1,0}(\gamma_0) &= -\frac{1}{2}\Big(-\frac{1}{\gamma^2}+2\frac{\gamma_0}{\gamma^3}\Big)\Big|_{\gamma=\gamma_0}=-\frac{1}{2\gamma_0^2}<0,
\end{aligned}
$$

[5]This can be derived by solving a feasible differential equation.

whereas the global maximum is justified by

$$\lim_{\gamma\to 0} \mathbb{Y}_{1,0}(\gamma) = -\infty,$$
$$\lim_{\gamma\to \infty} \mathbb{Y}_{1,0}(\gamma) = -\infty.$$

Now, we focus on the uniform convergence of $\mathbb{Y}_{1,n}$ and like in (3.16) we verify the tightness of $\mathbb{Y}_{1,n}$. Therefore, we calculate the derivative

$$\begin{aligned}
\left|\frac{\partial}{\partial\gamma}\mathbb{Y}_{1,n}(\gamma)\right| &= \left| -\frac{1}{2\gamma} - \frac{\gamma_0}{2\gamma^2 n}\sum_{j=1}^{n} \frac{\big(Z_{t_j} - \mu_{j-1}(\rho_0)\big)^2}{\sigma^2_{j-1}(\theta_0)}\right| \\
&= \frac{1}{2\gamma} + \frac{1}{2\gamma^2 n}\sum_{j=1}^{n} \frac{\big(Z_{t_j} - e^{-\beta_0 h}Z_{t_{j-1}} + \frac{\alpha_0}{\beta_0}(1-e^{-\beta_0 h})\big)^2}{\frac{1-e^{-\beta_0 h}}{\beta_0}\big[e^{-\beta_0 h}Z_{t_{j-1}} + \frac{\alpha_0}{2\beta_0}(1-e^{-\beta_0 h})\big]} \\
&\le \frac{1}{2\gamma} + \frac{1}{\gamma^2 n}\sum_{j=1}^{n} \frac{Z_{t_j}^2 + \big(e^{-\beta_0 h}Z_{t_{j-1}} - \frac{\alpha_0}{\beta_0}(1-e^{-\beta_0 h})\big)^2}{\frac{\alpha_0}{2\beta_0^2}(1-e^{-\beta_0 h})^2} \\
&\lesssim \frac{1}{2\gamma} + \frac{1}{\gamma^2}\cdot\frac{1}{n}\sum_{j=1}^{n}\big(Z_{t_j}^2 + Z_{t_{j-1}}^2 + Z_{t_{j-1}} + 1\big).
\end{aligned}$$

The last term has a common bound independent of γ and $n \in \mathbb{N}$ due to the parameter assumptions (3.3) and the law of large numbers, Lemma 1.8 (iii). Thus, we obtain the analogous formula from (3.16) using the mean value theorem.[6]

Case $\mathbb{Y}_{2,n}$: For the consistency of $\widehat{\rho}_n = (\widehat{\alpha}_n, \widehat{\beta}_n)$, we introduce

$$\begin{aligned}
\mathbb{Y}_{2,n}(\rho) &:= \frac{1}{T_n}\big(\mathbb{H}_n(\rho,\widehat{\gamma}_n) - \mathbb{H}_n(\rho_0,\widehat{\gamma}_n)\big), \\
\mathbb{Y}_{2,0}(\rho) &:= -\frac{1}{2\gamma_0}(\rho-\rho_0)^\top \int_{\mathbb{R}} \begin{pmatrix} z^{-1} & -1 \\ -1 & z \end{pmatrix} \pi_{\theta_0}(z)\,\mathrm{d}z(\rho-\rho_0) \\
&\overset{1.6}{=} -\frac{1}{2\gamma_0}(\rho-\rho_0)^\top \begin{pmatrix} \frac{2\beta_0}{2\alpha_0-\gamma_0} & -1 \\ -1 & \frac{\alpha_0}{\beta_0} \end{pmatrix}(\rho-\rho_0).
\end{aligned}$$

First, we validate argmax $\mathbb{Y}_{2,0} = \rho_0$. We compute its gradient

$$\nabla\mathbb{Y}_{2,0}(\rho) = -\frac{1}{\gamma_0}\begin{pmatrix} \frac{2\beta_0}{2\alpha_0-\gamma_0} & -1 \\ -1 & \frac{\alpha_0}{\beta_0} \end{pmatrix}(\rho-\rho_0)$$

[6] In the formula of (3.16) we need to replace ρ by γ and the exponent $2+\delta$ by $1+\delta$ due to a one dimensional parameter γ.

and its Hessian matrix

$$\mathrm{H}_{\mathbb{Y}_{2,0}}(\rho) = -\frac{1}{\gamma_0}\begin{pmatrix} \frac{2\beta_0}{2\alpha_0-\gamma_0} & -1 \\ -1 & \frac{\alpha_0}{\beta_0} \end{pmatrix},$$

which is negative definite at ρ_0:

$$\begin{aligned} \left(\mathrm{H}_{\mathbb{Y}_{2,0}}(\rho_0)\right)_{11} &= -\frac{1}{\gamma_0}\cdot\frac{2\beta_0}{2\alpha_0-\gamma_0} < 0, \\ \det\left(\mathrm{H}_{\mathbb{Y}_{2,0}}(\rho_0)\right) &= \frac{1}{\gamma_0^2}\Big(\frac{2\alpha_0}{2\alpha_0-\gamma_0}-1\Big) > \frac{1}{\gamma_0^2}\Big(\frac{2\alpha_0}{2\alpha_0}-1\Big) = 0. \end{aligned}$$

Hence, we get $\operatorname{argmax}\, \mathbb{Y}_{2,0} = \rho_0$ since additionally $\mathbb{Y}_{2,n}(\rho) \le 0$ and $\mathbb{Y}_{2,n}(\rho_0) = 0$ hold.

For a shorter notation we write $\mu_{j-1} := \mu_{j-1}(\rho_0)$ in the following to decompose:

$$\begin{aligned} \mathbb{Y}_{2,n}(\rho) &\overset{(3.6)}{=} -\frac{1}{2T_n}\sum_{j=1}^{n}\left[\log\left(\frac{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}{\sigma_{j-1}^2(\rho_0,\widehat{\gamma}_n)}\right) + \frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)} - \frac{\big(Z_{t_j}-\mu_{j-1}\big)^2}{\sigma_{j-1}^2(\rho_0,\widehat{\gamma}_n)}\right] \\ &=: \mathbb{Y}_{2,n}^{(1)}(\rho) - \frac{1}{2T_n}\sum_{j=1}^{n}\Bigg(\big[\sigma_{j-1}^{-2}(\rho,\widehat{\gamma}_n) - \sigma_{j-1}^{-2}(\rho_0,\widehat{\gamma}_n)\big]\big(Z_{t_j}-\mu_{j-1}\big)^2 \\ &\qquad\qquad + \frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)^2 - \big(Z_{t_j}-\mu_{j-1}\big)^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\Bigg) \\ &=: \mathbb{Y}_{2,n}^{(1)}(\rho) + \mathbb{Y}_{2,n}^{(2)}(\rho) - \frac{1}{2T_n}\sum_{j=1}^{n}\frac{\cancel{Z_{t_j}^2} - 2\mu_{j-1}(\rho)Z_{t_j} + \mu_{j-1}^2(\rho) - \cancel{Z_{t_j}^2} + 2\mu_{j-1}Z_{t_j} - \mu_{j-1}^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)} \\ &= \mathbb{Y}_{2,n}^{(1)}(\rho) + \mathbb{Y}_{2,n}^{(2)}(\rho) + \frac{1}{T_n}\sum_{j=1}^{n}\frac{\mu_{j-1}(\rho)-\mu_{j-1}}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}Z_{t_j} - \frac{1}{2T_n}\sum_{j=1}^{n}\frac{\mu_{j-1}^2(\rho)-\mu_{j-1}^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)} \\ &= \mathbb{Y}_{2,n}^{(1)}(\rho) + \mathbb{Y}_{2,n}^{(2)}(\rho) + \frac{1}{T_n}\sum_{j=1}^{n}\frac{\mu_{j-1}(\rho)-\mu_{j-1}}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\big(Z_{t_j}-\mu_{j-1}\big) \\ &\qquad\qquad - \frac{1}{2T_n}\sum_{j=1}^{n}\frac{\mu_{j-1}^2(\rho) - 2\mu_{j-1}(\rho)\mu_{j-1} + \mu_{j-1}^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)} \\ &= \mathbb{Y}_{2,n}^{(1)}(\rho) + \mathbb{Y}_{2,n}^{(2)}(\rho) + \mathbb{Y}_{2,n}^{(3)}(\rho) + \mathbb{Y}_{2,n}^{(4)}(\rho) \end{aligned}$$

with

$$\mathbb{Y}_{2,n}^{(1)}(\rho) := -\frac{1}{2T_n}\sum_{j=1}^{n}\log\left(\frac{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}{\sigma_{j-1}^2(\rho_0,\widehat{\gamma}_n)}\right),$$

$$\mathbb{Y}_{2,n}^{(2)}(\rho) := -\frac{1}{2T_n}\sum_{j=1}^{n}\big(\sigma_{j-1}^{-2}(\rho,\widehat{\gamma}_n) - \sigma_{j-1}^{-2}(\rho_0,\widehat{\gamma}_n)\big)\big(Z_{t_j} - \mu_{j-1}\big)^2,$$

$$\mathbb{Y}_{2,n}^{(3)}(\rho) := \frac{1}{T_n}\sum_{j=1}^{n}\frac{\mu_{j-1}(\rho) - \mu_{j-1}}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\big(Z_{t_j} - \mu_{j-1}\big),$$

$$\mathbb{Y}_{2,n}^{(4)}(\rho) := -\frac{1}{2T_n}\sum_{j=1}^{n}\frac{\big(\mu_{j-1}(\rho) - \mu_{j-1}\big)^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}.$$

Now, we examine these terms separately. Using Taylor expansion in h around zero leads to useful formulas:

$$\begin{aligned}\gamma^{-1}\sigma_{j-1}^2(\theta) &= \frac{\left(1-e^{-\beta h}\right)}{\beta}\left[e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{2\beta}\left(1-e^{-\beta h}\right)\right]\\ &= hZ_{t_{j-1}} + \frac{h^2}{2}\big(\alpha - 3\beta Z_{t_{j-1}}\big) + h^3R_1(\rho, Z_{t_{j-1}}),\end{aligned} \tag{3.18}$$

$$\begin{aligned}\left(\frac{\gamma^{-1}\sigma_{j-1}^2(\theta)}{h}\right)^{-1} &= \frac{1}{\frac{\left(1-e^{-\beta h}\right)}{\beta h}\left[e^{-\beta h}Z_{t_{j-1}} + \frac{\alpha}{2\beta}\left(1-e^{-\beta h}\right)\right]}\\ &= Z_{t_{j-1}}^{-1} + h\bigg(\frac{3\beta}{2}Z_{t_{j-1}}^{-1} - \frac{\alpha}{2}Z_{t_{j-1}}^{-2}\bigg) + h^2R_2(\rho, Z_{t_{j-1}}).\end{aligned} \tag{3.19}$$

In particular, the remainder R_1 is a polynomial in $Z_{t_{j-1}}$ of degree one and the remainder R_2 a polynomial in $Z_{t_{j-1}}^{-1}$, which are important for the application of Lemma 1.8 (iii). For $\mathbb{Y}_{2,n}^{(1)}$, we combine taylor expansion of the logarithm and Lemma 1.8 (iii) to achieve

$$\begin{aligned}-2\mathbb{Y}_{2,n}^{(1)}(\rho) &= \frac{1}{T_n}\sum_{j=1}^{n}\log\left(\frac{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}{\sigma_{j-1}^2(\rho_0,\widehat{\gamma}_n)}\right)\\ &\overset{(3.18)}{=} \frac{1}{T_n}\sum_{j=1}^{n}\log\left(\frac{hZ_{t_{j-1}} + \frac{h^2}{2}\big(\alpha - 3\beta Z_{t_{j-1}}\big) + h^3R_1(\rho, Z_{t_{j-1}})}{hZ_{t_{j-1}} + \frac{h^2}{2}\big(\alpha_0 - 3\beta_0 Z_{t_{j-1}}\big) + h^3R_1(\rho_0, Z_{t_{j-1}})}\right)\\ &= \frac{1}{T_n}\Bigg[\sum_{j=1}^{n}\log\bigg(1 + h\Big(\frac{\alpha}{2}Z_{t_{j-1}}^{-1} - \frac{3\beta}{2}\Big) + h^2Z_{t_{j-1}}^{-1}R_1(\rho, Z_{t_{j-1}})\bigg)\\ &\qquad - \sum_{j=1}^{n}\log\bigg(1 + h\Big(\frac{\alpha_0}{2}Z_{t_{j-1}}^{-1} - \frac{3\beta_0}{2}\Big) + h^2Z_{t_{j-1}}^{-1}R_1(\rho_0, Z_{t_{j-1}})\bigg)\Bigg]\\ &= \frac{1}{T_n}\Bigg[\sum_{j=1}^{n}\bigg(h\Big(\frac{\alpha}{2}Z_{t_{j-1}}^{-1} - \frac{3\beta}{2}\Big) + h^2\widetilde{R}_1(\rho, Z_{t_{j-1}})\bigg)\end{aligned}$$

$$
\begin{aligned}
&- \sum_{j=1}^{n} \Big(h\Big(\frac{\alpha_0}{2} Z_{t_{j-1}}^{-1} - \frac{3\beta_0}{2}\Big) + h^2 \widetilde{R}_1(\rho_0, Z_{t_{j-1}})\Big)\Big] \\
&= \frac{1}{n} \sum_{j=1}^{n} \Big(\Big(\frac{\alpha}{2} Z_{t_{j-1}}^{-1} - \frac{3\beta}{2}\Big) - \Big(\frac{\alpha_0}{2} Z_{t_{j-1}}^{-1} - \frac{3\beta_0}{2}\Big)\Big) + O_p^*(h) \\
&\overset{n\to\infty}{\to} \int_0^\infty \Big(\frac{\alpha}{2} z^{-1} - \frac{3\beta}{2}\Big) - \Big(\frac{\alpha_0}{2} z^{-1} - \frac{3\beta_0}{2}\Big) \pi_{\theta_0}(z)\,\mathrm{d}z
\end{aligned}
$$

in probability. For $\mathbb{Y}_{2,n}^{(2)}$ we derive

$$
\begin{aligned}
&\frac{1}{T_n} \sum_{j=1}^{n} \mathbb{E}\Big(\big(\widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho, \widehat{\gamma}_n) - \widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho_0, \widehat{\gamma}_n)\big) \big(Z_{t_j} - \mu_{j-1}\big)^2 \,\big|\, Z_{t_{j-1}} \Big) \\
&= \frac{1}{T_n} \sum_{j=1}^{n} \big(\widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho, \widehat{\gamma}_n) - \widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho_0, \widehat{\gamma}_n)\big)\, \mathbb{E}\Big(\big(Z_{t_j} - \mu_{j-1}\big)^2 \,\big|\, Z_{t_{j-1}} \Big) \\
&= \frac{1}{T_n} \sum_{j=1}^{n} \big(\widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho, \widehat{\gamma}_n) - \widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho_0, \widehat{\gamma}_n)\big) \sigma_{j-1}^{2}(\sigma_0) \\
&\overset{(3.18)}{=} \frac{\gamma_0}{T_n} \sum_{j=1}^{n} \big(h\widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho, \widehat{\gamma}_n) - h\widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho_0, \widehat{\gamma}_n)\big) \Big(Z_{t_{j-1}} + h\widetilde{R}_1(\rho_0, Z_{t_{j-1}})\Big) \\
&\overset{(3.19)}{=} \frac{\gamma_0}{T_n} \sum_{j=1}^{n} \Big[h\Big(\frac{3\beta}{2} Z_{t_{j-1}}^{-1} - \frac{\alpha}{2} Z_{t_{j-1}}^{-2}\Big) - h\Big(\frac{3\beta_0}{2} Z_{t_{j-1}}^{-1} - \frac{\alpha_0}{2} Z_{t_{j-1}}^{-2}\Big)\Big] Z_{t_{j-1}} + O_p^*(h) \\
&= \frac{\gamma_0}{n} \sum_{j=1}^{n} \Big[\Big(\frac{3\beta}{2} - \frac{\alpha}{2} Z_{t_{j-1}}^{-1}\Big) - \Big(\frac{3\beta_0}{2} - \frac{\alpha_0}{2} Z_{t_{j-1}}^{-1}\Big)\Big] + O_p^*(h) \\
&\overset{n\to\infty}{\to} \int_0^\infty \gamma_0\Big(\frac{3\beta}{2} - \frac{\alpha}{2} z^{-1}\Big) - \gamma_0\Big(\frac{3\beta_0}{2} - \frac{\alpha_0}{2} z^{-1}\Big) \pi_{\theta_0}(z)\,\mathrm{d}z
\end{aligned}
$$

in probability, where the last line is valid by virtue of Lemma 1.8 (iii) and by a similar argument we conclude

$$
\lim_{n\to\infty} \frac{1}{T_n^2} \sum_{j=1}^{n} \mathbb{E}\Big(\big(\widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho, \widehat{\gamma}_n) - \widehat{\gamma}_n^{-1} \sigma_{j-1}^{-2}(\rho_0, \widehat{\gamma}_n)\big)^2 \big(Z_{t_j} - \mu_{j-1}\big)^4 \,\Big|\, Z_{t_{j-1}} \Big) = 0
$$

in probability. Combining $\lim_{n\to\infty} \widehat{\gamma}_n = \gamma_0$ with Lemma 3.4 we conclude

$$
\lim_{n\to\infty} \Big(\mathbb{Y}_{2,n}^{(1)}(\rho) + \mathbb{Y}_{2,n}^{(2)}(\rho) \Big) = 0
$$

in probability. By using again Lemma 3.4 we can easily derive $\lim_{n\to\infty} \mathbb{Y}_{2,n}^{(3)}(\rho) = 0$ in probability. For this purpose, we specify

$$\frac{1}{T_n}\sum_{j=1}^{n} \mathbb{E}\left(\frac{\mu_{j-1}(\rho)-\mu_{j-1}}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\left(Z_{t_j}-\mu_{j-1}\right)\,\middle|\, Z_{t_{j-1}}\right) = 0$$

and

$$\begin{aligned}
&\frac{1}{T_n^2}\sum_{j=1}^{n} \mathbb{E}\left(\frac{\left(\mu_{j-1}(\rho)-\mu_{j-1}\right)^2}{\sigma_{j-1}^4(\rho,\widehat{\gamma}_n)}\left(Z_{t_j}-\mu_{j-1}\right)^2\,\middle|\, Z_{t_{j-1}}\right)\\
= \;& \frac{1}{T_n^2}\sum_{j=1}^{n} \frac{\left(\mu_{j-1}(\rho)-\mu_{j-1}\right)^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\cdot\frac{\sigma_{j-1}^2(\theta_0)}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\\
\overset{(3.5)}{\underset{(3.19)}{=}}\;& \frac{1}{T_n^2}\sum_{j=1}^{n} \left(\mu_{j-1}(\rho)-\mu_{j-1}\right)^2\frac{1}{h\widehat{\gamma}_n}Z_{t_{j-1}}\cdot\frac{\gamma_0}{\widehat{\gamma}_n}+O_p^*(h)\\
= \;& \frac{1}{T_n^2}\sum_{j=1}^{n} \Big(\left(\alpha-\beta Z_{t_{j-1}}\right)h-\left(\alpha_0-\beta_0 Z_{t_{j-1}}\right)h\Big)^2\frac{1}{h\widehat{\gamma}_n}Z_{t_{j-1}}^{-1}\frac{\gamma_0}{\widehat{\gamma}_n}+O_p^*(h)\\
= \;& \frac{\gamma_0}{\widehat{\gamma}_n^2}\frac{1}{nT_n}\sum_{j=1}^{n} \Big(\left(\alpha-\beta Z_{t_{j-1}}\right)-\left(\alpha_0-\beta_0 Z_{t_{j-1}}\right)\Big)^2 Z_{t_{j-1}}^{-1}+O_p^*(h) \overset{n\to\infty}{\to} 0
\end{aligned}$$

in probability. The last line is justified by Lemma 1.8 (iv) with $\varepsilon_n := T_n^{-1}$, whereas in the second-last line we applied Taylor expansion on μ_{j-1} in h around zero. Similarly, for $\mathbb{Y}_{2,n}^{(4)}$ we calculate

$$\begin{aligned}
\mathbb{Y}_{2,n}^{(4)}(\rho) \;&=\; -\frac{1}{2T_n}\sum_{j=1}^{n}\frac{\left(\mu_{j-1}(\rho)-\mu_{j-1}\right)^2}{\sigma_{j-1}^2(\rho,\widehat{\gamma}_n)}\\
&=\; -\frac{1}{2n\widehat{\gamma}_n}\sum_{j=1}^{n}\Big(\left(\alpha-\beta Z_{t_{j-1}}\right)-\left(\alpha_0-\beta_0 Z_{t_{j-1}}\right)\Big)^2 Z_{t_{j-1}}^{-1}+O_p^*(h)\\
&\overset{n\to\infty}{\to}\; -\frac{1}{2\gamma_0}\int_0^\infty \Big(\left(\alpha-\alpha_0\right)-\left(\beta-\beta_0\right)z\Big)^2 z^{-1}\pi_{\theta_0}(z)\,\mathrm{d}x\\
&=\; -\frac{1}{2\gamma_0}(\rho-\rho_0)^\top\int_{\mathbb{R}}\begin{pmatrix} z^{-1} & -1\\ -1 & z\end{pmatrix}\pi_{\theta_0}(z)\,\mathrm{d}z(\rho-\rho_0)\\
&=:\; \mathbb{Y}_{2,0}(\rho)
\end{aligned}$$

in probability. The uniform convergence is verified through the tightness argument as for

$\mathbb{Y}_{1,n}$ and in (3.16). Finally, we have proved $\mathbb{Y}_{2,n}(\rho) = \mathbb{Y}_{2,0}(\rho) + o_p^*(1)$, which completes the consistency. $\square$

To now prove the asymptotic normality we additionally specify the Fisher information matrix in case of uniformly elliptic diffusions [36]

$$\mathcal{I}(\theta) := \begin{pmatrix} \frac{1}{\gamma}\int\limits_0^\infty z^{-1}\pi_\theta(z)\,\mathrm{d}z & -\frac{1}{\gamma} & 0 \\ -\frac{1}{\gamma} & \frac{1}{\gamma}\int\limits_0^\infty z\pi_\theta(z)\,\mathrm{d}x & 0 \\ 0 & 0 & \frac{1}{2\gamma^2} \end{pmatrix} \overset{1.6}{=} \begin{pmatrix} \frac{1}{\gamma}\cdot\frac{2\beta}{2\alpha-\gamma} & -\frac{1}{\gamma} & 0 \\ -\frac{1}{\gamma} & \frac{1}{\gamma}\cdot\frac{\alpha}{\beta} & 0 \\ 0 & 0 & \frac{1}{2\gamma^2} \end{pmatrix}. \quad (3.20)$$

We notice that $\mathcal{I}(\theta)$ is invertible for $\theta \in \overline{\Theta}$, in formulas,

$$\mathcal{I}(\theta)^{-1} = \begin{pmatrix} \frac{\alpha(2\alpha-\gamma)}{\beta} & 2\alpha-\gamma & 0 \\ 2\alpha-\gamma & 2\beta & 0 \\ 0 & 0 & 2\gamma^2 \end{pmatrix}.$$

Lemma 3.6: For every true value $\theta_0 \in \Theta$ we have

$$\lim_{n\to\infty} \mathbb{D}_n(\widehat{\theta}_n - \theta_0) = \mathcal{N}\left(0, \mathcal{I}(\theta_0)^{-1}\right)$$

in distribution and

$$\lim_{n\to\infty} \mathcal{I}(\widehat{\theta}_n)^{\frac{1}{2}}\mathbb{D}_n(\widehat{\theta}_n - \theta_0) = \mathcal{N}\big(0, \mathbb{I}_3\big)$$

in distribution, where $\mathbb{I}_3$ denotes the 3-dimensional identity matrix.

Proof: The proof proceeds similar to the proof for the classical uniformly-elliptic diffusion model as in [58] though we have to take care of tightness and integrability issues caused by the diffusion-coefficient form.

We look at the third-order Taylor expansion of $\nabla\mathbb{H}_n(\widehat{\theta}_n)$ around $\theta_0 \in \Theta$. That is, we focus on the event $\big\{\nabla\mathbb{H}_n(\widehat{\theta}_n) = \mathbf{0}\big\}$, on which

$$\mathbf{0} = \mathbb{D}_n^{-1}\nabla\mathbb{H}_n(\theta_0) + \left(-\int\limits_0^1 \mathbb{D}_n^{-1}\mathrm{H}_{\mathbb{H}_n}\left(\theta_0 + s(\widehat{\theta}_n - \theta_0)\right)\mathbb{D}_n^{-1}ds\right)\big[\mathbb{D}_n\big(\widehat{\theta}_n - \theta_0\big)\big]$$

holds. Hence, having proved the consistency it suffices to verify the following statements:

(AN1) For $\Delta_n(\theta_0) := \mathbb{D}_n^{-1}\nabla\mathbb{H}_n(\theta_0)$ we have

$$\lim_{n\to\infty} \Delta_n(\theta_0) = \mathcal{N}_3\left(0, \mathcal{I}(\theta_0)\right)$$

in distribution.

(AN2) For $\mathcal{I}_n(\theta_0) := -\mathbb{D}_n^{-1}\big(\mathrm{H}_{\mathbb{H}_n}(\theta_0)\big)\mathbb{D}_n^{-1}$ we have

$$\lim_{n\to\infty} \mathcal{I}_n(\theta_0) = \mathcal{I}(\theta_0)$$

in probability.

(AN3) For any (non-random) positive sequence $\lim_{n\to\infty} \delta_n = 0$ we have

$$\lim_{n\to\infty} \sup_{\theta:\, |\theta-\theta_0|\le\delta_n} |\mathcal{I}_n(\theta) - \mathcal{I}(\theta_0)| = 0$$

in probability.

Proof of (AN1): The main tool here is a Lemma, [51, Lemmas 3.5 and 3.6], which we simplify in the following to fit our case:
For some d-dimensional Markov process $\left(\zeta_j^n\right)_{j=1,\dots,n}$ satisfying for $n \in \mathbb{N}$

$$\begin{aligned}
\mathbb{E}\left(\zeta_j^n \,|\, \zeta_{j-1}^n\right) &= 0,\\
\lim_{n\to\infty} \sum_{j=1}^n \mathbb{E}\left((\zeta_j^n)_k(\zeta_j^n)_l \,|\, \zeta_{j-1}^n\right) &= \Sigma_{kl},\\
\lim_{n\to\infty} \sum_{j=1}^n \mathbb{E}\left(\|\zeta_j^n\|^4 \,|\, \zeta_{j-1}^n\right) &= 0
\end{aligned}$$

in probability for every $k, l = 1, \dots, d$, where $\Sigma := \left(\Sigma_{kl}\right)_{1\le k,l\le d}$ is deterministic, $\sum_{j=1}^n \zeta_j^n$ converges in distribution to $\mathcal{N}(0, \Sigma)$.

Certainly, the aim is the application to Δ_n and hence we first determine the derivatives of $\mathbb{H}_n$. We recall (3.6) and compute

$$\frac{\partial}{\partial\gamma}\mathbb{H}_n(\theta) = -\frac{1}{2}\sum_{j=1}^n \left(\frac{1}{\gamma} - \frac{\left(Z_{t_j} - \mu_{j-1}(\rho)\right)^2}{\gamma\sigma_{j-1}^2(\theta)}\right) =: \sum_{j=1}^n \zeta_{\gamma,j}(\theta)$$

with

$$-\frac{1}{2}\,\mathbb{E}\left(\zeta_{\gamma,j}(\theta_0)\,|\,Z_{t_{j-1}}\right) = \mathbb{E}\left(\frac{1}{\gamma_0} - \frac{\left(Z_{t_j} - \mu_{j-1}(\rho_0)\right)^2}{\gamma_0 \sigma^2_{j-1}(\theta_0)}\,\Big|\,Z_{t_{j-1}}\right) = 0.$$

Next, we calculate

$$\begin{aligned}\frac{\partial}{\partial\alpha}\mathbb{H}_n(\theta) = -\frac{1}{2}\sum_{j=1}^{n}\Bigg[\frac{\frac{\partial}{\partial\alpha}\sigma^2_{j-1}(\theta)}{\sigma^2_{j-1}(\theta)} &- \frac{\left(Z_{t_j} - \mu_{j-1}(\rho)\right)^2\frac{\partial}{\partial\alpha}\sigma^2_{j-1}(\theta)}{\sigma^4_{j-1}(\theta)} \\ &- \frac{2\left(Z_{t_j} - \mu_{j-1}(\rho)\right)\frac{\partial}{\partial\alpha}\mu_{j-1}(\rho)}{\sigma^2_{j-1}(\theta)}\Bigg] \\ =: \sum_{j=1}^{n}\zeta_{\alpha,j}(\theta),\end{aligned}$$

for which obviously

$$\mathbb{E}\left(\zeta_{\alpha,j}(\theta_0)\,|\,Z_{t_{j-1}}\right) = 0$$

is valid and respectively

$$\begin{aligned}\frac{\partial}{\partial\beta}\mathbb{H}_n(\theta) = -\frac{1}{2}\sum_{j=1}^{n}\Bigg[\frac{\frac{\partial}{\partial\beta}\sigma^2_{j-1}(\theta)}{\sigma^2_{j-1}(\theta)} &- \frac{\left(Z_{t_j} - \mu_{j-1}(\rho)\right)^2\frac{\partial}{\partial\beta}\sigma^2_{j-1}(\theta)}{\sigma^4_{j-1}(\theta)} \\ &- \frac{2\left(Z_{t_j} - \mu_{j-1}(\rho)\right)\frac{\partial}{\partial\beta}\mu_{j-1}(\rho)}{\sigma^2_{j-1}(\theta)}\Bigg] \\ =: \sum_{j=1}^{n}\zeta_{\beta,j}(\theta)\end{aligned}$$

with

$$\mathbb{E}\left(\zeta_{\beta,j}(\theta_0)\,|\,Z_{t_{j-1}}\right) = 0.$$

With these considerations, the first property immediately follows

$$\mathbb{E}\left(\mathbb{D}_n^{-1}\zeta_j(\theta_0)\,|\,Z_{t_{j-1}}\right) = 0$$

by defining

$$\zeta_j(\theta) := \begin{pmatrix} \zeta_{\alpha,j}(\theta) \\ \zeta_{\beta,j}(\theta) \\ \zeta_{\gamma,j}(\theta) \end{pmatrix}.$$

In particular, we note that the condition in the expectation naturally carries over to $Z_{t_{j-1}}$. Furthermore, by the asymptotics in [58, Lemma 7] we conclude

$$\begin{aligned} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\left(\zeta_{\gamma,j}^2(\theta_0) \mid Z_{t_{j-1}}\right) &= \frac{1}{4n}\sum_{j=1}^{n}\left[\frac{1}{\gamma_0^2} + \mathbb{E}\left(\frac{\left(Z_{t_j} - \mu_{j-1}(\rho_0)\right)^4}{\gamma_0^2\sigma_{j-1}^4(\theta_0)} \,\Big|\, Z_{t_{j-1}}\right)\right] \\ &= \frac{1}{4n}\sum_{j=1}^{n}\left[\frac{1}{\gamma_0^2} + \frac{1}{\gamma_0^2}\left(1 + O_p(h)\right)\right] \\ &\overset{n\to\infty}{\to} \frac{1}{4\gamma_0^2} + \frac{1}{4\gamma_0^2} = \frac{1}{2\gamma_0^2} \end{aligned}$$

in probability. Analogously, as in the proof of the consistency, we can show by a feasible calculation

$$\begin{aligned} \lim_{n\to\infty}\sum_{j=1}^{n} \mathbb{E}\left(\frac{1}{nh}\zeta_{\alpha,j}^2(\theta_0) \,\Big|\, Z_{t_{j-1}}\right) &= \frac{1}{\gamma_0}\int_0^\infty z^{-1}\pi_{\theta_0}(z)\,\mathrm{d}z, \\ \lim_{n\to\infty}\sum_{j=1}^{n} \mathbb{E}\left(\frac{1}{nh}\zeta_{\beta,j}^2(\theta_0) \,\Big|\, Z_{t_{j-1}}\right) &= \frac{1}{\gamma_0}\int_0^\infty z\pi_{\theta_0}(z)\,\mathrm{d}z, \\ \lim_{n\to\infty}\sum_{j=1}^{n} \mathbb{E}\left(\frac{1}{nh}\zeta_{\alpha,j}(\theta_0)\zeta_{\beta,j}(\theta_0) \,\Big|\, Z_{t_{j-1}}\right) &= -\frac{1}{\gamma_0} \\ \lim_{n\to\infty}\sum_{j=1}^{n} \mathbb{E}\left(\frac{1}{n\sqrt{h}}\zeta_{\alpha,j}(\theta_0)\zeta_{\gamma,j}(\theta_0) \,\Big|\, Z_{t_{j-1}}\right) &= 0, \\ \lim_{n\to\infty}\sum_{j=1}^{n} \mathbb{E}\left(\frac{1}{n\sqrt{h}}\zeta_{\beta,j}(\theta_0)\zeta_{\gamma,j}(\theta_0) \,\Big|\, Z_{t_{j-1}}\right) &= 0, \\ \lim_{n\to\infty}\sum_{j=1}^{n} \mathbb{E}\left(\left|\mathbb{D}_n^{-1}\zeta_j(\theta_0)\right|^4 \mid Z_{t_{j-1}}\right) &= 0 \end{aligned}$$

in probability.

Proof of (AN2): In this proof we always use Lemma 3.4 without mentioning it. First, we compute the second derivatives. We start with the simplest derivatives where at least one is with respect to γ:

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha\partial\gamma}\mathbb{H}_n(\theta) &= -\frac{1}{2}\sum_{j=1}^{n}\left(\frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)\frac{\partial}{\partial\alpha}\mu_{j-1}(\rho)}{\gamma\sigma_{j-1}^2(\theta)} + \frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)\frac{\partial}{\partial\alpha}\sigma_{j-1}^2(\theta)}{\gamma\sigma_{j-1}^4(\theta)}\right) \\
&=: \sum_{j=1}^{n}\zeta_{\alpha\gamma,j}(\theta), \\
\frac{\partial^2}{\partial\beta\partial\gamma}\mathbb{H}_n(\theta) &= -\frac{1}{2}\sum_{j=1}^{n}\left(\frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)\frac{\partial}{\partial\beta}\mu_{j-1}(\rho)}{\gamma\sigma_{j-1}^2(\theta)} + \frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)\frac{\partial}{\partial\beta}\sigma_{j-1}^2(\theta)}{\gamma\sigma_{j-1}^4(\theta)}\right) \\
&=: \sum_{j=1}^{n}\zeta_{\beta\gamma,j}(\theta), \\
\frac{\partial^2}{\partial\gamma^2}\mathbb{H}_n(\theta) &= -\frac{1}{2}\sum_{j=1}^{n}\left(-\frac{1}{\gamma^2} + 2\frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)^2}{\gamma^2\sigma_{j-1}^2(\theta)}\right).
\end{aligned}$$

For the convergence we rewrite the last term to

$$\begin{aligned}
\frac{1}{n}\frac{\partial^2}{\partial\gamma^2}H_n(\theta) &= -\frac{1}{2n}\sum_{j=1}^{n}\left(-\frac{1}{\gamma^2} + 2\frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)^2}{\gamma^2\sigma_{j-1}^2(\theta)}\right) \\
&= \frac{1}{2\gamma^2} - \frac{1}{\gamma^2}\frac{1}{n}\sum_{j=1}^{n}\frac{\big(Z_{t_j}-\mu_{j-1}(\rho)\big)^2}{\sigma_{j-1}^2(\theta)}.
\end{aligned}$$

Recalling

$$\begin{aligned}
\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left(\frac{\big(Z_{t_j}-\mu_{j-1}(\rho_0)\big)^2}{\sigma_{j-1}^2(\theta_0)}\,\middle|\,Z_{t_{j-1}}\right) &= 1, \\
\lim_{n\to\infty}\frac{1}{n^2}\sum_{j=1}^{n}\mathbb{E}\left(\frac{\big(Z_{t_j}-\mu_{j-1}(\rho_0)\big)^4}{\sigma_{j-1}^4(\theta_0)}\,\middle|\,Z_{t_{j-1}}\right) &= 0,
\end{aligned}$$

we receive

$$\lim_{n\to\infty}\frac{1}{n}\frac{\partial^2}{\partial\gamma^2}\mathbb{H}_n(\theta_0) = -\frac{1}{2\gamma_0^2}$$

in probability. For the next terms, we immediately notice

$$\mathbb{E}\left(\zeta_{\alpha\gamma,j}(\theta)\,|\,Z_{t_{j-1}}\right) = 0$$

and observe by the mean value theorem

$$\begin{aligned} 0 &\le \frac{\partial}{\partial\alpha}\mu_{j-1}(\rho) = \frac{1-e^{-\beta h}}{\beta} = h\frac{1-e^{-\beta h}}{\beta h} \le h, \\ 0 &\le \frac{\partial}{\partial\alpha}\sigma^2_{j-1}(\theta) = \frac{\gamma}{\beta}\frac{\left(1-e^{-\beta h}\right)^2}{2\beta} \le \gamma h^2. \end{aligned} \tag{3.21}$$

Therefore, we conclude

$$\begin{aligned} \frac{1}{n^2h}\sum_{j=1}^{n}\mathbb{E}\left(\zeta^2_{\alpha\gamma,j}(\theta_0)\,|\,Z_{t_{j-1}}\right) &= \frac{1}{n^2h}\sum_{j=1}^{n}\left[\frac{\left(\frac{\partial}{\partial\alpha}\mu_{j-1}(\rho_0)\right)^2}{\gamma_0^2\sigma^2_{j-1}(\theta_0)} + 2\frac{\frac{\partial}{\partial\alpha}\mu_{j-1}(\rho_0)\frac{\partial}{\partial\alpha}\sigma^2_{j-1}(\theta_0)}{\gamma^2\sigma^4_{j-1}(\theta_0)}\right. \\ &\qquad\left. + \frac{\left(\frac{\partial}{\partial\alpha}\sigma_{j-1}(\theta_0)\right)^2}{\gamma_0^2\sigma^6_{j-1}(\theta_0)}\right] \\ &\overset{(3.21)}{\le} \frac{1}{n^2h}\sum_{j=1}^{n}\left[\frac{h^2}{\gamma_0^2\sigma^2_{j-1}(\theta_0)} + 2\frac{h^3}{\gamma_0\sigma^4_{j-1}(\theta_0)} + \frac{h^4}{\sigma^6_{j-1}(\theta_0)}\right] \\ &\overset{(3.19)}{\le} \frac{1}{n^2h}\sum_{j=1}^{n}\frac{h}{\gamma_0^3}\Big(Z^{-1}_{t_{j-1}} + hR(\theta_0, Z^{-1}_{t_{j-1}})\Big) \\ &= \frac{1}{n^2}\sum_{j=1}^{n}\frac{1}{\gamma_0^3}\Big(Z^{-1}_{t_{j-1}} + hR(\theta_0, Z^{-1}_{t_{j-1}})\Big) \\ &\overset{n\to\infty}{\to} 0 \end{aligned}$$

in probability. We note that the remainder R is a polynomial of $Z^{-1}_{t_{j-1}}$ and $Z^{-2}_{t_{j-1}}$ such that Lemma 1.8 (iv) is applicable. Consequently, we achieve

$$\lim_{n\to\infty}\frac{1}{n\sqrt{h}}\sum_{j=1}^{n}\zeta_{\alpha\gamma,j}(\theta_0) = 0$$

in probability and by similar arguments

$$\lim_{n\to\infty}\frac{1}{n\sqrt{h}}\sum_{j=1}^{n}\zeta_{\beta\gamma,j}(\theta_0) = 0$$

in probability as well. With analogous, feasible calculations we can show

$$\begin{aligned}\lim_{n\to\infty}\frac{\partial^2}{\partial\alpha^2}\mathbb{H}_n(\theta_0) &= -\frac{1}{\gamma_0}\int_0^\infty z^{-1}\pi_{\theta_0}(z)\,\mathrm{d}z,\\ \lim_{n\to\infty}\frac{\partial^2}{\partial\beta^2}\mathbb{H}_n(\theta_0) &= -\frac{1}{\gamma_0}\int_0^\infty z\pi_{\theta}(z)\,\mathrm{d}z,\\ \lim_{n\to\infty}\frac{\partial^2}{\partial\alpha\partial\beta}\mathbb{H}_n(\theta_0) = \frac{1}{\gamma_0}\end{aligned}$$

in probability which completes the proof of (AN2). In particular, we use the asymptotics found in [58, Lemma 7].

