Progress in Probability
72

Fabrice Baudoin
Jonathon Peterson
Editors

Stochastic Analysis and Related Topics

A Festschrift in Honor of Rodrigo Bañuelos

Birkhäuser

Birkhäuser

Progress in Probability
Volume 72

More information about this series at http://www.springer.com/series/4839

Fabrice Baudoin • Jonathon Peterson
Editors

Stochastic Analysis and Related Topics

A Festschrift in Honor of Rodrigo Bañuelos

Birkhäuser

Editors
Fabrice Baudoin
Department of Mathematics
University of Connecticut
Storrs, CT, USA

Jonathon Peterson
Department of Mathematics
Purdue University
West Lafayette, IN, USA

ISSN 1050-6977 ISSN 2297-0428 (electronic)
Progress in Probability
ISBN 978-3-319-59670-9 ISBN 978-3-319-59671-6 (eBook)
DOI 10.1007/978-3-319-59671-6

Library of Congress Control Number: 2017953194

Mathematics Subject Classification (2010): 35K05, 35K25, 60G18, 60G51, 60H07, 60H15, 82A57

Printed on acid-free paper

This book is published under the trade name Birkhäuser, www.birkhauser-science.com
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

A conference on *Stochastic Analysis and Related Topics* was held at Purdue University on May 20–22, 2015. The conference was supported by the National Science Foundation, the Institute for Mathematics and its Applications, and Purdue's Mathematics Department and College of Science. This Festschrift serves as a proceedings for this conference.

The conference was a wonderful success with 17 speakers, including international leaders in their fields, and more than 60 participants. A major reason for this success was that the conference was organized in part to serve as a celebration of the 60th birthday of Rodrigo Bañuelos. The quality of speakers in the conference and the number of attendees speak of the incredible impact (both professionally and personally) that Rodrigo has had throughout his career in mathematics. During the conference, a number of speeches and presentations were made highlighting his life story and his achievements in mathematics. Thus, while the remainder of this Festschrift will focus on the research topics discussed during the conference, it is appropriate in this preface to give a brief overview of Rodrigo's life story and career.

Rodrigo Bañuelos was born in a rural farming community of the state of Zacatecas, Mexico, to a Mexican–American father and a Mexican mother. As a child, he worked in the family farm and had no formal schooling. At the age of 15, along with his parents, grandmother, and six siblings, he moved to the USA. In spite of the late start Rodrigo had to his schooling, he went on to become a tremendous success academically. He was the first of his family (parents included) to attend high school. He enrolled at Pasadena City College (1973–1974) and received his BA in mathematics from UC Santa Cruz (1978), an MAT (master of arts in teaching) from UC Davis (1980), and then a PhD from UCLA (1984) under the direction of Richard Durrett. After graduating, he was awarded a Bantrell research fellowship at Caltech (1984–1986) and a National Science postdoctoral fellowship at the University of Illinois (1986–1987), before moving to Purdue University in 1987 where he has remained since. He served as head of the Mathematics Department at Purdue from 2007 to 2011, after serving as interim head in fall 2005.

While at Purdue, Rodrigo has supervised 11 PhD dissertations and mentored many postdocs. He has also served the profession as member of a large number of committees and editorial boards. Throughout his career, Rodrigo has been involved with many efforts to increase the number of underrepresented minority students in mathematics, and his continued commitment in this has impacted the lives of many. For his mathematical and professional contributions, Rodrigo has received several awards and recognitions, including a Young Presidential Investigator Award (1989–1993), the Blackwell-Tapia Prize in Mathematics (2004), an election to IMS fellow (2003) and AMS fellow (2013), and twice has been an AMS 1 hour invited speaker at the Joint Mathematics Meetings (1995, 2016).

Rodrigo's research is at the interface of probability, harmonic analysis, and spectral theory. All along his career, he has obtained spectacular and deep results by systematically applying martingale inequalities to various areas of analysis and in particular to L^p-estimates for singular integrals and Fourier multipliers which arise from martingale transforms. To name a few, these singular integrals and operators include the classical Hilbert transform, the Riesz transforms, the Beurling–Ahlfors operator, and versions of some of these (especially Riesz transforms) on manifolds, Lie groups, and in the setting of the Ornstein–Uhlenbeck operator in infinite dimensions. An advantage of the martingale methods is that they give sharp, or nearly sharp, bounds which are often universal in that they do not depend on the geometry or the dimension of the space where the operator is defined. These techniques have led Rodrigo (in work with his former student, Prabhu Janakiraman) to the best-known results on a longstanding conjecture of Tadeusz Iwaniec concerning the L^p-norm of the Beurling–Ahlfors operator.

Rodrigo has also made decisive contributions to other areas of probability, analysis, and spectral theory, including his work on intrinsic ultracontractivity for heat semigroups and applications to conditioned Brownian motion, his work on the hot spots conjecture of Jeff Rauch, his contributions to the spectral gap conjecture of Michiel van den Berg for Scrhödinger operators with convex potentials, and his work on eigenvalue estimates, isoperimetric inequalities, and spectral asymptotics for nonlocal operators arising as generators of Lévy processes. He has authored or coauthored over 100 papers and one book. His book, with Charles Moore, describes the many sharp estimates for the Lusin area integral and other classical functionals in harmonic analysis which led (in joint work with Ivo Klĕmes and Moore) to the solution of a problem posed by Richard Gundy in the early 1980s concerning a version of the law of the iterated logarithm for harmonic functions. The latter quantifies, in a precise way, the growth of the non-tangential maximal function in terms of the growth of the area integrals on the set where the harmonic function does not have non-tangential limits. Many of Rodrigo's papers have been published in the most demanding journals, and he is an indisputable worldwide leader in his field.

Storrs, CT, USA — Fabrice Baudoin
West Lafayette, IN, USA — Jonathon Peterson
January 30, 2017

Contents

Positive-Homogeneous Operators, Heat Kernel Estimates and the Legendre-Fenchel Transform

Evan Randles and Laurent Saloff-Coste

Dedicated to Professor Rodrigo Bañuelos on the occasion of his 60th birthday.

Abstract We consider a class of homogeneous partial differential operators on a finite-dimensional vector space and study their associated heat kernels. The heat kernels for this general class of operators are seen to arise naturally as the limiting objects of the convolution powers of complex-valued functions on the square lattice in the way that the classical heat kernel arises in the (local) central limit theorem. These so-called positive-homogeneous operators generalize the class of semi-elliptic operators in the sense that the definition is coordinate-free. More generally, we introduce a class of variable-coefficient operators, each of which is uniformly comparable to a positive-homogeneous operator, and we study the corresponding Cauchy problem for the heat equation. Under the assumption that such an operator has Hölder continuous coefficients, we construct a fundamental solution to its heat equation by the method of Levi, adapted to parabolic systems by Friedman and Eidelman. Though our results in this direction are implied by the long-known results of Eidelman for $2\vec{b}$-parabolic systems, our focus is to highlight the role played by the Legendre-Fenchel transform in heat kernel estimates. Specifically, we show that the fundamental solution satisfies an off-diagonal estimate, i.e., a heat kernel estimate, written in terms of the Legendre-Fenchel transform of the operator's principal symbol—an estimate which is seen to be sharp in many cases.

E. Randles (✉)
Department of Mathematics, University of California, Los Angeles, Box 951555, Los Angeles, CA 90025, USA
e-mail: randles@math.ucla.edu

L. Saloff-Coste
Department of Mathematics, Cornell University, Ithaca, NY 14853, USA
e-mail: lsc@math.cornell.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_1

Keywords $2\vec{b}$-parabolic operators • Heat kernel estimates • Legendre-Fenchel transform • Quasi-elliptic operators • Semi-elliptic operators

Mathematics Subject Classification Primary 35H30; Secondary 35K25

1 Introduction

In this article, we consider a class of homogeneous partial differential operators on a finite dimensional vector space and study their associated heat kernels. These operators, which we call nondegenerate-homogeneous operators, are seen to generalize the well-studied classes of semi-elliptic operators introduced by Browder [13], also known as quasi-elliptic operators [53], and a special "positive" subclass of semi-elliptic operators which appear as the spatial part of Eidelman's $2\vec{b}$-parabolic operators [26]. In particular, this class of operators contains all integer powers of the Laplacian.

1.1 Semi-Elliptic Operators

To motivate the definition of nondegenerate-homogeneous operators, given in the next section, we first introduce the class of semi-elliptic operators. Semi-elliptic operators are seen to be prototypical examples of nondegenerate-homogeneous operators; in fact, the definition of nondegenerate-homogeneous operators is given to formulate the following construction in a basis-independent way. Given d-tuple of positive integers $\mathbf{n} = (n_1, n_2, \dots, n_d) \in \mathbb{N}_+^d$ and a multi-index $\beta = (\beta_1, \beta_2, \dots, \beta_d) \in \mathbb{N}^d$, set $|\beta : \mathbf{n}| = \sum_{k=1}^d \beta_k/n_k$. Consider the constant coefficient partial differential operator

$$\Lambda = \sum_{|\beta:\mathbf{n}|\leq 1} a_\beta D^\beta$$

with principal part (relative to $\mathbf{n}$)

$$\Lambda_p = \sum_{|\beta:\mathbf{n}|=1} a_\beta D^\beta,$$

where $a_\beta \in \mathbb{C}$ and $D^\beta = (i\partial_{x_1})^{\beta_1}(i\partial_{x_2})^{\beta_2}\cdots(i\partial_{x_d})^{\beta_d}$ for each multi-index $\beta \in \mathbb{N}^d$. Such an operator Λ is said to be *semi-elliptic* if the symbol of Λ_p, defined by $P_p(\xi) = \sum_{|\beta:\mathbf{n}|=1} a_\beta \xi^\beta$ for $\xi \in \mathbb{R}^d$, is non-vanishing away from the origin. If Λ satisfies the stronger condition that $\operatorname{Re} P_p(\xi)$ is strictly positive away from the origin, we say that it is *positive-semi-elliptic*. What seems to be the most important property of semi-elliptic operators is that their principal part Λ_p is homogeneous in the following sense: If given any smooth function f we put $\delta_t(f)(x) = f(t^{1/n_1}x_1, t^{1/n_2}x_2, \ldots, t^{1/n_d}x_d)$ for all $t > 0$ and $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, then

$$t\Lambda = \delta_{1/t} \circ \Lambda_p \circ \delta_t$$

for all $t > 0$. This homogeneous structure was used explicitly in the work of Browder and Hörmander and, in this article, we generalize this notion. We note that our definition for the differential operators D^β is given to ensure a straightforward relationship between operators and symbols under our convention for the Fourier transform (defined in Sect. 1.3); this definition differs only slightly from the standard references [36, 37, 46, 48] in which i is replaced by $1/i$. In both conventions, the symbol of the operator $\Lambda = -\Delta = -\sum_{k=1}^d \partial_{x_k}^2$ is the positive polynomial $\xi \mapsto |\xi|^2 = \sum_{k=1}^d \xi_k^2$. In fact, the principal symbols of all positive-semi-elliptic operators agree in both conventions.

As mentioned above, the class of semi-elliptic operators was introduced by Browder in [13] who studied spectral asymptotics for a related class of variable-coefficient operators (operators of constant strength). Semi-elliptic operators appeared later in Hörmander's text [36] as model examples of hypoelliptic operators on $\mathbb{R}^d$ beyond the class of elliptic operators. Around the same time, Volevich [53] independently introduced the same class of operators but instead called them "quasi-elliptic". Since then, the theory of semi-elliptic operators, and hence quasi-elliptic operators, has reached a high level of sophistication and we refer the reader to the articles [1–5, 13, 34–38, 49, 51], which use the term semi-elliptic, and the articles [10–12, 14, 17–24, 31, 41, 43, 50, 52, 53], which use the term quasi-elliptic, for an account of this theory. We would also like to point to the 1971 paper of Troisi [50] which gives a more complete list of references (pertaining to quasi-elliptic operators).

Shortly after Browder's paper [13] appeared, Eidelman considered a subclass of semi-elliptic operators on $\mathbb{R}^{d+1} = \mathbb{R} \oplus \mathbb{R}^d$ (and systems thereof) of the form

$$\partial_t + \sum_{|\beta:2\mathbf{m}|\le 1} a_\beta D^\beta = \partial_t + \sum_{|\beta:\mathbf{m}|\le 2} a_\beta D^\beta, \tag{1}$$

where $\mathbf{m} \in \mathbb{N}_+^d$ and the coefficients a_β are functions of x and t. Such an operator is said to be $2\mathbf{m}$-*parabolic* if its spatial part, $\sum_{|\beta:2\mathbf{m}|\le 1} a_\beta D^\beta$, is (uniformly) positive-semi-elliptic. We note however that Eidelman's work and the existing literature refer exclusively to $2\vec{b}$-parabolic operators, i.e., where $\mathbf{m} = \vec{b}$, and for consistency we write $2\vec{b}$-parabolic henceforth [26, 28]. The relationship between positive-semi-elliptic operators and $2\vec{b}$-parabolic operators is analogous to the relationship between the Laplacian and the heat operator and, in the context of this article, the relationship between nondegenerate-homogeneous and positive-homogeneous operators described by Proposition 2.4. The theory of $2\vec{b}$-parabolic operators, which generalizes the theory of parabolic partial differential equations (and systems), has seen significant advancement by a number of mathematicians since Eidelman's original work. We encourage the reader to see the recent text [28] which provides an account of this theory and an exhaustive list of references. It should be noted however that the literature encompassing semi-elliptic operators and quasi-elliptic operators, as far as we can tell, has very few cross-references to the literature on $2\vec{b}$-parabolic operators beyond the 1960s. We suspect that the absence of cross-references is due to the distinctness of vocabulary.

1.2 Motivation: Convolution Powers of Complex-Valued Functions on $\mathbb{Z}^d$

We motivate the study of homogeneous operators by first demonstrating the natural appearance of their heat kernels in the study of convolution powers of complex-valued functions. To this end, consider a finitely supported function $\phi : \mathbb{Z}^d \to \mathbb{C}$ and define its convolution powers iteratively by

$$\phi^{(n)}(x) = \sum_{y\in\mathbb{Z}^d} \phi^{(n-1)}(x-y)\phi(y)$$

for $x \in \mathbb{Z}^d$ where $\phi^{(1)} = \phi$. In the special case that ϕ is a probability distribution, i.e., ϕ is non-negative and has unit mass, ϕ drives a random walk on $\mathbb{Z}^d$ whose nth-step transition kernels are given by $k_n(x,y) = \phi^{(n)}(y-x)$. Under certain mild conditions on the random walk, $\phi^{(n)}$ is well-approximated by a single Gaussian density; this is the classical local limit theorem. Specifically, for a symmetric, aperiodic and irreducible random walk, the theorem states that

$$\phi^{(n)}(x) = n^{-d/2} G_\phi(x/\sqrt{n}) + o(n^{-d/2}) \tag{2}$$

uniformly for $x \in \mathbb{Z}^d$, where G_ϕ is the generalized Gaussian density

$$G_\phi(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \exp\left(-\xi\cdot C_\phi \xi\right) e^{-ix\cdot\xi}\, d\xi = \frac{1}{(2\pi)^{d/2}\sqrt{\det C_\phi}} \exp\left(-\frac{x\cdot {C_\phi}^{-1}x}{2}\right); \tag{3}$$

here, C_ϕ is the positive definite covariance matrix associated to ϕ and $\cdot$ denotes the dot product [39, 44, 47]. The canonical example is that in which $C_\phi = I$ (e.g. Simple Random Walk) and in this case $\phi^{(n)}$ is approximated by the so-called heat kernel $K_{(-\Delta)} : (0,\infty)\times\mathbb{R}^d \to (0,\infty)$ defined by

$$K^t_{(-\Delta)}(x) = (2\pi t)^{-d/2} \exp\left(-\frac{|x|^2}{2t}\right)$$

for $t > 0$ and $x \in \mathbb{R}^d$. Indeed, we observe that $n^{-d/2}G_\phi(x/\sqrt{n}) = K^n_{(-\Lambda)}(x)$ for each positive integer n and $x \in \mathbb{Z}^d$ and so the local limit theorem (2) is written equivalently as

$$\phi^{(n)}(x) = K^n_{(-\Delta)}(x) + o(n^{-d/2})$$

uniformly for $x \in \mathbb{Z}^d$. In addition to its natural appearance as the *attractor* in the local limit theorem above, $K^t_{(-\Delta)}(x)$ is a fundamental solution to the heat equation

$$\partial_t + (-\Delta) = 0$$

on $(0, \infty) \times \mathbb{R}^d$. In fact, this connection to random walk underlies the heat equation's probabilistic/diffusive interpretation. Beyond the probabilistic setting, this link between convolution powers and fundamental solutions to partial differential equations persists as can be seen in the examples below. In what follows, the heat kernels $(t, x) \mapsto K^t_\Lambda(x)$ are fundamental solutions to the corresponding heat-type equations of the form

$$\partial_t + \Lambda = 0.$$

The appearance of K_Λ in local limit theorems (for $\phi^{(n)}$) is then found by evaluating $K^t_\Lambda(x)$ at integer time $t = n$ and lattice point $x \in \mathbb{Z}^d$.

Example 1.1 Consider $\phi : \mathbb{Z}^2 \to \mathbb{C}$ defined by

$$\phi(x_1, x_2) = \frac{1}{22 + 2\sqrt{3}} \times \begin{cases} 8 & (x_1, x_2) = (0, 0) \\ 5 + \sqrt{3} & (x_1, x_2) = (\pm 1, 0) \\ -2 & (x_1, x_2) = (\pm 2, 0) \\ i(\sqrt{3} - 1) & (x_1, x_2) = (\pm 1, -1) \\ -i(\sqrt{3} - 1) & (x_1, x_2) = (\pm 1, 1) \\ 2 \mp 2i & (x_1, x_2) = (0, \pm 1) \\ 0 & \text{otherwise.} \end{cases}$$

Analogous to the probabilistic setting, the large n behavior of $\phi^{(n)}$ is described by a generalized local limit theorem in which the attractor is a fundamental solution to a heat-type equation. Specifically, the following local limit theorem holds (see [44] for details):

$$\phi^{(n)}(x_1, x_2) = e^{-i\pi x_2/3} K^n_\Lambda(x_1, x_2) + o(n^{-3/4})$$

uniformly for $(x_1, x_2) \in \mathbb{Z}^2$ where $(t, x) \mapsto K_\Lambda^t(x)$ is the "heat" kernel for the heat-type equation $\partial_t + \Lambda = 0$ where

$$\Lambda = \frac{1}{22 + 2\sqrt{3}} \left(2\partial_{x_1}^4 - i(\sqrt{3} - 1)\partial_{x_1}^2 \partial_{x_2} - 4\partial_{x_2}^2 \right).$$

This local limit theorem is illustrated in Fig. 1 which shows $\mathrm{Re}(\phi^{(n)})$ and the approximation $\mathrm{Re}(e^{-i\pi x_2/3} K_\Lambda^n)$ when $n = 100$.

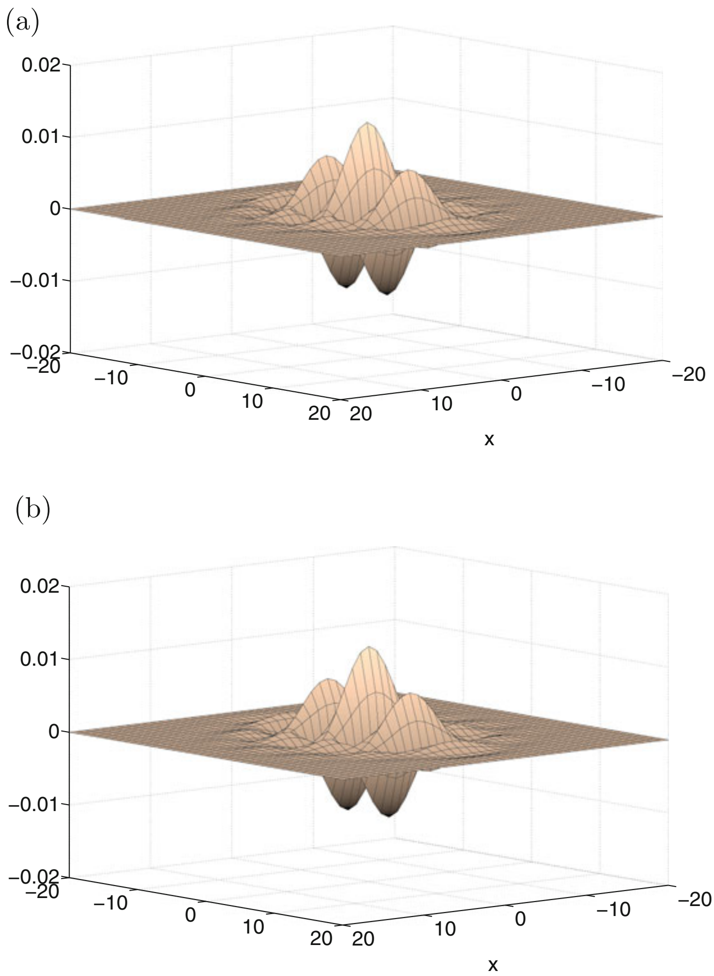

Fig. 1 The graphs of $\mathrm{Re}(\phi^{(n)})$ and $\mathrm{Re}(e^{-i\pi x_2/3} K_\Lambda^n)$ for $n = 100$. (**a**) $\mathrm{Re}(\phi^{(n)})$ for $n = 100$. (**b**) $\mathrm{Re}(e^{-i\pi x_2/3} K_\Lambda^n)$ for $n = 100$

Example 1.2 Consider $\phi : \mathbb{Z}^2 \to \mathbb{R}$ defined by $\phi = (\phi_1 + \phi_2)/512$, where

$$\phi_1(x_1,x_2) = \begin{cases} 326 & (x_1,x_2) = (0,0) \\ 20 & (x_1,x_2) = (\pm 2,0) \\ 1 & (x_1,x_2) = (\pm 4,0) \\ 64 & (x_1,x_2) = (0,\pm 1) \\ -16 & (x_1,x_2) = (0,\pm 2) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \phi_2(x_1,x_2) = \begin{cases} 76 & (x_1,x_2) = (1,0) \\ 52 & (x_1,x_2) = (-1,0) \\ \mp 4 & (x_1,x_2) = (\pm 3,0) \\ \mp 6 & (x_1,x_2) = (\pm 1,1) \\ \mp 6 & (x_1,x_2) = (\pm 1,-1) \\ \pm 2 & (x_1,x_2) = (\pm 3,1) \\ \pm 2 & (x_1,x_2) = (\pm 3,-1) \\ 0 & \text{otherwise.} \end{cases}$$

In this example, the following local limit theorem, which is illustrated by Fig. 2, describes the limiting behavior of $\phi^{(n)}$. We have

$$\phi^{(n)}(x_1,x_2) = K_\Lambda^n(x_1,x_2) + o(n^{-5/12})$$

uniformly for $(x_1,x_2) \in \mathbb{Z}^2$ where K_Λ is again a fundamental solution to $\partial_t + \Lambda = 0$ where, in this case,

$$\Lambda = \frac{1}{64}\left(-\partial_{x_1}^6 + 2\partial_{x_2}^4 + 2\partial_{x_1}^3\partial_{x_2}^2\right).$$

Example 1.3 Consider $\phi : \mathbb{Z}^2 \to \mathbb{R}$ defined by

$$\phi(x,y) = \begin{cases} 3/8 & (x_1,x_2) = (0,0) \\ 1/8 & (x_1,x_2) = \pm(1,1) \\ 1/4 & (x_1,x_2) = \pm(1,-1) \\ -1/16 & (x_1,x_2) = \pm(2,-2) \\ 0 & \text{otherwise.} \end{cases}$$

Here, the following local limit theorem is valid:

$$\phi^{(n)}(x_1,x_2) = \left(1 + e^{i\pi(x_1+x_2)}\right) K_\Lambda^n(x_1,x_2) + o(n^{-3/4})$$

uniformly for $(x_1,x_2) \in \mathbb{Z}^2$. Here again, the attractor K_Λ is the fundamental solution to $\partial_t + \Lambda = 0$ where

$$\Lambda = -\frac{1}{8}\partial_{x_1}^2 + \frac{23}{384}\partial_{x_1}^4 - \frac{1}{4}\partial_{x_1}\partial_{x_2} - \frac{25}{96}\partial_{x_1}^3\partial_{x_2} - \frac{1}{8}\partial_{x_2}^2 + \frac{23}{64}\partial_{x_1}^2\partial_{x_2}^2 - \frac{25}{96}\partial_{x_1}\partial_{x_2}^3 + \frac{23}{384}\partial_{x_2}^4.$$

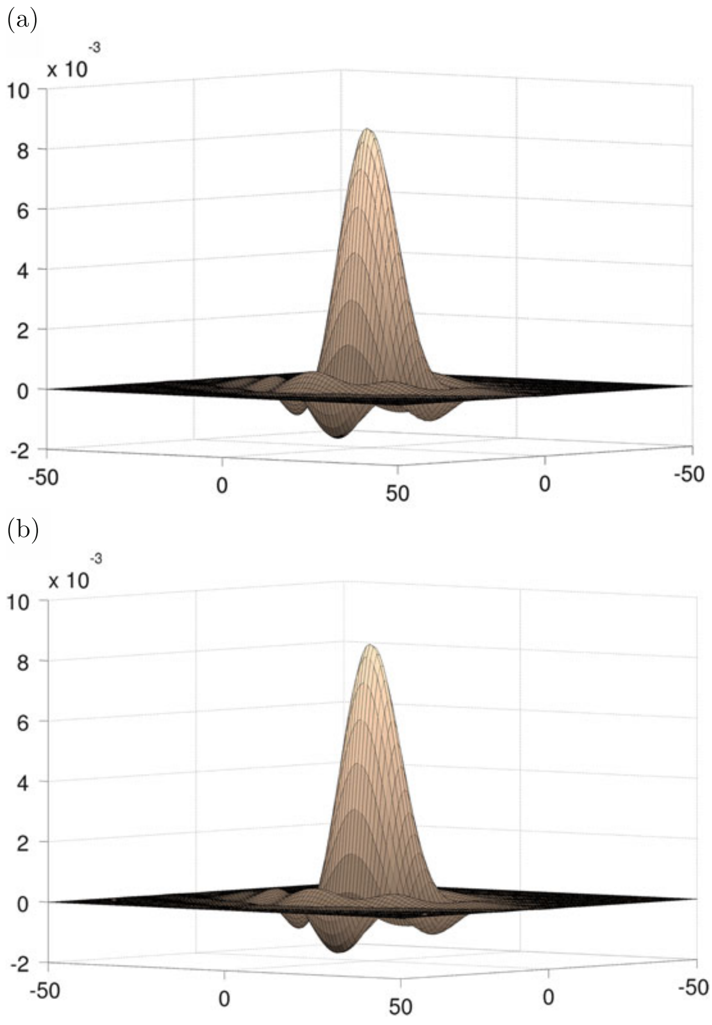

Fig. 2 The graphs of $\phi^{(n)}$ and K_Λ^n for $n = 10{,}000$. (**a**) $\phi^{(n)}$ for $n = 10{,}000$. (**b**) K_Λ^n for $n = 10{,}000$

Looking back at preceding examples, we note that the operators appearing in Examples 1.1 and 1.2 are both positive-semi-elliptic and consist only of their principal parts. This is easily verified, for $\mathbf{n} = (4, 2) = 2(2, 1)$ in Example 1.1 and $\mathbf{n} = (6, 4) = 2(3, 2)$ in Example 1.2. In contrast to Examples 1.1 and 1.2, the operator Λ which appears in Example 1.3 is not semi-elliptic in the given coordinate system. After careful study, the Λ appearing in Example 1.3 can be written equivalently as

$$\Lambda = -\frac{1}{8}\partial_{v_1}^2 + \frac{23}{384}\partial_{v_2}^4 \tag{4}$$

where ∂_{v_1} is the directional derivative in the $v_1 = (1, 1)$ direction and ∂_{v_2} is the directional derivative in the $v_2 = (1, -1)$ direction. In this way, Λ is seen to be semi-elliptic with respect to some basis $\{v_1, v_2\}$ of $\mathbb{R}^2$ and, with respect to this basis, we have $\mathbf{n} = (2, 4) = 2(1, 2)$. For this reason, our formulation of nondegenerate-homogeneous operators (and positive-homogeneous operators), given in the next section, is made in a basis-independent way.

