Brownian semi-stationary processes, turbulence and smooth processes

PhD Thesis
by
Jóse Ulises Márquez Urbina
Department of Mathematics

Aarhus University
Aarhus, Denmark
November 2015

Brownian semi-stationary processes,
turbulence and smooth processes

PhD thesis by José Ulises Márquez Urbina
Department of Mathematics, Aarhus University,
Ny Munkegade 118, 8000 Aarhus C, Denmark

Supervised by Jürgen Schmiegel

Submitted November 30, 2015

Thesis advisor: Jürgen Schmiegel Jóse Ulises Márquez Urbina

Brownian semi-stationary processes, turbulence and smooth processes

Summary

This thesis analysis the use of Brownian semi-stationary (BSS) processes to model the main statistical features present in turbulent time series, and some asymptotic properties of certain classes of smooth processes.

Turbulence is a complex phenomena governed by the Navier-Stokes equations. These equations do not represent a fully functional model and, consequently, it has been necessary to develop phenomenological models capturing main aspects of turbulent dynamics. The BSS processes were proposed as an option to model turbulent time series. In this thesis we proved, through a simulation-based approach, the potential of BSS processes to model turbulent velocity time series. It turns out that this family of processes reproduces accurately some of the main features present in turbulent time series, such as the distribution of the velocity increments and the statistics of the Kolmogorov variable. We also studied the distributional properties of the increments of BSS processes with the intent to better understand why the BSS processes seem to accurately reproduce the temporal turbulent dynamics.

BSS processes in general are not semimartingales. However, there are conditions which make a BSS process a bounded variation process with differentiable paths. It is natural to inquire if it is possible to obtain an asymptotic theory for this class of BSS processes. This problem is investigated and some partial results are presented. The asymptotic theory for BSS processes naturally leads to the study of the same problem for multiple Lebesgue integrals of Brownian motion. This thesis also presents some research about the asymptotic problem in the context of integrals of Brownian motion.

Thesis advisor: Jürgen Schmiegel Jóse Ulises Márquez Urbina

Brownian semi-stationary processes, turbulence and smooth processes

Dansk resumé

I denne afhandling analyseres anvendelsen af Brownske semi-stationære (BSS) processer til modellering af de primære statistiske egenskaber i turbulente tidsrækker, derudover studeres de asymptotiske egenskaber for bestemte klasser af glatte processer.

Turbulens er et komplekst fænomen beskrevet af Navier-Stokes ligningerne. Disse udgør dog ikke en funktionsdygtig model, og, som konsekvens heraf, har det været nødvendigt at udvikle fænomenologiske modeller til beskrivelse af de væsentlige aspekter i turbulent dynamik. BSS processerne er, af flere, blevet foreslået som en mulighed til modellering af turbulente tidsrækker. I denne afhandling har vi, ved en simuleringsbaseret tilgang, studeret potentialet ved at bruge BSS processer til modellering af turbulente hastighedstidsrækker. Studiet har vist, at denne familie af processer, præcist gengiver nogle af de primære statistiske egenskaber i turbulente tidsrækker, som eksempelvis fordelingen af hastighedstilvækster og de statiske egenskaber af Kolmogorov variablen. Derudover, har vi studeret de distributive egenskaber af tilvæksterne i BSS processer, med henblik på at opnå en større forståelse af hvorfor BSS processerne præcist gengiver den temporale turbulens dynamik.

BSS processerne er generelt ikke semi-martingaler. Dog findes der betingelser, under hvilke en BSS proces har begrænset variation og differentiable stier. Det er naturligt at spørge om det er muligt at udlede en asymptotisk teori for denne klasse af BSS processer. Dette problem er undersøgt og delvise resultater bliver præsenteret. Den asymptotiske teori for BSS processer leder naturligt til studiet af samme problem for flere Lebesgue integraler af Brownsk bevægelse. Derudover præsenteres der også i denne afhandling forskning i det asymptotiske problem i forbindelse med integraler af Brownsk bevægelse.

Contents

Listing of figures

Acknowledgments

In these lines I would like to express my gratitude to all those people who supported me during my PhD studies and my years in Aarhus. Without your company and backup, I would not have been able to complete this work. Since it is very unlikely that my friends read this dissertation, and the people reading this work do not know my friends, I prefer to avoid writing the names of my buddies. This has the advantage that no one will be forgotten. If we had a good time during my PhD studies, regardless if it was a very short moment, my gratitude is for you.

The above recognition also includes the people who worked with me during my PhD project. However, I consider important to make an explicit recognition to those who inspired me to give the best in my academic work and oriented me during the process. I would like to start thanking Emil Hedevang. I learned a lot from the discussions carried out with him and he was a source of motivation. I am especially in debt for his assistance in the programing part of my project.

A special mention goes to professor Ole E. Barndorff-Nielsen. His enthusiasm and love for the mathematics is contagious. I really appreciate having worked with someone as keen as he is. Besides, I warmly thank him for the opportunity to study in Aarhus.

And last, but not least, I would like to thank my supervisor Jürgen Schmiegel for his patience and help to produce this work. I am grateful for his time and effort.

Preface

The work behind this thesis was financed by the National Mexican Council of Science and Technology (Conacyt) through the PhD grant 217040. This thesis was supervised by Jürgen Schmiegel. It consists of three projects:

Paper A: Márquez, J.U. and Schmiegel, J. Modelling turbulent time series by BSS-processes. In: Podolskij, M., Stelzer, R., Thorbjørnsen, S. and Veraart, A.E.D (Eds.): The fascination of Probability, Statistics and their Applications: In Honour of Ole Barndorff-Nielsen. Springer, Berlin. 2015.

Paper B: Márquez, J.U. On the cumulants of increments for two classes of Brownian semi-stationary processes. Submitted to Scandinavian Journal of Statistics.

Paper C: Barndorff-Nielsen, O.E., Márquez, J.U. and Pakkanen, M. An asymptotic problem for smooth processes. Submitted as Research Report, Thiele Centre for Applied Mathematics in Natural Sciences, University of Aarhus.

In **Paper A** it is shown that Brownian semi-stationary (BSS) processes are able to reproduce the main characteristics of turbulent data. Furthermore, an algorithm is presented that allows to estimate the model parameters from second and third order statistics. In **Paper B** formulae for the cumulants of the increments of BSS processes are obtained, assuming that the volatility is a Lévy semi-stationary process and the exponential of an ambit process. These formulae are applied to some specific examples that are relevant as models for turbulent velocity time series. **Paper C** contains some efforts to solve an asymptotic problem for some smooth BSS process and for the integral of a Brownian motion.

1

Background on turbulence

In this Chapter we present some preliminaries about turbulence and its relation to BSS processes.

1.1 Basics about turbulence

Turbulent flows are characterized by low momentum diffusion, high moment convection, and rapid variation of pressure and velocity in space and time. Flow that is not turbulent is called laminar flow. The non-dimensional Reynolds number Re characterizes whether flow conditions lead to laminar or turbulent flow. Increasing the Reynolds number increases the turbulent character and the limit of infinite Reynolds number is called the fully developed turbulent state.

Turbulence, as part of hydrodynamics, is governed by the Navier-Stokes equations which has been known since 1823. In general there is no known unique solution for these equations, and it is not possible to describe the wide range of turbulent phenomena from basic principles. Consequently, a great deal of phenomenological models have emerged that are based on and designed for certain aspects of turbulent dynamics.

In general, turbulence concerns the dynamics in a fluid flow of the three dimensional velocity vector $\vec{v}(t,\mathbf{x}) = (v_1(t,\mathbf{x}), v_2(t,\mathbf{x}), v_3(t,\mathbf{x}))$ as a function of position $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ and time t. A derived quantity is the energy dissipation, defined as

$$\varepsilon(t,\mathbf{x}) = \frac{\nu}{2} \sum_{i,j=1,2,3} \left(\partial_{x_i} v_j(t,\mathbf{x}) + \partial_{x_j} v_i(t,\mathbf{x})\right)^2, \tag{1.1}$$

describing the loss of kinetic energy due to friction forces characterized by the viscosity ν.

Insight into the process that transforms kinetic energy into heat can be gained through the Richardson cascade ([31]). In this representation kinetic energy is injected into the flow at large scales through large scale forcing. Non-linear effects redistribute the kinetic energy towards smaller scales. This cascade of energy stops at small scales where dissipation transforms kinetic energy into heat. It is traditional to call the large scale L of energy input the integral scale and the small scale η of dissipation the dissipation scale or Kolmogorov scale. With increasing Reynolds number the fraction L/η increases, giving space for the so called inertial range $\eta \ll l \ll L$ where turbulent statistics are expected to have some universal character.

The resolution of all dynamically active scales in experiments is at present not achievable for the full three-dimensional velocity vector. Most high-resolution experiments measure a time-series of one component v (in direction of the mean flow) of the velocity vector at a fixed single location $\mathbf{x}_0$. Based on this restriction one defines the temporal energy dissipation

$$\varepsilon_t = \varepsilon_t(\mathbf{x}_0) = \frac{15\nu}{\overline{v}^2}\left(\frac{dv}{dt}(t,\mathbf{x}_0)\right)^2 \tag{1.2}$$

where $\overline{v}$ denotes the mean velocity. The temporal energy dissipation (1.2) is expected to have similar statistical properties as the true energy dissipation (1.1) for stationary, homogeneous and isotropic flows. Discrepancies appear at small scales and are termed surrogacy effects. The derivation of (1.2) from (1.1) is based on Taylor's Frozen Flow Hypothesis ([37]) which states that spatial structures of the flow are predominantly swept by the mean velocity $\overline{v}$ without relevant distortion. Under this hypothesis, widely used in analyzing stationary turbulent time series, spatial increments along the direction of the mean flow (in direction x_1) are expressed in terms of temporal increments

$$v(t+s,x_1) - v(t,x_1) = v(t,x_1 - \overline{v}s) - v(t,x_1).$$

1.2 The Kolmogorov-Obukhov statistical theory of turbulence

O. Reynolds, considered as the father of the scientific research in turbulence, had already realized that the deterministic study of turbulence is impracticable and that its analysis should be in statistical terms [30]. There were some early efforts in this direction but it was Kolmogorov ([23, 24]) who, for the first time, was able to introduce a statistical theory for turbulence with important implications.

Kolmogorov proposed a theoretical framework for turbulence, sometimes referred to as K41 theory, which applies to homogeneous and isotropic turbulence. In this framework, Kolmogorov made two assumptions:

1. The energy dissipation rate has a finite non-vanishing limit as the viscosity tends to zero while keeping the scale and velocity characteristic of the production of turbulence fixed.

2. In the limit of very large Reynolds numbers, there is a scaling exponent $h \in \mathbb{R}$ such that the velocity increments $\vec{u}(\mathbf{l}; t, \mathbf{x}) \equiv \vec{v}(t, \mathbf{x}) - \vec{v}(t, \mathbf{x} + \mathbf{l})$ satisfy

$$\vec{u}(\lambda \mathbf{l}; t, \mathbf{x}) \stackrel{d}{=} \lambda^h \vec{u}(\mathbf{l}; t, \mathbf{x}), \ \ \forall \lambda \in \mathbb{R}_+,$$

   for all $\mathbf{x}$ and all increments $\lambda \mathbf{l}$ and $\mathbf{l}$ small compared to the integral scale.

Since in a laminar flow the dissipation goes to zero with the viscosity, the first assumption is generally called "the existence of a dissipative anomaly" and is well supported by experimental and numerical results. The second assumption, generally called the "self-similarity" hypothesis, holds only in an approximate way.

The K41 scaling hypothesis immediately implies scaling laws for the structure functions $\mathbf{S}_p$,

$$\mathbf{S}_p(\mathbf{l}; t, \mathbf{x}) = E\left\{\left(\vec{u}(\mathbf{l}; t, \mathbf{x}) \cdot \frac{\mathbf{l}}{\|\mathbf{l}\|_2}\right)^p\right\}, \ \ p \in \mathbb{N}, \ \ (t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^3, \mathbf{l} \in \mathbb{R}^3, \tag{1.3}$$

namely,

$$\mathbf{S}_p(\mathbf{l}; t, \mathbf{x}) \propto \|\mathbf{l}\|_2^{hp}. \tag{1.4}$$

The case $p = 3$ is remarkable since Kolmogorov showed that the four-fifths relation

$$\mathbf{S}_3(\mathbf{l}; t, \mathbf{x}) = -\frac{4}{5} E\{\varepsilon(t, \mathbf{x})\} \|\mathbf{l}\|_2 \tag{1.5}$$

holds without any need to assume self-similarity ([23]). All these scaling relations are meant to apply within the inertial range. Equation (1.5) is reasonably well supported by experimental and numerical data.

The four-fifths law (1.5) and the self-similarity hypothesis imply that $h = 1/3$. As the units of $\vec{u}$, ε and $\|\mathbf{l}\|_2$ are m/s, m^2/s^3 and m, respectively, necessarily

$$\mathbf{S}_p(\mathbf{l}; t, \mathbf{x}) = C_p E\{\varepsilon(t, \mathbf{x})\}^{p/3} \|\mathbf{l}\|_2^{p/3}, \tag{1.6}$$

where C_p are dimensionless constants. The constants C_p cannot depend on the Reynolds number, since the limit of infinite Reynolds number is already taken.

Remark 1 *Equation (1.6) implies that the energy spectrum of turbulence follows a $k^{-5/3}$ law (where k is the wavenumber). This is the so-called 5/3rd-Kolmogorov law.*

Remark 2 *If $v_t(\mathbf{x}_0)$ denotes the component of the velocity vector in direction of the mean flow at time t at the fixed position $\mathbf{x}_0$, the Taylor Frozen Flow Hypothesis ([37]) implies that the structure functions for temporal* *increments*

$$S_p(l;t,\mathbf{x}_o) = E\left\{\left(v_{t+l}(\mathbf{x}_o) - v_t(\mathbf{x}_o)\right)^p\right\}, \quad p \in \mathbb{N}, \ (t,\mathbf{x}_o) \in \mathbb{R}_+ \times \mathbb{R}^3, l > 0, \tag{1.7}$$

also satisfy

$$S_p(l;t,\mathbf{x}_o) = C_p E\left\{\varepsilon(t,\mathbf{x}_o)\right\}^{p/3} l^{p/3} \tag{1.8}$$

for some constants C_p, and l within the temporal counterpart of the inertial range $\eta/\bar{v} \ll l \ll L/\bar{v}$.

Remark 3 *In 1945 the physicist L. Onsager made a profound observation ([28]): any fluid velocity $\vec{v}$ that satisfies the Kolmogorov-Obukhov scaling can not have a gradient continuous in $\mathbf{x}$ (i.e. space). More precisely, in mathematical terms, $\vec{v}$ is Hölder continuous in $\mathbf{x}$ with Hölder index 1/3. This implies, as pointed out by B. Birnir, that the turbulent solutions of the Navier-Stokes equations are not smooth [13].*

Careful experimental and numerical examinations of the scaling laws reveal small but measurable discrepancies from K41. Indeed, the structure functions $\mathbf{S}_p$ display a power-law behavior $\mathbf{S}_p(\mathbf{l}) \propto \|\mathbf{l}\|_2^{\zeta_p}$ within the inertial range, but the scaling exponents ζ_p do not exhibit the exact linear behavior $\zeta_p = p/3$ predicted by K41. Hence the self-similarity assumed in K41 may actually be broken. Nowadays the scaling exponents ζ_p are known with great accuracy and they seem to be universal, that is independent of the mechanism by which the turbulence is driven [19].

The K41 theory was criticized by Landau for not taking into account the influence of the large flow structure on the constants C_p and for not including the influence of the intermittency in the velocity fluctuations on the scaling exponents [26]. In 1962, in order to address these problems, Kolmogorov published two hypotheses ([25]), usually referred to as K62, about a quantity V that combines velocity increments and the energy dissipation. In its original formulation, Kolmogorov considered spatial averaging over spheres. In the presentation below, we reformulate the hypotheses using spatial averaging along the mean flow direction x which allows to translate the spatial structure into the time domain using the Taylor Frozen Flow Hypothesis. The hypotheses are:

(i) Let $\Delta v_x(r) = v(x + r/2) - v(x - r/2)$, where $v(x)$ denotes the x component of the velocity vector $\vec{v}(t,x,y,z)$ at fixed time t. Let $r\varepsilon_r$ be the integrated energy dissipation over a domain of linear size r, where

$$\varepsilon_r(t,\mathbf{x}) = \frac{1}{r}\int_{x-r/2}^{x+r/2} \varepsilon(t,\sigma,x_2,x_3)\mathrm{d}\sigma. \tag{1.9}$$

Then, for $r \ll L$ (where L is the integral scale), the pdf of the stochastic variable

$$V_r(t,\mathbf{x}) = \frac{\Delta v_x(r)}{(r\varepsilon_r(t,\mathbf{x}))^{1/3}} \tag{1.10}$$

only depends on the local Reynolds number

$$\mathrm{Re}_r = r(r\varepsilon_r(t,\mathbf{x}))^{1/3}/\nu. \tag{1.11}$$

(ii) For $Re_r \gg 1$, the pdf of V_r does not depend on Re_r, either, and is therefore universal.

Although, for small r, an additional r dependence of the pdf of V_r has been observed ([36]), the validity of several aspects of K62 has been verified experimentally and by numerical simulation of turbulence (see e.g. [20]). In particular it has been shown that the conditional densities $p(V_r|r\varepsilon_r)$ become independent of $r\varepsilon_r$ for scales r within the inertial range. However, the universality of the distribution of V has not been verified in the literature. In this respect, it is important to note that the experimental verification of the Kolmogorov hypotheses for high resolution data is restricted to temporal statistics and as such relies on the use of the temporal energy dissipation (1.2) instead of the true energy dissipation (1.1). In the time domain, the Kolmogorov variable V is defined as

$$V_s = \frac{u_s(t-s/2)}{(\overline{v}s\overline{\varepsilon}(s,t))^{1/3}} \tag{1.12}$$

where $u_s(t) = v(t,\mathbf{x}_0) - v(t-s,\mathbf{x}_0)$ denotes the temporal velocity increment at time scale s, $\overline{v}$ is the mean velocity and $\overline{\varepsilon}(s,t)$ is the coarse grained temporal energy dissipation

$$\overline{\varepsilon}(s,t) = \frac{1}{s}\int_{t-s/2}^{t+s/2} \varepsilon_r dr.$$

We have skipped reference to the spatial location in (1.12).

1.3 Scaling of the energy dissipation

Kolmogorov's refined similarity hypothesis (i) implies that the coarse grained energy dissipation shows the scaling behavior $E\left\{\varepsilon_l(t,\mathbf{x})^p\right\} \propto l^{\tau_p}$. More precisely, Kolmogorov's hypothesis (i) provides a relation between the scaling exponents ζ_p of the structure functions $\mathbf{S}_p$ and the scaling exponents τ_p, namely

$$\zeta_p = \frac{p}{3} + \tau_{p/3}.$$

In the time domain, under stationary flow conditions, the corresponding scaling of $E\left\{\overline{\varepsilon}(s,t)^2\right\}$ is related to a scaling law for $E\left\{\varepsilon_0\varepsilon_t\right\}$. Motivated by this fact, one considers the correlators of order (n,m) of the (surrogate) temporal energy dissipation ε_t (1.2), denoted by $c_{n,m}(t)$, and defined as

$$c_{n,m}(t) := \frac{E\left\{\varepsilon_t^n\varepsilon_0^m\right\}}{E\left\{\varepsilon_t^n\right\}E\left\{\varepsilon_0^m\right\}} \quad n,m\in\mathbb{N}, t>0. \tag{1.13}$$

In [32] it was reported that for time-scales t within the temporal counterpart of the inertial range, one observes scaling

$$c_{n,m}(t) \propto t^{\tau(n,m)}.$$

The length of the scaling range can be extended by plotting $c_{n,m}(t)$ as a function of $c_{n,m}(t)$. This improved scaling for the correlators is called self-scaling.

1.4 Intermittency and the pdf of velocity increments

The Kolmogorov-Obukhov theory had a major influence in the study of turbulence. Numerous efforts have been dedicated to the analysis of the structure functions S_p and $\mathbf{S}_p$, and their scaling behavior. However, it could be more meaningful and enlightening to understand the probability density functions (pdf) of the velocity increments, rather than a set of moments.

Experimental and numerical studies have provided evidence that the densities of the velocity increments have heavy or semi-heavy tails at small scales, while they are almost Gaussian at large scale where energy is fed into the flow ([19]). This evolution of the pdf across scales is traditionally referred to as "aggregational Gaussianity" in turbulence studies. A typical scenario is characterized by an approximate Gaussian shape for the large scales, turning to exponential tails for the intermediate scales and stretched exponential tails for dissipation scales. The deviation from Gaussianity at small scales is refered to as "intermittency".

In [6], it is reported that the evolution of the pdf of temporal velocity increments, for all amplitudes and all scales, can well be described within the class of normal inverse Gaussian (NIG) distributions (see Appendix D).

2

Ambit processes and Brownian semi-stationary processes

In this Chapter, we include a brief review of some aspects of ambit processes and Brownian semi-stationary (BSS) processes that are essential in Paper A for the modelling of turbulent velocity time series, and in Paper B for the analysis of the cumulants of increments of BSS processes. BSS processes are fundamental in this thesis.

This Chapter is divided in two Sections. Section 2.1 discusses the definition of a very simple version of an ambit process, relevant for the modelling of the turbulent energy dissipation. Section 2.2 presents the most relevant elements of BSS processes.

2.1 Integration with respect to a Lévy basis

Ambit processes were introduced in [9] as a framework for tempo-spatial modelling. These processes are defined in terms of integrals with respect to a Lévy basis. Here, we restrict our attention to those ambit processes defined as the stochastic integral of a deterministic function with respect to a homogeneous Lévy basis defined on $\mathbb{R}^2$.

Denote by $\mathcal{B}_b(\mathbb{R}^2)$ the set of bounded Borel subsets of $\mathbb{R}^2$. A Lévy basis Λ on $\mathbb{R}^2$ is an infinitely divisible, independently scattered random measure on $\mathbb{R}^2$, i.e. $(\Lambda(A))_{A\in\mathcal{B}_b(\mathbb{R}^2)}$ is a stochastic process such that: (i) $\Lambda(A)$ is infinitely divisible; (ii) $\Lambda(A)$ and $\Lambda(B)$ are independent if $A\cap B=\varnothing$; and, (iii) If $A_1,\ldots,A_n\in\mathcal{B}_b(\mathbb{R}^2)$ are disjoint and such that $\cup_{i=1}^n A_i\in\mathcal{B}_b(\mathbb{R}^2)$, then

$$\Lambda\left(\bigcup_{i=1}^n A_i\right)\overset{a.s.}{=}\sum_{i=1}^n \Lambda(A_i).$$

A Lévy basis Λ on $\mathbb{R}^2$ is called homogeneous if $\Lambda(A)\overset{d}{=}\Lambda(A+x_0)$, for $x_0\in\mathbb{R}^2$.

The stochastic integral $\int f d\Lambda$ of a deterministic measurable function $f:\mathbb{R}^2\to\mathbb{R}$ with respect to a homogeneous Lévy basis Λ is defined in two steps: (a) If $f=\sum_{i=1}^n a_i 1_{A_i}$ is a real simple function on $\mathbb{R}^2$ with $A_1,\ldots,A_n$ disjoint, for $A\in\mathcal{B}(\mathbb{R}^2)$, we define

$$\int_A f d\Lambda=\sum_{i=1}^n a_i\Lambda(A_i\cap A).$$

(b) If $f:\mathbb{R}^2\to\mathbb{R}$ can be approximated almost everywhere (with respect to the Lebesgue measure) by a sequence of simple functions $\{f_n\}$ as in (a), provided that the limit exist, we define

$$\int_A f d\Lambda = P-\lim\int_A f_n d\Lambda, \tag{2.1}$$

for $A\in\mathcal{B}(\mathbb{R}^2)$. We say that a measurable function $f:\mathbb{R}^2\to\mathbb{R}$ is Λ-integrable if the integral (2.1) exists.

Let $K\{z\ddagger X\}:=\log E\{\exp(zX)\}$ and $C\{z\ddagger X\}:=\log E\{\exp(izX)\}$ denote the log-moment generating function and the log-characteristic function, respectively, of the random variable X. The functions K and C will be called the kumulant and cumulant function, respectively. For each homogeneous Lévy basis Λ, we can associate a random variable Λ' to Λ such that

$$K\{z\ddagger\Lambda(da)\}=K\{z\ddagger\Lambda'\}\,da,$$

and

$$C\{z\ddagger\Lambda(da)\}=C\{z\ddagger\Lambda'\}\,da.$$

The random variable Λ' is called the Lévy seed of Λ.

The stochastic integral $\int f d\Lambda$ and the Lévy seed Λ' satisfy the next relation (see [29] for a proof).

Proposition 4 *Let Λ be a Lévy basis on $\mathbb{R}^2$ and $f:\mathbb{R}^2\to\mathbb{R}$ a Λ-integrable function. Then*

$$K\left\{z\ddagger\int_A f(a)\,d\Lambda(a)\right\}=\int_A K\{zf(a)\ddagger\Lambda'\}\,da$$

and

$$C\left\{z \ddagger \int_A f(a)\, d\Lambda\,(a)\right\} = \int_A C\left\{zf(a) \ddagger \Lambda'\right\} da.$$

For the purposes of this thesis, an ambit process is a stochastic process $(Y_t)_{t\geq 0}$ of the form

$$Y_t = \int_{A_t} f((0,t) - a)\, d\Lambda\,(a)\,,$$

where $A \in \mathcal{B}_b\left(\mathbb{R}^2\right)$ and $A_t = A + (0,t)$. For a more general definition of ambit processes and a discussion of their mathematical properties, we refer to [5].

2.2 Brownian semi-stationary processes

Brownian semi-stationary (BSS) processes, introduced in [11] as potential models for turbulent velocity time series, are stochastic processes of the form

$$Z_t = \mu + \int_{-\infty}^{t} g\,(t-s)\,\sigma_s \mathrm{d}B_s + \int_{-\infty}^{t} q\,(t-s)\,\varsigma_s \mathrm{d}s, \tag{2.2}$$

where μ is a constant, $(B_t)_{t\in\mathbb{R}}$ is standard Brownian motion, g and q are nonnegative deterministic functions on $\mathbb{R}$, with $g\,(t) = q\,(t) = 0$ for $t \leq 0$, and $(\sigma_t)_{t\in\mathbb{R}}$ and $(\varsigma_t)_{t\in\mathbb{R}}$ are càdlàg processes. When (σ, ς) is stationary and independent of B, then Z is stationary. In this thesis, only stationary BSS processes are considered.

In general, BSS processes are not necessarily semimartingales. A sufficient condition for Z to be a semimartingale is that σ and ς have second finite moments, $g, q \in L^1\left(\mathbb{R}_+\right) \bigcap L^2\left(\mathbb{R}_+\right)$, $g' \in L^2\left(\mathbb{R}_+\right)$ and $g\,(0+) < \infty$ (see [16]).

It is well-known that for any semimartingale X the limit

$$[X]_t = \lim_{n\to 0} \sum_{j=1}^{n} \left(X_{j\frac{t}{n}} - X_{(j-1)\frac{t}{n}}\right)^2 \tag{2.3}$$

exists as a limit in probability. The derived process $[X]$ expresses the cumulative variation exhibited by X and is called quadratic variation. For the case where (2.2) is a semimartingale, using the stochastic Fubini theorem and Itô algebra, we get

$$(\mathrm{d}Z_t)^2 = g^2\,(0+)\,\sigma_t^2 \mathrm{d}t$$

and

$$[Z]_t = \int_0^t (\mathrm{d}Z_s)^2 = g^2\,(0+) \int_0^t \sigma_s^2 \mathrm{d}s. \tag{2.4}$$

When X is not a semi-martingale, the limit in (2.3) might not exist. However, as has been proven in

[16], under certain assumptions on g it follows that

$$\lim_{n\to 0} \frac{1}{n \cdot R(1/n)} \sum_{j=1}^{n} \left(Z_{j\frac{t}{n}} - Z_{(j-1)\frac{t}{n}} \right)^2 = \int_0^t \sigma_s^2 ds, \tag{2.5}$$

where

$$R(s) = E\left[(G_s - G_0)^2 \right], \quad G_s = \int_{-\infty}^{s} g(s-r)\, dB_r.$$

Equation (2.5) is a particular case of the more general result established in [16].

The specification of the kernel g proportional to a gamma density,

$$g(x; a, \nu, \lambda) = a \cdot x^{\nu-1} \exp(-\lambda x)\, 1_{(0,\infty)}(x), \qquad \lambda > 0,\ \nu > 0, \tag{2.6}$$

is of particular interest since, for this class of kernels, the BSS process (2.2) has some nice mathematical and modelling properties. From the mathematical perspective, we have that, when g is a gamma kernel (2.6), the semimartingale character of (2.2) is determined by the value of the exponent ν: Z is a semimartingale if and only if $\nu > 3/2$. Besides, this type of kernels form the most simple class of functions that fulfill the assumptions necessary to satisfy (2.5). From the modelling point of view, we have that BSS processes with a gamma kernel reproduce Kolmogorov's 5/3rd law (see Remark 1).

The properties mentioned in the above paragraph make the BSS process with a gamma kernel a relevant object of study in this thesis.

3

Modelling turbulent time series by BSS-processes

This Chapter summarizes the main results of the application of BSS processes to model turbulent velocity time series (Paper A).

One of the main ingredients of BSS processes (2.2) is the volatility σ which, in the approach described here constitutes the energy dissipation process $\varepsilon = \sigma^2$, and will be modelled as the exponential of an ambit process. Section 3.1 briefly describes the ambit model for the energy dissipation. Section 3.2 presents the BSS model for turbulent velocity time series whose modelling performance will be illustrated in Section 3.4. Section 3.3 describes the algorithm used to estimate the model parameters.

3.1 A Lévy based model for the turbulent energy dissipation

The model for the stationary turbulent velocity field takes as an input the temporal energy dissipation which is modelled as the exponential of an ambit process. In [35] it is shown that this Lévy based approach is able to reproduce the main stylized features of the energy dissipation observed for a wide range of data sets, including the data analyzed in this thesis.

Specifically, we model the stationary temporal energy dissipation ε as the exponential of an integral with respect to a homogeneous Lévy basis L on $\mathbb{R}^2$,

$$\varepsilon_t = \exp\left(\int_{A(t)} L(\mathrm{d}y, \mathrm{d}s)\right) = \exp\left(L\left(A\left(t\right)\right)\right) \tag{3.1}$$

where $A(t) = A + (0, t)$ for a bounded set $A \subset \mathbb{R}^2$. The ambit set A is given as

$$A = \{(x, t) : 0 \le t \le T, -f(t) \le x \le f(t)\}. \tag{3.2}$$

For $T > 0$, $k > 1$, and $\theta > 0$, the function f is defined as

$$f(t) = \left(\frac{1 - (t/T)^\theta}{1 + (k \cdot t/T)^\theta}\right)^{1/\theta}, \quad 0 \le t \le T. \tag{3.3}$$

This specification of the ambit set is adapted to reproduce the empirically observed scaling of correlators (1.13).

In [35] it is shown that the density of the logarithm of the energy dissipation is well described by a normal inverse Gaussian distribution, i.e. $\log \varepsilon_t \sim \mathrm{NIG}(\alpha, \beta, \mu, \delta)$. For the correlators $c_{p,q}$ to exist it is necessary to assume that ε_t has exponential moments of order $p+q$ leading to the condition $p+q < \alpha - \beta$. The parameters of the underlying NIG-law of L and the corresponding ambit sets have been estimated in [35] for a number of turbulent data sets including the one used in this thesis.

3.2 A stochastic model for turbulent velocity time series

Inspired by [10], stationary time series of the main component v_t of the turbulent velocity field are modelled as a BSS process of the specific form

$$v_t = v_t(g, \sigma, \beta, x_0) = \int_{-\infty}^{t} g(t - s + x_0)\, \sigma_s \mathrm{d}B_s + \beta \int_{-\infty}^{t} g(t - s + x_0)\, \sigma_s^2 \mathrm{d}s \equiv R_t + \beta Q_t \tag{3.4}$$

where g is a convolution of gamma kernels (see (3.6)), $\sigma^2 = \varepsilon$ for ε given by (3.1), and β and x_0 are positive constants.

The introduction of a cut-off x_0 ensures that $[v]$ is positive. Equation (2.4) implies that σ^2 can be identified, up to a factor, with the temporal energy dissipation.

In this set-up, the Kolmogorov variable V (1.12) is given as

$$V_s = \frac{u_s(t - s/2)}{\left(\bar{v}\,[v]_s\right)^{1/3}}. \tag{3.5}$$

where $u_s(t - s/2) = v_{t+s/2} - v_{t-s/2}$ and $\bar{v} = E\{v_t\}$. The conditional independence of V_s refers to the independence of $p\left(V_s|\left[v\right]_s\right)$ on $\left[v\right]_s$. Here $\left[v\right]_s$ denotes the quadratic variation over the time horizon $[t-s/2, t+s/2]$.

The term $\beta \cdot Q$ in (3.4) determines the skewness of the density of velocity increments $u_t = v_t - v_0$. For this reason we refer to β as the skewness parameter.

The kernel $g \equiv g\left(\cdot, a_1, v_1, \lambda_1, a_2, v_2, \lambda_2\right)$ can be expressed as

$$g\left(x; a_1, v_1, \lambda_1, a_2, v_2, \lambda_2\right) = a_1 a_2 x^{v_1+v_2-1} 1_{\mathbb{R}_+}(x) \int_0^1 e^{-x(\lambda_1 u + \lambda_2(1-u))} u^{v_1-1} (1-u)^{v_2-1} \mathrm{d}u. \tag{3.6}$$

The relevant parameters are $\left(a, v_1, \lambda_1, v_2, \lambda_2\right)$ with $a = a_1 a_2$. From (3.6) it follows that $v\left(g, \sigma, \beta, x_0\right)$ is a semimartingale if $v_1 + v_2 > 3/2$. This choice for g in (3.4) reproduces accurately the spectral density function (sdf) of turbulent velocity time series and provides physical interpretation of the parameters involved. Consider the model $v\left(g, \sigma, \beta, x_0\right)$ and assume that $\lambda_1 < \lambda_2$. For $v_1 = 5/6$, λ_1 denotes the frequency where the inertial range starts and λ_2 denotes the frequency where the inertial range ends. Furthermore, $2v_1$ and $2\left(v_1 + v_2\right)$ give the slope in the sdf within the inertial and the dissipation ranges, respectively. Thus, Kolmogorov's 5/3rd law implies that $v_1 \approx 5/6$.

3.3 Estimation procedure

The modelling framework (3.4) has three degrees of freedom: the energy dissipation σ^2, the kernel g and the skewness parameter β. The estimation of the parameters of σ^2 from data has been performed in [35]. It remains to estimate g and β.

The third order structure function S_3 (1.8) of (3.4) can be written as

$$S_3\left(l\right) = 3\beta E\{(\Delta_l R)^2 \left(\Delta_l Q\right)\} + \beta^3 E\{(\Delta_l Q)^3\}, \tag{3.7}$$

where R and Q are given as in (3.4), and $\Delta_l R = R_l - R_0$, $\Delta_l Q = Q_l - Q_0$, for $l > 0$. Given paths of R and Q, our estimator for β is the value that minimizes the distance, in the sense of least squares, between the empirical third order structure function and (3.7) for a suitable range of scales l. Besides, given a value of β, our estimated parameters of g are those that minimize the distance, also in the sense of least squares, between the empirical sdf and the sdf of (3.4).

The complete estimation procedure can be described as follows: **1)** We neglect the skewness parameter β and we estimate the parameters of g from the sdf. **2)** Having a simulation of the σ process, we perform a simulation of (3.4). **3)** Using the simulation of the BSS process produced in 2), we estimate β as described above. **4)** We re-estimate the kernel g using the empirical sdf and the current value of β. **5)** With the same simulation of σ, we repeat steps 2) and 3) until we observe stabilization.

Under certain assumptions, we can ensure that the estimation procedure gives reasonable values for the skewness parameter β and the parameters of g. We refer to Paper A for a detailed presentation of these

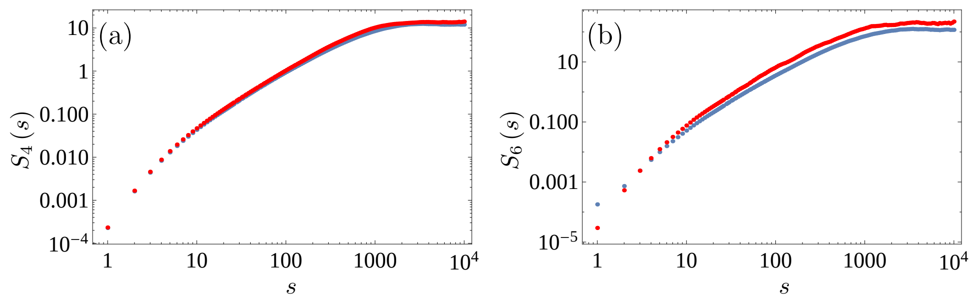

Figure 3.4.1: Comparison of the structure functions S_n, $n = 4, 6$ from the simulation of (3.4) (red) and the structure functions estimated from the data (blue). The time lags s are measured in units of the finest resolution of the empirical data.

conditions.

3.4 Model performance

In this Section, we present some exemplifying Figures illustrating the performance of (3.4) for modelling turbulent time series.

Figure 3.4.1 presents the structure functions S_4 and S_6 for empirical data and a simulation of the model (3.4). The data are measurements of time series of the main component of the turbulent velocity vector in a hellium jet experiment. The model shows very good agreement. It is important to note that the parameters of the model are completely estimated from the energy dissipation statistics and the structure functions S_2 and S_3 with no adjustable parameter for tuning the behavior of S_4 and S_6. Besides, the model reproduces excellently the structure order functions S_2 and S_3, as can be seen in Paper A.

Figure 3.4.2 shows the densities of velocity increments u_s for time lags $s = 1, 64$. In Paper A it is shown in more detail that the model (3.4) accurately reproduces the evolution of the densities of velocity increments across scales. NIG distributions fit these densities very well for different scales and amplitudes in full agreement with the results reported in [6].

Figure 3.4.3 presents the conditional densities $p\left(V_t | [v]_t\right)$ of the Kolmogorov variable V for some values of $[v]_t$. This Figure shows that, for t within the temporal counterpart of the inertial range, $p\left(V_t | [v]_t\right)$ is independent of $[v]_t$ for simulations and empirical data.

A similar excellent agreement between data and simulation is also obtained for energy dissipation correlators (see Paper A).