Proof of (AN3): If we take a closer look at the second order derivatives of $\mathbb{H}_n(\theta)$ and remind us again that all moments with order $p \geq -2$ are finite due to our parameter assumptions (3.3) and Lemma 1.8 (iii), we directly observe

$$\sup_{\theta\in\overline{\Theta}}\left|\frac{\partial}{\partial\nu}\mathcal{I}_n(\theta)\right| = O_p(1)$$

for $\nu \in \{\alpha,\beta,\gamma\}$. We note here that our parameter space Θ is bounded. Then, using the triangle inequality and mean value theorem leads to

$$\begin{aligned}\sup_{\theta:\,|\theta-\theta_0|\leq\delta_n}|\mathcal{I}_n(\theta)-\mathcal{I}(\theta_0)| &\leq |\mathcal{I}_n(\theta)-\mathcal{I}_n(\theta_0)| + |\mathcal{I}_n(\theta_0)-\mathcal{I}(\theta_0)|\\ &\lesssim \delta_n \max_{\nu\in\{\alpha,\beta,\gamma\}}\sup_{\theta\in\overline{\Theta}}\left|\frac{\partial}{\partial\alpha}\mathcal{I}_n(\theta)\right| + |\mathcal{I}_n(\theta_0)-\mathcal{I}(\theta_0)|\\ &\overset{\text{(AN2)}}{=} O_p(\delta_n)+o_p(1) = o_p(1)\end{aligned}$$

for any positive (non-random) sequence $(\delta_n)_{n\in\mathbb{N}}$ satisfying $\lim_{n\to\infty}\delta_n = 0$. □

3.2.2 One-step improvement

In Section 3.1 we first introduced a $\mathbb{D}_n$-consistent initial estimator $\widehat{\theta}_{0,n} = (\widehat{\alpha}_{0,n},\widehat{\beta}_{0,n},\widehat{\gamma}_{0,n})$ given in a closed form as in [72]. Using these results we now construct the Newton-Raphson one-step improvement

$$\widehat{\theta}_n^{(1,1)} = \widehat{\theta}_{0,n} - \left(\mathrm{H}_{\mathbb{H}_n}(\widehat{\theta}_{0,n})\right)^{-1}\nabla\mathbb{H}_n(\widehat{\theta}_{0,n}) \tag{3.22}$$

and the Fisher scoring one-step improvement

$$\widehat{\theta}_n^{(1,2)} = \widehat{\theta}_{0,n} + \mathbb{D}_n^{-1}\Big(\mathcal{I}\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathbb{D}_n^{-1}\nabla\mathbb{H}_n\big(\widehat{\theta}_{0,n}\big) \tag{3.23}$$

towards the Gaussian quasi-maximum likelihood estimator. We occasionally refer to them as one-step estimators. It is well-known that the Newton-Raphson method is used to successively obtain a better approximation of the zeros of a real-valued function. Our corresponding function here is $\nabla\mathbb{H}_n$. This method works better if we start close to a zero, c.f. [8, 2.3 Newton's method]. Therefore, it makes sense to use as the so-called initial guess an estimator of the true parameter. Replacing the Hessian matrix $-\mathrm{H}_{\mathbb{H}_n}$ by the Fisher information matrix leads to the Fisher scoring method, see [52, 62]. We now prove that the GQMLE $\widehat{\theta}_n$ and the one-step estimators $\widehat{\theta}_n^{(1,1)}$ and $\widehat{\theta}_n^{(1,2)}$ are all asymptotically equivalent at rate $\mathbb{D}_n$.

Theorem 3.7: We have

$$\mathbb{D}_n\big(\widehat{\theta}_n^{(1,i)} - \widehat{\theta}_n\big) = o_p(1)$$

for $i = 1, 2$ or equivalently

$$\lim_{n\to\infty} \mathbb{D}_n\big(\widehat{\theta}_n^{(1,i)} - \theta_0\big) = \mathcal{N}\big(0, \mathcal{I}(\theta_0)^{-1}\big)$$

in distribution for $i = 1, 2$.

Proof: First of all, to achieve asymptotic efficiency as well the asymptotic normality of the one-step estimators, we need to show the asymptotic equivalence between the one-step estimators $\widehat{\theta}_n^{(1,i)}$ and the joint GQMLE $\widehat{\theta}_n$. The arguments for both estimators are similar, so here we give the proof for $\widehat{\theta}_n^{(1,1)}$. We recall the setting and definitions from the asymptotic normality proof of $\widehat{\theta}_n$. We again focus on the event, where $\nabla\mathbb{H}_n\big(\widehat{\theta}_n\big) = 0$ holds, and use the object $\mathcal{I}_n(\theta) = -\mathbb{D}_n^{-1}\big(\mathrm{H}_{\mathbb{H}_n}(\theta)\big)\mathbb{D}_n^{-1}$. Then, we derive

$$\mathbb{D}_n\big(\widehat{\theta}_n^{(1,1)} - \widehat{\theta}_n\big) = \mathbb{D}_n\big(\widehat{\theta}_n^{(1,1)} - \widehat{\theta}_{0,n}\big) + \mathbb{D}_n\big(\widehat{\theta}_{0,n} - \widehat{\theta}_n\big)$$

and in particular

$$\begin{aligned}\mathbb{D}_n\big(\widehat{\theta}_n^{(1,1)} - \widehat{\theta}_{0,n}\big) &= -\mathbb{D}_n\Big(\mathrm{H}_{\mathbb{H}_n}\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\nabla_\theta\mathbb{H}_n\big(\widehat{\theta}_{0,n}\big)\\ &= \Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathbb{D}_n^{-1}\nabla\mathbb{H}_n\big(\widehat{\theta}_{0,n}\big)\end{aligned}$$

$$
\begin{aligned}
&= \Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\Big(\mathbb{D}_n^{-1}\nabla\mathbb{H}_n\big(\widehat{\theta}_{0,n}\big) - \mathbb{D}_n^{-1}\nabla\mathbb{H}_n\big(\widehat{\theta}_n\big)\Big)\\
&= \Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathbb{D}_n^{-1}\mathrm{H}_{\mathbb{H}_n}\big(\theta_n\big)\cdot\big(\widehat{\theta}_{0,n} - \widehat{\theta}_n\big)\\
&= -\Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathcal{I}_n\big(\theta_n\big)\mathbb{D}_n\big(\widehat{\theta}_{0,n} - \widehat{\theta}_n\big).
\end{aligned}
$$

We applied $\nabla\mathbb{H}_n\big(\widehat{\theta}_n\big) = 0$ in the third line and the mean value theorem in the fourth line, where we have $\theta_n = \widehat{\theta}_{0,n} + u(\widehat{\theta}_n - \widehat{\theta}_{0,n})$ for some $u \in [0,1]$ with $\lim_{n\to\infty}\theta_n = \theta_0$ in probability due to the consistency of $\widehat{\theta}_{0,n}$ and $\widehat{\theta}_n$. Combining these thoughts, we receive

$$
\begin{aligned}
\mathbb{D}_n\big(\widehat{\theta}_n^{(1,1)} - \widehat{\theta}_n\big) &= -\Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathcal{I}_n\big(\theta_n\big)\mathbb{D}_n\big(\widehat{\theta}_{0,n} - \widehat{\theta}_n\big) + \mathbb{D}_n\big(\widehat{\theta}_{0,n} - \widehat{\theta}_n\big)\\
&= \Big(\mathbb{I}_3 - \Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathcal{I}_n\big(\theta_n\big)\Big)\mathbb{D}_n\big(\widehat{\theta}_{0,n} - \widehat{\theta}_n\big)\\
&= \Big(\mathbb{I}_3 - \Big(\mathcal{I}_n\big(\widehat{\theta}_{0,n}\big)\Big)^{-1}\mathcal{I}_n\big(\theta_n\big)\Big)\cdot\Big(\mathbb{D}_n\big(\widehat{\theta}_{0,n} - \theta_0\big) + \mathbb{D}_n\big(\theta_0 - \widehat{\theta}_n\big)\Big)\\
&= o_p(1)\cdot\big(O_p(1) + O_p(1)\big) = o_p(1).
\end{aligned}
$$

The last line holds due to the asymptotic normality of $\widehat{\theta}_{0,n}$ and $\widehat{\theta}_n$ as well as (AN3) in the proof of Lemma 3.6. □

We can also conclude the asymptotic normality for the so-called standardized estimators, that is,

$$
\lim_{n\to\infty}\mathcal{I}(\widehat{\theta}_n)^{\frac{1}{2}}\mathbb{D}_n(\widehat{\theta}_n^{(1,i)} - \theta_0) = \mathcal{N}\big(0,\, \mathbb{I}_3\big)
$$

in distribution for $i = 1, 2$.

3.3 Numerical experiments

In this section, we compare simulation results of the preliminary estimator Lemma 3.1, (3.10) with the one step estimators based on the Newton-Raphson (3.22) and the Fisher scoring method (3.23). We use an exact Cox-Ingersoll-Ross simulator for $(X_{t_j})_{j=1,\dots,n}$ through non-central chi-squares [63]. The 1000 simulated estimators are performed for $n = 5000, 10000, 20000$ and $T = T_n = 500, 1000, 2000$. So that the condition $\frac{2\alpha}{\gamma} > 5$ is fulfilled, we choose as true value $\theta_0 = (\alpha_0, \beta_0, \gamma_0) = (3, 1, 1)$. Table 3.1 summarizes the mean and standard deviation (sd) of these estimators. As a comparison for the behaviour at smaller T and n, we contrast boxplots for $T = 500, n = 5000$ and $T = 50, n = 200$

(Figures 3.3 to 3.5). Additionally, the corresponding histograms are given in Figure 3.1 and 3.2.

Table 3.1 shows that the performance of the three estimators behaves quite similarly. Upon closer inspection, we recognize the smallest improvement in the estimates of γ_0. The two estimators $\widehat{\gamma}_{0,n}$ and $\widehat{\gamma}_{2,n}$ have even the same values on four decimal points except for two deviations of 10^{-4}. Comparing the estimators for α_0 and β_0, we detect, with one exception (Newton-Raphson method for $T = 1000$ and $n = 20000$), a small improvement in the one-step estimators compared to the preliminary estimator. Besides, the Fisher scoring method performs slightly better than the Newton-Raphson method.

Overall, the performances of the three estimators seem to be quite similar. This leads to the assumption that the preliminary estimator is already asymptotically optimal. The almost undetectable difference of the estimates of γ_0 is due to the faster convergence rate $\sqrt{n}$ instead of $\sqrt{T_n} = \sqrt{T}$.

Since we choose especially large values for n and T in Table 3.1, we can assume that the differences become more pronounced for smaller values. Therefore, comparing the boxplots for $T = 50, n = 200$ in Figures 3.3 to 3.5 the differences between the estimators are vanishingly small. This suggests that the preliminary estimator is already effective. The performance improvements become visually apparent when compared with the boxplots for $T = 500, n = 5000$.

If we take a closer look at the proofs, one essential aspect is that the moments with order at least -2 converge. To guarantee this, we choose $2\alpha > 5\gamma$ as our parameter assumption, see Lemma 1.8. However, we would expect a similar behaviour of the estimators for all positive true values when the origin is non attracting, that is, when $2\alpha > \gamma$. For this purpose, we consider the boundary case when $2\alpha = \gamma$ holds and hence choose $\theta_0 = (1, 1, 2)$ as the true value when using the same combinations of T and n as before. The boxplots presented in Figures 3.6 to 3.8 suggest that the presented estimators can converge in this case as well even if our proofs do not work here.

Table 3.1: The mean and the standard deviation (sd) of the estimators with true value $(\alpha_0, \beta_0, \gamma_0) = (3, 1, 1)$.

$n = 5000$		$T = 500$			$T = 1000$			$T = 2000$	
preliminary	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$
mean	3.0188	1.0079	0.9998	3.0200	1.0072	1.0023	3.0037	1.0015	1.0006
sd	0.2111	0.0752	0.0211	0.1570	0.0552	0.0235	0.1216	0.0421	0.0254
Newton	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$
mean	3.0141	1.0064	1.0001	3.0197	1.0071	1.0025	3.0027	1.0012	1.0010
sd	0.1872	0.0681	0.0220	0.1435	0.0515	0.0248	0.1156	0.0403	0.0270
scoring	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$
mean	3.0105	1.0052	0.9998	3.0182	1.0066	1.0023	3.0017	1.0008	1.0006
sd	0.1877	0.0682	0.0211	0.1434	0.0514	0.0235	0.1152	0.0400	0.0255
$n = 10000$									
preliminary	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$
mean	3.0351	1.0129	1.0001	3.0152	1.0060	1.0003	3.0097	1.0026	1.0012
sd	0.2143	0.0741	0.0145	0.1519	0.0541	0.0150	0.1135	0.0404	0.0167
Newton	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$
mean	3.0257	1.0099	1.0004	3.0121	1.0050	1.0005	3.0097	1.0026	1.0013
sd	0.1849	0.0654	0.0150	0.1319	0.0476	0.0156	0.1030	0.0372	0.0175
scoring	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$
mean	3.0219	1.0085	1.0001	3.0101	1.0043	1.0003	3.0089	1.0023	1.0012
sd	0.1849	0.0654	0.0145	0.1320	0.0475	0.0150	0.1028	0.0371	0.0167
$n = 20000$									
preliminary	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$	$\widehat{\alpha}_{0,n}$	$\widehat{\beta}_{0,n}$	$\widehat{\gamma}_{0,n}$
mean	3.0394	1.0140	1.0008	3.0102	1.0039	0.9996	3.0060	1.0023	1.0001
sd	0.2010	0.0717	0.0105	0.1464	0.0518	0.0104	0.1071	0.0374	0.0109
Newton	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$	$\widehat{\alpha}_{1,n}$	$\widehat{\beta}_{1,n}$	$\widehat{\gamma}_{1,n}$
mean	3.0263	1.0097	1.0010	3.0120	1.0045	0.9996	3.0043	1.0018	1.0002
sd	0.1756	0.0635	0.0106	0.1269	0.0460	0.0106	0.0946	0.0334	0.0112
scoring	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$	$\widehat{\alpha}_{2,n}$	$\widehat{\beta}_{2,n}$	$\widehat{\gamma}_{2,n}$
mean	3.0223	1.0083	1.0008	3.0101	1.0038	0.9996	3.0034	1.0015	1.0002
sd	0.1759	0.0636	0.0105	0.1270	0.0460	0.0104	0.0946	0.0333	0.0109

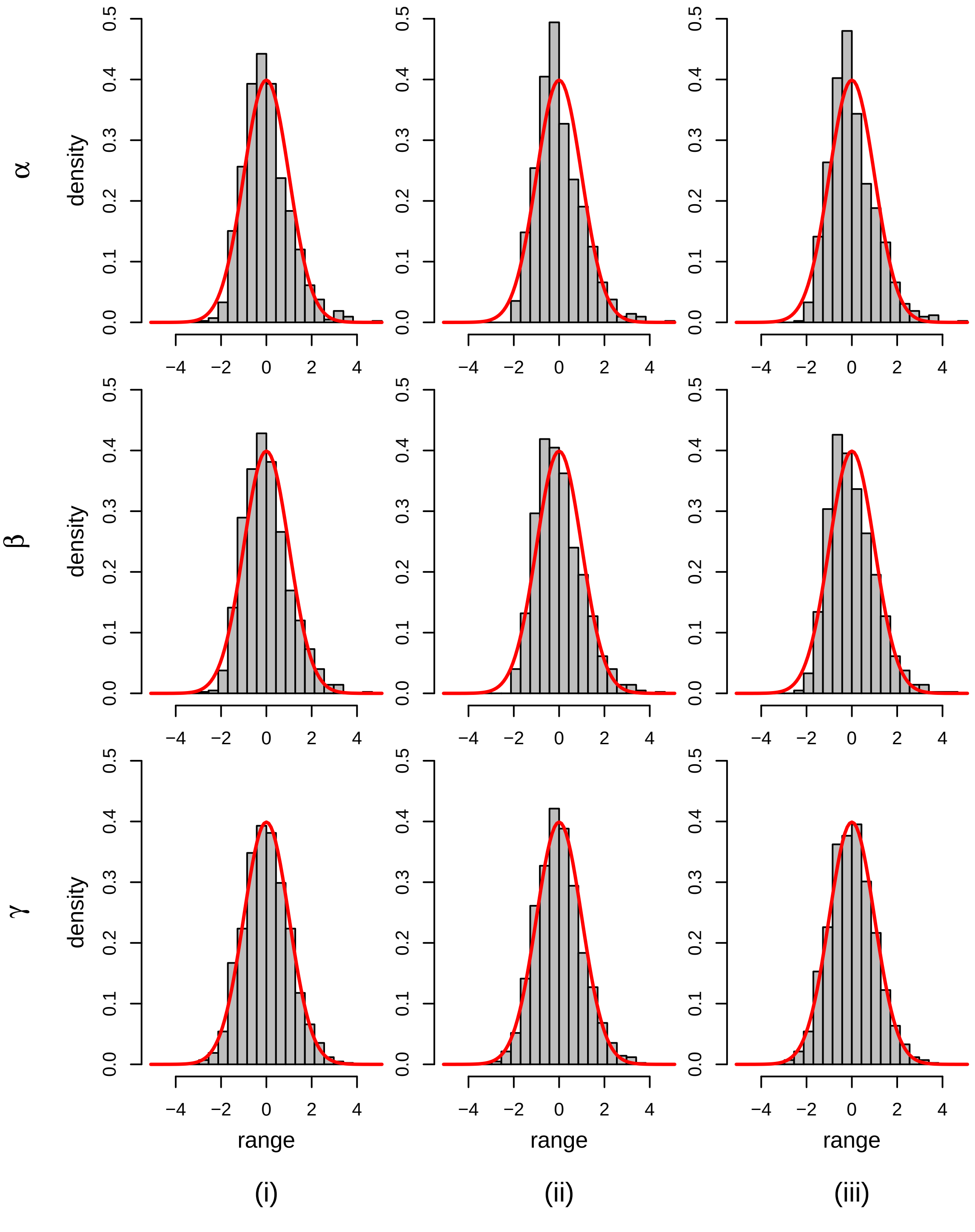

Figure 3.1: Histograms of the standardized estimators for $T = 50, n = 200$ with true value $\theta_0 = (3, 1, 1)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red curve indicates the density of the standard normal distribution.

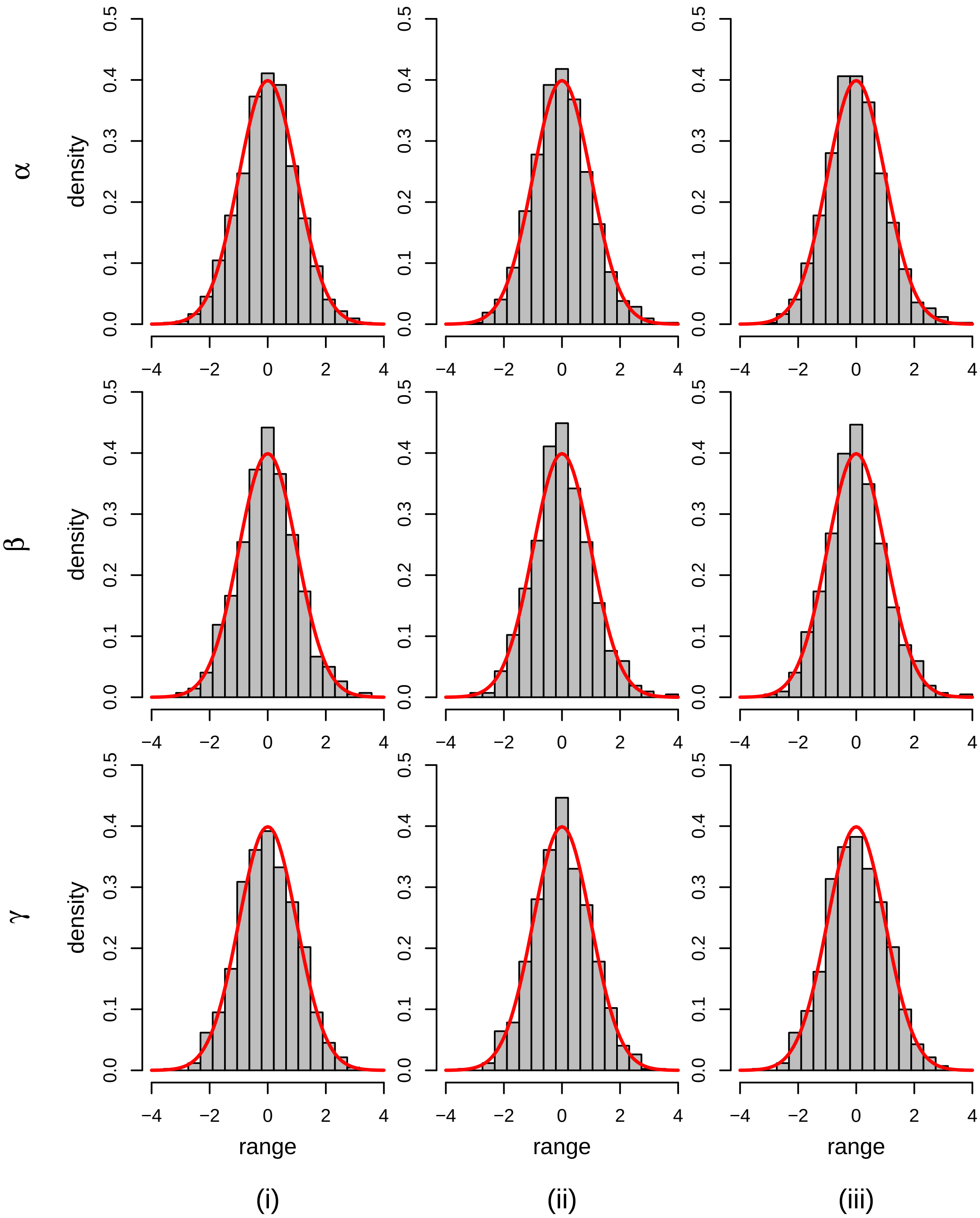

Figure 3.2: Histograms of the standardized estimators for $T = 500, n = 5000$ with true value $\theta_0 = (3, 1, 1)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red curve indicates the density of the standard normal distribution.

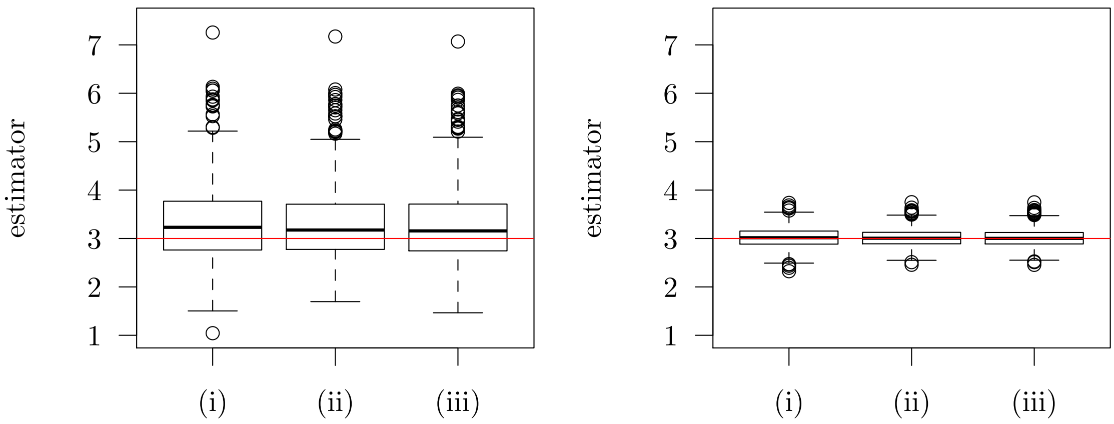

Figure 3.3: Results of estimating α_0 for $T = 50, n = 200$ (left) and $T = 500, n = 5000$ (right) with true value $\theta_0 = (3, 1, 1)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red line indicates the true value.

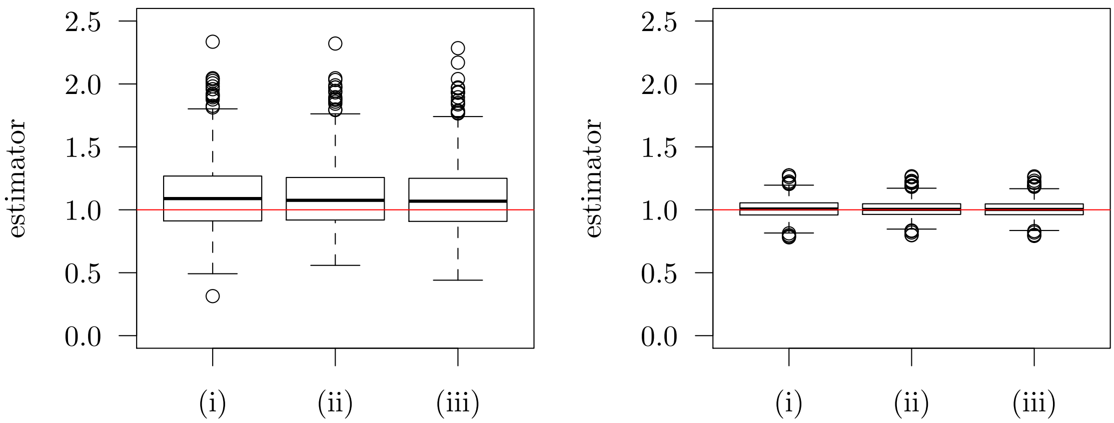

Figure 3.4: Results of estimating β_0 for $T = 50, n = 200$ (left) and $T = 500, n = 5000$ (right) with true value $\theta_0 = (3, 1, 1)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red line indicates the true value.

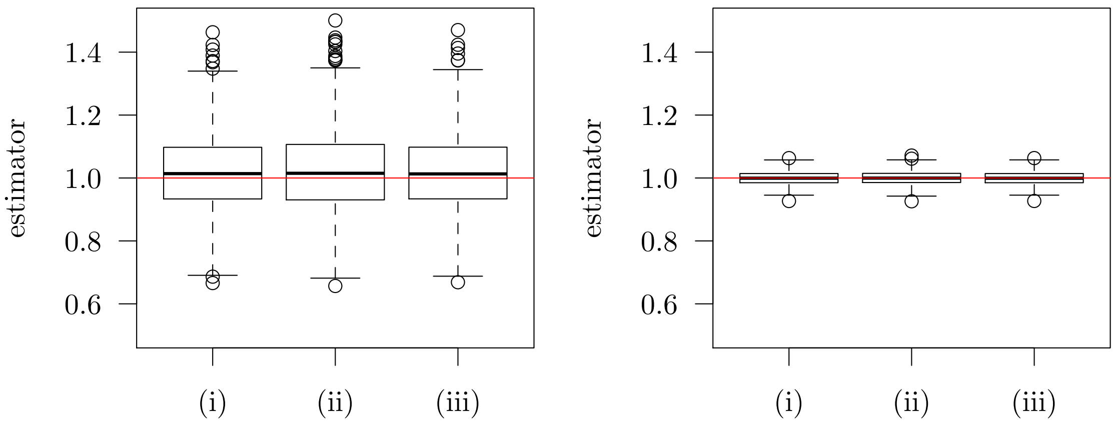

Figure 3.5: Results of estimating γ_0 for $T = 50, n = 200$ (left) and $T = 500, n = 5000$ (right) with true value $\theta_0 = (3, 1, 1)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red line indicates the true value.

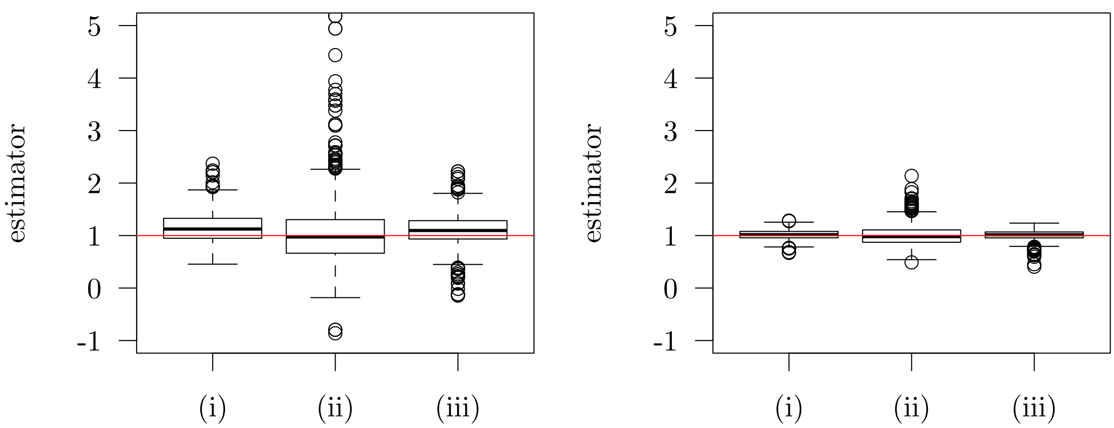

Figure 3.6: Results of estimating α_0 for $T = 50, n = 200$ (left) and $T = 500, n = 5000$ (right) with true value $\theta_0 = (1, 1, 2)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red line indicates the true value.

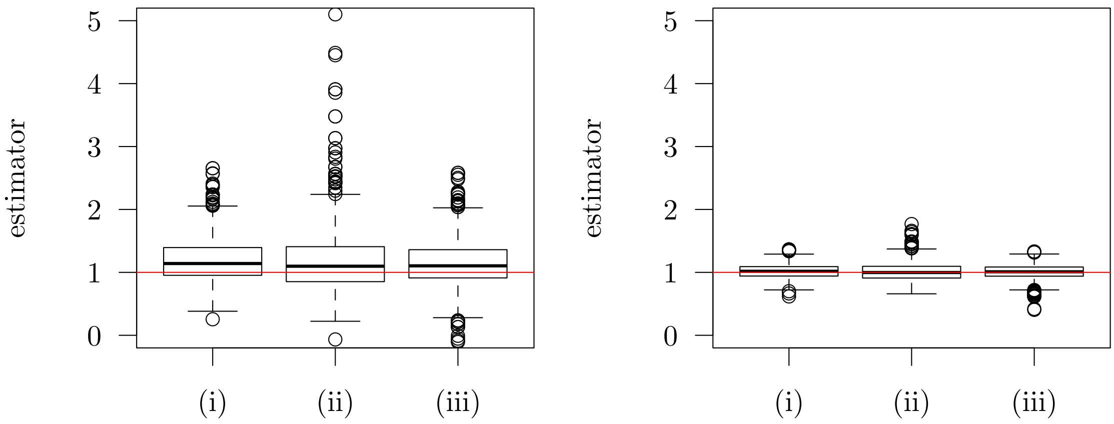

Figure 3.7: Results of estimating β_0 for $T = 50, n = 200$ (left) and $T = 500, n = 5000$ (right) with true value $\theta_0 = (1, 1, 2)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red line indicates the true value.

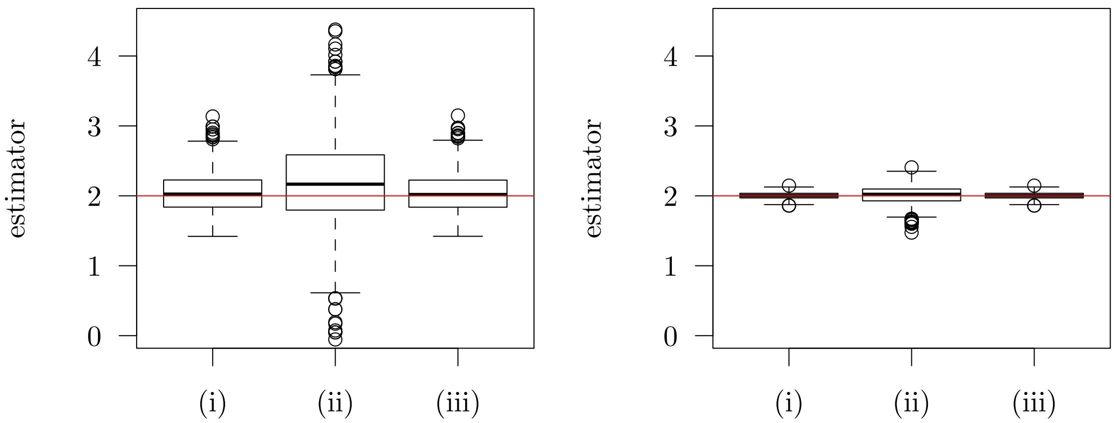

Figure 3.8: Results of estimating γ_0 for $T = 50, n = 200$ (left) and $T = 500, n = 5000$ (right) with true value $\theta_0 = (1, 1, 2)$: (i) preliminary estimator, (ii) Newton-Raphson method, (iii) Fisher scoring method. The red line indicates the true value.

3.4 Concluding remarks

We now return to our original problem of estimating the index parameter $\vartheta > -\frac{1}{2}$ of a classical Bessel process. For this we considered a space-time transformation in Chapter 2 so that the process becomes stationary. There, we noticed that our estimators depend only on the square of this modified process, which in turn is a Cox-Ingersoll-Ross process with parameter $\theta = \big(2\vartheta + 2, 2\alpha, 4\big)$. That is, we can now apply our estimators presented in this chapter to the problem as well. Analogous to Chapter 2, we can also apply all estimators to the Dunkl process and multidimensional Bessel process.

Comparing the estimators we should first consider the essential aspects. Primarily, we focus on low-frequency observations for the martingale estimators, whereas high-frequency observations are relevant here. This point on its own can affect the performance of the estimators significantly. Depending on the given situation, this may influence the choice of the estimator. The martingale estimators converge for each $\vartheta > -\frac{1}{2}$, although we additionally demand $\vartheta > 4$ and the boundedness of the parameter space for the estimators here in this situation. The most critical case arises when the origin is attracting for the Cox-Ingersoll-Ross process, that is, the origin is almost surely reached in finite time. In this case, our estimators presented in this chapter do not converge.

For further comparison, we now choose a value for ϑ such that all considered estimators converge. In Figure 3.9 we compare the Fisher information to the explicitly calculated asymptotic information, that is, the reciprocal of the asymptotic variance, of the previous chapter for $\vartheta = 5$. Here, the colored lines represent the asymptotic information for various Δ when considering one eigenfunction (solid line) respectively two eigenfunctions (dashed line) given the weights $(\omega_1, \omega_2) = (2, 1)$. The black solid line corresponds to the Fisher information where no Δ appears due to high frequency observations, that is, $\Delta \to 0$. The dashdotted lines represents the Fisher Information when including Δ. Here, the black solid line also represents the case $\Delta = 1$. We recall that the asymptotic information from the martingale estimators converge monotonically increasing for $\alpha\Delta \to \infty$, which can readily be seen in the figure. Depending on the value of Δ, we rapidly get infinitesimally close to the limit. In formulas, for one eigenfunction the limit is $\frac{1}{\vartheta+1}$ and in this specific case for two eigenfunctions the limit is $\frac{2\vartheta+5}{2(\vartheta+1)(\vartheta+2)}$, which in turn is greater than $\frac{1}{\vartheta+1}$. Looking at the shape of the Fisher information matrix (3.20), the asymptotic information when estimating ϑ is $\frac{\alpha}{\vartheta(\vartheta+1)}$. In particular the qualitative behaviour in terms of ϑ is the same for all three estimators, in that the asymptotic information increases as ϑ becomes smaller and vice versa. Moreover, the estimators improve in terms of asymptotic information with

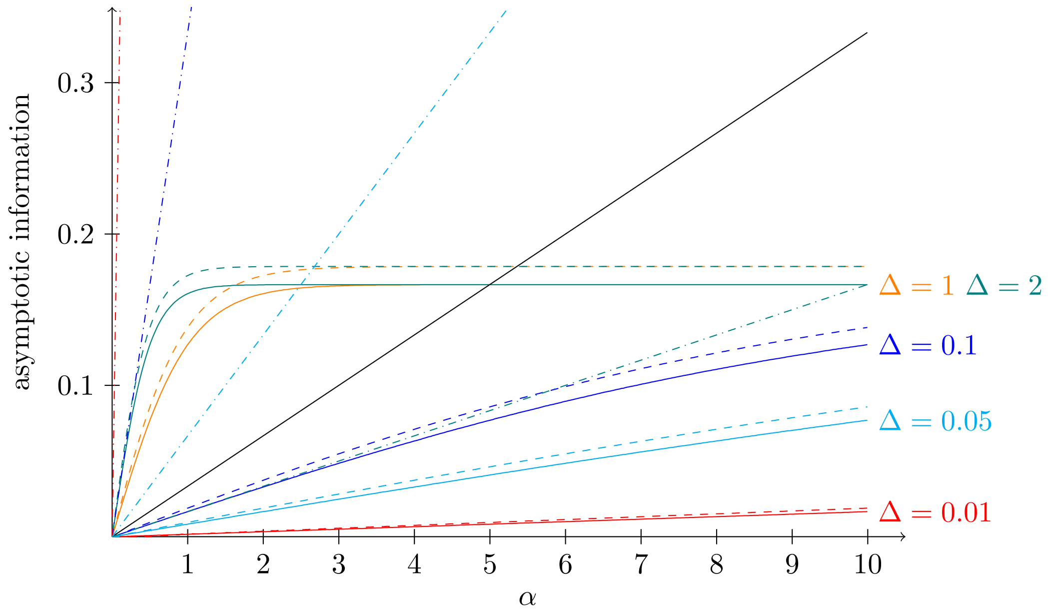

Figure 3.9: Comparison of the asymptotic information from Theorem 2.7 (colored solid line), the one from Theorem 2.11 (colored dashed line) for $\omega_1 = 2$ and $\omega_2 = 1$ to the Fisher information (black solid line and colored dashdotted line) for $\vartheta = 5$.

increasing α. This improvement is faster noticeable but limited for the martingale estimators, whereas it continues to grow linearly for the Fisher information. As the asymptotic information are symmetric in α and Δ, we observe a slower convergence in Figure 3.9 for small values of Δ. For this reason, the preliminary and one-step estimators perform better for small Δ values, that is, small distances between observations, whereas the martingale estimators perform better for large values of Δ, e.g., greater than 1. We can recognize this behaviour significantly in comparison to the dashdotted lines. For the latter, the value of alpha seems to be decisive. Above a value of $\alpha = 5$, the Fisher information has the largest value even in the case $\Delta = 1$. Note that the corresponding estimators do not behave well for large values of Δ since they are based on high-frequency observations.

As mentioned above, it is particularly interesting that the martingale estimators converge for all values $\vartheta > -\frac{1}{2}$, which is not the case for the estimators discussed in this chapter. Here, the distinction whether the origin is attractive was essential. In the next chapter, we will study this phenomenon for the classical and multivariate Bessel process.

4 Return times for the multivariate Bessel process

The content of this chapter is partially incorporated in the preprint

Hausdorff dimension of collision times in one-dimensional log-gases
arXiv:2109.08707 (2021)
Nicole Hufnagel, Sergio Andraus

and contains some additional thoughts which arose during this collaboration.