All of the operators appearing in Examples 1.1, 1.2 and 1.3 share two important properties: homogeneity and positivity (in the sense of symbols). While we make these notions precise in the next section, loosely speaking, homogeneity is the property that Λ "plays well" with some dilation structure on $\mathbb{R}^d$, though this structure is different in each example. Further, homogeneity for Λ is reflected by an analogous one for the corresponding heat kernel K_Λ; in fact, the specific dilation structure is, in some sense, selected by $\phi^{(n)}$ as $n \to \infty$ and leads to the corresponding local limit theorem. In further discussion of these examples, a very natural question arises: Given $\phi : \mathbb{Z}^d \to \mathbb{C}$, how does one compute the operator Λ whose heat kernel K_Λ appears as the attractor in the local limit theorem for $\phi^{(n)}$? In the examples we have looked at, one studies the Taylor expansion of the Fourier transform $\hat{\phi}$ of ϕ near its local extrema and, here, the symbol of the relevant operator Λ appears as certain scaled limit of this Taylor expansion. In general, however, this is a very delicate business and, at present, there is no known algorithm to determine these operators. In fact, it is possible that multiple (distinct) operators can appear by looking at the Taylor expansions about distinct local extrema of $\hat{\phi}$ (when they exist) and, in such cases, the corresponding local limit theorems involve sums of heat kernels—each corresponding to a distinct Λ. This study is carried out in the article [44] wherein local limit theorems involve the heat kernels of the positive-homogeneous operators studied in the present article. We note that the theory presented in [44] is not complete, for there are cases in which the associated Taylor approximations yield symbols corresponding to operators Λ which fail to be positive-homogeneous (and hence fail to be positive-semi-elliptic) and further, the heat kernels of these (degenerate) operators appear as limits of oscillatory integrals which correspond to the presence of "odd" terms in Λ, e.g., the Airy function. In one dimension, a complete theory of local limit theorems is known for the class of finitely supported functions $\phi : \mathbb{Z} \to \mathbb{C}$. Beyond one dimension, a theory for local limit theorems of complex-valued functions, in which the results of [44] will fit, remains open.

The subject of this paper is an account of positive-homogeneous operators and their corresponding heat equations. In Sect. 2, we introduce positive-homogeneous operators and study their basic properties; therein, we show that each positive-homogeneous operator is semi-elliptic in some coordinate system. Section 3 develops the necessary background to introduce the class of variable-coefficient operators studied in this article; this is the class of $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic operators introduced in Sect. 4—each of which is comparable to a constant-coefficient positive-homogeneous operator. In Sect. 5, we study the heat equations corresponding to uniformly $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic operators with Hölder continuous coefficients. Specifically, we use the famous method of Levi, adapted to parabolic systems by Friedman and Eidelman, to construct a fundamental solution to the corresponding heat equation. Our results in this direction are captured by those of Eidelman [26] and the works of his collaborators, notably Ivasyshen and Kochubei [28], concerning $2\vec{b}$-parabolic systems. Our focus in this presentation is to highlight the essential role played by the Legendre-Fenchel transform in heat kernel estimates which, to our knowledge, has not been pointed out in the context of semi-elliptic operators. In a forthcoming work, we study an analogous class of operators, written in divergence form, with measurable-coefficients and their corresponding heat kernels. This class of measurable-coefficient operators does not appear to have been previously studied. The results presented here, using the Legendre-Fenchel transform, provides the background and context for our work there.

1.3 Preliminaries

Fourier Analysis Our setting is a real d-dimensional vector space $\mathbb{V}$ equipped with Haar (Lebesgue) measure dx and the standard smooth structure; we do not affix $\mathbb{V}$ with a norm or basis. The dual space of $\mathbb{V}$ is denoted by $\mathbb{V}^*$ and the dual pairing is denoted by $\xi(x)$ for $x \in \mathbb{V}$ and $\xi \in \mathbb{V}^*$. Let $d\xi$ be the Haar measure on $\mathbb{V}^*$ which we take to be normalized so that our convention for the Fourier transform and inverse Fourier transform, given below, makes each unitary. Throughout this article, all functions on $\mathbb{V}$ and $\mathbb{V}^*$ are understood to be complex-valued. The usual Lebesgue spaces are denoted by $L^p(\mathbb{V}) = L^p(\mathbb{V}, dx)$ and equipped with their usual norms $\|\cdot\|_p$ for $1 \leq p \leq \infty$. In the case that $p = 2$, the corresponding inner product on $L^2(\mathbb{V})$ is denoted by $\langle\cdot,\cdot\rangle$. Of course, we will also work with $L^2(\mathbb{V}^*) := L^2(\mathbb{V}^*, d\xi)$; here the L^2-norm and inner product will be denoted by $\|\cdot\|_{2^*}$ and $\langle\cdot,\cdot\rangle_*$ respectively. The Fourier transform $\mathcal{F} : L^2(\mathbb{V}) \to L^2(\mathbb{V}^*)$ and inverse Fourier transform $\mathcal{F}^{-1}$: $L^2(\mathbb{V}^*) \to L^2(\mathbb{V})$ are initially defined for Schwartz functions $f \in \mathcal{S}(\mathbb{V})$ and $g \in \mathcal{S}(\mathbb{V}^*)$ by

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{V}} e^{i\xi(x)} f(x)\, dx \quad \text{and} \quad \mathcal{F}^{-1}(g)(x) = \check{g}(x) = \int_{\mathbb{V}^*} e^{-i\xi(x)} g(\xi)\, d\xi$$

for $\xi \in \mathbb{V}^*$ and $x \in \mathbb{V}$ respectively.

For the remainder of this article (mainly when duality isn't of interest), W stands for any real d-dimensional vector space (and so is interchangeable with $\mathbb{V}$ or $\mathbb{V}^*$). For a non-empty open set $\Omega \subseteq W$, we denote by $C(\Omega)$ and $C_b(\Omega)$ the set of continuous functions on Ω and bounded continuous functions on Ω, respectively. The set of smooth functions on Ω is denoted by $C^\infty(\Omega)$ and the set of compactly supported smooth functions on Ω is denoted by $C_0^\infty(\Omega)$. We denote by $\mathcal{D}'(\Omega)$ the space of distributions on Ω; this is dual to the space $C_0^\infty(\Omega)$ equipped with its usual topology given by seminorms. A partial differential operator H on W is said to be *hypoelliptic* if it satisfies the following property: Given any open set $\Omega \subseteq W$ and any distribution $u \in \mathcal{D}'(\Omega)$ which satisfies $Hu = 0$ in Ω, then necessarily $u \in C^\infty(\Omega)$.

Dilation Structure Denote by End(W) and Gl(W) the set of endomorphisms and isomorphisms of W respectively. Given $E \in \mathrm{End}(W)$, we consider the one-parameter group $\{t^E\}_{t>0} \subseteq \mathrm{Gl}(W)$ defined by

$$t^E = \exp((\log t)E) = \sum_{k=0}^{\infty} \frac{(\log t)^k}{k!} E^k$$

for $t > 0$. These one-parameter subgroups of Gl(W) allow us to define continuous one-parameter groups of operators on the space of distributions as follows: Given $E \in \mathrm{End}(W)$ and $t > 0$, first define $\delta_t^E(f)$ for $f \in C_0^\infty(W)$ by $\delta_t^E(f)(x) = f(t^E x)$ for $x \in W$. Extending this to the space of distribution on W in the usual way, the collection $\{\delta_t^E\}_{t>0}$ is a continuous one-parameter group of operators on $\mathcal{D}'(W)$; it will allow us to define homogeneity for partial differential operators in the next section.

Linear Algebra, Polynomials and the Rest Given a basis $\mathbf{w} = \{w_1, w_2, \dots, w_d\}$ of W, we define the map $\phi_{\mathbf{w}} : W \to \mathbb{R}^d$ by setting $\phi_{\mathbf{w}}(w) = (x_1, x_2, \dots, x_d)$ whenever $w = \sum_{l=1}^d x_l w_l$. This map defines a global coordinate system on W; any such coordinate system is said to be a linear coordinate system on W. By definition, a polynomial on W is a function $P : W \to \mathbb{C}$ that is a polynomial function in every (and hence any) linear coordinate system on W. A polynomial P on W is called a nondegenerate polynomial if $P(w) \neq 0$ for all $w \neq 0$. Further, P is called a positive-definite polynomial if its real part, $R = \mathrm{Re}\, P$, is non-negative and has $R(w) = 0$ only when $w = 0$. The symbols $\mathbb{R}, \mathbb{C}, \mathbb{Z}$ mean what they usually do, $\mathbb{N}$ denotes the set of non-negative integers and $\mathbb{I} = [0, 1] \subseteq \mathbb{R}$. The symbols $\mathbb{R}_+$, $\mathbb{N}_+$ and $\mathbb{I}_+$ denote the set of strictly positive elements of $\mathbb{R}$, $\mathbb{N}$ and $\mathbb{I}$ respectively. Likewise, $\mathbb{R}_+^d$, $\mathbb{N}_+^d$ and $\mathbb{I}_+^d$ respectively denote the set of d-tuples of these aforementioned sets. Given $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{R}_+^d$ and a basis $\mathbf{w} = \{w_1, w_2, \dots, w_d\}$ of W, we denote by $E_{\mathbf{w}}^{\alpha}$ the isomorphism of W defined by

$$E_{\mathbf{w}}^{\alpha} w_k = \frac{1}{\alpha_k} w_k \tag{5}$$

for $k = 1, 2, \dots, d$. We say that two real-valued functions f and g on a set X are comparable if, for some positive constant C, $C^{-1} f(x) \leq g(x) \leq C f(x)$ for all $x \in X$; in this case we write $f \asymp g$. Adopting the summation notation for semi-elliptic operators of Hörmander's treatise [37], for a fixed $\mathbf{n} = (n_1, n_2, \dots, n_d) \in \mathbb{N}_+^d$, we write

$$|\beta : \mathbf{n}| = \sum_{k=1}^{d} \frac{\beta_k}{m_k}$$

for all multi-indices $\beta = (\beta_1, \beta_2, \ldots, \beta_d) \in \mathbb{N}^d$. Finally, throughout the estimates made in this article, constants denoted by C will change from line to line without explicit mention.

2 Homogeneous Operators

In this section we introduce two important classes of homogeneous constant-coefficient on $\mathbb{V}$. These operators will serve as "model" operators in our theory in the way that integer powers of the Laplacian serves a model operators in the elliptic theory of partial differential equations. To this end, let Λ be a constant-coefficient partial differential operator on $\mathbb{V}$ and let $P : \mathbb{V}^* \to \mathbb{C}$ be its symbol. Specifically, P is the polynomial on $\mathbb{V}^*$ defined by $P(\xi) = e^{-i\xi(x)}\Lambda(e^{i\xi(x)})$ for $\xi \in \mathbb{V}^*$ (this is independent of $x \in \mathbb{V}$ precisely because Λ is a constant-coefficient operator). We first introduce the following notion of homogeneity of operators; it is mirrored by an analogous notion for symbols which we define shortly.

Definition 2.1 Given $E \in \mathrm{End}(\mathbb{V})$, we say that a constant-coefficient partial differential operator Λ is homogeneous with respect to the one-parameter group $\{\delta_t^E\}$ if

$$\delta_{1/t}^E \circ \Lambda \circ \delta_t^E = t\Lambda$$

for all $t > 0$; in this case we say that E is a member of the exponent set of Λ and write $E \in \mathrm{Exp}(\Lambda)$.

A constant-coefficient partial differential operator Λ need not be homogeneous with respect to a unique one-parameter group $\{\delta_t^E\}$, i.e., $\mathrm{Exp}(\Lambda)$ is not necessarily a singleton. For instance, it is easily verified that, for the Laplacian $-\Delta$ on $\mathbb{R}^d$,

$$\mathrm{Exp}(-\Delta) = 2^{-1}I + \mathfrak{o}_d$$

where I is the identity and $\mathfrak{o}_d$ is the Lie algebra of the orthogonal group, i.e., is given by the set of skew-symmetric matrices. Despite this lack of uniqueness, when Λ is equipped with a nondegenerateness condition (see Definition 2.2), we will find that trace is the same for each member of $\mathrm{Exp}(\Lambda)$ and this allows us to uniquely define an "order" for Λ; this is Lemma 2.10.

Given a constant coefficient operator Λ with symbol P, one can quickly verify that $E \in \mathrm{Exp}(\Lambda)$ if and only if

$$tP(\xi) = P(t^F\xi) \tag{6}$$

for all $t > 0$ and $\xi \in \mathbb{V}^*$ where $F = E^*$ is the adjoint of E. More generally, if P is any continuous function on W and (6) is satisfied for some $F \in \mathrm{End}(\mathbb{V}^*)$, we say that P *is homogeneous with respect to* $\{t^F\}$ and write $F \in \mathrm{Exp}(P)$. This admitted slight abuse of notation should not cause confusion. In this language, we see that $E \in \mathrm{Exp}(\Lambda)$ if and only if $E^* \in \mathrm{Exp}(P)$.

We remark that the notion of homogeneity defined above is similar to that put forth for homogeneous operators on homogeneous (Lie) groups, e.g., Rockland operators [29]. The difference is mostly a matter of perspective: A homogeneous group G is equipped with a fixed dilation structure, i.e., it comes with a one-parameter group $\{\delta_t\}$, and homogeneity of operators is defined with respect to this fixed dilation structure. By contrast, we fix no dilation structure on $\mathbb{V}$ and formulate homogeneity in terms of an operator Λ and the existence of a one-parameter group $\{\delta_t^E\}$ that "plays" well with Λ in sense defined above. As seen in the study of convolution powers on the square lattice (see [44]), it useful to have this freedom.

Definition 2.2 Let Λ be constant-coefficient partial differential operator on $\mathbb{V}$ with symbol P. We say that Λ is a nondegenerate-homogeneous operator if P is a nondegenerate polynomial and $\mathrm{Exp}(\Lambda)$ contains a diagonalizable endomorphism. We say that Λ is a positive-homogeneous operator if P is a positive-definite polynomial and $\mathrm{Exp}(\Lambda)$ contains a diagonalizable endomorphism.

For any polynomial P on a finite-dimensional vector space W, P is said to be *nondegenerate-homogeneous* if P is nondegenerate and $\mathrm{Exp}(P)$, defined as the set of $F \in \mathrm{End}(W)$ for which (6) holds, contains a diagonalizable endomorphism. We say that P is *positive-homogeneous* if it is a positive-definite polynomial and $\mathrm{Exp}(P)$ contains a diagonalizable endomorphism. In this language, we have the following proposition.

Proposition 2.3 *Let Λ be a positive homogeneous operator on $\mathbb{V}$ with symbol P. Then Λ is a nondegenerate-homogeneous operator if and only if P is a nondegenerate-homogeneous polynomial. Further, Λ is a positive-homogeneous operator if and only if P is a positive-homogeneous polynomial.*

Proof Since the adjectives "nondegenerate" and "positive", in the sense of both operators and polynomials, are defined in terms of the symbol P, all that needs to be verified is that $\mathrm{Exp}(\Lambda)$ contains a diagonalizable endomorphism if and only if $\mathrm{Exp}(P)$ contains a diagonalizable endomorphism. Upon recalling that $E \in \mathrm{Exp}(\Lambda)$ if and only if $E^* \in \mathrm{Exp}(P)$, this equivalence is verified by simply noting that diagonalizability is preserved under taking adjoints. □

Remark 1 To capture the class of nondegenerate-homogeneous operators (or positive-homogeneous operators), in addition to requiring that the symbol P of an operator Λ be nondegenerate (or positive-definite), one can instead demand only that $\mathrm{Exp}(\Lambda)$ contains an endomorphism whose characteristic polynomial factors over $\mathbb{R}$ or, equivalently, whose spectrum is real. This a priori weaker condition is seen to be sufficient by an argument which makes use of the Jordan-Chevalley decomposition. In the positive-homogeneous case, this argument is carried out in [44] (specifically Proposition 2.2) wherein positive-homogeneous operators are first defined by this (a priori weaker) condition. For the nondegenerate case, the same argument pushes through with very little modification.

We observe easily that all positive-homogeneous operators are nondegenerate-homogeneous. It is the "heat" kernels corresponding to positive-homogeneous operators that naturally appear in [44] as the attractors of convolution powers of complex-valued functions. The following proposition highlights the interplay between positive-homogeneity and nondegenerate-homogeneity for an operator Λ on $\mathbb{V}$ and its corresponding "heat" operator $\partial_t + \Lambda$ on $\mathbb{R} \oplus \mathbb{V}$.

Proposition 2.4 *Let Λ be a constant-coefficient partial differential operator on $\mathbb{V}$ whose exponent set* $\operatorname{Exp}(\Lambda)$ *contains a diagonalizable endomorphism. Let P be the symbol of Λ, set $R = \operatorname{Re} P$, and assume that there exists $\xi \in \mathbb{V}^*$ for which $R(\xi) > 0$. We have the following dichotomy: Λ is a positive-homogeneous operator on $\mathbb{V}$ if and only if $\partial_t + \Lambda$ is a nondegenerate-homogeneous operator on $\mathbb{R} \oplus \mathbb{V}$.*

Proof Given a diagonalizable endomorphism $E \in \operatorname{Exp}(\Lambda)$, set $E_1 = I \oplus E$ where I is the identity on $\mathbb{R}$. Obviously, E_1 is diagonalizable. Further, for any $f \in C_0^\infty(\mathbb{R} \oplus \mathbb{V})$,

$$\begin{aligned}\left((\partial_t + \Lambda) \circ \delta_s^{E_1}\right)(f)(t,x) &= \left(\partial_t \left(f\left(st, s^E x\right)\right) + \Lambda\left(f\left(st, s^E x\right)\right)\right)\\ &= s(\partial_t + \Lambda)(f)(st, s^E x) = s\left(\delta_s^{E_1} \circ (\partial_t + \Lambda)\right)(f)(t,x)\end{aligned}$$

for all $s > 0$ and $(t,x) \in \mathbb{R} \oplus \mathbb{V}$. Hence

$$\delta_{1/s}^{E_1} \circ (\partial_t + \Lambda) \circ \delta_s^{E_1} = s(\partial_t + \Lambda)$$

for all $s > 0$ and therefore $E_1 \in \operatorname{Exp}(\partial_t + \Lambda)$.

It remains to show that P is positive-definite if and only if the symbol of $\partial_t + \Lambda$ is nondegenerate. To this end, we first compute the symbol of $\partial_t + \Lambda$ which we denote by Q. Since the dual space of $\mathbb{R} \oplus \mathbb{V}$ is isomorphic to $\mathbb{R} \oplus \mathbb{V}^*$, the characters of $\mathbb{R} \oplus \mathbb{V}$ are represented by the collection of maps $(\mathbb{R} \oplus \mathbb{V}) \ni (t,x) \mapsto \exp(-i(\tau t + \xi(x)))$ where $(\tau, \xi) \in \mathbb{R} \oplus \mathbb{V}^*$. Consequently,

$$Q(\tau,\xi) = e^{-i(\tau t + \xi(x))}(\partial_t + \Lambda)(e^{i(\tau t + \xi(x))}) = i\tau + P(\xi)$$

for $(\tau, \xi) \in \mathbb{R} \oplus \mathbb{V}^*$. We note that $P(0) = 0$ because $E^* \in \operatorname{Exp}(P)$; in fact, this happens whenever $\operatorname{Exp}(P)$ is non-empty. Now if P is a positive-definite polynomial, $\operatorname{Re} Q(\tau,\xi) = \operatorname{Re} P(\xi) = R(\xi) > 0$ whenever $\xi \neq 0$. Thus to verify that Q is a nondegenerate polynomial, we simply must verify that $Q(\tau, 0) \neq 0$ for all non-zero $\tau \in \mathbb{R}$. This is easy to see because, in light of the above fact, $Q(\tau, 0) = i\tau + P(0) = i\tau \neq 0$ whenever $\tau \neq 0$ and hence Q is nondegenerate. For the other direction, we demonstrate the validity of the contrapositive statement. Assuming that P is not positive-definite, an application of the intermediate value theorem, using the condition that $R(\xi) > 0$ for some $\xi \in \mathbb{V}^*$, guarantees that $R(\eta) = 0$ for some non-zero $\eta \in \mathbb{V}^*$. Here, we observe that $Q(\tau, \eta) = i(\tau + \operatorname{Im} P(\eta)) = 0$ when $(\tau, \eta) = (-\operatorname{Im} P(\eta), \eta)$ and hence Q is not nondegenerate. □

We will soon return to the discussion surrounding a positive-homogeneous operator Λ and its heat operator $\partial_t + \Lambda$. It is useful to first provide representation formulas for nondegenerate-homogeneous and positive-homogeneous operators. Such representations connect our homogeneous operators to the class of semi-elliptic operators discussed in the introduction. To this end, we define the "base" operators on $\mathbb{V}$. First, for any element $u \in \mathbb{V}$, we consider the differential operator $D_u : \mathcal{D}'(\mathbb{V}) \to \mathcal{D}'(\mathbb{V})$ defined originally for $f \in C_0^\infty(\mathbb{V})$ by

$$(D_u f)(x) = i\frac{\partial f}{\partial u}(x) = i\left(\lim_{t\to 0}\frac{f(x+tu)-f(x)}{t}\right)$$

for $x \in \mathbb{V}$. Fixing a basis $\mathbf{v} = \{v_1, v_2, \ldots, v_d\}$ of $\mathbb{V}$, we introduce, for each multi-index $\beta \in \mathbb{N}^d$, $D_{\mathbf{v}}^\beta = (D_{v_1})^{\beta_1}(D_{v_2})^{\beta_2}\cdots(D_{v_d})^{\beta_d}$.

Proposition 2.5 *Let Λ be a nondegenerate-homogeneous operator on $\mathbb{V}$. Then there exist a basis $\mathbf{v} = \{v_1, v_2, \ldots, v_d\}$ of $\mathbb{V}$ and $\mathbf{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{N}_+^d$ for which*

$$\Lambda = \sum_{|\beta:\mathbf{n}|=1} a_\beta D_{\mathbf{v}}^\beta. \tag{7}$$

where $\{a_\beta\} \subseteq \mathbb{C}$. The isomorphism $E_{\mathbf{v}}^{\mathbf{n}} \in Gl(\mathbb{V})$, defined by (5), is a member of $\operatorname{Exp}(\Lambda)$. *Further, if Λ is positive-homogeneous, then $\mathbf{n} = 2\mathbf{m}$ for $\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{N}_+^d$ and hence*

$$\Lambda = \sum_{|\beta:\mathbf{m}|=2} a_\beta D_{\mathbf{v}}^\beta.$$

We will sometimes refer to the $\mathbf{n}$ and $\mathbf{m}$ of the proposition as *weights*. Before addressing the proposition, we first prove the following mirrored result for symbols.

Lemma 2.6 *Let P be a nondegenerate-homogeneous polynomial on a d-dimensional real vector space W. Then there exists a basis $\mathbf{w} = \{w_1, w_2, \ldots, w_d\}$ of W and $\mathbf{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{N}_+^d$ for which*

$$P(\xi) = \sum_{|\beta:\mathbf{n}|=1} a_\beta \xi^\beta$$

for all $\xi = \xi_1 w_1 + \xi_2 w_2 + \cdots + \xi_d w_d \in W$ where $\xi^\beta := (\xi_1)^{\beta_1}(\xi_2)^{\beta_2}\cdots(\xi_d)^{\beta_d}$ and $\{a_\beta\} \subseteq \mathbb{C}$. The isomorphism $E_{\mathbf{w}}^{\mathbf{n}} \in Gl(\mathbb{V})$, defined by (5), is a member of $\operatorname{Exp}(P)$. *Further, if P is a positive-definite polynomial, i.e., it is positive-homogeneous, then $\mathbf{n} = 2\mathbf{m}$ for $\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{N}_+^d$ and hence*

$$P(\xi) = \sum_{|\beta:\mathbf{m}|=2} a_\beta \xi^\beta$$

for $\xi \in W$.

Proof Let $E \in \operatorname{Exp}(P)$ be diagonalizable and select a basis $\mathbf{w} = \{w_1, w_2, \ldots, w_d\}$ which diagonalizes E, i.e., $Ew_k = \delta_k w_k$ where $\delta_k \in \mathbb{R}$ for $k = 1, 2, \ldots, d$. Because P is a polynomial, there exists a finite collection $\{a_\beta\} \subseteq \mathbb{C}$ for which

$$P(\xi) = \sum_\beta a_\beta \xi^\beta$$

for $\xi \in W$. By invoking the homogeneity of P with respect to E and using the fact that $t^E w_k = t^{\delta_k} w_k$ for $k = 1, 2, \ldots, d$, we have

$$t \sum_\beta a_\beta \xi^\beta = \sum_\beta a_\beta (t^E \xi)^\beta = \sum_\beta a_\beta t^{\delta \cdot \beta} \xi^\beta$$

for all $\xi \in W$ and $t > 0$ where $\delta \cdot \beta = \delta_1 \beta_1 + \delta_2 \beta_2 + \cdots + \delta_d \beta_d$. In view of the nondegenerateness of P, the linear independence of distinct powers of t and the polynomial functions $\xi \mapsto \xi^\beta$, for distinct multi-indices β, as C^∞ functions ensures that $a_\beta = 0$ unless $\beta \cdot \delta = 1$. We can therefore write

$$P(\xi) = \sum_{\beta \cdot \delta = 1} a_\beta \xi^\beta \tag{8}$$

for $\xi \in W$. We now determine $\delta = (\delta_1, \delta_2, \ldots, \delta_d)$ by evaluating this polynomial along the coordinate axes. To this end, by fixing $k = 1, 2, \ldots, d$ and setting $\xi = x w_k$ for $x \in \mathbb{R}$, it is easy to see that the summation above collapses into a single term $a_\beta x^{|\beta|}$ where $\beta = |\beta| e_k = (1/\delta_k) e_k$ (here e_k denotes the usual kth-Euclidean basis vector in $\mathbb{R}^d$). Consequently, $n_k := 1/\delta_k \in \mathbb{N}_+$ for $k = 1, 2, \ldots, d$ and thus, upon setting $\mathbf{n} = (n_1, n_2, \ldots, n_d)$, (8) yields

$$P(\xi) = \sum_{|\beta : \mathbf{n}| = 1} a_\beta \xi^\beta$$

for all $\xi \in W$ as was asserted. In this notation, it is also evident that $E_{\mathbf{w}}^{\mathbf{n}} = E \in \operatorname{Exp}(P)$. Under the additional assumption that P is positive-definite, we again evaluate P at the coordinate axes to see that $\operatorname{Re} P(x w_k) = \operatorname{Re}(a_{n_k e_k}) x^{n_k}$ for $x \in \mathbb{R}$. In this case, the positive-definiteness of P requires $\operatorname{Re}(a_{n_k e_k}) > 0$ and $n_k \in 2\mathbb{N}_+$ for each $k = 1, 2, \ldots, d$. Consequently, $\mathbf{n} = 2\mathbf{m}$ for $\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{N}_+^d$ as desired. □

Proof of Proposition 2.5 Given a nondegenerate-homogeneous Λ on $\mathbb{V}$ with symbol P, P is necessarily a nondegenerate-homogeneous polynomial on $\mathbb{V}^*$ in view of Proposition 2.3. We can therefore apply Lemma 2.6 to select a basis $\mathbf{v}^* = \{v_1^*, v_2^*, \ldots, v_d^*\}$ of $\mathbb{V}^*$ and $\mathbf{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{N}_+^d$ for which

$$P(\xi) = \sum_{|\beta : \mathbf{n}| = 1} a_\beta \xi^\beta \tag{9}$$

for all $\xi = \xi_1 v_1^* + \xi_2 v_2^* + \cdots \xi_d v_d^*$ where $\{a_\beta\} \subseteq \mathbb{C}$. We will denote by $\mathbf{v}$, the dual basis to $\mathbf{v}^*$, i.e., $\mathbf{v} = \{v_1, v_2, \ldots, v_d\}$ is the unique basis of $\mathbb{V}$ for which $v_k^*(v_l) = 1$ when $k = l$ and 0 otherwise. In view of the duality of the bases $\mathbf{v}$ and $\mathbf{v}^*$, it is straightforward to verify that, for each multi-index β, the symbol of $D_{\mathbf{v}}^\beta$ is ξ^β in the notation of Lemma 2.6. Consequently, the constant-coefficient partial differential operator defined by the right hand side of (7) also has symbol P and so it must be equal to Λ because operators and symbols are in one-to-one correspondence. Using (7), it is now straightforward to verify that $E_{\mathbf{v}}^{\mathbf{n}} \in \operatorname{Exp}(\Lambda)$. The assertion that $\mathbf{n} = 2\mathbf{m}$ when Λ is positive-homogeneous follows from the analogous conclusion of Lemma 2.6 by the same line of reasoning. □

In view of Proposition 2.5, we see that all nondegenerate-homogeneous operators are semi-elliptic in some linear coordinate system (that which is defined by $\mathbf{v}$). An appeal to Theorem 11.1.11 of [37] immediately yields the following corollary.

Corollary 2.7 *Every nondegenerate-homogeneous operator Λ on $\mathbb{V}$ is hypoelliptic.*

Our next goal is to associate an "order" to each nondegenerate-homogeneous operator. For a positive-homogeneous operator Λ, this order will be seen to govern the on-diagonal decay of its heat kernel K_Λ and so, equivalently, the ultracontractivity of the semigroup $e^{-t\Lambda}$. With the help of Lemma 2.6, the few lemmas in this direction come easily.