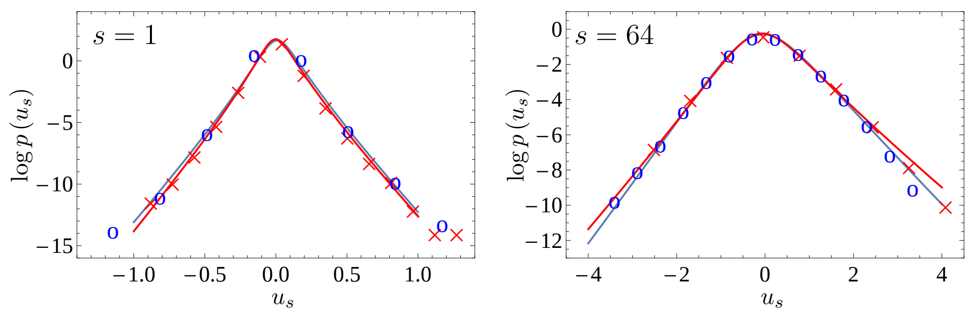

Figure 3.4.2: Comparison of the densities of velocity increments $p(u_s)$, $s = 1, 64$ from data (blue circles) and from the simulation of (3.4) (red crosses). The solid lines correspond to fitted NIG-distributions based on maximum likelihood estimation.

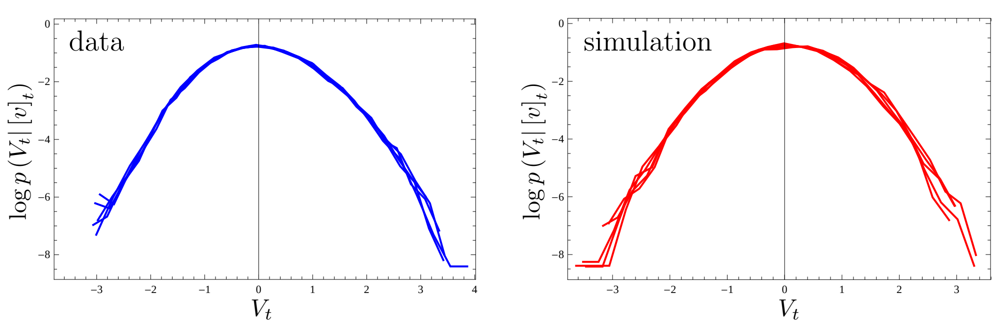

Figure 3.4.3: Comparison of the conditional densities $p\left(V_t | [v]_t\right)$ of the Kolmogorov variable from the data and from the simulation of (3.4) for $t = 128$ (in units of the finest resolution of the empirical data) and values $[v]_t = 0.8, 0.9, 1, 1.1, 1.2$.

4

On the cumulants of increments for two classes of Brownian semi-stationary processes

In this Chapter, we study the distribution of the increments of two classes of BSS processes via their cumulants.

More specifically, we find formulae for the cumulants of the increments of (2.2) assuming that $\sigma = \varepsilon$, $\varsigma = \varepsilon^2$ and where ε^2 has two forms: 1) ε^2 is a Lévy semi-stationary process (see, e.g., [40]); 2) ε^2 is the exponential of an ambit process driven by a homogeneous Lévy basis on $\mathbb{R}^2$. In the first case, the formulae we find are given in terms of the Lévy seed of the Lévy process that drives ε^2. In the second case, we obtain a formula for the n-th moments of the increments of Z in terms of the Lévy seed of the Lévy basis that drives $\log \varepsilon^2$. These formulae can be used to iteratively compute the cumulants of the increments of Z. The second case was used in Paper A and Chapter 3 to model turbulent velocity time series.

The increments of the process Z contain relevant information about Z itself, as well as of its elements. For instance, the increments may indicate how much the process Z varies, e.g. if it has finite quadratic variation (see [7]). The increments of Z also provide a way to estimate its parameters (see [16]; Paper A). In addition, the increments of Z are relevant for the theory of turbulence since the increments of the velocity field are the object of study in the Kolmogorov-Obukhov theory, which is probably the most important theory in turbulence.

This Chapter is organized as follows. In Section 4.1 we derive a formula for the cumulants of the increments of a BSS process assuming that ε^2 is a Lévy semi-stationary process (LSS). In Section 4.2 we apply the formula derived in Section 4.1 to some specific examples of Z. Section 4.3 presents a formula for the n-th moments of ε^2 assuming that it is the exponential of an ambit process. We use this formula to compute iteratively the cumulants of the increments of Z. In Section 4.4 we apply the formula derived in Section 4.3 to some specific examples of Z.

4.1 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as a LSS process: theory

Consider the BSS process

$$X_t = \int_{-\infty}^{t} g(t-s)\,\varepsilon_s dB_s + \int_{-\infty}^{t} q(t-s)\,\varepsilon_s^2 ds \tag{4.1}$$

where $g \in L^2(\mathbb{R})$ and $q \in L^1(\mathbb{R})$, with $g(x) = q(x) = 0$ for $x \leq 0$, and ε is a Lévy semi-stationary process (LSS) independent of B given by

$$\varepsilon_t^2 = \int_{-\infty}^{t} h(t-s)\,dL_s, \tag{4.2}$$

where, under the truncation function $\tau \equiv 0$, L is a subordinator with characteristic triplet $(m, 0, \nu)$, and $h \in L^1(\mathbb{R})$ is non-negative satisfying $h(x) = 0$ for $x \leq 0$ and

$$\int_{-\infty}^{t} \int_{\mathbb{R}_+} \left(1 \wedge h^2(t-s)\,x^2\right) \nu(dx)\,ds < \infty. \tag{4.3}$$

Besides, we assume that L has finite first moment, i.e. that

$$E[L_1] = \int_{\mathbb{R}_+} x\nu(dx) < \infty. \tag{4.4}$$

Under these requirements, the process X is well-defined. Condition (4.3) ensures that the process ε^2 is well-defined (see [12], Corollary 4.1). In the remaining part of this Section, we deduce a formula for the cumulants of the increments of X in terms of the cumulants of $L' \equiv L_1$.

Cumulants of $\Delta_t X$ relative to the cumulants of L'

For $t > 0$, let $\Delta_t X = X_t - X_0$ and

$$\phi_t(s) = (g(t-s) - g(-s)), \text{ and } \psi_t(s) = (q(t-s) - q(-s)). \tag{4.5}$$

Thus, $\Delta_t X$ can be expressed as

$$\Delta_t X = \int_{\mathbb{R}} \phi_t(s) \varepsilon_s dB_s + \int_{\mathbb{R}} \psi_t(s) \varepsilon_s^2 ds.$$

The cumulant function $C\{z \ddagger \Delta_t X\} := \log E(\exp\{iz\Delta_t X\})$ of $\Delta_t X$ is given as

$$\begin{aligned} C\{z \ddagger \Delta_t X\} &= \log E\left(\exp\left\{-\frac{1}{2}z^2 \int_{\mathbb{R}} \phi_t^2(s) \varepsilon_s^2 ds + iz \int_{\mathbb{R}} \psi_t(s) \varepsilon_s^2 ds\right\}\right) \\ &= \log E\left(\exp\left\{-\frac{1}{2}z^2 \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_t^2(s) h(s-r) dL_r ds + iz \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_t(s) h(s-r) dL_r ds\right\}\right). \end{aligned} \tag{4.6}$$

The formula for the cumulants of $\Delta_t X$ in terms of the cumulants of L' that we obtain in this Section is a consequence of the Fubini Theorem [3, Theorem 3.1] applied to the double integrals in (4.6).

Lemma 5 *The elements of the process X satisfy:*

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \phi_t^2(s) h(s-r) dL_r ds = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_t^2(s) h(s-r) ds dL_r, \tag{4.7}$$

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \psi_t(s) h(s-r) dL_r ds = \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_t(s) h(s-r) ds dL_r. \tag{4.8}$$

Proof. See Paper B. ■

For $t > 0$ and $t > r$, define $C_1(t,r)$ and $C_2(t,r)$ as

$$C_1(t,r) = \int_r^t \psi_t(s) h(s-r) ds, \quad C_2(t,r) = \int_r^t \phi_t^2(s) h(s-r) ds \tag{4.9}$$

when the integrals exist, and set $C_1(t,r) = C_2(t,r) = 0$ otherwise. Furthermore, for $(t,z) \in \mathbb{R}_+ \times \mathbb{R}$ and $t > r$, define $H_{t,z} \equiv izC_1(t,r) - \frac{1}{2}z^2 C_2(t,r)$. Using the functions C_1 and C_2, we shall express the cumulant function of $\Delta_t X$ in terms of the cumulants of L'.

Equation (4.6) and Lemma 5 imply

$$C\{z \ddagger \Delta_t X\} = \log E\left(\exp\left\{\int_{-\infty}^t H_{t,z}(r) dL_r\right\}\right)$$

Thus, assuming that all cumulants of the Lévy seed L' exist,

$$C\left\{z \ddagger \Delta_t X\right\} = \int_{-\infty}^{t} C\left\{H_{t,z}\left(r\right) \ddagger L'\right\} dr = \int_{-\infty}^{t} \sum_{m=1}^{\infty} \frac{\kappa_m\left(L'\right)}{m!} H_{t,z}^{m}\left(r\right) dr, \tag{4.10}$$

where $\kappa_m\left(V\right)$ denotes the m-th cumulant of the random variable V. From this last equation we can obtain a formula for the cumulant function of $\Delta_t X$ in terms of the cumulants of L'.

Proposition 6 *Let $t > 0$ and assume that all cumulants of L' exist. Besides, let $C_1\left(t,r\right)$ and $C_2\left(t,r\right)$ be as in (4.9) and, for $z \in \mathbb{R}$, set $H_{t,z}\left(r\right) = izC_1\left(t,r\right) - \frac{1}{2}z^2 C_2\left(t,r\right)$. If*

$$\sum_{m=1}^{\infty} \frac{\kappa_m\left(L'\right)}{m!} \int_{-\infty}^{t} \left|H_{t,z}\left(r\right)\right|^m dr < \infty \tag{4.11}$$

for $z \in D \subseteq \mathbb{R}$, then

$$C\left\{z \ddagger \Delta_t X\right\} = \sum_{j=1}^{\infty} \frac{(iz)^j}{j!} \left(\int_{-\infty}^{t} j! \sum_{m=\lceil j/2 \rceil}^{j} \binom{m}{2m-j} C_1^{2m-j}\left(t,r\right) C_2^{j-m}\left(t,r\right) \frac{\kappa_m\left(L'\right)}{2^{j-m} \cdot m!} dr \right), \tag{4.12}$$

for $z \in D$.

*Proof***.** See Paper B. ■

The next proposition provides a simple condition on the functions C_1 and C_2 to satisfy (4.11).

Proposition 7 *Let $t > 0$ and assume that all cumulants of L' exist. Besides, let $C_1\left(t,r\right)$ and $C_2\left(t,r\right)$ be as in (4.9) and, for $z \in \mathbb{R}$, set $H_{t,z}\left(r\right) = izC_1\left(t,r\right) - \frac{1}{2}z^2 C_2\left(t,r\right)$. Assume there exist $\alpha \geq 0$, $\eta > 0$ and $M\left(t\right) > 0$ such that*

$$\max\left\{\left|C_1\left(t,r\right)\right|, \left|C_2\left(t,r\right)\right|\right\} \leq M\left(t\right)\left|r\right|^{\alpha} e^{\eta r}. \tag{4.13}$$

Then $H_{t,z}$ satisfies condition (4.11) for $|z|$ sufficiently small.

*Proof***.** See Paper B. ■

Remark 8 *From the proof of Proposition 7, it is easy to see that the condition*

$$\max\left\{\left|C_1\left(t,r\right)\right|, \left|C_2\left(t,r\right)\right|\right\} \leq M\left(t\right)\left(\left|r\right|^{\alpha} \vee 1\right) e^{\eta r}$$

also implies that $H_{t,z}$ satisfies condition (4.11) for $|z|$ sufficiently small.

As a consequence of Proposition 6, we have that the cumulants of $\Delta_t X$ can be expressed in terms of the cumulants of L'.

Corollary 9 *Assume that all cumulants of L' exist. Let $t > 0$. If the conditions (4.11) or (4.13) are satisfied, then we have the relation*

$$\kappa_j(\Delta_t X) = j! \sum_{m=\lceil j/2 \rceil}^{j} \binom{m}{2m-j} \frac{\kappa_m(L')}{2^{j-m} \cdot m!} \int_{-\infty}^{t} C_1^{2m-j}(t,r)\, C_2^{j-m}(t,r)\, dr \quad j \in \mathbb{N}. \tag{4.14}$$

Remark 10 *If we assume that $q \equiv 0$ in (4.1), then $C_1(t,r) \equiv 0$. Under the condition (4.11), this implies that*

$$\kappa_j(\Delta_t X) = \begin{cases} j! \frac{\kappa_{j/2}(L')}{2^{j/2} \cdot \left(\frac{j}{2}\right)!} \int_{-\infty}^{t} C_2^{j/2}(t,r)\, dr & \text{if } j \text{ is even} \\ \\ 0 & \text{if } j \text{ is odd} \end{cases}, \tag{4.15}$$

for $j \in \mathbb{N}$.

In the derivation of the formulae (4.12) and (4.14), we have assumed that all cumulants of the Lévy seed L' exist. This is a very strong and restrictive condition. More generally, for the formulae (4.12) and (4.14) to hold it is only necessary to have that $E\left\{(L')^j\right\} < \infty$ for some $j \in \mathbb{N}$. Then, the cumulant function of L' can be written as

$$C\{z \ddagger L'\} = \sum_{m=1}^{j} \frac{\kappa_m(L')}{m!} (iz)^m + o\left(|z|^j\right) \quad \text{as } z \to 0. \tag{4.16}$$

Relation (4.16) can be used to reproduce the same arguments that led to the formulae (4.12) and (4.14). In this case, condition (4.11) can be replaced by

$$\sum_{m=1}^{j} \frac{\kappa_m(L')}{m!} \int_{-\infty}^{t} |H_{t,z}(r)|^m\, dr < \infty.$$

4.2 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as a LSS process: examples

In this Section we study the cumulants of $\Delta_t X$ for three examples of X given by (4.1), assuming that ε^2 is a LSS process given by (4.2). We assume that all cumulants of the Lévy seed L' exist. For the first example, $q \equiv 0$, $g(x) = e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, for $\lambda > 0$, and ε^2 is an Ornstein-Uhlenbeck process. This specification of X permits to obtain closed expressions for the cumulants of $\Delta_t X$ in terms of hypergeometric functions. In the second example, g and q are proportional to a gamma density and, again, ε^2 is an Ornstein-Uhlenbeck. In the last example, g, q and the kernel h of ε^2 are proportional to a gamma density. In this last case, the marginals of (4.1) are generalized hyperbolic distributions. The last three examples are relevant in the context of turbulence modelling.

Example 1: $g(x) = e^{\lambda x} 1_{\mathbb{R}_+}(x)$, Ornstein-Uhlenbeck process ε^2

For $\lambda, \rho > 0$ and $2\lambda \neq \rho$, consider the model (4.1), (4.2) with $g(x) = e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, $q \equiv 0$, $h(x) = e^{-\rho x} 1_{\mathbb{R}_+}(x)$ and L a subordinator. Then, for $t > 0$,

$$\phi_t^2(s) = e^{2\lambda s}\left(1 - e^{-\lambda t}\right)^2 1_{\mathbb{R}_-}(s) + e^{-2\lambda(t-s)} 1_{(0,t)}(s)$$

and

$$\begin{aligned} C_2(t,r) &= \frac{e^{-2\lambda t}}{2\lambda - \rho} e^{\rho r}\left(e^{(2\lambda-\rho)t} - e^{(2\lambda-\rho)r}\right) 1_{\mathbb{R}_+}(r) \\ &\quad + \frac{e^{\rho r}}{2\lambda - \rho}\left(\left(1 - e^{-\lambda t}\right)^2\left(1 - e^{(2\lambda-\rho)r}\right) + e^{-2\lambda t}\left(e^{(2\lambda-\rho)t} - 1\right)\right) 1_{\mathbb{R}_-}(r), \end{aligned}$$

where ϕ and C_2 are the functions defined in (4.5) and (4.9), respectively. The function C_2 satisfies the condition (4.13) in Proposition 7. According to the formula (4.15), assuming that L' has finite moments of all orders, we have that $\kappa_{2m-1}(\Delta_t X) = 0$, for $m \in \mathbb{N}$, and

$$\begin{aligned} \kappa_{2m}(\Delta_t X) = & \frac{\kappa_m(L')(2m)!}{2^m m!(2\lambda-\rho)^m \rho m}\left\{ d_1^m(t)\, {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1 + \frac{\rho m}{2\lambda-\rho}, \frac{\left(1-e^{-\lambda t}\right)^2}{d_1(t)}\right)\right. \\ & \left. + \frac{\Gamma(1+m)\Gamma\left(1+\frac{\rho m}{2\lambda-\rho}\right)}{\Gamma\left(1+m+\frac{\rho m}{2\lambda-\rho}\right)} - e^{-\rho t m}\, {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1 + \frac{\rho m}{2\lambda-\rho}, e^{-(2\lambda-\rho)t}\right)\right\} \end{aligned}$$

for $m \in \mathbb{N}$, where

$$d_1(t) \equiv \left(\left(1-e^{-\lambda t}\right)^2 + e^{-2\lambda t}\left(e^{(2\lambda-\rho)t} - 1\right)\right)$$

and ${}_2F_1$ is the Gaussian hypergeometric function. A detailed derivation of the formula for $\kappa_{2m}(\Delta_t X)$ can be found in Paper B.

In the following, for $m \in \mathbb{N}$ and $t \geq 0$, $\overline{\kappa}_m(\Delta_t X)$ denotes the normalized cumulant $\overline{\kappa}_m(\Delta_t X) \equiv \kappa_m(\Delta_t X)/\kappa_m(L')$, where L' is the Lévy seed of the process driving ε^2.

The first two normalized cumulants of $\Delta_t X$ are

$$\overline{\kappa}_2(\Delta_t X) = \frac{1 - e^{-\lambda t}}{\lambda \rho}$$

and

$$\overline{\kappa}_4(\Delta_t X) = \frac{3e^{-(3\lambda+\rho)t}\left(e^{\lambda t} - 2\right)\left(4e^{2\lambda t}\lambda - 2e^{\rho t}\rho - 2(2\lambda-\rho)e^{(\lambda+\rho)t} + e^{(2\lambda+\rho)t}(2\lambda-\rho)\right)}{4\lambda\rho\left(4\lambda^2 - \rho^2\right)}$$

$$+\frac{6\cdot\Gamma\left(\frac{2\rho}{2\lambda-\rho}\right)}{(2\lambda+\rho)^{3}\,\Gamma\left(\frac{6\lambda-\rho}{2\lambda-\rho}\right)}.$$

It is easy to check that

$$\overline{\kappa}_4\left(\Delta_t X\right)\underset{t\to\infty}{\longrightarrow}\frac{3}{4\lambda\rho\left(2\lambda+\rho\right)}+\frac{6\cdot\Gamma\left(\frac{2\rho}{2\lambda-\rho}\right)}{(2\lambda+\rho)^{3}\,\Gamma\left(\frac{6\lambda-\rho}{2\lambda-\rho}\right)}.$$

The kurtosis $\overline{\kappa}_4\left(\Delta_t X\right)/\overline{\kappa}_2^2\left(\Delta_t X\right)$ is decreasing as a function of t, and decreasing as a function of λ and ρ. This suggests that the non-normality of $\Delta_\infty X$ escalates as the value of λ and ρ increase.

Even in this simple case, despite it was possible to find closed expressions for the cumulants of $\Delta_t X$, it is not feasible to determine when the distribution of the increments of X belongs to a known class. This exemplifies the complex dynamics exhibited by the increments of a BSS process (4.1).

Example 2: $g\left(x\right)=x^{\alpha}e^{-\lambda x}1_{\mathbb{R}_+}\left(x\right)$, $q\left(x\right)=x^{\beta}e^{-\mu x}1_{\mathbb{R}_+}\left(x\right)$, Ornstein-Uhlenbeck process ε^2

For $\min\left\{\alpha,\beta/2\right\}>-1/2$ and $\lambda,\rho,\mu>0$ with $2\lambda\neq\rho$ and $\mu\neq\rho$, consider the model (4.1), (4.2) with $g\left(x\right)=x^{\alpha}e^{-\lambda x}1_{\mathbb{R}_+}\left(x\right)$, $q\left(x\right)=x^{\beta}e^{-\mu x}1_{\mathbb{R}_+}\left(x\right)$, $h\left(x\right)=e^{-\rho x}1_{\mathbb{R}_+}\left(x\right)$ and L a subordinator. BSS processes of this type have interesting mathematical properties and they are relevant as models for the temporal turbulent velocity field. Among the most remarkable mathematical properties for this class of BSS processes, we find multipower variation type limits (see, e.g., [16]). The parameter α controls the smoothness of X and determines when X is a semimartingale: X is a semimartingale if and only if $\alpha>1/2$.

From the modelling perspective, it has been shown in [34] that the BSS process (4.1), under the assumptions of the present example with $\alpha=-1/6$, reproduces the scaling behaviour of second order turbulent structure functions

$$E\left\{\left(X_t-X_0\right)^2\right\}\propto t^{2/3}, \tag{4.17}$$

for a certain range of scales t. Besides, the increments display non-vanishing odd cumulants. This is specially relevant in the context of turbulence.

For $t>0$, we have

$$\psi_t\left(s\right)=e^{\mu s}\left(\left(t-s\right)^{\beta}e^{-\mu t}-\left(-s\right)^{\beta}\right)1_{\mathbb{R}_-}\left(s\right)+e^{-\mu\left(t-s\right)}\left(t-s\right)^{\beta}1_{\left(0,t\right)}\left(s\right)$$

which yields

$$\begin{aligned}C_1\left(t,r\right) &= \frac{e^{-\rho(t-r)}}{\left(\mu-\rho\right)^{1+\beta}}\left(\Gamma\left(1+\beta\right)-\Gamma\left(1+\beta,-\left(t-r\right)\left(\mu-\rho\right)\right)\right)1_{\mathbb{R}_+}\left(r\right)\\ &\quad-\frac{e^{-\rho(t-r)}}{\left(\mu-\rho\right)^{1+\beta}}\left\{\left(e^{\rho t}-1\right)\Gamma\left(1+\beta\right)-\Gamma\left(1+\beta,\left(t-r\right)\left(\mu-\rho\right)\right)\right)\end{aligned}$$

$$-e^{\rho t}\Gamma\left(1+\beta,-r\left(\mu-\rho\right)\right)\Big\}\,1_{\mathbb{R}_-}\left(r\right).$$

The functions ϕ_t and C_2 are given as in (4.5) and (4.9), respectively. The functions C_1 and C_2 satisfy the condition (4.13) in Proposition 7. Therefore, assuming that L' has finite moments of all orders, the cumulants of $\Delta_t X$ are given by the formula (4.14).

To our knowledge, it does not seem to be possible to express $H_{t,z}$ in terms of simple functions. However, (4.14) can be evaluated numerically.

Example 3: generalized Hyperbolic marginals

For $c,\gamma\in\mathbb{R}$, $\bar{\lambda}>0$ and $-1/2<\alpha<0$, consider the model (4.1), (4.2) with

$$g\left(x\right)=c\frac{\bar{\lambda}^{\alpha+1/2}}{\Gamma\left(2\alpha+1\right)^{1/2}}x^{\alpha}e^{-\frac{\bar{\lambda}}{2}x}1_{\mathbb{R}_+}\left(x\right),$$

$$q\left(x\right)=\gamma\frac{\bar{\lambda}^{2\alpha+1}}{\Gamma\left(2\alpha+1\right)}x^{2\alpha}e^{-\bar{\lambda}x}1_{\mathbb{R}_+}\left(x\right),$$

$$h\left(x\right)=\frac{\bar{\lambda}^{-2\alpha-1}}{\Gamma\left(-2\alpha\right)}x^{-2\alpha-1}e^{-\bar{\lambda}x}1_{\mathbb{R}_+}\left(x\right),$$

and L a subordinator. Choosing L' such that the OU process

$$\varepsilon_t^2=\int_{-\infty}^{t}e^{-\bar{\lambda}(t-u)}dL_u$$

has the generalized inverse Gaussian law $\mathrm{GIG}(\lambda,\chi,\theta)$ (see Appendix D), it was reported in [4] that the law of (4.1) is the generalized hyperbolic $\mathrm{GH}(\lambda,\chi,\theta,0,c^2,\gamma)$ (see Appendix D). Notice that the distribution of X does not depend on the parameters $\left(\bar{\lambda},\alpha\right)$. However, the law of $\Delta_t X$ depends of these parameter as can be seen in the expression of C_i, $i=1,2$.

For the present example, for $t>0$,

$$\psi_t\left(s\right)=\gamma\frac{\bar{\lambda}^{2\alpha+1}}{\Gamma\left(2\alpha+1\right)}\left\{e^{\bar{\lambda}s}\left(\left(t-s\right)^{2\alpha}e^{-\bar{\lambda}t}-\left(-s\right)^{2\alpha}\right)1_{\mathbb{R}_-}\left(s\right)+e^{\bar{\lambda}s}\left(t-s\right)^{2\alpha}e^{-\bar{\lambda}t}1_{(0,t)}\left(s\right)\right\}$$

and

$$\phi_t^2\left(s\right)=\frac{c^2\bar{\lambda}^{2\alpha+1}}{\Gamma\left(2\alpha+1\right)}\left\{e^{\bar{\lambda}s}\left(\left(t-s\right)^{\alpha}e^{-\bar{\lambda}t/2}-\left(-s\right)^{\alpha}\right)^2 1_{\mathbb{R}_-}\left(s\right)+e^{\bar{\lambda}s}\left(t-s\right)^{2\alpha}e^{-\bar{\lambda}t}1_{(0,t)}\left(s\right)\right\}.$$

Consequently,

$$C_1(t,r) = \gamma e^{-\bar{\lambda}(t-r)} 1_{(0,t)}(r) + \gamma e^{\bar{\lambda} r} \left\{ -1 + \frac{(-1)^{2a+1} \sin(2a\pi) e^{-\bar{\lambda} t}}{\pi} \text{Beta}\left(\frac{r}{t}, -2a, 0\right) \right.$$
$$\left. - \frac{\sin(2a\pi) t^{2a+1} e^{-\bar{\lambda} t} (-r)^{-2a-1}}{\pi (2a+1)} {}_2F_1\left(1, 2a+1, 2a+2, \frac{t}{r}\right) \right\} 1_{\mathbb{R}_-}(r)$$

and

$$C_2(t,r) = c^2 e^{-\bar{\lambda}(t-r)} 1_{(0,t)}(r) + c^2 e^{\bar{\lambda} r} \left\{ 1 + \frac{(-1)^{2a+1} \sin(2a\pi) e^{-\bar{\lambda} t}}{\pi} \text{Beta}\left(\frac{r}{t}, -2a, 0\right) \right.$$
$$- \frac{2\Gamma(a+1) t^a e^{-\bar{\lambda} t/2}}{\Gamma(2a+1)\Gamma(1-a)} (-r)^{-a} {}_2F_1\left(-a, a+1, 1-a, \frac{r}{t}\right)$$
$$\left. - \frac{\sin(2a\pi) t^{2a+1} e^{-\bar{\lambda} t} (-r)^{-2a-1}}{\pi (2a+1)} {}_2F_1\left(1, 2a+1, 2a+2, \frac{t}{r}\right) \right\} 1_{\mathbb{R}_-}(r),$$

where $\text{Beta}(\cdot,\cdot,\cdot) : (-\infty, 0) \times (-\infty, 0) \times (-\infty, 0] \to \mathbb{C}$ denotes to the incomplete Beta function

$$\text{Beta}(x, a, b) = -\int_x^0 w^{a-1} (1-w)^{b-1} dw.$$

The functions C_1 and C_2 do not satisfy the condition (4.13) in Proposition 7. However, they fulfill the condition in Remark 8. Thus, when $E\{(L')^n\} < \infty$ for all $n \in \mathbb{N}$, the cumulants of $\Delta_t X$ are given by (4.14).

4.3 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as an exponential ambit process: theory

Consider the BSS process (4.1) with

$$\varepsilon_t^2 = \exp\{\Lambda(A + (0,t))\} \quad t \in \mathbb{R}, \tag{4.18}$$

where Λ is a homogeneous Lévy basis on $\mathbb{R}^2$ and $A \in \mathcal{B}_b(\mathbb{R}^2)$. In this Section, we deduce a formula for the n-th moments of ε^2 and the increments of X. These formulae can be used to compute the cumulants of the increments of X in terms of the cumulants of Lévy seed Λ'.

The formula for the cumulants

Throughout this Section, Leb will denote the Lebesgue measure on $\mathbb{R}^2$ and $K[z] \equiv K\{z \ddagger \Lambda'\}$.

For $(s_1, \ldots, s_n) \in \mathbb{R}^n$,

$$E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\} = \exp\left\{\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr\right\}. \tag{4.19}$$

For a general ambit set A, the evaluation of (4.19) might be difficult. The main obstacle is to split $\cup_{i=1}^{n} A_{s_i}$ into the sets $\left\{x : \sum_{i=1}^{n} 1_{A_{s_i}}(x) = j\right\}$, $j = 1, \ldots, n$. For the type of ambit sets we are interested in, it is easier to compute the intersections of $\left(A_{s_i}\right)_{i=1}^{n}$ than the previous partition.

Lemma 11 *For* $(s_1, \ldots, s_n) \in \mathbb{R}^n$,

$$\begin{aligned}\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr =& \sum_{m=1}^{n} \sum_{l=m}^{n} \sum_{1 \leq i_1 < \ldots < i_l \leq n} (-1)^{l-m} K[m] \binom{l}{m} \mathrm{Leb}\left(A_{s_{i_1}} \cap \cdots \cap A_{s_{i_l}}\right) \\ =& \sum_{l=1}^{n} \sum_{1 \leq i_1 < \ldots < i_l \leq n} \sum_{m=1}^{l} (-1)^{l-m} K[m] \binom{l}{m} \mathrm{Leb}\left(A_{s_{i_1}} \cap \cdots \cap A_{s_{i_l}}\right).\end{aligned} \tag{4.20}$$

Proof. See Paper B. ■

Formula (4.20) is easily evaluated numerically for n not too large. Furthermore, it permits an explicit computation of the cumulants of the increments of X when Λ' has a normal distribution and the ambit set A has a specific form (see Section 4.4).

Lemma 12 *Let* $n \in \mathbb{N}$. *Under the convention* $\prod_{j=1}^{0} \phi_t^2(s_j) = \prod_{j=1}^{0} \psi_t(s_j) = 1$, *we have that*

$$\begin{aligned}E\left\{(\Delta_t X)^n\right\} =& \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} (2i-1)!! \\ &\times \int_{-\infty}^{t} ds_1 \cdots \int_{-\infty}^{t} ds_i \int_{-\infty}^{t} dr_1 \cdots \int_{-\infty}^{t} dr_{n-2i} \prod_{j=1}^{i} \phi_t^2(s_j) \prod_{l=1}^{n-2i} \psi_t(s_l) E\left\{\varepsilon_{s_1}^2 \cdots \varepsilon_{s_n}^2 \varepsilon_{r_1}^2 \cdots \varepsilon_{r_{n-2i}}^2\right\},\end{aligned}$$

where $n!!$ *represents the double factorial of* $n \in \mathbb{N} \cup \{0, -1\}$, *and* ϕ *and* ψ *are the functions defined in (4.5).*

Proof. See Paper B. ■

In the turbulence modelling context, the n-th moments of increments of the temporal turbulent velocity field are relevant as they constitute a basic element in the Kolmogorov-Obukhov theory. In particular, for X to be a relevant model for the temporal turbulent velocity field, it should reproduce the 2/3rd-Kolmogorov law (4.17) for a certain range of scales t. Lemma 12 implies that, to satisfy the 2/3rd-Kolmogorov law, $\phi_t^2(s) \propto t^{-1/6} b(t)$ and $\psi_t(s) \propto t^{-1/3} b(t)$ when t is in a neighborhood of 0, where $b(t)$ is a bounded function in such a neighborhood with $b(0) \neq 0$. In particular, X with $q = 0$ and $g(x) = x^{-1/6} e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, for $\lambda > 0$, satisfies (4.17). This example of the BSS process (4.1) was used in [34] to model turbulent velocity time series. Lemma 12 can also be used to determine the behavior of $\phi_t^2(s)$ and $\psi_t(s)$ around 0 to satisfy other scaling laws, in addition to the 2/3rd-Kolmogorov law.

Lemmata 11 and 12 provide a way to compute the moments of the increments $\Delta_t X$. Furthermore, it is possible to compute the r-th cumulants $\kappa_r(\Delta_t X)$ by the recursive formula (see [33])

$$\kappa_r(\Delta_t X) = \mu_r - \sum_{j=1}^{r-1} \binom{r-1}{j} E\left\{(\Delta_t X)^j\right\} \kappa_{r-j}(\Delta_t X) \quad r \geq 2. \tag{4.21}$$

The generality in the shape of the ambit set, the distribution of Λ' and the form of the kernels g and q make it difficult to get closed expression for the n-th moments and the cumulants of $\Delta_t X$. However, the formulae in Lemmata 11 and 12 provide a simple way to evaluate the moments numerically. This might help to analyze the distribution of the increments of $\Delta_t X$.

4.4 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as an exponential ambit process: examples

In this Section we study the cumulants of $\Delta_t X$ for some examples assuming that ε^2 is the exponential of an ambit process. In these examples the cumulants of Λ' appear implicitly in the function $K[z] \equiv K\{z \ddagger \Lambda'\}$.

4.4.1 Normal Lévy basis example

We assume that $q \equiv 0$ and $\Lambda' \sim$Normal(μ, δ) (i.e. $K[z] = \mu z + \frac{1}{2}\delta^2 z^2$), and we analyze the cumulants of the increments of the model (4.1), (4.18) for a triangular ambit set.

The normality of Λ' allows to simplify the formula (4.20), and we get

$$\sum_{m=1}^{l} \binom{l}{m} (-1)^{l-m} K[m] = \begin{cases} \mu + \frac{\delta^2}{2} & \text{if } l = 1 \\ \delta^2 & \text{if } l = 2 \\ 0 & \text{otherwise} \end{cases}.$$

This provides a way to compute the n-moments $E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\}$ that we will use to calculate cumulants and n-moments of $\Delta_t X$.

Lemma 13 *Let $n \in \mathbb{N}$. Then,*

$$\log E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\} = n\left(\mu + \frac{\delta^2}{2}\right) \text{Leb}(A) + \delta^2 \sum_{1 \leq i_1 < i_2 \leq n} \text{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}}.$$

Let $a, T > 0$. Assume that $q \equiv 0$, $K[z] = \mu z + \frac{1}{2}\delta^2 z^2$ and that A is the ambit set given by

$$A = \left\{(x,t) \in \mathbb{R}^2 : 0 \leq t \leq T, |x| \leq \frac{a}{T}(T-t)\right\}. \tag{4.22}$$

For $s_1 < s_2 < s_1 + T$,

$$\text{Leb}\left(A_{s_1} \cap A_{s_2}\right) = \frac{a}{T} \cdot \left(T - |s_1 - s_2|\right)^2.$$

Since Leb $\left(A_{s_1} \cap A_{s_2}\right)$ has this simple expression, the present example also gives a simple expression for (4.19). Namely,

$$\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr = n\left(\mu + \frac{\delta^2}{2}\right) \text{Leb}(A) + \frac{a\delta^2}{T} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(T - |s_1 - s_2|\right)^2 1_{\{n>1\}}$$

which implies that

$$\exp C\{z \ddagger \Delta_t X\} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} \left\|\phi_t\right\|_2^{2n} + \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} a_n,$$

where $\|\cdot\|_2 = \|\cdot\|_{L^2(\mathbb{R})}$ and

$$a_n = -\left\|\phi_t\right\|_2^{2n} + \int_{\mathbb{R}} ds_1 \cdots \int_{\mathbb{R}} ds_n \phi_t^2(s_1) \cdots \phi_t^2(s_n) \exp\left\{\delta^2 \sum_{1 \le i_1 < i_2 \le n} \text{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}}\right\}.$$

In Paper B it is shown that, for $z \approx 0$, $a \cdot T \cdot \delta^2 \approx 0$,

$$C\{z \ddagger \Delta_t X\} \approx -\frac{z^2}{2} e^{Ta\left(\mu - \frac{\delta^2}{2}\right)} \left\|\phi_t\right\|_2^2.$$

Thus $\Delta_t X$ behaves similar to a normal distribution. (The distribution of $\Delta_t X$ is not exactly normal but gets more and more normal as $a \cdot T \cdot \delta^2 \downarrow 0$.)

4.4.2 Gamma Lévy basis example

Let $0 > \alpha > -1/2$ and $\beta, \gamma, \lambda > 0$. Consider the model (4.1), (4.18) with $q \equiv 0$, $g(x) = x^{\alpha} e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, A given as in (3.2) and $K[z] = \log(1 - z/\beta)^{-\gamma}$, $z < \beta$ (i.e. Λ' has a Gamma(γ, β) law). In general, it is of interest to determine specific distributional properties of ΔX_t. Of particular interest is the question of infinite divisibility of $\Delta_t X$. The present example provides a case where the distribution of $\Delta_t X$ is not infinitely divisible.

Figure 4.4.1 shows $\kappa_4(\Delta_t X) / \kappa_2^2(\Delta_t X)$ for $(\gamma, \beta) = (1, 5)$, $(\theta, L, T) = (1, 10, 1)$, $\alpha = -1/3, -1/6$ and $\lambda = 1, 2, 3$. It is well-know that, when the distribution of X is infinitely divisible, the cumulants $\kappa_n(X)$, for $n \geq 3$, are the moments of the Lévy measure of X. This implies that, when the distribution of X is infinitely divisible, $\kappa_4(X) \geq 0$. Figure 4.4.1 shows that $\kappa_4(\Delta_t X) < 0$ for $\lambda = 1, 2, 3$. Therefore, the law of $\Delta_t X$ cannot be infinitely divisible.