4.1 Setting and stopping times related to the Bessel process

In this chapter we deal with the multivariate Bessel process and both its hitting and return times. We can perform our initial considerations for the classical Bessel process which turns out to coincide with the case A_1. Since this process is already well studied, we gather known results from the literature on hitting times and infer new formulas. In this section we first refresh the most important aspects of the multivariate Bessel process $(Y_t)_{t\geq 0}$ of type A_{N-1} given via

$$\begin{cases} \mathrm{d}Y_{t,i} &= \mathrm{d}B_{t,i} + k \sum\limits_{\substack{j=1\\ j\neq i}}^{N} \frac{1}{Y_{t,i}-Y_{t,j}}\,\mathrm{d}t, \\ Y_0 &= y \in W_{A_{N-1}} \end{cases}$$

for $i = 1, \dots, N$, where $(B_t)_{t\geq 0}$ is a standard multivariate Brownian motion, for more details see Section 1.2. From a physical point of view, the dimension represents N particles, while $k > 0$, the inverse of the temperature, regulates the interaction. The process lives on the closure of the Weyl chamber $\overline{W}_{A_{N-1}} = \{x \in \mathbb{R}^N \,|\, x_1 \leq \dots \leq x_N\}$, which means the particles are ordered. Due to [19, Proposition 4.1], if $k > \frac{1}{2}$ no collisions with the boundary $\partial W_{A_{N-1}} := \big\{x \in \overline{W}_{A_{N-1}} \,|\, \exists i \in \{1, \dots, N-1\} : x_i = x_{i+1}\big\}$ almost surely will occur, despite when $k < \frac{1}{2}$ the first collision time of two particles is almost surely finite [26, Proposition 1]. From now on, we will consider the case $k < \frac{1}{2}$, high temperature, where we still have a unique strong solution, [26, Theorem 1]. As a starting point, we take a closer look at the case $N = 2$

$$\begin{aligned} \mathrm{d}Y_{t,1} &= \mathrm{d}B_{t,1} + k\frac{\mathrm{d}t}{Y_{t,1} - Y_{t,2}}, \\ \mathrm{d}Y_{t,2} &= \mathrm{d}B_{t,2} + k\frac{\mathrm{d}t}{Y_{t,2} - Y_{t,1}}. \end{aligned}$$

The closure of the Weyl chamber simplifies to $\overline{W}_{A_1} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1 \leq x_2\}$. Here, we can decompose the process into a classical Bessel process

$$\begin{aligned} \mathrm{d}\left(\frac{1}{\sqrt{2}}\big(Y_{t,2} - Y_{t,1}\big)\right) &= \frac{1}{\sqrt{2}}\left(\mathrm{d}B_{t,2} - \mathrm{d}B_{t,1} + k\frac{\mathrm{d}t}{Y_{t,2} - Y_{t,1}} - k\frac{\mathrm{d}t}{Y_{t,1} - Y_{t,2}}\right) \\ &=: \mathrm{d}\widetilde{B}_{t,1} + k\frac{\sqrt{2}}{Y_{t,2} - Y_{t,1}}\,\mathrm{d}t \end{aligned}$$

and a Brownian motion

$$\begin{aligned} \mathrm{d}\left(\frac{1}{\sqrt{2}}\big(Y_{t,1} + Y_{t,2}\big)\right) &= \frac{1}{\sqrt{2}}\left(\mathrm{d}B_{t,1} + \mathrm{d}B_{t,2} + k\frac{\mathrm{d}t}{Y_{t,1} - Y_{t,2}} + k\frac{\mathrm{d}t}{Y_{t,2} - Y_{t,1}}\right) \\ &=: \frac{1}{\sqrt{2}}\big(\mathrm{d}B_{t,1} + \mathrm{d}B_{t,2}\big) = \mathrm{d}\widetilde{B}_{t,2} \end{aligned}$$

by looking at the sum and difference of the two particles. Hitting the boundary of the Weyl chamber implies the two particles have the same value. Consequently, we analyze when the Bessel process $\big(X_t := \frac{1}{\sqrt{2}}\left(Y_{t,2} - Y_{t,1}\right)\big)_{t \geq 0}$ with index $k - \frac{1}{2}$ hits the origin. For a significantly shorter notation, we use $\vartheta := k - \frac{1}{2}$ in the following. As we are interested in the times when the classical Bessel process hits the origin, one intuition is to look at the first hitting time while starting in $z > 0$. We denote the corresponding probability measure by $\mathbb{P}^z$. Therefore, we define

$$\tau_x := \inf\big\{t > 0 : X_t = x\big\}$$

for every $x \geq 0$, which is obviously a stopping time. As mentioned, it is well-known that, given enough time, Bessel processes hit the origin with probability one, in formulas $\mathbb{P}^z(\tau_0 < \infty) = 1$, whenever $-\frac{1}{2} < \vartheta < 0$, [56, Theorem 1.1 (iv)]. This situation points to the intuition that for small parameter ϑ the repulsion from the origin, indicated by the drift, is not large enough to keep the process from returning to it, which makes this parameter region interesting. Whenever the starting point z is positive, the density with respect to the Lebesgue measure of τ_0 under $\mathbb{P}^z$ is known to be

$$f_{\tau_0}(t) = \frac{1}{t\Gamma(-\vartheta)} \left(\frac{z^2}{2t}\right)^{-\vartheta} e^{-\frac{z^2}{2t}} \mathbb{1}_{(0,\infty)}.$$

This expression was derived in [38, 2.1. First hitting times of Bessel processes, Eq. (15)]. For the sake of completeness, we can calculate the density of the elapsed time for the process to go from the state $0 \leq y < x$ up to state x,

$$\tau_{y,x} := \inf\left\{t > \tau_y : X_t = x\right\} - \tau_y.$$

Owing to the Markov property, $\tau_{y,x}$ has the same density under every $\mathbb{P}^z$ for $z \geq 0$. Based on [57] we are going to first determine the density of $\tau_{y,x}$ for $y > 0$ and then derive the density of $\tau_{0,x}$. Due to [57, Theorem 3.1] the Laplace transform of $\tau_{y,x}$ is

$$\phi_{\tau_{y,x}}(s) = \left(\frac{x}{y}\right)^{\vartheta} \frac{I_\vartheta\left(\sqrt{2s}y\right)}{I_\vartheta\left(\sqrt{2s}x\right)} \tag{4.1}$$

for $s > 0$, where I_ϑ is the modified Bessel function of the first kind defined as

$$I_\vartheta(x) = \left(\frac{x}{2}\right)^{\vartheta} \sum_{n=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2n}}{n!\Gamma(n+\vartheta+1)}.$$

We denote the corresponding density by $f_{\tau_{y,x}}$ and consequently for $s > 0$ the relation

$$\phi_{\tau_{y,x}}(s) = \int_0^\infty e^{-st} f_{\tau_{y,x}}(t)\,\mathrm{d}t$$

applies, see [15, Eq. (15.5)]. We recognize that the domain of $f_{\tau_{y,x}}$ is $(0,\infty)$, whereas the Laplace transform $\phi_{\tau_{y,x}}$ is always a function with complex argument s. In order to get the distribution of $\tau_{y,x}$ we use the inverse Laplace transform [15, Eq. (15.8)]

$$f_{\tau_{y,x}}(t) = \frac{1}{2\pi i} \oint_{\zeta-i\infty}^{\zeta+i\infty} e^{st} \phi_{\tau_{y,x}}(s)\,\mathrm{d}s \overset{(4.1)}{=} \frac{1}{2\pi i} \oint_{\zeta-i\infty}^{\zeta+i\infty} e^{st} \left(\frac{x}{y}\right)^{\vartheta} \frac{I_\vartheta\left(\sqrt{2s}y\right)}{I_\vartheta\left(\sqrt{2s}x\right)}\,\mathrm{d}s, \tag{4.2}$$

where ζ is a positive real number. The condition $\zeta > 0$ guarantees the applicability of the inverse Laplace transform, since the singularities i.e. zeros of I_ν are purely complex. In particular, this means that the real part of all singularities is zero, which is smaller than ζ. The zeros of I_ν are considered in more detail below. We deal with this integral by using the Residue theorem [3, 5.1. The Residue Theorem], hence we first determine the poles of the integrand. Several properties are known for the zeros of the Bessel function of the first kind J_ϑ. For this purpose, we need the relation formula

$$I_\vartheta(z) = e^{\mp\frac{\vartheta\pi i}{2}} J_\vartheta\left(ze^{\pm\frac{\pi i}{2}}\right), \tag{4.3}$$

cf. [69, Eq. 10.27.6]. In [83, 15.21 The non-repetition of zeros of cylinder functions] it was shown that all zeros of J_ϑ are simple with the origin as a possible exception. Additionally, in the case $\vartheta > -1$ all zeros are real, [82, 19. Der Fourier'sche Lehrsatz], and positive. We suppose $j_{\vartheta,\eta}$ to be the ordered zeros

$$0 < j_{\vartheta,1} < j_{\vartheta,2} < \cdots .$$

Accordingly, if we substitute

$$J_\vartheta(z) = \frac{\left(\frac{z}{2}\right)^\vartheta}{\Gamma(\vartheta+1)} \prod_{\eta=1}^{\infty} \left(1 - \frac{z^2}{j_{\vartheta,\eta}^2}\right), \tag{4.4}$$

cf. [40, p. 130 Eq. (8.)], into (4.3), we derive

$$\begin{aligned} I_\vartheta(z) &= e^{\mp\frac{\vartheta\pi i}{2}} J_\vartheta\left(ze^{\pm\frac{\pi i}{2}}\right) \\ &= \frac{e^{\mp\frac{\vartheta\pi i}{2}}}{\Gamma(\vartheta+1)} \left(\frac{ze^{\pm\frac{\pi i}{2}}}{2}\right)^\vartheta \prod_{\eta=1}^{\infty} \left(1 - \frac{\left(ze^{\pm\frac{\pi i}{2}}\right)^2}{j_{\vartheta,\eta}^2}\right) \\ &= \frac{\left(\frac{z}{2}\right)^\vartheta}{\Gamma(\vartheta+1)} \prod_{\eta=1}^{\infty} \left(1 + \frac{z^2}{j_{\vartheta,\eta}^2}\right) \end{aligned}$$

and hence receive $\pm i j_{\vartheta,\eta}$ as zeros of I_ϑ, which have no real part. As the zeros are pairwise disjoint, we just get simple poles in (4.2). By plugging in the calculated formula of I_ϑ we conclude

$$f_{\tau_{y,x}}(t) \overset{(4.2)}{=} \frac{1}{2\pi i} \oint_{\zeta-i\infty}^{\zeta+i\infty} e^{st} \left(\frac{x}{y}\right)^\vartheta \frac{I_\vartheta(\sqrt{2s}y)}{I_\vartheta(\sqrt{2s}x)} \, \mathrm{d}s$$

$$
\begin{aligned}
&= \frac{1}{2\pi i} \oint\limits_{\zeta - i\infty}^{\zeta + i\infty} e^{st} \prod_{\eta=1}^{\infty} \left(\frac{1 + \frac{2sy^2}{j^2_{\vartheta,\eta}}}{1 + \frac{2sx^2}{j^2_{\vartheta,\eta}}} \right) \mathrm{d}s \\
&= \frac{1}{2\pi i} \oint\limits_{\zeta - i\infty}^{\zeta + i\infty} \underbrace{e^{st} \prod_{\eta=1}^{\infty} \frac{1}{2x^2} \left(\frac{j^2_{\vartheta,\eta} + 2sy^2}{\frac{j^2_{\vartheta,\eta}}{2x^2} + s} \right)}_{=:g(s)} \mathrm{d}s.
\end{aligned}
$$

By the Residue theorem and because g has only simple poles, we can calculate the integral as the following sum of residues

$$
\begin{aligned}
f_{\tau_{y,x}}(t) &= \sum_{l=1}^{\infty} \lim_{z \to -\frac{j^2_{\vartheta,l}}{2x^2}} \left(z + \frac{j^2_{\vartheta,l}}{2x^2} \right) g(z) \\
&= \sum_{l=1}^{\infty} \lim_{z \to -\frac{j^2_{\vartheta,l}}{2x^2}} \frac{e^{zt}}{2x^2} \cdot \frac{\prod\limits_{\eta=1}^{\infty} \left(j^2_{\vartheta,\eta} + 2y^2 z \right)}{\prod\limits_{\substack{\eta=1 \\ \eta \neq l}}^{\infty} \left(j^2_{\vartheta,\eta} + 2x^2 z \right)} \\
&= \sum_{l=1}^{\infty} \frac{e^{-t\frac{j^2_{\vartheta,l}}{2x^2}}}{2x^2} \cdot \frac{\prod\limits_{\eta=1}^{\infty} \left(j^2_{\vartheta,\eta} - \frac{y^2}{x^2} j^2_{\vartheta,l} \right)}{\prod\limits_{\substack{\eta=1 \\ \eta \neq l}}^{\infty} \left(j^2_{\vartheta,\eta} - j^2_{\vartheta,l} \right)}.
\end{aligned}
$$

Now, we reformulate the last line by using some basic calculations and simplify the formula of the density:

$$
\begin{aligned}
f_{\tau_{y,x}}(t) &= \sum_{l=1}^{\infty} \frac{e^{-t\frac{j^2_{\vartheta,l}}{2x^2}}}{2x^2} \cdot \frac{\prod\limits_{\eta=1}^{\infty} \left(j^2_{\vartheta,\eta} - \frac{y^2}{x^2} j^2_{\vartheta,l} \right)}{\prod\limits_{\substack{\eta=1 \\ \eta \neq l}}^{\infty} \left(j^2_{\vartheta,\eta} - j^2_{\vartheta,l} \right)} \\
&= \sum_{l=1}^{\infty} \frac{e^{-t\frac{j^2_{\vartheta,l}}{2x^2}}}{2x^2} \cdot \frac{\prod\limits_{\eta=1}^{\infty} \left(\frac{j^2_{\vartheta,\eta}}{j^2_{\vartheta,\eta}} - \frac{y^2 j^2_{\vartheta,l}}{x^2 j^2_{\vartheta,\eta}} \right)}{\frac{1}{j^2_{\vartheta,l}} \prod\limits_{\substack{\eta=1 \\ \eta \neq l}}^{\infty} \left(\frac{j^2_{\vartheta,\eta}}{j^2_{\vartheta,\eta}} - \frac{j^2_{\vartheta,l}}{j^2_{\vartheta,\eta}} \right)}
\end{aligned}
$$

$$= \sum_{l=1}^{\infty} \frac{j_{\vartheta,l}^2}{2x^2} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \cdot \frac{\prod_{\eta=1}^{\infty}\left(1-\frac{y^2 j_{\vartheta,l}^2}{x^2 j_{\vartheta,\eta}^2}\right)}{\prod_{\substack{\eta=1\\ \eta\neq l}}^{\infty}\left(1-\frac{j_{\vartheta,l}^2}{j_{\vartheta,\eta}^2}\right)}.$$

We can rewrite the product in the numerator as

$$\prod_{\eta=1}^{\infty}\left(1-\frac{y^2 j_{\vartheta,l}^2}{x^2 j_{\vartheta,\eta}^2}\right) \overset{(4.4)}{=} \frac{\Gamma(\vartheta+1)}{\left(\frac{y}{2x} j_{\vartheta,l}\right)^{\vartheta}} J_\vartheta\left(\frac{y}{x} j_{\vartheta,l}\right).$$

The next step is to find a method of expressing the denominator in a similar way. For this purpose, we derive

$$\begin{aligned}\frac{\partial}{\partial s}\prod_{\eta=1}^{\infty}\left(1-\frac{s}{j_{\vartheta,\eta}^2}\right)\Bigg|_{s=j_{\vartheta,l}^2} &= -\sum_{m=1}^{\infty}\frac{1}{j_{\vartheta,m}^2}\prod_{\substack{\eta=1\\ \eta\neq m}}^{\infty}\left(1-\frac{s}{j_{\vartheta,\eta}^2}\right)\Bigg|_{s=j_{\vartheta,l}^2}\\ &= -\frac{1}{j_{\vartheta,l}^2}\prod_{\substack{\eta=1\\ \eta\neq l}}^{\infty}\left(1-\frac{j_{\vartheta,l}}{j_{\vartheta,\eta}^2}\right)\end{aligned}$$

to obtain the denominator. Secondly, we are looking for a way to display the product on the left-hand side differently and also differentiate this new representation. In this case, we also use the relation

$$\prod_{\eta=1}^{\infty}\left(1-\frac{s}{j_{\vartheta,\eta}^2}\right) \overset{(4.4)}{=} \frac{\Gamma(\vartheta+1)}{\left(\frac{\sqrt{s}}{2}\right)^{\vartheta}} J_\vartheta(\sqrt{s})$$

via the Bessel function J_ϑ. For the differentiation of this Bessel function there exists a simple formula,

$$\frac{\partial}{\partial z} z^{-\vartheta} J_\vartheta(z) = -z^{-\vartheta} J_{\vartheta+1}(z),$$

see [69, Eq. 10.29.4], which we can apply

$$\frac{\partial}{\partial s}\frac{\Gamma(\vartheta+1)}{\left(\frac{\sqrt{s}}{2}\right)^{\vartheta}} J_\vartheta(\sqrt{s})\Bigg|_{s=j_{\vartheta,l}^2} = 2^{\vartheta}\Gamma(\vartheta+1)\frac{\partial}{\partial s}(\sqrt{s})^{-\vartheta} J_\vartheta(\sqrt{s})\Bigg|_{s=j_{\vartheta,l}^2}$$

$$= -2^{\vartheta}\Gamma(\vartheta+1)\frac{(\sqrt{s})^{-\vartheta}J_{\vartheta+1}(\sqrt{s})}{2\sqrt{s}}\bigg|_{s=j_{\vartheta,l}^2}$$
$$= -2^{\vartheta-1}\Gamma(\vartheta+1)j_{\vartheta,l}^{-\vartheta-1}J_{\vartheta+1}(j_{\vartheta,l}).$$

Finally, by comparing both approaches we accomplish

$$\frac{1}{j_{\vartheta,l}^2}\prod_{\substack{\eta=1\\ \eta\neq l}}^{\infty}\left(1-\frac{j_{\vartheta,l}^2}{j_{\vartheta,\eta}^2}\right) = 2^{\vartheta-1}\Gamma(\vartheta+1)j_{\vartheta,l}^{-\vartheta-1}J_{\vartheta+1}(j_{\vartheta,l})$$

and infer

$$f_{\tau_{y,x}}(t) = \sum_{l=1}^{\infty}\frac{j_{\vartheta,l}^2}{2x^2}e^{-t\frac{j_{\vartheta,l}^2}{2x^2}}\cdot\frac{\prod\limits_{\eta=1}^{\infty}\left(1-\frac{y^2 j_{\vartheta,l}^2}{x^2 j_{\vartheta,\eta}^2}\right)}{\prod\limits_{\substack{\eta=1\\ \eta\neq l}}^{\infty}\left(1-\frac{j_{\vartheta,l}^2}{j_{\vartheta,\eta}^2}\right)}$$
$$= \sum_{l=1}^{\infty}\frac{j_{\vartheta,l}^2}{2x^2}e^{-t\frac{j_{\vartheta,l}^2}{2x^2}}\frac{\Gamma(\vartheta+1)\left(\frac{2x}{yj_{\vartheta,l}}\right)^{\vartheta}J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right)}{2^{\vartheta-1}\Gamma(\vartheta+1)j_{\vartheta,l}^{-\vartheta+1}J_{\vartheta+1}(j_{\vartheta,l})}$$
$$= \sum_{l=1}^{\infty}j_{\vartheta,l}\frac{x^{\vartheta-2}}{y^{\vartheta}}\cdot\frac{J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right)}{J_{\vartheta+1}(j_{\vartheta,l})}\cdot e^{-t\frac{j_{\vartheta,l}^2}{2x^2}}.$$

This expression is well-defined as for different parameters the zeros are disjoint and hence $J_{\vartheta+1}(j_{\vartheta,\eta}) \neq 0$ for all $\eta \in \mathbb{N}$, see [69, Section 10.21 Zeros]. Therefore, we proved the following statement.

Corollary 4.1: The density with respect to the Lebesgue measure of

$$\tau_{y,x} := \inf\left\{t > \tau_y : X_t = x\right\} - \tau_y$$

with $\tau_y := \inf\left\{t > 0 : X_t = y\right\}$, that is, the time that passes while a classical Bessel process with index $-\frac{1}{2} < \vartheta < 0$ proceeds from state y to state x with $0 < y < x$, is given via

$$f_{\tau_{y,x}}(t) = \sum_{l=1}^{\infty}j_{\vartheta,l}\frac{x^{\vartheta-2}}{y^{\vartheta}}\cdot\frac{J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right)}{J_{\vartheta+1}(j_{\vartheta,l})}\cdot e^{-t\frac{j_{\vartheta,l}^2}{2x^2}}\mathbb{1}_{(0,\infty)}(t).$$

Corollary 4.2: The density with respect to the Lebesgue measure of $\tau_{0,x}$ for $x > 0$ and

$-\frac{1}{2} < \vartheta < 0$ is determined by

$$f_{\tau_{0,x}}(t) = \sum_{l=1}^{\infty} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \cdot \frac{j_{\vartheta,l}^{\vartheta+1}}{2^{\vartheta}x^2\Gamma(\vartheta+1)J_{\vartheta+1}(j_{\vartheta,l})}\mathbb{1}_{(0,\infty)}(t).$$

Proof: We separate the proof in two steps.
Step 1: First of all, we show the validity of

$$\lim_{y\to 0} f_{\tau_{y,x}}(t) = \sum_{l=1}^{\infty} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \cdot \frac{j_{\vartheta,l}^{\vartheta+1}}{2^{\vartheta}x^2\Gamma(\vartheta+1)J_{\vartheta+1}(j_{\vartheta,l})}$$

for $t > 0$. For this, we use the explicit formula [69, Eq. 10.2.2] of the Bessel function

$$\begin{aligned}
\lim_{y\to 0}\left(\frac{x}{y}\right)^{\vartheta} J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right) &= \lim_{y\to 0}\left(\frac{x}{y}\right)^{\vartheta}\left(\frac{y}{2x}j_{\vartheta,l}\right)^{\vartheta}\sum_{n=0}^{\infty}(-1)^n\frac{\left(\frac{y}{2x}j_{\vartheta,l}\right)^{2n}}{n!\Gamma(n+\vartheta+1)} \\
&= \lim_{y\to 0}\left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\sum_{n=0}^{\infty}(-1)^n\frac{\left(\frac{y}{2x}j_{\vartheta,l}\right)^{2n}}{n!\Gamma(n+\vartheta+1)} \\
&= \left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\sum_{n=0}^{\infty}(-1)^n\frac{0^{2n}}{n!\Gamma(n+\vartheta+1)} \\
&= \left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\frac{1}{\Gamma(\vartheta+1)}.
\end{aligned}$$

The third line is valid due to the dominated convergence theorem which is applicable since the ratio test

$$\begin{aligned}
\lim_{n\to\infty}\frac{(n+1)!\Gamma(n+\vartheta+2)}{n!\Gamma(n+\vartheta+1)} &= \lim_{n\to\infty}\frac{(n+1)(n+\vartheta+1)\Gamma(n+\vartheta+1)}{\Gamma(n+\vartheta+1)} \\
&= \lim_{n\to\infty}(n+1)(n+\vartheta+1) = \infty
\end{aligned}$$

ensures that the power series

$$\sum_{n=0}^{\infty}\left(\frac{j_{\vartheta,l}}{2}\right)^{2n}\frac{1}{n!\Gamma(n+\vartheta+1)}$$

has convergence radius infinity. This indeed justifies the calculation above as $0 < \frac{y}{x} < 1$.

If we can now interchange the following limit and the series, we complete the first step

$$\begin{aligned}\lim_{y\to 0} f_{\tau_{y,x}}(t) &= \lim_{y\to 0}\sum_{l=1}^{\infty} j_{\vartheta,l}\frac{x^{\vartheta-2}}{y^{\vartheta}}\cdot\frac{J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right)}{J_{\vartheta+1}(j_{\vartheta,l})}\cdot e^{-t\frac{j^2_{\vartheta,l}}{2x^2}}\\ &= \sum_{l=1}^{\infty} e^{-t\frac{j^2_{\vartheta,l}}{2x^2}}\frac{j^{\vartheta+1}_{\vartheta,l}}{2^{\vartheta}x^2\Gamma(\vartheta+1)J_{\vartheta+1}(j_{\vartheta,l})}.\end{aligned}$$

Thus, we are left with the task of finding an upper bound for

$$\left(\frac{x}{y}\right)^{\vartheta} J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right) = a^{-\vartheta}J_{\vartheta}(aj_{\vartheta,l})$$

independent of $a := \frac{y}{x} \in (0,1)$ which justifies the exchange of limits by means of the dominated convergence theorem. By using the integral representation of J_{ϑ}, see [69, Eq. 10.9.4], we calculate

$$\begin{aligned}\left|a^{-\vartheta}J_{\vartheta}(aj_{\vartheta,l})\right| &= a^{-\vartheta}\left(\frac{aj_{\vartheta,l}}{2}\right)^{\vartheta}\frac{2}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma(\frac{1}{2})}\left|\int_0^1 \cos(aj_{\vartheta,l}t)(1-t^2)^{\vartheta-\frac{1}{2}}\,\mathrm{d}t\right|\\ &\le \left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\frac{2}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma(\frac{1}{2})}\int_0^1 (1-t^2)^{\vartheta-\frac{1}{2}}\,\mathrm{d}t\\ &= \left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma(\frac{1}{2})}\int_0^1 (1-s)^{\vartheta-\frac{1}{2}}s^{\frac{1}{2}}\,\mathrm{d}s\\ &= \left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma(\frac{1}{2})}\cdot\frac{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma(\frac{1}{2})}{\Gamma(\vartheta+1)}\\ &= \left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\frac{1}{\Gamma(\vartheta+1)}\end{aligned}$$

and conclude

$$\begin{aligned}\left|f_{\tau_{y,x}}(t)\right| &= \left|\sum_{l=1}^{\infty} e^{-t\frac{j^2_{\vartheta,l}}{2x^2}} j_{\vartheta,l}\frac{x^{\vartheta-2}}{y^{\vartheta}}\cdot\frac{J_{\vartheta}\left(\frac{y}{x}j_{\vartheta,l}\right)}{J_{\vartheta+1}(j_{\vartheta,l})}\right|\\ &\le \sum_{l=1}^{\infty} e^{-t\frac{j^2_{\vartheta,l}}{2x^2}}\frac{j_{\vartheta,l}}{x^2}\left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta}\frac{1}{\Gamma(\vartheta+1)|J_{\vartheta+1}(j_{\vartheta,l})|}.\end{aligned}$$

The right-hand side is now independent of y, leaving us merely to argue the convergence of the series. Therefore, we take a closer look at the asymptotic behaviour of the zeros

$j_{\vartheta,l}$ and the Bessel function J_ϑ. In [50, p. 174] the asymptotic

$$J_\vartheta(z) = \sqrt{\frac{2}{\pi z}}\left[\cos\left(z - \frac{\vartheta\pi}{2} - \frac{\pi}{4}\right) + O\big(z^{-1}\big)\right] \tag{4.5}$$

for $z \to \infty$ is shown, from which Mc Mahon's formula [1, Eq. 9.5.12]

$$j_{\vartheta,l} = \left(l + \frac{\vartheta}{2} - \frac{1}{4}\right)\pi + O\big(l^{-1}\big) \tag{4.6}$$

arises for $l \to \infty$. In particular, we observe from (4.6) that the zeros converge to infinity. This justifies the use of formula (4.5) when considering

$$\frac{1}{|J_{\vartheta+1}(j_{\vartheta,l})|} = \sqrt{\pi j_{\vartheta,l}} \cdot \frac{1}{\left|\sqrt{2}\cos\left(j_{\vartheta,l} - \frac{(\vartheta+1)\pi}{2} - \frac{\pi}{4}\right) + O\big(j_{\vartheta,l}^{-1}\big)\right|}.$$

The behaviour of the fraction depends on the values of the cosine

$$\begin{aligned}\cos\left(j_{\vartheta,l} - \frac{\vartheta\pi}{2} - \frac{\pi}{4}\right) &\overset{(4.6)}{=} \cos\left(\left(l + \frac{\vartheta}{2} - \frac{1}{4} - \frac{\vartheta+1}{2} - \frac{1}{4}\right)\pi + O(l^{-1})\right)\\ &= \cos\big(\pi(l-1) + O\big(l^{-1}\big)\big) =: f_l.\end{aligned}$$

Due to $|f_l| \to 1$ and $O(l^{-1}) \to 0$ when $l \to \infty$, there exists an $N \in \mathbb{N}$ such that $1 < \big|\sqrt{2} f_l + O\big(l^{-1}\big)\big|$ for all $l \geq N$. If we now combine all these calculations we achieve

$$\begin{aligned}\big|f_{\tau_{y,x}}(t)\big| &\leq \sum_{l=1}^{\infty} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \frac{j_{\vartheta,l}}{x^2}\left(\frac{j_{\vartheta,l}}{2}\right)^{\vartheta} \frac{1}{\Gamma(\vartheta+1)|J_{\vartheta+1}(j_{\vartheta,l})|}\\ &= \frac{1}{x^2 2^\vartheta \Gamma(\vartheta+1)}\left(\sum_{l=1}^{N-1} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \frac{j_{\vartheta,l}^{\vartheta+1}}{|J_{\vartheta+1}(j_{\vartheta,l})|}\right.\\ &\qquad \left. + \sqrt{\pi}\sum_{l=N}^{\infty} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \frac{j_{\vartheta,l}^{\vartheta+\frac{3}{2}}}{\left|\sqrt{2}\cos\left(j_{\vartheta,l} - \frac{(\vartheta+1)\pi}{2} - \frac{\pi}{4}\right) + O\big(j_{\vartheta,l}^{-1}\big)\right|}\right)\\ &\leq \frac{1}{x^2 2^\vartheta \Gamma(\vartheta+1)}\left(\sum_{l=1}^{N-1} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \frac{j_{\vartheta,l}^{\vartheta+1}}{|J_{\vartheta+1}(j_{\vartheta,l})|} + \sqrt{\pi}\sum_{l=N}^{\infty} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} j_{\vartheta,l}^{\vartheta+\frac{3}{2}}\right) \overset{(4.6)}{<} \infty\end{aligned}$$

and ultimately proved the first step

$$\lim_{y\to 0} f_{\tau_{y,x}f}(t) = \sum_{l=1}^{\infty} e^{-t\frac{j_{\vartheta,l}^2}{2x^2}} \frac{j_{\vartheta,l}^{\vartheta+1}}{2^\vartheta x^2 \Gamma(\vartheta+1) J_{\vartheta+1}(j_{\vartheta,l})}.$$

Step 2: Obviously, $\left\{\tau_{\frac{1}{n+1},x} \leq t\right\} \subset \left\{\tau_{\frac{1}{n},x} \leq t\right\}$ for all $t, x > 0$ follows from the definition of the stopping time and the Markov property, so with the continuity from above of probability measures we receive

$$\begin{aligned}\mathbb{P}(\tau_{0,x} < t) = \mathbb{P}\Big(\bigcap_{n\in\mathbb{N}} \left\{\tau_{\frac{1}{n},x} \leq t\right\}\Big) &= \lim_{n\to\infty} \mathbb{P}\Big(\tau_{\frac{1}{n},x} \leq t\Big) \\ &= \lim_{n\to\infty} \int_0^t f_{\tau_{\frac{1}{n},x}}(s)\,\mathrm{d}s = \int_0^t \lim_{n\to\infty} f_{\tau_{\frac{1}{n},x}}(s)\,\mathrm{d}s \\ &= \int_0^t f_{\tau_{0,x}}(s)\,\mathrm{d}s\end{aligned}$$

that $f_{\tau_{0,x}}$ is the density of $\tau_{0,x}$ since the right-hand side is everywhere continuous. We can exchange the limit and integral by using the same bound for $\left|f_{\tau_{\frac{1}{n},x}}\right|$ as in the first step. $\square$

In this section we dealt with hitting times and their densities. In the following section, we want to analyze the times when we hit the origin in more detail.

4.2 Existing results on return times of a classical Bessel process

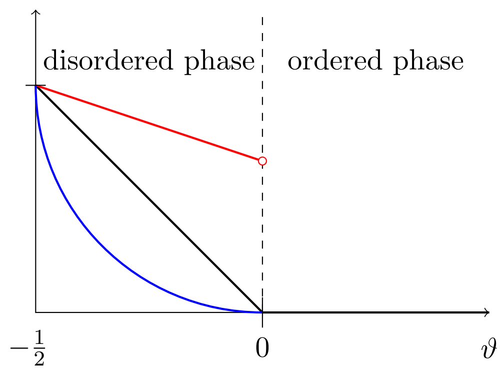

Figure 4.1: Phase transition

In physics, it is of interest to consider phase transitions, which occur when a system changes its behaviour noticeably at particular parameter values. The change in behaviour is usually represented by a so-called "order parameter", a quantity that changes from zero to non-zero at the point of transition. Here, $\vartheta + \frac{1}{2}$ represents the inverse temperature and we have seen that as long as $\vartheta > 0$ the Bessel process almost surely never hits the origin, while such collisions occur when $\vartheta < 0$. This observation suggests that the order parameter could be related to the occurence of collisions and that the transition distinguishes between a non-colliding (or ordered) phase at low temperature and a colliding (or disordered) phase at high temperature. This leads to the decision to consider ϑ as an order parameter. There are several types of phase transitions depending on the behaviour of the order parameter at the critical point. In Figure 4.1 we present three plausible classifications for this transition: discontinuous (red), continuous but not differentiable (black) and differentiable (blue).

Here, we still focus on a classical Bessel process and its return times to the origin. On the one hand, this set of return times has Lebesgue measure zero almost surely, but on the other hand, its cardinality is infinite almost surely. Of course, there are famous sets fulfilling both these properties, for example the Cantor set, but we want to measure this. This is why Luqin Liu and Yimin Xiao [61] considered the fractal Hausdorff dimension, defined below, for the times when self similar processes reach the origin. In particular, they cover a classical Bessel process hitting the origin. We will now present these calculations and proofs pertaining to the classical Bessel process. We reproduce the proof from Luqin Liu and Yimin Xiao but elaborate it in significantly more detail.

We will begin with a short introduction to the Hausdorff dimension and its properties, which we later apply to the Bessel process. We denote by $B(x, R) := \{y \in \mathbb{R}^d : \|y - x\| \leq R\}$ the closed d-dimensional ball centered at x with radius R.

Definition 4.3: For the monotonically increasing (on $[0, \infty)$) monomial of power $\alpha \geq 0$ the Hausdorff measure of $E \subset \mathbb{R}^d$ is defined as

$$m_\alpha(E) := \lim_{\varepsilon \to 0} \inf \left\{ \sum_{i=1}^{\infty} (2r_i)^\alpha : E \subset \bigcup_{i=1}^{\infty} B(x_i, r_i) \text{ for some } x_i \in \mathbb{R}^d \text{ and } 0 \leq r_i < \varepsilon \right\}$$

for $E \subset \mathbb{R}^d$.The radius r_i may be equal to zero and hence the covering can be a finite union.

The Hausdorff dimension is specified by the following lemma, [2, 8.1 Hausdorff dimension].

Lemma 4.4: For any set $E \subset \mathbb{R}^d$ there exists a unique number $\alpha^\star$, called the *Hausdorff dimension* of E, for which

$$\alpha < \alpha^\star \Rightarrow m_\alpha(E) = \infty, \quad \alpha > \alpha^\star \Rightarrow m_\alpha(E) = 0.$$

This number is denoted by $\dim(E)$ and satisfies

$$\alpha^\star = \dim(E) = \sup\{\alpha > 0 : m_\alpha(E) = \infty\} = \inf\{\alpha > 0 : m_\alpha(E) = 0\}.$$

In particular, if $m_\alpha(E)$ is finite we can conclude $\dim(E) \leq \alpha$. The Hausdorff dimension generalizes the well-known dimension concept. This means that familiar geometric objects such as straight lines and hyperplanes keep the same dimension. The Hausdorff dimension offers a finer distinction, since it admits not only natural numbers. The two main properties we are frequently using can be found in [33, 2.2 Hausdorff dimension].

Lemma 4.5: *Countable stability:* If we consider a sequence of sets $(F_i)_{i\in\mathbb{N}} \subset \mathbb{R}^d$ then the Hausdorff dimension of the union is

$$\dim\Big(\bigcup_{i=1}^{\infty} F_i\Big) = \sup_{i\in\mathbb{N}} \dim(F_i).$$

Monotonicity: If $E \subset F \subset \mathbb{R}^d$ holds then $\dim(E) \leq \dim(F)$ ensues.

Especially, we conclude $\dim(E) \in [0, d]$ for any set $E \subset \mathbb{R}^d$ from the monotonicity property. An often used tool to find a lower bound is the following capacity argument, [53, p. 133].

Lemma 4.6: For a compact set $E \subset \mathbb{R}^d$, we suppose there exists a positive measure μ concentrated on E , in formulas $\mu(\mathbb{R}^d) = \mu(E)$, and some $0 < \alpha < d$ such that the energy integrals

$$\|\mu\|^2_{\beta,E} := \int_E \int_E \frac{\mu(\,\mathrm{d}x)\mu(\,\mathrm{d}y)}{\|x-y\|^\beta}$$

are finite for all $0 < \beta < \alpha$, then $\dim(A) \geq \alpha$ is valid.

In the following, we give an extension of a lemma proved in [79]. This is a tool to find a measure with finite β-energy such that we can easily apply the capacity argument. The extension given here covers the case $t_1 > 0$ where the Lemma proven in [79] demands $t_1 = 0$. Therefore, we define $\mathcal{M}_\beta^+$ as the space of all non-negative measures on $[0,\infty)$ with finite β-energy. Due to [2, p. 207], $\big(\mathcal{M}_\beta^+, \|\cdot\|_{\beta,[0,\infty)}\big)$ is a complete metric space for every $\beta > 0$. Here, $\|\cdot\|_{\beta,[0,\infty)}$ defines a norm and we use the induced metric.

Lemma 4.7: We consider $(\mu_n)_{n\in\mathbb{N}}$ to be a sequence of random measures on $[0,\infty)$ concentrated on the compact set $[t_1,t_2] \subset [0,\infty)$.[7] If there exist finite constants $K_1, K_2 > 0$ such that

$$\mathbb{E}^x\big(\|\mu_n\|\big) \geq K_1, \quad \mathbb{E}^x\big(\|\mu_n\|^2\big) \leq K_2, \quad \mathbb{E}^x\big(\|\mu_n\|^2_{\beta,[t_1,t_2]}\big) < \infty$$

hold for $\beta > 0$ and $\|\mu_n\| := \mu_n\big([t_1,t_2],\cdot\big)$, then there exists a subsequence $\big(\mu_{n_k}\big)_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \mu_{n_k} = \mu$$

weakly in $\mathcal{M}_\beta^+$ and μ is strictly positive with $\mathbb{P}^x-$probability at least $\frac{K_1^2}{2K_2}$.