Lemma 2.8 *Let P be a nondegenerate-homogeneous polynomial on a d-dimensional real vector space W. Then $\lim_{\xi\to\infty} |P(\xi)| = \infty$; here $\xi \to \infty$ means that $|\xi| \to \infty$ in any (and hence every) norm on W.*

Proof The idea of the proof is to construct a function which bounds $|P|$ from below and obviously blows up at infinity. To this end, let $\mathbf{w}$ be a basis for W and take $\mathbf{n} \in \mathbb{N}_+^d$ as guaranteed by Lemma 2.6; we have $E_{\mathbf{w}}^{\mathbf{n}} \in \operatorname{Exp}(P)$ where $E_{\mathbf{w}}^{\mathbf{n}} w_k = (1/n_k) w_k$ for $k = 1, 2, \ldots, d$. Define $|\cdot|_{\mathbf{w}}^{\mathbf{n}} : W \to [0, \infty)$ by

$$|\xi|_{\mathbf{w}}^{\mathbf{n}} = \sum_{k=1}^{d} |\xi_k|^{n_k}$$

where $\xi = \xi_1 w_1 + \xi_2 w_2 + \cdots + \xi_d w_d \in W$. We observe immediately $E_{\mathbf{w}}^{\mathbf{n}} \in \operatorname{Exp}(|\cdot|_{\mathbf{w}}^{\mathbf{n}})$ because $t^{E_{\mathbf{w}}^{\mathbf{n}}} w_k = t^{1/n_k} w_k$ for $k = 1, 2, \ldots, d$. An application of Proposition 3.2 (a basic result appearing in our background section, Sect. 3), which uses the nondegenerateness of P, gives a positive constant C for which $|\xi|_{\mathbf{w}}^{\mathbf{n}} \leq C|P(\xi)|$ for all $\xi \in W$. The lemma now follows by simply noting that $|\xi|_{\mathbf{w}}^{\mathbf{n}} \to \infty$ as $\xi \to \infty$. □

Lemma 2.9 *Let P be a polynomial on W and denote by* $\operatorname{Sym}(P)$ *the set of $O \in End(W)$ for which $P(O\xi) = P(\xi)$ for all $\xi \in W$. If P is a nondegenerate-homogeneous polynomial, then* $\operatorname{Sym}(P)$*, called the symmetry group of P, is a compact subgroup of $Gl(W)$.*

Proof Our supposition that P is a nondegenerate polynomial ensures that, for each $O \in \operatorname{Sym}(P)$, $\operatorname{Ker}(O)$ is empty and hence $O \in \operatorname{Gl}(W)$. Consequently, given O_1 and

$O_2 \in \mathrm{Sym}(P)$, we observe that $P(O_1^{-1}\xi) = P(O_1 O_1^{-1}\xi) = P(\xi)$ and $P(O_1 O_2 \xi) = P(O_2\xi) = P(\xi)$ for all $\xi \in W$; therefore $\mathrm{Sym}(P)$ is a subgroup of $\mathrm{Gl}(W)$.

To see that $\mathrm{Sym}(P)$ is compact, in view of the finite-dimensionality of $\mathrm{Gl}(W)$ and the Heine-Borel theorem, it suffices to show that $\mathrm{Sym}(P)$ is closed and bounded. First, for any sequence $\{O_n\} \subseteq \mathrm{Sym}(P)$ for which $O_n \to O$ as $n \to \infty$, the continuity of P ensures that $P(O\xi) = \lim_{n\to\infty} P(O_n\xi) = \lim_{n\to\infty} P(\xi) = P(\xi)$ for each $\xi \in W$ and therefore $\mathrm{Sym}(P)$ is closed. It remains to show that $\mathrm{Sym}(P)$ is bounded; this is the only piece of the proof that makes use of the fact that P is nondegenerate-homogeneous and not simply homogeneous. Assume that, to reach a contradiction, that there exists an unbounded sequence $\{O_n\} \subseteq \mathrm{Sym}(P)$. Choosing a norm $|\cdot|$ on W, let S be the corresponding unit sphere in W. Then there exists a sequence $\{\xi_n\} \subseteq W$ for which $|\xi_n| = 1$ for all $n \in \mathbb{N}_+$ but $\lim_{n\to\infty} |O_n\xi_n| = \infty$. In view of Lemma 2.8,

$$\infty = \lim_{n\to\infty} |P(O_n\xi_n)| = \lim_{n\to\infty} |P(\xi_n)| \le \sup_{\xi\in S} |P(\xi)|,$$

which cannot be true for P is necessarily bounded on S because it is continuous. □

Lemma 2.10 *Let Λ be a nondegenerate-homogeneous operator. For any $E_1, E_2 \in$* $\mathrm{Exp}(\Lambda)$,

$$\operatorname{tr} E_1 = \operatorname{tr} E_2.$$

Proof Let P be the symbol of Λ and take $E_1, E_2 \in \mathrm{Exp}(\Lambda)$. Since $E_1^*, E_2^* \in \mathrm{Exp}(P)$, $t^{E_1^*} t^{-E_2^*} \in \mathrm{Sym}(P)$ for all $t > 0$. As $\mathrm{Sym}(P)$ is a compact group in view of the previous lemma, the determinant map $\det : \mathrm{Gl}(\mathbb{V}^*) \to \mathbb{C}^*$, a Lie group homomorphism, necessarily maps $\mathrm{Sym}(P)$ into the unit circle. Consequently,

$$1 = |\det(t^{E_1^*} t^{-E_2^*})| = |\det(t^{E_1^*}) \det(t^{-E_2^*}) = |t^{\operatorname{tr} E_1^*} t^{-\operatorname{tr} E_2^*}| = t^{\operatorname{tr} E_1^*} t^{-\operatorname{tr} E_2^*}$$

for all $t > 0$. Therefore, $\operatorname{tr} E_1 = \operatorname{tr} E_1^* = \operatorname{tr} E_2^* = \operatorname{tr} E_2$ as desired. □

By the above lemma, to each nondegenerate-homogeneous operator Λ, we define the *homogeneous order* of Λ to be the number

$$\mu_\Lambda = \operatorname{tr} E$$

for any $E \in \mathrm{Exp}(\Lambda)$. By an appeal to Proposition 2.5, $E_{\mathbf{v}}^{\mathbf{n}} \in \mathrm{Exp}(\Lambda)$ for some $\mathbf{n} \in \mathbb{N}_+$ and so we observe that

$$\mu_\Lambda = \frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_d}. \tag{10}$$

In particular, μ_Λ is a positive rational number. We note that the term "homogeneous-order" does not coincide with the usual "order" for a partial differential operator.

For instance, the Laplacian $-\Delta$ on $\mathbb{R}^d$ is a second order operator; however, because $2^{-1}I \in \operatorname{Exp}(-\Delta)$, its homogeneous order is $\mu_{(-\Delta)} = \operatorname{tr} 2^{-1}I = d/2$.

2.1 Positive-Homogeneous Operators and Their Heat Kernels

We now restrict our attention to the study of positive-homogeneous operators and their associated heat kernels. To this end, let Λ be a positive-homogeneous operator on $\mathbb{V}$ with symbol P and homogeneous order μ_Λ. The heat kernel for Λ arises naturally from the study of the following Cauchy problem for the corresponding heat equation $\partial_t + \Lambda = 0$: Given initial data $f : \mathbb{V} \to \mathbb{C}$ which is, say, bounded and continuous, find $u(t,x)$ satisfying

$$\begin{cases} (\partial_t + \Lambda)\, u = 0 & \text{in } (0,\infty) \times \mathbb{V} \\ u(0,x) = f(x) & \text{for } x \in \mathbb{V}. \end{cases} \tag{11}$$

The initial value problem (11) is solved by putting

$$u(t,x) = \int_{\mathbb{V}} K_\Lambda^t(x-y) f(y)\, dy$$

where $K_\Lambda^{(\cdot)}(\cdot) : (0,\infty) \times \mathbb{V} \to \mathbb{C}$ is defined by

$$K_\Lambda^t(x) = \mathcal{F}^{-1}\left(e^{-tP}\right)(x) = \int_{\mathbb{V}^*} e^{-i\xi(x)} e^{-tP(\xi)}\, d\xi$$

for $t > 0$ and $x \in \mathbb{V}$; we call K_Λ *the heat kernel* associated to Λ. Equivalently, K_Λ is the integral (convolution) kernel of the continuous semigroup $\{e^{-t\Lambda}\}_{t>0}$ of bounded operators on $L^2(\mathbb{V})$ with infinitesimal generator $-\Lambda$. That is, for each $f \in L^2(\mathbb{V})$,

$$\left(e^{-t\Lambda} f\right)(x) = \int_{\mathbb{V}} K_\Lambda^t(x-y) f(y)\, dy \tag{12}$$

for $t > 0$ and $x \in \mathbb{V}$. Let us make some simple observations about K_Λ. First, by virtue of Lemma 2.8, it follows that $K_\Lambda^t \in \mathcal{S}(\mathbb{V})$ for each $t > 0$. Further, for any $E \in \operatorname{Exp}(\Lambda)$,

$$\begin{aligned} K_\Lambda^t(x) &= \int_{\mathbb{V}^*} e^{-i\xi(x)} e^{-P(t^{E^*}\xi)}\, d\xi \\ &= \int_{\mathbb{V}^*} e^{-i(t^{-E^*})\xi(x)} e^{-P(\xi)} \det(t^{-E^*})\, d\xi \\ &= \frac{1}{t^{\operatorname{tr} E}} \int_{\mathbb{V}^*} e^{-i\xi(t^{-E}x)} e^{-P(\xi)} d\xi = \frac{1}{t^{\mu_\Lambda}} K_\Lambda^1(t^{-E}x) \end{aligned}$$

for $t > 0$ and $x \in \mathbb{V}$. This computation immediately yields the so-called on-diagonal estimate for K_Λ,

$$\|e^{-t\Lambda}\|_{1\to\infty} = \|K_\Lambda^t\|_\infty = \frac{1}{t^{\mu_\Lambda}}\|K_\Lambda^1\|_\infty \leq \frac{C}{t^{\mu_\Lambda}}$$

for $t > 0$; this is equivalently a statement of ultracontractivity for the semigroup $e^{-t\Lambda}$. As it turns out, we can say something much stronger.

Proposition 2.11 *Let Λ be a positive-homogeneous operator with symbol P and homogeneous order μ_Λ. Let $R^\# : \mathbb{V} \to \mathbb{R}$ be the Legendre-Fenchel transform of $R = \operatorname{Re} P$ defined by*

$$R^\#(x) = \sup_{\xi \in \mathbb{V}^*}\{\xi(x) - R(\xi)\}$$

for $x \in \mathbb{V}$. Also, let $\mathbf{v}$ and $\mathbf{m} \in \mathbb{N}_+^d$ be as guaranteed by Proposition 2.5. Then, there exist positive constants C_0 and M and, for each multi-index β, a positive constant C_β such that, for all $k \in \mathbb{N}$,

$$\left|\partial_t^k D_{\mathbf{v}}^\beta K_\Lambda^t(x-y)\right| \leq \frac{C_\beta C_0^k k!}{t^{\mu_\Lambda + k + |\beta:2\mathbf{m}|}} \exp\left(-tMR^\#\left(\frac{x-y}{t}\right)\right) \tag{13}$$

for all $x, y \in \mathbb{V}$ and $t > 0$. In particular,

$$\left|K_\Lambda^t(x-y)\right| \leq \frac{C_0}{t^{\mu_\Lambda}} \exp\left(-tMR^\#\left(\frac{x-y}{t}\right)\right) \tag{14}$$

for all $x, y \in \mathbb{V}$ and $t > 0$.

Remark 2 In view of (10), the exponent on the prefactor in (13) can be equivalently written, for any multi-index β and $k \in \mathbb{N}$, as $\mu_\Lambda + k + |\beta : 2\mathbf{m}| = k + |\mathbf{1} + \beta : 2\mathbf{m}| = |\mathbf{1} + 2k\mathbf{m} + \beta : 2\mathbf{m}|$ where $\mathbf{1} = (1, 1, \dots, 1)$.

Remark 3 We note that the estimates of Proposition 2.11 are written in terms of the difference $x - y$ and can (trivially) be expressed in terms of a single spatial variable x. The estimates are written in this way to emphasize the role that K plays as an integral kernel. We will later replace Λ in (22) by a comparable variable-coefficient operator H and, in that setting, the associated heat kernel is not a convolution kernel and so we seek estimates involving two spatial variables x and y. To that end, the estimates here form a template for estimates in the variable-coefficient setting.

We prove the proposition above in the Sect. 5; the remainder of this section is dedicated to discussing the result and connecting it to the existing theory. Let us first note that the estimate (13) is mirrored by an analogous space-time estimate, Theorem 5.3 of [44], for the convolution powers of complex-valued functions on

$\mathbb{Z}^d$ satisfying certain conditions (see Sect. 5 of [44]). The relationship between these two results, Theorem 5.3 of [44] and Proposition 2.11, parallels the relationship between Gaussian off-diagonal estimates for random walks and the analogous off-diagonal estimates enjoyed by the classical heat kernel [33].

Let us first show that the estimates (13) and (14) recapture the well-known estimates of the theory of parabolic equations and systems in $\mathbb{R}^d$—a theory in which the Laplacian operator $\Delta = \sum_{l=1}^{d} \partial_{x_l}^2$ and its integer powers play a central role. To place things into the context of this article, let us observe that, for each positive integer m, the partial differential operator $(-\Delta)^m$ is a positive-homogeneous operator on $\mathbb{R}^d$ with symbol $P(\xi) = |\xi|^{2m}$; here, we identify $\mathbb{R}^d$ as its own dual equipped with the dot product and Euclidean norm $|\cdot|$. Indeed, one easily observes that $P = |\cdot|^{2m}$ is a positive-definite polynomial and $E = (2m)^{-1}I \in \operatorname{Exp}((-\Delta)^m)$ where $I \in \operatorname{Gl}(\mathbb{R}^d)$ is the identity. Consequently, the homogeneous order of $(-\Delta)^m$ is $d/2m = (2m)^{-1}\operatorname{tr}(I)$ and the Legendre-Fenchel transform of $R = \operatorname{Re} P = |\cdot|^{2m}$ is easily computed to be $R^{\#}(x) = C_m|x|^{2m/(2m-1)}$ where $C_m = (2m)^{1/(2m-1)} - (2m)^{-2m/(2m-1)} > 0$. Hence, (14) is the well-known estimate

$$\left|K_{(-\Delta)^m}^t(x-y)\right| \le \frac{C_0}{t^{d/2m}} \exp\left(-M\frac{|x-y|^{2m/(2m-1)}}{t^{1/(2m-1)}}\right)$$

for $x, y \in \mathbb{R}^d$ and $t > 0$; this so-called off-diagonal estimate is ubiquitous to the theory of "higher-order" elliptic and parabolic equations [16, 27, 30, 45]. To write the derivative estimate (13) in this context, we first observe that the basis given by Proposition 2.5 can be taken to be the standard Euclidean basis, $\mathbf{e} = \{e_1, e_2, \dots, e_d\}$ and further, $\mathbf{m} = (m, m, \dots, m)$ is the (isotropic) weight given by the proposition. Writing $D^\beta = D_{\mathbf{e}}^\beta = (i\partial_{x_1})^{\beta_1}(i\partial_{x_2})^{\beta_2}\cdots(i\partial_{x_d})^{\beta_d}$ and $|\beta| = \beta_1 + \beta_2 + \cdots + \beta_d$ for each multi-index β, (13) takes the form

$$\left|\partial_t^k D^\beta K_{(-\Delta)^m}^t(x-y)\right| \le \frac{C_0}{t^{(d+|\beta|)/2m+k}} \exp\left(-M\frac{|x-y|^{2m/(2m-1)}}{t^{1/(2m-1)}}\right)$$

for $x, y \in \mathbb{R}^d$ and $t > 0$, c.f., [27, Property 4, p. 93].

The appearance of the 1-dimensional Legendre-Fenchel transform in heat kernel estimates was previously recognized and exploited in [8] and [9] in the context of elliptic operators. Due to the isotropic nature of elliptic operators, the 1-dimensional transform is sufficient to capture the inherent isotropic decay of corresponding heat kernels. Beyond the elliptic theory, the appearance of the full d-dimensional Legendre-Fenchel transform is remarkable because it sharply captures the general anisotropic decay of K_Λ. Consider, for instance, the particularly simple positive-homogeneous operator $\Lambda = -\partial_{x_1}^6 + \partial_{x_2}^8$ on $\mathbb{R}^2$ with symbol $P(\xi_1, \xi_2) = \xi_1^6 + \xi_2^8$. It is easily checked that the operator E with matrix representation $\operatorname{diag}(1/6, 1/8)$, in the standard Euclidean basis, is a member of the $\operatorname{Exp}(\Lambda)$ and so the homogeneous order of Λ is $\mu_\Lambda = \operatorname{tr}(\operatorname{diag}(1/6, 1/8)) = 7/24$. Here we can compute the

Legendre-Fenchel transform of $R = \operatorname{Re} P = P$ directly to obtain $R^{\#}(x_1, x_2) = c_1|x_1|^{6/5} + c_2|x_2|^{8/7}$ for $(x_1, x_2) \in \mathbb{R}^2$ where c_1 and c_2 are positive constants. In this case, Proposition 2.11 gives positive constants C_0 and M for which

$$|K_\Lambda^t(x_1 - y_1, x_2 - y_2)| \leq \frac{C_0}{t^{7/24}} \exp\left(-\left(M_1 \frac{|x_1 - y_1|^{6/5}}{t^{1/5}} + M_2 \frac{|x_2 - y_2|^{8/7}}{t^{1/7}}\right)\right) \tag{15}$$

for $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ and $t > 0$ where $M_1 = c_1 M$ and $M_2 = c_2 M$. We note however that Λ is "separable" and so we can write $K_\Lambda^t(x_1, x_2) = K_{(-\Delta)^3}^t(x_1) K_{(-\Delta)^4}^t(x_2)$ where Δ is the 1-dimensional Laplacian operator. In view of Theorem 8 of [8] and its subsequent remark, the estimate (15) is seen to be sharp (modulo the values of M_1, M_2 and C). To further illustrate the proposition for a less simple positive-homogeneous operator, we consider the operator Λ appearing in Example 1.3. In this case,

$$R(\xi_1, \xi_2) = P(\xi_1, \xi_2) = \frac{1}{8}(\xi_1 + \xi_2)^2 + \frac{23}{384}(\xi_1 - \xi_2)^4$$

and one can verify directly that the $E \in \operatorname{End}(\mathbb{R}^2)$, with matrix representation

$$E_{\mathbf{e}} = \begin{pmatrix} 3/8 & 1/8 \\ 1/8 & 3/8 \end{pmatrix}$$

in the standard Euclidean basis, is a member of $\operatorname{Exp}(\Lambda)$. From this, we immediately obtain $\mu_\Lambda = \operatorname{tr}(E) = 3/4$ and one can directly compute

$$R^{\#}(x_1, x_2) = c_1|x_1 + x_2|^2 + c_2|x_1 - x_2|^{4/3}$$

for $(x_1, x_2) \in \mathbb{R}^2$ where c_1 and c_2 are positive constants. An appeal to Proposition 2.11 gives positive constants C_0 and M for which

$$|K_\Lambda^t(x_1 - y_1, x_2 - y_2)| \leq \frac{C_0}{t^{3/4}} \exp\left(-\left(M_1 \frac{|(x_1 - y_1) + (x_2 - y_2)|^2}{t} + M_2 \frac{|(x_1 - y_1) - (x_2 - y_2)|^{4/3}}{t^{1/3}}\right)\right)$$

for $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ and $t > 0$ where $M_1 = c_1 M$ and $M_2 = c_2 M$. Furthermore, $\mathbf{m} = (1, 2) \in \mathbb{N}_+^2$ and the basis $\mathbf{v} = \{v_1, v_2\}$ of $\mathbb{R}^2$ given in discussion surrounding (4) are precisely those guaranteed by Proposition 2.5. Appealing to the full strength of Proposition 2.11, we obtain positive constants C_0, M and, for each multi-index β, a positive constant C_β such that, for each $k \in \mathbb{N}$,

$$\left|\partial_t^k D_{\mathbf{v}}^\beta K_\Lambda(x_1 - y_1, x_2 - y_2)\right|$$
$$\leq \frac{C_\beta C_0^k k!}{t^{3/4+k+|\beta:2\mathbf{m}|}} \exp\left(-\left(M_1 \frac{|(x_1-y_1)+(x_2-y_2)|^2}{t} + M_2 \frac{|(x_1-y_1)-(x_2-y_2)|^{4/3}}{t^{1/3}}\right)\right)$$

for $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ and $t > 0$ where $M_1 = c_1 M$ and $M_2 = c_2 M$.

In the context of homogeneous groups, the off-diagonal behavior for the heat kernel of a positive Rockland operator (a positive self-adjoint operator which is homogeneous with respect to the fixed dilation structure) has been studied in [6, 25, 32] (see also [5]). Given a positive Rockland operator Λ on homogeneous group G, the best known estimate for the heat kernel K_Λ, due to Auscher, ter Elst and Robinson, is of the form

$$|K_\Lambda^t(h^{-1}g)| \leq \frac{C_0}{t^{\mu_\Lambda}} \exp\left(-M\left(\frac{\|h^{-1}g\|^{2m}}{t}\right)^{1/(2m-1)}\right) \tag{16}$$

where $\|\cdot\|$ is a homogeneous norm on G (consistent with Λ) and $2m$ is the highest order derivative appearing in Λ. In the context of $\mathbb{R}^d$, given a symmetric and positive-homogeneous operator Λ with symbol P, the structure $G_D = (\mathbb{R}^d, \{\delta_t^D\})$ for $D = 2mE$ where $E \in \operatorname{Exp}(\Lambda)$ is a homogeneous group on which Λ becomes a positive Rockland operator. On G_D, it is quickly verified that $\|\cdot\| = R(\cdot)^{1/2m}$ is a homogeneous norm (consistent with Λ) and so the above estimate is given in terms of $R(\cdot)^{1/(2m-1)}$ which is, in general, dominated by the Legendre-Fenchel transform of R. To see this, we need not look further than our previous and simple example in which $\Lambda = -\partial_{x_1}^6 + \partial_{x_2}^8$. Here $2m = 8$ and so $R(x_1, x_2)^{1/(2m-1)} = (|x_1|^6 + |x_2|^8)^{1/7}$. In view of (15), the estimate (16) gives the correct decay along the x_2-coordinate axis; however, the bounds decay at markedly different rates along the x_1-coordinate axis. This illustrates that the estimate (16) is suboptimal, at least in the context of $\mathbb{R}^d$, and thus leads to the natural question: For positive-homogeneous operators on a general homogeneous group G, what is to replace the Legendre-Fenchel transform in heat kernel estimates?

Returning to the general picture, let Λ be a positive-homogeneous operator on $\mathbb{V}$ with symbol P and homogeneous order μ_Λ. To highlight some remarkable properties about the estimates (13) and (14) in this general setting, the following proposition concerning $R^\#$ is useful; for a proof, see Sect. 8.3 of [44].

Proposition 2.12 *Let Λ be a positive-homogeneous operator with symbol P and let $R^\#$ be the Legendre-Fenchel transform of $R = \operatorname{Re} P$. Then, for any $E \in \operatorname{Exp}(\Lambda)$, $I - E \in \operatorname{Exp}(R^\#)$. Moreover $R^\#$ is continuous, positive-definite in the sense that $R^\#(x) \geq 0$ and $R^\#(x) = 0$ only when $x = 0$. Further, $R^\#$ grows superlinearly in the* *sense that, for any norm* $|\cdot|$ *on* $\mathbb{V}$,

$$\lim_{x\to\infty}\frac{|x|}{R^{\#}(x)}=0;$$

in particular, $R^{\#}(x)\to\infty$ *as* $x\to\infty$.

Let us first note that, in view of the proposition, we can easily rewrite (14), for any $E\in\operatorname{Exp}(\Lambda)$, as

$$\left|K_{\Lambda}^{t}(x-y)\right|\leq\frac{C_0}{t^{\mu_\Lambda}}\exp\left(-MR^{\#}\left(t^{-E}(x-y)\right)\right)$$

for $x,y\in\mathbb{V}$ and $t>0$; the analogous rewriting is true for (13). The fact that $R^{\#}$ is positive-definite and grows superlinearly ensures that the convolution operator $e^{-t\Lambda}$ defined by (12) for $t>0$ is a bounded operator from L^p to L^q for any $1\leq p,q\leq\infty$. Of course, we already knew this because K_{Λ}^{t} is a Schwartz function; however, when replacing Λ with a variable-coefficient operator H, as we will do in the sections to follow, the validity of the estimate (14) for the kernel of the semigroup $\{e^{-tH}\}$ initially defined on L^2, guarantees that the semigroup extends to a strongly continuous semigroup $\{e^{-tH_p}\}$ on $L^p(\mathbb{R}^d)$ for all $1\leq p\leq\infty$ and, what's more, the respective infinitesimal generators $-H_p$ have spectra independent of p [15]. Further, the estimate (14) is key to establishing the boundedness of the Riesz transform, it is connected to the resolution of Kato's square root problem and it provides the appropriate starting point for uniqueness classes of solutions to $\partial_t+H=0$ [7, 42]. With this motivation in mind, following some background in Sect. 3, we introduce a class of variable-coefficient operators in Sect. 4 called $(2\mathbf{m},\mathbf{v})$-positive-semi-elliptic operators, each such operator H comparable to a fixed positive-homogeneous operator. In Sect. 5, under the assumption that H has Hölder continuous coefficients and this notion of comparability is uniform, we construct a fundamental solution to the heat equation $\partial_t+H=0$ and show the essential role played by the Legendre-Fenchel transform in this construction. As mentioned previously, in a forthcoming work we will study the semigroup $\{e^{-tH}\}$ where H is a divergence-form operator, which is comparable to a fixed positive-homogeneous operator, whose coefficients are at worst measurable. As the Legendre-Fenchel transform appears here by a complex change of variables followed by a minimization argument, in the measurable coefficient setting it appears quite naturally by an application of the so-called Davies' method, suitably adapted to the positive-homogeneous setting.

3 Contracting Groups, Hölder Continuity and the Legendre-Fenchel Transform

In this section, we provide the necessary background on one-parameter contracting groups, anisotropic Hölder continuity, and the Legendre-Fenchel transform and its interplay with the two previous notions.

3.1 One-Parameter Contracting Groups

In what follows, W is a d-dimensional real vector space with a norm $|\cdot|$; the corresponding operator norm on $\mathrm{Gl}(W)$ is denoted by $\|\cdot\|$. Of course, since everything is finite-dimensional, the usual topologies on W and $\mathrm{Gl}(W)$ are insensitive to the specific choice of norms.

Definition 3.1 Let $\{T_t\}_{t>0} \subseteq \mathrm{Gl}(W)$ be a continuous one-parameter group. $\{T_t\}$ is said to be contracting if

$$\lim_{t\to 0} \|T_t\| = 0.$$

We easily observe that, for any diagonalizable $E \in \mathrm{End}(W)$ with strictly positive spectrum, the corresponding one-parameter group $\{t^E\}_{t>0}$ is contracting. Indeed, if there exists a basis $\mathbf{w} = \{w_1, w_2, \dots, w_d\}$ of W and a collection of positive numbers $\lambda_1, \lambda_2, \dots, \lambda_d$ for which $Ew_k = \lambda_k w_k$ for $k = 1, 2, \dots, d$, then the one parameter group $\{t^E\}_{t>0}$ has $t^E w_k = t^{\lambda_k} w_k$ for $k = 1, 2, \dots, d$ and $t > 0$. It then follows immediately that $\{t^E\}$ is contracting.