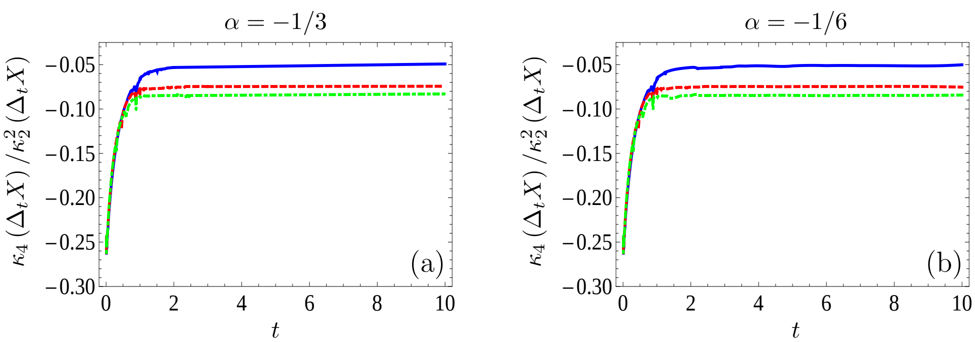

Figure 4.4.1: Standardized cumulant $\kappa_4(\Delta_t X)/\kappa_2^2(\Delta_t X)$ for Example 4.4.2 with $(\theta, L, T) = (1, 10, 1)$, $\Lambda' \sim$Gamma$(1, 5)$, and different values of α and λ. (a) Parameters $\alpha = -1/3$ and $\lambda = 1, 2, 3$. (b) Parameters $\alpha = -1/6$ and $\lambda = 1, 2, 3$. We use blue for $\lambda = 1$, red for $\lambda = 2$ and green for $\lambda = 3$.

5

An asymptotic problem for two classes of smooth processes

THIS CHAPTER summarizes the main ideas and results from Paper C.

Let $(X_t)_{t\geq 0}$ be a stochastic process. When X is a semimartingale, it is well-known that the limit

$$\lim_{n\to\infty} \sum_{i=1}^{\lfloor nt \rfloor} \left(X_{i/n} - X_{(i-1)/n}\right)^2 \tag{5.1}$$

exists in probability. In particular, when X is a bounded variation (BV) process the limit (5.1) is 0. It is natural to ask how we can rescale (5.1) to recover a non-trivial limit for the case where X is a BV process. We address this problem for two classes of smooth processes: the integrated Brownian motion (IBM) and the smooth Brownian semi-stationary (BSS) process with a gamma kernel. A smooth process is a stochastic process that is differentiable.

The integrated Brownian motion is a stochastic process of the form

$$J_t^n = \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} B_{s_n} ds_n ds_{n-1} \cdots ds_2 ds_1 \quad n \in \mathbb{N},$$

where $(B_t)_{t\in\mathbb{R}}$ is a standard Brownian motion. The index n indicates the number of iterated integrals and, therefore, the number of derivatives J has. We will denote the class of these smooth processes by $\mathcal{IBM}$.

Concerning BSS processes, we are interested in the asymptotics of (2.2) with $q \equiv \mu = 0$, $\sigma \equiv 1$ and g given by the gamma kernel

$$g(x) = x^{\alpha} \exp(-\lambda x) 1_{(0,\infty)}(x), \tag{5.2}$$

with $\alpha > 1/2$, $\lambda > 0$. We will refer to this subclass of smooth BSS processes as $\mathcal{SGKBSS}$.

5.1 Statement of the problem

Consider a stochastic process $(X_t)_{t\geq 0}$ such that $X \in \mathcal{IBM} \cup \mathcal{SGKBSS}$. The process X has differentiable paths. From the Mean Value Theorem, it follows that the normalized realized quadratic variation (NRQV) $[X_n]$ of X, defined as

$$[X_n]_t := n \sum_{i=1}^{\lfloor nt \rfloor} \left(X_{i/n} - X_{(i-1)/n} \right)^2, \tag{5.3}$$

satisfies

$$[X_n]_t \underset{n\to\infty}{\overset{a.s.}{\to}} \int_0^t (X_r')^2 \, dr \tag{5.4}$$

for $t \geq 0$. Motivated by this limit, we are interested in the asymptotic behavior of

$$n^{\beta} \left([X_n]_t - \int_0^t (X_r')^2 \, dr \right) \tag{5.5}$$

when $n \to \infty$, for some suitable $\beta > 0$. There are two immediate problems related to (5.5): 1) What is the correct value of $\beta > 0$, if any, to have a non-trivial limit? 2) What is the limit distribution given the correct β? This last question can be naturally extended to stronger concepts of convergence. For simplicity we only consider the case $t = 1$.

Paper C contains some approaches that partially answer these questions.

5.2 The $\mathcal{SGKBSS}$ class

Consider a stochastic process $(X_t)_{0\leq t\leq 1}$ whose paths are in $C^3[0,1]$. In this case, it is possible to find a β where (5.5) has an almost sure non-trivial limit. The proof of this result is an application of Taylor's Theorem and Theorem 5 in [15].

Proposition 14 *Let $(X_t)_{0\leq t\leq 1}$ be a stochastic process whose paths are almost surely in $C^3[0,1]$. Then,*

$$n^2 \left([X_n]_1 - \int_0^1 (X_r')^2 \, dr \right) \underset{n\to\infty}{\overset{a.s.}{\to}} -\frac{1}{12} \int_0^1 (X_r'')^2 \, dr = -\frac{1}{12} \|X''\|^2_{L^2(0,1)}.$$

Proof. See Paper C. ■

If $X \in \mathcal{SGKBSS}$ with $\alpha > 5/2$, the paths of X are almost surely in $C^3([0,\infty))$. The distribution of $Y = \frac{1}{12}\|X''\|^2_{L^2(0,1)}$ depends on the value of α.

Remark 15 *Since the index n represents the number of derivatives for the IBM J^n, $J^3 \in C^3([0,\infty))$. Consequently, Proposition 14 also applies to J^n, for $n > 3$.*

It still remains open to determine the rate of convergence and the limit distribution of (5.5) for the stochastic processes in $\mathcal{SGKBSS}$ with index $1/2 < \alpha < 5/2$, that is, for the BSS processes that are not in $C^3[0,1]$.

5.3 Integrated Brownian motion

In this Section, we summarize some investigations about the limit distribution of (5.5) for J^1_t.

5.3.1 Convergence of the variance

We have that

$$[J^m_n]_t \underset{n\to\infty}{\overset{a.s.}{\to}} \int_0^t \left(\frac{d}{dr} J^m_r \bigg|_{r=s} \right)^2 ds,$$

where $[J^m_n]$ denotes the NRQV of J^m.

Restricting to the case $m = 1$, we get

$$[J^1_n]_t \underset{n\to\infty}{\overset{a.s.}{\to}} \int_0^t B^2_s ds$$

for which we first analyze the variance

$$A_{n,t} := \mathrm{Var}\left([J^1_n]_t - \int_0^t B^2_s ds \right)$$

for $t \geq 0$. Isserlis' Theorem (see [21]) implies the next result.

Proposition 16 *For $t \geq 0$ and $n \in \mathbb{N}$,*

$$A_{n,t} = \frac{45\lfloor nt \rfloor^4 - 60nt\lfloor nt \rfloor^3 - 15\lfloor nt \rfloor^2 + 15nt\lfloor nt \rfloor + \lfloor nt \rfloor + 15n^4t^4}{45n^4}.$$

Proof. See Paper C. ■

An immediate consequence of Proposition 16 is the next result.

Proposition 17 *For $t \geq 0$,*

$$[J^1_n]_t \overset{L^2}{\to} \int_0^t B^2_s ds.$$

5.3.2 The limit distribution

We are interested in shedding light on the limit distribution of

$$\mathfrak{A}_{n,t} := A_{n,t}^{-1/2}\left([J_n^1]_t - \int_0^t B_s^2 ds\right).$$

To clarify the main aspects, we only consider the case $t = 1$.

It turns out (see Subsection 5.3.4) that the limit law of $\mathfrak{A}_{n,1}$ seems to be a Rosenblatt distribution [27, 39]. Using Maejima and Tudor's parametrization of the Rosenblatt distribution [27], we propose the next conjecture.

Conjecture 18 *Let $R(h)$ denote a Rosenblatt random variable with index $h \in (1/2, 1)$. Then, we have*

$$\mathfrak{A}_{n,1} \xrightarrow{d} R(h),$$

whith $h \approx 0.9$.

Numerical simulations of $\mathfrak{A}_{n,1}$ and $R(0.9)$ show strong evidence supporting the previous conjecture (see Subsection 5.3.4).

5.3.3 An expression for $\mathfrak{A}_{n,1}$

In relation to Conjecture 18, it is of interest to express

$$\mathfrak{X}_n := [J_n^1]_t - \int_0^t B_s^2 ds = n\sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} B_s \mathrm{d}s\right)^2 - \int_0^1 B_s^2 \mathrm{d}s,$$

in terms of a double Wiener integral with respect to $B.$.

Proposition 19 *We can rewrite the random variable $\mathfrak{X}_n$ as*

$$\mathfrak{X}_n = \int_0^1\int_0^1 F_n(r,s)\mathrm{d}B_r\mathrm{d}B_s - \frac{5}{6n}, \tag{5.6}$$

where

$$F_n(r,s) = n\sum_{i=1}^n f_i^{(n)}(r) f_i^{(n)}(s) - \left(1 - \max(r,s)\right)$$

and

$$f_i^{(n)}(s) = \begin{cases} \frac{1}{n}, & s \in \left[0, \frac{i-1}{n}\right), \\ \frac{i}{n} - s, & s \in \left[\frac{i-1}{n}, \frac{i}{n}\right), \\ 0, & s \in \left[\frac{i}{n}, 1\right]. \end{cases}$$

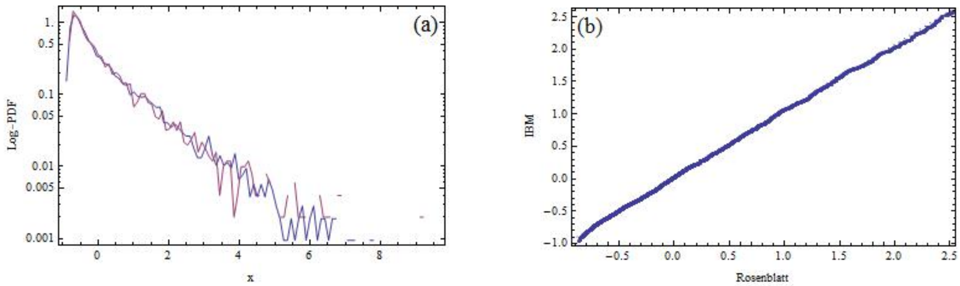

Figure 5.3.1: (a) Histograms in log-linear scale of the Rosenblatt sample (blue) and the $\mathfrak{A}_{100000,1}$ sample (purple). (b) QQ plot of the $\mathfrak{A}_{100000,1}$ sample and the Rosenblatt sample.

Proof. See Paper C. ■

The double integral expression (5.6) might provide evidence supporting Conjecture 18 since the Rosenblatt distribution can be expressed as a second order Wiener chaos (see, e.g. [38]). One possible way to prove Conjecture 18 would be to show that $F_n(r, s)$ converges in L_2 to the kernel that appears in Proposition 1 of [38]. However, this has not yet been done and is not part of this thesis.

5.3.4 Numerical results

Figure 5.3.1 illustrates the conjecture for $\mathfrak{A}_{n,1}$. These figures were obtained by simulating 10000 samples of *Rosenblatt* (0.9) and 10000 samples of $\mathfrak{A}_{n,1}$ for $n = 10^5$. Figure (a) shows two histograms in log-linear scale: the histogram of the Rosenblatt sample (blue) and the histogram of the $\mathfrak{A}_{100000,1}$ sample (purple). They are very similar. Figure (b) corresponds to the QQ plot of the $\mathfrak{A}_{100000,1}$ sample and the Rosenblatt sample.

6
Conclusion

The analysis performed in this thesis clearly demonstrates that Brownian semi-stationary processes are well adapted to reproduce key characteristics of turbulent time series. The parameters of the model are solely estimated from the marginal distribution and the correlator $c_{1,1}$ of the energy dissipation and from second and third order structure functions of velocity increments. This has been done under the specific model specification (3.4) with a shifted 2-gamma kernel g. The use of a shifted 2-gamma kernel is motivated by its ability to accurately reproduce the empirical sdf.

The data set analyzed here has a relatively high Reynolds number, with a visible inertial range. Investigations carried out for other data sets have shown that the BSS processes have the same potential for modelling turbulent velocity time series even for smaller Reynolds numbers.

The present work also provides a way to compute the cumulants of the increments of BSS processes for two specific classes of volatility processes. It is not possible to find closed expression for all the cases and examples presented here. However, the formulae are simple enough to be evaluated numerically.

Of particular interest is the discussion in Subsection 4.4.2 since it provides an example where our analysis of the cumulants shows that the distribution of the increments of a specific BSS process is not infinitely divisible. It remains open to determine conditions on the BSS processes such that their increments have an infinitely divisible law.

Our main purpose for studying the cumulants of increments of BSS processes was to establish a way that sheds some light on the distributions of increments via cumulants. This is a first step to understand why the BSS approach is able to model a great variety of stylized features in turbulence. The results discussed here allow to directly compare the models with data without time consuming simulations of the underlying processes.

We have not been able to solve the asymptotic problem for the smooth processes. The techniques used here do not provide the full answer since they do not permit to determine the limit distributions and the precise rate of convergence. For integrated Brownian motion, we have provided strong numerical evidence supporting a Rosenblatt limit for the asymptotic problem in the J^1 case. It remains open to determine the veracity of such a limit.

References

[1] O.E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. Lond. A*, **353**(1674):401–419, 1977.

[2] O.E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.*, **24**:1–13, 1997.

[3] O.E. Barndorff-Nielsen and A. Basse. Quasi Ornstein-Uhlenbeck processes. *Bernoulli*, **17**(3):916–941, 2011.

[4] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Modelling energy spot prices by volatility modulated Lévy-driven volterra processes. *Bernoulli*, **19**(3):803–845, 2013.

[5] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Recent advances in ambit stochastics with a view towards tempo-spatial stochastic volatility/intermittency. *Banach Center Publ.*, **104**:25–60, 2015.

[6] O.E. Barndorff-Nielsen, P. Blæsild, and J. Schmiegel. A parsimonious and universal description of turbulent velocity increments. *Eur. Phys. J. B*, **41**:345–363, 2004.

[7] O.E. Barndorff-Nielsen, J.M. Corcuera, and M. Podolskij. Multipower variation for Brownian semi-stationary processes. *Bernoulli*, **17**(4):1159–1194, 2011.

[8] O.E. Barndorff-Nielsen, T. Mikosch, and S.I. Resnick, editors. *Lévy processes: theory and applications*. Springer Science & Business Media, New York, 2012.

[9] O.E. Barndorff-Nielsen and J. Schmiegel. Tempo-spatial modelling; with applications to turbulence. *Uspekhi Mat. Nauk*, **159**:65–91, 2003.

[10] O.E. Barndorff-Nielsen and J. Schmiegel. A stochastic differential equation framework for the time-wise dynamics of turbulent velocities. *Theory Probab. Appl.*, **52**:372–388, 2008.

[11] O.E. Barndorff-Nielsen and J. Schmiegel. Brownian semistationary processes and volatility/intermittency. In H. Albrecher, W. Runggaldier, and W. Schachermayer, editors, *Advanced Financial Modelling*, pages 1–26. Walter de Gruyter, Berlin, 2009.

[12] A. Basse, S.E. Graversen, and J. Pedersen. A unified approach to stochastic integration on the real line. *Theory Probab. Appl.*, **58**(2):193–215, 2015.

[13] B. Birnir. The Kolmogorov-Obukhov statistical theory of turbulence. *J. Nonlinear Sci.*, **23**(4):657–688, 2013.

[14] W. Breymann and D. Lüthi. *ghyp: A package on generalized hyperbolic distributions. Manual for the R package ghyp.* CRAN project, 2010. https://cran.r-project.org/web/packages/ghyp/ghyp.pdf.

[15] C.K. Chui. Concerning rates of convergence of Riemann sums. *J. Approx. Theory*, **4**(3):279–287, 1971.

[16] J.M. Corcuera, E.H.L. Sørensen, M.S. Pakkanen, and M. Podolskij. Asymptotic theory for Brownian semi-stationary processes with application to turbulence. *Stochastic Process. Appl.*, **123**(7):2552–2574, 2013.

[17] E. Eberlein and E.A.V. Hammerstein. Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. In R.C. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications IV*, pages 221–264. Birkhäuser, Basel, 2004.

[18] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, **1**(3):281–299, 1995.

[19] U. Frisch. *Turbulence. The legacy of A.N. Kolmogorov*. Cambridge University Press, Cambridge, 1995.

[20] I. Hosokawa, C.W. Van Atta, and S.T. Thoroddsen. Experimental study of the Kolmogorov refined similarity variable. *Fluid Dyn. Res.*, **13**:329–333, 1994.

[21] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, **12**:134–139, 1918.

[22] B. Jørgensen. *Statistical Properties of the Generalized Inverse Gaussian Distribution*. Lecture Notes in Statistics **9**. Springer-Verlag, Heidelberg, 1982.

[23] A.N. Kolmogorov. Dissipation of energy in the locally isotropic turbulence. *Dokl. Akad. Nauk SSSR*, **32**:16–18, 1941.

[24] A.N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. *Dokl. Akad. Nauk SSSR*, **30**:16–18, 1941.

[25] A.N. Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, **13**:82–85, 1962.

[26] L.D. Landau and E.M. Lifshitz. *Fluid Mechanics,* Volume **6** of *Course of Theoretical Physics*. Pergamon Press, Oxford, 1959.

[27] M. Maejima and C.A. Tudor. On the distribution of the Rosenblatt process. *Stat. Probabil. Lett.*, **83**(6):1490–1495, 2013.

[28] L. Onsager. The distribution of energy in turbulence. *Phys. Rev.*, **68**:285, 1945.

[29] B.S. Rajput and J. Rosinski. Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields*, **82**(3):451–487, 1989.

[30] O. Reynolds. On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Phil. Trans. R. Soc. A*, **186**:123–164, 1895.

[31] L.F. Richardson. *Weather Prediction by Numerical Process*. Cambridge University Press, Cambridge, 1922.

[32] J. Schmiegel. Self-scaling of turbulent energy dissipation correlators. *Phys. Lett. A*, **337**(3):342–353, 2005.

[33] P.J. Smith. A recursive formulation of the old problem of obtaining moments from cumulants and vice-versa. *Amer. Statist.*, **49**(2):217–218, 1995.

[34] E.H.L. Sørensen. *Stochastic modelling of turbulence: With applications to wind energy.* PhD thesis, Aarhus University, 2012.

[35] E.H.L. Sørensen and J. Schmiegel. A causal continuous-time stochastic model for the turbulent energy cascade in a helium jet flow. *J. Turbul.*, **14**(11):1–26, 2013.

[36] G. Stolovitzky, P. Kailasnath, and K.R. Sreenivasan. Kolmogorov's refined similarity hypothesis. *Phys. Rev. Lett.*, **69**(2):1178–1181, 1992.

[37] G.I. Taylor. The spectrum of turbulence. *Proc. R. Soc. Lond. A*, **164**:476–490, 1938.

[38] C.A. Tudor. Analysis of the Rosenblatt process. *ESAIM Probab. Stat.*, **12**:230–257, 2008.

[39] M.S. Veillette and M.S. Taqqu. Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli*, **19**(3):721–1085, 2008.

[40] A.E.D. Veraart and L.A.M. Veraart. Modelling electricity day-ahead prices by multivariate Lévy semistationary processes. In F.E. Benth, V. Kholodnyi, and P. Laurence, editors, *Quantitative Energy Finance*, pages 157–188. Springer, Berlin, 2014.

A

Paper A: Modelling turbulent time series by BSS-processes

José Ulises Márquez* and Jürgen Schmiegel†
*Department of Mathematics, Aarhus University.
†Department of Engineering, Aarhus University

Abstract

Brownian semi-stationary processes have been proposed as a class of stochastic models for time series of the turbulent velocity field. We show, by detailed comparison, that these processes are able to reproduce the main characteristics of turbulent data. Furthermore, we present an algorithm that allows to estimate the model parameters from second and third order statistics. As an application we synthesise a turbulent time series measured in a helium jet flow.

A.1 Introduction

Stochastic modelling of the turbulent velocity field, understood as an explicit stochastic approach (in contrast to an implicit set up in terms of governing equations and/or in terms of related quantities like velocity increments or velocity derivatives) is generally speaking not well developed in the literature. Most of the existing literature on stochastic turbulence modelling deals with models for derived quantities like velocity increments, energy dissipation and accelerations.

Early attempts to model the rapid variation of the turbulent velocity field include [4, 10, 11, 18, 23, 36] (among many others). Such phenomenological approaches are sometimes called "synthetic turbulence" and can be divided into two classes. The first direction starts from modelling the velocity field and derives the model for the energy dissipation by taking squared small scale increments. The second line of investigation focuses on modelling the energy dissipation field and derives the velocity field by various, partly ad hoc, manipulations. The approach presented here conceives the energy dissipation as the fundamental field, entering directly the model for the velocity field and obeying the physical interpretation of the energy dissipation as the squared small scale fluctuations.

In [36], an iterative, geometric multi-affine model for the one-dimensional velocity process is constructed and some of the basic, global statistical quantities of the energy dissipation field are derived. However, this discrete, dyadic approach does not allow to give explicit expressions for more specific statistical quantities.

Another dyadic, iterative approach for the construction of the velocity field is discussed in [10]. Their model is based on a wavelet decomposition of the velocity field combined with a multiplicative cascading structure for the wavelet coefficients. As discussed in [18], such wavelet approaches are superior over discrete geometric approaches as they allow to model stationarity in a mathematical more rigorous way. The approach discussed here does not suffer from problems related to mathematical rigour and no iterative limit arguments are needed for the construction. A related and interesting wavelet-based approach is discussed in [11], which allows for a sequential construction of the field. A further wavelet-based approach [4] builds on random functions and their orthogonal wavelet transform. The authors show that to each such random function there is an associated cascade on a dyadic tree of wavelet coefficients.

The models [10, 36] fail to incorporate skewness for the velocity increments [23], a basic property of turbulent fields. As an alternative approach, [23] proposes a combination of a multiplicative cascade for the energy dissipation, the use of Kolmogorov's refined similarity hypothesis [24] and an appropriate summation rule for the increments to construct the velocity field. Here, again, only discrete iterative procedures are employed which make analytical statistical statements very difficult.

The stochastic models discussed in the present paper, called Brownian semi-stationary processes, have been proposed to be potentially suitable for turbulence modelling in [8, 9]. These processes define the turbulent velocity field explicitly and as such allow for analytic calculations and identification of the parameters of the model with physical quantities.

In [8] it has been shown that Brownian semi-stationary processes are able to qualitatively reproduce some aspects of turbulence statistics like the evolution of the densities of velocity increments across scales and the conditional statistics of the so-called Kolmogorov variable. Here we will extend and quantify in detail the comparison of the model with empirical data by including more stylized features of turbulent data. Our goal is to estimate the parameters entering Brownian semi-stationary processes from a given turbulent data set. Based on this estimation, a numerical simulation of the model is then compared in great detail with the turbulent data set at hand, including statistical properties not used for the estimation procedure.

The paper is organized as follows. In Section A.2 we list the main stylized features of turbulent time series we use to validate the model. Brownian semi-stationary processes as models for the turbuent velocity field along with cascades processes as models for the energy dissipation are presented in Section A.3. Section A.4 addresses the estimation procedure for the parameters of the model and briefly outlines the numerics behind the simulations. Finally, Section A.5 concludes and summarises the results.

A.2 Stylized features of turbulent time series

In general, turbulence concerns the dynamics in a fluid flow of the three-dimensional velocity vector $\vec{v}(\vec{r}, t) = (v_x(\vec{r}, t), v_y(\vec{r}, t), v_z(\vec{r}, t))$ as a function of position $\vec{r} = (x, y, z)$ and time t. A derived quantity is the energy dissipation, defined as

$$\varepsilon(\vec{r}, t) \equiv \frac{\nu}{2} \sum_{i,j=x,y,z} \left(\partial_i v_j(\vec{r}, t) + \partial_j v_i(\vec{r}, t) \right)^2 . \tag{A.1}$$

The energy dissipation describes the loss of kinetic energy due to friction forces characterized by the viscosity ν.

A pedagogical valuable illustration of a turbulent flow can be gained from the Kolmogorov cascade [20]. In this representation kinetic energy is injected into the flow at large scales through large scale forcing. Non-linear effects redistribute the kinetic energy towards smaller scales. This cascade of energy stops at small scales where dissipation transforms kinetic energy into heat. It is traditional to call the large scale I of energy input the integral scale and the small scale η of dominant dissipation the dissipation scale or Kolmogorov scale. With increasing Reynolds number the fraction I/η increases, giving space for the so called inertial range $\eta \ll l \ll I$ where turbulent statistics are expected to have some universal character. A more precise definition defines the inertial range as the range of scales $1/k$ where the spectrum $E(k)$ (the Fourier transform of the correlation function of the velocity field) displays a power law $E(k) \propto k^{-5/3}$ [20].

The resolution of all dynamically active scales in experiments is at present not achievable for the full three-dimensional velocity vector. Most experiments measure a time series of one component v (in direction of the mean flow) of the velocity vector at a fixed single location $\vec{r}_0$. Based on this restriction one defines the temporal (or surrogate) energy dissipation for stationary, homogeneous and isotropic flows

$$\varepsilon_t(\vec{r_0}) \equiv \frac{15\nu}{\overline{v}^2} \left(\frac{\mathrm{d}v(\vec{r_0}, t)}{\mathrm{d}t} \right)^2, \tag{A.2}$$

where $\overline{v}$ denotes the mean velocity (in direction of the mean flow).

The temporal energy dissipation (A.2) is expected to approximate basic statistical properties of the true energy dissipation (A.1) for stationary, homogeneous and isotropic flows. For other flow conditions, the temporal energy dissipation still contains important statistical information about the turbulent velocity field.

The transformation of the spatial derivatives in (A.1) to the temporal derivative in (A.2) is performed under the assumption of a stationary, homogeneous and isotropic flow and the assumption of Taylor's Frozen Flow Hypothesis [35] which states that spatial structures of the flow are predominantly swept by the mean velocity $\overline{v}$ without relevant distortion. Under this hypothesis, widely used in analyzing turbulent time series, spatial increments along the direction of the mean flow (in direction x) are expressed in terms of temporal increments

$$v(x, y, z, t + s) - v(x, y, z, t) = v(x - \overline{v}s, y, z, t) - v(x, y, z, t). \tag{A.3}$$

In the present paper, we only deal with homogeneous, isotropic and stationary turbulence. Furthermore, we restrict to temporal statistics at a fixed position in space and refer to the inertial range as the temporal counterpart of the spatial inertial range defined by time scales s where $\eta/\overline{v} \ll s \ll I/\overline{v}$. Time scales $s \lesssim \eta/\overline{v}$ are called dissipation time scales and time scales $s \gtrsim I/\overline{v}$ are called integral time scales.

In what follows, the notion energy dissipation refers to the temporal energy dissipation, unless otherwise stated. We also skip reference to the spatial location $\vec{r_0}$ in (A.2) and write ε_t for $\varepsilon_t(\vec{r_0})$.

The most striking feature of time series of the energy dissipation is the strong variability with localized and clustered outbursts of different size and duration. This strongly fluctuating behaviour, which is far away from what might be expected in a Gaussian framework, is called the *intermittency* of the energy dissipation.

The traditional characterization of the intermittent behaviour of the energy dissipation refers to the coarse grained field amplitude over a time horizon T

$$\overline{\varepsilon}(T, t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} \varepsilon_s \mathrm{d}s. \tag{A.4}$$

It has been shown in numerous publications that the moments of the coarse grained energy dissipation follow a scaling law in the inertial range for large Reynolds numbers

$$\mathrm{E}\left\{\bar{\varepsilon}(T,t)^n\right\} \propto T^{-\xi(n)} \tag{A.5}$$

where the positive multifractal scaling exponents $\xi(n)$ are expected to be universal in the limit of very large Reynolds number (cf. e.g. [20, 31] and references therein). The term multifractality refers to the non-linear dependence of the scaling exponents $\xi(n)$ on the order n. The notion of a Reynolds number refers to the time-wise defined Taylor micro-scale Reynolds number [20]

$$R = \frac{\mathrm{Var}\{v\}}{\nu\sqrt{\mathrm{E}\{\varepsilon_t\}}} \tag{A.6}$$

where Var denotes the variance.

An immediate consequence of the scaling relation (A.5) in second order $n = 2$ is scaling of correlators $c_{p,q}$ of order $(p, q) = (1, 1)$. These correlators are defined as

$$c_{p,q}(s) \equiv \frac{\mathrm{E}\left\{\varepsilon_t^p \varepsilon_{t+s}^q\right\}}{\mathrm{E}\left\{\varepsilon_t^p\right\}\mathrm{E}\left\{\varepsilon_{t+s}^q\right\}}. \tag{A.7}$$

The empirical analysis of $c_{p,q}$ revealed the existence of a range of scales s where

$$c_{p,q}(s) \propto s^{-\tau(p,q)} \tag{A.8}$$

and $\tau(1,1) = \xi(2)$ [2, 14, 15, 16, 21, 26, 27, 28, 32].

Intermittency of the velocity field refers to the fact that fluctuations around the mean velocity occur in clusters and are more violent than expected from Gaussian statistics. Furthermore, the frequency of large fluctuations increases with increasing resolution. In terms of moments of temporal velocity increments

$$u_s(t) \equiv v_{s+t} - v_t, \; s > 0 \tag{A.9}$$

intermittency is usually described by (approximate) multifractal scaling of structure functions (e.g. [1, 25])

$$S_n(s) = \mathrm{E}\{u_s(t)^n\} \propto s^{\tau(n)}. \tag{A.10}$$

Here, v_t is one component of the velocity (usually along the mean flow) at time t and at a fixed position and the time scale s is within the inertial range. When appropriate, we write u_s instead of $u_s\,(t)$ in (A.9) since we are only dealing with stationary time series.

Multifractal scaling of structure functions is assumed to hold in the limit of infinite Reynolds number [20]. However, experiments show that the scaling behaviour (A.10) might be poor, even for large Reynolds numbers [3, 31]. Furthermore, even if the scaling relation (A.10) holds, the inertial range still covers only part of the accessible scales where intermittency is observed.

From a probabilistic point of view, (A.10) expresses a scaling relation for the moments of the probability density function (pdf) of velocity increments. A proper estimation of higher-order moments requires an accurate estimation of the tails of the pdf. Thus it may be advantageous to directly work with the pdf. In terms of the pdf, intermittency refers to the increase of the non-Gaussian behaviour of the pdf of velocity increments with decreasing time scale.

A typical scenario is characterized by an approximate Gaussian shape for the large scales (larger than scales at the inertial range), turning to exponential tails within the inertial range and stretched exponential tails for dissipation scales (below the inertial range). This change of shape across all scales clearly reveals the inadequacy of a characterization of intermittency solely via multifractal scaling of structure functions (which is observed only within the inertial range). In [7, 6] it is shown that Normal inverse Gaussian (NIG) distributions are well adapted to accurately describe the densities of velocity increments at all scales and for a wide range of Reynolds numbers.

In 1962, Kolmogorov [24] published two hypotheses (usually refered to as K62) about a quantity V that combines velocity increments, being a large scale quantity, and the energy dissipation, being a small scale quantity. The first hypothesis states that the pdf of the stochastic variable

$$V_r = \frac{\Delta v_t(r)}{(r\varepsilon_r)^{1/3}} \tag{A.11}$$

depends, for $r \ll L$, only on the local Reynolds number

$$\mathrm{Re}_r = r(r\varepsilon_r)^{1/3}/\nu. \tag{A.12}$$

Here,

$$\Delta v_t(r) = v_t(x + r/2, y, z) - v_t(x - r/2, y, z) \tag{A.13}$$

denotes the increment of the component v of the velocity vector in direction of the mean flow (the x-direction) at scale r and $r\varepsilon_r$ is the integrated energy dissipation (A.1) over a domain of linear size r

$$\varepsilon_r = \frac{1}{r}\int_{x-r/2}^{x+r/2} \varepsilon(\sigma, y, z, t)\mathrm{d}\sigma. \tag{A.14}$$

The second hypothesis states that, for $\mathrm{Re}_r \gg 1$, the pdf of V_r does not depend on Re_r, either, and is therefore universal.

Although, for small r, an additional r dependence of the pdf of V_r has been observed [33], the validity of several aspects of K62 has been verified experimentally and by numerical simulation of turbulence [22, 33, 34, 37]. In particular it has been shown that the conditional densities $p(V_r|r\varepsilon_r)$ become independent of $r\varepsilon_r$ for a certain range of scales r within the inertial range. However, the universality of the distribution of V has not been verified in the literature. In this respect, it is important to note that the experimental verification of the Kolmogorov hypotheses is, with reasonable resolution of scales, restricted to temporal statistics and as such relies on the use of the temporal energy dissipation (A.2) instead of the true energy dissipation (A.1). In the time domain, the Kolmogorov variable V is defined as

$$V_{t,s} = \frac{u_s(t - s/2)}{(\overline{v}s\overline{\varepsilon}(s,t))^{1/3}} \tag{A.15}$$

where $u_s(t)$ denotes the temporal velocity increment (A.9) at time scale s, $\overline{v}$ is the mean velocity and $\overline{\varepsilon}(s,t)$ the coarse grained temporal energy dissipation (A.4).

A.3 Modelling framework

In this Section we present the stochastic framework for modelling turbulent time series. One of the main ingredients of the model is the surrogate energy dissipation which, in our approach, will be modelled as

data				Ambit set			NIG-law			
R	η	I	$\hat{f}$	T	k	θ	α	β	μ	δ
985	0.21	443.9	367500	880	10 000	2.20	2.50	-2.00	2.42	3.06

Table A.3.1: Parameters for the data set analysed in this paper. R denotes the Taylor micro-scale Reynolds number, η is the Kolmogorov scale (in units of the finest resolution), I denotes the integral scale (in units of the finest resolution) and $\hat{f}$ is the sampling frequency. The parameters T (in units of the finest resolution), k and θ characterize the ambit set (A.17). The parameters α, β, μ and δ specify the NIG-law of the Lévy seed of L in (A.16).

a continuous cascade process. We briefly discuss cascade models in Subsection A.3.1. Subsection A.3.2 presents the model for the temporal turbulent velocity field along with its most relevant properties.

A.3.1 The cascade model for the turbulent energy dissipation

Our model for the turbulent velocity field takes as an input the temporal energy dissipation which is modelled as a continuous cascade model [30]. In [30] it is shown that this Lévy based approach is able to reproduce the main stylized features of the energy dissipation observed for a wide range of data sets, including the data we analyze in the present paper.

Specifically, we model the temporal energy dissipation ε as the exponential of an integral with respect to a homogeneous Lévy basis L on $\mathbb{R}^2$,

$$\varepsilon_t = \exp\left(\int_{A(t)} L\left(\mathrm{d}y, \mathrm{d}s\right)\right) = \exp\left(L\left(A\left(t\right)\right)\right) \tag{A.16}$$

where $A\left(t\right) = A + \left(0, t\right)$ for a bounded set $A \subset \mathbb{R}^2$. The set $A\left(t\right)$ is called the *ambit* set. From the homogeneity of L it follows that (A.16) is a stationary stochastic process. For details about Lévy bases and the derivation of some of the properties of (A.16), we refer to [30] and the references therein.

The ambit set A is given as

$$A = \left\{(x, t) : 0 \leq t \leq T, -f(t) \leq x \leq f(t)\right\}, \tag{A.17}$$

where $T > 0$. For $T > 0$, $k > 1$, and $\theta > 0$, the function f is defined as

$$f(t) = \left(\frac{1 - (t/T)^{\theta}}{1 + (k \cdot t/T)^{\theta}}\right)^{1/\theta} \qquad 0 \leq t \leq T. \tag{A.18}$$

This specification of the ambit set is adapted to reproduce the empirically observed scaling of correlators.

In [30] it is shown that the density of the logarithm of the energy dissipation is well described by a normal inverse Gaussian distribution, i.e. $\log \varepsilon_t \sim \mathrm{NIG}\left(\alpha, \beta, \mu, \delta\right)$. For the correlators $c_{p,q}$ (A.7) to exist it is necessary to assume that ε_t has exponential moments of order $p + q$ leading to the condition $p + q < \alpha - \beta$. As discussed in [30], for a realistic modelling, it is enough to require existence of $c_{p,q}$ up to order $p + q = 4.5$. Furthermore, we set $E\{\varepsilon\} = 1$ for convenience. Under these constraints, the parameters of the underlying NIG-laws and the corresponding ambit sets have been estimated in [30] for a number of turbulent data sets including the one we use in our analysis. Table A.3.1 lists these parameters.

A.3.2 A stochastic model for turbulent velocity time series

Brownian semi-stationary (BSS) processes, introduced in [9] as potential models for turbulent velocity time series, are stochastic processes of the form

$$Z_t = \mu + \int_{-\infty}^{t} g(t-s)\,\sigma_s \mathrm{d}W_s + \int_{-\infty}^{t} q(t-s)\,a_s \mathrm{d}s, \tag{A.19}$$

where μ is a constant, $(W_t)_{t\in\mathbb{R}}$ is standard Brownian motion, g and q are nonnegative deterministic functions on $\mathbb{R}$, with $g(t) = q(t) = 0$ for $t \le 0$, and $(\sigma_t)_{t\in\mathbb{R}}$ and $(a_t)_{t\in\mathbb{R}}$ are càdlàg processes. When (σ, a) is stationary and independent of W, then Z is stationary.

In general, BSS processes are not necessarily semimartingales. However, in our modelling application, the choice of the ingredients of the model (A.19) ensures the semimartingale property. For that reason, we focus on the special case where (A.19) constitutes a semimartingale, keeping in mind that many of the arguments in the sequel are equally true for the general case. A sufficient condition for Z to be a semimartingale is that σ and a have second finite moment, $g, q \in L^1(\mathbb{R}_+) \bigcap L^2(\mathbb{R}_+)$, $g' \in L^2(\mathbb{R}_+)$ and $g(0+) < \infty$ (see [17]).

Following [8], we model time series of the main component v_t of the turbulent velocity field as a BSS process of the specific form

$$v_t = v_t(g,\sigma,\beta) = \int_{-\infty}^{t} g(t-s)\,\sigma_s \mathrm{d}W_s + \beta \int_{-\infty}^{t} g(t-s)\,\sigma_s^2 \mathrm{d}s \equiv R_t + \beta S_t \tag{A.20}$$

where $g \in L^1(\mathbb{R}_+) \cap L^2(\mathbb{R}_+)$ is a non-negative function, σ is a stationary process independent of W with $E[\sigma^6] < \infty$, and β is a constant.