Proof: We use the Paley-Zygmund inequality [89, p. 216 Lemma 8.26]: If X is a non-negative random variable with finite second moment then

$$\mathbb{P}^x\big(X > \theta\,\mathbb{E}\,(X)\big) \geq \big(1-\theta^2\big)\frac{\big(\mathbb{E}\,(X)\big)^2}{\mathbb{E}\,(X^2)}$$

holds for every $\theta \in [0,1]$. Setting $\theta = 0$ we observe

$$\mathbb{P}^x\big(\|\mu_n\| > 0\big) \geq \frac{K_1^2}{K_2} > 0.$$

By choosing a constant M large enough and a small $\varepsilon > 0$, we immediately receive

$$\mathbb{P}^x\big(\varepsilon < \|\mu_n\| \leq M\big) \geq \frac{K_1^2}{2K_2} > 0$$

which implies

$$\begin{aligned}\mathbb{P}^x(\varepsilon < \|\mu_n\| \leq M \text{ infinitely often}) &= \mathbb{P}^x\Big(\bigcap_{i=1}^{\infty}\bigcup_{i=n}^{\infty}\{\varepsilon < \|\mu_n\| \leq M\}\Big)\\ &= \lim_{n\to\infty}\mathbb{P}^x\Big(\bigcup_{i=n}^{\infty}\{\varepsilon < \|\mu_n\| \leq M\}\Big)\end{aligned}$$

[7] This means $\mu_n(\cdot,\omega)$ is a measure on $[0,\infty)$ for every $\omega \in \Omega$ with $\mu_n([0,\infty),\omega) = \mu_n([t_1,t_2],\omega)$.

$$\geq \mathbb{P}^x\Big(\{\varepsilon < \|\mu_n\| \leq M\}\Big) \geq \frac{K_1^2}{2K_2}.$$

On the event $\Omega_0 := \{\omega : \varepsilon < \|\mu_n\| \leq M \text{ infinitely often}\}$ there exists a sequence of measures $(\mu_{n_k})_{k\in\mathbb{N}}$ which is bounded and hence tight due to the compact set $[t_1, t_2]$. Therefore, by Prohorov's theorem [12, Chapter 1.5. Prohorov's theorem], there is a further subsequence $\big(\mu_{n_{k_l}}\big)_{l\in\mathbb{N}}$ that converges weakly to a measure μ with

$$\mathbb{P}^x(\|\mu\| > 0) \geq \frac{K_1^2}{2K_2}.$$

The $\varepsilon > 0$ ensures the positivity of μ. Next, we show that μ has finite β-energy. For any fixed $m > 0$ we have

$$\int_{t_1}^{t_2}\int_{t_1}^{t_2} \min\left\{m, \frac{1}{|x-y|^\beta}\right\}\mu_{n_{k_l}}(\,\mathrm{d}x,\cdot)\mu_{n_{k_l}}(\,\mathrm{d}y,\cdot) \to \int_{t_1}^{t_2}\int_{t_1}^{t_2} \min\left\{m, \frac{1}{|x-y|^\beta}\right\}\mu(\,\mathrm{d}x,\cdot)\mu(\,\mathrm{d}y,\cdot)$$

$\mathbb{P}^x$-almost surely as $l \to \infty$ due to the weak convergence. So for all $\omega \in \Omega_0$ we have

$$\int_{t_1}^{t_2}\int_{t_1}^{t_2} \min\left\{m, \frac{1}{|x-y|^\beta}\right\}\mu(\,\mathrm{d}x,\omega)\mu(\,\mathrm{d}y,\omega) \leq C(\omega) < \infty$$

since $\mathbb{E}^x\big(\|\mu_n\|^2_{\beta,[t_1,t_2]}\big) < \infty$. We let $m \nearrow \infty$ and conclude

$$\int_{t_1}^{t_2}\int_{t_1}^{t_2} \frac{1}{|x-y|^\beta}\mu(\,\mathrm{d}x)\mu(\,\mathrm{d}y) < \infty$$

by the monotone convergence theorem which implies $\mu(\cdot,\omega) \in \mathcal{M}_\beta^+$ for every $\omega \in \Omega_0$. □

In the main proof for the classical Bessel process this lemma would be enough for any finite, closed interval starting in zero. However, this extension will be important in the multivariate case. The proof in [79] covers only $t_1 = 0$.

Returning to the classical Bessel process $(X_t)_{t\geq 0}$ with index $-\frac{1}{2} < \vartheta < 0$, we require some properties for its distribution

$$Q_\vartheta(t,x,A) = \frac{2}{(2t)^\vartheta\Gamma(\vartheta+1)} \int_A j_\vartheta\left(\frac{ixy}{t}\right) e^{-\frac{x^2+y^2}{t}} y^{2\vartheta+1}\mathbb{1}_{(0,\infty)}\,\mathrm{d}y.$$

In this case the Bessel process will almost surely hit the origin in finite time, cf. [56, Theorem 1.1 (iv)]. First of all, we need the $\frac{1}{2}$-self similarity of $(X_t)_{t\geq 0}$, which means for every $a > 0$,

$$Q_\vartheta(t, x, A) = Q_\vartheta\big(at, a^{\frac{1}{2}}x, a^{\frac{1}{2}}A\big) \tag{4.7}$$

is true with $t > 0, x \geq 0$ and $A \in \mathcal{B}([0, \infty))$. This can easily be derived from the form of the density.

For a closer examination of the Hausdorff dimension we derive several inequalities. The following one is a requirement in Luqin Liu and Yimin Xiao [61]. They derive the Hausdorff dimension of $S^{-1}(0) = \{t \geq 0 \,|\, S_t = 0\}$ for self-similar Markov processes $(S_t)_{t\geq 0}$ satisfying Lemma 4.8 with a corresponding power of r. In particular, they mention strictly stable Lévy processes and the classical Bessel process obey this property. We additionally verify this condition pertaining to the classical Bessel process.

Lemma 4.8: There exist some $C_1 = C_1(\vartheta) > 0$ and $C_2 = C_2(\vartheta) > 0$ such that for every $r \geq 0$ and $0 \leq x \leq r$ the inequality

$$C_1 \min\left\{1, r^{2\vartheta+2}\right\} \leq Q_\vartheta\big(1, x, B(0, r)\big) \leq C_2 \min\left\{1, r^{2\vartheta+2}\right\}$$

is satisfied.

Proof: In order to evaluate

$$Q_\vartheta\big(1, x, B(0, r)\big) = \frac{2}{2^\vartheta \Gamma(\vartheta + 1)} \int_0^r j_\vartheta(ixy) e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1} \,\mathrm{d}y$$

we have a closer look at the spherical Bessel function

$$j_\vartheta(ixy) = \frac{\Gamma(\vartheta + 1)}{\Gamma(\vartheta + \frac{1}{2})\Gamma(\frac{1}{2})} \int_{-1}^1 e^{-sxy}\big(1 - s^2\big)^{\vartheta - \frac{1}{2}} \,\mathrm{d}s.$$

The inequalities,

$$\left.\begin{array}{ll} 1, & s \leq 0, \\ e^{-xy}, & s \geq 0. \end{array}\right\} \leq e^{-sxy} \leq \left\{\begin{array}{ll} e^{xy}, & s \leq 0, \\ 1, & s \geq 0. \end{array}\right.$$

for $x, y \in [0, \infty)$ imply

$$\frac{1}{2}\big(e^{-xy}+1\big) \leq j_\vartheta(ixy) \leq \frac{1}{2}\big(e^{xy}+1\big). \tag{4.8}$$

Upper bound: By (4.8) we simplify the probability and derive

$$\begin{aligned}
Q_\vartheta\big(1, x, B(0,r)\big) &\leq \frac{1}{2^{\vartheta-1}\Gamma(\vartheta+1)} \int_0^r \frac{1}{2}\big(e^{xy}+1\big)e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y \\
&= \frac{1}{2^{\vartheta-1}\Gamma(\vartheta+1)} \int_0^r \frac{1}{2}\Big(e^{-\frac{(x-y)^2}{2}} + e^{-\frac{x^2+y^2}{2}}\Big) y^{2\vartheta+1}\,\mathrm{d}y \\
&\leq \frac{1}{2^{\vartheta-1}\Gamma(\vartheta+1)} \int_0^r y^{2\vartheta+1}\,\mathrm{d}y \\
&= \frac{1}{2^{\vartheta-1}\Gamma(\vartheta+1)(2\vartheta+2)} r^{2\vartheta+2}.
\end{aligned}$$

We define $C_2(\vartheta) := \max\Big\{\frac{1}{2^{\vartheta-1}\Gamma(\vartheta+1)(2\vartheta+2)}, 1\Big\}$. If $r < 1$, our calculations above show the upper bound. If $r \geq 1$, we deduce

$$Q_\vartheta\big(1, x, B(0,r)\big) \leq 1 \leq C_2(\vartheta) = C_2(\vartheta) \min\big\{1, r^{2\vartheta+2}\big\}.$$

Lower bound: We first consider $r < 1$ and deduce

$$\begin{aligned}
Q_\vartheta\big(1, x, B(0,r)\big) &\overset{(4.8)}{\geq} \frac{1}{2^{\vartheta}\Gamma(\vartheta+1)} \int_0^r \big(e^{-xy}+1\big)e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y \\
&\geq \frac{1}{2^{\vartheta}\Gamma(\vartheta+1)} \int_0^r e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y \\
&\geq \frac{e^{-r^2}}{2^{\vartheta}\Gamma(\vartheta+1)} \int_0^r y^{2\vartheta+1}\,\mathrm{d}y \\
&\geq \frac{e^{-1}}{2^{\vartheta}\Gamma(\vartheta+1)} \int_0^r y^{2\vartheta+1}\,\mathrm{d}y \\
&= \frac{e^{-1}}{2^{\vartheta}\Gamma(\vartheta+1)} \cdot \frac{r^{2\vartheta+2}}{2\vartheta+2} \\
&= \frac{e^{-1}}{2^{\vartheta+1}\Gamma(\vartheta+2)} r^{2\vartheta+2}.
\end{aligned}$$

The third line holds as x and y are bounded by r, which itself is less than one justifying the fourth line. In the case $r > 1$, we assume $x > 1$. Otherwise, the inequality is trivial and works similar as in the case $r < 1$,

$$\begin{aligned}
Q_\vartheta\big(1, x, B(0,r)\big) &\geq \frac{1}{2^\vartheta \Gamma(\vartheta+1)} \int\limits_0^r e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y \\
&\geq \frac{1}{2^\vartheta \Gamma(\vartheta+1)} \int\limits_0^1 e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y \\
&\geq \frac{e^{-1}}{2^\vartheta \Gamma(\vartheta+1)} \int\limits_0^1 y^{2\vartheta+1}\,\mathrm{d}y \\
&= \frac{e^{-1}}{2^{\vartheta+1} \Gamma(\vartheta+2)}.
\end{aligned}$$

For $x > 1$ we attempt to get rid of the dependence on x and r by other means. For this purpose, we reduce the event whose probability we determine

$$\begin{aligned}
Q_\vartheta\big(1, x, B(0,r)\big) &\geq Q_\vartheta\big(1, x, B(0,x)\big) \\
&= \frac{1}{2^{\vartheta-1} \Gamma(\vartheta+1)} \int\limits_0^x j_\vartheta(ixy) e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y \\
&\geq \frac{1}{2^{\vartheta-1} \Gamma(\vartheta+1)} \int\limits_{x-1}^x j_\vartheta(ixy) e^{-\frac{x^2+y^2}{2}} y^{2\vartheta+1}\,\mathrm{d}y.
\end{aligned}$$

In the first line, we use $x \leq r$ and hence $B(0,r) \supset B(0,x)$, whereas $x > 1$ justifies the last line. We next attempt to find a suitable lower bound for the spherical Bessel function and rewrite it accordingly,

$$\begin{aligned}
j_\vartheta(ixy) &= \frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int\limits_{-1}^1 e^{-sxy}(1-s^2)^{\vartheta-\frac{1}{2}}\,\mathrm{d}s \\
&= e^{xy} \frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int\limits_{-1}^1 e^{-(s+1)xy}(1-s^2)^{\vartheta-\frac{1}{2}}\,\mathrm{d}s.
\end{aligned}$$

The substitution $t = (s+1)xy$ simplifies the integral

$$\begin{aligned}
j_\vartheta(ixy) &= e^{xy}\frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int\limits_0^{2xy} e^{-t}\left(1-\left(\frac{t}{xy}-1\right)^2\right)^{\vartheta-\frac{1}{2}}\mathrm{d}t\\
&= \frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\cdot\frac{e^{xy}}{xy}\int\limits_0^{2xy} e^{-t}\left(\not{1}-\left(\frac{t^2}{(xy)^2}-\frac{2t}{xy}+\not{1}\right)\right)^{\vartheta-\frac{1}{2}}\mathrm{d}t\\
&= \frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\cdot\frac{e^{xy}2^{\vartheta-\frac{1}{2}}}{(xy)^{\vartheta+\frac{1}{2}}}\int\limits_0^{2xy} e^{-t}t^{\vartheta-\frac{1}{2}}\left(1-\frac{t}{2xy}\right)^{\vartheta-\frac{1}{2}}\mathrm{d}t.
\end{aligned}$$

Using the binomial theorem and that the coefficient

$$\left(\vartheta-\frac{1}{2}\right)\cdots\left(\vartheta-\frac{1}{2}-m+1\right)(-1)^m = \left(\frac{1}{2}-\vartheta\right)\cdots\left(\frac{1}{2}-\vartheta+m-1\right) = \left(\frac{1}{2}-\vartheta\right)_m$$

is positive for every $m \in \mathbb{N}$ since $\vartheta < 0$, we obtain

$$\begin{aligned}
j_\vartheta(ixy) &= \frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\cdot\frac{e^{xy}2^{\vartheta-\frac{1}{2}}}{(xy)^{\vartheta+\frac{1}{2}}}\sum_{m=0}^{\infty}\int\limits_0^{2xy} e^{-t}t^{\vartheta-\frac{1}{2}}\frac{\left(\frac{1}{2}-\vartheta\right)_m}{m!}\left(\frac{t}{2xy}\right)^m\mathrm{d}t\\
&\geq \frac{\Gamma(\vartheta+1)}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\cdot\frac{e^{xy}2^{\vartheta-\frac{1}{2}}}{(xy)^{\vartheta+\frac{1}{2}}}\int\limits_0^{2xy} e^{-t}t^{\vartheta-\frac{1}{2}}\,\mathrm{d}t.
\end{aligned}$$

Combining these calculations, we derive

$$\begin{aligned}
Q_\vartheta\big(1,x,B(0,r)\big) &\geq \frac{1}{2^{\vartheta-1}\Gamma(\vartheta+1)}\int\limits_{x-1}^{x} j_\vartheta(ixy)e^{-\frac{x^2+y^2}{2}}y^{2\vartheta+1}\,\mathrm{d}y\\
&\geq \frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int\limits_{x-1}^{x} e^{-\frac{x^2+y^2}{2}}y^{2\vartheta+1}\frac{e^{xy}}{(xy)^{\vartheta+\frac{1}{2}}}\int\limits_0^{2xy} e^{-t}t^{\vartheta-\frac{1}{2}}\,\mathrm{d}t\,\mathrm{d}y\\
&\geq \frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int\limits_{x-1}^{x} e^{-\frac{(x-y)^2}{2}}\left(\frac{y}{x}\right)^{\vartheta+\frac{1}{2}}\mathrm{d}y\int\limits_0^{2x(x-1)} e^{-t}t^{\vartheta-\frac{1}{2}}\,\mathrm{d}t\\
&\geq \frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}e^{-\frac{1}{2}}\left(\frac{x-1}{x}\right)^{\vartheta+\frac{1}{2}}\int\limits_0^{2x(x-1)} e^{-t}t^{\vartheta-\frac{1}{2}}\,\mathrm{d}t
\end{aligned}$$

$$= \frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)} e^{-\frac{1}{2}} \left(1-\frac{1}{x}\right)^{\vartheta+\frac{1}{2}} \int\limits_{0}^{2x(x-1)} e^{-t} t^{\vartheta-\frac{1}{2}} \, \mathrm{d}t.$$

The third line is valid due to $y \geq x-1$. In the fourth line we used additionally $x-y \in [0,1]$ and $\vartheta + \frac{1}{2} > 0$. Now, we argue the function on the right-hand side has a positive lower bound that depends only on the parameter ϑ. Therefore, we regard

$$\begin{aligned} \lim_{x\to\infty} g(x) &:= \lim_{x\to\infty} \left(\frac{1}{\Gamma\left(\vartheta+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)} e^{-\frac{1}{2}} \left(1-\frac{1}{x}\right)^{\vartheta+\frac{1}{2}} \int\limits_{0}^{2x(x-1)} e^{-t} t^{\vartheta-\frac{1}{2}} \, \mathrm{d}t \right) \\ &= \frac{e^{-\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} = \frac{e^{-\frac{1}{2}}}{\sqrt{\pi}} > 0. \end{aligned}$$

Thus, we achieve

$$\lim_{x\to\infty} g(x) = C > 0,$$

which means equivalently for every $\varepsilon > 0$ there exists an $\tilde{x} > 0$ such that for every $x \geq \tilde{x}$

$$|g(x) - C| < \varepsilon$$

is valid. We choose $\varepsilon = \frac{C}{2}$ to receive

$$-\frac{C}{2} < g(x) - C < \frac{C}{2}$$

and especially infer

$$g(x) > \frac{C}{2} > 0$$

for all $x \in [\tilde{x}, \infty)$. Furthermore, g is continuous and hence takes its positive minimum on the compact interval $[0, \tilde{x}]$, where the function is positive. This means in particular we established

$$\begin{aligned} Q_\vartheta\big(1, x, B(0,r)\big) &\geq g(x) \\ &\geq \min\left\{ \frac{C}{2}, \min_{x\in[0,\tilde{x}]} g(x) \right\}. \end{aligned}$$

The assertion ensues by setting

$$\begin{aligned} C_1(\vartheta) &:= \min\Big\{\frac{e^{-1}}{2^{\vartheta+1}\Gamma(\vartheta+2)}, \frac{C}{2}, \min_{x\in[0,\tilde{x}]} g(x)\Big\} \\ &= \min\Big\{\frac{e^{-1}}{2^{\vartheta+1}\Gamma(\vartheta+2)}, \frac{e^{-\frac{1}{2}}}{2\sqrt{\pi}}, \min_{x\in[0,\tilde{x}]} g(x)\Big\}. \end{aligned}$$

□

Next, we provide a bound for the probability that, if we start in the origin, we will be back near the origin at some point within the time interval $[t_1, t_2]$. Using the strong Markov property of the Bessel process, we proceed to first obtain the following result, cf. [87, Proposition 2.1].

Lemma 4.9: For every $t_2 > t_1 > 0, x \geq 0$ and $r > 0$ we have

$$\mathbb{P}^x\big(\exists s \in [t_1, t_2] : |X_s| \leq r\big) \leq \frac{\int_{t_1}^{2t_2-t_1} Q_\vartheta\big(s, x, B(0,r)\big)\,\mathrm{d}s}{\inf_{|y|\leq r} \int_0^{t_2-t_1} Q_\vartheta\big(s, y, B(0,r)\big)\,\mathrm{d}s}.$$

Proof: According to [87, Proposition 2.1] we define the stopping time

$$T := \inf\big\{s \geq t_1 : |X_s| \leq r\big\}$$

to reformulate the probability on the left-hand side

$$\mathbb{P}^x(T \leq t_2) = \mathbb{P}^x\big(\exists s \in [t_1, t_2] : |X_s| \leq r\big).$$

We can also rewrite the term by using Fubini's theorem, which always holds for integrals over positive functions,

$$\begin{aligned} \int_{t_1}^{2t_2-t_1} Q_\vartheta\big(s, x, B(0,r)\big)\,\mathrm{d}s &= \int_{t_1}^{2t_2-t_1} \mathbb{P}^x\big(X_s \in B(0,r)\big)\,\mathrm{d}s \\ &= \int_{t_1}^{2t_2-t_1} \mathbb{E}^x\big(\mathbb{1}_{\{|X_s|\leq r\}}\big)\,\mathrm{d}s \end{aligned}$$

$$= \mathbb{E}^x\left(\int\limits_{t_1}^{2t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right)$$

to consequently reduce the proof to verifying

$$\mathbb{E}^x\left(\int\limits_{t_1}^{2t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right) \ge \mathbb{P}^x(T \le t_2) \inf_{|y|\le r} \int\limits_0^{t_2-t_1} Q_\vartheta\big(s, y, B(0,r)\big) \,\mathrm{d}s.$$

Based on the definition of the stopping time we recognize $T \ge t_1$. From this, we derive

$$\begin{aligned}
\mathbb{E}^x\left(\int\limits_{t_1}^{2t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right) &\ge \mathbb{E}^x\left(\int\limits_{T}^{2t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right) \\
&\ge \mathbb{E}^x\left(\mathbb{1}_{\{T\le t_2\}} \int\limits_{T}^{2t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right) \\
&= \mathbb{E}^x\left(\mathbb{1}_{\{T\le t_2\}} \,\mathbb{E}^{X_T}\left[\int\limits_{T}^{2t_2-t_1} \mathbb{1}_{\{|X_{s-T}|\le r\}} \,\mathrm{d}s \right] \right) \\
&= \mathbb{E}^x\left(\mathbb{1}_{\{T\le t_2\}} \,\mathbb{E}^{X_T}\left[\int\limits_{0}^{2t_2-t_1-T} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right] \right).
\end{aligned}$$

The second-last equation is valid because of the strong Markov property. Finally, we use the bound $T \le t_2$ and Fubini's theorem:

$$\begin{aligned}
\mathbb{E}^x\left(\int\limits_{t_1}^{2t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right) &\ge \mathbb{E}^x\left(\mathbb{1}_{\{T\le t_2\}} \,\mathbb{E}^{X_T}\left[\int\limits_{0}^{2t_2-t_1-T} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right] \right) \\
&\ge \mathbb{E}^x\left(\mathbb{1}_{\{T\le t_2\}} \,\mathbb{E}^{X_T}\left[\int\limits_{0}^{2t_2-t_1-t_2} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right] \right) \\
&\ge \mathbb{E}^x\left(\mathbb{1}_{\{T\le t_2\}}\right) \inf_{|y|\le r} \mathbb{E}^y\left(\int\limits_0^{t_2-t_1} \mathbb{1}_{\{|X_s|\le r\}} \,\mathrm{d}s \right) \\
&= \mathbb{P}^x(T \le t_2) \inf_{|y|\le r} \int\limits_0^{t_2-t_1} \mathbb{E}^y\left(\mathbb{1}_{\{|X_s|\le r\}}\right) \mathrm{d}s
\end{aligned}$$

$$= \mathbb{P}^x(T \le t_2) \inf_{|y|\le r} \int_0^{t_2-t_1} Q_\vartheta\big(s, y, B(0,r)\big)\, \mathrm{d}s.$$

□

Now, we combine Lemma 4.8 and Lemma 4.9 to receive the following bound, which will later give us the maximal possible Hausdorff dimension. This is only one part from [61, Lemma 4.2], but we specify further what values r can take instead of saying r has to be small.

Lemma 4.10: For every $\varepsilon > 0$ there exists a constant $C_3 = C_3(\vartheta, \varepsilon) > 0$ such that

$$\mathbb{P}^0\big(\exists s \in [t_1, t_2] : |X_s| \le r\big) \le C_3 (t_2 - t_1)^\vartheta$$

is valid for every $t_2 > t_1 \ge \varepsilon$ and $0 < r \le \sqrt{t_2 - t_1}$.

Proof: We use the inequality of Lemma 4.9 to search for an upper bound:

$$\mathbb{P}^0\big(\exists s \in [t_1, t_2] : |X_s| \le r\big) \le \frac{\int_{t_1}^{2t_2-t_1} Q_\vartheta\big(s, 0, B(0,r)\big)\, \mathrm{d}s}{\inf_{|y|\le r} \int_0^{t_2-t_1} Q_\vartheta\big(s, y, B(0,r)\big)\, \mathrm{d}s}.$$

The denominator easily simplifies to

$$\begin{aligned}
\int_0^{t_2-t_1} Q_\vartheta\big(s, y, B(0,r)\big)\, \mathrm{d}s &\overset{(4.7)}{=} \int_0^{t_2-t_1} Q_\vartheta\left(1, \frac{y}{\sqrt{s}}, B\left(0, \frac{r}{\sqrt{s}}\right)\right) \mathrm{d}s \\
&\overset{4.8}{\ge} C_1(\vartheta) \int_0^{t_2-t_1} \min\left\{1, \left(\frac{r}{\sqrt{s}}\right)^{2\vartheta+2}\right\} \mathrm{d}s \\
&\ge C_1(\vartheta) \int_0^{t_2-t_1} \min\left\{1, \left(\frac{r^2}{t_2-t_1}\right)^{\vartheta+1}\right\} \mathrm{d}s \\
&= C_1(\vartheta) \frac{r^{2\vartheta+2}}{(t_2-t_1)^{\vartheta+1}} (t_2 - t_1) \\
&= C_1(\vartheta) r^{2\vartheta+2} (t_2 - t_1)^{-\vartheta},
\end{aligned}$$

which indeed is independent of y. The second-last line holds since $r^2 \leq t_2 - t_1$. For the numerator we start with the same calculations

$$\begin{aligned}
\int_{t_1}^{2t_2-t_1} Q_\vartheta\big(s,0,B(0,r)\big)\,\mathrm{d}s &\overset{(4.7)}{=} \int_{t_1}^{2t_2-t_1} Q_\vartheta\left(1,0,B\left(0,\frac{r}{\sqrt{s}}\right)\right)\mathrm{d}s \\
&\overset{4.8}{\leq} C_2(\vartheta)\int_{t_1}^{2t_2-t_1} \min\left\{1,\frac{r^{2\vartheta+2}}{s^{\vartheta+1}}\right\}\mathrm{d}s \\
&\leq C_2(\vartheta) r^{2\vartheta+2}\int_{t_1}^{2t_2-t_1} s^{-\vartheta-1}\,\mathrm{d}s \\
&= \frac{C_2(\vartheta)}{-\vartheta} r^{2\vartheta+2}\left[(2t_2-t_1)^{-\vartheta} - t_1^{-\vartheta}\right].
\end{aligned}$$

By using Bernoulli's inequality [67, 2.4 Bernoulli's Inequality and its Generalizations], we conclude

$$\begin{aligned}
(2t_2-t_1)^{-\vartheta} - t_1^{-\vartheta} &= \big(2(t_2-t_1)+t_1\big)^{-\vartheta} - t_1^{-\vartheta} = t_1^{-\vartheta}\left(\frac{2(t_2-t_1)}{t_1}+1\right)^{-\vartheta} - t_1^{-\vartheta} \\
&\leq t_1^{-\vartheta}\left(1+(-2\vartheta)\frac{t_2-t_1}{t_1}\right) - t_1^{-\vartheta} = -2\vartheta(t_2-t_1)t_1^{-\vartheta-1} \\
&\leq \frac{-2\vartheta}{\varepsilon^{\vartheta+1}}(t_2-t_1).
\end{aligned}$$

Combining all calculations results in the desired bound

$$\begin{aligned}
\mathbb{P}^0(\exists s\in[t_1,t_2] : |X_s|\leq r) &\leq \frac{\int_{t_1}^{2t_2-t_1} Q_\vartheta\big(s,0,B(0,r)\big)\,\mathrm{d}s}{\inf\limits_{|y|\leq r}\int_0^{t_2-t_1} Q_\vartheta\big(s,y,B(0,r)\big)\,\mathrm{d}s} \\
&\leq \frac{\frac{C_2(\vartheta)}{-\vartheta} r^{2\vartheta+2}}{C_1(\vartheta) r^{2\vartheta+2}(t_2-t_1)^{-\vartheta}}\left[(2t_2-t_1)^{-\vartheta} - t_1^{-\vartheta}\right] \\
&= \frac{C_2(\vartheta)}{-\vartheta C_1(\vartheta)(t_2-t_1)^{-\vartheta}}\left[(2t_2-t_1)^{-\vartheta} - t_1^{-\vartheta}\right] \\
&\leq \frac{C_2(\vartheta)}{-\vartheta C_1(\vartheta)(t_2-t_1)^{-\vartheta}}\cdot\frac{-2\vartheta}{\varepsilon^{\vartheta+1}}(t_2-t_1) \\
&= \frac{2C_2(\vartheta)}{\varepsilon^{\vartheta+1}C_1(\vartheta)}(t_2-t_1)^{\vartheta+1}.
\end{aligned}$$

□

Finally, we analyze the set of times $X^{-1}(0) := \{t \geq 0 : X_t = 0\} \subset [0, \infty)$ when the process hits the origin. Obviously, we observe $\dim\left(X^{-1}(0)\right) \in [0,1]$. Using Lemmas 4.8 and 4.10, we proceed to calculate the Hausdorff dimension of this set, cf. [61, Theorem 4.1 and 4.2].

Theorem 4.11: For every $-\frac{1}{2} < \vartheta < 0$ the equality

$$\dim\left(X^{-1}(0)\right) = -\vartheta$$

is $\mathbb{P}^0$- almost surely valid.

Proof: Upper bound: To begin with, we intend to prove the upper bound. Therefore, we prove

$$\dim\left(X^{-1}(0) \cap [t_1, t_2]\right) \leq -\vartheta$$

for every interval $[t_1, t_2] \subset (0, \infty)$ $\mathbb{P}^0$-almost surely. If we recall the definition of the Hausdorff dimension, we observe that the balls $B(x, r)$ in the covering are intervals in the one-dimensional case whose length converges to zero. Thus, for $n \in \mathbb{N}$ we divide $[t_1, t_2]$ into n subintervals $I_{n,i} := \left[t_1 + (i-1)\frac{t_2-t_1}{n}, t_1 + i\frac{t_2-t_1}{n}\right]$. Thereby, it is essential that the length of $I_{n,i}$ converges to zero for $n \to \infty$ and $I_{n,1} \cup \ldots \cup I_{n,n} = [t_1, t_2]$ provides a covering for $X^{-1}(0) \cap [t_1, t_2]$. We conclude

$$\begin{aligned}
\mathbb{E}^0\left(m_{-\vartheta}(X^{-1}(0) \cap [t_1, t_2])\right) &\leq \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{t_2 - t_1}{n}\right)^{-\vartheta} \mathbb{P}^0\left(\exists t \in I_{n,i} : X_t = 0\right) \\
&\leq \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{t_2 - t_1}{n}\right)^{-\vartheta} \mathbb{P}^0\left(\exists t \in I_{n,i} : |X_t| \leq r_n\right) \\
&\overset{4.10}{\leq} C_3(\vartheta, t_1) \lim_{n\to\infty} \left(\frac{t_2 - t_1}{n}\right)^{-\vartheta} \left(\frac{t_2 - t_1}{n}\right)^{\vartheta+1} n \\
&= C_3(\vartheta, t_1)(t_2 - t_1) < \infty
\end{aligned}$$

for some $0 < r_n \leq \sqrt{\frac{t_2-t_1}{n}}$. The first inequality holds since we are considering one particular covering instead of the optimal cover. Consequently, we deduce

$$\dim\left(X^{-1}(0) \cap [t_1, t_2]\right) \leq -\vartheta$$

as $m_{-\vartheta}(E) < \infty$ implies $\dim(E) \leq -\vartheta$ for $E \subset \mathbb{R}^d$. Hence, with $\dim\big(\{0\}\big) = 0$ and the countable stability the upper bound immediately succeeds

$$\begin{aligned}\dim\big(X^{-1}(0)\big) &= \dim\Big(\{0\}\dot\cup\big(X^{-1}(0)\cap(0,\infty)\big)\Big)\\ &= \dim\left(\{0\}\dot\cup\bigcup_{n=1}^{\infty}\left(X^{-1}(0)\cap\left[\frac{1}{n},n\right]\right)\right)\\ &\overset{4.5}{=} \sup_{n\in\mathbb{N}}\dim\left(X^{-1}(0)\cap\left[\frac{1}{n},n\right]\right) \leq -\vartheta.\end{aligned}$$

Lower bound: For the lower bound we initially verify $\dim\big(X^{-1}(0)\cap[0,t]\big) \geq -\vartheta$ for every $t > 0$ with a positive $\mathbb{P}^0-$probability since we need a compact set for the capacity argument, Lemma 4.6. Therefore, we construct for every possible smaller dimension $0 < \beta < -\vartheta$ a positive measure μ concentrated on $X^{-1}(0)\cap[0,t]$ such that $\|\mu\|_{\beta,[0,t]} < \infty$, which implies $\dim\big(X^{-1}(0)\cap[0,t]\big) > \beta$ on $\{\mu > 0\}$ due to the capacity argument. We define a sequence of random positive measures on $\mathcal{B}([0,\infty))$ by

$$\mu_n(B,\omega) := n^{2\vartheta+2}\int\limits_{[0,t]\cap B}\mathbb{1}_{\left\{|X_s(\omega)|\leq\frac{1}{n}\right\}}\,\mathrm{d}s.$$

To apply Lemma 4.7 we search for constants $K_1 = K_1(\vartheta,t) > 0$ and $K_2 = K_2(\vartheta,t) > 0$ satisfying

$$\mathbb{E}^0(\|\mu_n\|) \geq K_1, \quad \mathbb{E}^0(\|\mu_n\|^2) \leq K_2, \quad \mathbb{E}^0(\|\mu_n\|^2_{\beta,[0,t]}) < \infty. \tag{4.9}$$

According to Lemma 4.7, we then receive a suitable μ as a limit of a subsequence. Since the Bessel process $(X_t)_{t\geq0}$ has continuous sample paths, such limit μ would be supported on $X^{-1}(0)\cap[0,t]$, see [64, p. 282]. Thus, our task is to prove (4.9). By using the $\frac{1}{2}-$self similarity of the Bessel process (4.7) and Lemma 4.8 we obtain

$$\begin{aligned}\mathbb{E}^0(\|\mu_n\|) = \mathbb{E}^0\left(n^{2\vartheta+2}\int_0^t\mathbb{1}_{\left\{|X_s|\leq\frac{1}{n}\right\}}\,\mathrm{d}s\right) &= n^{2\vartheta+2}\int_0^t\mathbb{P}^0\left(|X_s|\leq\frac{1}{n}\right)\mathrm{d}s\\ = n^{2\vartheta+2}\int_0^t Q_\vartheta\left(s,0,B\left(0,\frac{1}{n}\right)\right)\mathrm{d}s &\overset{(4.7)}{=} n^{2\vartheta+2}\int_0^t Q_\vartheta\left(1,0,B\left(0,\frac{1}{\sqrt{s}n}\right)\right)\mathrm{d}s\\ \geq n^{2\vartheta+2}\int_{n^{-2}}^t Q_\vartheta\left(1,0,B\left(0,\frac{1}{\sqrt{s}n}\right)\right)\mathrm{d}s &\overset{4.8}{\geq} C_1(\vartheta)n^{2\vartheta+2}n^{-2\vartheta-2}\int_{n^{-2}}^t s^{-\vartheta-1}\,\mathrm{d}s\end{aligned}$$

$$= \frac{C_1(\vartheta)}{-\vartheta}\left[t^{-\vartheta} - n^{2\vartheta}\right] \geq \frac{C_1(\vartheta)}{-\vartheta}\left[t^{-\vartheta} - \big(n_0(t)\big)^{2\vartheta}\right] =: K_1(\vartheta, t) > 0$$

for every $n \geq n_0(t) := \lfloor t^{-1} \rfloor + 1$. By means of analogous calculations, we receive

$$\begin{aligned}
\mathbb{E}^0(\|\mu_n\|^2) &= n^{4\vartheta+4}\int_0^t\int_0^t \mathbb{P}^0\Big(X_{s_1} \leq \frac{1}{n}, X_{s_2} \leq \frac{1}{n}\Big)\,\mathrm{d}s_1\,\mathrm{d}s_2 \\
&= 2n^{4\vartheta+4}\int_0^t\int_0^{s_2}\int_{B(0,n^{-1})} Q_\vartheta\Big(s_2-s_1, u, B\big(0,n^{-1}\big)\Big)Q_\vartheta\big(s_1,0,\,\mathrm{d}u\big)\,\mathrm{d}s_1\,\mathrm{d}s_2 \\
&\overset{(4.7)}{\underset{4.8}{\leq}} 2C_2^2(\vartheta)n^{4\vartheta+4}\int_0^t\int_{s_1}^t \min\big\{1,(s_2-s_1)^{-\vartheta-1}n^{-2\vartheta-2}\big\} \\
&\qquad\qquad \min\big\{1, s_1^{-\vartheta-1}n^{-2\vartheta-2}\big\}\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&\leq 2C_2^2(\vartheta)n^{4\vartheta+4}\int_0^t\int_{s_1}^t (s_2-s_1)^{-\vartheta-1}n^{-2\vartheta-2}s_1^{-\vartheta-1}n^{-2\vartheta-2}\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&= 2C_2^2(\vartheta)\int_0^t \frac{(t-s_1)^{-\vartheta}}{-\vartheta}s_1^{-\vartheta-1}\,\mathrm{d}s_1 \;\leq\; 2C_2^2(\vartheta)\frac{t^{-\vartheta}}{-\vartheta}\cdot\frac{t^{-\vartheta}}{-\vartheta} \\
&= 2C_2^2(\vartheta)\frac{t^{-2\vartheta}}{\vartheta^2} =: K_2(\vartheta,t) < \infty.
\end{aligned}$$

In the second-last line we applied $(t-s_1)^{-\vartheta} \leq t^{-\vartheta}$ since $-\vartheta > 0$. Similarly, we calculate

$$\begin{aligned}
\mathbb{E}^0(\|\mu_n\|^2_{\beta,[0,t]}) &= \mathbb{E}^0\left(\int_0^t\int_0^t \frac{\mu_n(\,\mathrm{d}s_1,\cdot)\mu_n(\,\mathrm{d}s_2,\cdot)}{|s_2-s_1|^\beta}\right) \\
&= 2n^{4\vartheta+4}\int_0^t\int_{s_1}^t \frac{\mathbb{P}^0\big(X_{s_1}\leq\frac{1}{n}, X_{s_2}\leq\frac{1}{n}\big)}{(s_2-s_1)^\beta}\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&= 2n^{4\vartheta+4}\int_0^t\int_{s_1}^t\int_{B(0,\frac{1}{n})} \frac{Q_\vartheta\big(s_2-s_1,u,B(0,\frac{1}{n})\big)Q_\vartheta\big(s_1,0,\,\mathrm{d}u\big)}{(s_2-s_1)^\beta}\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&\overset{(4.7)}{\underset{4.8}{\leq}} 2C_2^2(\vartheta)\int_0^t\int_{s_1}^t (s_2-s_1)^{-\vartheta-\beta-1}s_1^{-\vartheta-1}\,\mathrm{d}s_2\,\mathrm{d}s_1
\end{aligned}$$

$$\begin{aligned}
&= \frac{2C_2^2(\vartheta)}{-\vartheta-\beta}\int_0^t (t-s_1)^{-\vartheta-\beta} s_1^{-\vartheta-1}\,\mathrm{d}s_1 \\
&\le \frac{2C_2^2(\vartheta)t^{-\vartheta-\beta}}{-\vartheta-\beta}\int_0^t s_1^{-\vartheta-1}\,\mathrm{d}s_1 \;=\; \frac{2C_2^2(\vartheta)t^{-\vartheta-\beta}}{\vartheta(\vartheta+\beta)}t^{-\vartheta} < \infty
\end{aligned}$$

for $0<\beta<-\vartheta$. The inequality in the last line is valid since $-\vartheta-\beta>0$. Hence, we have proved

$$\mathbb{P}^0\Big(\dim\big(X^{-1}(0)\cap[0,t]\big)\ge -\vartheta\Big) \ge \frac{K_1^2(\vartheta,t)}{2K_2(\vartheta,t)} \tag{4.10}$$

for every $t>0$. Next, we define a subset for the hitting times of the origin as a sequence of well defined stopping times τ_n as $\tau_0 := 0$ and

$$\tau_n := \inf\big\{t\ge \tau_{n-1}+1 : X_t = 0\big\}. \tag{4.11}$$

By the strong Markov property, $X_{\tau_n}=0$ and (4.10) we deduce

$$\begin{aligned}
\mathbb{P}^0\Big(\dim\big(X^{-1}(0)\cap[0,\tau_{n+1}]\big)\ge -\vartheta \,\Big|\, \mathcal{F}_{\tau_n}\Big) &\ge \mathbb{P}^0\Big(\dim\big(X^{-1}(0)\cap[\tau_n,\tau_{n+1}]\big)\ge -\vartheta \,\Big|\, \mathcal{F}_{\tau_n}\Big)\\
&= \mathbb{P}^{X_{\tau_n}}\Big(\dim\big(X^{-1}(0)\cap[0,\tau_{n+1}-\tau_n]\big)\ge -\vartheta\Big)\\
&\ge \mathbb{P}^0\Big(\dim\big(X^{-1}(0)\cap[0,1]\big)\ge -\vartheta\Big)\\
&\ge \frac{K_1^2(\vartheta,t)}{2K_2(\vartheta,t)}.
\end{aligned}$$

The sequence $A_n := \big\{\dim\big(X^{-1}(0)\cap[0,\tau_n]\big)\ge -\vartheta\big\}\in\mathcal{F}_{\tau_n}$ fulfills

$$\sum_{n\in\mathbb{N}} \mathbb{P}^0(A_n\,|\,\mathcal{F}_{\tau_{n-1}}) \ge \sum_{n\in\mathbb{N}} \frac{K_1^2(\vartheta,t)}{2K_2(\vartheta,t)} = \infty$$

almost surely and hence by the conditional Borel-Cantelli lemma [14, Corollary 13.3.38]

$$\Bigg\{\sum_{n\in\mathbb{N}} \mathbb{P}^0(A_n\,|\,\mathcal{F}_{\tau_{n-1}}) = \infty\Bigg\} \equiv \Bigg\{\sum_{n\in\mathbb{N}} \mathbb{1}_{A_n} = \infty\Bigg\},$$

we conclude that

$$\mathbb{P}^0\Big(\dim\big(X^{-1}(0)\cap[0,\tau_n]\big)\ge -\vartheta \quad \text{for infinitely many } n\Big) = 1.$$

The property $\sup\{t : X_t = 0\} = \infty$ $\mathbb{P}^x-$almost surely, Lemma 4.12, completes the proof due to $[0, \tau_n] \subset [0, \tau_{n+1}]$. $\square$

We are left with the task of proving the following statement for the classical Bessel process.