Proposition 3.2 *Let Q and R be continuous real-valued functions on W. If $R(w) > 0$ for all $w \neq 0$ and there exists $E \in \mathrm{Exp}(Q) \cap \mathrm{Exp}(R)$ for which $\{t^E\}$ is contracting, then, for some positive constant C, $Q(w) \leq CR(w)$ for all $w \in W$. If additionally $Q(w) > 0$ for all $w \neq 0$, then $Q \asymp R$.*

Proof Let S denote the unit sphere in W and observe that

$$\sup_{w\in S} \frac{Q(w)}{R(w)} =: C < \infty$$

because Q and R are continuous and R is non-zero on S. Now, for any non-zero $w \in W$, the fact that t^E is contracting implies that $t^E w \in S$ for some $t > 0$ by virtue of the intermediate value theorem. Therefore, $Q(w) = Q(t^E w)/t \leq CR(t^E w)/t = CR(w)$. In view of the continuity of Q and R, this inequality must hold for all $w \in W$. When additionally $Q(w) > 0$ for all non-zero w, the conclusion that $Q \asymp R$ is obtained by reversing the roles of Q and R in the preceding argument. □

Corollary 3.3 *Let Λ be a positive-homogeneous operator on $\mathbb{V}$ with symbol P and let $R^\#$ be the Legendre-Fenchel transform of $R = \mathrm{Re}\, P$. Then, for any positive constant M, $R^\# \asymp (MR)^\#$.*

Proof By virtue of Proposition 2.5, let $\mathbf{m} \in \mathbb{N}_+^d$ and $\mathbf{v}$ be a basis for $\mathbb{V}$ and for which $E_{\mathbf{v}}^{2\mathbf{m}} \in \mathrm{Exp}(\Lambda)$. In view of Proposition 2.12, $R^\#$ and $(MR)^\#$ are both continuous, positive-definite and have $I - E_{\mathbf{v}}^{2\mathbf{m}} \in \mathrm{Exp}(R^\#) \cap \mathrm{Exp}((MR)^\#)$. In view of (5), it is easily verified that $I - E_{\mathbf{v}}^{2\mathbf{m}} = E_{\mathbf{v}}^{\omega}$ where

$$\omega := \left(\frac{2m_1}{2m_1 - 1}, \frac{2m_2}{2m_2 - 1}, \dots \frac{2m_d}{2m_d - 1} \right) \in \mathbb{R}_+^d \tag{17}$$

and so it follows that $\{t^{E^{\omega}_{\mathbf{v}}}\}$ is contracting. The corollary now follows directly from Proposition 3.2. □

Lemma 3.4 *Let P be a positive-homogeneous polynomial on W and let $\mathbf{n} = 2\mathbf{m} \in \mathbb{N}_+^d$ and $\mathbf{w}$ be a basis for W for which the conclusion of Lemma 2.6 holds. Let $R = \operatorname{Re} P$ and let β and γ be multi-indices such that $\beta \le \gamma$ (in the standard partial ordering of multi-indices); we shall assume the notation of Lemma 2.6.*

1. *For any $n \in \mathbb{N}_+$ such that $|\beta : \mathbf{m}| \le 2n$, there exist positive constants M and M' for which*
$$|\xi^{\gamma}\nu^{\beta-\gamma}| \le M(R(\xi) + R(\nu))^n + M'$$
for all $\xi, \nu \in W$.
2. *If $|\beta : \mathbf{m}| = 2$, there exist positive constants M and M' for which*
$$|\xi^{\gamma}\nu^{\beta-\gamma}| \le MR(\xi) + M'R(\nu)$$
for all $\nu, \xi \in W$.
3. *If $|\beta : \mathbf{m}| = 2$ and $\beta > \gamma$, then for every $\epsilon > 0$ there exists a positive constant M for which*
$$|\xi^{\gamma}\nu^{\beta-\gamma}| \le \epsilon R(\xi) + MR(\nu)$$
for all $\nu, \xi \in W$.

Proof Assuming the notation of Lemma 2.6, let $E = E_{\mathbf{w}}^{2\mathbf{m}} \in \operatorname{End}(W)$ and consider the contracting group $\{t^{E\oplus E}\} = \{t^E \oplus t^E\}$ on $W \oplus W$. Because R is a positive-definite polynomial, it immediately follows that $W \oplus W \ni (\xi, \nu) \mapsto R(\xi) + R(\nu)$ is positive-definite. Let $|\cdot|$ be a norm on $W \oplus W$ and respectively denote by B and S the corresponding unit ball and unit sphere in this norm.

To see Item 1, first observe that

$$\sup_{(\xi,\nu)\in S} \frac{|\xi^{\gamma}\nu^{\beta-\gamma}|}{(R(\xi) + R(\nu))^n} =: M < \infty$$

Now, for any $(\xi, \nu) \in W \oplus W \setminus B$, because $\{t^{E\oplus E}\}$ is contracting, it follows from the intermediate value theorem that, for some $t \ge 1$, $t^{-(E\oplus E)}(\xi, \nu) = (t^{-E}\xi, t^{-E}\nu) \in S$. Correspondingly,

$$\begin{aligned}
|\xi^{\gamma}\nu^{\beta-\gamma}| &= t^{|\beta:2\mathbf{m}|}|(t^{-E}\xi)^{\gamma}(t^{-E})^{\beta-\gamma}| \\
&\le t^{|\beta:2\mathbf{m}|}M(R(t^{-E}\xi) + R(t^{-E}\nu))^n \\
&\le t^{|\beta:\mathbf{m}|/2-n}M(R(\xi) + R(\nu))^n \\
&\le M(R(\xi) + R(\nu))^n
\end{aligned}$$

because $|\beta : \mathbf{m}|/2 \leq n$. One obtains the constant M' and hence the desired inequality by simply noting that $|\xi^\gamma \nu^{\beta-\gamma}|$ is bounded for all $(\xi, \nu) \in B$.

For Item 2, we use analogous reasoning to obtain a positive constant M for which $|\xi^\gamma \nu^{\beta-\gamma}| \leq M(R(\xi) + R(\nu))$ for all $(\xi, \nu) \in S$. Now, for any non-zero $(\xi, \nu) \in W \oplus W$, the intermediate value theorem gives $t > 0$ for which $t^{E\oplus E}(\xi, \nu) = (t^E \xi, t^E \nu) \in S$ and hence

$$|\xi^\gamma \nu^{\beta-\gamma}| \leq t^{-|\beta:2\mathbf{m}|} M(R(t^E \xi) + R(t^E \nu)) = M(R(\xi) + R(\nu))$$

where we have used the fact that $|\beta : 2\mathbf{m}| = |\beta : \mathbf{m}|/2 = 1$ and that $E \in \operatorname{Exp}(R)$. As this inequality must also trivially hold at the origin, we can conclude that it holds for all $\xi, \nu \in W$, as desired.

Finally, we prove Item 3. By virtue of Item 2, for any $\xi, \nu \in W$ and $t > 0$,

$$\begin{aligned}|\xi^\gamma \nu^{\beta-\gamma}| &= |(t^E t^{-E} \xi)^\gamma \nu^{\beta-\gamma}| = t^{|\gamma:2\mathbf{m}|} |(t^{-E}\xi)^\gamma \nu^{\beta-\gamma}| \\ &\leq t^{|\gamma:2\mathbf{m}|} \left(MR(t^{-E}\xi) + M'R(\nu)\right) = Mt^{|\gamma:2\mathbf{m}|-1} R(\xi) + M' t^{|\gamma:2\mathbf{m}|} R(\nu).\end{aligned}$$

Noting that $|\gamma : 2\mathbf{m}| - 1 < 0$ because $\gamma < \beta$, we can make the coefficient of $R(\xi)$ arbitrarily small by choosing t sufficiently large and thereby obtaining the desired result. □

3.2 *Notions of Regularity and Hölder Continuity*

Throughout the remainder of this article, $\mathbf{v}$ will denote a fixed basis for $\mathbb{V}$ and correspondingly we henceforth assume the notational conventions appearing in Proposition 2.5 and $\mathbf{n} = 2\mathbf{m}$ is fixed. For $\alpha \in \mathbb{R}^d_+$, consider the homogeneous norm $|\cdot|^\alpha_{\mathbf{v}}$ defined by

$$|x|^\alpha_{\mathbf{v}} = \sum_{i=1}^d |x_i|^{\alpha_i}$$

for $x \in \mathbb{V}$ where $\phi_{\mathbf{v}}(x) = (x_1, x_2, \ldots, x_d)$. As one can easily check,

$$|t^{E^\alpha_{\mathbf{v}}} x|^\alpha_{\mathbf{v}} = t|x|^\alpha_{\mathbf{v}}$$

for all $t > 0$ and $x \in \mathbb{V}$ where $E^\alpha_{\mathbf{v}} \in \operatorname{Gl}(\mathbb{V})$ is defined by (5).

Definition 3.5 Let $\mathbf{m} \in \mathbb{N}^d_+$. We say that $\alpha \in \mathbb{R}^d_+$ is consistent with $\mathbf{m}$ if

$$E^\alpha_{\mathbf{v}} = a(I - E^{2\mathbf{m}}_{\mathbf{v}}) \tag{18}$$

for some $a > 0$.

As one can check, α is consistent with $\mathbf{m}$ if and only if $\alpha = a^{-1}\omega$ where ω is defined by (17).

Definition 3.6 Let $\Omega \subseteq \Omega' \subseteq \mathbb{V}$ and let $f : \Omega' \to \mathbb{C}$. We say that f is $\mathbf{v}$-Hölder continuous on Ω if for some $\alpha \in \mathbb{I}_+^d$ and positive constant M,

$$|f(x) - f(y)| \leq M|x - y|_{\mathbf{v}}^{\alpha} \tag{19}$$

for all $x, y \in \Omega$. In this case we will say that α is the $\mathbf{v}$-Hölder exponent of f. If $\Omega = \Omega'$ we will simply say that f is $\mathbf{v}$-Hölder continuous with exponent α.
The following proposition essentially states that, for bounded functions, Hölder continuity is a local property; its proof is straightforward and is omitted.

Proposition 3.7 *Let $\Omega \subseteq \mathbb{V}$ be open and non-empty. If f is bounded and $\mathbf{v}$-Hölder continuous of order $\alpha \in \mathbb{I}_+^d$, then, for any $\beta < \alpha$, f is also $\mathbf{v}$-Hölder continuous of order β.*
In view of the proposition, we immediately obtain the following corollary.

Corollary 3.8 *Let $\Omega \subseteq \mathbb{V}$ be open and non-empty and $\mathbf{m} \in \mathbb{N}_+^d$. If f is bounded and $\mathbf{v}$-Hölder continuous on Ω of order $\beta \in \mathbb{I}_+^d$, there exists $\alpha \in \mathbb{I}_+^d$ which is consistent with $\mathbf{m}$ for which f is also $\mathbf{v}$-Hölder continuous of order α.*

Proof The statement follows from the proposition by choosing any α, consistent with $\mathbf{m}$, such that $\alpha \leq \beta$. □
The following definition captures the minimal regularity we will require of fundamental solutions to the heat equation.

Definition 3.9 Let $\mathbf{n} \in \mathbb{N}_+^d$, $\mathbf{v}$ be a basis of $\mathbb{V}$ and let $\mathcal{O}$ be a non-empty open subset of $[0, T] \times \mathbb{V}$. A function $u(t, x)$ is said to be $(\mathbf{n}, \mathbf{v})$-regular on $\mathcal{O}$ if on $\mathcal{O}$ it is continuously differentiable in t and has continuous (spatial) partial derivatives $D_{\mathbf{v}}^{\beta} u(t, x)$ for all multi-indices β for which $|\beta : \mathbf{n}| \leq 1$.

3.3 *The Legendre-Fenchel Transform and Its Interplay with* v*-Hölder Continuity*

Throughout this section, R is the real part of the symbol P of a positive-homogeneous operator Λ on $\mathbb{V}$. We assume the notation of Proposition 2.12 (and hence Proposition 2.5) and write $E = E_{\mathbf{v}}^{2\mathbf{m}}$. Let us first record two important results which follow essentially from Proposition 2.12.

Corollary 3.10

$$R^{\#} \asymp |\cdot|_{\mathbf{v}}^{\omega}.$$

where ω was defined in (17).

Proof In view of Propositions 2.5 and 2.12, $E_{\mathbf{v}}^{\omega} = I - E_{\mathbf{v}}^{2\mathbf{m}} \in \operatorname{Exp}(R^{\#}) \cap \operatorname{Exp}(|\cdot|_{\mathbf{v}}^{\omega})$. After recalling that $\{t^{E_{\mathbf{v}}^{\omega}}\}$ is contracting, Proposition 3.2 yields the desired result immediately. □

By virtue of Proposition 2.12, standard arguments immediately yield the following corollary.

Corollary 3.11 *For any $\epsilon > 0$ and polynomial $Q : \mathbb{V} \to \mathbb{C}$, i.e., Q is a polynomial in any coordinate system, then*

$$Q(\cdot)e^{-\epsilon R^{\#}(\cdot)} \in L^{\infty}(\mathbb{V}) \cap L^{1}(\mathbb{V}).$$

Lemma 3.12 *Let $\gamma = (2m_{\max} - 1)^{-1}$. Then for any $T > 0$, there exists $M > 0$ such that*

$$R^{\#}(x) \leq Mt^{\gamma}R^{\#}(t^{-E}x)$$

for all $x \in \mathbb{V}$ and $0 < t \leq T$.

Proof In view of Corollary 3.10, it suffices to prove the statement

$$|t^{E}x|_{\mathbf{v}}^{\omega} \leq Mt^{\gamma}|x|_{\mathbf{v}}^{\omega}$$

for all $x \in \mathbb{V}$ and $0 < t \leq T$ where $M > 0$ and ω is given by (17). But for any $0 < t \leq T$ and $x \in \mathbb{V}$,

$$|t^{E}x|_{\mathbf{v}}^{\omega} = \sum_{j=1}^{d} t^{1/(2m_j-1)}|x_j|^{\omega_j} \leq t^{\gamma}\sum_{j=1}^{d} T^{(1/(2m_j-1)-\gamma)}|x_j|^{\omega_j}$$

from which the result follows. □

Lemma 3.13 *Let $\alpha \in \mathbb{I}_{+}^{d}$ be consistent with* **m**. *Then there exists positive constants σ and θ such that $0 < \sigma < 1$ and for any $T > 0$ there exists $M > 0$ such that*

$$|x|_{\mathbf{v}}^{\alpha} \leq Mt^{\sigma}(R^{\#}(t^{-E}x))^{\theta}$$

for all $x \in \mathbb{V}$ and $0 < t \leq T$.

Proof By an appeal to Corollary 3.10 and Lemma 3.12,

$$|x|_{\mathbf{v}}^{\omega} \leq Mt^{\gamma}R^{\#}(t^{-E}x)$$

for all $x \in \mathbb{V}$ and $0 < t \leq T$. Since α is consistent with **m**, $\alpha = a^{-1}\omega$ where a is that of Definition 3.5, the desired inequality follows by setting $\sigma = \gamma/a$ and $\theta = 1/a$. Because $\alpha \in \mathbb{I}_{+}^{d}$, it is necessary that $a \geq 2m_{\min}/(2m_{\min} - 1)$ whence $0 < \sigma \leq (2m_{\min} - 1)/(2m_{\min}(2m_{\max} - 1)) < 1$. □

The following corollary is an immediate application of Lemma 3.13.

Corollary 3.14 *Let $f : \mathbb{V} \to \mathbb{C}$ be* **v***-Hölder continuous with exponent $\alpha \in \mathbb{I}_+^d$ and suppose that α is consistent with* **m***. Then there exist positive constants σ and θ such that $0 < \sigma < 1$ and, for any $T > 0$, there exists $M > 0$ such that*

$$|f(x) - f(y)| \leq Mt^{\sigma}(R^{\#}(t^{-E}))^{\theta}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$.

4 On (2m, v)-Positive-Semi-Elliptic Operators

In this section, we introduce a class of variable-coefficient operators on $\mathbb{V}$ whose heat equations are studied in the next section. These operators, in view of Proposition 2.5, generalize the class of positive-homogeneous operators. Fix a basis $\mathbf{v}$ of $\mathbb{V}$, $\mathbf{m} \in \mathbb{N}_+^d$ and, in the notation of the previous section, consider a differential operator H of the form

$$H = \sum_{|\beta:\mathbf{m}|\leq 2} a_\beta(x) D_{\mathbf{v}}^{\beta} = \sum_{|\beta:\mathbf{m}|=2} a_\beta(x) D_{\mathbf{v}}^{\beta} + \sum_{|\beta:\mathbf{m}|<2} a_\beta(x) D_{\mathbf{v}}^{\beta} := H_p + H_l$$

where the coefficients $a_\beta : \mathbb{V} \to \mathbb{C}$ are bounded functions. The symbol of H, $P : \mathbb{V} \times \mathbb{V}^* \to \mathbb{C}$, is defined by

$$P(y,\xi) = \sum_{|\beta:\mathbf{m}|\leq 2} a_\beta(y)\xi^{\beta} = \sum_{|\beta:\mathbf{m}|=2} a_\beta(y)\xi^{\beta} + \sum_{|\beta:\mathbf{m}|<2} a_\beta(y)\xi^{\beta} := P_p(y,\xi) + P_l(y,\xi).$$

for $y \in \mathbb{V}$ and $\xi \in \mathbb{V}^*$. We shall call H_p the principal part of H and correspondingly, P_p is its principal symbol. Let's also define $R : \mathbb{V}^* \to \mathbb{R}$ by

$$R(\xi) = \operatorname{Re} P_p(0,\xi) \tag{20}$$

for $\xi \in \mathbb{V}^*$. At times, we will freeze the coefficients of H and H_p at a point $y \in \mathbb{V}$ and consider the constant-coefficient operators they define, namely $H(y)$ and $H_p(y)$ (defined in the obvious way). We note that, for each $y \in \mathbb{V}$, $H_p(y)$ is homogeneous with respect to the one-parameter group $\{\delta_t^E\}_{t>0}$ where $E = E_{\mathbf{v}}^{2\mathbf{m}} \in \mathrm{Gl}(\mathbb{V})$ is defined by (5). That is, H_p is homogeneous with respect to the same one-parameter group of dilations at each point in space. This also allows us to uniquely define the *homogeneous order of H* by

$$\mu_H = \operatorname{tr} E = (2m_1)^{-1} + (2m_2)^{-1} + \cdots + (2m_d)^{-1}. \tag{21}$$

We remark that this is consistent with our definition of homogeneous-order for constant-coefficient operators and we remind the reader that this notion differs from the usual order a partial differential operator (see the discussion surrounding (10)). As in the constant-coefficient setting, $H_p(y)$ is not necessarily homogeneous with respect to a unique group of dilations, i.e., it is possible that $\operatorname{Exp}(H_p(y))$ contains members of $\mathrm{Gl}(\mathbb{V})$ distinct from E. However, we shall henceforth only work with the endomorphism E, defined above, for worrying about this non-uniqueness of dilations does not aid our understanding nor will it sharpen our results. Let us further observe that, for each $y \in \mathbb{V}$, $P_p(y, \cdot)$ and R are homogeneous with respect to $\{t^{E^*}\}_{t>0}$ where $E^* \in \mathrm{Gl}(\mathbb{V}^*)$.

Definition 4.1 The operator H is called $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic if for all $y \in \mathbb{V}$, $\operatorname{Re} P_p(y, \cdot)$ is a positive-definite polynomial. H is called uniformly $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic if it is $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic and there exists $\delta > 0$ for which

$$\operatorname{Re} P_p(y, \xi) \geq \delta R(\xi)$$

for all $y \in \mathbb{V}$ and $\xi \in \mathbb{V}^*$. When the context is clear, we will simply say that H is positive-semi-elliptic and uniformly positive-semi-elliptic respectively.

In light of the above definition, a semi-elliptic operator H is one that, at every point $y \in \mathbb{V}$, its frozen-coefficient principal part $H_p(y)$, is a constant-coefficient positive-homogeneous operator which is homogeneous with respect to the same one-parameter group of dilations on $\mathbb{V}$. A uniformly positive-semi-elliptic operator is one that is semi-elliptic and is uniformly comparable to a constant-coefficient positive-homogeneous operator, namely $H_p(0)$. In this way, positive-homogeneous operators take a central role in this theory.

Remark 4 In view of Proposition 2.5, the definition of R via (20) agrees with that we have given for constant-coefficient positive-homogeneous operators.

Remark 5 For an $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic operator H, uniform semi-ellipticity can be formulated in terms of $\operatorname{Re} P_p(y_0, \cdot)$ for any $y_0 \in \mathbb{V}$; such a notion is equivalent in view of Proposition 3.2.

5 The Heat Equation

For a uniformly positive-semi-elliptic operator H, we are interested in constructing a fundamental solution to the heat equation,

$$(\partial_t + H)u = 0 \tag{22}$$

on the cylinder $[0, T] \times \mathbb{V}$; here and throughout $T > 0$ is arbitrary but fixed. By definition, a fundamental solution to (22) on $[0, T] \times \mathbb{V}$ is a function $Z : (0, T] \times \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ satisfying the following two properties:

1. For each $y \in \mathbb{V}$, $Z(\cdot, \cdot, y)$ is $(2\mathbf{m}, \mathbf{v})$-regular on $(0, T) \times \mathbb{V}$ and satisfies (22).
2. For each $f \in C_b(\mathbb{V})$,

$$\lim_{t \downarrow 0} \int_{\mathbb{V}} Z(t, x, y) f(y) dy = f(x)$$

for all $x \in \mathbb{V}$.

Given a fundamental solution Z to (22), one can easily solve the Cauchy problem: Given $f \in C_b(\mathbb{V})$, find $u(t, x)$ satisfying

$$\begin{cases} (\partial_t + H)u = 0 & \text{on} \quad (0, T) \times \mathbb{V} \\ u(0, x) = f(x) & \text{for} \quad x \in \mathbb{V}. \end{cases}$$

This is, of course, solved by putting

$$u(t, x) = \int_{\mathbb{V}} Z(t, x, y) f(y)\, dy$$

for $x \in \mathbb{V}$ and $0 < t \leq T$ and interpreting $u(0, x)$ as that defined by the limit of $u(t, x)$ as $t \downarrow 0$. The remainder of this paper is essentially dedicated to establishing the following result:

Theorem 5.1 *Let H be uniformly $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic with bounded $\mathbf{v}$-Hölder continuous coefficients. Let R and μ_H be defined by (20) and (21) respectively and denote by $R^{\#}$ the Legendre-Fenchel transform of R. Then, for any $T > 0$, there exists a fundamental solution $Z : (0, T] \times \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ to (22) on $[0, T] \times \mathbb{V}$ such that, for some positive constants C and M,*

$$|Z(t, x, y)| \leq \frac{C}{t^{\mu_H}} \exp\left(-tMR^{\#}\left(\frac{x - y}{t}\right)\right) \tag{23}$$

for $x, y \in \mathbb{V}$ and $0 < t \leq T$.

We remark that, by definition, the fundamental solution Z given by Theorem 5.1 is $(2\mathbf{m}, \mathbf{v})$-regular. Thus Z is necessarily continuously differentiable in t and has continuous spatial derivatives of all orders β such that $|\beta : \mathbf{m}| \leq 2$.

As we previously mentioned, the result above is implied by the work of Eidelman for $2\vec{b}$-parabolic systems on $\mathbb{R}^d$ (where $\vec{b} = \mathbf{m}$) [26, 28]. Eidelman's systems, of the form (1), are slightly more general than we have considered here, for their coefficients are also allowed to depend on t (but in a uniformly Hölder continuous way). Admitting this t-dependence is a relatively straightforward matter and, for simplicity of presentation, we have not included it (see Remark 6). In this slightly more general situation, stated in $\mathbb{R}^d$ and in which $\mathbf{v} = \mathbf{e}$ is the standard Euclidean basis, Theorem 2.2 (p. 79) [28] guarantees the existence of a fundamental solution $Z(t, x, y)$ to (1), which has the same regularity appearing in Theorem 5.1 and satisfies

$$|Z(t,x,y)| \leq \frac{C}{t^{1/(2m_1)+1/(2m_2)+\cdots+1/(2m_d)}} \exp\left(-M\sum_{k=1}^{d}\frac{|x_k-y_k|^{2m_k/(2m_k-1)}}{t^{1/(2m_k-1)}}\right) \tag{24}$$

for $x, y \in \mathbb{R}^d$ and $0 < t \leq T$ where C and M are positive constants. By an appeal to Corollary 3.10, we have $R^{\#} \asymp |\cdot|_{\mathbf{v}}^{\omega}$ and from this we see that the estimates (23) and (24) are comparable.

In view of Corollary 3.8, the hypothesis of Theorem 5.1 concerning the coefficients of H immediately imply the following a priori stronger condition:

Hypothesis 5.2 *There exists $\alpha \in \mathbb{I}_+^d$ which is consistent with $\mathbf{m}$ and for which the coefficients of H are bounded and $\mathbf{v}$-Hölder continuous on $\mathbb{V}$ of order α.*

5.1 Levi's Method

In this subsection, we construct a fundamental solution to (22) under only the assumption that H, a uniformly $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic operator, satisfies Hypothesis 5.2. Henceforth, all statements include Hypothesis 5.2 without explicit mention. We follow the famous method of Levi, c.f., [40] as it was adopted for parabolic systems in [27] and [30]. Although well-known, Levi's method is lengthy and tedious and we will break it into three steps. Let's motivate these steps by first discussing the heuristics of the method.

We start by considering the auxiliary equation

$$\Big(\partial_t + \sum_{|\beta:\mathbf{m}|=2} a_\beta(y) D_{\mathbf{v}}^{\beta}\Big)u = (\partial_t + H_p(y))u = 0 \tag{25}$$

where $y \in \mathbb{V}$ is treated as a parameter. This is the so-called frozen-coefficient heat equation. As one easily checks, for each $y \in \mathbb{V}$,

$$G_p(t,x;y) := \int_{\mathbb{V}^*} e^{-i\xi(x)} e^{-tP_p(y,\xi)} d\xi \qquad (x \in \mathbb{V}, t > 0)$$

solves (25). By the uniform semi-ellipticity of H, it is clear that $G_p(t, \cdot; y) \in \mathcal{S}(\mathbb{V})$ for $t > 0$ and $y \in \mathbb{V}$. As we shall see, more is true: G_p is an approximate identity in the sense that

$$\lim_{t\downarrow 0} \int_{\mathbb{V}} G_p(t, x-y; y) f(y)\, dy = f(x)$$

for all $f \in C_b(\mathbb{V})$. Thus, it is reasonable to seek a fundamental solution to (22) of the form

$$\begin{aligned} Z(t,x,y) &= G_p(t,x-y;y) + \int_0^t \int_{\mathbb{V}} G_p(t-s,x-z;z)\phi(s,z,y)dzds \\ &= G_p(t,x-y;y) + W(t,x,y) \end{aligned} \tag{26}$$

where ϕ is to be chosen to ensure that the correction term W is $(2\mathbf{m},\mathbf{v})$-regular, accounts for the fact that G_p solves (25) but not (22), and is "small enough" as $t \to 0$ so that the approximate identity aspect of Z is inherited directly from G_p.

Assuming for the moment that W is sufficiently regular, let's apply the heat operator to (26) with the goal of finding an appropriate ϕ to ensure that Z is a solution to (22). Putting

$$K(t,x,y) = -(\partial_t + H)G_p(t,x-y;y),$$

we have formally,

$$\begin{aligned} (\partial_t + H)Z(t,x,y) &= -K(t,x,y) + (\partial_t + H)\int_0^t \int_{\mathbb{V}} G_p(t-s,x-z;z)\phi(s,z,y)\,dz\,ds \\ &= -K(t,x,y) + \lim_{s\uparrow t}\int_{\mathbb{V}} G_p(t-s,x-z;z)\phi(s,z,y)\,dz \\ &\qquad - \int_0^t \int_{\mathbb{V}} -(\partial_t + H)G_p(t-s,x-z;z)\phi(s,z,y)\,dz\,ds \\ &= -K(t,x,y) + \phi(t,x,y) - \int_0^t \int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds \end{aligned} \tag{27}$$

where we have made use of Leibniz' rule and our assertion that G_p is an approximate identity. Thus, for Z to satisfy (22), ϕ must satisfy the integral equation

$$\begin{aligned} K(t,x,y) &= \phi(t,x,y) - \int_0^t \int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds \\ &= \phi(t,x,y) - L(\phi)(t,x,y). \end{aligned} \tag{28}$$

Viewing L as a linear integral operator, (28) is the equation $K = (I-L)\phi$ which has the solution

$$\phi = \sum_{n=0}^{\infty} L^n K \tag{29}$$

provided the series converges in an appropriate sense.

Taking the above as purely formal, our construction will proceed as follows: We first establish estimates for G_p and show that G_p is an approximate identity;

this is Step 1. In Step 2, we will define ϕ by (29) and, after deducing some subtle estimates, show that ϕ's defining series converges whence (28) is satisfied. Finally in Step 3, we will make use of the estimates from Steps 1 and 2 to validate the formal calculation made in (27). Everything will be then pieced together to show that Z, defined by (26), is a fundamental solution to (22). Our entire construction depends on obtaining precise estimates for G_p and for this we will rely heavily on the homogeneity of P_p and the Legendre-Fenchel transform of R.

Remark 6 One can allow the coefficients of H to also depend on t in a uniformly continuous way, and Levi's method pushes though by instead taking G_p as the solution to a frozen-coefficient initial value problem [26, 28].

Step 1: Estimates for G_p and Its Derivatives

The lemma below is a basic building block used in our construction of a fundamental solution to (22) via Levi's method and it makes essential use of the uniform semi-ellipticity of H. We note however that the precise form of the constants obtained, as they depend on k and β, are more detailed than needed for the method to work. Also, the partial differential operators $D_{\mathbf{v}}^{\beta}$ of the lemma are understood to act of the x variable of $G_p(t, x; y)$.