It is well-known that for any semimartingale X the limit

$$[X]_t = \lim_{n\to 0} \sum_{j=1}^{n} \left(X_{j\frac{t}{n}} - X_{(j-1)\frac{t}{n}} \right)^2 \tag{A.21}$$

exists as a limit in probability. The derived process $[X]$ expresses the cumulative variation exhibited by X and is called quadratic variation. For the case where (A.20) is a semimartingale, using Itô algebra, we get

$$(\mathrm{d}v_t)^2 = g^2(0+)\,\sigma_t^2 \mathrm{d}t$$

and

$$[v]_t = \int_0^t (\mathrm{d}v_s)^2 = g^2(0+) \int_0^t \sigma_s^2 \mathrm{d}s. \tag{A.22}$$

In this setting, the quantity $(\mathrm{d}v_t)^2/\mathrm{d}t$ is the natural analogue of the squared first order derivative of v which in the classical formulation is taken to express the temporal energy dissipation (A.2). Consequently, the quadratic variation $[v]$ is the stochastic analogue of the integrated energy dissipation and σ^2 can be identified, up to a factor, with the temporal energy dissipation. We will therefore assume that $\sigma^2 = \varepsilon$, where ε is the process given by (A.16). Note that in this set-up the Kolmogorov variable V (A.15) is given as

$$V_{t,s} = \frac{u_s(t-s/2)}{\left(\bar{v}\,[v]_s^t\right)^{1/3}} \tag{A.23}$$

and the conditional independence refers to the independence of $p\left(V_{t,s}|\,[v]_s^t\right)$ on $[v]_s^t$. Here $[v]_s^t = [v]_{t+s/2} - [v]_{t-s/2}$ denotes the quadratic variation over the time horizon $[t-s/2, t+s/2]$.

The limit in (A.21) may not exist in the non-semimartingale case. However, even in the general case, σ^2 can still be identified, up to a normalisation, with the surrogate energy dissipation (see [17, Theorem 3.1]).

Note that $[v]$ in (A.22) is independent of the second term in (A.20). This second term determines the skewness of the density of velocity increments $u_t = v_t - v_0$. For this reason we refer to β as the skewness parameter.

For the specification of the kernel g we start the discussion following [29] where a convolution of gamma kernels was proposed to model the second order statistics of turbulent velocity time series. The gamma kernel is defined as

$$h(x; a, \nu, \lambda) = a \cdot x^{\nu-1} \exp(-\lambda x)\, 1_{(0,\infty)}(x), \tag{A.24}$$

with $a > 0$, $\nu > 0$ and $\lambda > 0$.

The convolution of two gamma kernels, $h_1(x; a_1, \nu_1, \lambda_1)$ and $h_2(x; a_2, \nu_2, \lambda_2)$, can be expressed as

$$\begin{aligned} g(x; a_1, \nu_1, \lambda_1, a_2, \nu_2, \lambda_2) &= h_1(x; a_1, \nu_1, \lambda_1) * h_2(x; a_2, \nu_2, \lambda_2) \\ &= a_1 a_2 x^{\nu_1+\nu_2-1} 1_{\mathbb{R}_+}(x) \int_0^1 e^{-x(\lambda_1 u + \lambda_2(1-u))} u^{\nu_1-1} (1-u)^{\nu_2-1} \mathrm{d}u. \end{aligned} \tag{A.25}$$

The relevant parameters are $(a, \nu_1, \lambda_1, \nu_2, \lambda_2)$ with $a = a_1 a_2$. We say that a function is a 2-gamma kernel if it can be written as the convolution of two gamma kernels.

In the following g will denote a 2-gamma kernel with parameters $(a, \nu_1, \lambda_1, \nu_2, \lambda_2)$, $\lambda_1 < \lambda_2$. In [29] it is shown that the sdf $\widehat{r}_v(\omega; g, \sigma, \beta)$ of (A.20) is then given as

$$\widehat{r}_v(\omega; g, \sigma, \beta) = a^2 \left(1 + \beta^2 \widehat{r}_{\sigma^2}(\omega)\right) \left(1 + \left(\frac{2\pi\omega}{\lambda_1}\right)^2\right)^{-\nu_1} \left(1 + \left(\frac{2\pi\omega}{\lambda_2}\right)^2\right)^{-\nu_2}, \tag{A.26}$$

where $\widehat{r}_{\sigma^2}$ is the sdf of the process σ^2. Referring to the parameters of g, we also write $\widehat{r}_v(\omega; a, \nu_1, \lambda_1, \nu_2, \lambda_2, \sigma, \beta)$ for $\widehat{r}_v(\omega; g, \sigma, \beta)$.

Ignoring the skewness term in (A.20), i.e. $\beta = 0$, the sdf $\widehat{r}_v$ of the velocity field v behaves as

$$\widehat{r}_v(\omega; a, \nu_1, \lambda_1, \nu_2, \lambda_2, \sigma, 0) \propto \begin{cases} 1 & \omega \ll \lambda_1/2\pi \\ \omega^{-2\nu_1} & \lambda_1/2\pi \ll \omega \ll \lambda_2/2\pi \\ \omega^{-2(\nu_1+\nu_2)} & \omega \gg \lambda_2/2\pi. \end{cases} \tag{A.27}$$

Thus, for $\nu_1 = 5/6$, λ_1 denotes the frequency where the inertial range starts and λ_2 denotes the frequency where the inertial range ends. The value $\nu_1 = 5/6$ reflects Kolmogorov's 5/3rd law [20] and $2(\nu_1 + \nu_2)$ gives the slope within the dissipation range (large frequencies). For the general case, i.e. $\beta \neq 0$, the previous interpretation remains essentially true since, for the data set we analyzed, $\beta^2 \widehat{r}_{\sigma^2}(\omega) \ll 1$ for $\lambda_1/2\pi \ll \omega$.

From (A.25) it follows that $v(g, \sigma, \beta)$ is a semimartingale for $\nu_1 + \nu_2 > 3/2$. Kolmogorov's 5/3rd law implies that $\nu_1 \approx 5/6$. Combining this with the estimated value of ν_2 (by fitting the empirical sdf at large frequencies) shows that $\nu_1 + \nu_2 > 3/2$ for the data set we analyzed (see Section A.4.2). It is for this reason that our focus is on the semimartingale case.

In the semimartingale case, the convolution of gamma kernels does not allow to identify the process

σ^2 with the energy dissipation since $g(0+) = 0$. Therefore, we propose for the kernel g a shifted 2-gamma kernel

$$g(x; x_0, a_1, a_2, \nu_1, \nu_2, \lambda_1, \lambda_2) = (h_1(\cdot; a_1, \nu_1, \lambda_1) * h_2(\cdot; a_2, \nu_2, \lambda_2))(x + x_0) 1_{\mathbb{R}_+}(x), \tag{A.28}$$

where $h_1(x; a_1, \nu_1, \lambda_1)$ and $h_2(x; a_2, \nu_2, \lambda_2)$ are gamma kernels and x_0 is a positive constant. We say that a function is a $(2, x_0)$-gamma kernel if it can be expressed as (A.28).

Figure A.4.1 shows, as an illustrating example, an excerpt of the empirical velocity time series and from a simulation of the model (A.20) using a $(2, 10^{-7})$-gamma kernel for g, a cascade model for σ^2 and the estimated β (see Section A.4.2). The sdf for the data and the sdf from the simulation are compared in Figure A.4.1(c). We can identify three characteristic regimes in the empirical sdf: a flat part at small frequencies, a scaling regime with approximate exponent $-5/3$ and a steeper part at the large frequencies. The central part reflects Kolmogorov's 5/3rd law [20], which is expected to hold in the inertial range. The influence of the shift x_0 on the sdf of v_t can be expected to be negligible for frequencies $\omega \ll x_0^{-1}$. To confirm this conjecture we include in Figure A.4.1(c) the sdf of the model with $x_0 = 0$ and all the other parameters unchanged. Differences only arise at frequencies around 10^5. This implies that for not too large frequencies the interpretation of the parameters of the model according to (A.27) remains valid. The inclusion of the shift x_0 bends the sdf away from the scaling $\omega^{-2(\nu_1+\nu_2)}$ at the very large frequencies.

A.4 Simulation results

In this Section we compare, in detail, the statistical properties of the model (A.20) to the stylized features described in Section A.2. We also briefly mention some aspects of the numerics behind the simulation and discuss how the skewness parameter β and the $(2, x_0)$-gamma kernel g can be estimated from empirical time series.

The data set we analysed consists of one-point time records of the longitudinal (along the mean flow) velocity component in a gaseous helium jet flow. We refer to [13] for more information about the data set. In Table A.3.1 we list the Taylor Reynolds number R, the Kolmogorov scale η, the integral scale I and the sampling frequency $\hat{f}$.

A.4.1 Model performance

The performance of the (A.20) for modelling turbulent time series is illustrated by comparing the marginal distributions of velocity increments, the structure functions and the conditional independence of the Kolmogorov variable. The estimation of the model parameters is based on the analysis of the sdf and the third order structure function and on the marginal distribution and the correlators of the derived energy dissipation.

Figures A.4.1(a) and A.4.1(b) show examples of time series of the velocity from data and from the simulation, respectively. The similarity between the characteristics of both time series is clearly present. A first quantitative result is given in Figure A.4.1(c) displaying the corresponding sdf. The model reproduces the empirical sdf for the whole range of observed frequencies.

The excellent agreement for the sdf translates directly to the corresponding second order structure functions that are shown in Figure A.4.2(a). Note that the sdf (or equivalently the second order structure function) is the basic observable that determines the parameters of the $(2, x_0)$-gamma kernel g used for the simulation. The excellent agreement for the sdf (and S_2) strongly indicates that the parametric choice of a $(2, x_0)$-gamma kernel is appropriate.

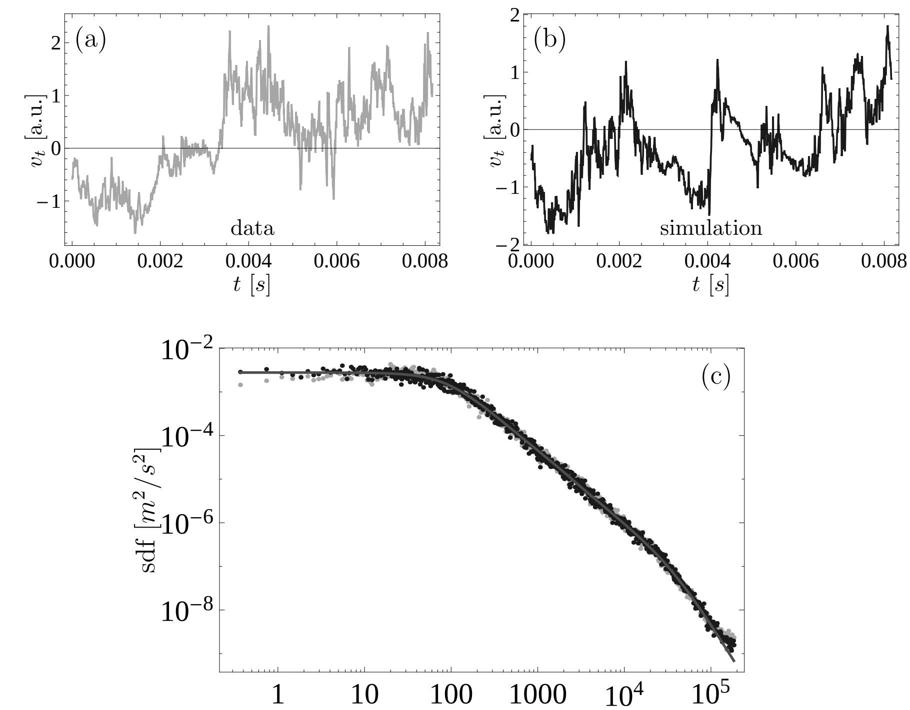

Figure A.4.1: (a) Excerpt of the empirical time series (in arbitrary units). (b) Excerpt of the simulated time series (in arbitrary units) using the model (A.20). (c) Comparision of the sdf from the data (gray dots) and from the simulation of the model (A.20) (black dots). The solid line corresponds to the sdf obtained from the simulation with $x_0 = 0$ and all other parameters unchanged.

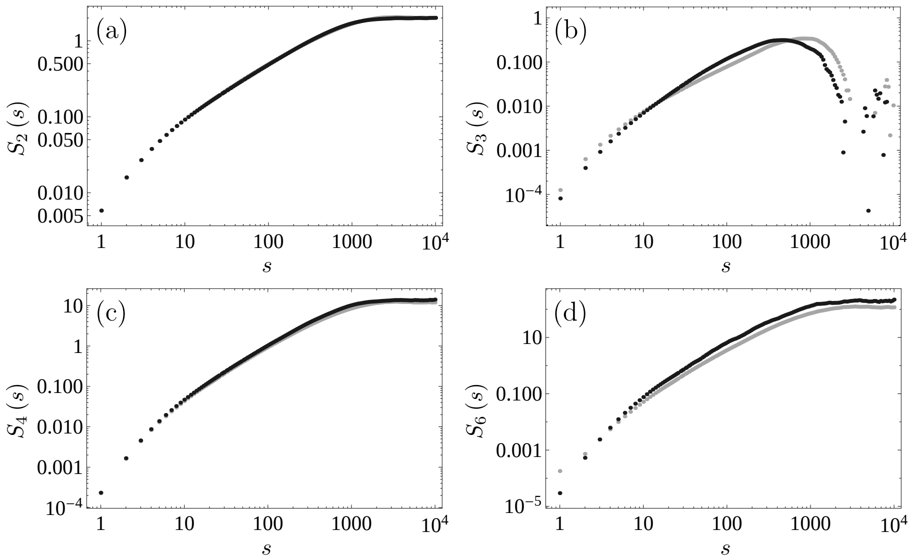

Figure A.4.2: Comparison of the structure functions S_n (A.10), $n = 2, 3, 4, 6$ from the simulation of (A.20) (black) and the structure functions estimated from the data (grey). The time lags s are measured in units of the finest resolution of the empirical data.

The estimation of the skewness parameter β is essentialy based on fitting the third order structure function S_3 (see Figure A.4.2(b)). Taking into account the notorious uncertainty for the estimation of S_3 from turbulent data, the model captures well the details of $S_3(t)$.

Examples of higher order structure functions are shown in Figures A.4.2(c) and A.4.2(d). Again the model shows excellent agreement. Only for S_6 some small systematic deviation is observed which is due to an amplification of small errors not visible for S_2 and slightly visible for S_3. It is important to note that the model is completely specified from the energy dissipation statistics and the structure functions S_2 and S_3 with no adjustable parameter for tuning the behavior of S_4 and S_6.

Figure A.4.3 shows the densities of velocity increments u_s for various time lags s. The densities evolve from semi-heavy tails at small time scales towards a Gaussian shape at the large time scales. NIG distributions fit these densities very well for all scales and all amplitudes in full agreement with the results reported in [7, 6]. The corresponding steepness and skewness parameters are shown in the NIG shape triangle in Figure A.4.4. Again, simulation and data show a good agreement.

Figure A.4.5 illustrates the performance of the model concerning the conditional independence of the densities of the Kolmogorov variable. For t within the inertial range, $p\left(V_{t,s}|\,[v]_s^t\right)$ is independent of $[v]_s^t$.

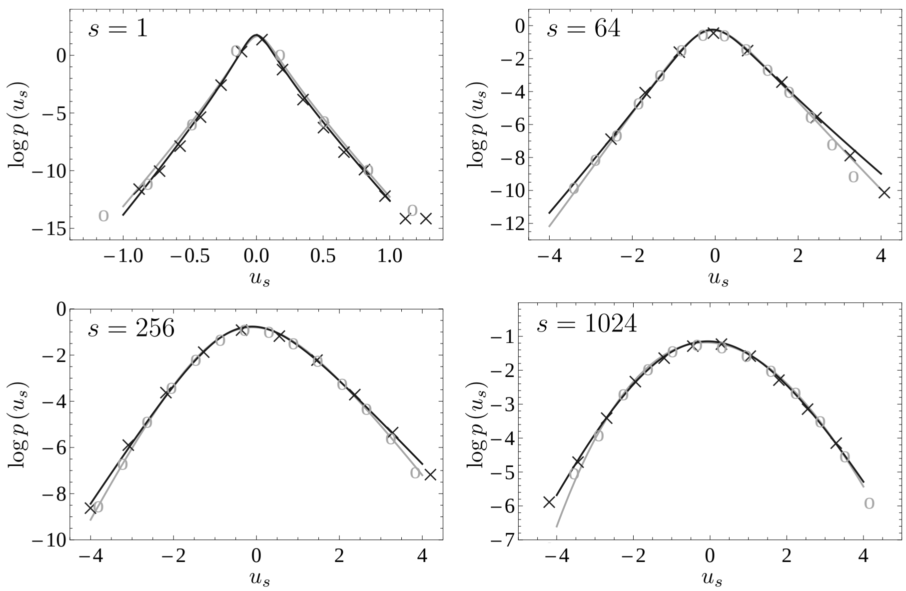

Figure A.4.3: Comparison of the densities of velocity increments $p(u_s)$, $s = 1, 64, 256, 1024$ from data (grey circles) and from the simulation of (A.20) (black crosses). The solid lines correspond to fitted NIG-distributions based on maximum likelihood estimation.

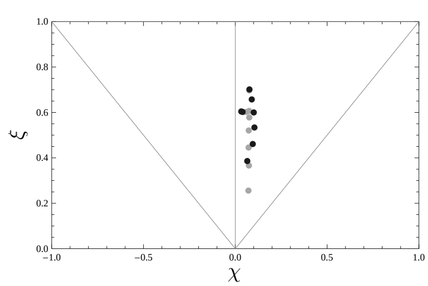

Figure A.4.4: NIG-shape triangle for the evolution of the pdf of velocity increments across lags for the data (grey) and for the simulation of (A.20) (black). Each point corresponds to a different time lag $s = 1, 4, 16, 32, 64, 128, 256, 512, 1024$, increasing from top to bottom.

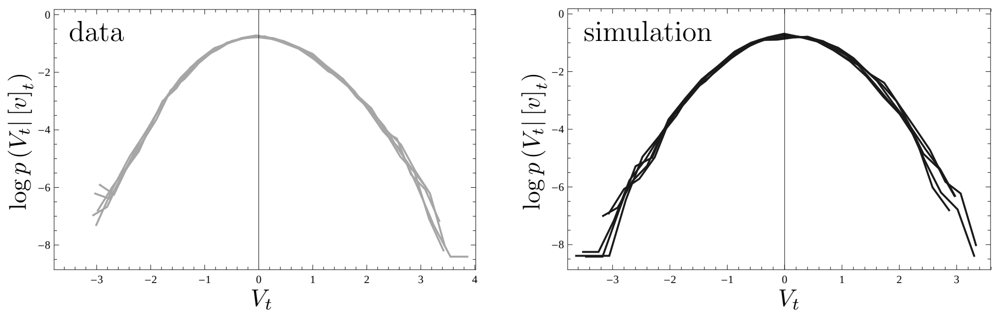

Figure A.4.5: Comparison of the conditional densities $p\left(V_{t,s} \middle| [v]_s^t\right)$ of the Kolmogorov variable (A.23) from the data and from the simulation of (A.20) for $t = 128$ (in units of the finest resolution) and values $[v]_s^t = 0.8, 0.9, 1, 1.1, 1.2$.

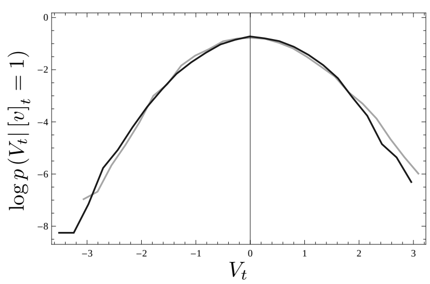

Figure A.4.6: Comparison of the conditional density $p\left(V_{t,s} \middle| [v]_s^t = 1\right)$ from the data (grey) and from the simulation of (A.20) (black).

The values of $[v]_s^t$ cover the core of the distribution of $[v]_s^t$ for which a sufficient sample size is ensured. Figure A.4.6 shows a direct comparison of $p\left(V_{t,s} \middle| [v]_s^t = 1\right)$ for the data and the simulation, showing the strong similarity of the distributions.

Finally, Figure A.4.7 shows the correlators of order $(1, 1)$ and $(1, 2)$ of the energy dissipation estimated from the empirical velocity time series and from the simulation of (A.20). Besides small scale scatter, data and simulations show (nearly) perfect agreement.

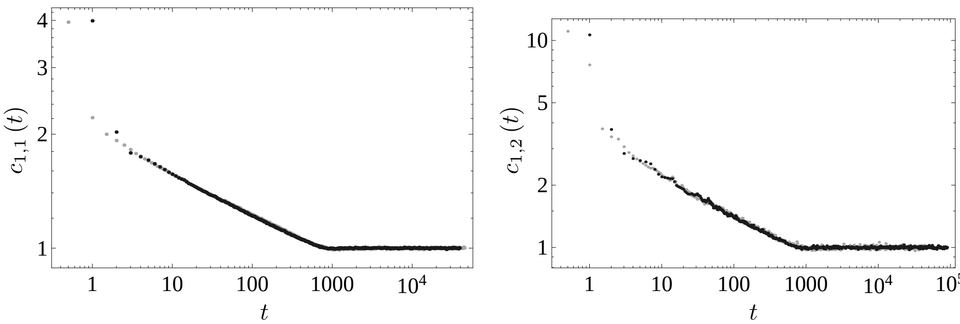

Figure A.4.7: Comparison of the correlators $c_{p,q}(t)$, $(p,q) = (1,1), (1,2)$, of the energy dissipation estimated from the data (grey) and from the simulation of (A.20) (black). The time lags t are measured in units of the finest resolution of the data.

A.4.2 Simulation of *BSS* processes

In this Subsection we briefly describe how we simulate the *BSS*-process (A.20) using time series of σ^2. The algorithm is based on

$$v_t(g,\sigma,\beta)\,|\sigma \sim N\left(\beta\int_0^\infty g(s)\,\sigma_{t-s}^2\mathrm{d}s, \int_0^\infty g^2(s)\,\sigma_{t-s}^2\mathrm{d}s\right) \tag{A.29}$$

which provides the conditional distribution for $v_t(g,\sigma,\beta)\,|\sigma^2$. In principle, to reproduce $v_t(g,\sigma,\beta)\,|\sigma^2$ a complete path for σ^2 is required. However, using a sufficiently small mesh, a linear interpolation on σ^2 gives the desired accuracy.

Let $\delta > 0$ and $N \in \mathbb{N}$. Assume that we know the values $\sigma_{i\cdot\delta}^2$ for $i = 0, \ldots, N$. The linear interpolation for σ^2 is given by the formula

$$\widetilde{\sigma}_s^2 = \sum_{i=1}^{N-1}\left\{\left(\frac{\sigma_{i+1}^2-\sigma_i^2}{\delta}\right)(s-i\cdot\delta)+\sigma_i^2\right\}1_{[i\cdot\delta,(i+1)\delta]}(s).$$

Thus,

$$v_t(g,\widetilde{\sigma},\beta) \sim N\left(\beta\int_0^\infty g(s)\,\widetilde{\sigma}_{t-s}^2\mathrm{d}s, \int_0^\infty g^2(s)\,\widetilde{\sigma}_{t-s}^2\mathrm{d}s\right).$$

Assuming that N is large enough, and since $g \in L_1(\mathbb{R}_+) \cap L_2(\mathbb{R}_+)$, we can, for $\varepsilon > 0$, choose $q \gg 0$ such that

$$\left|\int_0^\infty g(s)\,\widetilde{\sigma}_{t-s}^2\mathrm{d}s - \int_0^q g(s)\,\widetilde{\sigma}_{t-s}^2\mathrm{d}s\right| < \varepsilon.$$

Thus we replace the kernel g by the approximation $\widetilde{g} = g \cdot 1_{[0,q]}$. Let $q = n\cdot\delta$, for $n \in \mathbb{N}$. Expanding and applying a change of variable, we have that

$$\int_0^\infty \widetilde{g}(s)\,\widetilde{\sigma}_{t-s}^2\mathrm{d}s = \delta\sum_{k=0}^{n}\int_0^1 \widetilde{g}(\delta\{y+k\}+t)\left(\sigma_{j-k+1}^2(1-y)+\sigma_{j-k}^2 y\right)\mathrm{d}y, \tag{A.30}$$

$$\int_0^\infty \widetilde{g}^2(s)\,\widetilde{\sigma}_{t-s}^2 \mathrm{d}s \quad = \quad \delta \sum_{k=0}^{n} \int_0^1 \widetilde{g}^2(\delta\{y+k\}+t)\left(\sigma_{j-k+1}^2(1-y)+\sigma_{j-k}^2 y\right)\mathrm{d}y, \qquad \text{(A.31)}$$

for $t \geq \delta \cdot n$. Therefore, $v_{j\cdot\delta}(\widetilde{g},\widetilde{\sigma},\beta)\,|\widetilde{\sigma}, j = n,\ldots,N-1$, can be obtained by simulating a normal random variable with mean (A.30) and variance (A.31). Thus, by interpolation, we obtain a path for $v_t(g,\sigma,\beta)\,|\sigma$ on the interval $[\delta \cdot n, N \cdot \delta]$ through the approximation $v_t(\widetilde{g},\widetilde{\sigma},\beta)\,|\widetilde{\sigma}$.

A.4.3 Estimation procedure

Our modelling framework (A.20) has three degrees of freedom, the energy dissipation σ^2, the kernel g and the skewness parameter β. The energy dissipation can be estimated from velocity increments at the smallest time scale. This has been done in [30] for the data set in the present study. Following [30], we use a NIG Lévy basis in (A.16) and the ambit set to be of the form (A.17). The relevant parameters are listed in Table A.3.1.

It remains to estimate β and the kernel g within the class of $(2, x_0)$-gamma kernels. Given a value of β, our estimators for the parameters of g are those that minimize the distance, in the sense of least squares, between the empirical sdf and the sdf of (A.20) using $x_0 = 0$. As a consequence of the physical interpretation derived from (A.27), the minimization is performed restricting λ_1 to values around the initial frequencies of the inertial range, and restricting λ_2 to values near the end of the inertial range. These constraints have proven to produce good approximations to the empirical sdf. The shift x_0 is obtained by fitting the very large frequency behaviour of the sdf. It is important to note that the scatter of the data at large frequencies does not allow to estimate a precise value of x_0. We choose $x_0 = 10^{-7}$ by visual inspection.

For the estimation of the skewness parameter β, we consider the third order structure function S_3 of (A.20) rewritten as

$$S_3(l) = 3\beta E\{(\Delta_l R)^2(\Delta_l S)\} + \beta^3 E\{(\Delta_l S)^3\}, \qquad \text{(A.32)}$$

where $\Delta_l R = R_l - R_0$ and $\Delta_l S = S_l - S_0$, for $l > 0$. Given paths of R and S, our estimator for β is the value that minimizes the distance, in the sense of least squares, between the empirical third order structure function and (A.32) for a suitable range of scales l (between the smallest scale and the location of the peak, see Figure A.4.2(b)).

The complete estimation procedure can be described as follows. We first neglect the skewness parameter β and we estimate the parameters of g under this restriction from the sdf. Then, having a simulation of the σ process, we perform a simulation of (A.20). Using this simulation, we estimate β as described above. Next, we re-estimate the kernel g using the empirical sdf and the current value of β. We perform this algorithm until we observe stabilisation of β. This algorithm has proven to stabilise after 7 iterations. Figure A.4.8 shows the parameters of g and β obtained after each iteration. The resulting g function is depicted in Figure A.4.9.

The algorithm described above produces similar kernels in each iteration. Therefore it is reasonable to assume that the L^2-distance between these kernels is small.

The following Lemma provides some bound for the convergence of (A.20).

Lemma 20 *Consider the model (A.20). Let $\beta, \beta_1 > 0$ and $g, g_1 \in L_2(\mathbb{R})$. Assume that:*

1. $c \equiv E[\sigma^4] < \infty$.
2. $\max\{\|g_1 - g\|_{L_2(\mathbb{R})}, \|g_1 - g\|_{L_1(\mathbb{R})}\} < \frac{\varepsilon}{(1+c^{1/4}\beta_1)c^{1/4}}$, *for* $\varepsilon > 0$.

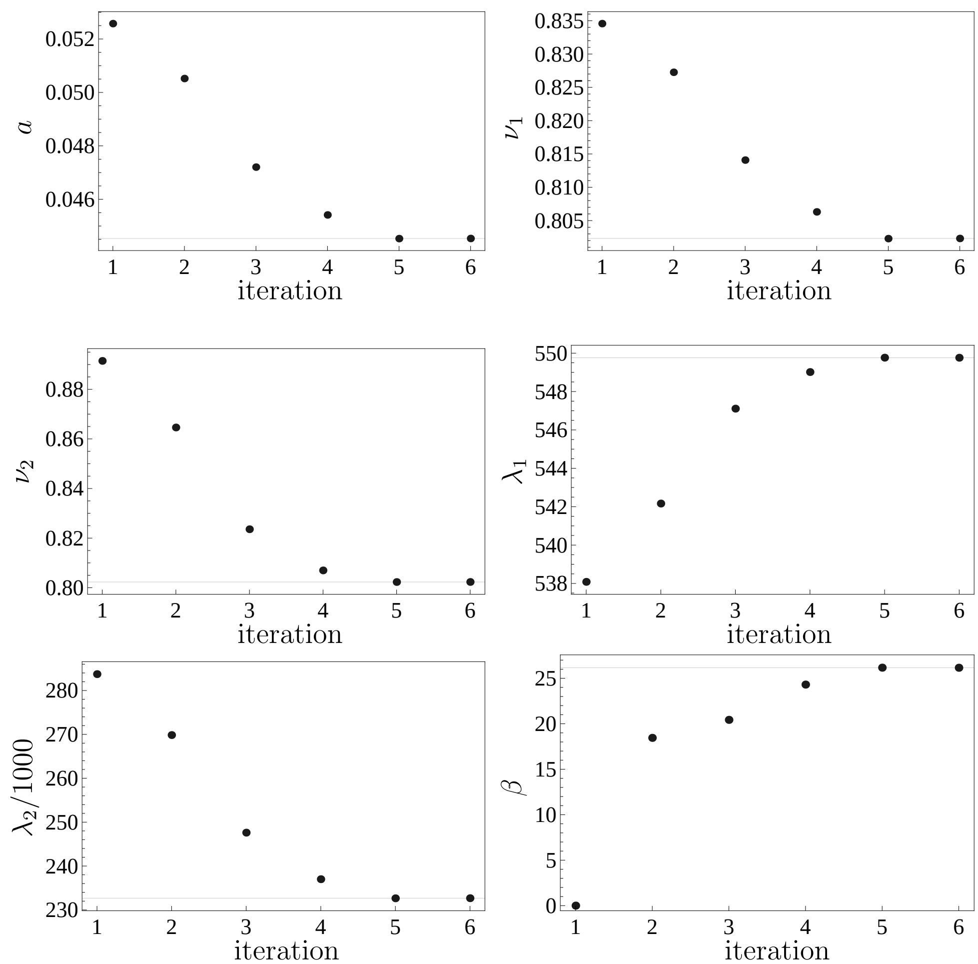

Figure A.4.8: Estimated values of the parameters of the $(2, x_0)$-gamma kernel g with $x_0 = 10^{-7}$ and estimated value of the skewness parameters β for each of the iterations performed.

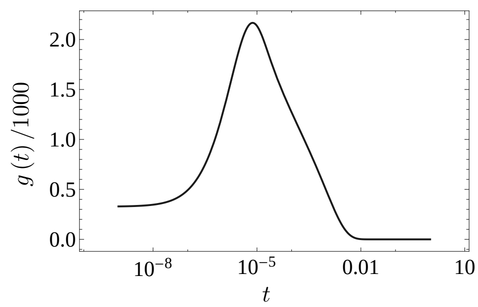

Figure A.4.9: The $(2, 10^{-7})$-gamma kernel g corresponding to the estimated parameters in Figure A.4.8 (iteration 7). Time t is measured in units of the finest resolution of the data.

Then, we have that

$$\left\| v_t\left(g_1, \sigma, \beta_1\right) - v_t\left(g, \sigma, \beta\right)\right\|_{L_2(\Omega)} \leq c^{1/2}|\beta_1 - \beta| \, \|g\|_{L_1(\mathbb{R})} + \varepsilon.$$

Proof. Let

$$\begin{aligned} P_t &= \int_{-\infty}^{t} \{g_1(t-s) - g(t-s)\} \sigma_s dB_s, \\ Q_t &= \int_{-\infty}^{t} \left\{\beta_1 g_1(t-s) - \beta g(t-s)\right\} \sigma_s^2 ds \end{aligned}$$

for $t > 0$. Observe that

$$Q_t = \beta_1 \int_{-\infty}^{t} \{g_1(t-s) - g(t-s)\} \sigma_s^2 ds + \left(\beta_1 - \beta\right) \int_{-\infty}^{t} g(t-s) \sigma_s^2 ds.$$

Then

$$\begin{aligned} \left\| v_t\left(g_1, \sigma, \beta_1\right) - v_t\left(g, \sigma, \beta\right)\right\|_{L_2(\Omega)} &\leq \|P_t\|_{L_2(\Omega)} + \left|\beta_1 - \beta\right| \left\| \int_{-\infty}^{t} g(t-s) \sigma_s^2 ds \right\|_{L_2(\Omega)} \\ &\quad + \beta_1 \left\| \int_{-\infty}^{t} \{g_1(t-s) - g(t-s)\} \sigma_s^2 ds \right\|_{L_2(\Omega)}. \end{aligned}$$

The Cauchy-Schwarz inequality implies that

$$E\left[\sigma_s^2 \sigma_r^2\right] \leq E\left[\sigma_s^4\right]^{1/2} E\left[\sigma_r^4\right]^{1/2} = E\left[\sigma^4\right] = c, \qquad \text{(A.33)}$$

for any $(s, r) \in \mathbb{R}^2$. Thus,

$$\left\| \int_{-\infty}^{t} g(t-s) \sigma_s^2 ds \right\|_{L_2(\Omega)}^2 = E\left[\left(\int_{-\infty}^{t} g(t-s) \sigma_s^2 ds\right)^2\right] \leq c \, \|g\|_{L_1(\mathbb{R})}^2$$

and

$$\left\| \int_{-\infty}^{t} \{g_1(t-s) - g(t-s)\} \sigma_s^2 ds \right\|_{L_2(\Omega)}^2 \leq c \|g_1 - g\|_{L_1(\mathbb{R})}^2 .$$

Itô isometry and (A.33) imply that

$$\|P_t\|_{L_2(\Omega)}^2 \leq c^{1/2} \|g_1 - g\|_{L_2(\mathbb{R})}^2 .$$

Thus,

$$\begin{aligned} \left\| v_t\left(g_1, \sigma, \beta_1\right) - v_t\left(g, \sigma, \beta\right) \right\|_{L_2(\Omega)} &\leq c^{1/4} \|g_1 - g\|_{L_2(\mathbb{R})} + c^{1/2} \left|\beta_1 - \beta\right| \|g\|_{L_1(\mathbb{R})} \\ &\quad + c^{1/2} \beta_1 \|g_1 - g\|_{L_1(\mathbb{R})} \\ &\leq \varepsilon + c^{1/2} \left|\beta_1 - \beta\right| \|g\|_{L_1(\mathbb{R})} , \end{aligned}$$

which concludes the proof. ■

Under certain assumptions, we can ensure that the estimation procedure gives reasonable estimators for the skewness parameter β. Assume that $v_t(g,\sigma,\beta)$ is the ideal model with g in the class of $(2, x_0)$-gamma kernels. Let $\{g_n\}_{n\in\mathbb{N}}$ be a sequence of $(2, x_0)$-gamma kernels and $\{\beta_n\}_{n\in\mathbb{N}}$ a sequence of positive numbers, and assume that $E[\sigma^6] < \infty$, $g_n \to g$ pointwise and g_n bounded by an integrable and square-integrable function.

Following the notation in (A.20), for each $n \in \mathbb{N}$, we decompose $v_t\left(g_n, \sigma, \beta_n\right) = R_t(g_n, \sigma) + \beta_n S_t(g_n, \sigma)$. For $l > 0$, let

$$\begin{aligned} a_l^n &= E\left[(\Delta_l R(g_n, \sigma))^2 (\Delta_l S(g_n, \sigma))\right], \\ a_l &= E\left[(\Delta_l R(g, \sigma))^2 (\Delta_l S(g_n, \sigma))\right], \\ b_l^n &= E\left[(\Delta_l S(g_n, \sigma))^3\right], \\ b_l &= E\left[(\Delta_l S(g_n, \sigma))^3\right], \\ f_n(l; \beta) &= 3\left(\beta a_l - \beta_n a_l^n\right) + \left(\beta^3 b_l - \beta_n^3 b_l^n\right), \end{aligned}$$

where $\Delta_l R(\cdot,\cdot) = R_l(\cdot,\cdot) - R_0(\cdot,\cdot)$ and $\Delta_l S(\cdot,\cdot) = S_l(\cdot,\cdot) - S_0(\cdot,\cdot)$. Observe that f is the difference of the third order structure functions of $v\left(g, \sigma, \beta_n\right)$ and $v(g_n, \sigma, \beta)$.

Lemma 21 *$a_l^n \to a_l$ and $b_l^n \to b_l$, for any $l > 0$.*

Proof. Define

$$\begin{aligned} \varphi_l^n(s) &:= g_n(l-s) - g_n(-s) 1_{(-\infty,0]}(s) \\ \varphi_l(s) &:= g(l-s) - g(-s) 1_{(-\infty,0]}(s) . \end{aligned}$$

$E[\sigma^6] < \infty$ implies that $(s_1, s_2, s_3) \mapsto E\left[\sigma_{s_1}^2 \sigma_{s_2}^2 \sigma_{s_3}^2\right]$ is a bounded function in $\mathbb{R}^3$. Thus, from the Dominated Convergence Theorem,

$$\begin{aligned} b_l^n &= E\left[(\Delta_l S(g_n, \sigma))^3\right] \\ &= \int_{-\infty}^{l} \int_{-\infty}^{l} \int_{-\infty}^{l} \varphi_l^n(s_1) \varphi_l^n(s_2) \varphi_l^n(s_3) E\left[\sigma_{s_1}^2 \sigma_{s_2}^2 \sigma_{s_3}^2\right] ds_1 ds_2 ds_3 \end{aligned}$$

$$\rightarrow \int_{-\infty}^{l}\int_{-\infty}^{l}\int_{-\infty}^{l} \varphi_l(s_1)\,\varphi_l(s_2)\,\varphi_l(s_3)\,E\left[\sigma_{s_1}^2\sigma_{s_2}^2\sigma_{s_3}^2\right] ds_1 ds_2 ds_3 = b_l.$$

On the other hand, from Itô isometry and the Dominated Convergence Theorem,

$$\begin{aligned}
a_l^n &= E\left[(\Delta_l R(g_n,\sigma))^2(\Delta_l S(g_n,\sigma))\right] \\
&= E\left[\Delta_l S(g_n,\sigma)\,E\left[(\Delta_l R(g_n,\sigma))^2|\sigma\right]\right] \\
&= E\left[\int_{-\infty}^{l}\varphi_l^n(s)\,\sigma_s^2 ds\cdot E\left[\left(\int_{-\infty}^{l}\varphi_l^n(r)\,\sigma_r dBr\right)^2|\sigma\right]\right] \\
&= \int_{-\infty}^{l}\int_{-\infty}^{l}\varphi_l^n(s)\left(\varphi_l^n(r)\right)^2 E\left[\sigma_s^2\sigma_r^2\right] dsdr \\
&\rightarrow \int_{-\infty}^{l}\int_{-\infty}^{l}\varphi_l(s)\left(\varphi_l(r)\right)^2 E\left[\sigma_s^2\sigma_r^2\right] dsdr = a_l.
\end{aligned}$$

This finishes the proof. ∎

The next Proposition ensures that a converging sequence β_n converges to the right value.