Lemma 4.12: For every classical Bessel process with index $-\frac{1}{2} < \vartheta < 0$

$$\mathbb{P}^x\big(\sup\{t : X_t = 0\} = \infty\big) = 1$$

is valid.

Proof: We use the stopping time τ_n defined in the proof of Theorem 4.11, see (4.11), and deduce

$$\begin{aligned}
\mathbb{P}^x\big(\tau_{n+1} < \infty\big) &= \mathbb{E}^x\big(\mathbb{1}_{\{\tau_{n+1}<\infty\}}\big)\\
&= \mathbb{E}^x\Big(\mathbb{E}\big(\mathbb{1}_{\{\tau_{n+1}<\infty\}}\,|\,\mathcal{F}_{\tau_n+1}\big)\Big)\\
&= \mathbb{E}^x\Big(\mathbb{E}\big(\mathbb{1}_{\{\inf\{t\geq 0\,|\,X_{t+\tau_n+1}=0\}<\infty\}}\,\big|\,\mathcal{F}_{\tau_n+1}\big)\Big)\\
&= \mathbb{E}^x\Big(\mathbb{E}^{X_{\tau_n+1}}\big(\mathbb{1}_{\{\inf\{t\geq 0\,|\,X_t=0\}<\infty\}}\big)\Big)\\
&= \mathbb{E}^x\big(\mathbb{1}_{\{\inf\{t\geq 0\,|\,X_t=0\}<\infty\}}\big)\\
&= \mathbb{P}^x\big(\inf\{t\geq 0\,|\,X_t=0\} < \infty\big) = 1.
\end{aligned}$$

The fourth line is valid due to the strong Markov property and the last line was proved in [56, Theorem 1.1 (iv)]. Consequently, using the continuity from above we derive

$$\begin{aligned}
\mathbb{P}^x\big(\forall n \in \mathbb{N} : \tau_n < \infty\big) &= \mathbb{P}^x\Bigg(\bigcap_{n\in\mathbb{N}}\{\tau_n < \infty\}\Bigg)\\
&= \lim_{n\to\infty} \mathbb{P}^x(\tau_n < \infty) = 1.
\end{aligned}$$

We consider $\omega \in \big\{\sup\{t : X_t = 0\} < \infty\big\}$ and define $t_0 := \sup\{t : X_t(\omega) = 0\}$. Hence, $\tau_{[t_0]+1}(\omega) = \infty$ yields

$$\big\{\sup\{t : X_t = 0\} = \infty\big\}^c \subset \big\{\forall n \in \mathbb{N} : \tau_n < \infty\big\}^c$$

and thus

$$\mathbb{P}^x\big(\sup\{t : X_t = 0\} = \infty\big) \geq \mathbb{P}^x\big(\forall n \in \mathbb{N} : \tau_n < \infty\big) = 1.$$

□

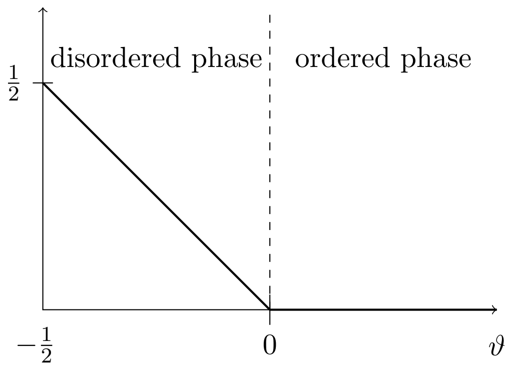

Figure 4.2: Behaviour of dim $\big(X^{-1}\{0\}\big)$.

Since the first hitting time of the origin is almost surely finite, Theorem 4.11 can be extended to arbitrary starting points $x > 0$ by using the countable stability of the Hausdorff dimension.

We give a schematic representation of the phase transition of the Hausdorff dimension in Figure 4.2. As every countable set has a Hausdorff dimension of zero, there exists a time interval in which the process hits the origin almost surely uncountably often when $-\frac{1}{2} < \vartheta < 0$.

4.3 Hausdorff dimension for the multivariate Bessel process

In this section we deal with the Hausdorff dimension of the times a multivariate Bessel process hits the Weyl chamber's boundary. The formulas here are slightly different than those presented in Section 4.2 as it was not successful proving the conditions in the multivariate case. Nevertheless, the proofs still work with a few modifications in the conditions. First of all, we state the basic facts: A multivariate Bessel process with positive multiplicity function k on the root system R hits ∂W_R almost surely if there exists an $\alpha \in R$ such that $k(\alpha) < \frac{1}{2}$ holds, [26, Proposition 1], whereas it never hits ∂W_R if $k(\alpha) > \frac{1}{2}$ for every $\alpha \in R$. We will focus on the A_{N-1} and B_N case, but nevertheless we hypothesize the following analogous result to the classical Bessel process for any multivariate Bessel process.

Conjecture 4.13: For every multivariate Bessel process $(Y_t)_{t\geq 0}$ starting in $x \in \overline{W}_R$ with multiplicity function $k : R \to (0,\infty)$ the equality

$$\dim\Big(Y^{-1}\big(\partial W_R\big)\Big) = \frac{1}{2} - \min\left\{\frac{1}{2}, \min_{\alpha\in R} k(\alpha)\right\}$$

is $\mathbb{P}^x$-almost surely valid.

We have already pointed out that we cover the case A_1 with the help of Section 4.2. For arbitrary dimensions we want to investigate this for the case A_{N-1} as well as B_N, which are quite similar models. In what follows, we cover the proofs for the A_{N-1} case. As things stand, we have not (yet) been able to give the complete proof for the B_N case. In Appendix A, we address the problems that arose by listing and proving the lemmas and the theorem which still work in this case, and explain the missing desired bound to finish the proof of Conjecture 4.13.

In the one-dimensional case, a key point for determining the Hausdorff dimension was finding bounds of the probability that the process is close to the origin. Therefore, we used the ball $B(0,r)$. Analogously, we define the edge set

$$E^r_{A_{N-1}} := \Big\{x \in \overline{W}_{A_{N-1}} \,\big|\, \exists\, i \in \{1,\ldots,N-1\} : 0 \leq x_{i+1} - x_i \leq r\Big\}$$

of thickness $r \geq 0$ which helps in later calculations of the hitting times with $\partial W_{A_{N-1}} = E^0_{A_{N-1}}$. Even if we focus on the A_{N-1} case we always suppress the Weyl chamber in the notation whenever the calculations work for every multivariate Bessel process. The following lemma is a basic property of our set $E^r_{A_{N-1}}$, which we regularly apply on the integration domain after a variable substitution without mentioning it.

Lemma 4.14: For any $r, c > 0$ the equality $cE^r_{A_{N-1}} = E^{cr}_{A_{N-1}}$ holds.

Proof: For the case A_{N-1} we derive immediately

$$\begin{aligned} cE^r_{A_{N-1}} &= \Big\{cx \in c\overline{W}_{A_{N-1}} \,\big|\, \exists\, i \in \{1,\ldots,N-1\} : 0 \leq x_{i+1} - x_i \leq r\Big\} \\ &= \Big\{cx \in \overline{W}_{A_{N-1}} \,\big|\, \exists\, i \in \{1,\ldots,N-1\} : 0 \leq c\big(x_{i+1} - x_i\big) \leq cr\Big\} \end{aligned}$$

$$= \left\{ y \in \overline{W}_{A_{N-1}} \,\middle|\, \exists\, i \in \{1,\dots,N-1\} : 0 \le y_{i+1} - y_i \le cr \right\}$$
$$= E^{cr}_{A_{N-1}}$$

by using the invariance under multiplication of the closed Weyl chamber in the second line. □

If we define the corresponding edge set

$$E^r_R := \left\{ x \in \overline{W}_R \,\middle|\, \exists \alpha \in R : 0 \le \langle \alpha, x \rangle \le r \right\}$$

of thickness $r > 0$, we preserve also the property of the previous lemma while

$$\partial W_R = E^0_R \subset E^r_R \subset W_R$$

holds. The main result of this section is the proof of Conjecture 4.13 for $R = A_{N-1}$.

Theorem 4.15 (Case A_{N-1}): The Hausdorff dimension of collision times for the multivariate Bessel process of type A_{N-1} starting in $x \in \overline{W}_{A_{N-1}}$ is given by

$$\dim \left(Y^{-1}(\partial W_{A_{N-1}}) \right) = \frac{1}{2} - \min \left\{ \frac{1}{2}, k \right\}$$

$\mathbb{P}^x$-almost surely.

We perform the proof for $x \in \partial W_{A_{N-1}}$. Having a finite hitting time of $\partial W_{A_{N-1}}$ in case $k < \frac{1}{2}$ and the countable stability of the Hausdorff dimension, we can easily transfer this to starting points $x \in \overline{W}_{A_{N-1}}$. Furthermore, we ignore the case $k > \frac{1}{2}$ since the dimension is obviously almost surely zero when no particles collide. As the proof is more involved, we split it into two parts, first proving the upper bound and then the lower bound for $x \in \partial W_{A_{N-1}}$.

4.3.1 Upper bound

The main work for the upper bound of the Hausdorff dimension lies in the proof of the following lemma, dealing with the probability of the process being in the edge set at time $t = 1$. The key point will be a useful inequality of Piotr Graczyk and Patrice Sawyer, [39].

Lemma 4.16 (Case A_{N-1}): For every $k>0$, $r>0$ and $\varrho>0$ there exist constants $C_1=C_1(k,N)>0$ and $C_2=C_2(k,N,\varrho)>0$ such that

$$C_1 \min\left\{1, r^{2k+1}\right\} \leq Q_k\left(1, x, E^r_{A_{N-1}}\right) \leq C_2 \min\left\{1, r^{2k+1}\right\}$$

holds for all $x \in E^r_{A_{N-1}} \cap B(0,\varrho)$.

Remark: The main difference to the classical Bessel process, apart from the more complicated set $E^r_{A_{N-1}}$, is that C_2 depends on the bound of the starting point x. This will not change the proof of the upper bound of the Hausdorff dimension but it does for the lower bound of the Hausdorff dimension considerably.

Proof: Lower bound: In this proof we use the behaviour of the density from [39, Theorem 3.1], which means there exists a constant $C=C(N,k)>0$ with

$$\mathbb{P}^x(Y_1 \in E^r_{A_{N-1}}) \geq C(N,k) \int\limits_{E^r_{A_{N-1}}} \frac{e^{-\frac{\|x-y\|^2}{2}} w_{A_{N-1}}(y)}{\prod\limits_{1\leq l<m\leq N} \left(\frac{1}{2}+(x_m-x_l)(y_m-y_l)\right)^k} \,\mathrm{d}y.$$

First, we consider the case $0<r\leq 1$. We can bound this further by choosing

$$i \in \underset{j=1,\dots,N-1}{\arg\min}\,(x_{j+1}-x_j)$$

and defining the set

$$E^{r,i}_{A_{N-1}} := \left\{y \in \overline{W}_{A_{N-1}} : 0 \leq y_{i+1}-y_i \leq r\right\}$$

in order to write

$$\mathbb{P}^x(Y_1 \in E^r_{A_{N-1}}) \geq C(N,k) \int\limits_{E^{r,i}_{A_{N-1}}} \frac{e^{-\frac{\|x-y\|^2}{2}} w_{A_{N-1}}(y)}{\prod\limits_{1\leq l<m\leq N} \left(\frac{1}{2}+(x_m-x_l)(y_m-y_l)\right)^k} \,\mathrm{d}y.$$

Next, we consider a variable substitution centered around x:

$$\begin{aligned} y_j &= z_j + x_j, \quad j \notin \{i, i+1\}, \\ y_i &= \frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2}, \\ y_{i+1} &= \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} \end{aligned} \tag{4.12}$$

or equivalently

$$\begin{aligned} z_j &= y_j - x_j, \quad j \notin \{i, i+1\}, \\ z_i &= \frac{y_i + y_{i+1}}{\sqrt{2}} - \frac{x_i + x_{i+1}}{\sqrt{2}}, \\ z_{i+1} &= \frac{y_{i+1} - y_i}{\sqrt{2}} \geq 0. \end{aligned}$$

Subsequently, we examine the new integration domain and make it independent of x by decreasing its size. In particular, for the following inequalities, we always begin by using the fact that y is ordered. Hereafter, we consider the upper bounds of the components of z separately. Owing to the ordering of y we automatically obtain the lower bounds as well. For $j \notin \{i-1, i, i+1\}$ we get

$$z_j + x_j \leq z_{j+1} + x_{j+1}$$

and hence

$$z_j \leq z_{j+1} + (x_{j+1} - x_j). \tag{4.13}$$

Since x is ordered as well, less combinations of (z_j, z_{j+1}) satisfy the inequality

$$z_j \leq z_{j+1}$$

than (4.13). For $j = i-1$, we similarly obtain

$$z_{i-1} + x_{i-1} \leq \frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2},$$

and reformulate

$$z_{i-1} \leq \frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{(x_i - x_{i-1}) + (x_{i+1} - x_{i-1})}{2}.$$

Coming from this upper bound we can make the integration domain smaller due to

$$\frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{(x_i - x_{i-1}) + (x_{i+1} - x_{i-1})}{2} \geq \frac{z_i - z_{i+1}}{\sqrt{2}} \geq \frac{z_i}{\sqrt{2}} - \frac{1}{2},$$

where we additionally use $0 \leq z_{i+1} = \frac{y_{i+1}-y_i}{\sqrt{2}} \leq \frac{r}{\sqrt{2}} \leq \frac{1}{\sqrt{2}}$. In the end, the more restrictive integration upper bound for z_{i-1} is specified by

$$z_{i-1} \leq \frac{z_i}{\sqrt{2}} - \frac{1}{2}.$$

For $j = i$, $y_{i+1} \leq y_{i+2}$ provides

$$\begin{aligned}\frac{z_i}{\sqrt{2}} &\leq z_{i+2} + x_{i+2} - \left(\frac{z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2}\right)\\ &= z_{i+2} - \frac{z_{i+1}}{\sqrt{2}} + \frac{(x_{i+2} - x_i) + (x_{i+2} - x_{i+1})}{2},\end{aligned}$$

so we get as new upper bound for the integration limit

$$z_{i+2} - \frac{z_{i+1}}{\sqrt{2}} + \frac{(x_{i+2} - x_i) + (x_{i+2} - x_{i+1})}{2} \geq z_{i+2} - \frac{1}{2}.$$

For $j = i + 1$, we have already observed

$$0 \leq z_{i+1} \leq \frac{r}{\sqrt{2}}.$$

Finally, we get as new integration limits

$$\begin{aligned} z_j &\leq z_{j+1},\ j \notin \{i-1, i, i+1\},\\ z_{i-1} &\leq \frac{z_i}{\sqrt{2}} - \frac{1}{2} \leq z_{i+2} - 1,\\ 0 &\leq z_{i+1} \leq \frac{r}{\sqrt{2}}. \end{aligned} \tag{4.14}$$

We now want to define a new integration domain imposing additionally a bound on

$$\sum_{\substack{j=1\\ j\neq i+1}}^{N} z_j^2$$

for later calculations. For this, every vector z within

$$\underbrace{\Big[0,\frac{1}{2}\Big]\times\cdots\times\Big[\frac{i-2}{2},\frac{i-1}{2}\Big]}_{\ni(z_1,\dots,z_{i-1})}\times\Big[\frac{i}{\sqrt{2}},\frac{i+1}{\sqrt{2}}\Big]\times\Big[0,\frac{r}{\sqrt{2}}\Big]\times\Big[\frac{i+2}{2},i+2\Big]$$
$$\times\underbrace{\Big[i+2,i+3\Big]\times\cdots\times\Big[N-1,N\Big]}_{\ni(z_{i+3},\dots,z_N)}$$

fulfills (4.14) and in addition

$$\begin{aligned}\sum_{\substack{j=1\\ j\neq i+1}}^{N} z_j^2 &\leq \sum_{j=1}^{i-1}\frac{j^2}{4}+\frac{(i+1)^2}{2}+\sum_{j=i+2}^{N} j^2\\ &\leq \sum_{j=1}^{N} j^2=\frac{N(N+1)(2N+1)}{6}\\ &\leq (N+1)^3.\end{aligned}$$

We use these considerations to define

$$\begin{aligned}\Lambda_{i,N}:=\Big\{z\in\mathbb{R}^N:\sum_{\substack{j=1\\ j\neq i+1}}^{N} z_j^2\leq (N+1)^3,\ z_j\leq z_{j+1}\ j\notin\{i-1,i,i+1\},\\ z_{i-1}\leq\frac{z_i}{\sqrt{2}}-\frac{1}{2}\leq z_{i+2}-1\Big\}\end{aligned}\tag{4.15}$$

so that the new integration domain is $\Lambda_{i,N}\cap\big\{0\leq z_{i+1}\leq\frac{r}{\sqrt{2}}\big\}$ since we observed in the calculations above that $Az+b\in E^{r,i}_{A_{N-1}}$ for every $z\in\Lambda_{i,N}\cap\big\{0\leq z_{i+1}\leq\frac{r}{\sqrt{2}}\big\}$, where A and b defines the transformation given by the equations (4.12). Since A maps all except of two components onto itself while the remaining ones are rotated, we note that $|\det(A)|=1$. In particular, we conclude

$$\lambda^{N-1}\big(p_{i+1}(\Lambda_{i,N})\big)\geq\frac{1}{2^{i-1}}\cdot\frac{1}{\sqrt{2}}\cdot\frac{i+2}{2}\cdot 1^{N-i-2}\geq\frac{1}{2^N}$$

with the projection

$$p_{i+1}:(\xi_1,\dots,\xi_N)\mapsto(\xi_1,\dots,\xi_i,\xi_{i+2},\dots,\xi_N)$$

which is important for the later calculations. Additionally, $\big[-(N+1)^{\frac{3}{2}},(N+1)^{\frac{3}{2}}\big]\supset$

$p_{i+1}(\Lambda_{i,N})$ implies $\lambda^{N-1}\big(p_{i+1}(\Lambda_{i,N})\big) \leq 2^{N-1}(N+1)^{\frac{3(N-1)}{2}}$. The bound on the sum in (4.15) may be chosen smaller so that $0 < \lambda^{N-1}\big(p_{i+1}(\Lambda_{i,N})\big) < \infty$ but we are not interested in the sharpest bound. We now examine each term of the integral separately to find an x-independent lower bound:

$$\begin{aligned}
\|x-y\|^2 &= \sum_{\substack{j=1\\ j\neq i,i+1}}^{N} (y_j - x_j)^2 + (y_i - x_i)^2 + (y_{i+1} - x_{i+1})^2 \\
&= \sum_{\substack{j=1\\ j\neq i,i+1}}^{N} z_j^2 + \Big(\frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} - x_i\Big)^2 + \Big(\frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} - x_{i+1}\Big)^2 \\
&= \sum_{\substack{j=1\\ j\neq i,i+1}}^{N} z_j^2 + \Big(\frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_i}{2}\Big)^2 + \Big(\frac{z_i + z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} - x_i}{2}\Big)^2 \\
&= \|z\|^2 + \frac{(x_{i+1}-x_i)^2}{2} + (x_{i+1}-x_i)\frac{z_i - z_{i+1}}{\sqrt{2}} - (x_{i+1}-x_i)\frac{z_i + z_{i+1}}{\sqrt{2}} \\
&= \|z\|^2 + \frac{(x_{i+1}-x_i)^2}{2} - \sqrt{2}(x_{i+1}-x_i)z_{i+1} \\
&\leq \|z\|^2 + \frac{(x_{i+1}-x_i)^2}{2} \\
&\leq (N+1)^3 + z_{i+1}^2 + \frac{(x_{i+1}-x_i)^2}{2} \\
&\leq (N+1)^3 + \frac{r^2}{2} + \frac{r^2}{2} \leq (N+1)^3 + 1.
\end{aligned}$$

The first and third inequality are valid since $0 \leq x_{i+1} - x_i \leq r$ and $0 \leq z_{i+1} \leq \frac{r}{\sqrt{2}}$ for $r < 1$, respectively, whereas we used the imposed bound on the sum in the second-last line. The remaining terms of the integral are given by the product

$$\prod_{1\leq l<m\leq N} \frac{(y_m - y_l)^2}{\frac{1}{2} + (x_m - x_l)(y_m - y_l)}$$

to the power k. To express this product in terms of z, we must consider six separate cases. The simplest case is that where $l = i$, $m = i+1$:

$$\frac{(y_{i+1} - y_i)^2}{\frac{1}{2} + (x_{i+1} - x_i)(y_{i+1} - y_i)} = \frac{(\sqrt{2}z_{i+1})^2}{\frac{1}{2} + \sqrt{2}z_{i+1}(x_{i+1} - x_i)} \geq \frac{4}{3}z_{i+1}^2.$$

Here, we have used the fact that both $\sqrt{2}z_{i+1}$ and $x_{i+1} - x_i$ are less than or equal to r, which in turn is bounded above by 1. Next, we look at the case $l, m \notin \{i, i+1\}$, where we get

$$\begin{aligned} \frac{(y_m - y_l)^2}{\frac{1}{2} + (y_m - y_l)(x_m - x_l)} &= \frac{(z_m - z_l + x_m - x_l)^2}{\frac{1}{2} + (z_m - z_l + x_m - x_l)(x_m - x_l)} \\ &\geq \frac{(z_m - z_l + x_m - x_l)^2}{\frac{1}{2} + (z_m - z_l + x_m - x_l)^2} \end{aligned}$$

since $z \in \Lambda_{i,N}$ implies $z_m - z_l \geq 0$. Observing that the function

$$\xi \mapsto \frac{\xi^2}{\frac{1}{2} + \xi^2} \tag{4.16}$$

is increasing in $\xi > 0$, we conclude that

$$\frac{(y_m - y_l)^2}{\frac{1}{2} + (y_m - y_l)(x_m - x_l)} \geq \frac{(z_m - z_l)^2}{\frac{1}{2} + (z_m - z_l)^2}$$

as $x \in \overline{W}_{A_{N-1}}$ means $x_m - x_l \geq 0$. In the other cases we proceed in the same way. First, we add a positive term in the numerator and then use the increasing behaviour of (4.16). For the second step we will use $x_{i+1} - x_i \leq 1$, $0 \leq z_{i+1} \leq \frac{1}{\sqrt{2}}$ and the ordering of x without further mentioning. The following case is $l = i$, $m > i+1$, where we have

$$\begin{aligned} \frac{(y_m - y_i)^2}{\frac{1}{2} + (y_m - y_i)(x_m - x_i)} &= \frac{\left(z_m + x_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_i + x_{i+1}}{2}\right)^2}{\frac{1}{2} + \left(z_m + x_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_i + x_{i+1}}{2}\right)(x_m - x_i)} \\ &= \frac{\left(z_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} - x_i}{2} + x_m - x_i\right)^2}{\frac{1}{2} + \left(z_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} - x_i}{2} + x_m - x_i\right)(x_m - x_i)} \\ &\geq \frac{\left(z_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} - x_i}{2} + x_m - x_i\right)^2}{\frac{1}{2} + \left(z_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} - x_i}{2} + x_m - x_i\right)^2} \\ &\geq \frac{\left(z_m - \frac{z_i}{\sqrt{2}} - \frac{1}{2}\right)^2}{\frac{1}{2} + \left(z_m - \frac{z_i}{\sqrt{2}} - \frac{1}{2}\right)^2}. \end{aligned}$$

In particular, we used for the first inequality

$$\begin{aligned} z_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} - x_i}{2} &\geq z_{i+1} - \frac{z_i - z_{i+1}}{\sqrt{2}} - \underbrace{\frac{x_{i+1} - x_i}{2}}_{\leq \frac{1}{2}} \\ &\geq \frac{z_i}{\sqrt{2}} + \frac{1}{2} - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{1}{2} \\ &= \frac{z_{i+1}}{\sqrt{2}} \geq 0 \end{aligned}$$

while reminding the form of $\Lambda_{i,N}$. Similarly, when $l = i+1$ and $m > i+1$ we have

$$\begin{aligned} \frac{(y_m - y_{i+1})^2}{\frac{1}{2} + (y_m - y_{i+1})(x_m - x_{i+1})} &= \frac{\left(z_m + x_m - \frac{z_i + z_{i+1}}{\sqrt{2}} - \frac{x_i + x_{i+1}}{2}\right)^2}{\frac{1}{2} + \left(z_m + x_m - \frac{z_i + z_{i+1}}{\sqrt{2}} - \frac{x_i + x_{i+1}}{2}\right)(x_m - x_{i+1})} \\ &= \frac{\left(z_m - \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_i}{2} + x_m - x_{i+1}\right)^2}{\frac{1}{2} + \left(z_m - \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_i}{2} + x_m - x_{i+1}\right)(x_m - x_{i+1})} \\ &\geq \frac{\left(z_m - \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_i}{2} + x_m - x_{i+1}\right)^2}{\frac{1}{2} + \left(z_m - \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_i}{2} + x_m - x_{i+1}\right)^2} \\ &\geq \frac{\left(z_m - \frac{z_i}{\sqrt{2}} - \frac{1}{2}\right)^2}{\frac{1}{2} + \left(z_m - \frac{z_i}{\sqrt{2}} - \frac{1}{2}\right)^2} \end{aligned}$$

since

$$\begin{aligned} z_m - \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_i}{2} &\geq z_{i+2} - \frac{z_i + z_{i+1}}{\sqrt{2}} + \underbrace{\frac{x_{i+1} - x_i}{2}}_{\geq 0} \\ &\geq \frac{z_i}{\sqrt{2}} + \frac{1}{2} - \frac{z_i + z_{i+1}}{\sqrt{2}} \\ &= \frac{1}{2} - \frac{z_{i+1}}{\sqrt{2}} \\ &\geq \frac{1}{2} - \frac{r}{2} > 0. \end{aligned}$$

The last two cases are those for which $l < i$ and $m = i$ or $i+1$. For the first one we derive

$$\begin{aligned}
\frac{(y_i - y_l)^2}{\frac{1}{2} + (y_i - y_l)(x_i - x_l)} &= \frac{\left(\frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} - z_l - x_l\right)^2}{\frac{1}{2} + \left(\frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} - z_l - x_l\right)(x_i - x_l)} \\
&= \frac{\left(\frac{z_i - z_{i+1}}{\sqrt{2}} - z_l + \frac{x_{i+1} - x_i}{2} + x_i - x_l\right)^2}{\frac{1}{2} + \left(\frac{z_i - z_{i+1}}{\sqrt{2}} - z_l + \frac{x_{i+1} - x_i}{2} + x_i - x_l\right)(x_i - x_l)} \\
&\geq \frac{\left(\frac{z_i - z_{i+1}}{\sqrt{2}} - z_l + \frac{x_{i+1} - x_i}{2} + x_i - x_l\right)^2}{\frac{1}{2} + \left(\frac{z_i - z_{i+1}}{\sqrt{2}} - z_l + \frac{x_{i+1} - x_i}{2} + x_i - x_l\right)^2} \\
&\geq \frac{\left(\frac{z_i}{\sqrt{2}} - \frac{1}{2} - z_l\right)^2}{\frac{1}{2} + \left(\frac{z_i}{\sqrt{2}} - \frac{1}{2} - z_l\right)^2}
\end{aligned}$$

since

$$\begin{aligned}
\frac{z_i - z_{i+1}}{\sqrt{2}} - z_l + \frac{x_{i+1} - x_i}{2} &\geq \frac{z_i - z_{i+1}}{\sqrt{2}} - z_{i-1} + \underbrace{\frac{x_{i+1} - x_i}{2}}_{\geq 0} \\
&\geq \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{z_i}{\sqrt{2}} + \frac{1}{2} \\
&= \frac{1}{2} - \frac{z_{i+1}}{\sqrt{2}} \\
&\geq \frac{1}{2} - \frac{r}{2} > 0.
\end{aligned}$$

In the same way, we get

$$\begin{aligned}
\frac{(y_{i+1} - y_l)^2}{\frac{1}{2} + (y_{i+1} - y_l)(x_{i+1} - x_l)} &= \frac{\left(\frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} - z_l - x_l\right)^2}{\frac{1}{2} + \left(\frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_i + x_{i+1}}{2} - z_l - x_l\right)(x_{i+1} - x_l)} \\
&= \frac{\left(\frac{z_i + z_{i+1}}{\sqrt{2}} - z_l - \frac{x_{i+1} - x_i}{2} + x_{i+1} - x_l\right)^2}{\frac{1}{2} + \left(\frac{z_i + z_{i+1}}{\sqrt{2}} - z_l - \frac{x_{i+1} - x_i}{2} + x_{i+1} - x_l\right)(x_{i+1} - x_l)} \\
&\geq \frac{\left(\frac{z_i + z_{i+1}}{\sqrt{2}} - z_l - \frac{x_{i+1} - x_i}{2} + x_{i+1} - x_l\right)^2}{\frac{1}{2} + \left(\frac{z_i + z_{i+1}}{\sqrt{2}} - z_l - \frac{x_{i+1} - x_i}{2} + x_{i+1} - x_l\right)^2}
\end{aligned}$$

$$\geq \frac{\left(\frac{z_i}{\sqrt{2}} - z_l - \frac{1}{2}\right)^2}{\frac{1}{2} + \left(\frac{z_i}{\sqrt{2}} - z_l - \frac{1}{2}\right)^2}$$

for $m = i+1$ and $l < i$ since

$$\begin{aligned}
\frac{z_i + z_{i+1}}{\sqrt{2}} - z_l - \frac{x_{i+1} - x_i}{2} &\geq \frac{z_i + z_{i+1}}{\sqrt{2}} - z_{i-1} - \underbrace{\frac{x_{i+1} - x_i}{2}}_{\leq \frac{1}{2}} \\
&\geq \frac{z_i + z_{i+1}}{\sqrt{2}} - \frac{z_i}{\sqrt{2}} + \frac{1}{2} - \frac{1}{2} \\
&= \frac{z_{i+1}}{\sqrt{2}} \geq 0.
\end{aligned}$$

For a more concise notation we define

$$z_{i+1}^* := p_{i+1}(z) = \big(z_1, \ldots, z_i, z_{i+2}, \ldots, z_N\big) \tag{4.17}$$

to finally obtain with all of these relations:

$$\begin{aligned}
&\mathbb{P}^x\big(X_1 \in E^r_{A_{N-1}}\big) \\
&\geq C(N,k) \int\limits_{p_{i+1}(\Lambda_{i,N})} \int\limits_0^{\frac{r}{\sqrt{2}}} e^{-\frac{(N+1)^3+1}{2}} \frac{4^k}{3^k} z_{i+1}^{2k} \Pi^{(1)}_{N,k,i}(z_{i+1}^*) \Pi^{(2)}_{k,i}(z_{i+1}^*) \Pi^{(3)}_{N,k,i}(z_{i+1}^*) \,\mathrm{d}z_{i+1} \,\mathrm{d}z_{i+1}^* \\
&= C(N,k) \frac{e^{-\frac{(N+1)^3+1}{2}} 2^k}{\sqrt{2}(2k+1)3^k} r^{2k+1} \int\limits_{p_{i+1}(\Lambda_{i,N})} \Pi^{(1)}_{N,k,i}(z_{i+1}^*) \Pi^{(2)}_{k,i}(z_i^*) \Pi^{(3)}_{N,k,i}(z_{i+1}^*) \,\mathrm{d}z_{i+1}^*,
\end{aligned}$$

where the products are given by

$$\begin{aligned}
\Pi^{(1)}_{N,k,i}(z_{i+1}^*) &:= \prod_{\substack{1 \leq l < m \leq N \\ l,m \neq i,i+1}} \frac{(z_m - z_l)^{2k}}{\left(\frac{1}{2} + (z_m - z_l)^2\right)^k}, \\
\Pi^{(2)}_{k,i}(z_{i+1}^*) &:= \prod_{1 \leq l < i} \frac{\left(\frac{z_i}{\sqrt{2}} - z_l - \frac{1}{2}\right)^{4k}}{\left[\frac{1}{2} + \left(\frac{z_i}{\sqrt{2}} - z_l - \frac{1}{2}\right)^2\right]^{2k}}, \\
\Pi^{(3)}_{N,k,i}(z_{i+1}^*) &:= \prod_{i+1 < m \leq N} \frac{\left(z_m - \frac{z_i}{\sqrt{2}} - \frac{1}{2}\right)^{4k}}{\left[\frac{1}{2} + \left(z_m - \frac{z_i}{\sqrt{2}} - \frac{1}{2}\right)^2\right]^{2k}}.
\end{aligned}$$

We notice that the cases $l = i+1$ and $l = i$ respectively $m = i+1$ and $m = i$ lead to the same bound. Setting

$$C_1(N,k) := C(N,k)\frac{e^{-\frac{(N+1)^3+1}{2}}2^k}{\sqrt{2}(2k+1)3^k}\cdot \min_{i=1,\dots,N-1} \int\limits_{p_{i+1}(\Lambda_{i,N,R})} \Pi^{(1)}_{N,k,i}(z^*_{i+1})\Pi^{(2)}_{k,i}(z^*_{i+1})\Pi^{(3)}_{N,k,i}(z^*_{i+1})\,\mathrm{d}z^*_{i+1}$$

and noting that all factors in the integrand are positive in the interior of $\Lambda_{i,N}$, which has a non-zero finite Lebesgue measure, we conclude that $C_1(N,k) > 0$ with

$$\mathbb{P}^x(Y_1 \in E^r) \geq C_1(N,k)r^{2k+1}.$$

In the case $r > 1$ we again start from the same point:

$$\mathbb{P}^x\big(Y_1 \in E^r_{A_{N-1}}\big) \geq C(N,k)\int\limits_{E^r_{A_{N-1}}} \frac{e^{-\frac{\|x-y\|^2}{2}}w_{A_{N-1}}(y)}{\prod\limits_{1\leq l<m\leq N}\big(\frac{1}{2}+(x_m-x_l)(y_m-y_l)\big)^k}\,\mathrm{d}y.$$

If $x \in E^1_{A_{N-1}}$, then we use the same arguments as in the the case $r \leq 1$ by making use of

$$\mathbb{P}^x\big(Y_1 \in E^r_{A_{N-1}}\big) \geq \mathbb{P}^x\big(Y_1 \in E^1_{A_{N-1}}\big).$$

By this observation, we can apply the calculations for the case $r \leq 1$ specialized to $r = 1$. Otherwise, $x \in E^r_{A_{N-1}} \backslash E^1_{A_{N-1}}$ which means that after choosing

$$i = \operatorname*{arg\,min}_{j=1,\dots,N-1}(x_{j+1}-x_j)$$

we have

$$x_{i+1} - x_i \geq 1.$$

We consider the simple transformation $y = z+x$ with integration domain $z \in B(0,\frac{1}{2}) \cap \{z : z_{i+1}-z_i \leq 0\}$. To ensure this leads to a smaller integration domain we need to verify that $z + x \in E^r_{A_{N-1}}$ for any such z. Therefore, we examine bounds for $z_m - z_l$ for any two distinct $m, l \in \{1, \ldots, N\}$ under the restriction

$$\sum_{j=1}^{N} z_j^2 \leq \frac{1}{4}.$$

In particular, for all $l \neq m$ the simple condition

$$z_m^2 + z_l^2 \leq \frac{1}{4}$$

holds, which implies

$$-\sqrt{\frac{1}{4} - z_l^2} \leq z_m \leq \sqrt{\frac{1}{4} - z_l^2}$$

or equivalently

$$-\sqrt{\frac{1}{4} - z_l^2} - z_l \leq z_m - z_l \leq \sqrt{\frac{1}{4} - z_l^2} - z_l$$

for $z_l \in \big[-\frac{1}{2}, \frac{1}{2}\big]$. When maximizing the right-hand side, we calculate

$$\frac{\partial}{\partial z_l}\left(\sqrt{\frac{1}{4} - z_l^2} - z_l\right) = -\frac{z_l}{\sqrt{\frac{1}{4} - z_l^2}} - 1$$

which is zero for

$$z_l = -\sqrt{\frac{1}{4} - z_l^2}.$$

By squaring we receive as the only possible solution

$$z_l = -\frac{1}{\sqrt{8}} = -\frac{1}{2\sqrt{2}} \in \Big[-\frac{1}{2}, \frac{1}{2}\Big].$$

Comparing with the value at $\pm\frac{1}{2}$, that is,

$$\begin{aligned}\sqrt{\frac{1}{4} - \Big(\pm\frac{1}{2}\Big)^2} \mp \frac{1}{2} = \mp\frac{1}{2}& \\ \leq \frac{1}{\sqrt{2}} = \sqrt{\frac{1}{4} - \Big(-\frac{1}{\sqrt{8}}\Big)^2} + \frac{1}{2\sqrt{2}}&\end{aligned}$$

we obtain $\frac{1}{2\sqrt{2}}$ as the global maximum value. By an analogous calculation we get

$$-\frac{1}{\sqrt{2}} \leq z_m - z_l \leq \frac{1}{\sqrt{2}}. \tag{4.18}$$

Having these observations in mind, we receive

$$\begin{aligned} 0 < -\frac{1}{\sqrt{2}} + 1 &\leq -\frac{1}{\sqrt{2}} + x_{i+1} - x_i \\ &\leq z_{i+1} - z_i + x_{i+1} - x_i \\ &= y_{i+1} - y_i \\ &\leq 0 + r = r. \end{aligned}$$

due to $1 \leq x_{i+1} - x_i \leq r$ and $z_{i+1} - z_i \geq 0$. For the rest of the pairs $j \neq i$, we get

$$\begin{aligned} y_{j+1} - y_j &= z_{j+1} - z_j + x_{j+1} - x_j \\ &\geq -\frac{1}{\sqrt{2}} + x_{i+1} - x_i \\ &\geq -\frac{1}{\sqrt{2}} + 1 > 0. \end{aligned}$$