Lemma 5.3 *There exist positive constants M and C_0 and, for each multi-index β, a positive constant C_β such that, for any $k \in \mathbb{N}$,*

$$|\partial_t^k D_{\mathbf{v}}^{\beta} G_p(t, x; y)| \leq \frac{C_\beta C_0^k k!}{t^{\mu_H + k + |\beta:2\mathbf{m}|}} \exp\left(-tMR^{\#}\left(x/t\right)\right) \tag{30}$$

for all $x, y \in \mathbb{V}$ and $t > 0$.

Before proving the lemma, let us note that $tR^{\#}(x/t) = R^{\#}(t^{-E}x)$ for all $t > 0$ and $x \in \mathbb{V}$ in view of Proposition 2.12. Thus the estimate (30) can be written equivalently as

$$|\partial_t^k D_{\mathbf{v}}^{\beta} G_p(t, x; y)| \leq \frac{C_\beta C_0^k k!}{t^{\mu_H + k + |\beta:2\mathbf{m}|}} \exp(-MR^{\#}(t^{-E}x)) \tag{31}$$

for $x, y \in \mathbb{V}$ and $t > 0$. We will henceforth use these forms interchangeably and without explicit mention.

Proof Let us first observe that, for each $x, y \in \mathbb{V}$ and $t > 0$,

$$\begin{aligned}
\partial_t^k D_{\mathbf{v}}^{\beta} G_p(t, x; y) &= \int_{\mathbb{V}^*} (P_p(y, \xi))^k \xi^{\beta} e^{-i\xi(x)} e^{-tP_p(y,\xi)}\, d\xi \\
&= \int_{\mathbb{V}^*} (P_p(y, t^{-E^*}\xi))^k (t^{-E^*}\xi)^{\beta} e^{-i\xi(t^{-E}x)} e^{-P_p(y,\xi)} t^{-\operatorname{tr} E}\, d\xi \\
&= t^{-\mu_H - k - |\beta:2\mathbf{m}|} \int_{\mathbb{V}^*} (P_p(y, \xi))^k \xi^{\beta} e^{-i\xi(t^{-E}x)} e^{-P_p(y,\xi)} d\xi
\end{aligned}$$

where we have used the homogeneity of P_p with respect to $\{t^{E^*}\}$ and the fact that $\mu_H = \operatorname{tr} E$. Therefore

$$t^{\mu_H+k+|\beta:2\mathbf{m}|}(\partial_t^k D_{\mathbf{v}}^\beta G_p(t,\cdot\,;y))(t^E x) = \int_{\mathbb{V}^*} (P_p(y,\xi))^k \xi^\beta e^{-i\xi(x)} e^{-P_p(y,\xi)}\,d\xi \tag{32}$$

for all $x, y \in \mathbb{V}$ and $t > 0$. Thus, to establish (30) (equivalently (31)) it suffices to estimate the right hand side of (32) which is independent of t.

The proof of the desired estimate requires making a complex change of variables and for this reason we will work with the complexification of $\mathbb{V}^*$, whose members are denoted by $z = \xi - i\nu$ for $\xi, \nu \in \mathbb{V}^*$; this space is isomorphic to $\mathbb{C}^d$. We claim that there are positive constants C_0, M_1, M_2 and, for each multi-index β, a positive constant C_β such that, for each $k \in \mathbb{N}$,

$$|(P_p(y,\xi - i\nu))^k (\xi - i\nu)^\beta e^{-P_p(y,\xi - i\nu)}| \le C_\beta C_0^k k! e^{-M_1 R(\xi)} e^{M_2 R(\nu)} \tag{33}$$

for all $\xi, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$. Let us first observe that

$$P_p(y, \xi - i\nu) = P_p(y,\xi) + \sum_{|\beta:\mathbf{m}|=2} \sum_{\gamma<\beta} a_{\beta,\gamma}\xi^\gamma(-i\nu)^{\beta-\gamma}$$

for all $z, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$, where $a_{\beta,\gamma}$ are bounded functions of y arising from the coefficients of H and the coefficients of the multinomial expansion. By virtue of the uniform semi-ellipticity of H and the boundedness of the coefficients, we have

$$-\operatorname{Re} P_p(y,\xi - i\nu) \le -\delta R(\xi) + C \sum_{|\beta:\mathbf{m}|=2} \sum_{\gamma<\beta} |\xi^\gamma \nu^{\beta-\gamma}|$$

for all $\xi, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$ where C is a positive constant. By applying Lemma 3.4 to each term $|\xi^\gamma \nu^{\beta-\gamma}|$ in the summation, we can find a positive constant M for which the entire summation is bounded above by $\delta/2R(\xi) + MR(\nu)$ for all $\xi, \nu \in \mathbb{V}^*$. By setting $M_1 = \delta/6$, we have

$$-\operatorname{Re} P_p(y,\xi - i\nu) \le -3M_1 R(\xi) + MR(\nu) \tag{34}$$

for all $\xi, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$. By analogous reasoning (making use of item 1 of Lemma 3.4), there exists a positive constant C for which

$$|P_p(y,\xi - i\nu)| \le C(R(\xi) + R(\nu))$$

for all $\xi, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$. Thus, for any $k \in \mathbb{N}$,

$$|P_p(y,\xi - i\nu)|^k \le \frac{C^k k!}{M_1^k} \frac{(M_1(R(\xi) + R(\nu)))^k}{k!} \le C_0^k k! e^{M_1(R(\xi)+R(\nu))} \tag{35}$$

for all $\xi, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$ where $C_0 = C/M_1$. Finally, for each multi-index β, another application of Lemma 3.4 gives $C' > 0$ for which

$$|(\xi - i\nu)^\beta| \leq |\xi^\beta| + |\nu^\beta| + \sum_{0<\gamma<\beta} c_{\gamma,\beta}|\xi^\gamma \nu^{\beta-\gamma}| \leq C' \left((R(\xi) + R(\nu))^n + 1\right)$$

for all $\xi, \nu \in \mathbb{V}^*$ where $n \in \mathbb{N}$ has been chosen to satisfy $|\beta : 2n\mathbf{m}| < 1$. Consequently, there is a positive constant C_β for which

$$|(\xi - i\nu)^\beta| \leq C_\beta e^{M_1(R(\xi)+R(\nu))} \tag{36}$$

for all $\xi, \nu \in \mathbb{V}^*$. Upon combining (34)–(36), we obtain the inequality

$$\left|P_p(y, \xi - i\nu)^k (\xi - i\nu)^\beta e^{-P_p(y,\xi-i\nu)}\right| \leq C_\beta C_0^k k! e^{-M_1 R(\xi)+(M+2M_1)R(\nu)}$$

which holds for all $\xi, \nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$. Upon paying careful attention to the way in which our constants were chosen, we observe the claim is established by setting $M_2 = M + 2M_1$.

From the claim above, it follows that, for any $\nu \in \mathbb{V}^*$ and $y \in \mathbb{V}$, the following change of coordinates by means of a $\mathbb{C}^d$ contour integral is justified:

$$\begin{aligned}\int_{\mathbb{V}^*} (P_p(y,\xi))^k \xi^\beta e^{-i\xi(x)} e^{-P_p(y,\xi)}\, d\xi &= \int_{\xi\in\mathbb{V}^*} (P_p(y, \xi - i\nu)^k (\xi - i\nu)^\beta e^{-i(\xi-i\nu)(x)} e^{-P_p(y,\xi-i\nu)}\, d\xi \\ &= e^{-\nu(x)} \int_{\xi\in\mathbb{V}^*} (P_p(y, \xi - i\nu)^k (\xi - i\nu)^\beta e^{-i\xi(x)} e^{-P_p(y,\xi-i\nu)}\, d\xi.\end{aligned}$$

Thus, by virtue of the estimate (33),

$$\begin{aligned}\left|\int_{\mathbb{V}^*} (P_p(y,\xi))^k \xi^\beta e^{-i\xi(x)} e^{-P_p(y,\xi)}\, d\xi\right| &\leq C_\beta C_0^k k! e^{-\nu(x)} e^{M_2 R(\nu)} \int_{\mathbb{V}^*} e^{-M_1 R(\xi)}\, d\xi \\ &\leq C_\beta C_0^k k! e^{-(\nu(x)-M_2 R(\nu))}\end{aligned}$$

for all $x, y \in \mathbb{V}$ and $\nu \in \mathbb{V}^*$ where we have absorbed the integral of $\exp(-M_1 R(\xi))$ into C_β. Upon minimizing with respect to $\nu \in \mathbb{V}^*$, we have

$$\left|\int_{\mathbb{V}^*} (P_p(y,\xi))^k \xi^\beta e^{-i\xi(x)} e^{-P_p(y,\xi)} d\xi\right| \leq C_\beta C_0^k k! e^{-(M_2 R)^\#(x)} \leq C_\beta C_0^k k! e^{-MR^\#(x)} \tag{37}$$

for all x and $y \in \mathbb{V}$ because

$$-(M_2 R)^\#(x) = -\sup_\nu\{\nu(x) - M_2 R(\nu)\} = \inf_\nu\{-(\nu(x) - M_2 R(\nu))\};$$

in this we see the natural appearance of the Legendre-Fenchel transform. The replacement of $(M_2R)^{\#}(x)$ by $MR^{\#}(x)$ is done using Corollary 3.3 and, as required, the constant M is independent of k and β. Upon combining (32) and (37), we obtain the desired estimate (30). □

As a simple corollary to the lemma, we obtain Proposition 2.11.

Proof of Proposition 2.11 Given a positive-homogeneous operator Λ, we invoke Proposition 2.5 to obtain $\mathbf{v}$ and $\mathbf{m}$ for which $\Lambda = \sum_{|\beta:\mathbf{m}|=2} a_\beta D_{\mathbf{v}}^{\beta}$. In other words, Λ is an $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic operator which consists only of its principal part. Consequently, the heat kernel K_Λ satisfies $K_\Lambda^t(x) = G_p(t, x; 0)$ for all $x \in \mathbb{V}$ and $t > 0$ and so we immediately obtain the estimate (13) from the lemma. □

Making use of Hypothesis 5.2, a similar argument to that given in the proof of Lemma 5.3 yields the following lemma.

Lemma 5.4 *There is a positive constant M and, to each multi-index β, a positive constant C_β such that*

$$|D_{\mathbf{v}}^{\beta}[G_p(t, x; y+h) - G_p(t, x; y)]| \leq C_\beta t^{-(\mu_H + |\beta:2\mathbf{m}|)} |h|_{\mathbf{v}}^{\alpha} \exp(-tMR^{\#}(x/t))$$

for all $t > 0$, $x, y, h \in \mathbb{V}$. Here, in view of Hypothesis 5.2, α is the $\mathbf{v}$-Hölder continuity exponent for the coefficients of H.

Lemma 5.5 *Suppose that $g \in C_b((t_0, T] \times \mathbb{V})$ where $0 \leq t_0 < T < \infty$. Then, on any compact set $Q \subseteq (t_0, T] \times \mathbb{V}$,*

$$\int_{\mathbb{V}} G_p(t, x-y; y) g(s-t, y)\, dy \to g(s, x)$$

uniformly on Q as $t \to 0$. In particular, for any $f \in C_b(\mathbb{V})$,

$$\int_{\mathbb{V}} G_p(t, x-y; y) f(y)\, dy \to f(x)$$

uniformly on all compact subsets of $\mathbb{V}$ as $t \to 0$.

Proof Let Q be a compact subset of $(t_0, T] \times \mathbb{V}$ and write

$$\begin{aligned}
&\int_{\mathbb{V}} G_p(t, x-y; y) g(s-t, y)\, dy \\
&= \int_{\mathbb{V}} G_p(t, x-y; x) g(s-t, y)\, dy + \int_{\mathbb{V}} [G_p(t, x-y; y) - G_p(t, x-y; x)] g(s-t, y)\, dy \\
&:= I_t^{(1)}(s, x) + I_t^{(2)}(s, x).
\end{aligned}$$

Let $\epsilon > 0$ and, in view of Corollary 3.11, let K be a compact subset of $\mathbb{V}$ for which

$$\int_{\mathbb{V}\setminus K} \exp(-MR^{\#}(z))\,dz < \epsilon$$

where the constant M is that given in (30) of Lemma 5.3. Using the continuity of g, we have for sufficiently small $t > 0$,

$$\sup_{\substack{(s,x)\in Q\\ z\in K}} |g(s-t, x-t^E z) - g(s,x)| < \epsilon.$$

We note that, for any $t > 0$ and $x \in \mathbb{V}$,

$$\int_{\mathbb{V}} G_p(t, x-y; x)\,dy = e^{-tP_p(x,\xi)}\Big|_{\xi=0} = 1.$$

Appealing to Lemma 5.3 we have, for any $(s,x) \in Q$,

$$\begin{aligned}
|I_t^{(1)}(s,x) - g(s,x)| &\le \left| \int_{\mathbb{V}} G_p(t, x-y; x)(g(s-t,y) - g(s,x))\,dy \right| \\
&\le \int_{\mathbb{V}} |G_p(1,z;x)(g(s-t, x-t^E z) - g(s,x))|\,dz \\
&\le 2\|g\|_\infty C \int_{\mathbb{V}\setminus K} \exp(-MR^{\#}(z))\,dz \\
&\qquad + C\int_K \exp(-MR^{\#}(z))|(g(s-t, x-t^E z) - g(s,x))|\,dz \\
&\le \epsilon C\left(2\|g\|_\infty + \|e^{-MR^{\#}}\|_1\right);
\end{aligned}$$

here we have made the change of variables: $y \mapsto t^E(x-y)$ and used the homogeneity of P_p to see that $t^{\mu_H} G_p(t, t^E z; x) = G_p(1, z; x)$. Therefore $I_t^{(1)}(s,x) \to g(s,x)$ uniformly on Q as $t \to 0$.

Let us now consider $I^{(2)}$. With the help of Lemmas 3.13 and 5.4 and by making similar arguments to those above we have

$$\begin{aligned}
|I_t^{(2)}(s,x)| &\le C\|g\|_\infty \int_{\mathbb{V}} t^{-\mu_H} |x-y|_{\mathbf{v}}^{\alpha} \exp(-MR^{\#}(t^{-E}(x-y))\,dy \\
&\le \|g\|_\infty C t^{\sigma} \int_{\mathbb{V}} t^{-\operatorname{tr} E} (R^{\#}(t^{-E}(x-y)))^{\theta} \exp(-MR^{\#}(t^{-E}(x-y)))\,dy \\
&\le \|g\|_\infty C t^{\sigma} \int_{\mathbb{V}} (R^{\#}(x))^{\theta} \exp(-MR^{\#}(z))\,dz \le \|g\|_\infty C' t^{\sigma}
\end{aligned}$$

for all $s \in (t_0, T]$, $0 < t < s - t_0$ and $x \in \mathbb{V}$; here $0 < \sigma < 1$. Consequently, $I_t^{(2)}(s,x) \to 0$ uniformly on Q as $t \to 0$ and the lemma is proved. □

Combining the results of Lemmas 5.3 and 5.5 yields at once:

Corollary 5.6 *For each $y \in \mathbb{V}$, $G_p(\cdot, \cdot - y; y)$ is a fundamental solution to* (25).

Step 2: Construction of ϕ and the Integral Equation
For $t > 0$ and $x, y \in \mathbb{V}$, put

$$\begin{aligned} K(t,x,y) &= -(\partial_t + H)G_p(t, x-y; y) \\ &= \big(H_p(y) - H\big)G_p(t, x-y; y) \\ &= \int_{\mathbb{V}^*} e^{-i\xi(x-y)}\big(P_p(y,\xi) - P(x,\xi)\big)e^{-tP_p(y,\xi)}\, d\xi \end{aligned}$$

and iteratively define

$$K_{n+1}(t,x,y) = \int_0^t \int_{\mathbb{V}} K_1(t-s,x,z)K_n(s,z,y)\, dz ds$$

where $K_1 = K$. In the sense of (29), note that $K_{n+1} = L^n K$.

We claim that for some $0 < \rho < 1$ and positive constants C and M,

$$|K(t,x,y)| \le Ct^{-(\mu_H + 1 - \rho)} \exp(-MR^\#(t^{-E}(x-y))) \tag{38}$$

for all $x, y \in \mathbb{V}$ and $0 < t \le T$. Indeed, observe that

$$|K(t,x,y)| \le \sum_{|\beta:\mathbf{m}|=2} |a_\beta(y) - a_\beta(x)||D_{\mathbf{v}}^\beta G_p(t,x-y;y)| + C\sum_{|\beta:\mathbf{m}|<2} |D_{\mathbf{v}}^\beta G_p(t,x-y;y)|$$

for all $x, y \in \mathbb{V}$ and $t > 0$ where we have used the fact that the coefficients of H are bounded. In view of Lemma 5.3, we have

$$\begin{aligned} |K(t,x,y)| \le \sum_{|\beta:\mathbf{m}|=2} &|a_\beta(y) - a_\beta(x)|Ct^{-(\mu_H+1)} \exp(-MR^\#(t^{-E}(x-y))) \\ &+Ct^{-(\mu_H+\eta)} \exp(-MR^\#(t^{-E}(x-y))) \end{aligned}$$

for all $x, y \in \mathbb{V}$ and $0 < t \le T$ where

$$\eta = \max\{|\beta : 2\mathbf{m}| : |\beta : \mathbf{m}| \ne 2 \text{ and } a_\beta \ne 0\} < 1.$$

Using Hypothesis 5.2, an appeal to Corollary 3.14 gives $0 < \sigma < 1$ and $\theta > 0$ for which

$$\begin{aligned} |K(t,x,y)| \le Ct^{\sigma-(\mu_H+1)}&(R^\#(t^{-E}(x-y)))^\theta \exp(-MR^\#(t^{-E}(x-y))) \\ &+Ct^{-(\mu_H+\eta)} \exp(-MR^\#(t^{-E}(x-y))) \end{aligned}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$. Our claim is then justified by setting

$$\rho = \max\{\sigma, 1 - \eta\} \tag{39}$$

and adjusting the constants C and M appropriately to absorb the prefactor $(R^{\#}(t^{-E}(x - y)))^{\theta}$ into the exponent. It should be noted that the constant ρ is inherently dependent on H. For it is clear that η depends on H. The constants σ and θ are specified in Lemma 3.13 and are defined in terms of the Hölder exponent of the coefficients of H and the weight $\mathbf{m}$.

Taking cues from our heuristic discussion, we will soon form a series whose summands are the functions K_n for $n \geq 1$. In order to talk about the convergence of this series, our next task is to estimate these functions and in doing this we will observe two separate behaviors: a finite number of terms will exhibit singularities in t at the origin; the remainder of the terms will be absent of such singularities and will be estimated with the help of the Gamma function. We first address the terms with the singularities.

Lemma 5.7 *Let* $0 < \rho < 1$ *be given by* (39) *and* $M > 0$ *be any constant for which* (38) *is satisfied. For any positive natural number* n *such that* $\rho(n-1) \leq \mu_H + 1$ *and* $\epsilon > 0$ *for which* $\epsilon n < 1$, *there is a constant* $C_n(\epsilon) \geq 1$ *such that*

$$|K_n(t, x, y)| \leq C_n(\epsilon) t^{-(\mu_H + 1 - n\rho)} \exp(-M(1 - \epsilon n) R^{\#}(t^{-E}(x - y)))$$

for all $x, y \in \mathbb{V}$ *and* $0 < t \leq T$.

Proof In view of (38), it is clear that the estimate holds when $n = 1$. Let us assume the estimate holds for $n \geq 1$ such that $\rho n < 1 + \mu_H$ and $\epsilon > 0$ for which $\epsilon n < \epsilon(n + 1) < 1$. Then

$$\begin{aligned} |K_{n+1}(t, x, y)| \leq \int_0^t \int_{\mathbb{V}} & C_1(\epsilon)(t-s)^{-(\mu_H + 1 - \rho)} C_n(\epsilon) s^{-(\mu_H + 1 - n\rho)} \\ & \times \exp(-M R^{\#}((t-s)^{-E}(x - z))) \exp(-M_{\epsilon,n} R^{\#}(s^{-E}(z - y)))\, dz\, ds \end{aligned} \tag{40}$$

for $x, y \in \mathbb{V}$ and $0 < t \leq T$ where we have set $M_{\epsilon,n} = M(1 - \epsilon n)$. Observe that

$$\begin{aligned} R^{\#}(t^{-E}(x - y)) &= \sup\{\xi(x - y) - tR(\xi)\} \\ &= \sup\{\xi(x - z) - (t - s)R(\xi) + \xi(z - y) - sR(\xi)\} \\ &\leq R^{\#}((t - s)^{-E}(x - z)) + R^{\#}(s^{-E}(z - y)) \end{aligned} \tag{41}$$

for all $x, y, z \in \mathbb{V}$ and $0 < s \leq t$. Using the fact that $0 < \epsilon n < \epsilon(n+1) < 1$, (41) guarantees that

$$\begin{aligned}
&(1-\epsilon(n+1))R^{\#}(t^{-E}(x-y)) + \epsilon\left(R^{\#}((t-s)^{-E}(x-z)) + R^{\#}(s^{-E}(z-y))\right) \\
&\leq (1-\epsilon(n+1))\left(R^{\#}((t-s)^{-E}(x-z)) + R^{\#}(s^{-E}(z-y))\right) \\
&\qquad\qquad +\epsilon\left(R^{\#}((t-s)^{-E}(x-z)) + R^{\#}(s^{-E}(z-y))\right) \\
&\leq (1-\epsilon n)R^{\#}((t-s)^{-E}(x-z)) + (1-\epsilon n)R^{\#}(s^{-E}(z-y)) \\
&\leq R^{\#}((t-s)^{-E}(x-z)) + (1-\epsilon n)R^{\#}(s^{-E}(z-y))
\end{aligned}$$

or equivalently

$$\begin{aligned}
&-MR^{\#}((t-s)^{-E}(x-z)) - M_{\epsilon,n}R^{\#}(s^{-E}(z-y)) \\
&\leq -M_{\epsilon,n+1}R^{\#}(t^{-E}(x-y)) - \epsilon M\left(R^{\#}((t-s)^{-E}(x-z)) + R^{\#}(s^{-E}(z-y))\right)
\end{aligned} \tag{42}$$

for all $x, y, z \in \mathbb{V}$ and $0 < s \leq t$. Combining (40) and (42) yields

$$\begin{aligned}
&|K_{n+1}(t,x,y)| \\
&\leq C_1(\epsilon)C_n(\epsilon)\exp(-M_{\epsilon,n+1}R^{\#}(t^{-E}(x-y)))\int_0^t\int_{\mathbb{V}}(t-s)^{-(\mu_H+1-\rho)}s^{-(\mu_H+1-n\rho)} \\
&\qquad\qquad \times\exp(-\epsilon M(R^{\#}((t-s)^{-E}(x-z)) + R^{\#}(s^{-E}(z-y)))\,dz\,ds \\
&\leq C_1(\epsilon)C_n(\epsilon)\exp(-M_{\epsilon,n+1}R^{\#}(t^{-E}(x-y))) \\
&\qquad \times\Big[\int_0^{t/2}\int_{\mathbb{V}}(t-s)^{-(\mu_H+1-\rho)}s^{-(\mu_H+1-n\rho)}\times\exp(-\epsilon MR^{\#}(s^{-E}(z-y)))\,dz\,ds \\
&\qquad + \int_{t/2}^{t}\int_{\mathbb{V}}(t-s)^{-(\mu_H+1-\rho)}s^{-(\mu_H+1-n\rho)}\exp(-\epsilon MR^{\#}((t-s)^{-E}(x-z)))\,dz\,ds\Big]
\end{aligned} \tag{43}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$, where we have used the fact that $R^{\#}$ is non-negative. Let us focus our attention on the first term above. For $0 \leq s \leq t/2$,

$$(t-s)^{-(\mu_H+1-\rho)}s^{-(\mu_H+1-n\rho)} \leq (t/2)^{-(\mu_H+1-\rho)}s^{-(\mu_H+1-n\rho)}$$

because $\mu_H + 1 - \rho > 0$. Consequently,

$$\begin{aligned}
&\int_0^{t/2}\int_{\mathbb{V}}(t-s)^{-(\mu_H+1-\rho)}s^{-(\mu_H+1-n\rho)}\exp(-\epsilon MR^{\#}(s^{-E}(z-y)))\,dz\,ds \\
&\leq (t/2)^{-(\mu_H+1-\rho)}\int_0^{t/2}s^{-(\mu_H+1-n\rho)}\int_{\mathbb{V}}\exp(-\epsilon MR^{\#}(s^{-E}(z-y)))\,dz\,ds
\end{aligned}$$

$$\leq (t/2)^{-(\mu_H+1-\rho)} \int_0^{t/2} s^{n\rho-1} \int_{\mathbb{V}} \exp(-\epsilon M R^{\#}(z))\, dz\, ds$$

$$\leq t^{-(\mu_H+1-(n+1)\rho)} \frac{2^{(\mu_H+1-(n+1)\rho)}}{n\rho} \int_{\mathbb{V}} \exp(-\epsilon M R^{\#}(z))\, dz\, ds \tag{44}$$

for all $y \in \mathbb{V}$ and $t > 0$. We note that the second inequality is justified by making the change of variables $z \mapsto s^{-E}(z-y)$ (thus canceling the term $s^{-\operatorname{tr} E} = s^{-\mu_H}$ in the integral over s) and the final inequality is valid because $n\rho - 1 > \rho - 1 > -1$. By similar reasoning, we obtain

$$\int_{t/2}^{t} \int_{\mathbb{V}} (t-s)^{-(\mu_H+1-n\rho)} s^{-(\mu_H+1-\rho)} \exp(-\epsilon M R^{\#}((t-s)^{-E}(x-z)))\, dz\, ds$$

$$\leq t^{-(\mu_H+1-(n+1)\rho)} \frac{2^{(\mu_H+1-(n+1)\rho)}}{\rho} \int_{\mathbb{V}} \exp(-\epsilon M R^{\#}(z))\, dz\, ds \tag{45}$$

for all $x \in \mathbb{V}$ and $t > 0$. Upon combining the estimates (43), (44) and (45), we have

$$|K_{n+1}(t,x,y)| \leq C_{n+1}(\epsilon) t^{-(\mu_H+1-(n+1)\rho)} \exp(-M_{\epsilon,n+1} R^{\#}(t^{-E}(x-y))$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$ where we have put

$$C_{n+1}(\epsilon) = C_1(\epsilon) C_n(\epsilon) \frac{n+1}{n\rho} 2^{\mu_H+(1-(n+1)\rho)} \int_{\mathbb{V}} \exp(-\epsilon M R^{\#}(z))\, dz$$

and made use of Corollary 3.11. □

Remark 7 The estimate (41) is an important one and will be used again. In the context of elliptic operators, i.e., where $R^{\#}(x) = C_m |x|^{2m/(2m-1)}$, the analogous result is captured in Lemma 5.1 of [27]. It is interesting to note that Eidelman worked somewhat harder to prove it. Perhaps this is because the appearance of the Legendre-Fenchel transform wasn't noticed.

It is clear from the previous lemma that for sufficiently large n, K_n is bounded by a positive power of t. The first such n is $\bar{n} := \lceil \rho^{-1}(\operatorname{tr} E + 1) \rceil$. In view of the previous lemma,

$$|K_{\bar{n}}(t,x,y)| \leq C_{\bar{n}}(\epsilon) \exp(-M(1-\epsilon\bar{n}) R^{\#}(t^{-E}(x-y)))$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$ where we have adjusted $C_{\bar{n}}(\epsilon)$ to account for this positive power of t. Let $\delta < 1/2$ and set

$$\epsilon = \frac{\delta}{\bar{n}}, \qquad M_1 = M(1-\delta) \quad \text{and} \quad C_0 = \max_{1 \leq n \leq \bar{n}} C_n(\epsilon).$$

Upon combining preceding estimate with the estimates (38) and (41), we have

$$
\begin{aligned}
&|K_{\bar{n}+1}(t,x,y)| \\
&\leq C_0^2 \int_0^t \int_{\mathbb{V}} (t-s)^{-(\mu_H+(1-\rho))} \\
&\qquad\qquad \times \exp(-MR^{\#}((t-s)^{-E}(x-z)) \exp(-M(1-\epsilon\bar{n})R^{\#}(s^{-E}(z-y))) \, ds \, dz \\
&\leq C_0^2 \exp(-M_1 R^{\#}(t^{-E}(x-y))) \int_0^t \int_{\mathbb{V}} (t-s)^{-(\mu_H+(1-\rho))} \exp(-C\delta R^{\#}((t-s)^{-E}(z))) \, dz \, ds \\
&\leq C_0(C_0 F) \frac{t^{\rho}}{\rho} \exp(-M_1 R^{\#}(t^{-E}(x-y)))
\end{aligned}
$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$ where

$$F = \int_{\mathbb{V}} \exp(-M\delta R^{\#}(z)) \, dz < \infty.$$

Let us take this a little further.