Proposition 22 *Assume there is some $l_0 > 0$ such that $f_n(l_0;\beta) = 0$ and that β_n converges. Then, $\beta_n \rightarrow \beta$.*

Proof. Assume that $l_0 > 0$ satisfies $f_n(l_0;\beta) = 0$. Then,

$$\lim_{n\rightarrow\infty}\left\{3\left(\beta a_{l_0} - \beta_n a_{l_0}^n\right) + \left(\beta^3 b_{l_0} - \beta_n^3 b_{l_0}^n\right)\right\} = 0.$$

This implies

$$3\left(\beta - \beta_*\right) a_{l_0} + \left(\beta^3 - \beta_*^3\right) b_{l_0} = 0,$$

where $\beta_* = \lim \beta_n$. Therefore,

$$\left(\beta - \beta_*\right)\left(3a_{l_0} + \left(\beta^2 + \beta\beta_* + \beta_*^2\right) b_{l_0}\right) = 0,$$

Thus, necessarily, $\beta_* = \beta$. ∎

A.5 Conclusion

The analysis performed in this paper clearly demonstrates that Brownian semi-stationary processes are well adapted to reproduce key characteristics of turbulent time series. The parameters of the model are solely estimated from the marginal distribution and the correlator $c_{1,1}$ of the energy dissipation [30] and from second and third order structure functions of velocity increments. This has been done under the specific model specification (A.20) with a $(2, x_0)$-gamma kernel g. The use of a $(2, x_0)$-gamma kernel is motivated by its ability to reproduce the empirical sdf. The fact that, starting from second order and third order structure functions, higher order structure functions, the evolution of the densities of velocity increments across scales and the essential statistics of the Kolmogorov variable are also reproduced clearly indicates the appropriateness of the semi-parametric model (A.20).

In [12, 19] a similar approach for modelling turbulent velocity time series is suggested. They propose a causal continuous-time moving average of the form

$$Y_t = \int_{-\infty}^{t} g(t-s)\, dL_s \tag{A.34}$$

where L is a Lévy process with zero mean and finite second moment. A non-parametric estimation of the kernel g from second order statistics of turbulent data shows the same qualitative behaviour as the kernel estimated in the current study. The extraction of the driving noise L from velocity time series is addressed in [19] showing that the autocorrelation of the energy dissipation resulting from the model agrees well with empirical findings. The performance of (A.34) for turbulence modelling beyond second order statistics is, however, not discussed. It would be interesting to compare the two approaches in more detail, including more of the stylised features discussed in the present paper.

The data set analyzed here has a relatively high Reynolds number, with a visible inertial range. It is important to investigate how the model performs for lower Reynolds numbers, where inertial range scaling is not observed. Concerning the model for the energy dissipation, this has been done in [30] where it was shown that continuous cascades are equally suitable for a wide range of Reynolds numbers. For the velocity field itself this is work in progress.

Acknowledgements

The authors are much indebted to B. Chabaud for allowing the use of the data set. The first author is grateful to the Mexican National Council of Science and Technology CONACYT for its support through the Ph.D. grant 217040. Finally, the authors are very thankful to Emil Hedevang Sørensen for the many stimulating discussions and suggestions in the preparation of this work. The authors are also grateful to the referee of this paper for helpful comments.

Appendix

The normal inverse Gaussian (NIG) distribution is a four-parameter family of continuous probability distributions whose probability density function is given by

$$f_{\mathrm{NIG}}(x; \alpha, \beta, \mu, \delta) = \frac{\alpha e^{\delta\gamma}}{\pi} e^{\beta(x-\mu)} \frac{K_1\left(\delta\alpha q\left(\frac{x-\mu}{\delta}\right)\right)}{q\left(\frac{x-\mu}{\delta}\right)}, \tag{A.35}$$

where $\gamma = \alpha^2 - \beta^2$, $q(x) = \sqrt{1+x^2}$ and K_1 denotes the modified Bessel function of the second kind with index 1. The domain of variation of the parameters is given by $\mu \in \mathbb{R}$, $\delta \in \mathbb{R}_+$, and $0 \leq |\beta| < \alpha$. The parameters α and β are shape parameters, μ determines the location, and δ determines the scale. The distribution is denoted by $\mathrm{NIG}(\alpha, \beta, \mu, \delta)$.

The normal inverse Gaussian distribution is a subclass of the generalised hyperbolic distribution. These distributions were introduced by Barndorff-Nielsen [5] to describe the law of the logarithm of the size of sand particles.

The cumulant function $K(z; \alpha, \beta, \mu, \delta) = \log E[\exp\{zX\}]$ of a random variable X with distribution $\mathrm{NIG}(\alpha, \beta, \mu, \delta)$ is given by

$$K(z; \alpha, \beta, \mu, \delta) = z\mu + \delta\left(\gamma - \sqrt{\alpha^2 - (\beta+z)^2}\right). \tag{A.36}$$

It follows immediately from this that the normal inverse Gaussian distribution is infinitely divisible. Namely, if $X_i \sim \text{NIG}\left(\alpha, \beta, \mu_i, \delta_i\right)$, $i = 1, 2$, are independent random variables, then we have $X_1 + X_2 \sim \text{NIG}(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2)$.

It is often of interest to consider alternative parametrisations of the normal inverse Gaussian laws. In particular, letting $\overline{\alpha} = \delta\alpha$ and $\overline{\beta} = \delta\beta$, we have that $\overline{\alpha}$ and $\overline{\beta}$ are invariant under location-scale changes.

Sometimes it is useful to represent NIG distributions in the so-called shape triangle. Consider the alternative asymmetry and steepness parameters χ and ξ defined by

$$\xi = (1 + \overline{\gamma})^{-1/2}, \quad \chi = \rho\xi,$$

where $\rho = \beta/\alpha$ and $\overline{\gamma} = \delta\gamma = \delta\sqrt{\alpha^2 - \beta^2}$. These parameters are invariant under location-scale changes. Their range defines the NIG shape triangle

$$\{(\chi, \xi) : 0 < \xi < 1, -\xi < \chi < \xi\}.$$

When $\chi = 0$ the NIG distribution is symmetric. Values $\chi > 0$ indicate a positively skewed distribution and $\chi < 0$ a negatively skewed law. The steepness parameter ξ measures the heaviness of the tails of the NIG distribution. The limiting case $\xi = 0$ corresponds to a normal distribution.

References

[1] F. Anselmet, Y. Gagne, E.J. Hopfinger, and R.A. Antonia. Higher-order velocity structure functions in turbulent shear flows. *J. Fluid Mech.*, **140**:63–89, 1984.

[2] R.A. Antonia, B.R. Satyaprakash, and A.K.M.F. Hussain. Statistics of fine-scale velocity in turbulent plane and circular jets. *J. Fluid Mech.*, **119**:55–89, 1982.

[3] A. Arneodo, C. Baudet, F. Belin, R. Benzi, B. Castaing, R. Chavarria, S. Ciliberto, R. Camussi, F. Chillà, B. Dubrulle, Y. Gagne, B. Hebral, J. Herweijer, M. Marchand, J. Maurer, J.F. Muzy, A. Naert, A. Noullez, J. Peinke, F. Roux, P. Tabeling, W. van de Water, and H. Willaime. Structure functions in turbulence, in various flow configurations, at Reynolds numbers between 30 and 5000, using extended self-similarity. *Europhys. Lett.*, **34**:411–416, 1996.

[4] Arneodo.A, E. Bacry, and J.F. Muzy. Random cascades on wavelet dyadic trees. *J. Math. Phys.*, **39**:4142–4164, 1998.

[5] O.E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. Lond. A*, **353**(1674):401–419, 1977.

[6] O.E. Barndorff-Nielsen, P. Blæsild, and J. Schmiegel. A parsimonious and universal description of turbulent velocity increments. *Eur. Phys. J. B*, **41**:345–363, 2004.

[7] O.E. Barndorff-Nielsen and J. Schmiegel. Time change and universality in turbulence. Technical report, Thiele Centre, University of Aarhus, Denmark, 2006.

[8] O.E. Barndorff-Nielsen and J. Schmiegel. A stochastic differential equation framework for the timewise dynamics of turbulent velocities. *Theory Probab. Appl.*, **52**:372–388, 2008.

[9] O.E. Barndorff-Nielsen and J. Schmiegel. Brownian semistationary processes and volatility/intermittency. In H. Albrecher, W. Runggaldier, and W. Schachermayer, editors, *Advanced Financial Modelling*, pages 1–26. Walter de Gruyter, Berlin, 2009.

[10] R. Benzi, L. Biferale, A. Crisanti, G. Paladin, M. Vergassola, and A. Vulpiani. A random process for the construction of multiaffine fields. *Physica D*, **65**:352–358, 1993.

[11] L. Biferale, G. Boffetta, A. Celani, A. Crisanti, and A. Vulpiani. Mimicking a turbulent signal: Sequential multiaffine processes. *Phys. Rev. E*, **57**:6261–6264, 1998.

[12] P.J. Brockwell, V. Ferrazano, and C. Klüppelberg. High frequency sampling and kernel estimation for continuous-time moving average processes. *J. Time Ser. Anal.*, **34**:385–404, 2013.

[13] O. Chanal, B. Chebaud, B. Castaing, and B. Hébral. Intermittency in a turbulent low temperature gaseous helium jet. *Eur. Phys. J. B*, **17**:309–317, 2000.

[14] J. Cleve, T. Dziekan, J. Schmiegel, O.E. Barndorff-Nielsen, B.R. Pearson, K.R. Sreenivasan, and M. Greiner. Finite-size scaling of two-point statistics and the turbulent energy cascade generators. *Phys. Rev. E*, **71**:026309, 2005.

[15] J. Cleve, M. Greiner, B.R. Pearson, and K.R. Sreenivasan. Intermittency exponent of the turbulent energy cascade. *Phys. Rev. E*, **69**:066316, 2004.

[16] J. Cleve, J. Schmiegel, and M. Greiner. To be and not to be: scale correlations in random multifractal processes. Technical report, Thiele Centre, University of Aarhus, Denmark, 2006.

[17] J.M. Corcuera, E.H.L. Sørensen, M.S. Pakkanen, and M. Podolskij. Asymptotic theory for Brownian semi-stationary processes with application to turbulence. *Stochastic Process. Appl.*, **123**(7):2552–2574, 2013.

[18] F. Elliot, A. Majda, D. Horntrop, and R. McLaughlin. Hierarchical monte carlo methods for fractal random fields. *J. Stat. Phys.*, **81**:717–736, 1995.

[19] V. Ferrazano and C. Klüppelberg. Turbulence modelling by time-series methods. *ArXiv e-prints*, **1205**(6614), 2012.

[20] U. Frisch. *Turbulence. The legacy of A.N. Kolmogorov*. Cambridge University Press, Cambridge, 1995.

[21] M. Greiner, J. Cleve, J. Schmiegel, and K.R. Sreenivasan. Data-driven stochastic processes in fully developed turbulence. In Waymire E.C. and Duan J., editors, *IMA Volume 140: Probability and Partial Differential Equations in Modern Applied Mathematics.* Springer, New York, 2004.

[22] I. Hosokawa, C.W. Van Atta, and S.T. Thoroddsen. Experimental study of the Kolmogorov refined similarity variable. *Fluid Dyn. Res.*, **13**:329–333, 1994.

[23] A. Juneja, D. Lathrop, K.R. Sreenivasan, and G. Stolovitzky. Synthetic turbulence. *Phys. Rev. E*, **49**:5179–5194, 1994.

[24] A.N. Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, **13**:82–85, 1962.

[25] J. Maurer, P. Tabeling, and G. Zocchi. Statistics of turbulence between two counter-rotating disks in low temperature helium gas. *Europhys. Lett.*, **26**:31–36, 1994.

[26] J. Schmiegel. Self-scaling of turbulent energy dissipation correlators. *Phys. Lett. A*, **337**(3):342–353, 2005.

[27] J. Schmiegel, O.E. Barndorff-Nielsen, and H.C. Eggers. A class of spatio-temporal and causal stochastic processes, with application to multiscaling and multifractality. *South African J. of Science*, **101**:513–519, 2005.

[28] J. Schmiegel, J. Cleve, H.C. Eggers, B.R. Pearson, and M. Greiner. Stochastic energy-cascade model for $(1+1)$-dimensional fully developed turbulence. *Phys. Lett. A*, **320**:247–253, 2004.

[29] E.H.L. Sørensen. *Stochastic modelling of turbulence: With applications to wind energy.* PhD thesis, Aarhus University, 2012.

[30] E.H.L. Sørensen and J. Schmiegel. A causal continuous-time stochastic model for the turbulent energy cascade in a helium jet flow. *J. Turbul.*, **14**(11):1–26, 2013.

[31] K.R. Sreenivasan and R.A. Antonia. The phenomenology of small-scale turbulence. *Ann. Rev. Fluid Mech.*, **29**:435–472, 1997.

[32] K.R. Sreenivasan and P. Kailasnath. An update on the intermittency exponent in turbulence. *Phys. Fluids A*, **5**:512–514, 1992.

[33] G. Stolovitzky, P. Kailasnath, and K.R. Sreenivasan. Kolmogorov's refined similarity hypothesis. *Phys. Rev. Lett.*, **69**(2):1178–1181, 1992.

[34] G. Stolovitzky and K.R. Sreenivasan. Kolmogorov's refined similarity hypotheses for turbulence and general stochastic processes. *Rev. Mod. Phys.*, **66**:229–239, 1994.

[35] G.I. Taylor. The spectrum of turbulence. *Proc. R. Soc. Lond. A*, **164**:476–490, 1938.

[36] T. Vicsek and A.L. Barabási. Multi-affine model for the velocity distribution in fully turbulent flows. *J. Phys. A*, **24**:L845, 1991.

[37] Y. Zhu, R.A. Antonia, and I. Hosokawa. Refined similarity hypotheses for turbulent velocity and temperature fields. *Phys. Fluids*, **7**:1637–1648, 1995.

B

Paper B: On the cumulants of increments for two classes of Brownian semi-stationary processes

José Ulises Márquez
Department of Mathematics, Aarhus University.

Abstract

In this article we obtain formulae for the cumulants of the increments of two classes of Brownian semi-stationary (BSS) processes. The first class corresponds to BSS processes where the volatility is a Lévy semi-stationary process and the second class consists in BSS processes where the volatility is given as the exponential of an ambit process. We analyse and apply these formulae having in mind applications in the context of turbulence. Specifically, we apply these formulae to some particular examples of BSS processes that are relevant as models for turbulent velocity time series.

B.1 Introduction

In this note, we study the cumulants of the increments of Brownian semi-stationary processes (BSS). The BSS processes, introduced in [9], are stochastic processes of the form

$$Z_t = \mu + \int_{-\infty}^{t} g(t-s)\,\varepsilon_s dB_s + \int_{-\infty}^{t} q(t-s)\,\varsigma_s ds \tag{B.1}$$

where g and q are deterministic functions with $g(t) = q(t) = 0$ for $t \leq 0$, ε and ς are adapted stochastic processes, $(B)_{t\in\mathbb{R}}$ is standard Brownian motion on $\mathbb{R}$ and μ is a constant. When the process $(\varepsilon_t, \varsigma_t)$ is stationary and independent of B, Z is stationary itself. In general, the BSS processes are not semimartingales which make them a very interesting class of stochastic processes. For a more detailed and extensive discussion about the BSS processes and their use as modelling framework, we refer to [7, 9, 5, 14, 24, 27].

BSS processes have been used to model time series of the turbulent velocity field for stationary and isotropic flows (see [24, 27]). In this paper we restrict our attention to the particular case where $\varsigma = \varepsilon^2$, which arises in the context of turbulence. In this setting, the process ε^2 is interpreted as the turbulent energy dissipation, a process that measures the loss of kinetic energy in a turbulent flow due to friction forces.

The increments of the process Z contain relevant information about the process Z and of its elements (g, q, ε and ς). For instance, the increments may indicate how much the process Z varies, e.g. if it has finite quadratic variation (see, e.g. [5]). The increments of Z also provide a way to estimate its parameters (see [14, 24]). In addition, the increments of Z are relevant for the theory of turbulence since the increments of the velocity field are the object of study in the Kolmogorov-Obukhov theory ([18]), which is probably the most important theory in turbulence.

In this paper we find formulae for the cumulants of the increments of Z assuming that $\varsigma = \varepsilon^2$ and where ε^2 has two forms: 1) ε^2 is a Lévy semi-stationary process (see, e.g., [29]); 2) ε^2 is the exponential of an ambit process driven by a homogeneous Lévy basis on $\mathbb{R}^2$. In the first case, the formulae we find are given in terms of the Lévy seed of the Lévy process that drives ε^2. In the second case, we obtain a formula for the n-th moments of the increments of Z in terms of the Lévy seed of the Lévy basis that drives $\log \varepsilon^2$. These formulae can be used to iteratively compute the cumulants of the increments of Z. The second case is pertinent since it has been used to model the energy dissipation and turbulent velocity time series (see [24, 27]).

This paper is organized as follows. Section 2 contains the necessary background to understand the role of the increments and their importance in turbulence. In Section 3 we derive a formula for the cumulants of the increments of BSS processes assuming that ε^2 is a Lévy semi-stationary process. In Section 4 we apply the formulae derived in Section 3 to some specific examples of Z. Section 5 presents a formula for the n-th moments of ε^2 assuming that it is the exponential of an ambit process. We use this formula to compute iteratively the cumulants of the increments of Z. In Section 6 we apply the formula derived in Section 5 to some specific examples of Z. Section 7 concludes.

B.2 BSS processes and turbulence

BSS processes have been used as phenomenological models for the temporal velocity field in turbulence ([8, 24]). The strength of the BSS modelling framework lies in the fact that the volatility term $\varepsilon^2 = \varsigma$ and the functions g and q can, to a large extent, be chosen arbitrarily.

In these BSS models $q = \beta \cdot g$, where $\beta > 0$. The parameter β controls the skewness of increments of the velocity field and is called the skewness parameter (see [8, 24]). In [27] it is shown that, taking g proportional to the convolution of gamma densities, BSS processes accurately reproduce the spectral density function found for turbulent time series (including the 5/3rd-Kolmogorov law).

According to [8], $\varepsilon^2 = \varsigma$ can be identified with the energy dissipation. The literature provides a number of different phenomenological models for the energy dissipation (see, e.g., [18, 20, 28], and the references there). Of particular interest here are ambit models. Ambit models are a continuous generalization of cascade models ([11, 13, 19, 21, 20, 22, 28]), which are an important and well-studied class of phenomenological models for the energy dissipation. Ambit models are able to reproduce a very rich class of probability laws and their ingredients can be chosen in a way such that the empirically observed scaling and self-scaling of certain moments of the energy dissipation are reproduced. This potential to model the energy dissipation has been illustrated in [28].

In [24], it is shown that the BSS models are able to reproduce the main stylized features of turbulent velocity time series. Assuming that ε^2 is given by the ambit model estimated in [28] and that g and $q = \beta \cdot g$ are proportional to a shifted convolution of gamma densities, the parameters of the BSS model (B.1) are estimated from data obtained in a helium jet experiment. The shift in the convolution of gamma kernels was introduced to reproduce the behavior of the spectral density function at very small scales. The results in [24] show that BSS processes are able to reproduce: 1) the probability law of the increments of the velocity field; 2) the structure functions of order $p = 2, 4, 6$; 3) scaling and self-scaling observed for the energy dissipation; 4) the conditional independence of the Kolmogorov variable; and 5) the skewness of velocity increments. All these features are directly related to the increments of the velocity field. For this reason, it is essential to better understand the increments of BSS processes and to have at hand formulae that allow to compute cumulants without time-consuming simulations of the underlying process.

B.3 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as a LSS process: theory

Consider the BSS process

$$X_t = \int_{-\infty}^{t} g(t-s)\,\varepsilon_s dB_s + \int_{-\infty}^{t} q(t-s)\,\varepsilon_s^2 ds \tag{B.2}$$

where $g \in L^2(\mathbb{R})$ and $q \in L^1(\mathbb{R})$, with $g(x) = q(x) = 0$ for $x \leq 0$, and ε is a Lévy semi-stationary process (LSS) independent of B given by

$$\varepsilon_t^2 = \int_{-\infty}^{t} h(t-s)\,dL_s, \tag{B.3}$$

where, under the truncation function $\tau \equiv 0$, L is a subordinator with characteristic triplet $(m, 0, \nu)$, and $h \in L^1(\mathbb{R})$ is non-negative satisfying $h(x) = 0$ for $x \leq 0$ and

$$\int_{-\infty}^{t} \int_{\mathbb{R}_+} \left(1 \wedge h^2(t-s)\,x^2\right) \nu(dx)\,ds < \infty. \tag{B.4}$$

Besides, we also assume that L has finite first moment, i.e. that

$$E[L_1] = \int_{\mathbb{R}_+} x\nu(dx) < \infty. \tag{B.5}$$

Under these assumptions, the process X is well-defined. Condition (B.4) ensures that the process ε^2 is well-defined (see [10], Corollary 4.1). In the remaining part of this Section, we deduce a formula for the cumulants of the increments of X in terms of the cumulants of $L' \equiv L_1$.

B.3.1 Preliminary calculations

Let $\Delta_t X = X_t - X_0$ for $t > 0$. We have that

$$\Delta_t X = \int_{\mathbb{R}} \phi_t(s)\,\varepsilon_s dB_s + \int_{\mathbb{R}} \psi_t(s)\,\varepsilon_s^2 ds$$

where, for $t > 0$,

$$\phi_t(s) = (g(t-s) - g(-s)), \text{ and } \psi_t(s) = (q(t-s) - q(-s)). \tag{B.6}$$

Then, the cumulant function $C\{z \ddagger \Delta_t X\} := \log E(\exp\{iz\Delta_t X\})$ of $\Delta_t X$ is given as

$$\begin{aligned} C\{z \ddagger \Delta_t X\} &= \log E\left(\exp\left\{-\frac{1}{2}z^2\int_{\mathbb{R}}\phi_t^2(s)\,\varepsilon_s^2 ds + iz\int_{\mathbb{R}}\psi_t(s)\,\varepsilon_s^2 ds\right\}\right) \\ &= \log E\left(\exp\left\{-\frac{z^2}{2}\int_{\mathbb{R}}\int_{\mathbb{R}}\phi_t^2(s)\,h(s-r)\,dL_r ds + iz\int_{\mathbb{R}}\int_{\mathbb{R}}\psi_t(s)\,h(s-r)\,dL_r ds\right\}\right). \end{aligned} \tag{B.7}$$

The formula for the cumulants of $\Delta_t X$ in terms of the cumulants of L' that we obtain in this Section is a consequence of the Fubini Theorem [2, Theorem 3.1] applied to the double integrals in (B.7). In the next lines we check that ϕ, ψ, h, and L satisfy the conditions stated there.

Remark 23 *Since $h, q \geq 0$, Tonelli's Theorem implies that*

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\psi_t(s)\,h(s-r)\,dsdr = \int_{\mathbb{R}}\int_{\mathbb{R}}\psi_t(s)\,h(s-r)\,drds < \infty$$

and

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\phi_t^2(s)\,h(s-r)\,dsdr = \int_{\mathbb{R}}\int_{\mathbb{R}}\phi_t^2(s)\,h(s-r)\,drds < \infty.$$

The Fubini Theorem [2, Theorem 3.1] applies to centered Lévy processes. Therefore, for the moment, we will consider the process $Y_t = L_t^j - tE[L_1^j]$, where L^j is the pure jump part of L under the truncation function $\tau \equiv 0$ (i.e. the characteristic triplet of L^j under the truncation function $\tau \equiv 0$ is $(0, 0, \nu)$). Let $\varphi : \mathbb{R} \to \mathbb{R}$ be given by

$$\varphi(y) = \int_{\mathbb{R}_+}\left((yu)^2 1_{\{|yu|\leq 1\}} + (2|yu| - 1)\,1_{\{|yu|>1\}}\right)\nu(du),$$

and for all measurable functions $f : \mathbb{R} \to \mathbb{R}$ define

$$\|f\|_\varphi = \inf\left\{c > 0 : \int_{\mathbb{R}}\varphi\left(c^{-1}f(s), s\right)ds \leq 1\right\}.$$

Moreover, let $L^\varphi = L^\varphi(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ denote the Musielak-Orlicz space of measurable functions f with

$$\int_{\mathbb{R}}\int_{\mathbb{R}_+}|f(w)\,u|^2 \wedge |f(w)\,u|\,\nu(du)\,dw < \infty,$$

equipped with the Luxemburg norm $\|f\|_\varphi$ (see e.g. [15]). For $t > 0$, define

$$f_t^g(x,s) = \phi_t^2(x)\,h(x-s) \quad \text{and} \quad f_t^q(x,s) = \psi_t(x)\,h(x-s),$$

and

$$f_{x,t}^g = f_t^g(x,\cdot) \quad \text{and} \quad f_{x,t}^q = f_t^q(x,\cdot).$$

We want to change the order of integration in $I_1 = \int_{\mathbb{R}} \int_{\mathbb{R}} f_t^g(x,s)\,dL_s dx$ and $I_2 = \int_{\mathbb{R}} \int_{\mathbb{R}} f_t^q(x,s)\,dL_s dx$. Thanks to Remark 23, we do not need to worry about the drift term of Y and the work is done if we change the order of integration in $\mathcal{I}_1 = \int_{\mathbb{R}} \int_{\mathbb{R}} f_t^g(x,s)\,dY_s dx$ and $\mathcal{I}_2 = \int_{\mathbb{R}} \int_{\mathbb{R}} f_t^q(x,s)\,dY_s dx$. We apply the Fubini Theorem to $\mathcal{I}_1$ and $\mathcal{I}_2$. There are only two conditions to check in the Fubini Theorem [2, Theorem 3.1]: 1) $f_{x,t}^g$, $f_{x,t}^q \in L^\varphi$ for $x \in \mathbb{R}$; and, 2) $\int_{\mathbb{R}} \left\|f_{x,t}^g\right\|_\varphi dx < \infty$ and $\int_{\mathbb{R}} \left\|f_{x,t}^q\right\|_\varphi dx < \infty$. Since $h \in L^1(\mathbb{R})$, from (B.5) follows that $f_{x,t}^g$ and $f_{x,t}^q$ belong to the Musielak-Orlicz space L^φ, i.e. $f_{x,t}^g, f_{x,t}^q \in L^\varphi$ (condition 1). Besides,

$$\begin{aligned} E\left(\left|\int_{\mathbb{R}} f_{x,t}^g(s)\,dY_s\right|\right) &\leq 2\phi_t^2(x) \int_{\mathbb{R}} \int_{\mathbb{R}_+} h(x-s)\,uv(du)\,ds \\ &= 2\phi_t^2(x) \int_{\mathbb{R}} \int_{\mathbb{R}_+} h(w)\,uv(du)\,dw \\ &\equiv 2c\phi_t^2(x) < \infty, \end{aligned}$$

and, similarly,

$$\begin{aligned} E\left(\left|\int_{\mathbb{R}} f_{x,t}^q(s)\,dY_s\right|\right) &\leq 2\left|\psi_t(x)\right| \int_{\mathbb{R}} \int_{\mathbb{R}_+} h(w)\,uv(du)\,dw \\ &= 2c\left|\psi_t(x)\right| < \infty. \end{aligned}$$

Since $f_{x,t}^g$ and $f_{x,t}^q$ are in L^φ, Theorem 2.1 in [23] shows that

$$\left\|f_{x,t}^g\right\|_\varphi \leq 8E\left(\left|\int_{\mathbb{R}} f_{x,t}^g(s)\,dY_s\right|\right) \quad \text{and} \quad \left\|f_{x,t}^q\right\|_\varphi \leq 8E\left(\left|\int_{\mathbb{R}} f_{x,t}^q(s)\,dY_s\right|\right).$$

Therefore,

$$\left\|f_{x,t}^g\right\|_\varphi \leq 16c\phi_t^2(x) \quad \text{and} \quad \left\|f_{x,t}^q\right\|_\varphi \leq 16c\left|\psi_t(x)\right|.$$

Since $g \in L^2(\mathbb{R})$ and $q \in L^1(\mathbb{R})$, this implies that

$$\int_{\mathbb{R}} \left\|f_{x,t}^g\right\|_\varphi dx < \infty \quad \text{and} \quad \int_{\mathbb{R}} \left\|f_{x,t}^q\right\|_\varphi dx < \infty.$$

This proves the second condition in the stochastic Fubini Theorem. The Fubini Theorem implies that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \phi_t^2(s)\,h(s-r)\,dY_r ds = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_t^2(s)\,h(s-r)\,ds dY_r = \int_{-\infty}^t \int_r^t \phi_t^2(s)\,h(s-r)\,ds dY_r$$

and

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \psi_t(s)\,h(s-r)\,dY_r ds = \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_t(s)\,h(s-r)\,ds dY_r = \int_{-\infty}^t \int_r^t \psi_t(s)\,h(s-r)\,ds dY_r.$$

It follows that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \phi_t^2(s)\,h(s-r)\,dL_r ds = \int_{-\infty}^t \int_r^t \phi_t^2(s)\,h(s-r)\,ds dL_r$$

and

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \psi_t(s)\, h(s-r)\, dL_r ds = \int_{-\infty}^{t}\int_{r}^{t} \psi_t(s)\, h(s-r)\, ds dL_r.$$

Remark 24 *Fubini Theorem 3.1 in [2] also states that $f_t^g(\cdot,s), f_t^q(\cdot,s) \in L^1(\mathbb{R})$, for almost all $s \in \mathbb{R}$ (with respect to the Lebesgue measure); i.e.*

$$\int_{r}^{t} \psi_t(s)\, h(s-r)\, ds < \infty \quad \text{and} \quad \int_{r}^{t} \phi_t^2(s)\, h(s-r)\, ds < \infty,$$

for almost all $r \in \mathbb{R}$. However, in our case, this is also a consequence of Remark 23.

B.3.2 Cumulants of $\Delta_t X$ relative to the cumulants of L'

Now we can proceed to find a formula for the cumulants of $\Delta_t X$. For $(t,z) \in \mathbb{R}_+ \times \mathbb{R}$, let $H_{t,z} : (-\infty, t) \to \mathbb{C}$ be the function given by

$$H_{t,z}(r) = \int_{r}^{t} \left\{ iz\psi_t(s)\, h(s-r) - \frac{1}{2} z^2 \phi_t^2(s)\, h(s-r) \right\} ds, \tag{B.8}$$

when the integral exists, and 0 otherwise. Then, from (B.7),

$$\begin{aligned} C\{z \ddagger \Delta_t X\} &= \log E\left(\exp\left\{\int_{-\infty}^{t} H_{t,z}(r)\, dL_r\right\}\right) = \int_{-\infty}^{t} C\{H_{t,z}(r) \ddagger L'\}\, dr \\ &= \int_{-\infty}^{t} \sum_{m=1}^{\infty} \frac{\kappa_m(L')}{m!} H_{t,z}^m(r)\, dr, \end{aligned} \tag{B.9}$$

where $\kappa_m(X)$ denotes the cumulant of order m of the random variable X.

Proposition 25 *Let $t > 0$ and assume that all cumulants of L' exist. Define*

$$C_1(t,r) = \int_{r}^{t} \psi_t(s)\, h(s-r)\, ds \quad \text{and} \quad C_2(t,r) = \int_{r}^{t} \phi_t^2(s)\, h(s-r)\, ds, \tag{B.10}$$

when the integrals exist, and $C_1(t,r) = C_2(t,r) = 0$ otherwise. If $H_{t,z}(r) = izC_1(t,r) - \frac{1}{2}z^2 C_2(t,r)$ and

$$\sum_{m=1}^{\infty} \frac{\kappa_m(L')}{m!} \int_{-\infty}^{t} |H_{t,z}(r)|^m\, dr < \infty \tag{B.11}$$

for $z \in D \subseteq \mathbb{R}$, then

$$C\{z \ddagger \Delta_t X\} = \sum_{j=1}^{\infty} \frac{(iz)^j}{j!} \left(\int_{-\infty}^{t} j! \sum_{m=\lceil j/2 \rceil}^{j} \binom{m}{2m-j} C_1^{2m-j}(t,r)\, C_2^{j-m}(t,r)\, \frac{\kappa_m(L')}{2^{j-m} \cdot m!} dr \right), \tag{B.12}$$

for $z \in D$. In particular, if there exist $\eta > 0$ and $M(t) > 0$ such that

$$\max\{|C_1(t,r)|, |C_2(t,r)|\} \leq M(t)\, e^{\eta r}, \tag{B.13}$$

then (B.12) is true for $|z|$ sufficiently small.

Proof. Changing variables and using the regular Fubini Theorem, for $z \in D$, we can rewrite (B.9) as

$$\begin{aligned} C\{z \ddagger \Delta_t X\} &= \int_{-\infty}^{t} \sum_{m=1}^{\infty} \frac{\kappa_m(L')}{m!} \sum_{l=0}^{m} \binom{m}{l} (iz)^l C_1^l(t,r)(-1)^{m-l} \frac{1}{2^{m-l}} (z^2)^{m-l} C_2^{m-l}(t,r)\, dr \\ &= \int_{-\infty}^{t} \sum_{m=1}^{\infty} \sum_{l=0}^{m} \binom{m}{l} C_1^l(t,r) C_2^{m-l}(t,r) \frac{\kappa_m(L')}{2^{m-l} \cdot m!} (iz)^{2m-l} dr \\ &= \int_{-\infty}^{t} \sum_{m=1}^{\infty} \sum_{j=m}^{2m} \binom{m}{2m-j} C_1^{2m-j}(t,r) C_2^{j-m}(t,r) \frac{\kappa_m(L')}{2^{j-m} \cdot m!} (iz)^j dr \\ &= \int_{-\infty}^{t} \sum_{j=1}^{\infty} \sum_{m=\lceil j/2 \rceil}^{j} \binom{m}{2m-j} C_1^{2m-j}(t,r) C_2^{j-m}(t,r) \frac{\kappa_m(L')}{2^{j-m} \cdot m!} (iz)^j dr \\ &= \sum_{j=1}^{\infty} \frac{(iz)^j}{j!} \left(\int_{-\infty}^{t} j! \sum_{m=\lceil j/2 \rceil}^{j} \binom{m}{2m-j} C_1^{2m-j}(t,r) C_2^{j-m}(t,r) \frac{\kappa_m(L')}{2^{j-m} \cdot m!} dr \right). \end{aligned}$$

This proves the first part of the proposition. Now, assume that the functions C_1 and C_2 (B.10) satisfy (B.13). Then, there is a constant $K > 0$ such that

$$\begin{aligned} \sum_{m=1}^{\infty} \frac{\kappa_m(L')}{m!} \int_{-\infty}^{t} |H_{t,z}(r)|^m dr &\leq \sum_{m=1}^{\infty} \frac{\kappa_m(L')}{m!} |z|^m \int_{-\infty}^{t} \left(|C_1(t,r)| + \frac{1}{2} z |C_2(t,r)| \right)^m dr \\ &\leq \sum_{m=1}^{\infty} K \cdot |z|^m \int_{-\infty}^{t} \left(|C_1(t,r)| + \frac{1}{2} |z| |C_2(t,r)| \right)^m dr \\ &\leq K \sum_{m=1}^{\infty} |z|^m \int_{-\infty}^{t} \left(2^{m-1} |C_1(t,r)|^m + \frac{1}{2} |z|^m |C_2(t,r)|^m \right) dr \\ &\leq K \sum_{m=1}^{\infty} |z|^m \left(2^{m-1} + \frac{1}{2} |z|^m \right) M(t)^m \int_{-\infty}^{t} e^{m\eta r} dr \\ &\leq K \sum_{m=1}^{\infty} |z|^m \left(2^{m-1} + \frac{1}{2} |z|^m \right) M(t)^m \frac{e^{m\eta t}}{m\eta} \\ &\leq \frac{K}{2\eta} \sum_{m=1}^{\infty} \frac{(2|z| M(t) e^{\eta t})^m}{m} + \frac{K}{2\eta} \sum_{m=1}^{\infty} \frac{(z^2 M(t)^m e^{\eta t})^m}{m} < \infty, \end{aligned}$$

for small $|z|$. This finishes the proof. ∎

As a consequence of Proposition 25, we have that the cumulants of $\Delta_t X$ can be expressed in terms of the cumulants of L'.

Corollary 26 *Assume that all cumulants of L' exist. Let $t > 0$. If the conditions (B.11) or (B.13) are satisfied,* *then we have the relation*

$$\kappa_j(\Delta_t X) = j! \sum_{m=\lceil j/2 \rceil}^{j} \binom{m}{2m-j} \frac{\kappa_m(L')}{2^{j-m} \cdot m!} \int_{-\infty}^{t} C_1^{2m-j}(t,r)\, C_2^{j-m}(t,r)\, dr \quad j \in \mathbb{N}. \tag{B.14}$$

Remark 27 *If we assume that $q \equiv 0$ in (B.2), then $C_1(t,r) \equiv 0$. Under the condition (B.11), this implies that*

$$\kappa_j(\Delta_t X) = \begin{cases} j! \frac{\kappa_{j/2}(L')}{2^{j/2} \cdot \left(\frac{j}{2}\right)!} \int_{-\infty}^{t} C_2^{j/2}(t,r)\, dr & \text{if } j \text{ is even} \\ 0 & \text{if } j \text{ is odd} \end{cases}, \tag{B.15}$$

for $j \in \mathbb{N}$.