Therefore, the components of y are still ordered and in $E^r_{A_{N-1}}$. Consequently, by choosing a smaller integration domain we conclude

$$\begin{aligned} &\mathbb{P}^x(Y_1 \in E^r_{A_{N-1}}) \\ &\quad \geq C(N,k) \int\limits_{B\left(0,\frac{1}{2}\right) \cap \{z_{i+1} - z_i \leq 0\}} \frac{e^{-\frac{\|z\|^2}{2}} w_k(z+x)}{\prod\limits_{1 \leq l < m \leq N} \left(\frac{1}{2} + (x_m - x_l)(z_m - z_l + x_m - x_l)\right)^k} \, \mathrm{d}z \\ &\quad \geq C(N,k) \int\limits_{B\left(0,\frac{1}{2}\right) \cap \{z_{i+1} - z_i \leq 0\}} \frac{e^{-\frac{1}{8}} w_k(z+x)}{\prod\limits_{1 \leq l < m \leq N} \left(\frac{1}{2} + (x_m - x_l)(z_m - z_l + x_m - x_l)\right)^k} \, \mathrm{d}z. \end{aligned}$$

We take a look at each term in the product separately:

$$\begin{aligned} \frac{(z_m - z_l + x_m - x_l)^2}{\frac{1}{2} + (z_m - z_l + x_m - x_l)(x_m - x_l)} &\geq \frac{\left(x_m - x_l - \frac{1}{\sqrt{2}}\right)^2}{\frac{1}{2} + \left(x_m - x_l + \frac{1}{\sqrt{2}}\right)(x_m - x_l)} \\ &\geq \frac{\left(1 - \frac{1}{\sqrt{2}}\right)^2}{\frac{1}{2} + \left(1 + \frac{1}{\sqrt{2}}\right) \cdot 1}. \end{aligned}$$

First, we used the known bounds from (4.18) of $z_m - z_l$ and then applied that the function

$$\xi \mapsto \frac{\left(\xi - \frac{1}{\sqrt{2}}\right)^2}{\frac{1}{2} + \left(\xi + \frac{1}{\sqrt{2}}\right)\xi}$$

is increasing for $\xi \geq 1$ while $x_m - x_l \geq \min_{j=1,\dots,N-1}(x_{j+1} - x_j) = x_{i+1} - x_i \geq 1$. Therefore, we can write

$$\begin{aligned}
\mathbb{P}^x(Y_1 \in E^r) &\geq C(N,k) \left(\frac{\left(1 - \frac{1}{\sqrt{2}}\right)^2}{\frac{3}{2} + \frac{1}{\sqrt{2}}} \right)^{\frac{kN(N-1)}{2}} e^{-\frac{1}{8}} \lambda^N \Big(B\Big(0, \frac{1}{2}\Big) \cap \{z : z_{i+1} - z_i \leq 0\} \Big) \\
&= C(N,k) \left(\frac{\left(1 - \frac{1}{\sqrt{2}}\right)^2}{\frac{3}{2} + \frac{1}{\sqrt{2}}} \right)^{\frac{kN(N-1)}{2}} e^{-\frac{1}{8}} \frac{\lambda^N \Big(B\big(0, \frac{1}{2}\big)\Big)}{2} \\
&= C(N,k) \left(\frac{\left(1 - \frac{1}{\sqrt{2}}\right)^2}{\frac{3}{2} + \frac{1}{\sqrt{2}}} \right)^{\frac{kN(N-1)}{2}} \frac{\pi^{\frac{N}{2}} e^{-\frac{1}{8}}}{\Gamma\left(\frac{N}{2} + 1\right) 2^{N+1}}.
\end{aligned}$$

Since the Lebesgue measure is independent of i, we obtain the desired lower bound. Upper bound: The proof for the upper bound is significantly shorter, allowing the constant to depend on ϱ, the bound of the starting point x. This additional condition ensures that we have the necessary property

$$0 \leq x_m - x_l \leq \sqrt{2}\varrho \tag{4.19}$$

for every $1 \leq l < m \leq N$ since $x \in E^r_{A_{N-1}} \cap B(0, \varrho)$. This follows by making an analogous calculation as for the ball $B\big(0, \frac{1}{2}\big)$, see (4.18). First, if we choose the constant $C_2 \geq 1$ the inequality is trivial in the case $r > 1$ due to

$$\mathbb{P}^x\big(Y_1 \in E^r_{A_{N-1}}\big) \leq 1 \leq C_2 = C_2 \min\big\{1, r^{2k+1}\big\}.$$

Hence, we assume $r \leq 1$ and using again [39, Theorem 3.1] we obtain

$$\mathbb{P}^x\big(Y_1 \in E^r_{A_{N-1}}\big) \leq C(N,k) \int_{E^r_{A_{N-1}}} \frac{e^{-\frac{\|x-y\|^2}{2}} w_k(y)}{\prod_{1 \leq l < m \leq N} \left(\frac{1}{2} + (x_m - x_l)(y_m - y_l)\right)^k} \, \mathrm{d}y$$

$$\leq C(N,k) \sum_{i=1}^{N-1} \int\limits_{E^{r,i}_{A_{N-1}}} \frac{e^{-\frac{\|x-y\|^2}{2}} w_k(y)}{\prod\limits_{1\leq l<m\leq N} \left(\frac{1}{2} + (x_m - x_l)(y_m - y_l)\right)^k} \, \mathrm{d}y.$$

Here, it is evident that

$$E^r_{A_{N-1}} = \bigcup_{i=1}^{N-1} E^{r,i}_{A_{N-1}} = \bigcup_{i=1}^{N-1} \left\{ y \in \overline{W}_{A_{N-1}} : 0 \leq y_{i+1} - y_i \leq r \right\}$$

and we use the variable transformation (4.12). In particular, the constant $C(N,k) > 0$ differs from the one in the lower bound case but for simplicity we use the same term. Therefore, we again receive

$$\begin{aligned} \|x-y\|^2 &= \|z\|^2 + \frac{(x_{i+1}-x_i)^2}{2} - \sqrt{2}(x_{i+1}-x_i) z_{i+1} \\ &\geq \|z\|^2 - (x_{i+1}-x_i) r \\ &\geq \|z\|^2 - (x_{i+1}-x_i) \\ &\geq \sum_{\substack{j=1 \\ j\neq i+1}}^{N} z_j^2 - \sqrt{2}\varrho. \end{aligned}$$

since $\sqrt{2} z_{i+1} \leq r \leq 1$. The rest of the terms come from the product

$$\prod_{1\leq l<m\leq N} \left(\frac{(y_m - y_l)^2}{\frac{1}{2} + (x_m - x_l)(y_m - y_l)} \right)^k \leq 2^{\frac{kN(N-1)}{2}} \prod_{1\leq l<m\leq N} (y_m - y_l)^{2k}.$$

We have taken away the term $(x_m - x_l)(y_m - y_l) \geq 0$ in the denominator. The simplest term is $(l,m) = (i, i+1)$:

$$0 \leq y_{i+1} - y_i = \sqrt{2} z_{i+1} \leq r$$

which ensures the order r^{2k+1} of the probability. Now, we bound the other terms in which z_{i+1} occurs. In the following cases, we always use $0 \leq \frac{z_{i+1}}{\sqrt{2}} \leq \frac{r}{2} \leq \frac{1}{2}$ and (4.19). We bound the terms with $l, m \notin \{i, i+1\}$ by

$$0 \leq y_m - y_l = z_m - z_l + x_m - x_l \leq z_m - z_l + \sqrt{2}\varrho.$$

For $l = i$ and $m > i+1$, we obtain

$$\begin{aligned} 0 \le y_m - y_i &= z_m + x_m - \frac{z_i - z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} + x_i}{2} \\ &= z_m - \frac{z_i}{\sqrt{2}} + \frac{z_{i+1}}{\sqrt{2}} + \frac{x_m - x_{i+1}}{2} + \frac{x_m - x_i}{2} \\ &\le z_m - \frac{z_i}{\sqrt{2}} + \frac{1}{2} + \sqrt{2}\varrho \end{aligned}$$

and similarly, if $l = i+1$ and $m > i+1$,

$$\begin{aligned} 0 \le y_m - y_{i+1} &= z_m + x_m - \frac{z_i + z_{i+1}}{\sqrt{2}} - \frac{x_{i+1} + x_i}{2} \\ &= z_m - \frac{z_i}{\sqrt{2}} - \frac{z_{i+1}}{\sqrt{2}} + \frac{x_m - x_{i+1}}{2} + \frac{x_m - x_i}{2} \\ &\le z_m - \frac{z_i}{\sqrt{2}} + \sqrt{2}\varrho. \end{aligned}$$

When $m = i+1$ and $l < i$, we derive

$$\begin{aligned} 0 \le y_{i+1} - y_l &\le \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} + x_i}{2} - z_l - x_l \\ &= \frac{z_i}{\sqrt{2}} - z_l + \frac{z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_l}{2} + \frac{x_i - x_l}{2} \\ &\le \frac{z_i}{\sqrt{2}} - z_l + \frac{1}{2} + \sqrt{2}\varrho. \end{aligned}$$

The last case we must examine is that where $m = i$ and $l < i$:

$$\begin{aligned} 0 \le y_i - y_l &= \frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} + x_i}{2} - z_l - x_l \\ &= \frac{z_i}{\sqrt{2}} - z_l - \frac{z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} - x_l}{2} + \frac{x_i - x_l}{2} \\ &\le \frac{z_i}{\sqrt{2}} - z_l + \sqrt{2}\varrho. \end{aligned}$$

Next, we recall the integration domain. Since $y \in \overline{W}_{A_{N-1}}$ the components of z must satisfy

$$\begin{aligned} z_j + x_j &\le z_{j+1} + x_{j+1}, \qquad j \notin \{i-1, i, i+1\}, \\ z_{i-1} + x_{i-1} &\le \frac{z_i - z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} + x_i}{2}, \\ \frac{z_i + z_{i+1}}{\sqrt{2}} + \frac{x_{i+1} + x_i}{2} &\le z_{i+2} + x_{i+2}, \\ 0 &\le \sqrt{2} z_{i+1} \le r. \end{aligned}$$

In particular, these should now become independent of z_{i+1} and r except for the component z_{i+1} itself. Because of $z_{i+1} \geq 0$, there exist more possible z that satisfy the following conditions:

$$\begin{aligned} z_j + x_j &\leq z_{j+1} + x_{j+1}, \qquad j \notin \{i-1, i, i+1\}, \\ z_{i-1} + x_{i-1} &\leq \frac{z_i}{\sqrt{2}} + \frac{x_{i+1} + x_i}{2}, \\ \frac{z_i}{\sqrt{2}} + \frac{x_{i+1} + x_i}{2} &\leq z_{i+2} + x_{i+2}, \\ 0 &\leq \sqrt{2} z_{i+1} \leq r. \end{aligned}$$

Furthermore, to get rid of the dependence of x we reformulate

$$\begin{aligned} z_j &\leq z_{j+1} + (x_{j+1} - x_j), \qquad j \notin \{i-1, i, i+1\}, \\ z_{i-1} &\leq \frac{z_i}{\sqrt{2}} + \frac{x_{i+1} - x_{i-1}}{2} + \frac{x_i - x_{i-1}}{2} \\ \frac{z_i}{\sqrt{2}} + &\leq z_{i+2} + \frac{x_{i+2} - x_{i+1}}{2} + \frac{x_{i+2} - x_i}{2} \\ 0 &\leq \sqrt{2} z_{i+1} \leq r. \end{aligned} \tag{4.20}$$

and then enlarge the integration domain given by (4.20) due to (4.19). In summary, we define the set

$$\mathcal{M}_{i,\varrho,N} := \Big\{ z \in \mathbb{R}^N : z_j \leq z_{j+1} + \sqrt{2}\varrho \text{ for } j \notin \{i-1, i, i+1\}, \\ z_{i-1} - \sqrt{2}\varrho \leq \frac{z_i}{\sqrt{2}} \leq z_{i+2} + \sqrt{2}\varrho \Big\}$$

in order to obtain the new, greater integration domain

$$\mathcal{M}_{i,\varrho,N} \cap \Big\{ 0 \leq z_{i+1} \leq \frac{r}{\sqrt{2}} \Big\}.$$

With these definitions and inequalities, we obtain

$$\mathbb{P}^x\big(Y_1 \in E^r_{A_{N-1}}\big) \leq C(N,k) \sum_{i=1}^{N-1} \int_{E^{r,i}_{A_{N-1}}} \frac{e^{-\frac{\|x-y\|^2}{2}} w_k(y)}{\prod\limits_{1 \leq l < m \leq N} \big(\frac{1}{2} + (x_m - x_l)(y_m - y_l)\big)^k} \, \mathrm{d}y$$

$$
\begin{aligned}
&\leq C(N,k)\sum_{i=1}^{N-1}\int\limits_{E^{r,i}_{A_{N-1}}} 2^{\frac{kN(N-1)}{2}}e^{-\frac{\|x-y\|^2}{2}}w_k(y)\,\mathrm{d}y\\
&\leq C(N,k)\sum_{i=1}^{N-1}\int\limits_{p_{i+1}(\mathcal{M}_{i,\varrho,N})}\int\limits_{0}^{\frac{r}{\sqrt{2}}} 2^{\frac{kN(N-1)}{2}}e^{-\frac{\|z^*_{i+1}\|^2}{2}+\frac{\varrho}{\sqrt{2}}}\cdot\\
&\qquad\qquad 2^k z_{i+1}^{2k}\widetilde{\Pi}^{(1)}_{N,k,i,\varrho}(z^*_{i+1})\widetilde{\Pi}^{(2)}_{N,k,i,\varrho}(z^*_{i+1})\widetilde{\Pi}^{(3)}_{k,i,\varrho}(z^*_{i+1})\,\mathrm{d}z^*_{i+1}\\
&= C(N,k)\frac{2^{\frac{kN(N-1)}{2}+k}r^{2k+1}}{2k+1}\sum_{i=1}^{N-1}\int\limits_{p_{i+1}(\mathcal{M}_{i,\varrho,N})} e^{-\frac{\|z^*_{i+1}\|^2}{2}+\frac{\varrho}{\sqrt{2}}}\cdot\\
&\qquad\qquad \widetilde{\Pi}^{(1)}_{N,k,i,\varrho}(z^*_{i+1})\widetilde{\Pi}^{(2)}_{N,k,i,\varrho}(z^*_{i+1})\widetilde{\Pi}^{(3)}_{k,i,\varrho}(z^*_{i+1})\,\mathrm{d}z^*_{i+1}\\
&=: C_2(N,k,\varrho)r^{2k+1}
\end{aligned}
$$

with

$$
\begin{aligned}
\widetilde{\Pi}^{(1)}_{N,k,i,\varrho}(z^*_{i+1}) &:= \prod_{\substack{1\leq l<m\leq N\\ l,m\neq i,i+1}}(z_m-z_l+\sqrt{2}\varrho)^{2k},\\
\widetilde{\Pi}^{(2)}_{N,k,i,\varrho}(z^*_{i+1}) &:= \prod_{i+1<m\leq N}\left[\left(z_m-\frac{z_i}{\sqrt{2}}+\frac{1}{2}+\sqrt{2}\varrho\right)\left(z_m-\frac{z_i}{\sqrt{2}}+\sqrt{2}\varrho\right)\right]^{2k},\\
\widetilde{\Pi}^{(3)}_{k,i,\varrho}(z^*_{i+1}) &:= \prod_{1\leq l<i}\left[\left(\frac{z_i}{\sqrt{2}}-z_l+\frac{1}{2}+\sqrt{2}\varrho\right)\left(\frac{z_i}{\sqrt{2}}-z_l+\sqrt{2}\varrho\right)\right]^{2k},
\end{aligned}
$$

which are all quantities independent of r and z_{i+1}. □

The previous lemma allows us to find the following bound.

Lemma 4.17 (Case A_{N-1}): For every $\varepsilon>0$ and $x\in\partial W_{A_{N-1}}$ there exists a constant $C_3=C_3(N,x,k,\varepsilon)>0$ such that

$$\mathbb{P}^x\big(\exists s\in[t_1,t_2]: Y_s\in E^r_{A_{N-1}}\big)\leq C_3(t_2-t_1)^{k+\frac{1}{2}}$$

for every $t_2>t_1\geq\varepsilon$ with $0<r\leq\sqrt{t_2-t_1}$.

Proof: Since we have already carried out the proof in detail for the classical Bessel process in the previous section, see Lemmas 4.9 and 4.10, we now abbreviate it significantly. The main difference is the more complicated set E^r, therefore we define the stopping time alternatively as follows

$$T := \inf\left\{s \geq t_1 : Y_t \in E^r\right\}$$

so that

$$\mathbb{P}^x\big(T \leq t_2\big) = \mathbb{P}^x\big(\exists s \in [t_1, t_2] : Y_s \in E^r\big).$$

The further calculations are analogous. We can easily derive

$$\mathbb{P}^x\big(\exists s \in [t_1, t_2] : Y_s \in E^r\big) \leq \frac{\int\limits_{t_1}^{2t_2 - t_1} \mathbb{P}^x\big(Y_s \in E^r\big)\,\mathrm{d}s}{\inf\limits_{y \in E^r} \int\limits_0^{t_2 - t_1} \mathbb{P}^y\big(Y_u \in E^r\big)\,\mathrm{d}u} \tag{4.21}$$

and then use the $\frac{1}{2}$-semi stability to apply Lemma 4.16 with $\varrho := \frac{\|x\|}{\sqrt{\varepsilon}} \geq \frac{\|x\|}{\sqrt{t_1}} \geq \frac{\|x\|}{\sqrt{s}}$. The additional dependence on ϱ or, in particular on x, in the upper bound of Lemma 4.16 does not cause any problems since we only take an infimum in the denominator. In case of a dependence only on the starting point x, the constant becomes dependent on $\frac{x}{\sqrt{s}}$ and we cannot take it out of the integral. Here, we immediately see the importance of the additional constant ϱ and that it enables more applications, seeing that we can always set $\varrho := \|x\|$. □

Remark: In particular, (4.21) is true for every multivariate Bessel process for any set $E^r \subset \overline{W}_R$, e.g. E^r_R, since we just need the strong Markov property. The main work lies in Lemma 4.17. Thus, to prove this upper bound of the Hausdorff dimension for a different root system, we need an analogous lemma with a suitable edge set $E^r_R \subset \overline{W}_R$ such that $\partial W_R \subset E^r_R$ holds while the power of r in the lemma determines the upper bound. The easiest solution for later calculations, especially for the lower bound of the Hausdorff dimension, is if the constants are independent of the starting point x, however it is sufficient if only C_1 is independent of x and C_2 depends on ϱ additionally.

By means of Lemma 4.17, the proof of the upper bound is analogous to the one for the classical Bessel process, see Theorem 4.11.

Theorem 4.18 (Case A_{N-1}): For every $0 < k < \frac{1}{2}$ and $x \in \partial W_{A_{N-1}}$ the inequality

$$\dim\Big(Y^{-1}\big(\partial W_{A_{N-1}}\big)\Big) \leq \frac{1}{2} - k$$

is $\mathbb{P}^x$-almost surely valid.

4.3.2 Lower bound

In this subsection, we focus on proving the following statement, which completes the proof of Theorem 4.15.

Theorem 4.19 (Case $\boldsymbol{A_{N-1}}$): For every $0 < k < \frac{1}{2}$ and $x \in \partial W_{A_{N-1}}$ the inequality

$$\dim\Big(Y^{-1}\big(\partial W_{A_{N-1}}\big)\Big) \geq \frac{1}{2} - k$$

is $\mathbb{P}^x$-almost surely valid.

If we could manage the proof of Lemma 4.16 so that C_2 is independent of x and its norm's bound ϱ, we would not have to do any further work. In that case, the proof would proceed in the same way as for the classical Bessel process. This small detail causes that we have to determine one more bound. We first derive the following auxiliary bound.

Lemma 4.20: For $0 < \varepsilon \leq s_1 < s_2$ and $n \in \mathbb{N}$ we have

$$\mathbb{P}^x\big(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in E^{\frac{1}{n}}\big) \leq c_k^{-2}\left(\frac{s_2 - s_1}{s_2}\right)^{\frac{N}{2}+\kappa} \int\limits_{E^{\sqrt{\frac{s_2}{n^2 s_1(s_2-s_1)}}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|u\|^2}{2}-\frac{\|v\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|v\|\right)} w(v)w(u)\,\mathrm{d}v\,\mathrm{d}u,$$

where in particular κ is the sum of multiplicities as defined in Section 1.1.

Proof: We start from

$$\mathbb{P}^x\left(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in E^{\frac{1}{n}}\right) = \int\limits_{E^{\frac{1}{n}}} Q_k\Big(s_2 - s_1, y, E^{\frac{1}{n}}\Big) Q_k\big(s_1, x, \mathrm{d}y\big)$$

and first rewrite this term

$$
\begin{aligned}
&\mathbb{P}^x\big(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in E^{\frac{1}{n}}\big) \\
&\quad = c_k^{-2} \int\limits_{E^{\frac{1}{n}}} \int\limits_{E^{\frac{1}{n}}} \frac{e^{-\frac{\|y\|^2+\|z\|^2}{2(s_2-s_1)}}}{(s_2-s_1)^{\frac{N}{2}}} J_k\left(\frac{y}{\sqrt{s_2-s_1}}, \frac{z}{\sqrt{s_2-s_1}}\right) w\left(\frac{z}{\sqrt{s_2-s_1}}\right) \mathrm{d}z \\
&\qquad\qquad \cdot \frac{e^{-\frac{\|x\|^2+\|y\|^2}{2s_1}}}{s_1^{\frac{N}{2}}} J_k\left(\frac{x}{\sqrt{s_1}}, \frac{y}{\sqrt{s_1}}\right) w\left(\frac{y}{\sqrt{s_1}}\right) \mathrm{d}y \\
&\quad = c_k^{-2} \int\limits_{E^{\frac{1}{n}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|y\|^2}{2(s_2-s_1)}-\frac{\|v\|^2}{2}} J_k\left(\frac{y}{\sqrt{s_2-s_1}}, v\right) w(v)\, \mathrm{d}v \\
&\qquad\qquad \cdot \frac{e^{-\frac{\|x\|^2+\|y\|^2}{2s_1}}}{s_1^{\frac{N}{2}}} J_k\left(\frac{x}{\sqrt{s_1}}, \frac{y}{\sqrt{s_1}}\right) w\left(\frac{y}{\sqrt{s_1}}\right) \mathrm{d}y.
\end{aligned}
$$

In the last step, we performed the substitution $z = \sqrt{s_2 - s_1}v$. Using $J_k(x, y) \le e^{\|x\|\|y\|}$ yields

$$
\begin{aligned}
&\mathbb{P}^x\big(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in E^{\frac{1}{n}}\big) \\
&\quad \le \frac{1}{c_k^2 s_1^{\frac{N}{2}}} \int\limits_{E^{\frac{1}{n}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|y\|^2}{2(s_2-s_1)}-\frac{\|v\|^2}{2}+\frac{\|v\|\|y\|}{\sqrt{s_2-s_1}}} e^{-\frac{\|x\|^2+\|y\|^2}{2s_1}+\frac{\|x\|\|y\|}{s_1}} \\
&\qquad\qquad \cdot w(v)\, w\left(\frac{y}{\sqrt{s_1}}\right) \mathrm{d}v\, \mathrm{d}y \\
&\quad = \frac{e^{-\frac{\|x\|^2}{2s_1}}}{c_k^2 s_1^{\frac{N}{2}}} \int\limits_{E^{\frac{1}{n}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|y\|^2}{2}\cdot\frac{s_2}{s_1(s_2-s_1)}-\frac{\|v\|^2}{2}+\|y\|\left(\frac{\|x\|}{s_1}+\frac{\|v\|}{\sqrt{s_2-s_1}}\right)} \\
&\qquad\qquad \cdot w(v)\, w\left(\frac{y}{\sqrt{s_1}}\right) \mathrm{d}v\, \mathrm{d}y \\
&\quad \le \frac{1}{c_k^2 s_1^{\frac{N}{2}}} \int\limits_{E^{\frac{1}{n}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|y\|^2}{2}\cdot\frac{s_2}{s_1(s_2-s_1)}-\frac{\|v\|^2}{2}+\|y\|\left(\frac{\|x\|}{s_1}+\frac{\|v\|}{\sqrt{s_2-s_1}}\right)} \\
&\qquad\qquad \cdot w(v)\, w\left(\frac{y}{\sqrt{s_1}}\right) \mathrm{d}v\, \mathrm{d}y.
\end{aligned}
$$

Next, we substitute $y = \sqrt{\frac{s_1(s_2-s_1)}{s_2}}u$ to obtain

$$
\begin{aligned}
&\mathbb{P}^x\big(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in E^{\frac{1}{n}}\big) \\
&\quad \leq c_k^{-2}\left(\frac{s_2-s_1}{s_2}\right)^{\frac{N}{2}} \int\limits_{E^{\sqrt{\frac{s_2}{n^2 s_1(s_2-s_1)}}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|u\|^2}{2}-\frac{\|v\|^2}{2}+\sqrt{\frac{s_1(s_2-s_1)}{s_2}}\|u\|\left(\frac{\|x\|}{s_1}+\frac{\|v\|}{\sqrt{s_2-s_1}}\right)} \\
&\qquad\qquad \cdot w(v)w\left(\sqrt{\frac{s_2-s_1}{s_2}}u\right) \mathrm{d}v\,\mathrm{d}u \\
&\quad = c_k^{-2}\left(\frac{s_2-s_1}{s_2}\right)^{\frac{N}{2}+\kappa} \int\limits_{E^{\sqrt{\frac{s_2}{n^2 s_1(s_2-s_1)}}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|u\|^2}{2}-\frac{\|v\|^2}{2}+\|u\|\left(\sqrt{\frac{s_2-s_1}{s_1 s_2}}\|x\|+\sqrt{\frac{s_1}{s_2}}\|v\|\right)} \\
&\qquad\qquad \cdot w(v)w(u)\,\mathrm{d}v\,\mathrm{d}u \\
&\quad \leq c_k^{-2}\left(\frac{s_2-s_1}{s_2}\right)^{\frac{N}{2}+\kappa} \int\limits_{E^{\sqrt{\frac{s_2}{n^2 s_1(s_2-s_1)}}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|u\|^2}{2}-\frac{\|v\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|v\|\right)} w(v)w(u)\,\mathrm{d}v\,\mathrm{d}u.
\end{aligned}
$$

The equality holds due to $w(cy) = c^{2\kappa}w(y)$ for any constant $c > 0$. In the last line we used $\varepsilon \leq s_1 < s_2$. □

Besides, Lemma 4.20 holds for every multivariate Bessel processes meaning that this can be helpful when dealing with Conjecture 4.13 for arbitary root systems R in more detail. Now, we specify this bound further for the A_{N-1} case. For the B_N case, we elaborate on this in Appendix A.

Lemma 4.21 (Case A_{N-1}): For $0 < \varepsilon \leq s_1 < s_2$ and $n \in \mathbb{N}$ there exists a constant $C(x, \varepsilon, N, k) > 0$ such that

$$
\mathbb{P}^x\Big(Y_{s_1} \in E^{\frac{1}{n}}_{A_{N-1}}, Y_{s_2} \in E^{\frac{1}{n}}_{A_{N-1}}\Big) \leq \frac{C(x,\varepsilon,N,k)}{n^{4k+2}(s_2-s_1)^{k+\frac{1}{2}} s_1^{k+\frac{1}{2}}}
$$

is valid.

Proof: We start the proof with the bound for this probability given in Lemma 4.20. For $r > 0$ we recall

$$E^r_{A_{N-1}} = \bigcup_{i=1}^{N-1} E^{r,i}_{A_{N-1}} = \bigcup_{i=1}^{N-1} \left\{y \in \overline{W}_{A_{N-1}} : 0 \leq y_{i+1} - y_i \leq r\right\}$$

so the inner integral over v from Lemma 4.20 is bounded above by the sum over $i \in \{1, \ldots, N-1\}$ of

$$\int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}},i}_{A_{N-1}}} e^{-\frac{\|u\|^2}{2} - \frac{\|v\|^2}{2} + \|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}} + \|v\|\right)} w_{A_{N-1}}(v)\,\mathrm{d}v. \tag{4.22}$$

For the moment, we disregard the term $w_{A_{N-1}}(u)$ since it is not important for bounding the integral over v. For each i we perform the substitution

$$\begin{aligned} z_j &= v_j && \text{for } j \notin \{i, i+1\}, \\ z_i &= \tfrac{v_{i+1}+v_i}{\sqrt{2}}, \\ z_{i+1} &= \tfrac{v_{i+1}-v_i}{\sqrt{2}}, \end{aligned}$$

a simple rotation to preserve norms and receive the distance between the corresponding particles that are about to collide. In particular, observing

$$\begin{aligned} z_i - z_{i+1} &= \sqrt{2}v_i, \\ z_i + z_{i+1} &= \sqrt{2}v_{i+1}, \\ \sqrt{2}z_{i+1} &= v_{i+1} - v_i, \end{aligned}$$

we rewrite and bound the weight function in terms of z:

$$\begin{aligned} w_{A_{N-1}}(v) &= \prod_{1 \leq l < m \leq N} (v_m - v_l)^{2k} \\ &= 2^k z_{i+1}^{2k} \prod_{\substack{1 \leq l < m \leq N \\ m,l \neq i,i+1}} (z_m - z_l)^{2k} \prod_{l=1}^{i-1} \left(\frac{z_i - z_{i+1}}{\sqrt{2}} - z_l\right)^{2k} \left(\frac{z_i + z_{i+1}}{\sqrt{2}} - z_l\right)^{2k} \\ &\qquad \prod_{m=i+2}^{N} \left(z_m - \frac{z_i - z_{i+1}}{\sqrt{2}}\right)^{2k} \left(z_m - \frac{z_i + z_{i+1}}{\sqrt{2}}\right)^{2k} \end{aligned}$$

$$
\begin{aligned}
&= 2^k z_{i+1}^{2k} \prod_{\substack{1\le l<m\le N\\ m,l\neq i,i+1}} (z_m - z_l)^{2k} \prod_{l=1}^{i-1} \left(\left(\frac{z_i}{\sqrt{2}} - z_l \right)^2 - \frac{z_{i+1}^2}{2} \right)^{2k} \\
&\qquad \prod_{m=i+2}^{N} \left(\left(z_m - \frac{z_i}{\sqrt{2}} \right)^2 - \frac{z_{i+1}^2}{2} \right)^{2k} \\
&\le 2^k z_{i+1}^{2k} \prod_{\substack{1\le l<m\le N\\ m,l\neq i,i+1}} (z_m - z_l)^{2k} \prod_{l=1}^{i-1} \left(\frac{z_i}{\sqrt{2}} - z_l \right)^{4k} \prod_{m=i+2}^{N} \left(z_m - \frac{z_i}{\sqrt{2}} \right)^{4k} \\
&=: 2^k z_{i+1}^{2k} \Pi_{N,k,i}(z_{i+1}^*).
\end{aligned}
$$

Next, we focus on the integration domain. Especially, we observe how the ordering of v transfers to z:

$$
\begin{aligned}
z_j &\le z_{j+1}, && \text{for } j \notin \{i-1, i, i+1\}, \\
z_{i-1} = v_{i-1} &\le v_i = \frac{z_i - z_{i+1}}{\sqrt{2}}, \\
\frac{z_i + z_{i+1}}{\sqrt{2}} = v_{i+1} &\le v_{i+2} = z_{i+2},
\end{aligned}
$$

which imposes the definition

$$
\Lambda_{i,N} := \left\{ z \in \mathbb{R}^N : z_j \le z_{j+1} \ j \notin \{i-1,i,i+1\}, z_{i-1} + \frac{z_{i+1}}{\sqrt{2}} \le \frac{z_i}{\sqrt{2}} \le z_{i+2} - \frac{z_{i+1}}{\sqrt{2}} \right\}.
$$

The additional condition $v \in E_{A_{N-1}}^{\frac{1}{n\sqrt{s_2-s_1}},i}$ indicates

$$
0 \le z_{i+1} \le \frac{1}{n\sqrt{2(s_2-s_1)}}.
$$

Now, we want to have the terms in the exponential function in (4.22) be independent of z_{i+1}. For this purpose, we recall z_{i+1}^*, (4.17), and introduce the vector $\psi(z_{i+1}^*)$, which differs just in the $(i+1)$th component of z, by $\left(\psi(z_{i+1}^*)\right)_{i+1} := z_i - \sqrt{2} z_{i-1}$. The inequality

$$
\|z_{i+1}^*\| \le \|z\| \le \|\psi(z_{i+1}^*)\|
$$

holds for every z within the new integration domain

$$
\Lambda_{i,N} \cap \left\{ 0 \le z_{i+1} \le \frac{1}{n\sqrt{2(s_2-s_1)}} \right\}.
$$

Since the variable substitution $v \to z$ is a rotation, norms are preserved and thus we can proceed to bound the integral as follows:

$$\int\limits_{E_{A_{N-1}}^{\frac{1}{n\sqrt{s_2-s_1}},i}} e^{-\frac{\|u\|^2}{2}-\frac{\|v\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|v\|\right)} w_{A_{N-1}}(v)\,\mathrm{d}v$$

$$\leq \int\limits_{\Lambda_{i,N}\cap\left\{0\leq z_{i+1}\leq\frac{1}{n\sqrt{2(s_2-s_1)}}\right\}} e^{-\frac{\|u\|^2}{2}-\frac{\|z_{i+1}^*\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|\psi(z_{i+1}^*)\|\right)} 2^k z_{i+1}^{2k}\Pi_{N,k,i}(z_{i+1}^*)\,\mathrm{d}(z_{i+1}^*, z_{i+1}).$$

In a next step, we make the constraints in $\Lambda_{i,N}$ independent of z_{i+1}. Due to the positivity of z_{i+1} we receive

$$z_{i-1} \leq z_{i-1} + \frac{z_{i+1}}{\sqrt{2}} \leq \frac{z_i}{\sqrt{2}} \leq z_{i+2} - \frac{z_{i+1}}{\sqrt{2}} \leq z_{i+2}$$

and hence the set

$$\widetilde{\Lambda}_{i,N} := \left\{z \in \mathbb{R}^N : z_j \leq z_{j+1}\ j \notin \{i-1, i, i+1\}, z_{i-1} \leq \frac{z_i}{\sqrt{2}} \leq z_{i+2}\right\}$$

fulfills

$$\widetilde{\Lambda}_{i,N} \cap \{0 \leq z_{i+1}\} \supset \Lambda_{i,N} \cap \{0 \leq z_{i+1}\}.$$

Moreover, we can immediately derive as upper bound of (4.22):

$$\frac{1}{\sqrt{2}n^{2k+1}(s_2-s_1)^{k+\frac{1}{2}}} \int\limits_{p_{i+1}(\widetilde{\Lambda}_{i,N})} e^{-\frac{\|u\|^2}{2}-\frac{\|z_{i+1}^*\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|\psi(z_{i+1}^*)\|\right)}\Pi_{N,k,i}(z_{i+1}^*)\,\mathrm{d}z_{i+1}^*.$$

We note that this expression depends on the index $i = 1, \ldots, N-1$. Having successfully dealt with the inner integral over v, we can now analogously derive a bound and decomposition of the integral over u, this time using the index $j = 1, \ldots, N-1$. We denote this resulting integral, depending on i and j, by C_{ij}. Hence, we receive

$$\frac{C_{ij}(x,\varepsilon,N,k)}{\sqrt{2}n^{2k+1}(s_2-s_1)^{k+\frac{1}{2}}} \cdot \frac{s_2^{k+\frac{1}{2}}}{\sqrt{2}n^{2k+1}s_1^{k+\frac{1}{2}}(s_2-s_1)^{k+\frac{1}{2}}}$$

$$=\frac{C_{ij}(x,\varepsilon,N,k)s_2^{k+\frac{1}{2}}}{2n^{4k+2}s_1^{k+\frac{1}{2}}(s_2-s_1)^{2k+1}}.$$

To obtain a bound on the whole integral, it suffices to sum over both i and j and finally adding the prefactor from Lemma 4.20 we yield

$$\begin{aligned}
&\mathbb{P}^x\Big(Y_{s_1}\in E^{\frac{1}{n}}_{A_{N-1}}, Y_{s_2}\in E^{\frac{1}{n}}_{A_{N-1}}\Big)\\
&\qquad\le \sum_{i,j=1}^{N-1} C_{ij}(x,\varepsilon,N,k)\frac{1}{2c_k^2}\underbrace{\left(\frac{s_2-s_1}{s_2}\right)^{\frac{(kN+1)(N-1)}{2}-k}}_{\le 1}\frac{1}{n^{4k+2}(s_2-s_1)^{k+\frac{1}{2}}s_1^{k+\frac{1}{2}}}\\
&\qquad\le \sum_{i,j=1}^{N-1} C_{ij}(x,\varepsilon,N,k)\frac{1}{2c_k^2}\cdot\frac{1}{n^{4k+2}(s_2-s_1)^{k+\frac{1}{2}}s_1^{k+\frac{1}{2}}}\\
&\qquad=: C(x,\varepsilon,N,k)\frac{1}{n^{4k+2}(s_2-s_1)^{k+\frac{1}{2}}s_1^{k+\frac{1}{2}}}
\end{aligned}$$

with $\varepsilon\le s_1<s_2$. The dominant negative term in the exponential function of the integrand ensures that the constants C_{ij} are finite and thus the same applies to C. □

Finally, we can give the proof for the Hausdorff dimension's lower bound, Theorem 4.19, which completes Theorem 4.18.