Lemma 5.8 *For every* $k \in \mathbb{N}_+$,

$$|K_{\bar{n}+k}(t,x,y)| \leq \frac{C_0}{\Gamma(\rho)} \frac{(C_0 F \Gamma(\rho))^k}{k!} t^{\rho k} \exp(-M_1 R^{\#}(t^{-E}(x-y))) \tag{46}$$

for all $x, y \in \mathbb{V}$ *and* $0 < t \leq T$. *Here* $\Gamma(\cdot)$ *denotes the Gamma function.*

Proof The Euler-Beta function $B(\cdot,\cdot)$ satisfies the well-known identity $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. Using this identity, one quickly obtains the estimate

$$\prod_{j=1}^{k-1} B(\rho, 1+j\rho) = \frac{\Gamma(\rho)^{k-1}}{\Gamma(1+k\rho)} \leq \frac{\Gamma(\rho)^{k-1}}{k!}.$$

It therefore suffices to prove that

$$|K_{\bar{n}+k}(t,x,y)| \leq C_0(C_0 F)^k \prod_{j=0}^{k-1} B(\rho, 1+j\rho) t^{k\rho} \exp(-M_1 R^{\#}(t^{-E}(x-y))) \tag{47}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$.

We first note that $B(\rho, 1) = \rho^{-1}$ and so, for $k = 1$, (47) follows directly from the calculation proceeding the lemma. We shall induct on k. By another application of

(38) and (41), we have

$$\begin{aligned}
J_{k+1}(t,x,y) &:= \left[C_0^2(C_0F)^k \prod_{j=0}^{k-1} B(\rho, 1+j\rho)\right]^{-1} |K_{\bar{n}+k+1}(t,x,y)| \\
&\leq \int_0^t \int_{\mathbb{V}} (t-s)^{-(\mu_H+(1-\rho))} s^{-k\rho} \exp(-MR^{\#}((t-s)^{-E}(x-z))) \\
&\qquad\qquad \times \exp(-M_1 R^{\#}(s^{-E}(z-y)))\, dz\, ds \\
&\leq \exp(-M_1 R^{\#}(t^{-E}(x-y))) \\
&\qquad \times \int_0^t \int_{\mathbb{V}} (t-s)^{-(\mu_H+(1-\rho))} s^{-k\rho} \exp(-M\delta R^{\#}((t-s)^{-E}(x-z)))\, dz\, ds
\end{aligned}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$. Upon making the changes of variables $z \to (t-s)^{-E}(x-z)$ followed by $s \to s/t$, we have

$$\begin{aligned}
J_{k+1}(t,x,y) &\leq \exp(-M_1 R^{\#}(t^{-E}(x-y))) F \int_0^1 (t-st)^{\rho-1} (st)^{k\rho} t\, ds \\
&\leq \exp(-M_1 R^{\#}(t^{-E}(x-y))) F t^{(k+1)\rho} B(\rho, 1+k\rho)
\end{aligned}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$. Therefore (47) holds for $k+1$ as required. □

Proposition 5.9 *Let* $\phi : (0,T] \times \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ *be defined by*

$$\phi = \sum_{n=1}^{\infty} K_n.$$

This series converges uniformly for $x, y \in \mathbb{V}$ *and* $t_0 \leq t \leq T$ *where* t_0 *is any positive constant. There exists* $C \geq 1$ *for which*

$$|\phi(t,x,y)| \leq \frac{C}{t^{\mu_H+(1-\rho)}} \exp(-M_1 R^{\#}(t^{-E}(x-y))) \tag{48}$$

for all $x, y \in \mathbb{V}$ *and* $0 < t \leq T$ *where* M_1 *and* ρ *are as in the previous lemmas. Moreover, the identity*

$$\phi(t,x,y) = K(t,x,y) + \int_0^t \int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\, dz\, ds \tag{49}$$

holds for all $x, y \in \mathbb{V}$ *and* $0 < t \leq T$.

Proof Using Lemmas 5.7 and 5.8 we see that

$$\sum_{k=1}^{\infty}|K_n(t,x,y)| \le C_0\Big[\sum_{n=1}^{\bar{n}} t^{-(\mu_H+(1-n\rho))} + \frac{1}{\Gamma(\rho)}\sum_{k=1}^{\infty}\frac{(C_0F\Gamma(\rho))^k}{k!}t^{k\rho}\Big]\exp(-M_1R^{\#}(t^{-E}(x-y)))$$

for all $x,y \in \mathbb{V}$ and $0 < t \le T$ from which (48) and our assertion concerning uniform convergence follow. A similar calculation and an application of Tonelli's theorem justify the following use of Fubini's theorem: For $x,y \in \mathbb{V}$ and $0 < t \le T$,

$$\begin{aligned}\int_0^t\int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,ds\,dz &= \sum_{n=1}^{\infty}\int_0^t\int_{\mathbb{V}} K(t-s,x,z)K_n(s,z,y)\,dz\,ds\\ &= \sum_{n=1}^{\infty}K_{n+1}(t,z,y) = \phi(t,x,y) - K(t,x,y)\end{aligned}$$

as desired. □

The following Hölder continuity estimate for ϕ is obtained by first showing the analogous estimate for K and then deducing the desired result from the integral formula (49). As the proof is similar in character to those of the preceding two lemmas, we omit it. A full proof can be found in [28, p. 80]. We also note here that the result is stronger than is required for our purposes (see its use in the proof of Lemma 5.12). All that is really required is that $\phi(\cdot,\cdot,y)$ satisfies the hypotheses (for f) in Lemma 5.11 for each $y \in \mathbb{V}$.

Lemma 5.10 *There exists $\alpha \in \mathbb{I}_+^d$ which is consistent with* $\mathbf{m}$*, $0 < \eta < 1$ and $C \ge 1$ such that*

$$|\phi(t,x+h,y) - \phi(t,x,y)| \le \frac{C}{t^{\mu_H+(1-\eta)}}|h|_{\mathbf{v}}^{\alpha}\exp(-M_1R^{\#}(t^{-E}(x-y)))$$

for all $x,y,h \in \mathbb{V}$ and $0 < t \le T$.

Step 3: Verifying That Z Is a Fundamental Solution to (22)

Lemma 5.11 *Let $\alpha \in \mathbb{I}_+^d$ be consistent with* $\mathbf{m}$ *and, for $t_0 > 0$, let $f : [t_0,T]\times\mathbb{V} \to \mathbb{C}$ be bounded and continuous. Moreover, suppose that f is uniformly* $\mathbf{v}$*-Hölder continuous in x on $[t_0,T]\times\mathbb{V}$ of order α, by which we mean that there is a constant $C > 0$ such that*

$$\sup_{t\in[t_0,T]}|f(t,x)-f(t,y)| \le C|x-y|_{\mathbf{v}}^{\alpha}$$

for all $x, y \in \mathbb{V}$*. Then* $u : [t_0, T] \times \mathbb{V} \to \mathbb{C}$ *defined by*

$$u(t,x) = \int_{t_0}^{t} \int_{\mathbb{V}} G_p(t-s, x-z; z) f(s,z)\, dz\, ds$$

is $(2\mathbf{m}, \mathbf{v})$*-regular on* $(t_0, T) \times \mathbb{V}$*. Moreover,*

$$\partial_t u(t,x) = f(t,x) + \lim_{h \downarrow 0} \int_{t_0}^{t-h} \int_{\mathbb{V}} \partial_t G_p(t-s, x-z; z) f(s,z)\, dz\, ds \tag{50}$$

and for any β *such that* $|\beta : \mathbf{m}| \leq 2$*, we have*

$$D_{\mathbf{v}}^{\beta} u(t,x) = \lim_{h \downarrow 0} \int_{t_0}^{t-h} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) f(s,z)\, dz\, ds \tag{51}$$

for $x \in \mathbb{V}$ *and* $t_0 < t < T$*.*

Before starting the proof, let us observe that, for each multi-index β, $D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) f(s,z)$ is locally uniformly (in x) dominated by the function

$$h_\beta(s,z;t,x) := C(t-s)^{-(\mu_H + |\beta : 2\mathbf{m}|)} \exp(-MR^{\#}((t-s)^{-E}(x-z)))$$

for $x, z \in \mathbb{V}$ and $t_0 \leq s \leq t \leq T$, where the constant $C > 0$ depends on β, $\|f\|_\infty$ and the bound for $D_{\mathbf{v}}^{\beta} G_p$ yielded by Proposition 2.11 which can be seen to hold locally uniformly by an argument analogous to that given in the proof of Proposition 8.10 of [44]. We observe that

$$\begin{aligned}
\int_{t_0}^{t} \int_{\mathbb{V}} h_\beta(s,z;t,x)\, dz\, ds &= C \int_{t_0}^{t} \int_{\mathbb{V}} (t-s)^{-(\mu_H + |\beta : 2\mathbf{m}|)} \exp(-MR^{\#}((t-s)^{-E}(x-z)))\, dz\, ds \\
&\leq C \int_{t_0}^{t} \int_{\mathbb{V}} (t-s)^{-|\beta : 2\mathbf{m}|} \exp(-MR^{\#}(z))\, dz\, ds \\
&\leq C \int_{t_0}^{t} (t-s)^{-|\beta : 2\mathbf{m}|}\, ds
\end{aligned}$$

for all $t_0 \leq t \leq T$ and $x \in \mathbb{V}$. When $|\beta : \mathbf{m}| < 2$,

$$\int_{t_0}^{t} (t-s)^{-|\beta : 2\mathbf{m}|}\, ds \tag{52}$$

converges and so, in view of the fact that $D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) f(s,z)$ is locally uniformly dominated by h_β, we may conclude that

$$D_{\mathbf{v}}^{\beta} u(t,x) = \int_{t_0}^{t} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, z-x; z) f(s,z)\, dz\, ds$$

for all $t_0 \leq t \leq T$ and $x \in \mathbb{V}$. From this it follows that $D_{\mathbf{v}}^{\beta} u(t,x)$ is continuous on $(t_0, T) \times \mathbb{V}$ and moreover (51) holds for such an β in view of Lebesgue's dominated convergence theorem. When $|\beta : \mathbf{m}| = 2$, (52) does not converge and hence the above argument fails. The main issue in the proof below centers around using **v**-Hölder continuity to remove this obstacle.

Proof Our argument proceeds in two steps. The fist step deals with the spatial derivatives of u. Therein, we prove the asserted x-regularity and show that the formula (51) holds. In fact, we only need to consider the case where $|\beta : \mathbf{m}| = 2$; the case where $|\beta : \mathbf{m}| < 2$ was already treated in the paragraph proceeding the proof. In the second step, we address the time derivative of u. As we will see, (50) and the asserted t-regularity are partial consequences of the results proved in Step 1; this is, in part, due to the fact that the time derivative of G_p can be exchanged for spatial derivatives. The regularity shown in the two steps together will automatically ensure that u is $(2\mathbf{m}, \mathbf{v})$-regular on $(t_0, T) \times \mathbb{V}$.

Step 1 Let β be such that $|\beta : \mathbf{m}| = 2$. For $h > 0$ write

$$u_h(t,x) = \int_{t_0}^{t-h} \int_{\mathbb{V}} G_p(t-s, x-z; z) f(s,z)\, dz\, ds$$

and observe that

$$D_{\mathbf{v}}^{\beta} u_h(t,x) = \int_{t_0}^{t-h} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) f(s,z)\, dz\, ds$$

for all $t_0 \leq t - h < t \leq T$ and $x \in \mathbb{V}$; it is clear that $D_{\mathbf{v}}^{\beta} u_h(t,x)$ is continuous in t and x. The fact that we can differentiate under the integral sign is justified because t has been replaced by $t-h$ and hence the singularity in (52) is avoided in the upper limit. We will show that $D_{\mathbf{v}}^{\beta} u_h(t,x)$ converges uniformly on all compact subsets of $(t_0, T) \times \mathbb{V}$ as $h \to 0$. This, of course, guarantees that $D_{\mathbf{v}}^{\beta} u(x,t)$ exists, satisfies (51) and is continuous on $(t_0, T) \times \mathbb{V}$. To this end, let us write

$$\begin{aligned} D_{\mathbf{v}}^{\beta} u_h(t,x) &= \int_{t_0}^{t-h} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)(f(s,z) - f(s,x))\, dz\, ds \\ &\quad + \int_{t_0}^{t-h} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) f(s,x)\, dz\, ds \\ &=: I_h^{(1)}(t,x) + I_h^{(2)}(t,x). \end{aligned}$$

Using our hypotheses, Corollary 3.8 and Lemma 3.13, for some $0 < \sigma < 1$ and $\theta > 0$, there is $M > 0$ such that

$$|f(s,z) - f(s,x)| \leq C(t-s)^{\sigma} (R^{\#}((t-s)^{-E}(x-z)))^{\theta}$$

for all $x, z \in \mathbb{V}$, $t \in [t_0, T]$ and $s \in [t_0, t]$. In view of the preceding estimate and Lemma 5.3, we have

$$\begin{aligned}
&|D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)(f(s,z) - f(s,x))| \\
&\quad \le C(t-s)^{-(\mu_H+1)}(t-s)^{\sigma}(R^{\#}((t-s)^{-E}(x-z)))^{\theta} \exp(-MR^{\#}((t-s)^{-E}(x-z))) \\
&\quad \le C(t-s)^{-(\mu_H+(1-\sigma))} \exp(-MR^{\#}(t-s)^{-E}(x-z))
\end{aligned}$$

for all $x, z \in \mathbb{V}$, $t \in [t_0, T]$ and $s \in [t_0, t]$, where C and M are positive constants. We then observe that

$$\begin{aligned}
&\int_{t_0}^{t} \int_{\mathbb{V}} |D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)(f(s,z) - f(s,x))| \, dz \, ds \\
&\qquad \le C \int_{t_0}^{t} (t-s)^{-(\mu_H+(1-\sigma))} \int_{\mathbb{V}} \exp(-MR^{\#}((t-s)^{-E}(x-z))) \, dz \, ds \\
&\qquad \le C \int_{t_0}^{t} (t-s)^{\sigma-1} \int_{\mathbb{V}} \exp(-MR^{\#}(z)) \, dz \, ds \\
&\qquad \le \frac{C(t-t_0)^{\sigma}}{\sigma} \int_{\mathbb{V}} \exp(-MR^{\#}(z)) \, dz \\
&\qquad \le \frac{C(T-t_0)^{\sigma}}{\sigma} \int_{\mathbb{V}} \exp(-MR^{\#}(z)) \, dz < \infty
\end{aligned}$$

for all $t \in [t_0, T]$ and $x \in \mathbb{V}$, where the validity of the second inequality is seen by making the change of variables $z \mapsto (t-s)^{-E}(x-z)$ and canceling the term $(t-s)^{-\mu_H} = (t-s)^{-\operatorname{tr} E}$. Consequently,

$$I^{(1)}(t,x) := \int_{t_0}^{t} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)(f(s,z) - f(s,x)) \, dz \, ds$$

exists for each $t \in [t_0, T]$ and $x \in \mathbb{V}$. Moreover, for all $t_0 \le t-h < t \le T$ and $x \in \mathbb{V}$,

$$\begin{aligned}
|I^{(1)}(t,x) - I_h^{(1)}(t,x)| &\le \int_{t-h}^{t} \int_{\mathbb{V}} |D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)(f(s,z) - f(s,x))| \, dz \, ds \\
&\le C \int_{t-h}^{t} \int_{\mathbb{V}} (t-s)^{\sigma-1} \exp(-MR^{\#}(z)) \, dz \, ds \le Ch^{\sigma}.
\end{aligned}$$

From this we see that $\lim_{h \downarrow 0} I_h^{(1)}(t,x)$ converges uniformly on all compact subsets of $(t_0, T) \times \mathbb{V}$.

We claim that for some $0 < \rho < 1$, there exists $C > 0$ such that

$$\left| \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) \, dz \right| \le C(t-s)^{-(1-\rho)} \tag{53}$$

for all $x \in \mathbb{V}$ and $s \in [t_0, t]$. Indeed,

$$\int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)\, dz$$
$$= \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} [G_p(t-s, x-z; z) - G_p(t-s, x-z; y)]\big|_{y=x}\, dz + \big[D_{\mathbf{v}}^{\beta} \int_{\mathbb{V}} G_p(t-s, x-z; y)\, dz\big]\big|_{y=x}.$$

The first term above is estimated with the help of Lemma 5.4 and by making arguments analogous to those in the previous paragraph; the appearance of ρ follows from an obvious application of Lemma 3.13. This term is bounded by $C(t-s)^{-(1-\rho)}$. The second term is clearly zero and so our claim is justified.

By essentially repeating the arguments made for $I_h^{(1)}$ and making use of (53), we see that

$$\lim_{h \downarrow 0} I_h^{(2)}(t, x) = I^{(2)}(t, x) =: \int_{t_0}^{t} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) f(s, x)\, dz\, ds$$

where this limit converges uniformly on all compact subsets of $(t_0, T) \times \mathbb{V}$.

Step 2 It follows from Leibnitz' rule that

$$\begin{aligned} \partial_t u_h(x, t) &= \int_{\mathbb{V}} G_p(h, x-z; z) f(t-h, z)\, dz + \int_{t_0}^{t-h} \int_{\mathbb{V}} \partial_t G_p(t-s, x-z; z) f(s, z)\, dz\, ds \\ &=: J_h^{(1)}(t, x) + J_h^{(2)}(t, x) \end{aligned}$$

for all $t_0 < t-h < t < T$ and $x \in \mathbb{V}$. Now, in view of Lemma 5.5 and our hypotheses concerning f,

$$\lim_{h \downarrow 0} J_h^{(1)}(t, x) = f(t, x)$$

where this limit converges uniformly on all compact subsets of $(t_0, T) \times \mathbb{V}$.

Using the fact that $\partial_t G_p(t-s, x-z; z) = -H_p(z) G_p(t-s, x-z; z)$, we see that

$$\begin{aligned} \lim_{h \downarrow 0} J_h^{(2)}(t, x) &= \lim_{h \downarrow 0} \int_0^{t-h} \int_{\mathbb{V}} \Big(- \sum_{|\beta:\mathbf{m}|=2} a_{\beta}(z) D_{\mathbf{v}}^{\beta} \Big) G_p(t-s, x-z; z) f(s, z)\, dz\, ds \\ &= - \sum_{|\beta:\mathbf{m}|=2} \lim_{h \downarrow 0} \int_0^{t-h} \int_{\mathbb{V}} D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z) (a_{\beta}(z) f(s, z))\, dz\, ds \end{aligned}$$

for all $t \in (t_0, T)$ and $x \in \mathbb{V}$. Because the coefficients of H are $\mathbf{v}$-Hölder continuous and bounded, for each β, $a_{\beta}(z) f(s, z)$ satisfies the same condition we have required for f and so, in view of Step 1, it follows that $J_h^{(2)}(t, x)$ converges uniformly on all compact subsets of $(t_0, T) \times \mathbb{V}$ as $h \to 0$. We thus conclude that $\partial_t u(t,x)$ exists, is continuous on $(t_0, T) \times \mathbb{V}$ and satisfies (50). □

Lemma 5.12 *Let* $W : (0,T] \times \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ *be defined by*

$$W(t,x,y) = \int_0^t \int_{\mathbb{V}} G_p(t-s, x-z; z)\phi(s,z,y)\, dz\, ds,$$

for $x, y \in \mathbb{V}$ *and* $0 < t \le T$. *Then, for each* $y \in \mathbb{V}$, $W(\cdot,\cdot,y)$ *is* $(2\mathbf{m}, \mathbf{v})$*-regular on* $(0,T) \times \mathbb{V}$ *and satisfies*

$$(\partial_t + H)W(t,x,y) = K(t,x,y). \tag{54}$$

for all $x, y \in \mathbb{V}$ *and* $t \in (0,T)$. *Moreover, there are positive constants* C *and* M *for which*

$$|W(t,x,y)| \le Ct^{-\mu_H + \rho} \exp(-MR^{\#}(t^{-E}(x-y))) \tag{55}$$

for all $x, y \in \mathbb{V}$ *and* $0 < t \le T$ *where* ρ *is that which appears in Lemma 5.7.*

Proof The estimate (55) follows from (30) and (48) by an analogous computation to that done in the proof of Lemma 5.7. It remains to show that, for each $y \in \mathbb{V}$, $W(\cdot,\cdot,y)$ is $(2\mathbf{m}, \mathbf{v})$-regular and satisfies (54) on $(0,T) \times \mathbb{V}$. These are both local properties and, as such, it suffices to examine them on $(t_0, T) \times \mathbb{V}$ for an arbitrary but fixed $t_0 > 0$. Let us write

$$\begin{aligned} W(t,x,y) &= \int_{t_0}^t \int_{\mathbb{V}} G_p(t-s, x-z; z)\phi(s,z,y)\, dz\, ds \\ &\quad + \int_0^{t_0} \int_{\mathbb{V}} G_p(t-s, x-z; z)\phi(s,z,y)\, dz\, ds \\ &=: W_1(t,x,y) + W_2(t,x,y) \end{aligned}$$

for $x, y \in \mathbb{V}$ and $t_0 < t < T$. In view of Lemmas 5.10 and 5.11, for each $y \in \mathbb{V}$, $W_1(\cdot,\cdot,y)$ is $(2\mathbf{m}, \mathbf{v})$-regular on $(t_0, T) \times \mathbb{V}$ and

$$\begin{aligned} (\partial_t + H)W_1(t,x,y) &= \partial_t W_1(t,x,y) + \sum_{|\beta:\mathbf{m}| \le 2} a_\beta(x) D_{\mathbf{v}}^\beta W_1(t,x,y) \\ &= \phi(t,x,y) + \lim_{h \downarrow 0} \int_{t_0}^{t-h} \int_{\mathbb{V}} \partial_t G_p(t-s, x-z; z)\phi(s,z,y)\, dz\, dy \\ &\quad + \lim_{h \downarrow 0} \int_{t_0}^{t-h} \int_{\mathbb{V}} \sum_{|\beta:\mathbf{m}| \le 2} a_\beta(x) D_{\mathbf{v}}^\beta G_p(t-s, x-z; z)\phi(s,z,y)\, dz\, ds \end{aligned}$$

$$= \phi(t,x,y) + \lim_{h\downarrow 0}\int_{t_0}^{t-h}\int_{\mathbb{V}} (\partial_t + H)G_p(t-s, x-z; z)\phi(s,z,y)\,dz\,ds$$

$$= \phi(t,x,y) - \lim_{h\downarrow 0}\int_{t_0}^{t-h}\int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds \tag{56}$$

for all $x \in \mathbb{V}$ and $t_0 < t < T$; here we have used the fact that

$$(\partial_t + H)G_p(t-s, x-z; z) = -K(t-s,x,z).$$

Treating W_2 is easier because $\partial_t G_p(t-s, x-z; z)$ and, for each multi-index β, $D_{\mathbf{v}}^{\beta} G_p(t-s, x-z; z)$ are, as functions of s and z, absolutely integrable on $(0, t_0] \times \mathbb{V}$ for every $t \in (t_0, T]$ and $x \in \mathbb{V}$ by virtue of Lemma 5.3. Consequently, derivatives may be taken under the integral sign and so it follows that, for each $y \in \mathbb{V}$, $W_2(\cdot,\cdot,y)$ is $(2\mathbf{m}, \mathbf{v})$-regular on $(t_0, T) \times \mathbb{V}$ and

$$(\partial_t + H)W_2(t,x,y) = -\int_0^{t_0}\int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds \tag{57}$$

for $x \in \mathbb{V}$ and $t_0 < t < T$. We can thus conclude that, for each $y \in \mathbb{V}$, $W(\cdot,\cdot,y)$ is $(2\mathbf{m}, \mathbf{v})$-regular on $(t_0, T) \times \mathbb{V}$ and, by combining (56) and (57),

$$(\partial_t + H)W(t,x,y) = \phi(t,x,y) - \lim_{h\downarrow 0}\int_0^{t-h}\int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds$$

for $x \in \mathbb{V}$ and $t_0 < t < T$. By (38), Proposition 5.9 and the Dominated Convergence Theorem,

$$\lim_{h\downarrow 0}\int_0^{t-h}\int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds = \int_0^{t}\int_{\mathbb{V}} K(t-s,x,z)\phi(s,z,y)\,dz\,ds$$
$$= \phi(t,x,y) - K(t,x,y)$$

and therefore

$$(\partial_t + H)W(t,x,y) = K(t,x,y)$$

for all $x, y \in \mathbb{V}$ and $t_0 < t < T$. □

The theorem below is our main result. It is a more refined version of Theorem 5.1 because it gives an explicit formula for the fundamental solution Z; in particular Theorem 5.1 is an immediate consequence of the result below.

Theorem 5.13 *Let H be a uniformly $(2\mathbf{m}, \mathbf{v})$-positive-semi-elliptic operator. If H satisfies Hypothesis 5.2 then $Z : (0, T] \times \mathbb{V} \times \mathbb{V} \to \mathbb{C}$, defined by*

$$Z(t,x,y) = G_p(t, x-y; y) + W(t,x,y) \tag{58}$$

for $x, y \in \mathbb{V}$ and $0 < t \leq T$, is a fundamental solution to (22). Moreover, there are positive constants C and M for which

$$|Z(t,x,y)| \leq \frac{C}{t^{\mu_H}} \exp\left(-tMR^{\#}\left(\frac{x-y}{t}\right)\right) \tag{59}$$

for all $x, y \in \mathbb{V}$ and $0 < t \leq T$.

Proof As $0 < \rho < 1$, (55) and Lemma 5.3 imply the estimate (59). In view of Lemma 5.12 and Corollary 5.6, for each $y \in \mathbb{V}$, $Z(\cdot,\cdot,y)$ is $(2\mathbf{m},\mathbf{v})$-regular on $(0,T)\times\mathbb{V}$ and

$$\begin{aligned}(\partial_t + H)Z(t,x,y) &= (\partial_t + H)G_p(t,x-y,y) + (\partial_t + H)W(t,x,y)\\ &= -K(t,x,y) + K(t,x,y) = 0\end{aligned}$$

for all $x \in \mathbb{V}$ and $0 < t < T$. It remains to show that for any $f \in C_b(\mathbb{V})$,

$$\lim_{t\to 0}\int_{\mathbb{V}} Z(t,x,y)f(y)\,dy = f(x)$$

for all $x \in \mathbb{V}$. Indeed, let $f \in C_b(\mathbb{V})$ and, in view of (55), observe that

$$\begin{aligned}\left|\int_{\mathbb{V}} W(t,x,y)f(y)\right| &\leq Ct^{\rho}\|f\|_\infty \int_{\mathbb{V}} t^{-\mu_H}\exp(-MR^{\#}(t^{-E}(x-y)))dy\\ &\leq Ct^{\rho}\|f\|_\infty \int_{\mathbb{V}} \exp(-MR^{\#}(y))\,dy \leq Ct^{\rho}\|f\|_\infty\end{aligned}$$

for all $x \in \mathbb{V}$ and $0 < t \leq T$. An appeal to Lemma 5.5 gives, for each $x \in \mathbb{V}$,

$$\begin{aligned}\lim_{t\to 0}\int_{\mathbb{V}} Z(t,x,y)f(y)dy &= \lim_{t\to 0}\int_{\mathbb{V}} G_p(t,x-y;y)f(y)\,dy + \lim_{t\to 0}\int_{\mathbb{V}} W(t,x,y)f(y)\,dy\\ &= f(x) + 0 = f(x)\end{aligned}$$

as required. In fact, the above argument guarantees that this convergence happens uniformly on all compact subsets of $\mathbb{V}$. □

We remind the reader that implicit in the definition of fundamental solution to (22) is the condition that Z is $(2\mathbf{m},\mathbf{v})$-regular. In fact, one can further deduce estimates for the spatial derivatives of Z, $D_{\mathbf{v}}^{\beta}Z$, of the form (13) for all β such that $|\beta : 2\mathbf{m}| \leq 1$ (see [28, p. 92]). Using the fact that Z satisfies (22) and H's coefficients are bounded, an analogous estimate is obtained for a single t derivative of Z.

Acknowledgements We thank the anonymous referee for many helpful comments. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144153 (Evan Randles) and the National Science Foundation under Grant No. DMS-1404435 (Laurent Saloff-Coste).