In the derivation of the formulae (B.12) and (B.14), we have assumed that all cumulants of the Lévy seed L' exist. This is a very strong and restrictive condition. More generally, for the formulae (B.12) and (B.14) to hold it is only necessary to have that $E\left\{(L')^j\right\} < \infty$ for some $j \in \mathbb{N}$. In this way, the cumulant function of L' can be written as

$$C\{z \ddagger L'\} = \sum_{m=1}^{j} \frac{\kappa_m(L')}{m!} (iz)^m + o\left(|z|^j\right) \quad \text{as } z \to 0. \tag{B.16}$$

Relation (B.16) can be used to reproduce the same arguments that led to the formulae (B.12) and (B.14). In this case, condition (B.11) can be replaced by

$$\sum_{m=1}^{j} \frac{\kappa_m(L')}{m!} \int_{-\infty}^{t} |H_{t,z}(r)|^m\, dr < \infty.$$

B.4 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as a LSS process: examples

In this Section we study the cumulants of $\Delta_t X$ for four examples of X given by (B.2) and assuming that ε^2 is a LSS process given by (B.3). We assume that all cumulants of the Lévy seed L' exist. For the first example, $q \equiv 0$, $g(x) = e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, for $\lambda > 0$, and ε^2 is an Ornstein-Uhlenbeck process. This specification of X permits to obtain closed expressions for the cumulants of $\Delta_t X$ in terms of hypergeometric functions. In the second example, $q \equiv 0$, g is proportional to a gamma density and ε^2 is an Ornstein-Uhlenbeck process. For the third example, g and q are proportional to a gamma density and, again, ε^2 is an Ornstein-Uhlenbeck. In the last example, g, q and the kernel h of ε^2 are proportional to a gamma density. In this last case, the marginals of (B.2) are generalized hyperbolic distributions. The last three examples are relevant in the context of turbulence modelling.

B.4.1 Example 1: $g(x) = e^{\lambda x}1_{\mathbb{R}_+}(x)$, Ornstein-Uhlenbeck process ε^2

For $\lambda, \rho > 0$ and $2\lambda \neq \rho$, consider the model (B.2), (B.3) with $g(x) = e^{-\lambda x}1_{\mathbb{R}_+}(x)$, $q \equiv 0$, $h(x) = e^{-\rho x}1_{\mathbb{R}_+}(x)$ and L a subordinator. Then, for $t > 0$,

$$\phi_t^2(s) = e^{2\lambda s}\left(1 - e^{-\lambda t}\right)^2 1_{\mathbb{R}_-}(s) + e^{-2\lambda(t-s)}1_{(0,t)}(s)$$

and

$$\begin{aligned} C_2(t,r) &= \frac{e^{-2\lambda t}}{2\lambda - \rho} e^{\rho r}\left(e^{(2\lambda-\rho)t} - e^{(2\lambda-\rho)r}\right) 1_{\mathbb{R}_+}(r) \\ &\quad + \frac{e^{\rho r}}{2\lambda - \rho}\left(\left(1 - e^{-\lambda t}\right)^2\left(1 - e^{(2\lambda-\rho)r}\right) + e^{-2\lambda t}\left(e^{(2\lambda-\rho)t} - 1\right)\right) 1_{\mathbb{R}_-}(r), \end{aligned}$$

where ϕ and C_2 are the functions defined in (B.6) and (B.10), respectively. For $m \in \mathbb{N}$, we define

$$\begin{aligned} C_2^m(t,r) &= \frac{e^{-2m\lambda t}e^{m\rho r}}{(2\lambda-\rho)^m}\left(e^{(2\lambda-\rho)t} - e^{(2\lambda-\rho)r}\right)^m 1_{\mathbb{R}_+}(r) \\ &\quad + \frac{e^{m\rho r}}{(2\lambda-\rho)^m}\left(\left(1 - e^{-\lambda t}\right)^2\left(1 - e^{(2\lambda-\rho)r}\right) + e^{-2\lambda t}\left(e^{(2\lambda-\rho)t} - 1\right)\right)^m 1_{\mathbb{R}_-}(r), \\ &\equiv \frac{1}{(2\lambda-\rho)^m}\left(D_1(m,t,r) + D_2(m,t,r)\right). \end{aligned}$$

We expand D_1 and D_2 as

$$\begin{aligned} D_1(m,t,r) &= e^{-2m\lambda t}e^{m\rho r}\left(e^{(2\lambda-\rho)t} - e^{(2\lambda-\rho)r}\right)^m 1_{\mathbb{R}_+}(r) \\ &= e^{-2m\lambda t}\sum_{i=0}^{m}\binom{m}{i}(-1)^i e^{((2\lambda-\rho)i+m\rho)r}e^{(2\lambda-\rho)t(m-i)}1_{\mathbb{R}_+}(r) \end{aligned}$$

and

$$\begin{aligned} D_2(m,t,r) &= e^{m\rho r}\left(\left(1 - e^{-\lambda t}\right)^2\left(1 - e^{(2\lambda-\rho)r}\right) + e^{-2\lambda t}\left(e^{(2\lambda-\rho)t} - 1\right)\right)^m 1_{\mathbb{R}_-}(r) \\ &= e^{m\rho r}\sum_{i=0}^{m}\binom{m}{i}\left(1 - e^{-\lambda t}\right)^{2i}\left(1 - e^{(2\lambda-\rho)r}\right)^i\left(\frac{e^{(2\lambda-\rho)t} - 1}{e^{-2\lambda t}}\right)^{m-i} 1_{\mathbb{R}_-}(r) \\ &= \sum_{i=0}^{m}\binom{m}{i}\left(1 - e^{-\lambda t}\right)^{2i}\left(\frac{e^{(2\lambda-\rho)t} - 1}{e^{-2\lambda t}}\right)^{m-i}\sum_{n=0}^{i}\binom{i}{n}(-1)^n e^{((2\lambda-\rho)n+m\rho)r}1_{\mathbb{R}_-}(r). \end{aligned}$$

We get

$$\begin{aligned} I_{D_1}(m,t) &\equiv \int_{-\infty}^{t} D_1(m,t,r)\,dr \\ &= e^{-2m\lambda t}\sum_{i=0}^{m}\binom{m}{i}(-1)^i e^{(2\lambda-\rho)t(m-i)}\int_0^t e^{((2\lambda-\rho)i+m\rho)r} \end{aligned}$$

$$= \quad e^{-2m\lambda t} \sum_{i=0}^{m} \binom{m}{i} (-1)^i e^{(2\lambda-\rho)t(m-i)} \frac{e^{((2\lambda-\rho)i+m\rho)t} - 1}{(2\lambda-\rho)\, i + m\rho}$$

and

$$\begin{aligned}
I_{D_2}(m,t) &\equiv \int_{-\infty}^{t} D_2(m,t,r)\, dr \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(1 - e^{-2\lambda t}\right)^{2i} \left(\frac{e^{(2\lambda-\rho)t} - 1}{e^{-2\lambda t}}\right)^{m-i} \sum_{n=0}^{i} \binom{i}{n} \frac{(-1)^n}{(2\lambda-\rho)\, n + m\rho} \\
&= \sum_{i=0}^{m} \sum_{n=0}^{i} \binom{m}{i} \binom{i}{n} \frac{(-1)^n \left(1 - e^{-\lambda t}\right)^{2i} e^{-2\lambda t(m-i)} \left(e^{(2\lambda-\rho)t} - 1\right)^{(m-i)}}{(2\lambda-\rho)\, n + m\rho} \\
&= \sum_{n=0}^{m} \sum_{i=n}^{m} \binom{m}{i} \binom{i}{n} \frac{(-1)^n \left(1 - e^{-\lambda t}\right)^{2i} e^{-2\lambda t(m-i)} \left(e^{(2\lambda-\rho)t} - 1\right)^{(m-i)}}{(2\lambda-\rho)\, n + m\rho} \\
&= \sum_{n=0}^{m} \sum_{j=0}^{m-n} \binom{m}{j+n} \binom{j+n}{n} \frac{(-1)^n \left[\left(1 - e^{-\lambda t}\right)^2\right]^{j+n} \left[e^{-2\lambda t}\left(e^{(2\lambda-\rho)t} - 1\right)\right]^{m-j-n}}{(2\lambda-\rho)\, n + m\rho}.
\end{aligned}$$

The expressions for I_{D_1} and I_{D_2} can be rewritten in terms of hypergeometric functions. For this, we need the next lemma.

Lemma 28 *Let $a, b \in \mathbb{R}$, $c, d > 0$ and $m \in \mathbb{N}$. Then,*

$$\sum_{j=0}^{m-n} \binom{m}{j+n} \binom{j+n}{n} a^{j+n} b^{m-j-n} = \binom{m}{n} a^n (a+b)^{m-n}, \quad n \in \mathbb{N} \cup \{0\}$$

and

$$\sum_{n=0}^{m} \binom{m}{n} \frac{a^n b^{m-n}}{cn + dm} = \frac{b^m}{dm} {}_2F_1\left(-m, \frac{dm}{c}, 1 + \frac{dm}{c}, -\frac{a}{b}\right), \quad \text{if } |a/b| < 1.$$

In particular,

$$\sum_{n=0}^{m} \sum_{j=0}^{m-n} \binom{m}{j+n} \binom{j+n}{n} \frac{(-1)^n a^{j+n} b^{m-j-n}}{cn + dm} = \frac{b^m (a+b)^m}{dm} {}_2F_1\left(-m, \frac{dm}{c}, 1 + \frac{dm}{c}, \frac{a}{a+b}\right).$$

Lemma 28 implies that

$$\begin{aligned}
I_{D_1}(m,t) &= e^{-2m\lambda t} \left\{ \sum_{i=0}^{m} \binom{m}{i} (-1)^i e^{(2\lambda-\rho)t(m-i)} \frac{e^{((2\lambda-\rho)i+m\rho)t}}{(2\lambda-\rho)\, i + m\rho} \right. \\
&\qquad \left. - \sum_{i=0}^{m} \binom{m}{i} \frac{(-1)^i e^{(2\lambda-\rho)t(m-i)}}{(2\lambda-\rho)\, i + m\rho} \right\} \\
&= \frac{1}{\rho m} {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1 + \frac{\rho m}{2\lambda-\rho}, 1\right)
\end{aligned}$$

$$-\frac{e^{-\rho tm}}{\rho m} {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, e^{-(2\lambda-\rho)t}\right)$$

$$=\frac{1}{\rho m}\frac{\Gamma(1+m)\Gamma\left(1+\frac{\rho m}{2\lambda-\rho}\right)}{\Gamma\left(1+m+\frac{\rho m}{2\lambda-\rho}\right)}-\frac{e^{-\rho tm}}{\rho m} {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, e^{-(2\lambda-\rho)t}\right)$$

and

$$I_{D_2}(m,t)=\frac{d_1^m(t)}{\rho m} {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, \frac{\left(1-e^{-\lambda t}\right)^2}{d_1(t)}\right)$$

where

$$d_1(t) \equiv \left(\left(1-e^{-\lambda t}\right)^2+e^{-2\lambda t}\left(e^{(2\lambda-\rho)t}-1\right)\right).$$

Therefore,

$$\begin{aligned}\int_{-\infty}^{t} C_2^m(t,r)\,dr &= \frac{1}{(2\lambda-\rho)^m}\int_{-\infty}^{t}\left(D_1(m,t,r)+D_2(m,t,r)\right)dr\\ &= \frac{1}{(2\lambda-\rho)^m \rho m}\left\{\frac{\Gamma(1+m)\Gamma\left(1+\frac{\rho m}{2\lambda-\rho}\right)}{\Gamma\left(1+m+\frac{\rho m}{2\lambda-\rho}\right)}\right.\\ &\quad -e^{-\rho tm} {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, e^{-(2\lambda-\rho)t}\right)\\ &\quad \left.+d_1^m(t)\, {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, \frac{\left(1-e^{-\lambda t}\right)^2}{d_1(t)}\right)\right\}.\end{aligned}$$

The function C_2 satisfies the condition (B.13) in Proposition 25. According to (B.15), assuming that L' has finite moments of all orders, we have that

$$\begin{aligned}\kappa_{2m}(\Delta_t X) = \frac{\kappa_m(L')(2m)!}{2^m m!(2\lambda-\rho)^m \rho m}&\left\{d_1^m(t)\, {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, \frac{\left(1-e^{-\lambda t}\right)^2}{d_1(t)}\right)\right.\\ &\left.+\frac{\Gamma(1+m)\Gamma\left(1+\frac{\rho m}{2\lambda-\rho}\right)}{\Gamma\left(1+m+\frac{\rho m}{2\lambda-\rho}\right)}-e^{-\rho tm} {}_2F_1\left(-m, \frac{\rho m}{2\lambda-\rho}, 1+\frac{\rho m}{2\lambda-\rho}, e^{-(2\lambda-\rho)t}\right)\right\}\end{aligned}$$

and $\kappa_{2m-1}(\Delta_t X)=0$, for $m \in \mathbb{N}$.

In the following, for $m \in \mathbb{N}$ and $t \geq 0$, $\overline{\kappa}_m(\Delta_t X)$ denotes the normalized cumulant $\overline{\kappa}_m(\Delta_t X) \equiv \kappa_m(\Delta_t X)/\kappa_m(L')$, where L' is the Lévy seed of the process driving ε^2. We reserve the term *normalized cumulant* only for the cumulants $\overline{\kappa}$ and *standardized cumulant* for the cumulants $\overline{\kappa}_m(\Delta_t X)/\overline{\kappa}_2^{m/2}(\Delta_t X)$ or $\kappa_m(\Delta_t X)/\kappa_2^{m/2}(\Delta_t X)$.

We have that

$$\overline{\kappa}_2(\Delta_t X)=\frac{1-e^{-\lambda t}}{\lambda\rho}$$

and

$$\bar{\kappa}_4(\Delta_t X) = \frac{3e^{-(3\lambda+\rho)t}\left(e^{\lambda t}-2\right)\left(4e^{2\lambda t}\lambda - 2e^{\rho t}\rho - 2(2\lambda-\rho)e^{(\lambda+\rho)t} + e^{(2\lambda+\rho)t}(2\lambda-\rho)\right)}{4\lambda\rho\left(4\lambda^2-\rho^2\right)} + \frac{6\cdot\Gamma\left(\frac{2\rho}{2\lambda-\rho}\right)}{(2\lambda+\rho)^3\,\Gamma\left(\frac{6\lambda-\rho}{2\lambda-\rho}\right)}.$$

It is easy to check that

$$\bar{\kappa}_4(\Delta_t X) \underset{t\to\infty}{\longrightarrow} \frac{3}{4\lambda\rho(2\lambda+\rho)} + \frac{6\cdot\Gamma\left(\frac{2\rho}{2\lambda-\rho}\right)}{(2\lambda+\rho)^3\,\Gamma\left(\frac{6\lambda-\rho}{2\lambda-\rho}\right)}.$$

The kurtosis $\bar{\kappa}_4(\Delta_t X)/\bar{\kappa}_2^2(\Delta_t X)$ is decreasing as a function of t, and decreasing as a function of λ and ρ. This suggests that the non-normality of $\Delta_\infty X$ escalates as the value of λ and ρ increase.

Even in this simple case, despite it was possible to find closed expressions for the cumulants of $\Delta_t X$, it is not feasible to determine when the distribution of the increments of X belongs to a known class. This exemplifies the complex dynamics exhibited by the increments of a BSS process (B.2).

B.4.2 Example 2: $g(x) = x^a e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, Ornstein-Uhlenbeck process ε^2

For $a > -1/2$, $\lambda, \rho > 0$ and $2\lambda \neq \rho$, consider the model (B.2), (B.3) with $g(x) = x^a e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, $q \equiv 0$, $h(x) = e^{-\rho x} 1_{\mathbb{R}_+}(x)$ and L a subordinator. BSS processes of this type have interesting mathematical properties and they are relevant as models for the temporal turbulent velocity field. Among the most remarkable mathematical properties for this class of BSS processes, we find multipower variation type limits and that, in general, X is not a semimartingale (see, e.g., [14]). The parameter a controls the smoothness of X and determines when X is a semimartingale: X is a semimartingale if and only if $a > 1/2$.

From the modelling perspective, it has been shown in [27] that the BSS process (B.2), under the assumptions of Example B.4.2 with $a = -1/6$, reproduces the so-called 2/3rd-Kolmogorov law, according to which turbulent velocity increments obey

$$E\left\{(X_t - X_0)^2\right\} \propto t^{2/3}, \tag{B.17}$$

for a certain range of scales t.

For $t > 0$,

$$\phi_t^2(s) = e^{2\lambda s}\left((t-s)^a e^{-\lambda t} - (-s)^a\right)^2 1_{\mathbb{R}_-}(s) + e^{2\lambda s}(t-s)^{2a} e^{-2\lambda t} 1_{(0,t)}(s) \tag{B.18}$$

and

$$\begin{aligned} C_2(t,r) = {} & \frac{e^{\rho(r-t)}}{(2\lambda-\rho)^{2a+1}}\left(\Gamma(1+2a) - \Gamma(1+2a,(t-r)(2\lambda-\rho))\right) 1_{\mathbb{R}_+}(r) \\ & + \frac{e^{\rho(r-t)}}{(2\lambda-\rho)^{2a+1}}\left\{-2e^{-\lambda t}e^{\rho t}R(r) + \left(1+e^{\rho t}\right)\Gamma(1+2a)\right. \\ & \left. - e^{\rho t}\Gamma(1+2a,-r(2\lambda-\rho)) - \Gamma(1+2a,(t-r)(2\lambda-\rho))\right\} 1_{\mathbb{R}_-}(r), \end{aligned} \tag{B.19}$$

where the function $R : \mathbb{R}^- \to \mathbb{R}$ is defined as

$$R(r) = \int_r^0 (t-s)^\alpha (-s)^\alpha e^{(2\lambda-\rho)s} ds.$$

It is difficult to get a general expression for the cumulant of arbitrary order. For $m = 2$, we have that

$$\int_{-\infty}^t C_2^{m/2}(t,r)\, dr = I_1 + I_2 + I_3 + I_4,$$

where

$$\begin{aligned}
I_1 &= \int_0^t \frac{e^{\rho(r-t)}}{(2\lambda-\rho)^{2\alpha+1}} \left(\Gamma(1+2\alpha) - \Gamma(1+2\alpha,(t-r)(2\lambda-\rho))\right) dr \\
&= \frac{1}{2\rho}\Big(4^{-\alpha}\lambda^{-2\alpha-1}\left(\Gamma(1+2\alpha)-\Gamma(1+2\alpha,2t\lambda)\right) \\
&\quad +2e^{-\rho t}(2\lambda-\rho)^{-2\alpha-1}\left(\Gamma(1+2\alpha)-\Gamma(1+2\alpha,(2\lambda-\rho)t)\right)\Big) \\
I_2 &= -\int_{-\infty}^0 \frac{e^{\rho(r-t)}}{(2\lambda-\rho)^{2\alpha+1}} 2e^{-\lambda t}e^{\rho t} R(r)\, dr = \frac{2^{\frac{1}{2}-\alpha}\sqrt{\pi}t^{\alpha+\frac{1}{2}}\lambda^{-\frac{1}{2}-\alpha}\csc(\pi\alpha)}{\rho\Gamma(-\alpha)(2\lambda-\rho)^{2\alpha+1}} K_{\frac{1}{2}+\alpha}(t\lambda) \\
I_3 &= \int_{-\infty}^0 \frac{e^{\rho(r-t)}}{(2\lambda-\rho)^{2\alpha+1}} e^{\rho t}\left(\Gamma(1+2\alpha)-\Gamma(1+2\alpha,-r(2\lambda-\rho))\right) dr = -\frac{4^{-\alpha}\alpha}{\rho}\lambda^{-1-2\alpha}\Gamma(2\alpha) \\
I_4 &= \int_{-\infty}^0 \frac{e^{\rho(r-t)}}{(2\lambda-\rho)^{2\alpha+1}} \left(\Gamma(1+2\alpha)-\Gamma(1+2\alpha,(t-r)(2\lambda-\rho))\right) dr \\
&= \frac{2^{-1-2\alpha}\lambda^{-1-2\alpha}}{\rho}\Gamma(1+2\alpha,2\lambda t) + \frac{e^{-\rho t}(2\lambda-\rho)^{-1-2\alpha}}{\rho}\left(\Gamma(1+2\alpha)-\Gamma(1+2\alpha,t(2\lambda-\rho))\right).
\end{aligned}$$

It seems not possible to express the sum of the I_j's in a simple nice expression. However, the results presented here permit to numerically compute the cumulants.

B.4.3 Example 3: $g(x) = x^\alpha e^{-\lambda x} 1_{\mathbb{R}_+}(x)$,$q(x) = x^\beta e^{-\mu x} 1_{\mathbb{R}_+}(x)$, Ornstein-Uhlenbeck process ε^2

For $\min\{\alpha, \beta/2\} > -1/2$ and $\lambda, \rho, \mu > 0$ with $2\lambda \neq \rho$ and $\mu \neq \rho$, consider the model (B.2), (B.3) with $g(x) = x^\alpha e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, $q(x) = x^\beta e^{-\mu x} 1_{\mathbb{R}_+}(x)$, $h(x) = e^{-\rho x} 1_{\mathbb{R}_+}(x)$ and L a subordinator. This class of BSS processes is a generalization of the class described in Section B.4.2. The main difference between Example B.4.2 and Example B.4.3 is that the increments display non-vanishing odd cumulants (which is of particular interest for turbulence modelling).

For $t > 0$, we have

$$\psi_t(s) = e^{\mu s}\left((t-s)^\beta e^{-\mu t} - (-s)^\beta\right) 1_{\mathbb{R}_-}(s) + e^{-\mu(t-s)}(t-s)^\beta 1_{(0,t)}(s)$$

which yields

$$C_1(t,r) \quad = \quad \frac{e^{-\rho(t-r)}}{(\mu-\rho)^{1+\beta}}\left(\Gamma(1+\beta) - \Gamma(1+\beta,-(t-r)(\mu-\rho))\right) 1_{\mathbb{R}_+}(r)$$

$$
-\frac{e^{-\rho(t-r)}}{(\mu-\rho)^{1+\beta}}\left\{\left(e^{\rho t}-1\right)\Gamma(1+\beta)-\Gamma(1+\beta,(t-r)(\mu-\rho))\right.
$$
$$
\left.-e^{\rho t}\Gamma(1+\beta,-r(\mu-\rho))\right\}1_{\mathbb{R}_-}(r).
$$

The functions ϕ_t and C_2 are given as in (B.18) and (B.19), respectively. The functions C_1 and C_2 satisfy the condition (B.13) in Proposition 25. Therefore, assuming that L' has finite moments of all orders, the cumulants of $\Delta_t X$ satisfy (B.14) and the cumulants of $\Delta_t X$ are given by the formula (B.14).

To our knowledge, it does not seem to be possible to express $H_{t,z}$ in terms of simple functions. However, (B.14) can be evaluated numerically.

B.4.4 Example 4: generalized hyperbolic marginals

For $c, \gamma \in \mathbb{R}, \bar{\lambda} > 0$ and $-1/2 < a < 0$, consider the model (B.2), (B.3) with

$$
g(x) = c\frac{\bar{\lambda}^{a+1/2}}{\Gamma(2a+1)^{1/2}}x^a e^{-\frac{\bar{\lambda}}{2}x}1_{\mathbb{R}_+}(x),
$$

$$
q(x) = \gamma\frac{\bar{\lambda}^{2a+1}}{\Gamma(2a+1)}x^{2a}e^{-\bar{\lambda}x}1_{\mathbb{R}_+}(x),
$$

$$
h(x) = \frac{\bar{\lambda}^{-2a-1}}{\Gamma(-2a)}x^{-2a-1}e^{-\bar{\lambda}x}1_{\mathbb{R}_+}(x),
$$

and L a subordinator. Choosing L' such that the OU process

$$
z_t = \int_{-\infty}^{t} e^{-\bar{\lambda}(t-u)}dL_u
$$

has the generalized inverse Gaussian law GIG(λ, χ, θ), the law of (B.2) is generalized hyperbolic GH$(\lambda, \chi, \theta, 0, c^2, \gamma)$ (see [3], Section 5.4). The density of the GIG(λ, χ, θ) distribution is given by

$$
f_{\mathrm{GIG}(\lambda,\chi,\theta)}(x) = \left(\frac{\theta}{\lambda}\right)^{\lambda/2}\frac{x^{\lambda-1}}{2K_\lambda\left(\sqrt{\chi\theta}\right)}\exp\left\{-\frac{1}{2}\left(\frac{\chi}{x}+\theta x\right)\right\}, \quad x \in \mathbb{R}
$$

where K_λ denotes the modified Bessel function of the third kind and the parameters (λ, χ, θ) have to satisfy one of the following three restrictions

$$
\chi > 0, \theta \geq 0, \lambda < 0 \quad \text{or} \quad \chi > 0, \theta > 0, \lambda = 0 \quad \text{or} \quad \chi \geq 0, \theta > 0, \lambda > 0.
$$

The density of the GH$(\lambda, \chi, \theta, \mu, \Sigma, \gamma)$ law is

$$
f_{\mathrm{GH}(\lambda,\chi,\theta,\mu,\Sigma,\gamma)}(x) \quad = \quad \frac{\left(\sqrt{\theta/\chi}\right)^{\lambda}\left(\theta+\Sigma^{-1}\gamma^2\right)^{\frac{1}{2}-\gamma}}{(2\pi)^{\frac{1}{2}}\Sigma^{1/2}K_\lambda\left(\sqrt{\theta\gamma}\right)}
$$

$$\times \frac{K_{\lambda-\frac{1}{2}}\left(\sqrt{\left(\chi+\Sigma^{-1}(x-\mu)^2\right)\left(\theta+\Sigma^{-1}\gamma^2\right)}\right)e^{\Sigma^{-1}\gamma(x-\mu)}}{\sqrt{\left(\chi+\Sigma^{-1}(x-\mu)^2\right)\left(\theta+\Sigma^{-1}\gamma^2\right)}}$$

for $(\lambda, \chi, \theta, \mu, \Sigma, \gamma) \in \mathbb{R} \times \mathbb{R}_+^2 \times \mathbb{R}^3$ and $x \in \mathbb{R}$.

The above parametrization for the GH law is the so-called $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$-parametrization (in order to avoid confusion with the increment functions (B.6), we replaced ψ by θ). We refer to [12] for more information about the different parameterizations of the GH distribution and their generalizations to the $\mathbb{R}^d$ case.

Notice that the distribution of X does not depend on the parameters $(\overline{\lambda}, a)$. However, the law of $\Delta_t X$ depends of these parameter as can be seen in the expression of C_i, $i = 1, 2$.

The GH class is a very rich and flexible family of distributions that nests several other distributions and has found applications in different areas. For more information about this distribution and its applications, we refer, for instance, to [1, 17, 16] and the papers cited in [12].

For the present example, for $t > 0$,

$$\psi_t(s) = \gamma \frac{\overline{\lambda}^{2a+1}}{\Gamma(2a+1)}\left\{e^{\overline{\lambda}s}\left((t-s)^{2a}e^{-\overline{\lambda}t} - (-s)^{2a}\right)1_{\mathbb{R}_-}(s) + e^{\overline{\lambda}s}(t-s)^{2a}e^{-\overline{\lambda}t}1_{(0,t)}(s)\right\}$$

and

$$\phi_t^2(s) = \frac{c^2\overline{\lambda}^{2a+1}}{\Gamma(2a+1)}\left\{e^{\overline{\lambda}s}\left((t-s)^{a}e^{-\overline{\lambda}t/2} - (-s)^{a}\right)^2 1_{\mathbb{R}_-}(s) + e^{\overline{\lambda}s}(t-s)^{2a}e^{-\overline{\lambda}t}1_{(0,t)}(s)\right\}.$$

Consequently,

$$\begin{aligned}C_1(t,r) = {} & \gamma e^{-\overline{\lambda}(t-r)}1_{(0,t)}(r) + \gamma e^{\overline{\lambda}r}\left\{-1 + \frac{(-1)^{2a+1}\sin(2a\pi)e^{-\overline{\lambda}t}}{\pi}\mathrm{Beta}\left(\frac{r}{t}, -2a, 0\right)\right.\\ & \left. - \frac{\sin(2a\pi)t^{2a+1}e^{-\overline{\lambda}t}(-r)^{-2a-1}}{\pi(2a+1)}{}_2F_1\left(1, 2a+1, 2a+2, \frac{t}{r}\right)\right\}1_{\mathbb{R}_-}(r)\end{aligned}$$

and

$$\begin{aligned}C_2(t,r) = {} & c^2 e^{-\overline{\lambda}(t-r)}1_{(0,t)}(r) + c^2 e^{\overline{\lambda}r}\left\{1 + \frac{(-1)^{2a+1}\sin(2a\pi)e^{-\overline{\lambda}t}}{\pi}\mathrm{Beta}\left(\frac{r}{t}, -2a, 0\right)\right.\\ & - \frac{2\Gamma(a+1)t^a e^{-\overline{\lambda}t/2}}{\Gamma(2a+1)\Gamma(1-a)}(-r)^{-a}{}_2F_1\left(-a, a+1, 1-a, \frac{r}{t}\right)\\ & \left. - \frac{\sin(2a\pi)t^{2a+1}e^{-\overline{\lambda}t}(-r)^{-2a-1}}{\pi(2a+1)}{}_2F_1\left(1, 2a+1, 2a+2, \frac{t}{r}\right)\right\}1_{\mathbb{R}_-}(r),\end{aligned}$$

where $\mathrm{Beta}(\cdot,\cdot,\cdot) : (-\infty, 0) \times (-\infty, 0) \times (-\infty, 0] \to \mathbb{C}$ denotes to the incomplete Beta function

$$\mathrm{Beta}(x, a, b) = -\int_x^0 w^{a-1}(1-w)^{b-1}\,dw.$$

In general, Beta produces complex values; however, $(-1)^{-a+1}$ Beta $(x, a, 0)$ is always real.

The functions C_1 and C_2 do not satisfy the condition (B.13) in Proposition 25 but they satisfy a similar inequality. Since $-1/2 < \alpha < 0$, the functions $(-1)^{2\alpha+1}$ Beta $\left(\frac{r}{t}, -2\alpha, 0\right)$, ${}_2F_1\left(1, 2\alpha+1, 2\alpha+2, \frac{t}{r}\right)$ and ${}_2F_1\left(-\alpha, \alpha+1, 1-\alpha, \frac{r}{t}\right)$ are bounded. Therefore, there exists $M(\alpha, t) > 0$ such that

$$\max\left\{\left|C_1(t,r)\right|, \left|C_2(t,r)\right|\right\} \leq M(\alpha, t)\, e^{\bar{\lambda} r}\left(|r|^{-2\alpha-1} \vee |r|^{-\alpha} \vee 1\right).$$

Similarly to condition (B.13), since $-1/2 < \alpha < 0$, the previous inequality implies that $H_{t,z}$ satisfies condition (B.11). Thus, when $E\left\{(L')^n\right\} < \infty$ for all $n \in \mathbb{N}$, the cumulants of $\Delta_t X$ are given by (B.14).

Formula (B.14) implies

$$\begin{aligned}
\kappa_3(\Delta_t X) &= \gamma^4 c^2 \kappa_2(L')\,\mathfrak{K}_{3,1}\left(t, \alpha, \bar{\lambda}\right) + \gamma^3 \kappa_3(L')\,\mathfrak{K}_{3,2}\left(t, \alpha, \bar{\lambda}\right),\\
\kappa_4(\Delta_t X) &= c^4 \kappa_2(L')\,\mathfrak{K}_{4,1}\left(t, \alpha, \bar{\lambda}\right) + \gamma^4 c^2 \kappa_3(L')\,\mathfrak{K}_{4,2}\left(t, \alpha, \bar{\lambda}\right) + \gamma^4 \kappa_4(L')\,\mathfrak{K}_{4,3}\left(t, \alpha, \bar{\lambda}\right),
\end{aligned}$$

for some functions $\mathfrak{K}_{j,i}$, $j = 3, 4$, $i = \lceil j/2 \rceil, \ldots, j$. The parameter γ determines the magnitude and the sign of the skewness $\kappa_3(\Delta_t X)/\kappa_2^{3/2}(\Delta_t X)$, and the kurtosis $\kappa_4(\Delta_t X)/\kappa_2^2(\Delta_t X)$ is controlled by c and γ.

B.5 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as an exponential ambit process: theory

This Section provides a formula for the n-th moment of the increments of a BSS process of the type (B.2) assuming that ε^2 is an exponential ambit process. In order to have a self-contained presentation, Subsection 5.1 introduces some basic preliminaries about ambit stochastics. Subsection 5.2 presents the formula we have derived.

B.5.1 Ambit stochastics: a short review

Ambit processes were introduced in [6] as a framework for tempo-spatial modeling. These processes are defined in terms of integrals with respect to a Lévy basis. In the present paper, we restrict our attention to those ambit processes defined as the stochastic integral of a deterministic function with respect to a homogeneous Lévy basis defined in $\mathbb{R}^2$.

We denote by $\mathcal{B}_b(\mathbb{R}^2)$ the set of bounded Borel subsets of $\mathbb{R}^2$. A Lévy basis Λ on $\mathbb{R}^2$ is an infinitely divisible, independently scattered random measure on $\mathbb{R}^2$, i.e. $(\Lambda(A))_{A\in\mathcal{B}_b(\mathbb{R}^2)}$ is a stochastic process such that: (i) $\Lambda(A)$ is infinitely divisible; (ii) $\Lambda(A)$ and $\Lambda(B)$ are independent if $A \cap B = \varnothing$; and, (iii) If $A_1, \ldots, A_n \in \mathcal{B}_b(\mathbb{R}^2)$ are disjoint and such that $\cup_{i=1}^n A_i \in \mathcal{B}_b(\mathbb{R}^2)$, then

$$\Lambda\left(\bigcup_{i=1}^n A_i\right) \overset{a.s.}{=} \sum_{i=1}^n \Lambda(A_i).$$

A Lévy basis Λ on $\mathbb{R}^2$ is called homogeneous if $\Lambda(A) \overset{d}{=} \Lambda(A + x_0)$, for $x_0 \in \mathbb{R}^2$, $A \in \mathcal{B}_b(\mathbb{R}^2)$.

The stochastic integral $\int f d\Lambda$ of a deterministic measurable function $f : \mathbb{R}^2 \to \mathbb{R}$ with respect to a homogeneous Lévy basis Λ is defined in two steps: (a) If $f = \sum_{i=1}^n a_i 1_{A_i}$ is a real simple function on $\mathbb{R}^2$

with $A_1, \ldots, A_n$ disjoint, for $A \in \mathcal{B}(\mathbb{R}^2)$, we define

$$\int_A f d\Lambda = \sum_{i=1}^{n} a_i \Lambda(A_i \cap A).$$

(b) If $f : \mathbb{R}^2 \to \mathbb{R}$ can be approximated almost everywhere (with respect to the Lebesgue measure) by a sequence of simple functions $\{f_n\}$ as in (a), provided that the limit exist, we define

$$\int_A f d\Lambda = P - \lim \int_A f_n d\Lambda, \tag{B.20}$$

for $A \in \mathcal{B}(\mathbb{R}^2)$. We say that a measurable function $f : \mathbb{R}^2 \to \mathbb{R}$ is Λ-integrable if the integral (B.20) exists.

Let $K\{z \ddagger X\} = \log E\{\exp(sX)\}$ denote the log-moment generating function of the random variable X. As before, $C\{z \ddagger X\} := \log E\{\exp(isX)\}$ denotes the cumulant function of the random variable X. To each homogeneous Lévy basis Λ, we can associate a random variable Λ' such that

$$K\{z \ddagger \Lambda(da)\} = K\{z \ddagger \Lambda'\} da,$$

and

$$C\{z \ddagger \Lambda(da)\} = C\{z \ddagger \Lambda'\} da.$$

The random variable Λ' is called the Lévy seed of Λ. Notice that this concept is analogous to the concept of Lévy seed that we have used for Lévy processes.

The stochastic integral $\int f d\Lambda$ and the Lévy seed Λ' satisfy the next relation.

Proposition 29 *Let Λ be a Lévy basis on $\mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}$ a Λ-integrable function. Then*

$$K\left\{z \ddagger \int_A f(a)\, d\Lambda(a)\right\} = \int_A K\{zf(a) \ddagger \Lambda'\}\, da$$

and

$$C\left\{z \ddagger \int_A f(a)\, d\Lambda(a)\right\} = \int_A C\{zf(a) \ddagger \Lambda'\}\, da.$$

It follows from Proposition 29 that the distribution of the stochastic integral is determined by the function f and the cumulant function of the Lévy seed Λ'. We use Proposition 29 to obtain a formula for the cumulants of the BSS process when ε^2 is an ambit process.

For our purposes, an ambit process is a stochastic process $(Y_t)_{t \geq 0}$ of the form

$$Y_t = \int_{A_t} f((0,t) - a)\, d\Lambda(a),$$

where $A \in \mathcal{B}_b(\mathbb{R}^2)$ and $A_t = A + (0,t)$. For a more general definition of ambit processes and a discussion of their mathematical properties, we refer to [4].

B.5.2 The formula for the cumulants

Consider the BSS process (B.2) with

$$\varepsilon_t^2 = \exp\{\Lambda(A + (0,t))\} \quad t \in \mathbb{R}, \tag{B.21}$$

where Λ is a homogeneous Lévy basis on $\mathbb{R}^2$ and $A \in \mathcal{B}_b(\mathbb{R}^2)$. In this Section, we deduce a formula for the n-th moments of ε^2 and the increments of X. These formulae can be used to compute the cumulants of the increments of X in terms of the cumulants of Λ'.

Throughout this Section, Leb will denote the Lebesgue measure on $\mathbb{R}^2$ and $K[z] \equiv K\{z \ddagger \Lambda'\}$. We start by finding an explicit formula for the n-th moments of ε^2. The next result, taken from [25], is the starting point to deduce such a formula.

Lemma 30 *For* $(s_1, \ldots, s_n) \in \mathbb{R}^n$,

$$E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\} = \exp\left\{\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr\right\}. \tag{B.22}$$

Since the form of the ambit set A (see e.g. [28]) can be very general, the evaluation of (B.22) might be difficult. The main obstacle is to split $\cup_{i=1}^{n} A_{s_i}$ into the sets $\left\{x : \sum_{i=1}^{n} 1_{A_{s_i}}(x) = j\right\}, j = 1, \ldots, n$. For the type of ambit sets we are interested in, it is easier to compute the intersections of $(A_{s_i})_{i=1}^{n}$ than the previous partition. To proceed, we need the next well-known result.