Proof of Theorem 4.19: The idea of the proof remains precisely as for the classical Bessel process. The main differences are the additional calculations in the previous Lemmas 4.20 and 4.21, we found necessary for the following. This is a substitute for the upper bound from Lemma 4.16 since the constant depends on the starting point. Before starting with the calculations, we repeat the setting within the proof of Theorem 4.11. Therefore, we consider an interval $[t_1,t_2]\subset(0,\infty)$ for verifying

$$\dim\big(Y^{-1}(\partial W_{A_{N-1}})\cap[t_1,t_2]\big)\ge\frac{1}{2}-k.$$

The proof relies on Lemma 4.7. As Lemma 4.21 can not be applied in the case $t_1=0$, we required Lemma 4.7 to extend to $t_1>0$, which is a key difference to the proof for the classical Bessel process, Theorem 4.11. Similarly, we construct for every possible smaller dimension $0<\beta<\frac{1}{2}-k$ a positive measure μ on $Y^{-1}\big(\partial W_{A_{N-1}}\big)\cap[t_1,t_2]$ such that $\|\mu\|_{\beta,[t_1,t_2]}$ is finite, which implies $\dim\big(Y^{-1}(\partial W)\cap[t_1,t_2]\big)>\beta$ on $\{\mu>0\}$ due to the capacity argument, Lemma 4.6. We define an analogous sequence of random positive measures on $\mathcal{B}([0,\infty))$ by

$$\mu_n(B,\omega) := n^{2k+1} \int\limits_{[t_1,t_2]\cap B} \mathbb{1}_{\left\{Y_s(\omega)\in E^{\frac{1}{n}}\right\}} \,\mathrm{d}s$$

and are again searching for constants $K_1 > 0$ and $K_2 > 0$ such that Lemma 4.7 is applicable. By using the $\frac{1}{2}$-semi stability of the multivariate Bessel process and Lemma 4.14 we obtain

$$\begin{aligned}
\mathbb{E}^x(\|\mu_n\|) &= n^{2k+1} \int\limits_{t_1}^{t_2} \mathbb{P}^x\big(Y_s \in E^{\frac{1}{n}}\big) \,\mathrm{d}s \\
&= n^{2k+1} \int\limits_{t_1}^{t_2} Q_k\big(s, x, E^{\frac{1}{n}}\big) \,\mathrm{d}s \\
&= n^{2k+1} \int\limits_{t_1}^{t_2} Q_k\left(1, \frac{x}{\sqrt{s}}, E^{\frac{1}{n\sqrt{s}}}\right) \mathrm{d}s \\
&\overset{4.16}{\geq} C_1(k,N) n^{2k+1} n^{-2k-1} \int\limits_{t_1}^{t_2} s^{-k-\frac{1}{2}} \,\mathrm{d}s \\
&= \frac{C_1(k,N)}{\frac{1}{2}-k} \left(t_2^{\frac{1}{2}-k} - t_1^{\frac{1}{2}-k} \right) =: K_1(t_1,t_2,N,k) > 0.
\end{aligned}$$

Furthermore, we calculate the following expressions from Lemma 4.21:

$$\begin{aligned}
\mathbb{E}^x\big(\|\mu_n\|^2\big) &= n^{4k+2} \int\limits_{t_1}^{t_2}\int\limits_{t_1}^{t_2} \mathbb{P}^x\left(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in \mathbb{E}^{\frac{1}{n}}\right) \mathrm{d}s_2 \,\mathrm{d}s_1 \\
&\leq 2n^{4k+2} \int\limits_{t_1}^{t_2}\int\limits_{s_1}^{t_2} \frac{C(x,t_1,N,k)}{n^{4k+2}(s_2-s_1)^{k+\frac{1}{2}} s_1^{k+\frac{1}{2}}} \,\mathrm{d}s_2 \,\mathrm{d}s_1 \\
&= 2C(x,t_1,N,k) \int\limits_{t_1}^{t_2}\int\limits_{s_1}^{t_2} (s_2-s_1)^{-k-\frac{1}{2}} s_1^{-k-\frac{1}{2}} \,\mathrm{d}s_2 \,\mathrm{d}s_1 \\
&= \frac{2C(x,t_1,N,k)}{\frac{1}{2}-k} \int\limits_{t_1}^{t_2} (t-s_1)^{\frac{1}{2}-k} s_1^{-k-\frac{1}{2}} \,\mathrm{d}s_1
\end{aligned}$$

$$\leq \frac{2C(x,t_1,N,k)}{\frac{1}{2}-k} t_2^{\frac{1}{2}-k} \int_0^{t_2} s_1^{-k-\frac{1}{2}} \, \mathrm{d}s_1$$
$$= \frac{2C(x,t_1,N,k)}{\left(\frac{1}{2}-k\right)^2} t_2^{1-2k} =: K_2(x,t_1,t_2,N,k).$$

This object is clearly positive and finite for $t_1 > 0$ and $0 < k < \frac{1}{2}$. Now, we turn to the β-capacity for $0 < \beta < \frac{1}{2} - k$ and we again make use of Lemma 4.21:

$$\begin{aligned}
\mathbb{E}^x\left(\|\mu_n\|_{\beta,[t_1,t_2]}\right) &= \mathbb{E}^x\left(\int_{t_1}^{t_2}\int_{t_1}^{t_2} \frac{\mu(\mathrm{d}s_1)\mu(\mathrm{d}s_2)}{|s_2-s_1|^\beta}\right) \\
&= 2n^{4k+2} \int_{t_1}^{t_2}\int_{s_1}^{t_2} \frac{\mathbb{P}^x\left(Y_{s_1} \in E^{\frac{1}{n}}, Y_{s_2} \in E^{\frac{1}{n}}\right)}{(s_2-s_1)^\beta} \, \mathrm{d}s_2 \, \mathrm{d}s_1 \\
&\leq 2C(x,t_1,N,k) \int_{t_1}^{t_2}\int_{s_1}^{t_2} (s_2-s_1)^{-k-\beta-\frac{1}{2}} s_1^{-k-\frac{1}{2}} \, \mathrm{d}s_2 \, \mathrm{d}s_1 \\
&= \frac{2C(x,t_1,N,k)}{\frac{1}{2}-k-\beta} \int_{t_1}^{t_2} (t-s_1)^{\frac{1}{2}-k-\beta} s_1^{-k-\frac{1}{2}} \, \mathrm{d}s_1 \\
&\leq \frac{2C(x,t_1,N,k) t_2^{\frac{1}{2}-k-\beta}}{\frac{1}{2}-k-\beta} \int_{t_1}^{t_2} s_1^{-k-\frac{1}{2}} \, \mathrm{d}s_1 \\
&= \frac{2C(x,t_1,N,k) t_2^{\frac{1}{2}-k-\beta}}{\left(\frac{1}{2}-k-\beta\right)\left(\frac{1}{2}-k\right)} \left(t_2^{\frac{1}{2}-k} - t_1^{\frac{1}{2}-k}\right) \\
&\leq \frac{2C(x,t_1,N,k) t_2^{1-2k-\beta}}{\left(\frac{1}{2}-k-\beta\right)\left(\frac{1}{2}-k\right)} < \infty.
\end{aligned}$$

Hence, we have proved for every $0 < t_1 < t_2 < \infty$ that

$$\mathbb{P}^x\Big(\dim\left(Y^{-1}(\partial W_{A_{N-1}}) \cap [t_1,t_2]\right) \geq \frac{1}{2} - k\Big) \geq \frac{K_1^2(t_1,t_2,N,k)}{2K_2^2(x,t_1,t_2,N,k)}. \tag{4.23}$$

We define the stopping time τ_n

$$\tau_n := \inf\left\{t \geq \tau_{n-1} + 1 : Y_t \in \partial W_{A_{N-1}}\right\}$$

with $\tau_0 := 0$. By the strong Markov property, $Y_{\tau_n} \in \partial W_{A_{N-1}}$ almost surely and by (4.23) we determine

$$\begin{aligned}
&\mathbb{P}^x\Big(\dim\big(Y^{-1}(\partial W_{A_{N-1}})\cap[0,\tau_{n+1}]\big)\geq\frac{1}{2}-k\,\Big|\,\mathcal{F}_{\tau_n}\Big)\\
&\geq\mathbb{P}^x\Big(\dim\big(Y^{-1}(\partial W_{A_{N-1}})\cap[\tau_n+\varepsilon,\tau_{n+1}]\big)\geq\frac{1}{2}-k\,\Big|\,\mathcal{F}_{\tau_n}\Big)\\
&=\mathbb{P}^{Y_{\tau_n}}\Big(\dim\big(Y^{-1}(\partial W_{A_{N-1}})\cap[\varepsilon,\tau_{n+1}-\tau_n]\big)\geq\frac{1}{2}-k\Big)\\
&\geq\mathbb{P}^{Y_{\tau_n}}\Big(\dim\big(Y^{-1}(\partial W)\cap[\varepsilon,1]\big)\geq\frac{1}{2}-k\Big)\geq\frac{K_1^2(t_1,t_2,N,k)}{2K_2(x,t_1,t_2,N,k)}
\end{aligned}$$

for some $0<\varepsilon<1$, which we need to add in the multivariate case, due to the positivity of t_1 in the calculations above. From here, we omit the rest of the proof since it works exactly as in the classical case with small modifications as well as the proof of the analogous result to Lemma 4.12. □

Looking at the results of the first hitting time, we observe that for any $\alpha\in R$ with $k(\alpha)<\frac{1}{2}$ the time such that $\langle\alpha,Y\rangle=0$ holds is almost surely finite, [26, Proposition 1]. This result transfers in the A_{N-1} case to the sets

$$\partial W^i_{A_{N-1}}:=\big\{x\in\partial W_{A_{N-1}}\,|\,x_{i+1}=x_i\big\}.$$

On closer inspection, we find that we have performed all the proofs on the sets $E^{r,i}_{A_{N-1}}\subset\partial W^i_{A_{N-1}}$. In particular, we have already proved all the corresponding lemmas to derive the following corollary.

Corollary 4.22: The Hausdorff dimension of collision times for the multivariate Bessel process of type A_{N-1} starting in $x\in\overline{W}_{A_{N-1}}$ is given by

$$\dim\big(Y^{-1}(\partial W^i_{A_{N-1}})\big)=\frac{1}{2}-\min\Big\{\frac{1}{2},k\Big\}$$

$\mathbb{P}^x$-almost surely.

If we would have proved Conjecture 4.13, we suppose that likewise the following ensues.

Conjecture 4.23: For every multivariate Bessel process $(Y_t)_{t\geq0}$ starting in $x\in\overline{W}_R$ with multiplicity function $k:R\to(0,\infty)$ the equality

$$\dim\Big(Y^{-1}\big(\partial W^\alpha_R\big)\Big)=\frac{1}{2}-\min\Big\{\frac{1}{2},k(\alpha)\Big\}$$

with

$$\partial W_R^\alpha := \big\{ y \in \partial W_R \,|\, \langle \alpha, Y \rangle = 0 \big\}$$

is $\mathbb{P}^x$-almost surely valid for every $\alpha \in R$.

A Hausdorff dimension of the multivariate Bessel process of type B_N

We could not fully prove Conjecture 4.13 for case B_N by now. Therefore, we now elaborate on the results so far and the problems that have arisen. In Section 4.3 we have proven this result for the A_{N-1} case, which is a quite similar model. The main differences are that the process of type B_N lives in the positive half-line and there exists an additional repulsion from the origin with a different strength, hence this process is indexed by a two dimensional parameter $k = (k_1, k_2) \in (0, \infty)^2$ given by

$$\begin{cases} \mathrm{d}Y_{t,i}^B &= \mathrm{d}B_{t,i} + k_1 \frac{1}{Y_{t,i}^B}\,\mathrm{d}t + k_2 \sum\limits_{\substack{j=1\\ j\neq i}}^{N} \left(\frac{1}{Y_{t,i}^B - Y_{t,j}^B} + \frac{1}{Y_{t,i}^B + Y_{t,j}^B} \right) \mathrm{d}t, \\ Y_0^B &= y \in \overline{W}_{B_N} \end{cases}$$

for $i = 1, \dots, N$ with a standard multivariate Brownian motion $(B_t)_{t\geq 0}$ living on the closure of the Weyl chamber $W_{B_N} := \{x \in \mathbb{R}^N \,|\, 0 < x_1 < \cdots < x_N\}$. We assume $N \geq 2$ for the calculations, seeing that $N = 1$ is a classical Bessel process. As an auxiliary tool we define here, analogously to the A_{N-1} case, the edge set

$$E_{B_N}^r := \Big\{ x \in \overline{W}_{B_N} \,|\, \exists\, i \in \{1, \dots, N-1\} : 0 \leq x_{i+1} - x_i \leq r \Big\} \cup \Big\{ x \in \overline{W}_{B_N} \,|\, 0 \leq x_1 \leq r \Big\}$$

of thickness $r \geq 0$ such that $\partial W_{B_N} = E_{B_N}^0 \subset E_{B_N}^r$. After a simple calculation it is easily seen that $E_{B_N}^{rc} = cE_{B_N}^r$ still works for each $r, c > 0$.

Lemma A.1 (Case B_N): For every $k = (k_1, k_2) \in (0, \infty)^2$ and $\varrho > r > 0$ there exist constants $C_1 = C_1(k, N, \varrho)$ and $C_2 = C_2(k, N, \varrho) > 0$ such that

$$C_1 r^{2\min\{k_1,k_2\}+1} \leq Q_k\big(1, x, E_{B_N}^r\big) \leq C_2 \min\big\{1, r^{2\min\{k_1,k_2\}+1}\big\}$$

holds for all $x \in E_{B_N}^r \cap B(0, \varrho)$.

Proof: Upper bound: First of all, the case $r \geq 1$ is trivial when choosing $C_2 \geq 1$. Therefore, we focus on the case $r < 1$. For this case we consider the variable transformation

$$z_0 = y_1,\ z_i = y_{i+1} - y_i$$

or equivalently

$$y_1 = z_0,\ y_i = \sum_{j=0}^{i} z_j$$

for $i = 1, \ldots, N-1$. Therefore, this simple transformation leads to a classification: $z \in E^r_{B_N}$ implies there exists an $i \in \{1, \ldots, N\}$ such that $0 \leq z_i \leq r$. The corresponding Jacobian is 1 and the weight function becomes

$$\begin{aligned}
w_{B_N}(y) &= \prod_{m=1}^{N} y_m^{2k_1} \prod_{1\leq m<n\leq N} \big(y_n - y_m\big)^{2k_2} \prod_{1\leq m<n\leq N} \big(y_n + y_m\big)^{2k_2} \\
&= \prod_{m=0}^{N-1} \Big(\sum_{l=0}^{m} z_l\Big)^{2k_1} \prod_{0\leq m<n\leq N-1} \Big(\sum_{l=m+1}^{n} z_l\Big)^{2k_2} \\
&\qquad \prod_{1\leq m<n\leq N} \Big(\sum_{l=1}^{m} z_l + \sum_{l=1}^{n} z_l\Big) \\
&= z_0^{2k_1} \prod_{m=1}^{N-1} \Big(\sum_{l=0}^{m} z_m\Big)^{2k_1} \prod_{m=1}^{N} z_m^{2k_2} \prod_{1\leq m<n\leq N-1} \Big(\sum_{l=m}^{n} z_l\Big)^{2k_2} \\
&\qquad \prod_{1\leq m<n\leq N} \Big(2\sum_{l=1}^{m} z_l + \sum_{l=m+1}^{n} z_l\Big) \\
&=: \Pi_{N,k}(z) z_0^{2k_1} \prod_{m=1}^{N} z_m^{2k_2}.
\end{aligned}$$

We later will use the inequalities

$$\Pi_{N,k}(z)\Big|_{z_i=0} \leq \Pi_{N,k}(z) \leq \Pi_{N,k}(z)\Big|_{z_i=\varrho}$$

for all $i = 0, \ldots, N-1$ to bound the weight function. The norm of y is rewritten as

$$\|y\|^2 = \sum_{m=0}^{N-1} \Big(\sum_{n=0}^{m} z_n\Big)^2.$$

Now, we proceed to derive

$$\begin{aligned}Q_k(1,x,E^r_{B_N}) &= c_k^{-1} \int\limits_{E^r_{B_N}} e^{-\frac{\|x\|^2+\|y\|^2}{2}} J_k(x,y) w_{B_N}(y)\,\mathrm{d}y \\ &\le c_k^{-1} \int\limits_{E^r_{B_N}} e^{-\frac{\|y\|^2}{2}+\|x\|\|y\|} w_{B_N}(y)\,\mathrm{d}y \\ &\le c_k^{-1} \int\limits_{E^r_{B_N}} e^{-\frac{\|y\|^2}{2}+\varrho\|y\|} w_{B_N}(y)\,\mathrm{d}y\end{aligned}$$

since the relation $J_k(x,y) \le e^{\|x\|\|y\|}$ holds while $x \in B(0,\varrho)$. Making use of the transformation and noting

$$\begin{aligned}E^r_{B_N} &= \bigcup_{i=1}^{N-1} \left\{y \in \overline{W}_{B_N} : 0 \le y_{i+1} - y_i \le r\right\} \cup \left\{y \in \overline{W}_{B_N} : 0 \le y_1 \le r\right\} \\ &=: \bigcup_{i=0}^{N-1} E^{r,i}_{B_N},\end{aligned} \tag{A.1}$$

we find that the integral on the right-hand side is bounded above by the sum over $i \in \{0,\dots,N-1\}$ of

$$\int\limits_0^r \int\limits_0^\infty \cdots \int\limits_0^\infty \exp\left(-\frac{1}{2} \sum_{m=0}^{N-1} \left(\sum_{n=0}^{m} z_n \right)^2 \Big|_{z_i=0} + \varrho \sqrt{\sum_{m=0}^{N-1} \left(\sum_{n=0}^{m} z_m \right)^2} \Big|_{z_i=\varrho} \right) \\ \cdot \Big(\Pi_{N,k}(z)\Big)\Big|_{z_i=\varrho} \left(z_0^{2k_1} \prod_{j=1}^{N-1} z_j^{2k_2} \right)\Big|_{z_i=r} \prod_{\substack{j=0\\ j\neq i}}^{N-1} \mathrm{d}z_j\,\mathrm{d}z_i.$$

In the case $i=0$, this integral is equal to $K_0(N,\varrho,k)r^{2k_1+1}$, while for $i>0$ we get $K_i(N,\varrho,k)r^{2k_2+1}$. Obviously, the constants K_i are finite since the second order terms in the exponential have negative coefficients. We now write

$$\begin{aligned}Q_k(1,x,E^r_{B_N}) &\le c_k^{-1}\left(K_0 r^{2k_1+1} + \sum_{i=1}^{N-1} K_i r^{2k_2+1} \right) \\ &\le c_k^{-1}\left(K_0 r^{2\min\{k_1,k_2\}+1} + \sum_{i=1}^{N-1} K_i r^{2\min\{k_1,k_2\}+1} \right)\end{aligned}$$

$$= c_k^{-1}\left(K_0 + \sum_{i=1}^{N-1} K_i\right) r^{2\min\{k_1,k_2\}+1}.$$

Lower bound: The lower bound is derived similarly. Using the lower bound $J_k(x,y) \geq e^{-\|x\|\|y\|}$ as well as $x \in B(0,\varrho)$, we rewrite

$$\begin{aligned} Q_k(1,x,E^r_{B_N}) &= c_k^{-1}\int_{E^r_{B_N}} e^{-\frac{\|x\|^2+\|y\|^2}{2}} J_k(x,y) w_{B_N}(y)\,\mathrm{d}y \\ &\geq c_k^{-1}\int_{E^r_{B_N}} e^{-\frac{(\|y\|+\varrho)^2}{2}} w_{B_N}(y)\,\mathrm{d}y. \end{aligned}$$

Using the same transformation, the integral is bounded below by the sum over $i \in \{0,\dots,N-1\}$ of

$$\begin{aligned} \int_0^r \int_\varrho^\infty \cdots \int_\varrho^\infty \exp\left(-\frac{1}{2}\left(\sqrt{\sum_{m=0}^{N-1}\left(\sum_{n=0}^{m} z_n\right)^2}\Bigg|_{z_i=\varrho} + \varrho\right)^2\right) \\ \cdot \left(\Pi_{N,k}(z)\right)\Big|_{z_i=0}\left(z_0^{2k_1}\prod_{m=1}^{N} z_m^{2k_2}\right)\prod_{\substack{j=0\\ j\neq i}}^{N-1} \mathrm{d}z_j\,\mathrm{d}z_i. \end{aligned}$$

Since $\varrho > r > 0$ we ensure by starting the integrals at ϱ that either only two particles are close to each other or y_1 is close to the origin. The integral over z_i gives a factor of $\frac{r^{2k_1+1}}{2k_1+1}$ for $i=0$ and $\frac{r^{2k_2+1}}{2k_2+1}$ for $i>0$, while the rest of the integrals give constants that depend only on N, ϱ and k. In other words, the multiple integral is equal to $\widetilde{K}_0 r^{2k_1+1}$ for $i=0$ and $\widetilde{K}_i r^{2k_2+1}$ for $i \in \{1,\dots,N-1\}$. Again, all the $\widetilde{K}_i$'s are positive and finite, hence we receive

$$\begin{aligned} Q_k\big(1,x,E^r_{B_N}\big) &\geq c_k^{-1}\left(\widetilde{K}_0 r^{2k_1+1} + \sum_{i=1}^{N-1}\widetilde{K}_i r^{2k_2+1}\right) \\ &\geq c_k^{-1}\sum_{i=1}^{N-1}\widetilde{K}_i r^{2k_2+1} \end{aligned}$$

for $k_1 \geq k_2 > 0$ and

$$Q_k\big(1,x,E^r_{B_N}\big) \geq \frac{1}{c_k}\widetilde{K}_0 r^{2k_1+1}$$

for $0 < k_1 < k_2$. Finally, we can write $C_1 r^{2\min(k_1,k_2)+1} \leq Q_k(1,x,E^r_{B_N})$ with C_1 depending on N, ϱ and k as claimed. □

Even this more restrictive condition on the upper bound would not produce additional problems within the proofs for the Hausdorff dimension even though C_1 needs to be independent of ϱ and x. In the B_N case, it was not possible to make the constant dependent on x but independent of r without introducing ϱ. Here, the cornerstone is $\varrho > r$. To preserve all the essential statements for the proof of the upper bound in Conjecture 4.13 for $R = B_N$, we are only missing either the inequality

$$Q_k\big(1, x, E^r_{B_N}\big) \leq C_1 r^{2\min\{k_1,k_2\}+1}$$

or

$$Q_k\big(1, x, E^r_{B_N}\big) \leq C_1 \min\left\{1, r^{2\min\{k_1,k_2\}+1}\right\}$$

for every $k = (k_1, k_2) \in (0, \infty)^2$ and $x \in E^r_{B_N}$ with $C_1 = C_1(k, N) > 0$. In the proof of the lower bound in Conjecture 4.13, the proved statement, Lemma A.1, is enough even if we are (just) using the lower bound of the previous lemma.

Lemma A.2 (Case B_N): For $0 < \varepsilon \leq s_1 < s_2$ and $n \in \mathbb{N}$, there exists a constant $C(x, \varepsilon, N, k) > 0$ such that

$$\begin{aligned}\mathbb{P}^x\big(X_{s_1} \in E^{\frac{1}{n}}, X_{s_2} \in E^{\frac{1}{n}}\big) \leq{}& \frac{C(x, \varepsilon, N, k)}{n^{4\min\{k_1,k_2\}+2}} \\ &\cdot \big((s_2 - s_1)^{2k_2} + 2(s_2 - s_1)^{k_1+k_2} + (s_2 - s_1)^{2k_1}\big)\end{aligned}$$

is valid.

Proof: We proceed as before, starting from the integral

$$\int\limits_{E^{\sqrt{\frac{s_2}{n^2 s_1(s_2-s_1)}}}} \int\limits_{E^{\frac{1}{n\sqrt{s_2-s_1}}}} e^{-\frac{\|u\|^2}{2} - \frac{\|v\|^2}{2} + \|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}} + \|v\|\right)} w_{B_N}(v) w_{B_N}(u) \, \mathrm{d}v \, \mathrm{d}u. \tag{A.2}$$

Here, recalling the relationship (A.1) we can see that (A.2) is bounded above by

$$\sum_{i=0}^{N-1} \int_{E\sqrt{\frac{s_2}{n^2 s_1(s_2-s_1)}}} \int_{E^{\frac{1}{n\sqrt{s_2-s_1}},i}_{B_N}} e^{-\frac{\|u\|^2}{2}-\frac{\|v\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|v\|\right)} w_{B_N}(v) w_{B_N}(u)\,\mathrm{d}v\,\mathrm{d}u.$$

We first focus on the integrals for $i \in \{1,\dots,N-1\}$. After performing the same rotation as in the proof of Lemma 4.21

$$\begin{aligned} z_j &= v_j, && \text{for } j \neq i, i+1,\\ z_i &= \tfrac{v_{i+1}+v_i}{\sqrt{2}}, \\ z_{i+1} &= \tfrac{v_{i+1}-v_i}{\sqrt{2}}, \end{aligned}$$

we recall the relation

$$\begin{aligned} z_i - z_{i+1} &= \sqrt{2} v_i, \\ z_i + z_{i+1} &= \sqrt{2} v_{i+1}, \\ \sqrt{2} z_{i+1} &= v_{i+1} - v_i \end{aligned} \tag{A.3}$$

and derive from it

$$\begin{aligned} v_i v_{i+1} &= \frac{z_i^2 - z_{i+1}^2}{2}, \\ v_{i+1}^2 - v_i^2 &= \frac{(z_i+z_{i+1})^2}{2} - \frac{(z_i - z_{i+1})^2}{2} \\ &= \frac{z_i^2 + 2z_i z_{i+1} + z_{i+1}^2 - z_i^2 + 2 z_i z_{i+1} - z_{i+1}^2}{2} \\ &= 2 z_i z_{i+1}. \end{aligned}$$

Hence, we conclude

$$\begin{aligned} w_{B_N}(v) &= \prod_{l=1}^{N} v_l^{2k_1} \prod_{1\le l<m\le N} (v_m^2 - v_l^2)^{2k_2} \\ &= \big(2 z_i z_{i+1}\big)^{2k_2} \frac{(z_i^2 - z_{i+1}^2)^{2k_1}}{2^{2k_1}} \prod_{\substack{l=1:\\ l\neq i,i+1}}^{N} z_l^{2k_1} \prod_{\substack{1\le l<m\le N:\\ l,m\neq i,i+1}} (z_m^2 - z_l^2)^{2k_2} \\ &\quad \prod_{l=1}^{i-1} \left(\frac{(z_i - z_{i+1})^2}{2} - z_l^2\right)^{2k_2} \left(\frac{(z_i + z_{i+1})^2}{2} - z_l^2\right)^{2k_2} \end{aligned}$$

$$\prod_{m=i+2}^{N}\left(z_m^2-\frac{(z_i-z_{i+1})^2}{2}\right)^{2k_2}\left(z_m^2-\frac{(z_i+z_{i+1})^2}{2}\right)^{2k_2}.$$

Obviously, $v\in\overline{W}_{B_N}\subset[0,\infty)^N$ guarantees that all terms in the products are positive and thus we can bound every term separately. We further summarize parts of this via

$$\begin{aligned}
w_{B_N}(v)&=\big(2z_iz_{i+1}\big)^{2k_2}\frac{\big(z_i^2-z_{i+1}^2\big)^{2k_1}}{2^{2k_1}}\prod_{\substack{l=1:\\ l\neq i,i+1}}^{N}z_l^{2k_1}\prod_{\substack{1\le l<m\le N:\\ l,m\neq i,i+1}}\big(z_m^2-z_l^2\big)^{2k_2}\\
&\prod_{l=1}^{i-1}\left(\left(\frac{z_i^2-z_{i+1}^2}{2}\right)^2-z_l^2\big(z_i^2+z_{i+1}^2\big)+z_l^4\right)^{2k_2}\\
&\prod_{m=i+2}^{N}\left(z_m^4-z_m^2\big(z_i^2+z_{i+1}^2\big)+\left(\frac{z_i^2-z_{i+1}^2}{2}\right)^2\right)^{2k_2}
\end{aligned}$$

and combining (A.3) with the ordering of the vector $v\in\overline{W}_{B_N}$

$$\begin{aligned}
0\le z_l=v_l&\le\frac{z_i-z_{i+1}}{\sqrt{2}}=v_{i+1}\le v_m=z_m,\\
0\le z_l=v_l&\le\frac{z_i+z_{i+1}}{\sqrt{2}}=v_i\quad\ \le v_m=z_m
\end{aligned}$$

for $l=1,\dots,i-1$ and $m=i+2,\dots,N$ as well as

$$0\le v_iv_{i+1}=\frac{z_i^2-z_{i+1}^2}{2}\le\frac{z_i^2}{2},$$

we bound the weight function

$$\begin{aligned}
w_{B_N}(v)&\le 2^{2(k_2-k_1)}z_i^{2(2k_1+k_2)}z_{i+1}^{2k_2}\prod_{\substack{l=1:\\ l\neq i,i+1}}^{N}z_l^{2k_1}\prod_{\substack{1\le l<m\le N:\\ l,m\neq i,i+1}}\big(z_m^2-z_l^2\big)^{2k_2}\\
&\prod_{l=1}^{i-1}\left(\frac{z_i^4}{4}-z_l^2z_i^2+z_l^4\right)^{2k_2}\prod_{m=i+2}^{N}\left(z_m^4-z_m^2z_i^2+\frac{z_i^4}{4}\right)^{2k_2}\\
&=2^{2(k_2-k_1)}z_i^{2(2k_1+k_2)}z_{i+1}^{2k_2}\prod_{\substack{l=1:\\ l\neq i,i+1}}^{N}z_l^{2k_1}\prod_{\substack{1\le m<n\le N:\\ m,n\neq i,i+1}}\big(z_n^2-z_m^2\big)^{2k_2}\\
&\prod_{l=1}^{i-1}\left(\frac{z_i^2}{2}-z_l^2\right)^{4k_2}\prod_{m=i+2}^{N}\left(z_m^2-\frac{z_i^2}{2}\right)^{4k_2}
\end{aligned}$$

$$=: z_{i+1}^{2k_2}\Pi_{N,k,i}(z_{i+1}^*).$$

Once more, we obtain an integral of greater value by changing the integration domain similar as in the A_{N-1} case in Lemma 4.21. Therefore, we define

$$\Lambda_{i,N} := \left\{ z \in [0,\infty)^N : z_j \le z_{j+1}\ j \notin \{i-1,i,i+1\}, z_{i-1} + \frac{z_{i+1}}{\sqrt{2}} \le \frac{z_i}{\sqrt{2}} \le z_{i+2} - \frac{z_{i+1}}{\sqrt{2}} \right\}$$

and the larger set

$$\widetilde{\Lambda}_{i,N} := \left\{ z \in [0,\infty)^N : z_j \le z_{j+1}\ j \notin \{i-1,i,i+1\}, z_{i-1} \le \frac{z_i}{\sqrt{2}} \le z_{i+2} \right\}$$

to achieve as new integration domains

$$\Lambda_{i,N} \cap \left\{ 0 \le z_{i+1} \le \frac{1}{n\sqrt{2(s_2-s_1)}} \right\} \subset \widetilde{\Lambda}_{i,N} \cap \left\{ 0 \le z_{i+1} \le \frac{1}{n\sqrt{2(s_2-s_1)}} \right\}.$$

We define $\psi(z_{i+1}^*)$ in the same manner as in the A_{N-1} case to find that the inner integral is bounded by

$$\int\limits_{\Lambda_{i,N}\cap\left\{0\le z_{i+1}\le \frac{1}{n\sqrt{2(s_2-s_1)}}\right\}} e^{-\frac{\|u\|^2}{2}-\frac{\|z_{i+1}^*\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|\psi(z_{i+1}^*)\|\right)} z_{i+1}^{2k_2}\Pi_{N,k,i}(z_{i+1}^*)\,\mathrm{d}z$$

$$\le \int\limits_{\widetilde{\Lambda}_{i,N}\cap\left\{0\le z_{i+1}\le \frac{1}{n\sqrt{2(s_2-s_1)}}\right\}} e^{-\frac{\|u\|^2}{2}-\frac{\|z_{i+1}^*\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|\psi(z_{i+1}^*)\|\right)} z_{i+1}^{2k_2}\Pi_{N,k,i}(z_{i+1}^*)\,\mathrm{d}z$$

$$= \int\limits_{p_{i+1}(\widetilde{\Lambda}_{i,N})} e^{-\frac{\|u\|^2}{2}-\frac{\|z_{i+1}^*\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|\psi(z_{i+1}^*)\|\right)}\Pi_{N,k,i}(z_{i+1}^*)\,\mathrm{d}z_{i+1}^* \int\limits_0^{\frac{1}{n\sqrt{2(s_2-s_1)}}} z_{i+1}^{2k_2}\,\mathrm{d}z_{i+1}$$

$$= \frac{1}{n^{2k_2+1}2^{k_2+\frac{1}{2}}(s_2-s_1)^{k_2+\frac{1}{2}}} \int\limits_{p_{i+1}(\widetilde{\Lambda}_{i,N})} e^{-\frac{\|u\|^2}{2}-\frac{\|z_{i+1}^*\|^2}{2}+\|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}}+\|\psi(z_{i+1}^*)\|\right)}\Pi_{N,k,i}(z_{i+1}^*)\,\mathrm{d}z_{i+1}^*.$$

The case when v_1 is close to zero is treated more simply. The weight function can be bounded above by noting that

$$w_{B_N}(v) = \prod_{l=1}^N v_l^{2k_1} \prod_{1\le l<m\le N} (v_m^2 - v_l^2)^{2k_2}$$

$$
\begin{aligned}
&\le v_1^{2k_1} \prod_{l=2}^{N} v_l^{2(k_1+2k_2)} \prod_{2\le l<m\le N} (v_m^2 - v_l^2)^{2k_2} \\
&=: v_1^{2k_1} \Pi'_{N,k}(v_1^*),
\end{aligned}
$$

accordingly we recognize that the inner integral for $i = 0$ is bounded above by

$$
\begin{aligned}
&\int_0^{\frac{1}{n\sqrt{s_2-s_1}}} \int_{v_1}^{\infty} \cdots \int_{v_{N-1}}^{\infty} e^{-\frac{\|u\|^2}{2} - \frac{\|v\|^2}{2} + \|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}} + \|v\|\right)} v_1^{2k_1} \Pi'_{N,k}(v_1^*) \, \mathrm{d}v_N \cdots \mathrm{d}v_2 \, \mathrm{d}v_1 \\
\le &\int_{W_{B_{N-1}}} e^{-\frac{\|u\|^2}{2} - \frac{\|v_1^*\|^2}{2} + \|u\|\left(\frac{\|x\|}{\sqrt{\varepsilon}} + \|\phi(v_1^*)\|\right)} \frac{1}{n^{2k_1+1}(s_2-s_1)^{k_1+\frac{1}{2}}} \Pi'_{N,k}(v_1^*) \, \mathrm{d}v_1^*
\end{aligned}
$$

by using the integration limit of v_1 and starting the integration over v_2 in zero instead of v_1, which guarantees $W_{B_{N-1}}$ as new integration domain. Additionally, we define $\phi(v_1^*) := (v_2, v_2, v_3, \dots, v_N)$. Now, we employ a similar strategy to bound the integral over u with a sum of integrals over sets of the form given in (A.1). The summation index is here denoted by j. Then, we must distinguish four different cases depending on the type of integral for u and v: whether $i = j = 0$, both i and j are positive or if they are mixed. The result is

$$
\begin{aligned}
\mathbb{P}^x\big(X_{s_1} \in E^{\frac{1}{n}}, X_{s_2} \in E^{\frac{1}{n}}\big) \le c_k^{-2} \left(\frac{s_2 - s_1}{s_2}\right)^{\frac{N}{2}+k_1N+k_2N(N-1)} \Bigg(&\frac{C_{00}(x,\varepsilon,N,k) s_2^{k_1+\frac{1}{2}}}{n^{4k_1+2} s_1^{k_1+\frac{1}{2}} (s_2-s_1)^{2k_1+1}} \\
+ \sum_{j=1}^{N-1} \frac{C_{0j}(x,\varepsilon,N,k) s_2^{k_2+\frac{1}{2}}}{n^{2(k_1+k_2)+2} s_1^{k_2+\frac{1}{2}} (s_2-s_1)^{k_1+k_2+1}} &+ \sum_{i=1}^{N-1} \frac{C_{i0}(x,\varepsilon,N,k) s_2^{k_1+\frac{1}{2}}}{n^{2(k_1+k_2)+2} s_1^{k_1+\frac{1}{2}} (s_2-s_1)^{k_1+k_2+1}} \\
+ \sum_{i,j=1}^{N-1} \frac{C_{ij}(x,\varepsilon,N,k) s_2^{k_2+\frac{1}{2}}}{n^{4k_2+2} s_1^{k_2+\frac{1}{2}} (s_2-s_1)^{2k_2+1}} \Bigg)&.
\end{aligned}
$$

If $i = j = 0$, repeating the calculation that led to the previous expression yields an upper bound given by the first term on the right hand side here. The single sums over i and j come from the cases $j > 0$, $i = 0$ and vice versa. The last term occurs from the case $i, j > 0$. Since $C_{ij} > 0$ we can write

$$
C(x,\varepsilon,N,k) := c_k^{-2} \Bigg(C_{00}(x,\varepsilon,N,k) + \sum_{i,j=1}^{N-1} C_{ij}(x,\varepsilon,N,k)
$$

$$+\sum_{i=1}^{N-1}\Big(C_{i0}(x,\varepsilon,N,k)+C_{0i}(x,\varepsilon,N,k)\Big)\Big)$$

and finally simplify to

$$\begin{aligned}
&\mathbb{P}^x\big(X_{s_1}\in E^{\frac{1}{n}},X_{s_2}\in E^{\frac{1}{n}}\big)\le C(x,\varepsilon,N,k)\left(\frac{s_2-s_1}{s_2}\right)^{\frac{N}{2}+k_1N+k_2N(N-1)}\\
&\quad\cdot\left(\frac{s_2^{k_1+\frac12}}{n^{4k_1+2}s_1^{k_1+\frac12}(s_2-s_1)^{2k_1+1}}+\frac{s_2^{k_2+\frac12}}{n^{2k_1+2k_2+2}s_1^{k_2+\frac12}(s_2-s_1)^{k_1+k_2+1}}\right.\\
&\quad\quad\left.+\frac{s_2^{k_1+\frac12}}{n^{2k_1+2k_2+2}s_1^{k_1+\frac12}(s_2-s_1)^{k_1+k_2+1}}+\frac{s_2^{k_2+\frac12}}{n^{4k_2+2}s_1^{k_2+\frac12}(s_2-s_1)^{2k_2+1}}\right)\\
&\quad\le\frac{C(x,\varepsilon,N,k)}{n^{4\min\{k_1,k_2\}+2}}\left(\frac{s_2-s_1}{s_2}\right)^{\frac{N}{2}+k_1N+k_2N(N-1)}\\
&\quad\quad\cdot\left(\frac{s_2^{k_1+\frac12}}{s_1^{k_1+\frac12}(s_2-s_1)^{2k_1+1}}+\frac{s_2^{k_2+\frac12}}{s_1^{k_2+\frac12}(s_2-s_1)^{k_1+k_2+1}}\right.\\
&\quad\quad\quad\left.+\frac{s_2^{k_1+\frac12}}{s_1^{k_1+\frac12}(s_2-s_1)^{k_1+k_2+1}}+\frac{s_2^{k_2+\frac12}}{s_1^{k_2+\frac12}(s_2-s_1)^{2k_2+1}}\right)\\
&\quad=\frac{C(x,\varepsilon,N,k)}{n^{4\min\{k_1,k_2\}+2}}\\
&\quad\Bigg[\Bigg(\underbrace{\frac{s_2-s_1}{s_2}}_{\le 1}\Bigg)^{\frac{N-1}{2}+k_1(N-1)+k_2N(N-1)}\left(\frac{1}{s_1^{k_1+\frac12}(s_2-s_1)^{k_1+\frac12}}+\frac{1}{s_1^{k_1+\frac12}(s_2-s_1)^{k_2+\frac12}}\right)\\
&\quad+\Bigg(\underbrace{\frac{s_2-s_1}{s_2}}_{\le 1}\Bigg)^{\frac{N-1}{2}+k_1N+k_2N(N-1)-k_2}\left(\frac{1}{s_1^{k_2+\frac12}(s_2-s_1)^{k_1+\frac12}}+\frac{1}{s_1^{k_2+\frac12}(s_2-s_1)^{k_2+\frac12}}\right)\Bigg].
\end{aligned}$$

Since $n\ge 1$ we can bound k_1 and k_2 by $\min\{k_1,k_2\}$ in the exponents of n whereas $\frac{s_2-s_1}{s_2}$ is less then 1 so we can choose the smallest possible exponent, which is obtained for $N=2$, to complete the proof

$$\begin{aligned}
&\mathbb{P}^x\big(X_{s_1}\in E^{\frac{1}{n}},X_{s_2}\in E^{\frac{1}{n}}\big)\le\frac{C(x,\varepsilon,N,k)}{n^{4\min\{k_1,k_2\}+2}}\\
&\quad\Bigg[\left(\frac{s_2-s_1}{s_2}\right)^{\frac{2-1}{2}+k_1(2-1)+k_22(2-1)}\left(\frac{1}{s_1^{k_1+\frac12}(s_2-s_1)^{k_1+\frac12}}+\frac{1}{s_1^{k_1+\frac12}(s_2-s_1)^{k_2+\frac12}}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\left(\frac{s_2-s_1}{s_2}\right)^{\frac{2-1}{2}+k_1 2+k_2 2(2-1)-k_2}\left(\frac{1}{s_1^{k_2+\frac{1}{2}}(s_2-s_1)^{k_1+\frac{1}{2}}}+\frac{1}{s_1^{k_2+\frac{1}{2}}(s_2-s_1)^{k_2+\frac{1}{2}}}\right)\Bigg]\\
&=\frac{C(x,\varepsilon,N,k)}{n^{4\min\{k_1,k_2\}+2}}\\
&\left(\frac{(s_2-s_1)^{2k_2}}{s_1^{k_1+\frac{1}{2}}s_2^{k_1+2k_2+\frac{1}{2}}}+\frac{(s_2-s_1)^{k_1+k_2}}{s_1^{k_1+\frac{1}{2}}s_2^{k_1+2k_2+\frac{1}{2}}}+\frac{(s_2-s_1)^{k_1+k_2}}{s_1^{k_2+\frac{1}{2}}s_2^{2k_1+k_2+\frac{1}{2}}}+\frac{(s_2-s_1)^{2k_1}}{s_1^{k_2+\frac{1}{2}}s_2^{2k_1+k_2+\frac{1}{2}}}\right)\\
&\le\frac{C(x,\varepsilon,N,k)}{n^{4\min\{k_1,k_2\}+2}}\\
&\left(\frac{(s_2-s_1)^{2k_2}}{\varepsilon^{k_1+\frac{1}{2}}\varepsilon^{k_1+2k_2+\frac{1}{2}}}+\frac{(s_2-s_1)^{k_1+k_2}}{\varepsilon^{k_1+\frac{1}{2}}\varepsilon^{k_1+2k_2+\frac{1}{2}}}+\frac{(s_2-s_1)^{k_1+k_2}}{\varepsilon^{k_2+\frac{1}{2}}\varepsilon^{2k_1+k_2+\frac{1}{2}}}+\frac{(s_2-s_1)^{2k_1}}{\varepsilon^{k_2+\frac{1}{2}}\varepsilon^{2k_1+k_2+\frac{1}{2}}}\right)\\
&=\frac{C(x,\varepsilon,N,k)}{n^{4\min\{k_1,k_2\}+2}}\cdot\frac{(s_2-s_1)^{2k_2}+2(s_2-s_1)^{k_1+k_2}+(s_2-s_1)^{2k_1}}{\varepsilon^{2k_1+2k_2+1}}.
\end{aligned}
$$

□

Owing to this lemma and the lower bound from Lemma A.1, the lower bound is valid in Conjecture 4.13 for $R=B_N$ by an analogous calculation as in the A_{N-1} case, see the proof of Theorem 4.19. We have already discussed which kind of bound is lacking to achieve the upper bound of Conjecture 4.13.