References

1. T. Apel, T.G. Flaig, S. Nicaise, A priori error estimates for finite element methods for H (2,1)-elliptic equations. Numer. Funct. Anal. Optim. **35**(2), 153–176 (2014)
2. R.A. Artino, On semielliptic boundary value problems. J. Math. Anal. Appl. **42**(3), 610–626 (1973)
3. R.A. Artino, Completely semielliptic boundary value problems. Port. Math. **50**(2), 185–192 (1993)
4. R.A. Artino, J. Barros-Neto. Regular semielliptic boundary value problems. J. Math. Anal. Appl. **61**(1), 40–57 (1977)
5. R.A. Artino, J. Barros-Neto, Semielliptic pseudodifferential operators. J. Funct. Anal. **129**(2), 471–496 (1995)
6. P. Auscher, A. ter Elst, D. Robinson, On positive Rockland operators. Colloq. Math. **67**(2), 197–216 (1994)
7. P. Auscher, S. Hofmann, A. McIntosh, P. Tchamitchian, The Kato square root problem for higher order elliptic operators and systems on $\mathbb{R}^n$. J. Evol. Equ. **1**(4), 361–385 (2001)
8. G. Barbatis, E.B. Davies, Sharp bounds on heat kernels of higher order uniformly elliptic operators. J. Oper. Theory **36**(1), 179–198 (1996)
9. S. Blunck, P.C. Kunstmann, Generalized Gaussian estimates and the Legendre transform. J. Oper. Theory **53**(2), 351–365 (2005)
10. L.N. Bondar', Conditions for the solvability of boundary value problems for quasi-elliptic systems in a half-space. Differ. Equ. **48**(3), 343–353 (2012)
11. L.N. Bondar', Solvability of boundary value problems for quasielliptic systems in weighted Sobolev spaces. J. Math. Sci. **186**(3), 364–378 (2012)
12. L.N. Bondar, G.V. Demidenko, Boundary value problems for quasielliptic systems. Sib. Math. J. **49**(2), 202–217 (2008)
13. F.E. Browder, The asymptotic distribution of eigenfunctions and eigenvalues for semi-elliptic differential operators. Proc. Nat. Acad. Sci. U.S.A. **43**(3), 270–273 (1957)
14. A. Cavallucci, Sulle proprietà differenziali delle soluzioni delle equazioni quasi-ellittiche. Ann. Mat. Pura Appl. Ser. 4 **67**(1), 143–167 (1965)
15. E.B. Davies, Uniformly elliptic operators with measurable coefficients. J. Funct. Anal. **132**(1), 141–169 (1995)
16. E.B. Davies, Lp spectral theory of higher-order elliptic differential operators. Bull. Lond. Math. Soc. **29**(5), 513–546 (1997)
17. G.V. Demidenko, Correct solvability of boundary-value problems in a halfspace for quasielliptic equations. Sib. Math. J. **29**(4), 555–567 (1989)
18. G.V. Demidenko, Integral operators determined by quasielliptic equations. I. Sib. Math. J. **34**(6), 1044–1058 (1993)
19. G.V. Demidenko, Integral operators determined by quasielliptic equations. II. Sib. Math. J. **35**(1), 37–61 (1994)
20. G.V. Demidenko, On quasielliptic operators in Rn. Sib. Math. J. **39**(5), 884–893 (1998)
21. G.V. Demidenko, Isomorphic properties of one class of differential operators and their applications. Sib. Math. J. **42**(5), 865–883 (2001)
22. G.V. Demidenko, Mapping properties of quasielliptic operators and applications. Int. J. Dyn. Syst. Differ. Equ. **1**(1), 58 (2007)
23. G.V. Demidenko, Quasielliptic operators and Sobolev type equations. Sib. Math. J. **49**(5), 842–851 (2008)
24. G.V. Demidenko, Quasielliptic operators and Sobolev type equations. II. Sib. Math. J. **50**(5), 838–845 (2009)
25. J. Dziubański, A. Hulanicki, On semigroups generated by left-invariant positive differential operators on nilpotent Lie groups. Stud. Math. **94**(1), 81–95 (1989)
26. S.D. Eidelman, On a class of parabolic systems. Sov. Math. Dokl. **1**, 85–818 (1960)
27. S.D. Eidelman, *Parabolic Systems* (North-Holland/Wolters-Noordhoff, Amsterdam, 1969)

28. S.D. Eidelman, A.N. Kochubei, S.D. Ivasyshen, *Analytic Methods in the Theory of Differential and Pseudo-Differential Equations of Parabolic Type*, vol. 1 (Birkhäuser, Basel, 2004)
29. G.B. Folland, E.M. Stein, *Hardy Spaces on Homogeneous Groups* (Princeton University Press, Princeton, 1982)
30. A. Friedman, *Partial Differential Equations of Parabolic Type* (Prentice-Hall, Englewood Cliffs, 1964)
31. E. Giusti, Equazioni quasi ellittiche e spazi $L^{p,\theta}(\Omega,\delta)$ (I). Ann. Mat. Pura Appl. Ser. 4 **75**(1), 313–353 (1967)
32. W. Hebisch, Sharp pointwise estimate for the heat kernels of the semigroup generated by sums of even powers of vector fields on homogeneous groups. Studia Math. **95**, 93–106 (1989)
33. W. Hebisch, L. Saloff-Coste, Gaussian estimates for Markov chains and random walks on groups. Ann. Probab. **21**(2), 673–709 (1993)
34. G.N. Hile, Fundamental solutions and mapping properties of semielliptic operators. Math. Nachr. **279**(13–14), 1538–1564 (2006)
35. G.N. Hile, C.P. Mawata, C. Zhou, A priori bounds for semielliptic operators. J. Differ. Equ. **176**(1), 29–64 (2001)
36. L. Hörmander, *Linear Partial Differential Operators* (Springer, Berlin, 1963)
37. L. Hörmander, *The Analysis of Linear Partial Differential Operators II* (Springer, Berlin, Heidelberg, 1983)
38. Y. Kannai, On the asymptotic behavior of resolvent kernels, spectral functions and eigenvalues of semi-elliptic systems. Ann. Sc. Norm. Super. Pisa Cl. Sci. **23**(4), 563–634 (1969)
39. G.F. Lawler, V. Limic, *Random Walk: A Modern Introduction*, vol. 123 (Cambridge University Press, Cambridge, 2010)
40. E.E. Levi, Sulle equazioni lineari totalmente ellittiche alle derivate parziali. Rend. Circ. Mat. Palermo **24**(1), 275–317 (1907)
41. T. Matsuzawa, On quasi-elliptic boundary problems. Trans. Am. Math. Soc. **133**(1), 241–241 (1968)
42. E.-M. Ouhabaz, *Analysis of Heat Equations on Domains (LMS-31)* (Princeton University Press, Princeton, 2009)
43. B. Pini, Proprietà locali delle soluzioni di una classe di equazioni ipoellittiche. Rend. Semin. Mat. Univ. Padova **32**, 221–238 (1962)
44. E. Randles, L. Saloff-Coste, Convolution powers of complex functions on $\mathbb{Z}^d$. Rev. Mat. Iberoam. (2016, to appear)
45. D.W. Robinson, *Elliptic Operators and Lie Groups* (Oxford University Press, Oxford, 1991)
46. W. Rudin, *Functional Analysis*, 2nd edn. (McGraw-Hill, New York, 1991)
47. F. Spitzer, *Principles of Random Walk*. Graduate Texts in Mathematics, vol. 34 (Springer, New York, 1964)
48. M.E. Taylor, *Pseudodifferential Operators* (Princeton University Press, Princeton, 1981)
49. H. Triebel, A priori estimates and boundary value problems for semi-elliptic differential equations: a model case. Commun. Partial Differ. Equ. **8**(15), 1621–1664 (1983)
50. M. Troisi, Problemi al contorno con condizioni omogenee per le equazioni quasi-ellittiche. Ann. Mat. Pura Appl. Ser. 4 **90**(1), 331–412 (1971)
51. A. Tsutsumi, On the asymptotic behavior of resolvent kernels and spectral functions for some class of hypoelliptic operators. J. Differ. Equ. **18**(2), 366–385 (1975)
52. L.R. Volevich, A class of hypoelliptic systems. Sov. Math. Dokl. **136**(6), 1275–1278 (1960)
53. L.R. Volevich, Local properties of solutions of quasi-elliptic systems. Mat. Sb. **59**(101), 3–52 (1962)

Strong Stability of Heat Kernels of Non-symmetric Stable-Like Operators

Zhen-Qing Chen and Xicheng Zhang

Abstract Let $d \geqslant 1$ and $\alpha \in (0, 2)$. Consider the following non-local and non-symmetric Lévy-type operator on $\mathbb{R}^d$:

$$\mathscr{L}_\alpha^\kappa f(x) := \text{p.v.} \int_{\mathbb{R}^d} (f(x+z) - f(x)) \frac{\kappa(x,z)}{|z|^{d+\alpha}} \mathrm{d}z,$$

where $0 < \kappa_0 \leqslant \kappa(x,z) \leqslant \kappa_1$, $\kappa(x,z) = \kappa(x,-z)$, and $|\kappa(x,z) - \kappa(y,z)| \leqslant \kappa_2 |x-y|^\beta$ for some $\beta \in (0,1)$. In Chen and Zhang (Probab Theory Relat Fields 165:267–312, 2016), we obtained two-sided estimates on the fundamental solution (also called heat kernel) $p_\alpha^\kappa(t,x,y)$ of $\mathscr{L}_\alpha^\kappa$. In this note, we establish pointwise estimate on $|p_\alpha^\kappa(t,x,y) - p_\alpha^{\tilde\kappa}(t,x,y)|$ in terms of $\|\kappa - \tilde\kappa\|_\infty$.

Keywords Heat kernel estimate • Levi's method • Non-symmetric stable-like operator • Strong stability

AMS 2010 Mathematics Subject Classification Primary 60J35, 47G20, 60J75; Secondary 47D07

1 Introduction

There are many interest recently in studying non-local operators or discontinuous Markov processes as many phenomena can be modeled by these objects. Quite a lot progress has been made during the last decade in developing DeGiorgi-Nash-

Z.-Q. Chen (✉)
Department of Mathematics, University of Washington, C-341 Padelford Hall, Seattle, WA 98195, USA
e-mail: zqchen@uw.edu

X. Zhang
School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, People's Republic of China
e-mail: xichengzhang@gmail.com

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_2

Moser-Aronson type theory for symmetric non-local operators; see, e.g., [1, 3–6] and the references therein. In particular, it is shown in Chen and Kumagai [5] that, for every $0 < \alpha < 2$ and for any symmetric measurable function $c(x, y)$ on $\mathbb{R}^d \times \mathbb{R}^d$ that is bounded between two positive constants c_1 and c_2, the symmetric non-local operator

$$\mathscr{L} f(x) = \lim_{\varepsilon \to 0} \int_{\{y \in \mathbb{R}^d : |y-x| > \varepsilon\}} (f(y) - f(x)) \frac{c(x, y)}{|x-y|^{d+\alpha}} \mathrm{d}y, \tag{1.1}$$

defined in the weak sense, admits a jointly Hölder continuous heat kernel $p(t, x, y)$ with respect to the Lebesgue measure on $\mathbb{R}^d$, which satisfies

$$C^{-1} \frac{t}{(t^{1/\alpha} + |x-y|)^{d+\alpha}} \leqslant p(t, x, y) \leqslant C \frac{t}{(t^{1/\alpha} + |x-y|)^{d+\alpha}} \tag{1.2}$$

for every $t > 0$ and $x, y \in \mathbb{R}^d$, where $C \geqslant 1$ is a constant that depends only on d, α, c_1 and c_2. When $c(x, y)$ is a positive constant, $\mathscr{L}$ above is a constant multiple of the fractional Laplacian $\Delta^{\alpha/2} := -(-\Delta)^{\alpha/2}$ on $\mathbb{R}^d$, which is the infinitesimal generator of a (rotationally) symmetric α-stable process on $\mathbb{R}^d$.

Recently, heat kernels and their sharp two-sided estimates for non-symmetric and non-local stable-like operators of the following form are studied in Chen and Zhang [7]:

$$\mathscr{L}_\alpha^\kappa f(x) := \text{p.v.} \int_{\mathbb{R}^d} (f(x+z) - f(x)) \frac{\kappa(x, z)}{|z|^{d+\alpha}} \mathrm{d}z, \tag{1.3}$$

where p.v. stands for the Cauchy principle value; that is

$$\mathscr{L}_\alpha^\kappa f(x) = \lim_{\varepsilon \to 0} \int_{\{z \in \mathbb{R}^d : |z| > \varepsilon\}} (f(x+z) - f(x)) \frac{\kappa(x, z)}{|z|^{d+\alpha}} \mathrm{d}z.$$

Here $d \geqslant 1$, $0 < \alpha < 2$, and $\kappa(x, z)$ is a measurable function on $\mathbb{R}^d \times \mathbb{R}^d$ satisfying

$$0 < \kappa_0 \leqslant \kappa(x, z) \leqslant \kappa_1, \qquad \kappa(x, z) = \kappa(x, -z), \tag{1.4}$$

and for some $\beta \in (0, 1)$

$$|\kappa(x, z) - \kappa(y, z)| \leqslant \kappa_2 |x-y|^\beta. \tag{1.5}$$

That $\kappa(x, z)$ is symmetric in z is a commonly assumed condition in the literature of non-local operators. Due to this symmetric, we can rewrite $\mathscr{L}_\alpha^\kappa$ as

$$\mathscr{L}_\alpha^\kappa f(x) = \int_{\mathbb{R}^d} (f(x+z) - f(x) - \mathbf{1}_{\{|z| \leqslant 1\}} z \cdot \nabla f(x)) \frac{\kappa(x, z)}{|z|^{d+\alpha}} \mathrm{d}z$$

for every bounded C^2-smooth function f on $\mathbb{R}^d$. We can also write $\mathscr{L}_\alpha^\kappa$ by

$$\mathscr{L}_\alpha^\kappa f(x) = \frac{1}{2}\int_{\mathbb{R}^d} (f(x+z) + f(x-z) - 2f(x))\frac{\kappa(x,z)}{|z|^{d+\alpha}}\mathrm{d}z.$$

Note that operators $\mathscr{L}_\alpha^\kappa$ defined by (1.3) are typically non-symmetric. They can be regarded as the non-local counterpart of elliptic operators of non-divergence form. Hölder continuity assumption (1.5) for non-symmetric operator $\mathscr{L}_\alpha^\kappa$ is quite natural. Unlike the symmetric case, even for elliptic differential operators, certain Dini-type continuity assumption on the coefficients is needed for the existence and for the two-sided estimates of the fundamental solution of non-divergence form operators.

Under conditions (1.4)–(1.5), it is established in [7] that $\mathscr{L}_\alpha^\kappa$ has a jointly continuous heat kernel $p_\alpha^\kappa(t,x,y)$, and there is a constant $c = c(d,\alpha,\beta,\kappa_0,\kappa_1,\kappa_2) \geqslant 1$ so that

$$c^{-1}\frac{t}{(t^{1/\alpha}+|x-y|)^{d+\alpha}} \leqslant p_\alpha^\kappa(t,x,y) \leqslant c\frac{t}{(t^{1/\alpha}+|x-y|)^{d+\alpha}} \tag{1.6}$$

for all $t \in (0,1]$ and $x,y \in \mathbb{R}^d$. Estimates in (1.2) and (1.6) can be viewed as the counterpart for non-local operators of Anroson estimates for elliptic differential operators.

In modeling, state-dependent parameters $c(x,y)$ of (1.1) and $\kappa(x,z)$ of (1.3) are approximations of real data. So a natural question is how reliable the conclusion is when using such an approximation. In this note, we study the pointwise strong stability on the heat kernel $p_\alpha^\kappa(t,x,y)$ in $\kappa(x,z)$. The following is the main result of this note. We use $:=$ as a way of definition. For $a,b \in \mathbb{R}$, $a \wedge b := \min\{a,b\}$ and $a \vee b := \max\{a,b\}$.

Theorem 1.1 *Suppose $\beta \in (0,\alpha/4]$, and κ and $\tilde\kappa$ are two functions satisfying (1.4) and (1.5). Then for every $\gamma \in (0,\beta)$ and $\eta \in (0,1)$, there exists a constant $C = C(d,\alpha,\beta,\kappa_0,\kappa_1,\kappa_2,\gamma,\eta) > 0$ so that for all $t \in (0,1]$ and $x,y \in \mathbb{R}^d$,*

$$|p_\alpha^\kappa(t,x,y) - p_\alpha^{\tilde\kappa}(t,x,y)| \leqslant C\|\kappa-\tilde\kappa\|_\infty^{1-\eta}\left(1+t^{-\gamma/\alpha}(|x-y|^\gamma \wedge 1)\right) t(t^{1/\alpha}+|x-y|)^{-d-\alpha}. \tag{1.7}$$

Here $\|\kappa-\tilde\kappa\|_\infty := \sup_{x,z\in\mathbb{R}^d}|\kappa(x,z)-\tilde\kappa(x,z)|$.

Observe that by (1.6), the $t(t^{1/\alpha}+|x-y|)^{-d-\alpha}$ term in (1.7) is comparable to $p_\alpha^\kappa(t,x,y)$ and to $p_\alpha^{\tilde\kappa}(t,x,y)$. So the error bound (1.7) is also a relative error bound, which is good even in the region when $|x-y|$ is large. When both $\kappa(x,z)$ and $\tilde\kappa(x,z)$ are functions of z only, estimate (1.7) has been established in [7, Theorem 2.6] with $\eta = 0$.

Let $\{P_t^\kappa; t \geqslant 0\}$ and $\{P_t^{\tilde\kappa}; t \geqslant 0\}$ be the semigroups generated by $\mathscr{L}_\alpha^\kappa$ and $\mathscr{L}_\alpha^{\tilde\kappa}$, respectively. For $p \geqslant 1$, denote by $\|P_t^\kappa - P_t^{\tilde\kappa}\|_{p,p}$ the operator norm of $P_t^\kappa - P_t^{\tilde\kappa}$ in Banach space $L^p(\mathbb{R}^d;\mathrm{d}x)$.

Corollary 1.2 *Suppose $\beta \in (0,\alpha/4]$, and κ and $\tilde\kappa$ are two functions satisfying (1.4) and (1.5). Then for every $\gamma \in (0,\beta)$ and $\eta \in (0,1)$, there exists a constant*

$C = C(d, \alpha, \beta, \kappa_0, \kappa_1, \kappa_2, \gamma, \eta) > 0$ *so that for every* $p \geqslant 1$ *and* $t \in (0, 1]$,

$$\|P_t^{\kappa} - P_t^{\tilde{\kappa}}\|_{p,p} \leqslant C t^{-\gamma/\alpha} \|\kappa - \tilde{\kappa}\|_{\infty}^{1-\eta}. \tag{1.8}$$

For uniformly elliptic divergence form operators $\mathscr{L}$ and $\tilde{\mathscr{L}}$ on $\mathbb{R}^d$, pointwise estimate on $|p(t, x, y) - \tilde{p}(t, x, y)|$ and the L^p-operator norm estimates on $P_t - \tilde{P}_t$ are obtained in Chen et al. [8] in terms of the local L^2-distance between the diffusion matrix of $\mathscr{L}$ and $\tilde{\mathscr{L}}$. Recently, Bass and Ren [2] extended the above result to symmetric α-stable-like non-local operators of (1.1), with error bound expressed in terms of the L^q-norm on the function $c(x) := \sup_{y \in \mathbb{R}^d} |c(x, y) - \tilde{c}(x, y)|$.

2 Proofs

We recall from [7] that the heat kernel $p_\alpha^\kappa(t, x, y)$ is constructed as follows:

$$p_\alpha^\kappa(t, x, y) = p_\alpha^{\kappa(y)}(t, x - y) + \int_0^t \int_{\mathbb{R}^d} p_\alpha^{\kappa(z)}(t - s, x - z) q^\kappa(s, z, y) \mathrm{d}z \mathrm{d}s, \tag{2.1}$$

where $q^\kappa(t, x, y)$ solves the following integral equation:

$$q^\kappa(t, x, y) = q_0^\kappa(t, x, y) + \int_0^t \int_{\mathbb{R}^d} q_0^\kappa(t - s, x, z) q^\kappa(s, z, y) \mathrm{d}z \mathrm{d}s \tag{2.2}$$

with

$$\begin{aligned} q_0^\kappa(t, x, y) &:= (\mathscr{L}_\alpha^{\kappa(x)} - \mathscr{L}_\alpha^{\kappa(y)}) p_\alpha^{\kappa(y)}(t, x - y) \\ &= \int_{\mathbb{R}^d} \delta_{p_\alpha^{\kappa(y)}}(t, x - y; z)(\kappa(x, z) - \kappa(y, z))|z|^{-d-\alpha} \mathrm{d}z. \end{aligned} \tag{2.3}$$

Here for a function f defined on $\mathbb{R}_+ \times \mathbb{R}^d$,

$$\delta_f(t, x; z) := f(t, x + z) + f(t, x - z) - 2f(t, x).$$

For $n \in \mathbb{N}$, define $q_n^\kappa(t, x, y)$ recursively by

$$q_n^\kappa(t, x, y) := \int_0^t \int_{\mathbb{R}^d} q_0^\kappa(t - s, x, z) q_{n-1}^\kappa(s, z, y) \mathrm{d}z \mathrm{d}s. \tag{2.4}$$

It is established in [7, Theorem 3.1] that

$$q^\kappa(t, x, y) := \sum_{n=0}^\infty q_n^\kappa(t, x, y),$$

where the series converges absolutely and locally uniformly on $(0, 1] \times \mathbb{R}^d \times \mathbb{R}^d$. Moreover, by Chen and Zhang [7, (3.8) and (3.10)], there is a constant $C = C(d, \alpha, \beta, \kappa_0, \kappa_1, \kappa_2) > 0$ so that for every $t \in (0, 1]$ and $x, y \in \mathbb{R}^d$,

$$|q_0^\kappa(t, x, y)| \leqslant C\varrho_0^\beta(t, x - y), \tag{2.5}$$

and for $n \geqslant 1$,

$$|q_n^\kappa(t, x, y)| \leqslant \frac{(C\,\Gamma(\frac{\beta}{\alpha}))^{n+1}}{\Gamma(\frac{(n+1)\beta}{\alpha})} \left(\varrho^0_{(n+1)\beta} + \varrho^\beta_{n\beta}\right)(t, x - y). \tag{2.6}$$

Here for $\gamma, \beta \in \mathbb{R}$,

$$\varrho_\gamma^\beta(t, x) := t^{\frac{\gamma}{\alpha}}(|x|^\beta \wedge 1)(t^{1/\alpha} + |x|)^{-d-\alpha}. \tag{2.7}$$

Obviously, for $\gamma, \beta \geqslant 0$, we have

$$\varrho_\gamma^\beta(t, x) \leqslant \varrho^0_{\beta+\gamma}(t, x) + \varrho_0^{\beta+\gamma}(t, x). \tag{2.8}$$

To prove Theorem 1.1, we first establish a continuity result for $q^\kappa(t, x, y)$ in κ.

Theorem 2.1 *Suppose $\beta \in (0, \alpha/4]$, and κ and $\tilde\kappa$ are two functions satisfying (1.4) and (1.5). For every $\gamma \in (0, \beta)$ and $\eta \in (0, 1)$, there is a constant $C = C(d, \alpha, \kappa_0, \kappa_1, \kappa_2, \gamma, \eta) > 0$ so that*

$$|q^\kappa(t, x, y) - q^{\tilde\kappa}(t, x, y)| \leqslant C\|\kappa - \tilde\kappa\|_\infty^{1-\eta}\left(\varrho^{\beta+\gamma}_{-\gamma} + \varrho_0^{\eta\beta} + \varrho^0_\beta\right)(t, x - y). \tag{2.9}$$

Proof Since $q^\kappa(t, x, y) = \sum_{n=0}^\infty q_n^\kappa(t, x, y)$, we estimate $|q_n^\kappa(t, x, y) - q_n^{\tilde\kappa}(t, x, y)|$ for each $n \geqslant 0$. First, observe that by (1.4) and (1.5), for $\eta \in (0, 1)$,

$$|(\kappa - \tilde\kappa)(x, z) - (\kappa - \tilde\kappa)(y, z)| \leqslant C(|x - y|^\beta \wedge 1)^\eta\, \|\kappa - \tilde\kappa\|_\infty^{1-\eta} \tag{2.10}$$

with $C = C(\kappa_1, \kappa_2) > 0$. When $n = 0$, by (2.3),

$$\begin{aligned}
&|q_0^\kappa(t, x, y) - q_0^{\tilde\kappa}(t, x, y)| \\
&\leqslant \int_{\mathbb{R}^d} |\delta_{p_\alpha^{\kappa(y)}}(t, x - y; z) - \delta_{p_\alpha^{\tilde\kappa(y)}}(t, x - y; z)|\, |\kappa(x, z) - \kappa(y, z)|\, |z|^{-d-\alpha}\mathrm{d}z \\
&\quad + \int_{\mathbb{R}^d} |\delta_{p_\alpha^{\tilde\kappa(y)}}(t, x - y; z)|\, |(\kappa - \tilde\kappa)(x, z) - (\kappa - \tilde\kappa)(y, z)|\, |z|^{-d-\alpha}\mathrm{d}z \\
&\preceq (|x - y|^\beta \wedge 1) \int_{\mathbb{R}^d} |\delta_{p_\alpha^{\kappa(y)}}(t, x - y; z) - \delta_{p_\alpha^{\tilde\kappa(y)}}(t, x - y; z)|\, |z|^{-d-\alpha}\mathrm{d}z
\end{aligned}$$

$$
\begin{aligned}
&\quad + \|\kappa-\tilde\kappa\|_\infty^{1-\eta}(|x-y|^{\eta\beta}\wedge 1)\int_{\mathbb{R}^d}|\delta_{p_\alpha^{\tilde\kappa(y)}}(t,x-y;z)|\,|z|^{-d-\alpha}\mathrm{d}z\\
&\preceq \|\kappa-\tilde\kappa\|_\infty(|x-y|^{\beta}\wedge 1)\left(\varrho_0^0+\varrho_{-\gamma}^{\gamma}\right)(t,x-y)\\
&\quad + \|\kappa-\tilde\kappa\|_\infty^{1-\eta}(|x-y|^{\eta\beta}\wedge 1)\varrho_0^0(t,x-y)\\
&\preceq \|\kappa-\tilde\kappa\|_\infty^{1-\eta}\left(\varrho_0^{\beta}+\varrho_{-\gamma}^{\beta+\gamma}+\varrho_0^{\eta\beta}\right)(t,x-y)\\
&\preceq \|\kappa-\tilde\kappa\|_\infty^{1-\eta}\left(\varrho_{-\gamma}^{\beta+\gamma}+\varrho_0^{\eta\beta}\right)(t,x-y), \qquad (2.11)
\end{aligned}
$$

where the third inequality is due to [7, (2.32) and (2.28)], respectively. Here and below, "$f\preceq g$" means that $f\leqslant Cg$ for some positive constant C.