Lemma 31 *Let* $(\Omega, \mathcal{G}, Q)$ *be a probability space. If* $A_1, \ldots, A_n$ *are events, define*

$$S_m(A_1, \ldots, A_n) = \sum_{1 \le i_1 < \ldots < i_m \le n} Q\left(A_{i_1} \cap \cdots \cap A_{i_m}\right).$$

Then, for $m = 1, \ldots, n$,

$$Q\left(\sum_{i=1}^{n} 1_{A_i} = m\right) = \sum_{i=m}^{n} (-1)^{i-m} \binom{m + (i-m)}{m} S_i(A_1, \ldots, A_n).$$

Applying Lemma 31 to the formula in Lemma 30 gives the following result for $\log E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\}$.

Lemma 32 *For* $(s_1, \ldots, s_n) \in \mathbb{R}^n$,

$$\begin{aligned}\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr &= \sum_{m=1}^{n} \sum_{l=m}^{n} \sum_{1 \le i_1 < \ldots < i_l \le n} (-1)^{l-m} K[m] \binom{l}{m} \mathrm{Leb}\left(A_{s_{i_1}} \cap \cdots \cap A_{s_{i_l}}\right) \\ &= \sum_{l=1}^{n} \sum_{1 \le i_1 < \ldots < i_l \le n} \sum_{m=1}^{l} (-1)^{l-m} K[m] \binom{l}{m} \mathrm{Leb}\left(A_{s_{i_1}} \cap \cdots \cap A_{s_{i_l}}\right). \end{aligned} \tag{B.23}$$

Proof. Define

$$V = \bigcup_{i=1}^{n} A_{s_i}.$$

For $m = 1, \ldots, n$, let $(B_m)_{m=1}^{n}$ be the sequence of sets given by

$$B_m = \left\{r \in \mathbb{R}^2 : \sum_{i=1}^{n} 1_{A_{s_i}}(r) = m\right\}.$$

Since $K[0] = 0$, we have that

$$\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr = \sum_{m=1}^{n} \int_{B_m} K[m]\, dr = \sum_{m=1}^{n} K[m]\, \mathrm{Leb}(B_m). \tag{B.24}$$

Define

$$Q(A) := \frac{\mathrm{Leb}(A)}{\mathrm{Leb}(V)}.$$

The measure Q defines a probability measure on $\mathcal{B}(V)$. Thus, from Lemma 31, it follows that

$$Q(B_m) = Q\left(\left\{r : \sum_{i=1}^{n} 1_{A_{s_i}}(r) = m\right\}\right) = \sum_{l=m}^{n} (-1)^{l-m} \binom{m+(l-m)}{m} S_l\left(A_{s_1}, \ldots, A_{s_n}\right), \tag{B.25}$$

where

$$S_l(A_1, \ldots, A_n) = \sum_{1 \le i_1 < \ldots < i_l \le n} \frac{\mathrm{Leb}\left(A_{s_{i_1}} \cap \cdots \cap A_{s_{i_l}}\right)}{\mathrm{Leb}(V)}.$$

The first equality on the right hand side of (B.23) follows from (B.24) and (B.25). The second equality is a consequence of Fubini's Theorem applied to the first equality. ■

Formula (B.23) is easily evaluated numerically for n not too large. Furthermore, it permits an explicit computation of the cumulants of the increments of X when Λ' has a normal distribution and the ambit set A has a specific form (see Section B.6).

Lemma 33 *Let $n \in \mathbb{N}$. Under the convention $\prod_{j=1}^{0} \phi_t^2(s_j) = \prod_{j=1}^{0} \psi_t(s_j) = 1$, we have that*

$$E\left\{(\Delta_t X)^n\right\} = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} (2i-1)!!$$
$$\times \int_{-\infty}^{t} ds_1 \cdots \int_{-\infty}^{t} ds_i \int_{-\infty}^{t} dr_1 \cdots \int_{-\infty}^{t} dr_{n-2i} \prod_{j=1}^{i} \phi_t^2(s_j) \prod_{l=1}^{n-2i} \psi_t(s_l)\, E\left\{\varepsilon_{s_1}^2 \cdots \varepsilon_{s_n}^2 \varepsilon_{r_1}^2 \cdots \varepsilon_{r_{n-2i}}^2\right\},$$

where $n!!$ represents the double factorial of $n \in \mathbb{N} \cup \{0, -1\}$, and ϕ and ψ are the functions defined in (B.6).

Proof. For $t \ge 0$, define

$$X_t^1 = \int_{-\infty}^{t} g(t-s)\, \varepsilon_s dB_s, \quad X_t^2 = \int_{-\infty}^{t} q(t-s)\, \varepsilon_s^2 ds,$$

and $\Delta_t X^i = X_t^i - X_0^i$, for $i = 1, 2$. Since

$$\left(\Delta_t X^1\right)^i | \varepsilon \sim N\left(0, \int_{-\infty}^{t} \phi_t^2(s)\, \varepsilon_s^2 ds\right),$$

we have

$$E\left\{(\Delta_t X^1)^i \,|\varepsilon\right\} = \begin{cases} (i-1)!! \left(\int_{-\infty}^{t} \phi_t^2(s)\, \varepsilon_s^2 ds\right)^{i/2} & \text{if } i \text{ is even} \\ 0 & \text{if } i \text{ is odd.} \end{cases}$$

Thus,

$$\begin{aligned} E\left\{(\Delta_t X)^n\right\} &= E\left\{(\Delta_t X^1 + \Delta_t X^2)^n\right\} = \sum_{i=0}^{n} \binom{n}{i} E\left\{E\left\{(\Delta_t X^1)^i \,|\varepsilon\right\} (\Delta_t X^2)^{n-i}\right\} \\ &= \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} (2i-1)!! E\left\{\left(\int_{-\infty}^{t} \phi_t^2(s)\, \varepsilon_s^2 ds\right)^i \left(\int_{-\infty}^{t} \psi_t(s)\, \varepsilon_s^2 ds\right)^{n-2i}\right\}. \end{aligned}$$

The desired result is a consequence of the Fubini Theorem applied to the last term in the above equation. ■

In the turbulence modelling context, the n-moments of increments of the temporal turbulent velocity field are relevant as they constitute a basic element in the so-called Kolmogorov-Obukhov theory. In particular, for X to be a relevant model for the temporal turbulent velocity field, it should reproduce the so-called 2/3rd-Kolmogorov law (B.17) for a certain range of scales t. Lemma 33 implies that, to satisfy the 2/3rd-Kolmogorov law, $\phi_t^2(s) \propto t^{-1/6} b(t)$ and $\psi_t(s) \propto t^{-1/3} b(t)$ when t is in a neighborhood of 0, where $b(t)$ is a bounded function in such a neighborhood with $b(0) \neq 0$. In particular, X with $q = 0$ and $g(x) = x^{-1/6} e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, for $\lambda > 0$, satisfies (B.17). This example of the BSS process (B.2) was used in [27] to model turbulent velocity time series. Lemma 33 can also be used to determine the behavior of $\phi_t^2(s)$ and $\psi_t(s)$ around 0 to satisfy other scaling rules, in addition to the 2/3rd-Kolmogorov law.

Lemmata 32 and 33 provide a way to compute the moments of the increments $\Delta_t X$. Furthermore, it is possible to compute the r-cumulants $\kappa_r(\Delta_t X)$ by the recursive formula (see [26])

$$\kappa_r(\Delta_t X) = \mu_r - \sum_{j=1}^{r-1} \binom{r-1}{j} E\left\{(\Delta_t X)^j\right\} \kappa_{r-j}(\Delta_t X) \quad r \geq 2. \tag{B.26}$$

The generality in the shape of the ambit set, the distribution of Λ' and the form of the kernels g and q make it difficult to get closed expression for the n-th moments and the cumulants of $\Delta_t X$. However, the formulae in Lemmata 32 and 33 provide a simple way to evaluate the moments numerically. This might help to analyze the distribution of the increments of $\Delta_t X$.

B.6 Formula for the cumulants of the increments of a BSS-process for ε^2 specified as an exponential ambit process: examples

In this Section we study the cumulants of $\Delta_t X$ for some examples of X assuming that ε^2 is the exponential of an ambit process. In these examples the cumulants of Λ' appear implicitly in the function $K[z] \equiv K\{z \ddagger \Lambda'\}$. It is possible to rewrite the formulae we found in terms of the cumulants of Λ', but for simplicity of presentation we avoid to do so.

B.6.1 Normal Lévy basis example

In this Subsection, assuming that $q \equiv 0$ and $\Lambda' \sim$Normal(μ, δ) (i.e. $K[z] = \mu z + \frac{1}{2}\delta^2 z^2$), we analyze the cumulants of the increments of the model (B.2), (B.21) for two different types of ambit sets. In the first example, we assume that A has a triangular form. The second example deals with the case where A is an ambit set coming from a modelling framework for the turbulent energy dissipation ([28]).

The normality of Λ' allows to simplify the formula (B.23). For this, we need the next result.

Lemma 34 *Let $m, r \in \mathbb{N}$. Then,*

$$\sum_{m=0}^{l} (-1)^m \binom{l}{m}\binom{m}{r} = \begin{cases} (-1)^l & \text{if } l = r \\ 0 & \text{if } l \neq r \end{cases}.$$

Lemma 34 implies that

$$\begin{aligned} \sum_{m=1}^{l} \binom{l}{m} (-1)^{l-m} K[m] &= (-1)^l \sum_{m=1}^{l} \binom{l}{m} (-1)^m \left(\mu m + \frac{1}{2}\delta^2 m^2\right) \\ &= (-1)^l \sum_{m=1}^{l} \binom{l}{m} (-1)^m \left(\mu \binom{m}{1} + \delta^2 \left[\binom{m}{2} + \frac{1}{2}\binom{m}{1}\right]\right) \\ &= \begin{cases} \mu + \frac{\delta^2}{2} & \text{if } l = 1 \\ \delta^2 & \text{if } l = 2 \\ 0 & \text{otherwise} \end{cases}. \end{aligned}$$

Therefore, when $\Lambda' \sim$Normal(μ, δ),

$$\begin{aligned} \log E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\} &= \int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr \\ &= \sum_{l=1}^{n} \sum_{1 \le i_1 < \ldots < i_l \le n} \sum_{m=1}^{l} (-1)^{l-m} K[m] \binom{l}{m} \text{Leb}\left(A_{s_{i_1}} \cap \cdots \cap A_{s_{i_l}}\right) \\ &= \left(\mu + \frac{\delta^2}{2}\right) \sum_{1 \le i \le n} \text{Leb}\left(A_{s_i}\right) + \delta^2 \sum_{1 \le i_1 < i_2 \le n} \text{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}} \\ &= n\left(\mu + \frac{\delta^2}{2}\right) \text{Leb}(A) + \delta^2 \sum_{1 \le i_1 < i_2 \le n} \text{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}}. \end{aligned}$$

This provides a way to compute the n-moments $E\left\{\varepsilon_{s_1}^2 \varepsilon_{s_2}^2 \cdots \varepsilon_{s_n}^2\right\}$ that we will use to calculate cumulants and n-moments of $\Delta_t X$.

Normal Lévy basis example with a triangular ambit set

Let $a, T > 0$. Assume that $q \equiv 0$, $K[z] = \mu z + \frac{1}{2}\delta^2 z^2$ and that A (see Figure B.6.1) is the ambit set given by

$$A = \left\{(x, t) \in \mathbb{R}^2 : 0 \le t \le T, |x| \le \frac{a}{T}(T - t)\right\}. \tag{B.27}$$

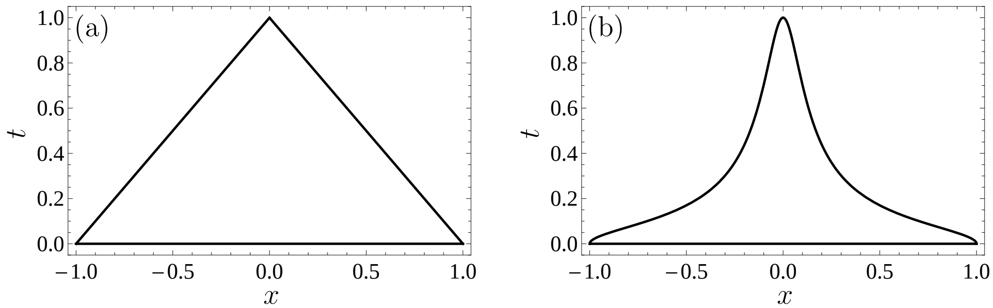

Figure B.6.1: Examples of ambit sets. (a) Ambit set (4.22) with parameters $(a, T) = (1, 1)$. (b) Ambit set (3.2) with parameters $(\theta, L, T) = (2, 10, 1)$.

For $s_1 < s_2 < s_1 + T$,

$$\text{Leb}\left(A_{s_1} \cap A_{s_2}\right) = \frac{a}{T} \cdot \left(T - |s_1 - s_2|\right)^2 .$$

Since $\text{Leb}\left(A_{s_1} \cap A_{s_2}\right)$ has this simple expression, the present example also produces a simple expression for (B.22). Namely,

$$\begin{aligned}\int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r)\right] dr =& \left(\mu + \frac{\delta^2}{2}\right) \sum_{1\le i\le n} \text{Leb}\left(A_{s_i}\right) + \delta^2 \sum_{1\le i_1<i_2\le n} \text{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}} \\ =& n\left(\mu + \frac{\delta^2}{2}\right) \text{Leb}(A) + \frac{a\delta^2}{T} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(T - |s_1 - s_2|\right)^2 1_{\{n>1\}}\end{aligned}$$

which implies that

$$\begin{aligned}\exp C\left\{z \ddagger \Delta_t X\right\} =& E\left(\exp\left\{-\frac{1}{2} z^2 \int_{\mathbb{R}} \phi_t^2(s)\, \varepsilon_s^2 ds\right\}\right) \\ =& \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} \\ & \times \int_{\mathbb{R}} ds_1 \cdots \int_{\mathbb{R}} ds_n \phi_t^2(s_1) \cdots \phi_t^2(s_n) \exp\left\{\delta^2 \sum_{1\le i_1<i_2\le n} \text{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}}\right\} \\ =& \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} \left\|\phi_t\right\|_2^{2n} + \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} a_n,\end{aligned}$$

where $\|\cdot\|_2 = \|\cdot\|_{L^2(\mathbb{R})}$ and

$$a_n = -\left\|\phi_t\right\|_2^{2n} + \int_{\mathbb{R}} ds_1 \cdots \int_{\mathbb{R}} ds_n \phi_t^2(s_1) \cdots \phi_t^2(s_n) \exp\left\{\delta^2 \sum_{1 \le i_1 < i_2 \le n} \mathrm{Leb}\left(A_{s_{i_1}} \cap A_{s_{i_2}}\right) 1_{\{n>1\}}\right\}.$$

Since $\phi_t^2(s) = 0$ for $s > t$ and by counting the intersections of A_{s_i}, we get that

$$\begin{aligned}
a_n &= \sum_{k=2}^{n} \left\|\phi_t\right\|_2^{2(n-k)} \binom{n}{k} \int_{-\infty}^{t} ds_1 \int_{s_1-T}^{s_1+T} ds_2 \cdots \int_{s_1-T}^{s_1+T} ds_k \prod_{i=1}^{k} \phi_t^2(s_k) \left(e^{\sum_{j=1}^{k} \delta^2 \mathrm{Leb}\left(A_{s_{j+1}} \cap A_{s_j}\right)} - 1 \right) \\
&\le \sum_{k=2}^{n} \left\|\phi_t\right\|_2^{2(n-k)} \binom{n}{k} \int_{-\infty}^{t} ds_1 \int_{s_1-T}^{s_1+T} ds_2 \cdots \int_{s_1-T}^{s_1+T} ds_k \phi_t^2(s_1) \cdots \phi_t^2(s_n) \left(e^{\delta^2 aTk} - 1\right) \\
&\le \left\|\phi_t\right\|_2^{2n} \sum_{k=2}^{n} \binom{n}{k} \left(e^{\delta^2 aTk} - 1\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu - \frac{\delta^2}{2}\right)} \left\|\phi_t\right\|_2^{2n} &\le \exp C\{z \ddagger \Delta_t X\} \\
&\le \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} \left\|\phi_t\right\|_2^{2n} \left(1 + \sum_{k=2}^{n} \binom{n}{k} \left(e^{\delta^2 aTk} - 1\right)\right).
\end{aligned}$$

The series on the right hand side converges, namely

$$\begin{aligned}
s &\equiv \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} z^{2n} e^{nTa\left(\mu + \frac{\delta^2}{2}\right)} \left(1 + \sum_{k=2}^{n} \binom{n}{k} \left(e^{\delta^2 aTk} - 1\right)\right) \\
&= \frac{1}{2} e^{-\left\|\phi_t\right\|_2^2 e^{aT\left(\delta^2/2+\mu\right)} z^2} \left(-2 + 2e^{\frac{\left\|\phi_t\right\|_2^2}{2} e^{aT\left(\delta^2/2+\mu\right)} z^2} + 2e^{\frac{\left\|\phi_t\right\|_2^2}{2} e^{aT\left(\delta^2/2+\mu\right)} \left(1 - e^{aT\delta^2}\right) z^2} \right. \\
&\quad \left. - \left\|\phi_t\right\|_2^2 e^{\frac{\left\|\phi_t\right\|_2^2}{2} e^{aT\left(\delta^2/2+\mu\right)} z^2 + aT\left(\delta^2/2+\mu\right)} z^2 + \left\|\phi_t\right\|_2^2 e^{\frac{\left\|\phi_t\right\|_2^2}{2} e^{aT\left(\delta^2/2+\mu\right)} z^2 + aT\left(3\delta^2/2+\mu\right)} z^2 \right).
\end{aligned}$$

For $z \approx 0$ and $a \cdot T \cdot \delta^2 \approx 0$, we have that

$$s \approx \exp\left\{ -\frac{\left\|\phi_t\right\|_2^2 z^2}{2} \exp\left\{ aT\left(\mu + \frac{\delta^2}{2}\right) \right\} \right\}.$$

Thus, for $z \approx 0$ and $a \cdot T \cdot \delta^2 \approx 0$, the previous inequality implies that

$$C\{z \ddagger \Delta_t X\} \approx -\frac{z^2}{2} e^{Ta\left(\mu - \frac{\delta^2}{2}\right)} \left\|\phi_t\right\|_2^2.$$

This means that, in this case, $\Delta_t X$ behaves similar to a normal distribution. (The distribution of $\Delta_t X$ is not exactly normal but gets more and more normal as $a \cdot T \cdot \delta^2 \downarrow 0$.)

For $T > 0, L > 1$ and $\theta > 0$, define

$$h(t) = \left(\frac{1 - (t/T)^{\theta}}{1 + (t/(T/L))^{\theta}} \right)^{1/\theta} \quad 0 \leq t \leq T. \tag{B.28}$$

Consider the model (B.2), (B.21) assuming that $K[z] = \mu z + \frac{1}{2}\delta^2 z^2$, $q \equiv 0$ and that A is the ambit set (see Figure B.6.1) given by

$$A = \left\{ (x,t) \in \mathbb{R}^2 : 0 \leq t \leq T, |x| \leq h(t) \right\}. \tag{B.29}$$

The exponential ambit process (B.21) with an ambit set of the form (B.29) has been used to model the energy dissipation in a turbulent flow (see [28]). The process (B.21) with A given by (B.29) accurately approximates the distribution of the empirical energy dissipation ε_t and is able to reproduce the scaling and self-scaling behavior observed in the empirical correlators $c_{n,m}(t)$ defined as

$$c_{n,m}(t) := \frac{E\left\{\varepsilon_t^n \varepsilon_0^m\right\}}{E\left\{\varepsilon_t^n\right\} E\left\{\varepsilon_0^m\right\}} \quad n, m \in \mathbb{N}, t > 0.$$

For $s_1 < s_2 < s_1 + T$,

$$\text{Leb}\left(A_{s_1} \cap A_{s_2}\right) = 2 \int_{s_2 - s_1}^{T} h(t)\, dt.$$

In general, it is difficult to find simple expressions for $\text{Leb}\left(A_{s_1} \cap A_{s_2}\right)$. However, for some special cases, it is possible to deduce some expressions. For instance, assuming that $\theta = 1$, we have that

$$\text{Leb}\left(A_{s_1} \cap A_{s_2}\right) = \frac{2}{L}\left(|s_2 - s_1| - T\right) - \frac{2T(1+L)}{L^2} \log\left(\frac{L|s_2 - s_1| + T}{(1+L)T} \right).$$

This implies that, for $\theta = 1$,

$$\begin{aligned} \int_{\mathbb{R}^2} K\left[\sum_{i=1}^{n} 1_{A_{s_i}}(r) \right] dr = & n\left(\mu + \frac{\delta^2}{2}\right) \text{Leb}(A) - \frac{\delta^2}{2L} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| s_i - s_j \right| + \frac{n(n-1)T\delta^2}{L} \\ & + \frac{2\delta^2 T(1+L)}{L^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \log\left(L|s_2 - s_1| + T\right) 1_{\{n>1\}} \\ & - \frac{n(n-1)\log\left((1+L)T\right) T (1+L)\delta^2}{L^2}. \end{aligned}$$

The case $(\theta, L, T) = (1.96, 10^4, 583)$ is pertinent for the modelling of turbulence since these are the parameters estimated in [28] from empirical data. We have not been able to find a closed expression for $\text{Leb}\left(A_{s_1} \cap A_{s_2}\right)$ in this case, but we have the approximation

$$\text{Leb}\left(A_{s_1} \cap A_{s_2}\right) \approx e^{559483.79 \cdot 10^{-6}} \left(\frac{1}{|s_2 - s_1|^{1/23}} - 0.195 \right).$$

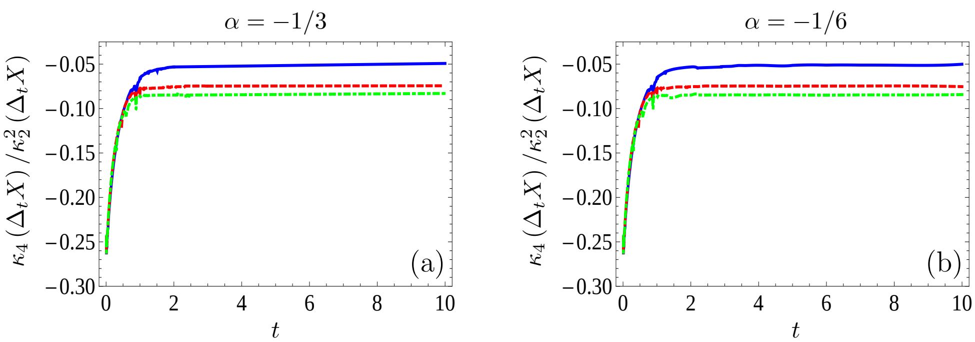

Figure B.6.2: Standardized cumulant $\kappa_4(\Delta_t X)/\kappa_2^2(\Delta_t X)$ for Example 4.4.2 with $(\theta, L, T) = (1, 10, 1)$, $\Lambda' \sim$Gamma$(1, 5)$, and different values of α and λ. (a) Parameters $\alpha = -1/3$ and $\lambda = 1, 2, 3$. (b) Parameters $\alpha = -1/6$ and $\lambda = 1, 2, 3$. We use blue for $\lambda = 1$, red for $\lambda = 2$ and green for $\lambda = 3$.

B.6.2 Gamma Lévy basis example

Let $0 > \alpha > -1/2$ and $\beta, \gamma, \lambda > 0$. Consider the model (B.2), (B.21) with $q \equiv 0$, $g(x) = x^{\alpha} e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, A given as in (B.29) and $K[z] = \log(1 - z/\beta)^{-\gamma}$, $z < \beta$ (i.e. Λ' has a Gamma(γ, β) law). In general, it is of interest to determine specific distributional properties of ΔX_t. Of particular interest is the question of infinite divisibility of $\Delta_t X$. The present example provides a case where the distribution of $\Delta_t X$ is not infinitely divisible.

Figure B.6.2 shows $\kappa_4(\Delta_t X)/\kappa_2^2(\Delta_t X)$ for $(\gamma, \beta) = (1, 5)$, $(\theta, L, T) = (1, 10, 1)$, $\alpha = -1/3, -1/6$ and $\lambda = 1, 2, 3$. The small peaks are due to numerical effects.

It is well-know that, when the distribution of X is infinitely divisible, the cumulants $\kappa_n(X)$, for $n \geq 3$, are the moments of the Lévy measure of X. This implies that, when the distribution of X is infinitely divisible, $\kappa_4(X) \geq 0$. For $q \equiv 0$, $g(x) = x^{\alpha} e^{-\lambda x} 1_{\mathbb{R}_+}(x)$, $\alpha = -1/3$, A given as in (B.29) and $\Lambda' \sim$Gamma$(1, 5)$, Figure B.6.2 shows that $\kappa_4(\Delta_t X) < 0$ for $\lambda = 1, 2, 3$. Therefore, the law of $\Delta_t X$ cannot be infinitely divisible.

B.7 Conclusion

The present work provides a way to compute the cumulants of the increments of BSS processes for two specific classes of ε^2 processes. It is not possible to find closed expression for all the cases and examples presented here. However, the formulae are simple enough to be evaluated numerically.

Of particular interest is the discussion in Subsection B.6.2 since it provides an example where our analysis of the cumulants shows that the distribution of the increments of a BSS process is not infinitely divisible. It remains open to determine conditions on the BSS processes such that their increments have an infinitely divisible law.

Our main purpose to study the cumulants of increments of BSS processes was to establish a way that sheds some light on the distributions of increments via cumulants. This is a first step to understand why the BSS approach is able to model a great variety of stylized features in turbulence. The results discussed here allow to directly compare the models with data without time consuming simulations of the underlying processes.

Acknowledgements

The author thanks Ole E. Barndorff-Nielsen and Jürgen Schmiegel for their valuable comments during the preparation of this paper. The author also acknowledges financial support from the Mexican National Science Council (Conacyt) through the PhD grant 217040.

References

[1] O.E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. Lond. A*, **353**(1674):401–419, 1977.

[2] O.E. Barndorff-Nielsen and A. Basse. Quasi Ornstein-Uhlenbeck processes. *Bernoulli*, **17**(3):916–941, 2011.

[3] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Modelling energy spot prices by volatility modulated Lévy-driven volterra processes. *Bernoulli*, **19**(3):803–845, 2013.

[4] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Recent advances in ambit stochastics with a view towards tempo-spatial stochastic volatility/intermittency. *Banach Center Publ.*, **104**:25–60, 2015.

[5] O.E. Barndorff-Nielsen, J.M. Corcuera, and M. Podolskij. Multipower variation for Brownian semi-stationary processes. *Bernoulli*, **17**(4):1159–1194, 2011.

[6] O.E. Barndorff-Nielsen and J. Schmiegel. Tempo-spatial modelling; with applications to turbulence. *Uspekhi Mat. Nauk*, **159**:65–91, 2003.

[7] O.E. Barndorff-Nielsen and J. Schmiegel. Ambit processes; with applications to turbulence and cancer growth. In F.E. Benth, G.D. Nunno, T. Linstrøm, B. Øksendal, and T. Zhang, editors, *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 93–124. Springer, Berlin Heidelberg, 2007.

[8] O.E. Barndorff-Nielsen and J. Schmiegel. A stochastic differential equation framework for the timewise dynamics of turbulent velocities. *Theory Probab. Appl.*, **52**:372–388, 2008.

[9] O.E. Barndorff-Nielsen and J. Schmiegel. Brownian semistationary processes and volatility/intermittency. In H. Albrecher, W. Runggaldier, and W. Schachermayer, editors, *Advanced Financial Modelling*, pages 1–26. Walter de Gruyter, Berlin, 2009.

[10] A. Basse, S.E. Graversen, and J. Pedersen. A unified approach to stochastic integration on the real line. *Theory Probab. Appl.*, **58**(2):193–215, 2015.

[11] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani. On the multifractal nature of fully developed turbulence and chaotic systems. *J. Phys. A*, **23**(4):3521, 1984.

[12] W. Breymann and D. Lüthi. *ghyp: A package on generalized hyperbolic distributions. Manual for the R package ghyp.* CRAN project, 2010. https://cran.r-project.org/web/packages/ghyp/ghyp.pdf.

[13] J. Cleve and M. Greiner. The markovian metamorphosis of a simple turbulent cascade model. *Phys. Lett. A*, **273**:104–108, 2000.

[14] J.M. Corcuera, E.H.L. Sørensen, M.S. Pakkanen, and M. Podolskij. Asymptotic theory for Brownian semi-stationary processes with application to turbulence. *Stochastic Process. Appl.*, **123**(7):2552–2574, 2013.

[15] Y. Cui, H. Hudzik, L. Szymaszkiewicz, and T. Wang. Criteria for monotonicity properties of Musielak–Orlicz spaces equipped with the Amemiya norm. *J. Math. Anal. Appl.*, **303**(2):376–390, 2005.

[16] E. Eberlein and E.A.V. Hammerstein. Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. In R.C. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications IV*, pages 221–264. Birkhäuser, Basel, 2004.

[17] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, **1**(3):281–299, 1995.

[18] U. Frisch. *Turbulence. The legacy of A.N. Kolmogorov*. Cambridge University Press, Cambridge, 1995.

[19] U. Frisch, P.L. Sulem, and M. Nelkin. A simple dynamical model of intermittent fully developed turbulence. *J. Fluid Mech.*, **87**(4):719–736, 1978.

[20] B. Jouault, M. Greiner, and P. Lipa. Fix-point multiplier distributions in discrete turbulent cascade models. *Phys. D*, **136**:125–144, 2000.

[21] B. Jouault, P. Lipa, and M. Greiner. Multiplier phenomenology in random multiplicative cascade processes. *Phys. Rev. E*, **59**:2451–2454, 1999.

[22] B. Mandelbrot. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, **62**(2):331–358, 1984.

[23] M.B. Marcus and J. Rosinski. L^1-norms of infinitely divisible random vectors and certain stochastic integrals. *Electron. Commun. Probab.*, **6**:15–29, 2001.

[24] J.U. Marquez and J. Schmiegel. Modelling turbulent time series by BSS-processes. In M. Podolskij, R. Stelzer, S. Thorbjørnsen, and A.E.D Veraart, editors, *The fascination of Probability, Statistics and their Applications: In Honour of Ole Barndorff-Nielsen*. Springer, Berlin, 2015.

[25] J. Schmiegel. Self-scaling of turbulent energy dissipation correlators. *Phys. Lett. A*, **337**(3):342–353, 2005.

[26] P.J. Smith. A recursive formulation of the old problem of obtaining moments from cumulants and vice-versa. *Amer. Statist.*, **49**(2):217–218, 1995.

[27] E.H.L. Sørensen. *Stochastic modelling of turbulence: With applications to wind energy.* PhD thesis, Aarhus University, 2012.

[28] E.H.L. Sørensen and J. Schmiegel. A causal continuous-time stochastic model for the turbulent energy cascade in a helium jet flow. *J. Turbul.*, **14**(11):1–26, 2013.

[29] A.E.D. Veraart and L.A.M. Veraart. Modelling electricity day-ahead prices by multivariate Lévy semistationary processes. In F.E. Benth, V. Kholodnyi, and P. Laurence, editors, *Quantitative Energy Finance*, pages 157–188. Springer, Berlin, 2014.

C

Paper C: An asymptotic problem for two classes of smooth processes

Ole E. Barndorff-Nielsen*, José Ulises Márquez*, and Mikko Pakkanen†
*Department of Mathematics, Aarhus University.
†Imperial College, London

Abstract

In this note we study an asymptotic problem for two classes of bounded variation processes: the smooth Brownian semi-stationary process and the integrated Brownian motion.

C.1 Introduction

Let $(X_t)_{t\geq 0}$ be a stochastic process. When X is a semimartingale, it is well-known that the limit

$$\lim_{n\to\infty}\sum_{i=1}^{\lfloor nt\rfloor}\left(X_{i/n}-X_{(i-1)/n}\right)^2 \tag{C.1}$$

exists in probability. In particular, when X is a bounded variation (BV) process the limit (C.1) is 0. It is natural to ask how we can rescale (C.1) to recover a non-trivial limit for the case where X is a BV process. We address this problem for two classes of smooth processes: the integrated Brownian motion (IBM) and the smooth Brownian semi-stationary ($\mathcal{BSS}$) process with a gamma kernel. A smooth process is a stochastic process that is differentiable.

The integrated Brownian motion is a stochastic process of the form

$$J_t^n = \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} B_{s_n} ds_n ds_{n-1} \cdots ds_2 ds_1 \quad n \in \mathbb{N}, \tag{C.2}$$

where $(B_t)_{t\in\mathbb{R}}$ is a standard Brownian motion. The index n indicates the number of iterated integrals and, therefore, the number of derivatives J has. We will denote by $\mathcal{IBM}$ the class of these smooth processes.

The $\mathcal{BSS}$ processes, which were introduced in [1], are stochastic processes of the form

$$Z_t = \mu + \int_{-\infty}^t g(t-s)\,\sigma_s dW_s + \int_{-\infty}^t q(t-s)\,a_s ds, \tag{C.3}$$

where μ is a constant, $(W_t)_{t\in\mathbb{R}}$ is a standard Brownian motion, g and q are nonnegative deterministic functions on $\mathbb{R}$, with $g(t) = q(t) = 0$ for $t \leq 0$, and $(\sigma_t)_{t\in\mathbb{R}}$ and $(a_t)_{t\in\mathbb{R}}$ are càdlàg processes. When (σ, a) is stationary and independent of W, then Z is stationary. This motivates the name Brownian semi-stationary. The specification of g in the gamma form

$$g(x) = x^\alpha \exp(-\lambda x)\, 1_{(0,\infty)}(x), \tag{C.4}$$

with $\alpha > -1$ and $\lambda > 0$, is of some particular interest, because of its simplicity and because it allows explicit analytic calculations. When $\alpha \in (-1/2, 1/2) \setminus \{0\}$ the process Z is not a semimartingale (see [1]). For $\alpha > 1/2$, the process Z has differentiable paths and, therefore, is a bounded variation process (see [1]). We are interested in the asymptotics of the class of smooth $\mathcal{BSS}$ processes with $q \equiv \mu = 0$, $\sigma \equiv 1$ and g given by (C.4) with $\alpha > 1/2$. We will refer to this subclass of smooth $\mathcal{BSS}$ processes as $\mathcal{SGKBSS}$.

Originally, we expected to obtain some standard central limit theorems for this problem. Soon we realized that for the considered processes the standard techniques do not work. Therefore, it is necessary to develop new techniques to obtain a satisfactory and full theory. This note contains some approaches that partially answer the questions we were interested in.

C.2 Statement of the problem

Consider a stochastic process $(X_t)_{t\geq 0}$ such that $X \in \mathcal{IBM} \cup \mathcal{SGKBSS}$. The process X has differentiable paths. From the Mean Value Theorem, it follows that the normalized realized quadratic variation (NRQV) $[X_n]$ of X, defined as

$$[X_n]_t := n \sum_{i=1}^{\lfloor nt \rfloor} \left(X_{i/n} - X_{(i-1)/n}\right)^2, \tag{C.5}$$

satisfies

$$[X_n]_t \underset{n\to\infty}{\overset{a.s.}{\to}} \int_0^t (X_r')^2\, dr \tag{C.6}$$

for $t \geq 0$. Partly motivated by this limit, we are interested in the asymptotic behavior of

$$n^\beta \left([X_n]_t - \int_0^t (X_r')^2\, dr \right) \tag{C.7}$$

when $n \to \infty$, for some suitable $\beta > 0$. There are two immediate problems related to (C.7): 1) What is the correct value of $\beta > 0$, if any, to have a non-trivial limit? 2) What is the limit distribution given the correct β? This last question can be naturally extended to stronger concepts of convergence. For simplicity we only consider the case $t = 1$.

C.3 The $\mathcal{SGKBSS}$ class

In this Section we analyse the problem stated in Section C.2 for the $\mathcal{SGKBSS}$ class.

Consider a stochastic process $(X_t)_{0\leq t\leq 1}$ whose paths are in $C^3[0,1]$. In this case, it is possible to find a β where (C.7) has an almost sure non-trivial limit. The proof of this result is an application of Taylor's Theorem and Theorem 5 in [2].

The next theorem was taken from [2], Theorem 5. The original formulation in [2] has a typo mistake, corrected below.

Theorem 35 *If f is twice differentiable and f' is bounded and almost everywhere continuous on $[0,1]$, then*

$$\lim_{n\to\infty} n^2\left(\int_0^1 f(x)\,dx - \sum_{i=1}^{n} f\left(\frac{k-\frac{1}{2}}{n}\right)\frac{1}{n}\right) = \frac{1}{24}\int_0^1 f''(x)\,dx = \frac{f'(1)-f'(0)}{24}.$$

Using Theorem 35 we can prove the next result.

Proposition 36 *Let $f \in C^3([0,1])$. Then,*

$$\lim_{n\to\infty} n^2\left(n\sum_{i=1}^{n}(\Delta_i^n f)^2 - \int_0^1 (f'(x))^2\,dx\right) = -\frac{1}{12}\int_0^1 (f''(x))^2\,dx,$$

where $\Delta_i^n f = f(i/n) - f((i-1)/n)$, $i = 1,\ldots,n$, $n\in\mathbb{N}$.