List of symbols

General

$\mathbb{N}$	natural numbers: $\{1, 2, \dots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$
$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\langle \cdot, \cdot \rangle$	standard Euclidean inner product
$\\| \cdot \\|$	$\\|x\\| := \langle x, x \rangle = \sqrt{x_1^2 + \cdots + x_d^2}$
$\langle \cdot \rangle^\perp$	hyperplane $\langle x \rangle^\perp := \{y \in \mathbb{R}^d \mid \langle x, y \rangle = 0\}$
S_N	symmetric group
$\lambda^d(A)$	d-dimensional Lebesgue measure of $A \subset \mathbb{R}^d$
∂A	boundary of the set $A \subset \mathbb{R}^d$
$\overline{A}$	closure of the set $A \subset \mathbb{R}^d$

Stochastic

$\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}\big)$	underlying filtered probability space with $\mathcal{F}_t := \sigma\big(X_s \mid s \leq t\big)$
$\mathbb{E}\,(X_t \mid X_s)$	conditional mean
$\mathbb{V}\mathrm{ar}\,(X_t \mid X_s)$	conditional variance
$[X]_t$	quadratic variation of the process X
$[X, Y]_t$	covariation process of X and Y
$\mathcal{N}(\mu, \Sigma)$	normal distribution with mean vector μ and covariance matrix Σ
$\phi(\cdot, \mu, \Sigma)$	corresponding Gaussian density of $\mathcal{N}(\mu, \Sigma)$
$X \sim \mu$	X has distribution μ
$\delta.$	Dirac measure
$a_n \lesssim b_n$	$\sup\big\{\frac{a_n}{b_n} \mid n \in \mathbb{N}\big\} \leq c$ almost surely

Spaces of functions

$\mathcal{C}^1(A)$	set of continuous differentiable functions on $A \subset \mathbb{R}^d$
$\mathcal{C}^2(A)$	set of two times continuous differentiable functions on $A \subset \mathbb{R}^d$

Special functions

Γ	Gamma function
I_ν	modified Bessel function of the first kind
J_ν	Bessel function of the first kind
j_ν	spherical Bessel function of the first kind
$\lfloor\cdot\rfloor$	floor function $\lfloor C\rfloor := \inf\{n\in\mathbb{N}\,|\,n\le C<n+1\}$ for some $C\in\mathbb{R}$
$\mathbb{1}$	indicator function
$\mathbb{I}_n$	n dimensional identity matrix
$\operatorname{diag}(\cdots)$	diagonal matrix

Derivatives

Δ	Laplacian
∇	gradient
∇_x	gradient with respect to x
$\frac{\partial}{\partial x}$	derivative with respect to x
D	total differential
H_f	Hessian matrix of the function f

Symbols belonging to the Dunkl theory

The page belongs to the definition of the symbol.

A_{N-1}, B_N	Weyl chambers, pages 9 and 10
$w_{A_{N-1}}$	weight function belonging to A_{N-1}, page 10
w_{B_N}	weight function belonging to B_N, page 10
$\kappa(k,R)$	sum of multiplicities, pages 8 and 11
$J_{k,R}$	Dunkl Bessel function, page 14
$q_{k,R}, Q_{k,R}$	density and distribution of the multivariate Bessel process, page 14

Landau symbols

We consider a positive function $r_n > 0$ and a random function ζ_n on $\Theta\subset\mathbb{R}^d$.

$\zeta_n = O_p(r_n)$	If $\limsup_{n\to\infty} \lvert r_n^{-1}\zeta_n\rvert < \infty$ in distribution.
$\zeta_n = O_p^*(r_n)$	If $\sup_{\theta\in\Theta} \lvert\zeta_n(\theta)\rvert = O_p(r_n)$.
$\zeta_n = O^*(r_n)$	If $\zeta_n(\theta) = O_p^*(r_n)$ and ζ_n is non-random.
$\zeta_n = o_p(r_n)$	If $\lim_{n\to\infty} r_n^{-1}\zeta_n = 0$ in distribution.
$\zeta_n = o_p^*(r_n)$	If $\sup_{\theta\in\Theta} \lvert\zeta_n(\theta)\rvert = o_p(r_n)$.
$\zeta_n = o^*(r_n)$	If $\zeta_n = o_p^*(r_n)$ and ζ_n is non-random.

R Source Code

Martingale estimator based on one eigenfunction

This code uses the Euler Maruyama method to simulate the modified Bessel process. These simulations belongs to the estimator (2.8) and the corresponding optimal estimator on page 51.

```
# approximation of a Brownian motion on [lower,upper]=[0,1] with
    length.out=100 equidistant observations
get_Wiener <- function(mu=0, sigma=1, lower=0, upper=1, length.out=100){
  l <- length.out-1
  c(0, cumsum(rnorm(l, mean=mu*(upper-lower)/length.out,
      sd=sqrt(sigma*(upper-lower)/length.out))))
}

# alpha, theta specified by the stochastic differential equation defining the
    modified Bessel process.
# generates data on the interval [lower,upper] with n observations and
    starting point x_0
get_modified_Bessel <- function(alpha,theta,x_0,lower,upper,n){
  Z <- x_0 # starting point
  is_viable <- FALSE
  broken_process <- FALSE
  length.out <- n*upper/10
  length.out_temp <- n*upper/10 # Using more time points to generate the
      process than we actually use for the estimator
  while( !is_viable ) {
    B <- get_Wiener(lower=lower,upper=upper,length.out=length.out_temp+1)
    for(i in 1 : length.out_temp){
      Z[i+1] <- Z[i]+B[i+1]-B[i]+((theta+0.5)/Z[i]-alpha*Z[i])*resolution
      if (Z[i+1] <= 0 ) break
    }
    if(tail(Z,1)>0&&length(Z)==length.out_temp+1){
      is_viable = TRUE
    }
    else {
      length.out_temp <- 2*length.out_temp # if the distance between the
          observations is too small, we enlarge the number of observations
          for the simulation of the Brownian motion
```

```
         resolution <- resolution / 2
         Z <- x_0
         if (log2(length.out_temp/length.out)>10){
            is_viable = TRUE
            broken_process = TRUE
         }
      }
   }
   if(broken_process) {
      return(NA)
   } else {
      Y <- numeric(n+1)
      Y[1] <- x_0
      for (i in 1:n){
         # save the desired observations
         Y[i+1] <- Z[upper/10*2^log2(l.out_temp/length.out)*i+1]
         return(Y)
      }}
}

estimator_one_eigenfunction <- function(alpha,theta,x_0=0.1,delta,n){
   Y <- get_modified_Bessel(alpha=alpha, theta=theta, x_0=x_0, lower=0,
       upper=delta*n, n=n)
   return((alpha*(sum(Y[-1]^2)-exp(-2*alpha*delta)*sum(Y[-n-1]^2)))/
       (n*(1-exp(-2*alpha*delta)))-1)
}

optimal_estimator_one_eigenfunction <- function(alpha,theta,x_0=0.1,delta,n){
   Y <- get_modified_Bessel(alpha=alpha, theta=theta, x_0=x_0, lower=0,
       upper=delta*n, n=n)
   estimator <- function(vartheta){
      sum(1/((vartheta+1)/alpha*(1-exp(-2*alpha*delta))+2*exp(-2*alpha*delta)*
         Y[-n-1]^2)*(Y[-1]^2-Y[-n-1]^2*exp(-2*alpha*delta)-(vartheta+1)/alpha*
         (1-exp(-2*alpha*delta))))^2
   }
   # this function searches for the value theta such that estimator(theta)=0
   return(optim(par=theta, fn=estimator, method="L-BFGS-B", lower=theta-5,
       upper=theta+5)$par)
}
```

Simulating a Cox-Ingersoll-Ross process

This code is available as simCIR through the yuima package in R, cf. [45, 46], and generates a Cox-Ingersoll-Ross process.

```
## Simulate Cox-Ingersoll-Ross process with parameters alpha, beta and gamma
    at times specified via time.points
simCIR <- function(time.points,n,h,alpha,beta,gamma,equi.dist=FALSE){
   # generate an equidistant time vector of length n+1 and distant h between
```

```
    observations
  if(equi.dist==TRUE){time.points <- 0:n*h}
  # must start in t=0, otherwise t_vec is adjusted
  if(time.points[1]!=0){time.points <- c(0,time.points)}
  # define auxiliary variables, following notation of (1.9)
  nu <- 4*beta*alpha/(beta*gamma) # degrees of freedom
  # auxiliary vector for the computation of the non-centrality parameter in
      each step
  eta_vec <- 4*beta*exp(-beta*diff(time.points))/
     (gamma*(1-exp(-beta*diff(time.points))))
  # sample X_0 from stationary distribution
  X <- rgamma(1,scale=gamma/(2*beta),shape=2*alpha/gamma)
  # compute X_t iteratively, using Proposition 1 of Malham and Wiese
  for(i in seq_along(eta_vec)){
     # non-centrality parameter of the conditional distribution
     lambda <- tail(X,1)*eta_vec[i]
     # calculate next step of the CIR
     X <- c(X,rchisq(1,df=nu,ncp=lambda)*exp(-beta*diff(time.points)[i])/
         eta_vec[i])
  }
  # return data; first row: time points, second row: CIR at time point
  return(rbind(t=time.points,X=X))
}
```

Simulations for the preliminary estimator and one-step improvements

This code is available through the yuima package as fitCIR in R, cf. [45, 46], and is the code to simulate the preliminary estimator, cf. Lemma 3.1 and (3.10), and the one-step improvements based on the Newton-Raphson (3.22) and the Fisher scoring method (3.23) using the function simCIR above.

```
## function to provide the preliminary explicit estimator
get_preliminaryEstimators <- function(data){
  # use data returned from calling simCIR()
  n <- dim(data)[2]-1 # we have observations at t_j for j=0,...,n, this
      equals n+1 observations in total
  # -> therefore, n is the number of observations minus 1
  h <- as.numeric(data[1,2]) # as the vector containing the t always starts
      with 0, this equals diff(data[1,])[1]
  # and thus gives h if the t_j are chosen to be equidistant
  X_major <- data[2,-1] # the observations of the CIR process starting at
      t_1=h.
  X_minor <- head(data[2,],-1) # the observations of the CIR process
      discarding j=n.
  X_mean_major <- mean(X_major) # mean of all observations without j=0.
  X_mean_minor <- mean(X_minor) # mean of all observations without j=n.
  # calculate estimate for beta from Lemma 3.1
  beta_0n <- -1/h*log(sum((X_minor-X_mean_minor)*(X_major - X_mean_major))/
```

```
        sum((X_minor-X_mean_minor)^2))
   # calculate estimate for alpha from Lemma 3.1
   alpha_0n <- (X_mean_major-exp(-beta_0n*h)* X_mean_minor)*beta_0n/
        (1-exp(-beta_0n*h))
   # calculate estimate for gamma from Eq. (3.10), where first the numerator
        and denominator are computed separately.
   to_gamma_numerator <- (X_major-exp(-beta_0n*h)*X_minor-alpha_0n/beta_0n*
        (1-exp(-beta_0n*h)))^2
   to_gamma_denominator <-
        (1-exp(-beta_0n*h))/beta_0n*(exp(-beta_0n*h)*X_minor+
        alpha_0n*(1-exp(-beta_0n*h))/(2*beta_0n))
   gamma_0n <- 1/n*sum(to_gamma_numerator/to_gamma_denominator)
   # return the estimated parameter vector
   return(as.matrix(c(alpha_0n,beta_0n,gamma_0n)))
}


### ONE STEP IMPROVEMENT
# first we need a few auxiliary functions

# the mean of the CIR with parameter (alpha,beta,gamma) conditioned on X_h=x,
    Eq. (3.1)
mu<- function(alpha,beta,gamma,x,h){
   return(exp(-beta*h)*x+alpha/beta*(1-exp(-beta*h)))
}

# derivatives of the mean of CIR with parameter (alpha,beta,gamma)
    conditioned on X_h=x
mu_alpha <-function(alpha,beta,gamma,x,h)
{
   return((1-exp(-beta*h))/beta)
}

mu_alpha_alpha<-function(alpha,beta,gamma,x,h)
{
   return(0)
}

mu_alpha_beta <- function(alpha,beta,gamma,x,h)
{
   return((exp(-beta*h)*(beta*h+1)-1)/beta^2)
}

mu_alpha_gamma <- function(alpha,beta,gamma,x,h)
{
   return(0)
}

mu_beta <- function(alpha,beta,gamma,x,h)
```

```
{
  return(-h*exp(-beta*h)*x+alpha*(exp(-beta*h)*(beta*h+1)-1)/beta^2)
}

mu_beta_beta <- function(alpha,beta,gamma,x,h)
{
  return(h^2*exp(-beta*h)*x+alpha*(exp(-beta*h)*(-beta^2*h^2-2*beta*h-2)+2)/beta^3)
}

mu_beta_gamma <- function(alpha,beta,gamma,x,h)
{
  return(0)
}

mu_gamma <- function(alpha,beta,gamma,x,h)
{
  return(0)
}

mu_gamma_gamma <- function(alpha,beta,gamma,x,h)
{
  return(0)
}

# the variance of CIR with parameter (alpha,beta,gamma) conditioned on X_h=x,
    Eq. (3.1)

sigma_mod <- function(alpha,beta,gamma,x,h){
  return(gamma/beta*(1-exp(-beta*h))*(exp(-beta*h)*x+alpha/(2*beta)*
      (1-exp(-beta*h))))
}

# derivatives of the variance of CIR with parameter (alpha,beta,gamma)
    conditioned on X_h=x

sigma_alpha <- function(alpha,beta,gamma,x,h)
{
  return(gamma*(1-exp(-beta*h))^2/(2*beta^2))
}

sigma_alpha_alpha <- function(alpha,beta,gamma,x,h)
{
  return(0)
}

sigma_alpha_beta <- function(alpha,beta,gamma,x,h)
{
  return(-gamma*exp(-2*beta*h)*(exp(beta*h)-1)*(-beta*h+exp(beta*h)-1)/
      beta^3)
```

```
}

sigma_alpha_gamma <- function(alpha,beta,gamma,x,h)
{
   return((1-exp(-beta*h))^2/(2*beta^2))
}

sigma_beta <- function(alpha,beta,gamma,x,h)
{
   exp(-2*h*beta)/beta^3*(x*beta*(2*h*beta-exp(h*beta)*(h*beta+1)+1)-
       alpha*(exp(beta*h)-1)*(-h*beta+exp(h*beta)-1))
}

sigma_beta_beta <- function(alpha,beta,gamma,x,h)
{
   exp(-2*beta*h)/(beta^4)*(alpha*(2*h*beta*(h*beta+2)+3*exp(2*beta*h)-
       exp(h*beta)*(h*beta*(h*beta+4)+6)+3)+x*beta*(exp(beta*h)*
       (h^2*beta^2+2*h*beta+2)-2*(2*h^2*beta^2+2*h*beta+1)))
}

sigma_beta_gamma <- function(alpha,beta,gamma,x,h)
{
   return(h*exp(-beta*h)*(exp(-beta*h)*x/beta+alpha*(1-exp(-beta*h))/(2*beta^2))+
       (1-exp(-beta*h))*(-exp(-beta*h)*(beta*h+1)/beta^2*x+
       alpha*exp(-beta*h)* (beta*h-2*exp(beta*h)+2)/(2*beta^3)))
}

sigma_gamma <- function(alpha,beta,gamma,x,h)
{
   return((1-exp(-beta*h))*(exp(-beta*h)*x+alpha/(2*beta)*(1-exp(-beta*h)))/beta)
}

sigma_gamma_gamma <- function(alpha,beta,gamma,x,h)
{
   return(0)
}

# the inverse of the Hessian matrix of H_n
H_n_Hessian <- function(alpha,beta,gamma,x,h)
{
   mu <- mu(alpha,beta,gamma,x,h)

   mu_1 <- mu_alpha(alpha,beta,gamma,x,h)
   mu_2 <- mu_beta(alpha,beta,gamma,x,h)
   mu_3 <- mu_gamma(alpha,beta,gamma,x,h)

   mu_11 <- mu_alpha_alpha(alpha,beta,gamma,x,h)
   mu_12 <- mu_alpha_beta(alpha,beta,gamma,x,h)
   mu_13 <- mu_alpha_gamma(alpha,beta,gamma,x,h)
```

```
mu_22 <- mu_beta_beta(alpha,beta,gamma,x,h)
mu_23 <- mu_beta_gamma(alpha,beta,gamma,x,h)
mu_33 <- mu_gamma_gamma(alpha,beta,gamma,x,h)

sigma <- sigma_mod(alpha,beta,gamma,x,h)

sigma_1 <- sigma_alpha(alpha,beta,gamma,x,h)
sigma_2 <- sigma_beta(alpha,beta,gamma,x,h)
sigma_3 <- sigma_gamma(alpha,beta,gamma,x,h)

sigma_11 <- sigma_alpha_alpha(alpha,beta,gamma,x,h)
sigma_12 <- sigma_alpha_beta(alpha,beta,gamma,x,h)
sigma_13 <- sigma_alpha_gamma(alpha,beta,gamma,x,h)
sigma_22 <- sigma_beta_beta(alpha,beta,gamma,x,h)
sigma_23 <- sigma_beta_gamma(alpha,beta,gamma,x,h)
sigma_33 <- sigma_gamma_gamma(alpha,beta,gamma,x,h)

H_11 <- -0.5*sum((sigma_11*sigma-sigma_1*sigma_1)/sigma^2-
    (sigma_11*sigma-2*sigma_1*sigma_1)/sigma^3*(x-mu)^2+
    2*sigma_1/sigma^2*(x-mu)*mu_1-
    2*(mu_11*sigma-mu_1*sigma_1)*(x-mu)/sigma^2+2*mu_1*mu_1/sigma)


H_22 <- -0.5*sum((sigma_22*sigma-sigma_2*sigma_2)/sigma^2-
    (sigma_22*sigma-2*sigma_2*sigma_2)/sigma^3*(x-mu)^2+
    2*sigma_2/sigma^2*(x-mu)*mu_2-
    2*(mu_22*sigma-mu_2*sigma_2)*(x-mu)/sigma^2+ 2*mu_2*mu_2/sigma)

H_33 <- -0.5*sum((sigma_33*sigma-sigma_3*sigma_3)/sigma^2-
    (sigma_33*sigma-2*sigma_3*sigma_3)/sigma^3*(x-mu)^2+
    2*sigma_3/sigma^2*(x-mu)*mu_3-
    2*(mu_33*sigma-mu_3*sigma_3)*(x-mu)/sigma^2+2*mu_3*mu_3/sigma)


H_12 <- -0.5*sum((sigma_12*sigma-sigma_1*sigma_2)/sigma^2-
    (sigma_12*sigma-2*sigma_1*sigma_2)/sigma^2*(x-mu)^2+
    2*sigma_1/sigma^2*(x-mu)*mu_2-
    2*(mu_12*sigma-mu_1*sigma_2)*(x-mu)/sigma^2+2*mu_1*mu_2/sigma)

H_13 <- -0.5*sum((sigma_13*sigma-sigma_1*sigma_3)/sigma^2-
    (sigma_13*sigma-2*sigma_1*sigma_3)/sigma^3*(x-mu)^2+
    2*sigma_1/sigma^2*(x-mu)*mu_3-
    2*(mu_13*sigma-mu_1*sigma_3)*(x-mu)/sigma^2+2*mu_1*mu_3/sigma)

H_23 <- -0.5*sum((sigma_23*sigma-sigma_2*sigma_3)/sigma^2-
    (sigma_23*sigma-2*sigma_2*sigma_3)/sigma^3*(x-mu)^2+
    2*sigma_2/sigma^2*(x-mu)*mu_3-
    2*(mu_23*sigma-mu_2*sigma_3)*(x-mu)/sigma^2+ 2*mu_2*mu_3/sigma)
```

```
  H_det <- H_11*H_22*H_33+H_12*H_23*H_13+H_13*H_12*H_23-
      H_13*H_22*H_13-H_11*H_23*H_23-H_12*H_12*H_33

  return(matrix(1/H_det*c(H_22*H_33-H_23^2,H_13*H_23-H_12*H_33,
      H_12*H_23-H_13*H_22,H_13*H_23-H_12*H_33,H_11*H_33-H_13^2,
      H_13*H_12-H_11*H_23,H_12*H_23-H_13*H_22,H_13*H_12-H_11*H_23,
      H_11*H_22-H_12^2),nrow=3))

}

# derivatives of the function H_n(theta)

H_n_alpha <- function(alpha,beta,gamma,x,h){
  x_minor <- head(x,-1)
  x_major <- x[-1]
  sigma_deriv <- sigma_alpha(alpha,beta,gamma,x_minor,h)
  sigma <- sigma_mod(alpha,beta,gamma,x_minor,h)
  mu <- mu(alpha,beta,gamma,x_minor,h)
  mu_deriv <- mu_alpha(alpha,beta,gamma,x_minor,h)
  return(sum(-0.5*(sigma_deriv/sigma-sigma_deriv/sigma^2*(x_major-mu)^2-
      2/sigma*(x_major-mu)*mu_deriv)))
}

H_n_beta <- function(alpha,beta,gamma,x,h){
  x_minor <- head(x, -1)
  x_major <- x[-1]
  sigma_deriv <- sigma_beta(alpha,beta,gamma,x_minor,h)
  sigma <- sigma_mod(alpha,beta,gamma,x_minor,h)
  mu <- mu(alpha,beta,gamma,x_minor,h)
  mu_deriv <- mu_beta(alpha,beta,gamma,x_minor,h)
  return(sum(-0.5*(sigma_deriv/sigma-sigma_deriv/sigma^2*(x_major-mu)^2-
      2/sigma*(x_major-mu)*mu_deriv)))
}

H_n_gamma <- function(alpha,beta,gamma,x,h){
  x_minor <- head(x,-1)
  x_major <- x[-1]
  sigma_deriv <- sigma_gamma(alpha,beta,gamma,x_minor,h)
  sigma <- sigma_mod(alpha,beta,gamma,x_minor,h)
  mu <- mu(alpha,beta,gamma,x_minor,h)
  return (sum(-0.5*(sigma_deriv/sigma-sigma_deriv/sigma^2*(x_major-mu)^2)))
}

# The level-a confidence intervals for the one-step improvements calculated
    from the estimated asymptotic covariance:
get_confidenceIntervals <- function(alpha,beta,gamma,a,n,T){
  z<- qnorm(1-a/2)
  alpha_lower <- alpha-z*sqrt(alpha*(2*alpha-gamma)/(T*beta))
```

```r
   alpha_upper <- alpha+z*sqrt( alpha*(2*alpha-gamma)/(T*beta))
   beta_lower <- beta-z*sqrt(2*beta/T)
   beta_upper <- beta+z*sqrt(2*beta/T)
   gamma_lower <- gamma-z*sqrt(2*gamma^2/n)
   gamma_upper <- gamma+z*sqrt(2*gamma^2/n)
   return(matrix(c(alpha_lower,beta_lower,gamma_lower,
       alpha_upper,beta_upper,gamma_upper),3,2,
       dimnames=list(c("alpha","beta","gamma"),c("lower","upper"))))
}

get_finalEstimators_scoring <- function(data,a) {
   # We need the preliminary estimator for the one step improvement, Eq.
       (3.23)
   prelim_estim <- get_preliminaryEstimators(data)
   alpha <- prelim_estim[1]
   beta <- prelim_estim[2]
   gamma <- prelim_estim[3]
   T <- tail(data[1,],1) # T is the time of the last observation n*h
   n<- length(data[1,])-1 # number of observations minus 1
   x<-data[2,-1]
   h<-data[1,2]

   H_n<-c(H_n_alpha(alpha,beta,gamma,x,h),H_n_beta(alpha,beta,gamma,x,h),
       H_n_gamma(alpha,beta,gamma,x,h)) #calculate estimate from Eq.
   finalEstimators_scoring <- c(alpha,beta,gamma)+diag(c(1/T,1/T,1/n))%*%
       matrix(c(alpha*(2*alpha-gamma)/beta,2*alpha-gamma,0,
       2*alpha-gamma,2*beta,0,0,0,2*gamma^2),nrow=3,byrow=TRUE)%*%H_n
   confidenceIntervals <- get_confidenceIntervals(finalEstimators_scoring[1],
       finalEstimators_scoring[2],finalEstimators_scoring[3],a,n,T)
   return(list(estimation=finalEstimators_scoring,
       confidence=confidenceIntervals))
}

# Newton-Raphson Eq. (3.22)
get_finalEstimators_newton <- function(data,a) {
   prelim_estim <- get_preliminaryEstimators(data)
   alpha <- prelim_estim[1]
   beta <- prelim_estim[2]
   gamma <- prelim_estim[3]
   T <- tail(data[1,],1)
   n<- length(data[1,])-1
   x<-data[2,-1]
   h<-data[1,2]

   H_n<-c(H_n_alpha(alpha,beta,gamma,x,h),H_n_beta(alpha,beta,gamma,x,h),
       H_n_gamma(alpha,beta,gamma,x,h))
   finalEstimators_newton <- c(alpha,beta,gamma)-
       H_n_Hessian(alpha,beta,gamma,x,h)%*%H_n
   confidenceIntervals <- get_confidenceIntervals(finalEstimators_newton[1],
```

```
        finalEstimators_newton[2],finalEstimators_newton[3],a,n,T)
    return(list(estimation=finalEstimators_newton,
        confidence=confidenceIntervals))
}


fitCIR<- function(data,a=0.05){
    if((max(diff(data[1,]))/min(diff(data[1,]))-1)>1e-10) stop('Please use
        equidistant sampling points')
    if(a>1) stop('Please set the number a less than 1')
    newtonEstimators <- get_finalEstimators_newton(data,a)
    return(list(preliminary=get_preliminaryEstimators(data),
        NewtonRaphson=get_finalEstimators_newton(data,a),
        scoring=get_finalEstimators_scoring(data,a)))
}


# ### Examples of usage
# ## If the sampling points are not equidistant, there will be an
    corresponding output.
# data <- simCIR(alpha=3,beta=1,gamma=1,time.points = c(0,0.1,0.2,0.25,0.3))
# fitCIR(data)
# ## Otherwise it calculates the three estimators
# data <- simCIR(alpha=3,beta=1,gamma=1,n=1000,h=0.1,equi.dist=TRUE)
# fitCIR(data)
```

Bibliography

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables.* Applied mathematics series. Dover Publications, 1965.

[2] R. Adler. *The Geometry of Random Fields.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1981.

[3] L. Ahlfors and K. M. R. Collection. *Complex Analysis: An Introduction to The Theory of Analytic Functions of One Complex Variable.* International series in pure and applied mathematics. McGraw-Hill Education, 1979.

[4] M. B. Alaya and A. Kebaier. Parameter estimation for the square-root diffusions: Ergodic and nonergodic cases. *Stochastic Models*, 28(4):609–634, 2012.

[5] M. B. Alaya and A. Kebaier. Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. *Stochastic Analysis and Applications*, 31(4):552–573, 2013.

[6] M. B. Alaya, A. Kebaier, and N. K. Tran. Local asymptotic properties for Cox-Ingersoll-Ross process with discrete observations. *Scandinavian Journal of Statistics*, 47(4):1401–1464, 2020.

[7] A. Ambrosetti and P. H. Rabinowitz. Dual variational methods in critical point theory and applications. *Journal of Functional Analysis*, 14(4):349–381, 1973.

[8] K. Atkinson. *An Introduction to Numerical Analysis.* Wiley, 1989.

[9] J. Aubin and I. Ekeland. *Applied Nonlinear Analysis.* Dover Books on Mathematics Series. Dover Publications, 2006.

[10] T. Baker and P. Forrester. The Calogero-Sutherland Model and Generalized Classical Polynomials. *Communications in Mathematical Physics*, 188:175–216, 1997.

[11] B. M. Bibby and M. Sørensen. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, 1(1–2):17–39, 1995.

[12] P. Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Statistics. Wiley, 2nd edition, 1999.

[13] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[14] P. Brémaud. *Probability Theory and Stochastic Processes.* Universitext. Springer International Publishing, 2020.

[15] I. Bronstein, K. Semendjajew, G. Musiol, and H. Mühlig. *Taschenbuch der Mathematik.* Harri Deutsch, 8th edition, 2008.

[16] R. Brown. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

[17] M.-F. Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751, 1991.

[18] R. Carroll and D. Ruppert. *Transformation and Weighting in Regression.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1988.

[19] E. Cépa and D. Lépingle. Diffusing particles with electrostatic repulsion. *Probability Theory and Related Fields*, 107:429–449, 1997.

[20] Y. Cheng. Gaussian quasi-likelihood estimation of the square-root diffusion from high-frequency data. Master thesis, Kyushu University, 2021.

[21] Y. Cheng, N. Hufnagel, and H. Masuda. Estimation of ergodic square-root diffusion under high-frequency sampling. *Econometrics and Statistics*, DOI 10.1016/j.ecosta.2022.05.003, 2022.

[22] O. Chybiryakov, L. Gallardo, and M. Yor. Dunkl processes and their radial parts relative to a root system. In *Harmonic and stochastic analysis of Dunkl processes.* Hermann, 2008.

[23] J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[24] E. Cépa. *Equations différentielles stochastiques multivoques.* PhD thesis, Lille, 1994.

[25] N. Demni. A guided tour in the world of radial Dunkl processes. In *Harmonic and stochastic analysis of Dunkl processes.* Hermann, 2008.

[26] N. Demni. Radial Dunkl processes: Existence, uniqueness and hitting time. *Comptes Rendus de l'Académie des Sciences de Paris, Série I*, 347:1125–1128, 2009.

[27] C. F. Dunkl. Differential-difference operators associated to reflection groups. *Transactions of the American Mathematical Society*, 311(1):167–183, 1989.

[28] C. F. Dunkl. Integral kernels with reflection group invariance. *Canadian Journal of Mathematics*, 43(6):1213–1227, 1991.

[29] C. F. Dunkl and Y. Xu. *Orthogonal Polynomials of Several Variables.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2001.

[30] F. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962.

[31] A. Einstein. Investigations on the Theory of the Brownian Movement. *Dover Publications*, 7(10):281–281, 1956.

[32] S. N. Ethier and T. G. Kurtz. *Markov Processes: Charakterization and Convergence.* Wiley Series in Probability and Statistics, 1986.

[33] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications.* Wiley, 2013.

[34] V. Genon-Catalot and J. Jacob. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l'Institut Henri Poincaré, Probabilités et statistiques*, 29(1):119–151, 1993.

[35] V. Genon-Catalot, T. Jeantheau, and C. Larédo. Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli*, 6(6):1051–1079, 2000.

[36] E. Gobet. LAN property for ergodic diffusions with discrete observations. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 38(5):711–737, 2002.

[37] V. P. Godambe and C. C. Heyde. Quasi-likelihood and optimal estimation, correspondent paper. *International Statistical Review / Revue Internationale de Statistique*, 55(3):231–244, 1987.

[38] A. Göing-Jaeschke and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[39] P. Graczyk and P. Sawyer. Sharp estimates for W-invariant Dunkl and heat kernels in the A_n case, arXiv:2111.13529, 2021.

[40] J. Graf and E. Gubler. *Einleitung in die Theorie der Bessel'schen Funktionen.* K. J. Wyss, 1898.

[41] C. Heyde. Fixed sample and asymptotic optimality for classes of estimating functions. *Contemporary Mathematics*, 80:241–247, 1988.

[42] R. Höpfner. *Asymptotic Statistics: With a View to Stochastic Processes.* De Gruyter, 2014.

[43] N. Hufnagel and S. Andraus. Hausdorff dimension of collision times in one-dimensional log-gases, arXiv:2109.08707, 2021.

[44] N. Hufnagel and J. H. C. Woerner. Martingale estimation functions for Bessel processes. *Statistical Inference for Stochastic Processes*, 25:337–353, 2022.

[45] S. Iacus, Y. Koike, and et al. Yuima project: Project home - r-forge. https://r-forge.r-project.org/projects/yuima/. Last visited: 16/11/2022, 2009.

[46] S. Iacus and N. Yoshida. *Simulation and Inference for Stochastic Processes with YUIMA: A Comprehensive R Framework for SDEs and Other Stochastic Processes.* Use R! Springer Cham, 2018.

[47] I. A. Ibragimov and R. Z. Hasminskii. *Statistical estimation*, volume 16 of *Applications of Mathematics.* Springer-Verlag, New York-Berlin, 1981. Asymptotic theory, Translated from the Russian by Samuel Kotz.

[48] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, volume 24 of *North-Holland Mathematical Library.* Elsevier, 1981.

[49] K. Itô and H. P. McKean. *Diffusion Processes and their Sample Paths.* Springer, 1974.

[50] C. G. J. Jacobi. *Versuch einer Berechnung der grossen Ungleichheit des Saturns nach einer strengen Entwickelung*, volume 7 of *Cambridge Library Collection - Mathematics*, page 145–174. Cambridge University Press, 2013.

[51] J. Jacod. Statistics and high-frequency data. In M. K. et al., editor, *Statistical Methods for Stochastic Differential Equations.* Taylor & Francis Group, Paris, 2012.

[52] R. I. Jennrich and P. F. Sampson. Newton-raphson and related algorithms for maximum likelihood variance component estimation. *Technometrics*, 18(1):11–17, 1976.

[53] J.-P. Kahane. *Some random series of functions.* Cambridge University Press, 1994.

[54] I. Karatzas and Shreve. *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics (113). Springer New York, 2nd edition, 1991.

[55] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes.* Elsevier Science, 1981.

[56] M. Katori. *Bessel processes Schramm-Loewner evolution and the Dyson model.* Springer, 2015.

[57] J. Kent. Some Probabilistic Properties of Bessel Functions. *The Annals of Probability*, 6(5):760–770, 1978.

[58] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24(2):211–229, 1997.

[59] M. Kessler and M. Sørensen. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5:299–314, 1999.

[60] W. König and N. O'Connell. Eigenvalues of the Laguerre process as non-colliding squared Bessel processes. *Electronic Communications in Probability*, 6:107–114, 2001.

[61] L. Liu and Y. Xiao. Hausdorff dimension theorem for self-similar Markov processes. *Probability and Mathematical Statistics*, 18, 1998.

[62] N. T. Longford. A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, 74(4):817–827, 1987.

[63] S. J. A. Malham and A. Wiese. Chi-square simulation of the CIR process and the Heston model. *International Journal of Theoretical and Applied Finance*, 16(3):1–38, 2013.

[64] M. B. Marcus. Capacity of Level Sets of Certain Stochastic Processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 34(4):279–284, 1976.

[65] H. Masuda. Non-Gaussian quasi-likelihood estimation of SDE driven by locally stable Lévy process. *Stochastic Processes and their Applications*, 129(3):1013–1059, 2019.

[66] H. P. McKean Jr. The Bessel motion and a singular integral equation. *Memoirs of the College of Science, University of Kyoto. Series A: Mathematics*, 33(2):317 – 322, 1960.

[67] D. S. Mitrinović. *Analytic Inequalities.* Springer Berlin, Heidelberg, 1st edition, 1970.

[68] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Universitext. Springer-Verlag, 6th corrected printing edition, 2003.

[69] F. Olver, D. Lozier, R. Boisvert, and C. Clark. *The NIST Handbook of Mathematical Functions.* Cambridge University Press, New York, 2010.

[70] E. M. Opdam. Dunkl operators, Bessel functions and the discriminant of a finite Coxeter group. *Compositio Mathematica*, 85(3):333–373, 1993.

[71] L. Overbeck. Estimation for continuous branching processes. *Scandinavian Journal of Statistics. Theory and Applications*, 25(1):111–126, 1998.

[72] L. Overbeck and T. Rydén. Estimation in the Cox-Ingersoll-Ross model. *Econometric Theory*, 13(3):430–461, 1997.

[73] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 3rd corrected version 3rd edition, 2004.

[74] M. Rösler. Generalized Hermite Polynomials and the Heat Equation for Dunkl Operators. *Communications in Mathematical Physics*, 192(3):519–542, 1998.

[75] M. Rösler and M. Voit. Dunkl theory, convolution algebras, and related Markov processes. In P. G. et al., editor, *Harmonic and stochastic analysis of Dunkl processes.* Hermann, Paris, 2008.

[76] M. Rösler and M. Voit. Markov Processes Related with Dunkl Operators. *Advances in Applied Mathematics*, 21(4):575–643, 1998.

[77] O. X. Schlömilch. Über die Besselsche Funktion. *Zeitschrift für Mathematik und Physik*, 2:137–165, 1857.

[78] M. Sørensen. Estimation functions for diffusion-type processes. In M. K. et al., editor, *Statistical Methods for Stochastic Differential Equations*. Taylor & Francis Group, Paris, 2012.

[79] F. Testard. *Polarité points multiples et géométrie de certains processus gaussiens*. PhD thesis, Orsay, 1987.

[80] J. van Diejen and L. Vinet. *Calogero—Moser— Sutherland Models*. CRM Series in Mathematical Physics. Springer New York, 1st edition, 2000.

[81] J. F. van Diejen. Confluent hypergeometric orthogonal polynomials related to the rational quantum Calogero system with harmonic confinement. *Communications in Mathematical Physics*, 188:467–497, 1997.

[82] E. von Lommel. *Studien über die Bessel'schen Funktionen*. Teubner, 1868.

[83] G. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library. Cambridge University Press, 1995.

[84] N. Wiener. Differential-Space. *Journal of Mathematical Physics*, 2:131–174, 1923.

[85] N. Wiener. Un problème de probabilité dénombrables. *Bulletin de la Société Mathématique de France*, 52:569–578, 1924.

[86] H. Witting and U. Müller-Funk. *Mathematische Statistik II: Asymptotische Statistik: Parametrische Modelle und nichtparametrische Funktionale*. Vieweg+Teubner Verlag Wiesbaden, 2012.

[87] Y. Xiao. Asymptotic Results for Self-Similar Markov Processes. In B. Szyszkowicz, editor, *Asymptotic Methods in Probability and Statistics*, pages 323–340. North-Holland, Amsterdam, 1998.

[88] N. Yoshida. Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Annals of the Institute of Statistical Mathematics*, 63(3):431–479, 2011.

[89] A. Zygmund and R. Fefferman. *Trigonometric Series*. Cambridge Mathematical Library. Cambridge University Press, 3rd edition, 2003.

Mischief managed