Proof. Using Taylor's Theorem, we have that

$$\Delta_i^n f = f\left(\frac{i}{n}\right) - f\left(\frac{i-1}{n}\right) = f'\left(\frac{i-1}{n}\right)\frac{1}{n} + \frac{1}{2}f''\left(\frac{i-1}{n}\right)\frac{1}{n^2} + \frac{1}{6}f'''(\xi_{i,n})\frac{1}{n^3},$$

for $(i-1)/n \leq \xi_{i,n} \leq i/n$, $i = 1,\ldots,n$, $n\in\mathbb{N}$. This implies

$$\begin{aligned}(\Delta_i^n f)^2 &= \left(f'\left(\frac{i-1}{n}\right)\right)^2\frac{1}{n^2} + \frac{1}{4}\left(f''\left(\frac{i-1}{n}\right)\right)^2\frac{1}{n^4} + \frac{1}{36}\left(f'''(\xi_{i,n})\right)^2\frac{1}{n^6}\\ &+f'\left(\frac{i-1}{n}\right)f''\left(\frac{i-1}{n}\right)\frac{1}{n^3} + \frac{1}{3}f'\left(\frac{i-1}{n}\right)f'''(\xi_{i,n})\frac{1}{n^4}\\ &+\frac{1}{6}f''\left(\frac{i-1}{n}\right)f'''(\xi_{i,n})\frac{1}{n^5}.\end{aligned}$$

From the last equation we get

$$\begin{aligned}S_n &\equiv n^2\left(n\sum_{i=1}^{n}(\Delta_i^n f)^2 - \int_0^1 (f'(x))^2\,dx\right)\\ &= n^2\left(\sum_{i=1}^{n}\left(f'\left(\frac{i-1}{n}\right)\right)^2\frac{1}{n} - \int_0^1 (f'(x))^2\,dx\right) + \frac{1}{4}\sum_{i=1}^{n}\left(f''\left(\frac{i-1}{n}\right)\right)^2\frac{1}{n}\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=1}^{n} f'\left(\frac{i-1}{n}\right) f''\left(\frac{i-1}{n}\right)+\frac{1}{3}\sum_{i=1}^{n} f'\left(\frac{i-1}{n}\right) f'''\left(\xi_{i,n}\right)\frac{1}{n}\\
&+\frac{1}{6}\sum_{i=1}^{n} f''\left(\frac{i-1}{n}\right) f'''\left(\xi_{i,n}\right)\frac{1}{n^2}+\frac{1}{36}\sum_{i=1}^{n}\left(f'''\left(\xi_{i,n}\right)\right)^2\frac{1}{n^3}\\
\equiv{}& \mathsf{A}_n+\frac{1}{4}\mathsf{B}_n+\mathsf{C}_n+\frac{1}{3}\mathsf{D}_n+\frac{1}{6}\mathsf{E}_n+\frac{1}{36}\mathsf{F}_n.
\end{aligned}
$$

We rewrite

$$
\begin{aligned}
\mathsf{A}_n &= n^2\left(\sum_{i=1}^{n}\left(f'\left(\frac{i-1}{n}\right)\right)^2\frac{1}{n}-\sum_{i=1}^{n}\left(f'\left(\frac{i-1/2}{n}\right)\right)^2\frac{1}{n}\right)\\
&\quad+n^2\left(\sum_{i=1}^{n}\left(f'\left(\frac{i-1/2}{n}\right)\right)^2\frac{1}{n}-\int_0^1\left(f'(x)\right)^2 dx\right)\\
&\equiv \mathsf{A}_n^1+\mathsf{A}_n^2.
\end{aligned}
$$

Taylor's Theorem applied to $(f')^2$ implies

$$
\begin{aligned}
\left(f'\left(\frac{i-1/2}{n}\right)\right)^2-\left(f'\left(\frac{i-1}{n}\right)\right)^2 &= \frac{d}{dx}\left(f'(x)\right)^2\bigg|_{x=\frac{i-1}{n}}\frac{1}{2n}+\frac{1}{8n^2}\frac{d^2}{dx^2}\left(f'(x)\right)^2\bigg|_{x=\eta_{i,n}}\\
&= f'\left(\frac{i-1}{n}\right) f''\left(\frac{i-1}{n}\right)\frac{1}{n}+\left(f''\left(\eta_{i,n}\right)\right)^2\frac{1}{(2n)^2}\\
&\quad+f'\left(\eta_{i,n}\right) f'''\left(\eta_{i,n}\right)\frac{1}{(2n)^2}
\end{aligned}
$$

for $(i-1)/n \le \eta_{i,n} \le i/n$, $i=1,\ldots,n$, $n\in\mathbb{N}$. Thus,

$$
\begin{aligned}
\mathsf{A}_n^1 &= n\sum_{i=1}^{n}\left\{\left(f'\left(\frac{i-1}{n}\right)\right)^2-\left(f'\left(\frac{i-1/2}{n}\right)\right)^2\right\}\\
&= -\sum_{i=1}^{n} f'\left(\frac{i-1}{n}\right) f''\left(\frac{i-1}{n}\right)-\frac{1}{4}\sum_{i=1}^{n}\left\{\left(f''\left(\eta_{i,n}\right)\right)^2+f'\left(\eta_{i,n}\right) f'''\left(\eta_{i,n}\right)\right\}\frac{1}{n}.
\end{aligned}
$$

Since $f^{(j)}f^{(k)}$ and $f^{(j)}$ are integrable for $j,k=1,2,3$, we have that

$$
\begin{aligned}
S_n-\mathsf{A}_n^2 &= \mathsf{A}_n^1+\frac{1}{4}\mathsf{B}_n+\mathsf{C}_n+\frac{1}{3}\mathsf{D}_n+\frac{1}{6}\mathsf{E}_n+\frac{1}{36}\mathsf{F}_n\\
&= -\sum_{i=1}^{n} f'\left(\frac{i-1}{n}\right) f''\left(\frac{i-1}{n}\right)-\frac{1}{4}\sum_{i=1}^{n}\left\{\left(f''\left(\eta_{i,n}\right)\right)^2+f'\left(\eta_{i,n}\right) f'''\left(\eta_{i,n}\right)\right\}\frac{1}{n}\\
&\quad+\frac{1}{4}\sum_{i=1}^{n}\left(f''\left(\frac{i-1}{n}\right)\right)^2\frac{1}{n}+\sum_{i=1}^{n} f'\left(\frac{i-1}{n}\right)\left\{f''\left(\frac{i-1}{n}\right)+\frac{1}{3}f'''\left(\xi_{i,n}\right)\frac{1}{n}\right\}\\
&\quad+\frac{1}{6}\sum_{i=1}^{n} f''\left(\frac{i-1}{n}\right) f'''\left(\xi_{i,n}\right)\frac{1}{n^2}+\frac{1}{36}\sum_{i=1}^{n}\left(f'''\left(\xi_{i,n}\right)\right)^2\frac{1}{n^3}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \left\{ \frac{1}{3} f\left(\frac{i-1}{n}\right) f''\left(\xi_{i,n}\right) - \frac{1}{4} f\left(\eta_{i,n}\right) f''\left(\eta_{i,n}\right) \right\} \frac{1}{n} \\
&+ \frac{1}{4} \sum_{i=1}^{n} \left\{ \left(f'\left(\frac{i-1}{n}\right)\right)^2 - \left(f'\left(\eta_{i,n}\right)\right)^2 \right\} \frac{1}{n} + \frac{1}{6} \sum_{i=1}^{n} f'\left(\frac{i-1}{n}\right) f''\left(\xi_{i,n}\right) \frac{1}{n^2} \\
&+ \frac{1}{36} \sum_{i=1}^{n} \left(f''\left(\xi_{i,n}\right)\right)^2 \frac{1}{n^3} \\
&\underset{n\to\infty}{\to} \int_0^1 \left(\frac{1}{3} f(x) f''(x) - \frac{1}{4} f(x) f''(x) \right) dx = \frac{1}{12} \int_0^1 f(x) f''(x)\, dx. \qquad \text{(C.8)}
\end{aligned}
$$

On the other hand, Theorem 35 implies

$$
\begin{aligned}
\mathsf{A}_n^2 &= n^2 \left(\sum_{i=1}^{n} \left(f\left(\frac{i-1/2}{n}\right)\right)^2 \frac{1}{n} - \int_0^1 \left(f(x)\right)^2 dx \right) \\
&\underset{n\to\infty}{\to} -\frac{1}{24} \int_0^1 \frac{d^2}{dx^2} \left(f(x)\right)^2 dx = \frac{f(0) f'(0) - f(1) f'(1)}{12}. \qquad \text{(C.9)}
\end{aligned}
$$

Combining equations (C.8) and (C.9) gives

$$
\begin{aligned}
\lim_{n\to\infty} \mathsf{S}_n &= \frac{1}{12} \int_0^1 f(x) f''(x)\, dx - \frac{1}{24} \int_0^1 \frac{d^2}{dx^2} \left(f(x)\right)^2 dx \\
&= \frac{1}{12} \left(\int_0^1 f(x) f''(x)\, dx - \int_0^1 \left(\left(f'(x)\right)^2 + f(x) f''(x) \right) dx \right) \\
&= -\frac{1}{12} \int_0^1 \left(f'(x)\right)^2 dx,
\end{aligned}
$$

which concludes the proof. ■

Proposition 36 implies the main result of this Section.

Corollary 37 *Let $(X_t)_{0\le t\le 1}$ be a stochastic process whose paths are almost surely in $C^3[0,1]$. Then,*

$$
n^2 \left([X_n]_1 - \int_0^1 \left(X_r'\right)^2 dr \right) \underset{n\to\infty}{\overset{a.s.}{\to}} -\frac{1}{12} \int_0^1 \left(X_r''\right)^2 dr = -\frac{1}{12} \left\| X'' \right\|_{L^2(0,1)}^2.
$$

If $X \in \mathcal{SGKBSS}$ with $\alpha > 5/2$, the paths of X are almost surely in $C^3([0,\infty))$. The distribution of $Y = \frac{1}{12} \|X''\|_{L^2(0,1)}^2$ depends on the value of α. It might possible to find the characteristic function of Y reproducing the arguments used in [5] to obtain the characteristic function of the Rosenblatt distribution.

Remark 38 *Since the index n represents the number of derivatives for the IBM J^n, $J^3 \in C^3([0,\infty))$. Consequently, Corollary 37 also applies to J^n, for $n > 3$. Integrating by parts we get*

$$
\frac{d^2}{ds^2} J_s^3 = J_s^1 = \int_0^s B_{s_1} ds_1 = \frac{1}{2} \int_0^1 (s-1)\, dB_s.
$$

The technique used to determine the limit in Proposition 36 requires $f \in C^3[0,1]$. This method cannot be extended to functions in $C^2[0,1] \setminus C^3[0,1]$ or $C^1[0,1] \setminus C^3[0,1]$. It is necessary a new technique to determine the rate of convergence for functions in these sets. When $f = X$ is a stochastic process, a broadly used technique consists in using moments to estimate the rate of convergence of quantities similar to (C.7). Since the terms in (C.7) display a complicated dependence, the use of moments did not help us to find the rate.

It remains open to determine the rate of convergence and the limit distribution of (C.7) for the stochastic processes in $\mathcal{SGKBSS}$ with index $1/2 < \alpha < 5/2$, that is, for the BSS processes that are not in $C^3[0,1]$. There is evidence suggesting that $n^{\alpha-1/2}$ might be the correct rate of convergence.

C.4 Integrated Brownian motion

In this Section, we perform some investigations about the limit distribution of (C.7) for J_t^1. In Subsection C.4.1 we establish a L_2 convergence together with its respective rate of convergence. Subsection C.4.2 discusses a conjecture about the limit distribution for J_t^1. Subsection C.4.3 presents a useful decomposition that partially motivates the conjectured limit distribution for J_t^1. Subsection C.4.4 contains the numerical validation of the conjecture of Subsection C.4.2.

C.4.1 Convergence of the variance

We have that

$$[J_n^m]_t \underset{n\to\infty}{\overset{a.s.}{\to}} \int_0^t \left(\frac{d}{dr} J_r^m \bigg|_{r=s} \right)^2 ds,$$

where $[J_n^m]$ denotes the NRQV of J^m as defined in Subsection C.2.

Restricting to the case $m = 1$, we get

$$[J_n^1]_t \underset{n\to\infty}{\overset{a.s.}{\to}} \int_0^t B_s^2 ds$$

for which we first analyze the variance

$$A_{n,t} := \text{Var}\left([J_n^1]_t - \int_0^t B_s^2 ds \right)$$

for $t \geq 0$. Isserlis' Theorem (see [3]) implies the next result.

Proposition 39 *For $t \geq 0$ and $n \in \mathbb{N}$,*

$$A_{n,t} = \frac{45\lfloor nt \rfloor^4 - 60nt\lfloor nt \rfloor^3 - 15\lfloor nt \rfloor^2 + 15nt\lfloor nt \rfloor + \lfloor nt \rfloor + 15n^4t^4}{45n^4}.$$

Proof. We will compute $A_{n,t}$ using

$$\text{Var}\left([J_n^1]_t - \int_0^t B_s^2 ds \right) = E\left\{ \left([J_n^1]_t - \int_0^t B_s^2 ds \right)^2 \right\} - E\left\{ [J_n^1]_t - \int_0^t B_s^2 ds \right\}^2.$$

We have that

$$E\left\{\left(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^n J^1)^2\right)^2\right\} = \sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\left\{(\Delta_i^n J^1)^2\left(\Delta_j^n J^1\right)^2\right\} + \sum_{i=1}^{\lfloor nt\rfloor} E\left\{(\Delta_i^n J^1)^4\right\}$$
$$\equiv A_1 + A_2.$$

Fubini's Theorem implies that

$$A_1 = 2\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{j-1}{n}}^{\frac{j}{n}}\int_{\frac{j-1}{n}}^{\frac{j}{n}} E\left\{B_{s_1}B_{s_2}B_{r_1}B_{r_2}\right\} dr_2 dr_1 ds_2 ds_1,$$
$$A_2 = \sum_{i=1}^{\lfloor nt\rfloor}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}} E\left\{B_{r_1}B_{r_2}B_{r_3}B_{r_4}\right\} dr_4 dr_3 dr_2 dr_1.$$

From Isserlis' Theorem (see [3]) it follows that

$$E\left\{B_{r_1}B_{r_2}B_{r_3}B_{r_4}\right\} = (r_1\wedge r_2)(r_3\wedge r_4) + (r_1\wedge r_3)(r_2\wedge r_4) + (r_1\wedge r_4)(r_2\wedge r_3).$$

Then, we have that

$$A_1 = 2\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{j-1}{n}}^{\frac{j}{n}}\int_{\frac{j-1}{n}}^{\frac{j}{n}} \left[(s_1\wedge s_2)(r_1\wedge r_2) + 2s_1s_2\right] dr_2 dr_1 ds_2 ds_1$$
$$= 2\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\left[\frac{(3i-2)(3j-2)}{9n^6} + \frac{(1-2i)^2}{2n^6}\right],$$
$$A_2 = 3\sum_{i=1}^{\lfloor nt\rfloor}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}(r_1\wedge r_2)\, dr_2 dr_1\right)^2 = \sum_{i=1}^{\lfloor nt\rfloor}\frac{(3i-2)^2}{3n^6}.$$

Therefore,

$$E\left\{\left(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^n J^1)^2\right)^2\right\} = 2\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\left[\frac{(3i-2)(3j-2)}{9n^6} + \frac{(1-2i)^2}{2n^6}\right] + \sum_{i=1}^{\lfloor nt\rfloor}\frac{(3i-2)^2}{3n^6}.$$

On the other hand, we have that

$$E\left\{\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^n J^1)^2\int_0^t B_s^2 ds\right\} = \sum_{i=1}^{\lfloor nt\rfloor}\int_0^t E\left\{(\Delta_i^n J^1)^2 B_s^2\right\} ds$$
$$= \sum_{i=1}^{\lfloor nt\rfloor}\int_0^t\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}} E\left\{B_{r_1}B_{r_2}B_s^2\right\} dr_2 dr_1 ds.$$

Isserlis' Theorem implies that

$$E\left\{B_{r_1}B_{r_2}B_s^2\right\} = s\left(r_1 \wedge r_2\right) + 2\left(s \wedge r_1\right)\left(s \wedge r_2\right),$$

and, therefore,

$$\begin{aligned}
E\left\{\sum_{i=1}^{\lfloor nt\rfloor}\left(\Delta_i^n J^1\right)^2 \int_0^t B_s^2 ds\right\} &= \sum_{i=1}^{\lfloor nt\rfloor}\left(\int_0^t s ds\right)\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\left(r_1 \wedge r_2\right) dr_2 dr_1\right) \\
&\quad + 2\sum_{i=1}^{\lfloor nt\rfloor}\int_0^t\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\left(s \wedge r_1\right)\left(s \wedge r_2\right) dr_2 dr_1 ds\right) \\
&= \sum_{i=1}^{\lfloor nt\rfloor}\frac{\left(60i^2nt - 60int + 15nt - 40i^3 + 60i^2 - 35i + 8\right)}{30n^5} \\
&\quad + \sum_{i=1}^{\lfloor nt\rfloor}\frac{\left(3i-2\right)t^2}{6n^3}.
\end{aligned}$$

Then, combining the previous expectations, we get

$$\begin{aligned}
E\left\{\left(\left[J_n^1\right]_t - \int_0^t B_s^2 ds\right)^2\right\} &= E\left\{\left(n\sum_{i=1}^{\lfloor nt\rfloor}\left(\Delta_i^n J^1\right)^2 - \int_0^t B_s^2 ds\right)^2\right\} \\
&= 2\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\left[\frac{\left(3i-2\right)\left(3j-2\right)}{9n^5} + \frac{\left(1-2i\right)^2}{2n^5}\right] \\
&\quad + \sum_{i=1}^{\lfloor nt\rfloor}\frac{\left(3i-2\right)^2}{3n^5} + \frac{7}{12}t^4 - \sum_{i=1}^{\lfloor nt\rfloor}\frac{\left(3i-2\right)t^2}{3n^3} \\
&\quad - \sum_{i=1}^{\lfloor nt\rfloor}\frac{\left(60i^2nt - 60int + 15nt - 40i^3 + 60i^2 - 35i + 8\right)}{15n^5} \\
&= \frac{105n^4t^4 + 30n^2t^2\left(1 - 3\lfloor nt\rfloor\right)\lfloor nt\rfloor - 60nt\lfloor nt\rfloor\left(4\lfloor nt\rfloor^2 - 1\right)}{180n^4} \\
&\quad + \frac{\lfloor nt\rfloor\left(4 - 55\lfloor nt\rfloor - 30\lfloor nt\rfloor^2 + 225\lfloor nt\rfloor^3\right)}{180n^4}.
\end{aligned}$$

Furthermore, we have that

$$E\left\{\sum_{i=1}^{\lfloor nt\rfloor}\left(\Delta_i^n J^1\right)^2\right\} = \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{i-1}{n}}^{\frac{i}{n}} E\left\{B_{r_1}B_{r_2}\right\} dr_2 dr_1\right) = \frac{\lfloor nt\rfloor}{6n^3}\left(3\lfloor nt\rfloor - 1\right).$$

Thus,

$$E\left\{\left[J_n^1\right]_t - \int_0^t B_s^2 ds\right\}^2 = \frac{\lfloor nt\rfloor^2}{36n^4}\left(3\lfloor nt\rfloor - 1\right)^2 - \frac{t^2\lfloor nt\rfloor}{6n^3}\left(3\lfloor nt\rfloor - 1\right) + \frac{t^4}{4}$$

and

$$\operatorname{Var}\left([J_n^1]_t - \int_0^t B_s^2 ds\right) = \frac{45\lfloor nt\rfloor^4 - 60nt\lfloor nt\rfloor^3 - 15\lfloor nt\rfloor^2 + 15nt\lfloor nt\rfloor + \lfloor nt\rfloor + 15n^4t^4}{45n^4}.$$

This finishes the proof. ■

An interesting observation about $A_{n,t}$ is that we do not have the same order of convergence for different t's. To exemplify this, consider two different values $t_1 = 1$ and $t_2 = 1/3$. Notice that, for $n, k \in \mathbb{N}$,

$$A_{n,t_1} = \frac{1}{45n^3}, \quad A_{3k,t_2} = \frac{1}{45k^3}, \quad A_{3k+1,t_2} = \frac{270k^2 + 222k + 5}{1215\,(1+3k)^4}.$$

Clearly, A_{n_k,t_1} tends to 0 with the same order for any subsequence $\{n_k\}$. However, the last equations show that A_{n,t_2} has different rates of convergence for different subsequences.

An immediate consequence of Proposition 39 is the next result.

Proposition 40 *For $t \geq 0$,*

$$[J_n^1]_t \overset{L^2}{\to} \int_0^t B_s^2 ds.$$

C.4.2 The limit distribution

We are interested in shedding light on the limit distribution of

$$\mathfrak{A}_{n,t} := A_{n,t}^{-1/2}\left([J_n^1]_t - \int_0^t B_s^2 ds\right).$$

To clarify the main aspects, we only consider the case $t = 1$.

It turns out (see Subsection C.4.4) that the limit law of $\mathfrak{A}_{n,1}$ seems to be a Rosenblatt distribution [4, 7]. Using Maejima and Tudor's parametrization of the Rosenblatt distribution [4], we have the next conjecture.

Conjecture 41 *Let $R(h)$ denote a Rosenblatt random variable with index $h \in (1/2, 1)$. Then, we have*

$$\mathfrak{A}_{n,1} \overset{d}{\to} R(h),$$

where $h \approx 0.9$.

Numerical simulations of $\mathfrak{A}_{n,1}$ and $R(0.9)$ show strong evidence supporting the previous conjecture (see Subsection 5.4).

C.4.3 An expression for $\mathfrak{A}_{n,1}$

In this Subsection we derive an expression for the random variable

$$\mathfrak{X}_n := [J_n^1]_t - \int_0^t B_s^2 ds = n\sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} B_s \mathrm{d}s\right)^2 - \int_0^1 B_s^2 \mathrm{d}s,$$

in terms of a double Wiener integral with respect to B. This partly allows to heuristically justify the Rosenblatt distribution for $\mathfrak{A}_{n,1}$.

Proposition 42 *We can rewrite the random variable $\mathfrak{X}_n$ as*

$$\mathfrak{X}_n = \int_0^1 \int_0^1 F_n(r,s)\mathrm{d}B_r\mathrm{d}B_s - \frac{5}{6n}, \tag{C.10}$$

where

$$F_n(r,s) = n\sum_{i=1}^{n} f_i^{(n)}(r) f_i^{(n)}(s) - \big(1-\max(r,s)\big)$$

and

$$f_i^{(n)}(s) = \begin{cases} \frac{1}{n}, & s \in \left[0, \frac{i-1}{n}\right), \\ \frac{i}{n} - s, & s \in \left[\frac{i-1}{n}, \frac{i}{n}\right), \\ 0, & s \in \left[\frac{i}{n}, 1\right]. \end{cases} \tag{C.11}$$

Proof. Using integration by parts (twice), we obtain

$$\begin{aligned}\int_0^1 B_s^2\mathrm{d}s &= B_1^2 - \int_0^1 s\mathrm{d}B_s^2 = 2\int_0^1 B_s\mathrm{d}B_s + 1 - 2\int_0^1 sB_s\mathrm{d}B_s - \int_0^1 s\mathrm{d}s \\ &= 2\int_0^1 (1-s)B_s\mathrm{d}B_s + \frac{1}{2} = 2\int_0^1 (1-s)\int_0^s \mathrm{d}B_r\mathrm{d}B_s + \frac{1}{2} \\ &= 2\int_0^1\int_0^s \big(1-\max(r,s)\big)\mathrm{d}B_r\mathrm{d}B_s + \frac{1}{2} = \int_0^1\int_0^1 \big(1-\max(r,s)\big)\mathrm{d}B_r\mathrm{d}B_s + \frac{1}{2}, \end{aligned} \tag{C.12}$$

where the last equality follows because $(r,s) \mapsto 1-\max(r,s)$ is a symmetric function.

Suppose that $0 \le a < b \le 1$. Integration by parts yields

$$\begin{aligned}\int_a^b B_s\mathrm{d}s &= bB_b - aB_a - \int_a^b s\mathrm{d}B_s = \int_0^1 \left(b1_{(0,b)}(s) - a1_{(0,a)}(s) - s1_{(a,b)}(s)\right)\mathrm{d}B_s \\ &= \int_0^1 \max\left(b - \max(a,s), 0\right)\mathrm{d}B_s.\end{aligned}$$

We can thus write

$$\int_{\frac{i-1}{n}}^{\frac{i}{n}} B_s\mathrm{d}s = \int_0^1 f_i^{(n)}(s)\mathrm{d}B_s,$$

where $f_i^{(n)}(s)$ is as defined in (C.11). A straightforward calculation yields

$$\int_0^1 f_i^{(n)}(s)^2\mathrm{d}s = \int_0^{\frac{i-1}{n}} \frac{\mathrm{d}s}{n^2} + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left(\frac{i}{n} - s\right)^2\mathrm{d}s = \frac{i-1}{n^3} + \int_0^{\frac{1}{n}} s^2\mathrm{d}s = \frac{i-1}{n^3} - \frac{1}{3n^3} = \frac{i-\frac{4}{3}}{n^3}.$$

Using the multiplication formula for Wiener integrals, we find that

$$\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} B_s\mathrm{d}s\right)^2 = \left(\int_0^1 f_i^{(n)}(s)\mathrm{d}B_s\right)^2 = \int_0^1\int_0^1 f_i^{(n)}(r)f_i^{(n)}(s)\mathrm{d}B_r\mathrm{d}B_s + \int_0^1 f_i^{(n)}(s)^2\mathrm{d}s.$$

Thus,

$$n\sum_{i=1}^{n}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}}B_s\mathrm{d}s\right)^2=\int_0^1\int_0^1\left(n\sum_{i=1}^{n}f_i^{(n)}(r)f_i^{(n)}(s)\right)\mathrm{d}B_r\mathrm{d}B_s+\frac{1}{2}-\frac{5}{6n}. \tag{C.13}$$

Combining C.12 and C.13, we obtain

$$\mathfrak{X}_n=\int_0^1\int_0^1 F_n(r,s)\mathrm{d}B_r\mathrm{d}B_s-\frac{5}{6n},$$

where

$$F_n(r,s)=n\sum_{i=1}^{n}f_i^{(n)}(r)f_i^{(n)}(s)-\big(1-\max(r,s)\big).$$

This concludes the proof. ■

The double integral expression (C.10) might provide evidence supporting Conjecture 41 since the Rosenblatt distribution can be expressed as a second order Wiener chaos (see, e.g. [6]). One possible way to prove Conjecture 41 would be to show that $F_n(r,s)$ converges in L_2 to the kernel that appears in Proposition 1 of [6]. However, this is work in progress.

C.4.4 Numerical results

In this subsection we present some numerical results that motivate Conjecture 41.

To test our conjecture about $\mathfrak{A}_{n,1}$, we need to simulate the Rosenblatt distribution. The next asymptotic result provides a way to do so.

Proposition 43 *Let ξ_i, $i=1,2,\ldots$, be a stationary Gaussian process with $E[\xi_i]=0$, $E[\xi_i^2]=1$, and*

$$E\left[\xi_i\xi_j\right]=\frac{1}{2}\left((j+1)^{2H}-2j^{2H}+(j-1)^{2H}\right),\quad j=1,2,\ldots,$$

where $H\in(0,1)$. If $H\in(3/4,1)$ then

$$\frac{1}{n^{2H-1}}\sum_{i=1}^{n}(\xi_i^2-1)\xrightarrow{d}\lim_{n\to\infty}\mathit{Rosenblatt}\,(2H-1)\,.$$

Remark 44 *The random variables ξ_i have the same distribution as the normalized increments of fractional Brownian motion with Hurst parameter H.*

Also, in order to test the conjecture about $\mathfrak{A}_{n,1}$, it is necessary to simulate such a random variable. This means that we need to simulate the increments of J^1 and the integral $\int_0^1 B_s^2 ds$. We simulate the increments of J^1 by approximating Riemann sums. Being more precise, we partition the interval $[0,1]$ in $k\gg n$ equidistant points, i.e. we use the partition $\{0,1/k,\ldots,(k-1)/k,1\}$. Then, we approximate the increments

$$\Delta_m^n J^1:=J^1_{m/n}-J^1_{(m-1)/n}=\int_{(m-1)/n}^{m/n}B_s ds,\quad \text{for } m=1,\ldots,n,$$

through the Riemann sum $\sum_{i=0}^{\lfloor k/n\rfloor}B_{(i+m)/k}/k$. Besides, we approximate the value of $\int_0^1 B_s^2 ds$ by the Riemann sum $\sum_{i=0}^{k}B_{i/k}^2/k$, where we consider the previous partition of $[0,1]$. Here, it is important to note

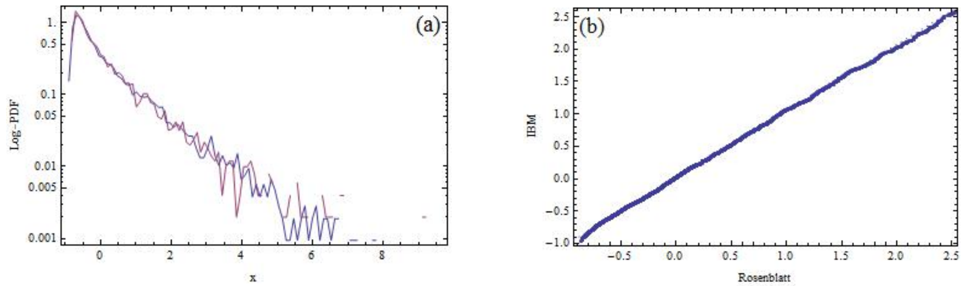

Figure C.4.1: (a) Histograms in log-linear scale of the Rosenblatt sample (dashed) and the $\mathfrak{A}_{100000,1}$ sample (solid). (b) QQ plot of the $\mathfrak{A}_{100000,1}$ sample and the Rosenblatt sample.

that $[J_n^1]_1$ and $\int_0^1 B_s^2 ds$ are not independent.

Figure 1 illustrates the conjecture for $\mathfrak{A}_{n,1}$. These figures were obtained by simulating 10000 samples of *Rosenblatt* (0.9) and 10000 samples of $\mathfrak{A}_{n,1}$ with a discretization of $k = 10^7$ and an increment size of $n^{-1} = 10^{-5}$. We standardized the samples by subtracting the sample mean and dividing by the sample variance. Figure (a) shows two histograms in log-linear scale: the histogram of the Rosenblatt sample (blue) and the histogram of the $\mathfrak{A}_{100000,1}$ sample (purple). They are very similar. Figure (b) corresponds to the QQ plot of the $\mathfrak{A}_{100000,1}$ sample and the Rosenblatt sample.

C.5 Conclusion

In this note, we have partially solved the problem stated in Section C.2 for the class $\mathcal{SGKBSS}$. The techniques used here do not provide the full answer since they do not permit to determine the limit distributions. We have also provided strong evidence supporting a Rosenblatt limit for the asymptotic problem in the J^1 case. It remains open to determine the veracity of such a limit.

Acknowledgments

The authors thank Jürgen Schmiegel for his valuable comments in the preparation of this work.

References

[1] O.E. Barndorff-Nielsen and J. Schmiegel. Brownian semistationary processes and volatility/intermittency. In H. Albrecher, W. Runggaldier, and W. Schachermayer, editors, *Advanced Financial Modelling*, pages 1–26. Walter de Gruyter, Berlin, 2009.

[2] C.K. Chui. Concerning rates of convergence of Riemann sums. *J. Approx. Theory*, **4**(3):279–287, 1971.

[3] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, **12**:134–139, 1918.

[4] M. Maejima and C.A. Tudor. On the distribution of the Rosenblatt process. *Stat. Probabil. Lett.*, **83**(6):1490–1495, 2013.

[5] M. Rosenblatt. On the distribution of the Rosenblatt process. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.*, volume **II**, pages 157–188, Berkeley, 1961. Univ. California Press.

[6] C.A. Tudor. Analysis of the Rosenblatt process. *ESAIM Probab. Stat.*, **12**:230–257, 2008.

[7] M.S. Veillette and M.S. Taqqu. Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli*, **19**(3):721–1085, 2008.

D

Some relevant probability distributions

THIS APPENDIX briefly discusses some probability distributions that are used in this thesis.

D.1 Generalized inverse Gaussian distribution

The generalized inverse Gaussian (GIG) distribution is a three-parameter family of continuous probability distributions whose probability density function is given by

$$f_{\mathrm{GIG}(\lambda,\chi,\theta)}(x) = \left(\frac{\theta}{\lambda}\right)^{\lambda/2} \frac{x^{\lambda-1}}{2K_\lambda\left(\sqrt{\chi\theta}\right)} \exp\left\{-\frac{1}{2}\left(\frac{\chi}{x}+\theta x\right)\right\}, \quad x \in \mathbb{R}$$

where K_λ denotes the modified Bessel function of the third kind with index λ and the parameters (λ, χ, θ) have to satisfy one of the following three restrictions

$$\chi > 0, \theta \geq 0, \lambda < 0 \;\text{ or }\; \chi > 0, \theta > 0, \lambda = 0 \;\text{ or }\; \chi \geq 0, \theta > 0, \lambda > 0. \tag{D.1}$$

The distribution is denoted by $\mathrm{GIG}(\lambda, \chi, \theta)$.

The GIG law nests the inverse Gaussian (IG) distribution ($\lambda = -1/2$) and contains many special cases as limits, among others the Gamma and Inverse Gamma distributions.

The GIG distribution is infinitely divisible. Besides, the moment generating function of a random variable $V \sim \mathrm{GIG}(\lambda, \chi, \theta)$ is determined by

$$E\left[\exp zV\right] = \left(\frac{\psi}{\psi - 2z}\right)^{\lambda/2} \frac{K_\lambda\left(\sqrt{\chi(\psi - 2z)}\right)}{K_\lambda\left(\sqrt{\chi\psi}\right)}.$$

For a more extensive discussion of the GIG law, we refer to [22].

D.2 Generalized hyperbolic distribution

The generalized hyperbolic (GH) distribution was introduced in [1] to describe the law of the logarithm of the size of sand particles. The GH class is a very rich and flexible family of distributions. For an extensive discussion of this distribution and its applications, we refer, for instance, to [1, 17, 18] and the papers cited in [14].

Different parameterizations are known for the GH law. A summary of the most used parameterizations, and the relation between them, can be found in [14]. Here, we focus our attention to the so-called $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$-parametrization. The density of the $\mathrm{GH}(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ law is

$$\begin{aligned} f_{\mathrm{GH}(\lambda,\chi,\psi,\mu,\Sigma,\gamma)}(x) &= \frac{\left(\sqrt{\psi/\chi}\right)^{\lambda}\left(\psi+\Sigma^{-1}\gamma^{2}\right)^{\frac{1}{2}-\gamma}}{(2\pi)^{\frac{1}{2}}\,\Sigma^{1/2}K_{\lambda}\left(\sqrt{\psi\gamma}\right)} \\ &\quad\times\frac{K_{\lambda-\frac{1}{2}}\left(\sqrt{\left(\chi+\Sigma^{-1}(x-\mu)^{2}\right)\left(\psi+\Sigma^{-1}\gamma^{2}\right)}\right)e^{\Sigma^{-1}\gamma(x-\mu)}}{\sqrt{\left(\chi+\Sigma^{-1}(x-\mu)^{2}\right)\left(\psi+\Sigma^{-1}\gamma^{2}\right)}} \end{aligned}$$

where K_λ denotes the modified Bessel function of the third kind with index λ, $(\lambda, \chi, \psi, \mu, \Sigma, \gamma) \in \mathbb{R} \times \mathbb{R}^2_+ \times \mathbb{R}^3$ and $x \in \mathbb{R}$. The parameters (λ, χ, ψ) satisfy one of the restrictions (D.1).

The parameters λ, χ, and ψ are shape parameters that determine the weight in the tails. In general, the larger those parameters the closer the distribution is to the normal distribution. The parameter μ determines the location, Σ partially regulates the variance, and γ regulates the skewness. If $\gamma = 0$, then the distribution is symmetric around μ.

The $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$-parametrization is obtained as a normal mean-variance mixture distribution where the mixing distribution is the generalized inverse Gaussian distribution $\mathrm{GIG}(\lambda, \chi, \psi)$.

The GH distribution is infinitely divisible. Besides, there are closed expression for their characteristic and moment generating functions (see [14, 18] for more details).

D.3 Normal inverse Gaussian distribution

The normal inverse Gaussian (NIG) distribution is a four-parameter family of continuous probability distributions whose probability density function is given by

$$f_{\mathrm{NIG}(\alpha,\beta,\mu,\delta)}(x) = \frac{\alpha e^{\delta\gamma}}{\pi}e^{\beta(x-\mu)}\frac{K_1\left(\delta\alpha q\left(\frac{x-\mu}{\delta}\right)\right)}{q\left(\frac{x-\mu}{\delta}\right)}, \tag{D.2}$$

where $\gamma = \alpha^2 - \beta^2$, $q(x) = \sqrt{1+x^2}$ and K_1 denotes the modified Bessel function of the second kind with index 1. The domain of variation of the parameters is given by $\mu \in \mathbb{R}$, $\delta \in \mathbb{R}_+$, and $0 \leq |\beta| < \alpha$. The parameters α and β are shape parameters, μ determines the location, and δ determines the scale. The distribution is denoted by $\mathrm{NIG}(\alpha, \beta, \mu, \delta)$.

The NIG distribution is a particular case of the GH law that arises when $\lambda = 1/2$ in $f_{\mathrm{GH}(\lambda,\chi,\theta,\mu,\Sigma,\gamma)}$. However, the $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$-parametrization of the GH law produces a parametrization for the NIG distribution which differs from the above parametrization $f_{\mathrm{NIG}(\alpha,\beta,\mu,\delta)}$. The relation between these two parameterizations can be found in [14].

The cumulant function $K(z;\alpha,\beta,\mu,\delta) = \log E[\exp\{zV\}]$ of a random variable V with distribution NIG $(\alpha,\beta,\mu,\delta)$ is given by

$$K(z;\alpha,\beta,\mu,\delta) = z\mu + \delta\left(\gamma - \sqrt{\alpha^2 - (\beta+z)^2}\right). \tag{D.3}$$

It follows immediately from this that the normal inverse Gaussian distribution is infinitely divisible. Namely, if $X_i \sim$ NIG $(\alpha,\beta,\mu_i,\delta_i)$, $i = 1, 2$, are independent random variables, then we have $X_1 + X_2 \sim$ NIG$(\alpha,\beta,\mu_1+\mu_2,\delta_1+\delta_2)$.

It is often of interest to consider alternative parameterizations of the normal inverse Gaussian laws. In particular, letting $\overline{\alpha} = \delta\alpha$ and $\overline{\beta} = \delta\beta$, we have that $\overline{\alpha}$ and $\overline{\beta}$ are invariant under location-scale changes.

Sometimes it is useful to represent NIG distributions in the so-called shape triangle. Consider the alternative asymmetry and steepness parameters χ and ξ defined by

$$\xi = (1+\overline{\gamma})^{-1/2}, \quad \chi = \rho\xi,$$

where $\rho = \beta/\alpha$ and $\overline{\gamma} = \delta\gamma = \delta\sqrt{\alpha^2-\beta^2}$. These parameters are invariant under location-scale changes. Their range defines the NIG shape triangle

$$\{(\chi,\xi) : 0 < \xi < 1, -\xi < \chi < \xi\}.$$

When $\chi = 0$ the NIG distribution is symmetric. Values $\chi > 0$ indicate a positively skewed distribution and $\chi < 0$ a negatively skewed law. The steepness parameter ξ measures the heaviness of the tails of the NIG distribution. The limiting case $\xi = 0$ corresponds to a normal distribution.

The NIG law has a wide range of applications. For more details about this distribution and their applications, we refer to [1, 2, 6, 8].