When $n=1$, by (2.4)–(2.5) and (2.11),

$$
\begin{aligned}
&|q_1^{\kappa}(t,x,y)-q_1^{\tilde\kappa}(t,x,y)|\\
&\leqslant \int_0^t\!\!\int_{\mathbb{R}^d}\left|\left(q_0^{\kappa}(t-s,x,z)-q_0^{\tilde\kappa}(t-s,x,z)\right)q_0^{\kappa}(s,z,y)\right|\mathrm{d}z\mathrm{d}s\\
&\quad+\int_0^t\!\!\int_{\mathbb{R}^d}\left|q_0^{\tilde\kappa}(t-s,x,z)\left(q_0^{\kappa}(s,z,y)-q_0^{\tilde\kappa}(s,z,y)\right)\right|\mathrm{d}z\mathrm{d}s\\
&\preceq \|\kappa-\tilde\kappa\|_\infty^{1-\eta}\int_0^t\!\!\int_{\mathbb{R}^d}\left(\varrho_{-\gamma}^{\beta+\gamma}+\varrho_0^{\eta\beta}\right)(t-s,x-z)\varrho_0^{\beta}(s,z-y)\mathrm{d}z\mathrm{d}s\\
&\preceq \|\kappa-\tilde\kappa\|_\infty^{1-\eta}\Bigg(\mathcal{B}(\tfrac{\beta}{\alpha},\tfrac{\beta}{\alpha})\left(\varrho_{2\beta}^0+\varrho_{\beta-\gamma}^{\beta+\gamma}+\varrho_{\beta}^{\beta}\right)(t,x-y)\\
&\quad+\mathcal{B}(\tfrac{\eta\beta}{\alpha},\tfrac{\beta}{\alpha})\left(\varrho_{(1+\eta)\beta}^0+\varrho_{\beta}^{\eta\beta}+\varrho_{\eta\beta}^{\beta}\right)(t,x-y)\Bigg)\\
&\preceq \|\kappa-\tilde\kappa\|_\infty^{1-\eta}\mathcal{B}(\tfrac{\eta\beta}{\alpha},\tfrac{\beta}{\alpha})\left(\varrho_{(1+\eta)\beta}^0+\varrho_0^{(1+\eta)\beta}\right)(t,x-y). \qquad (2.12)
\end{aligned}
$$

Here in the second to the last inequality we used [7, (2.4)], while in the last inequality we used (2.8) and the monotonicity of ϱ_γ^β with respect to β and γ. Thus we have shown that there are constants $C_0=C_0(d,\alpha,\eta,\kappa_0,\kappa_1,\kappa_2,\beta,\gamma)\geqslant 1$ (we can eliminate the dependence on γ and η by taking, for example, $\gamma=\beta/2$ and $\eta=1/2$) and $C_1=C_1(d,\alpha)\geqslant 1$ so that for every $t\in(0,1]$ and $x,y\in\mathbb{R}^d$,

$$
|q_0^{\kappa}(t,x,y)-q_0^{\tilde\kappa}(t,x,y)|\leqslant C_0\|\kappa-\tilde\kappa\|_\infty^{1-\eta}\left(\varrho_{-\gamma}^{\beta+\gamma}+\varrho_0^{\eta\beta}\right)(t,x-y) \qquad (2.13)
$$

and

$$
|q_1^{\kappa}(t,x,y)-q_1^{\tilde\kappa}(t,x,y)|\leqslant C_0C_1\|\kappa-\tilde\kappa\|_\infty^{1-\eta}\frac{\Gamma(\frac{\eta\beta}{\alpha})\Gamma(\frac{\beta}{\alpha})}{\Gamma(\frac{(1+\eta)\beta}{\alpha})}\left(\varrho_{(1+\eta)\beta}^0+\varrho_0^{(1+\eta)\beta}\right)(t,x-y). \qquad (2.14)
$$

Without loss of generality, we may assume that C_0 is larger than the constant C in (2.5)–(2.6), and C_1 is larger than the comparison constant in [7, (2.4)]. Suppose for $n \geqslant 1$,

$$\begin{aligned}&|q_n^{\kappa}(t,x,y) - q_n^{\tilde{\kappa}}(t,x,y)|\\&\quad\leqslant \|\kappa - \tilde{\kappa}\|_{\infty}^{1-\eta}(20C_0C_1)^{n+1}\frac{\Gamma(\frac{\eta\beta}{\alpha})\Gamma(\frac{\beta}{\alpha})^n}{\Gamma(\frac{(n+\eta)\beta}{\alpha})}\left(\varrho^0_{(n+\eta)\beta} + \varrho_0^{(n+\eta)\beta}\right)(t,x-y).\end{aligned} \tag{2.15}$$

Then by (2.4), (2.5)–(2.6), [7, (2.4)], (2.13)–(2.14) and (2.8), we have

$$\begin{aligned}&|q_{n+1}^{\kappa}(t,x,y) - q_{n+1}^{\tilde{\kappa}}(t,x,y)|\\
&\leqslant \int_0^t\!\!\int_{\mathbb{R}^d}\left|q_0^{\kappa}(t-s,x,z)\left(q_n^{\kappa}(s,z,y) - q_n^{\tilde{\kappa}}(s,z,y)\right)\right|\mathrm{d}z\mathrm{d}s\\
&\quad + \int_0^t\!\!\int_{\mathbb{R}^d}\left|\left(q_0^{\kappa}(t-s,x,z) - q_0^{\tilde{\kappa}}(t-s,x,z)\right)q_n^{\tilde{\kappa}}(s,z,y)\right|\mathrm{d}z\mathrm{d}s\\
&\leqslant \|\kappa-\tilde{\kappa}\|_{\infty}^{1-\eta}\Bigg[C_0(20C_0C_1)^{n+1}\frac{\Gamma(\frac{\eta\beta}{\alpha})\Gamma(\frac{\beta}{\alpha})^n}{\Gamma(\frac{(n+\eta)\beta}{\alpha})}\varrho_0^{\beta} * \left(\varrho^0_{(n+\eta)\beta} + \varrho_0^{(n+\eta)\beta}\right)(t,x-y)\\
&\qquad + \frac{C_0(C_0\Gamma(\frac{\beta}{\alpha}))^{n+1}}{\Gamma(\frac{(n+1)\beta}{\alpha})}\left(\varrho_{-\gamma}^{\beta+\gamma} + \varrho_0^{\eta\beta}\right) * \left(\varrho^0_{(n+1)\beta} + \varrho_0^{(n+1)\beta}\right)(t,x-y)\Bigg]\\
&\leqslant \|\kappa-\tilde{\kappa}\|_{\infty}^{1-\eta}\Bigg[C_0(20C_0C_1)^{n+1}\frac{\Gamma(\frac{\eta\beta}{\alpha})\Gamma(\frac{\beta}{\alpha})^n}{\Gamma(\frac{(n+\eta)\beta}{\alpha})}10C_1\mathcal{B}\Big(\tfrac{\beta}{\alpha},\tfrac{(n+\eta)\beta}{\alpha}\Big)\\
&\qquad + \frac{C_0(C_0\Gamma(\frac{\beta}{\alpha}))^{n+1}}{\Gamma(\frac{(n+1)\beta}{\alpha})}10C_1\mathcal{B}\Big(\tfrac{\eta\beta}{\alpha},\tfrac{(n+1)\beta}{\alpha}\Big)\Bigg]\left(\varrho^0_{(n+1+\eta)\beta} + \varrho_0^{(n+1+\eta)\beta}\right)(t,x-y)\\
&\leqslant \|\kappa-\tilde{\kappa}\|_{\infty}^{1-\eta}(20C_0C_1)^{n+2}\frac{\Gamma(\frac{\eta\beta}{\alpha})\Gamma(\frac{\beta}{\alpha})^{n+1}}{\Gamma(\frac{(n+1+\eta)\beta}{\alpha})}\left(\varrho^0_{(n+1+\eta)\beta} + \varrho_0^{(n+1+\eta)\beta}\right)(t,x-y),\end{aligned} \tag{2.16}$$

where we have used the notation in the second inequality: for $f, g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$,

$$(f * g)(t,x) := \int_0^t\!\!\int_{\mathbb{R}^d} f(t-s,x-z)g(s,z)\mathrm{d}z.$$

This proves by induction that (2.15) holds for every integer $n \geqslant 1$. Consequently, for every $t \in (0, 1]$ and $x, y \in \mathbb{R}^d$,

$$
\begin{aligned}
&|q^\kappa(t,x,y) - q^{\tilde\kappa}(t,x,y)| \\
&\leqslant |q_0^\kappa(t,x,y) - q_0^{\tilde\kappa}(t,x,y)| + \sum_{n=1}^\infty |q_n^\kappa(t,x,y) - q_n^{\tilde\kappa}(t,x,y)| \\
&\leqslant C_0\|\kappa - \tilde\kappa\|_\infty^{1-\eta}\left(\varrho_{-\gamma}^{\beta+\gamma} + \varrho_0^{\eta\beta}\right)(t, x-y) \\
&\quad + C_2\|\kappa - \tilde\kappa\|_\infty^{1-\eta}\left(\varrho^0_{(1+\eta)\beta} + \varrho_0^{(1+\eta)\beta}\right)(t, x-y) \\
&\leqslant C_3\|\kappa - \tilde\kappa\|_\infty^{1-\eta}\left(\varrho_{-\gamma}^{\beta+\gamma} + \varrho_0^{\eta\beta} + \varrho_\beta^0\right)(t, x-y).
\end{aligned}
$$

This proves the theorem. □

We now proceed to prove the main result of this note.

Proof of Theorem 1.1 For $t \in (0, 1]$ and $x, y \in \mathbb{R}^d$, by (2.1), [7, (2.30) and (3.6)] and Theorem 2.1, we have

$$
\begin{aligned}
|p_\alpha^\kappa(t,x,y) - p_\alpha^{\tilde\kappa}(t,x,y)| &\leqslant |p_\alpha^{\kappa(y)}(t,x,y) - p_\alpha^{\tilde\kappa(y)}(t,x,y)| \\
&\quad + \int_0^t\!\!\int_{\mathbb{R}^d} |p_\alpha^{\kappa(z)}(t-s,x-z) - p_\alpha^{\tilde\kappa(z)}(t-s,x-z)|\,|q^\kappa(s,z,y)|\mathrm{d}z\mathrm{d}s \\
&\quad + \int_0^t\!\!\int_{\mathbb{R}^d} p_\alpha^{\tilde\kappa(z)}(t-s,x-z)\,|q^\kappa(s,z,y) - q^{\tilde\kappa}(s,z,y)|\mathrm{d}z\mathrm{d}s \\
&\preceq \|\kappa - \tilde\kappa\|_\infty \left(\varrho_\alpha^0 + \varrho_{\alpha-\gamma}^\gamma + (\varrho_\alpha^0 + \varrho_{\alpha-\gamma}^\gamma) * (\varrho_0^\beta + \varrho_\beta^0)\right)(t, x-y) \\
&\quad + \|\kappa - \tilde\kappa\|_\infty^{1-\eta}\left(\varrho_{-\gamma}^{\beta+\gamma} + \varrho_0^{\eta\beta} + \varrho_\beta^0\right) * \varrho_\alpha^0(t, x-y) \\
&\preceq \|\kappa - \tilde\kappa\|_\infty \left(\varrho_\alpha^0 + \varrho_{\alpha-\gamma}^\gamma + \varrho_{\alpha+\beta}^0 + \varrho_\alpha^\beta + \varrho_{\alpha+\beta-\gamma}^\gamma\right)(t, x-y) \\
&\quad + \|\kappa - \tilde\kappa\|_\infty^{1-\eta}\left(\varrho_{\alpha+\beta}^0 + \varrho_{\alpha-\gamma}^{\beta+\gamma} + \varrho_{\alpha+\eta\beta}^0 + \varrho_\alpha^{\eta\beta}\right)(t, x-y) \\
&\preceq \|\kappa - \tilde\kappa\|_\infty^{1-\eta}(\varrho_\alpha^0 + \varrho_{\alpha-\gamma}^\gamma)(t, x-y).
\end{aligned}
$$

This proves (1.7). □

Proof of Corollary 1.2 Let $\{P_t; t \geqslant 0\}$ be the transition semigroup of the rotationally symmetric α-stable process on $\mathbb{R}^d$. For $p \geqslant 1$ and any $f \in L^p(\mathbb{R}^d; \mathrm{d}x)$, we have by (1.6) and (1.7) that

$$
|(P_t^\kappa - P_t^{\tilde\kappa})f(x)| \leqslant \int_{\mathbb{R}^d} |p_\alpha^\kappa(t,x,y) - p_\alpha^{\tilde\kappa}(t,x,y)|\,|f(y)|dy \leqslant Ct^{-\gamma/\alpha}\|\kappa - \tilde\kappa\|_\infty^{1-\eta} P_t|f|(x).
$$

Hence

$$\|(P_t^\kappa - P_t^{\tilde\kappa})f\|_p \leqslant Ct^{-\gamma/\alpha}\|\kappa - \tilde\kappa\|_\infty^{1-\eta}\|P_t|f|\|_p \leqslant Ct^{-\gamma/\alpha}\|\kappa - \tilde\kappa\|_\infty^{1-\eta}\|f\|_p.$$

This establishes the corollary. □

Acknowledgements The research of ZC is partially supported by NSF Grants DMS-1206276. The research of XZ is partially supported by NNSFC grant of China (Nos. 11271294, 11325105).

References

1. R.F. Bass, D.A. Levin, Transition probabilities for symmetric jump processes. Trans. Am. Math. Soc. **354**, 2933–2953 (2002)
2. R.F. Bass, H. Ren, Meyers inequality and strong stability for stable-like operators. J. Funct. Anal. **265**, 28–48 (2013)
3. L. Caffarelli, L. Silvestre, The Evans-Krylov theorem for nonlocal fully nonlinear equations. Ann. Math. **174**, 1163–1187 (2011)
4. Z.-Q. Chen, Symmetric jump processes and their heat kernel estimates. Sci. China Ser. A Math. **52**, 1423–1445 (2009)
5. Z.-Q. Chen, T. Kumagai, Heat kernel estimates for stable-like processes on *d*-sets. Stoch. Process. Appl. **108**, 27–62 (2003)
6. Z.-Q. Chen, T. Kumagai, Heat kernel estimates for jump processes of mixed types on metric measure spaces. Probab. Theory Relat. Fields **140**, 277–317 (2008)
7. Z.-Q. Chen, X. Zhang, Heat kernels and analyticity of non-symmetric jump diffusion semigroups. Probab. Theory Relat. Fields **165**, 267–312 (2016)
8. Z.-Q. Chen, Y. Hu, Z. Qian, W. Zheng, Stability and approximations of symmetric diffusion semigroups and kernels. J. Funct. Anal. **152**, 255–280 (1998)

Multiplicative Functional for the Heat Equation on Manifolds with Boundary

Cheng Ouyang

Abstract The multiplicative functional for the heat equation on k-forms with absolute boundary condition is constructed and a probabilistic representation of the solution is obtained. As an application, we prove a heat kernel domination that was previously discussed by Donnelly and Li, and Shigekawa.

Keywords Absolute boundary condition • Gradient inequality • Heat kernel domination • Hodge-de Rham Laplacian • Riemannian manifold with boundary

1 Introduction

Throughout this paper, we assume that M is an n-dimensional compact Riemannian manifold with boundary ∂M. Denote by $\Box$ the Hodge-de Rham Laplacian. Let θ_0 be a differential k-form on M and consider the following initial boundary valued problem on M:

$$\begin{cases} \frac{\partial\theta}{\partial t} = \frac{1}{2}\Box\theta, \\ \theta(\cdot, 0) = \theta_0, \\ \theta_{norm} = 0, (d\theta)_{norm} = 0. \end{cases} \tag{1.1}$$

The well known Weitzenböck formula shows that the difference between the Hodge-de Rham Laplacian and the covariant Laplacian for the differential forms on a Riemannian manifold M is a linear transformation at each $x \in M$. So the heat equation for differential forms is naturally associated with a matrix-valued Feynman-Kac multiplicative functional determined by the curvature tensor. The boundary condition

$$\theta_{norm} = 0, \text{ and } (d\theta)_{norm} = 0,$$

C. Ouyang (✉)
Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, MSCS (M/C 249), 851 South Morgan Street, Chicago, IL 60607, USA
e-mail: couyang@math.uic.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_3

is called the absolute boundary condition. The significance of the absolute boundary condition stems from the well-know work [7]. Since it is Dirichlet in the normal direction and Neumann in the tangential directions, the associated multiplicative functional is discontinuous and therefore difficult to handle. Ikeda and Watanabe [5, 6] have dealt with this situation by using an excursion theory. Later, Hsu [3] constructed the discontinuous multiplicative functional M_t for 1-forms by an approximating argument inspired by Ariault [1]. Following a similar idea, the same multiplicative functional M_t has been constructed for non-compact manifolds with boundary by Wang [9]. The solution to Eq. (1.1) for 1-forms thus can be represented in terms of M_t as

$$\theta(x,t) = u_0 \mathbb{E}_x \{ M_t u_t^{-1} \theta_0(x_t) \}, \tag{1.2}$$

where $\{x_t\}$ is a reflecting Brownian motion on M, and $\{u_t\}$ its horizontal lift process to the orthonormal fame bundle $\mathscr{O}(M)$ starting from a frame $u_0 : \mathbb{R}^n \to T_x M$, which we will use to identify $T_x M$ with $\mathbb{R}^n$. As a direct consequence, a gradient estimate

$$|\nabla P_t f(x)| \le \mathbb{E}_x \left\{ |\nabla f(x_t)| \exp\left[-\frac{1}{2} \int_0^t \kappa(x_s) ds - \int_0^t h(x_s) dl_s \right] \right\}$$

was obtained. Here l is the boundary local time for $\{x_t\}$, $\kappa(x)$ the lower bound of the Ricci curvature at $x \in M$, and $h(x)$ the lower bound of the second fundamental form at $x \in \partial M$.

The present paper extends Hsu's work [3] to multiplicative functional on the full exterior algebra $\wedge^* M$. We lift the absolute boundary condition onto the frame bundle $\mathscr{O}(M)$ and clarify the action of second fundamental form on k-forms in the absolute boundary condition. Then the multiplicative functional M_t for the heat equation (1.1) is constructed. With this M_t, the representation (1.2) still holds for k-forms, and we have the following estimate

$$|M_t|_{2,2} \le \exp\left[\frac{1}{2} \int_0^t \lambda(x_s) ds - \int_0^t \sigma_k(x_s) dl_s \right]. \tag{1.3}$$

Here

$$\lambda(x) = \sup_{\theta \in \wedge_x^k M, \langle \theta, \theta \rangle = 1} \langle D^* R(x) \theta, \theta \rangle, \tag{1.4}$$

with $D^* R \theta$ the curvature tensor acting on θ as the Lie algebra action, and $\sigma_k(x)$, $k = 1, 2, \ldots, n$ being combinations of eigenvalues of the second fundamental form at $x \in \partial M$, which we will specify later. It follows immediately with (1.2) and (1.3) our generalized gradient inequality

$$|dP_t \theta(x)| \le \mathbb{E}_x \left\{ \left| d\theta \right| \exp\left[\frac{1}{2} \int_0^t \lambda(x_s) ds - \int_0^t \sigma_{k+1}(x_s) dl_s \right] \right\}.$$

Let $\bar{\lambda} = \sup_{x\in\partial M} \lambda(x)$, we also prove the heat kernel domination

$$|p_M^k(t,x,y)|_{2,2} \le e^{\frac{1}{2}\bar{\lambda}t} p_M(t,x,y)\mathbb{E}_x\{e^{-\int_0^t \sigma_k(x_s)dl_s} | x_t = y\}.$$

Here $p^k(t,x,y)$ is the heat kernel on k-forms with absolute boundary condition and $p_M(t,x,y)$ is the heat kernel on functions with Neumann boundary condition. Note that when $\sigma_k \ge 0$, the above inequality reduces to

$$|p_M^k(t,x,y)|_{2,2} \le e^{\frac{1}{2}\bar{\lambda}t} p_M(t,x,y). \tag{1.5}$$

This special case was proved by Donnelly and Li [2]. We remark that the heat kernel domination was also discussed in Shigekawa [8] by an approach using theory of Dirichlet form. Inequality (1.5) was obtained as an example for 1-forms in [8].

Finally, we would like to remark that although the present work focuses on compact manifolds, we believe similar results can be obtained for non-compact manifolds (under suitable conditions on curvatures and second fundamental forms) by using the treatment discussed in Wang [9, Chap. 3].

The rest of the paper is organized as follows. In Sect. 2, we briefly recall the Weitzenböck formula and corresponding actions on differential forms. In Sect. 3, we give an explicit expression for the absolute boundary condition. The reflecting Brownian motion with Neumann boundary condition is briefly introduced in Sect. 4. Then, we focus on the construction of the multiplicative functional on k-forms for heat equation (1.1) in Sect. 5. Finally we provide some applications in Sect. 6.

2 Weitzenböck Formula on Orthonormal Frame Bundle

For our purpose, it is more convenient to lift equation (1.1) onto the orthonormal frame bundle $\mathscr{O}(M)$. In this section, we give a brief review of Weitzenböck formula and it's lift on the frame bundle $\mathscr{O}(M)$. More detailed discussion can be found in [4].

Let $\triangle = \text{trace}\nabla^2$ be the Laplace-Beltrami operator and $\square = -(dd^* + d^*d)$ the Hodge-de Rham Laplacian. They are related by the Weitzenböck formula

$$\square = \triangle + D^*R.$$

We first explain the action of the curvature tensor R on differential forms in the above formula. Suppose that $T : T_xM \to T_xM$ is a linear transformation and $T^* : \wedge_x^1 M \to \wedge_x^1 M$ its dual. The linear map T^* on $\wedge_x^1 M$ can be extended to the full exterior algebra $\wedge_x^* M = \sum_{k=0}^n \bigoplus \wedge_x^k M$ as a Lie algebra action (derivation) D^*T by

$$D^*T(\theta_1 \wedge \theta_2) = D^*T\theta_1 \wedge \theta_2 + \theta_1 \wedge D^*T\theta_2.$$

Let $\mathrm{End}(T_xM)$ be the space of linear maps from T_xM to itself. We define a bilinear map

$$D^* : \mathrm{End}(T_xM) \otimes \mathrm{End}(T_xM) \to \mathrm{End}(\wedge_x^* M)$$

by

$$D^*(T_1 \otimes T_2) = D^* T_1 \circ D^* T_2.$$

From elementary algebra we know that $\mathrm{End}(T_xM)=(T_xM)^* \otimes T_xM$. By the definition of the curvature tensor R and using the isometry $(T_xM)^* \to T_xM$ induced by the inner product, we can identify R as an element in $\mathrm{End}(T_xM) \otimes \mathrm{End}(T_xM)$. Thus by the above definition, we obtain a linear map

$$D^* R : \wedge_x^* M \to \wedge_x^* M,$$

which, by the Weiztenböck formula, is the difference between the covariance Laplacian and the Hodge-de Rham Laplacian.

A frame $u \in \mathscr{O}(M)$ is an isometry $u : \mathbb{R}^n \to T_xM$, where $x = \pi u$ and $\pi : \mathscr{O}(M) \to M$ is the canonical projection. A curve $\{u_t\}$ in $\mathscr{O}(M)$ is horizontal if, for any $e \in \mathbb{R}^n$, the vector field $\{u_t e\}$ is parallel along the curve $\{\pi u_t\}$. A vector on $\mathscr{O}(M)$ is horizontal if it is the tangent vector of a horizontal curve. For each $v \in T_xM$ and a frame $u \in \mathscr{O}(M)$ such that $\pi u = x$, there is a unique horizontal vector V, called the horizontal lift of v, such that $\pi_* V = v$. For each $i = 1, \ldots, n$, let $H_i(u)$ be the horizontal lift of $ue_i \in T_xM$. Each H_i is a horizontal vector field on $\mathscr{O}(M)$, and $H_1, \ldots, H_n$ are called the fundamental horizontal vector fields on $\mathscr{O}(M)$.

On the orthonormal frame bundle $\mathscr{O}(M)$, a k-form θ is lifted to its scalarization $\tilde{\theta}$ defined by

$$\tilde{\theta}(u) = u^{-1} \theta(\pi u).$$

Here a frame $u : \mathbb{R}^n \to T_xM$ is assumed to be extended canonically to an isometry $u : \wedge^* \mathbb{R}^n \to \wedge_x^* M$. By definition, $\tilde{\theta}$ is a function on $\mathscr{O}(M)$ taking values in the vector space $\wedge^k \mathbb{R}^n$ and is $O(n)$-invariant in the sense that $\tilde{\theta}(gu) = g\tilde{\theta}(u)$ for $g \in O(n)$. We remark that through the isometry $u : \wedge^* \mathbb{R}^n \to \wedge_x^* M$, a linear transformation $T(x) : \wedge_x^* M \to \wedge_x^* M$ can also be lifted onto $\mathscr{O}(M)$ as a linear map

$$\tilde{T}(u) = u^{-1} H(\pi u) u : \wedge^* \mathbb{R}^n \to \wedge^* \mathbb{R}^n.$$

To simplify the notation, whenever feasible, we still use T for the more precise $\tilde{T}$ throughout our discussion.

Bochner's horizontal Laplacian on the frame bundle $\mathscr{O}(M)$ is defined to be $\triangle_{\mathscr{O}(M)} = \sum_{i=1}^{n} H_i^2$. It is the lift of the Laplace-Beltrami operator $\triangle$ in the sense that

$$\triangle_{\mathscr{O}(M)}\tilde{\theta}(u) = \widetilde{\triangle\theta(x)}, \quad \pi u = x.$$

To write the Weitzenbök formula on the frame bundle, we lift $D^*R : \wedge_x^* M \to \wedge_x^* M$ to the frame bundle $\mathscr{O}(M)$, which will be denoted by $D^*\Omega$, and let

$$\Box_{\mathscr{O}(M)} = \triangle_{\mathscr{O}(M)} + D^*\Omega. \tag{2.1}$$

Then $\Box_{\mathscr{O}(M)}$ is the lift of the Hodged-de Rham Laplacian in the sense that $\Box_{\mathscr{O}(M)}\tilde{\theta}(u) = \widetilde{\Box\theta(x)}$, where $\pi u = x$. The identity (2.1) is the lifted Weiztenböck formula on the orthonormal frame bundle $\mathscr{O}(M)$.

3 Absolute Boundary Condition

The purpose of this section is to give an explicit expression for the absolute boundary condition on forms. Once the boundary condition is identified, the multiplicative functional M_t could be constructed accordingly.

Fix an $x \in \partial M$, we let $n(x)$ be the inward unit normal vector at x. For a k-form θ, we may decompose θ into its tangential and normal component, $\theta = \theta_{tan} + n(x) \wedge \beta$, with $\theta_{tan} \in \wedge_x^k \partial M$ and $\beta \in \wedge_x^{k-1} \partial M$. We denote $\theta_{norm} = \theta - \theta_{tan}$. The form θ is said to satisfy the absolute boundary condition if

$$\theta_{norm} = 0 \text{ and } (d\theta)_{norm} = 0.$$

Let $Q(x) : \wedge_x^* M \to \wedge_x^* M$ be the orthogonal projection to the tangential component, i.e., $Q(x)\theta = \theta_{tan}$. We extend Q (indeed $\tilde{Q}$) to a smooth, projection linear map on the whole bundle $\mathscr{O}(M)$ and let $P(x) = I - Q(x)$. $P(x)$ is the orthogonal projection to the normal component.

Recall that the second fundamental form $H : T_x\partial M \otimes_{\mathbb{R}} T_x\partial M \to \mathbb{R}$ is defined by

$$H(x)(X, Y) = \langle \nabla_X Y, n(x) \rangle, \quad X, Y \in T_x \partial M.$$

By duality, $H(x)$ can also be regarded as a linear map $H(x) : T_x\partial M \to T_x\partial M$ via the relation

$$\langle HX, Y \rangle = H\langle X, Y \rangle.$$

It is clear that $H(x)$ is symmetric on $T_x\partial M$. We extend H to the whole tangent space $T_x M$ by letting $H(x)n(x) = 0$, and denote the dual of H still by $H : \wedge_x^1 M \to \wedge_x^1 M$.

The following lemma gives an explicit expression for the absolute boundary condition on differential forms. Let

$$\partial\mathscr{O}(M) = \{u \in \mathscr{O}(M) : \pi u \in \partial M\}.$$

Lemma 3.1 *For any k-form θ on M, it satisfies the absolute boundary condition if and only if*

$$Q[N - H]\tilde{\theta} - P\tilde{\theta} = 0 \ \ on \ \ \partial\mathscr{O}(M).$$

Note that $\tilde{\theta}$ is the scalarization of θ, and N is the horizontal lift of n along the boundary ∂M.

Before we proceed to the proof of the above lemma, let us explain the various actions that appear in the above expression. Recall that N is a vector field on $\partial\mathscr{O}(M)$ and $\tilde{\theta}$ is a $\wedge^k\mathbb{R}^n$-valued function on $\mathscr{O}(M)$, thus $N\tilde{\theta}$ is naturally understood as the vector field acting on functions. The action $H\tilde{\theta}$ is more important. We know that H is a linear transformation on $\wedge^1_x M$ for $x \in \partial M$. For $\theta \in \wedge^k_x M$, the action $H\theta$ is the extension of H to $\wedge^* M$ as the Lie-algebra action(derivation) specified in Sect. 2. More specifically,

$$H(\theta_1 \wedge \ldots \wedge \theta_k) = \sum_{i=1}^{k} \theta_1 \wedge \ldots \wedge H\theta_i \wedge \ldots \wedge \theta_k,$$

where θ_i are 1-forms. Now $H\tilde{\theta}$ is simply $\tilde{H}\tilde{\theta}$.

Proof It is enough to show that

$$\theta_{norm} = 0 \Leftrightarrow P\tilde{\theta} = 0$$

and that, if $\theta_{norm} = 0$, then

$$(d\theta)_{norm} = 0 \Leftrightarrow Q[N - H]\tilde{\theta} = 0.$$

Fix any $x \in \partial M$. Let $\{E_i\}$ be a frame in a neighborhood of x with $E_1 = n$, the inward pointing unit normal vector field along the boundary and all other E_i's being tangent to the boundary. Furthermore we can choose the frame such that$\{E_i\}$ are orthonormal at x and $\nabla_{E_1} E_i = 0$ for all $i = 2, \ldots, n$ in a small neighborhood of x in M. To illustrate, we only prove the case when θ is a 2-form. The proof for k-forms will be clear, and actually identical when we understand what happens to 2-forms.

Let $\theta = \theta_{ij} E^i \wedge E^j$ be any 2-form, where $\{E^i\}$ is the dual of $\{E_i\}$. It's easy to see that $\theta_{norm} = 0$ is equivalent to $\theta_{1j} = \theta_{i1} = 0$ for all i, j, i.e., $P\tilde{\theta} = 0$.

Now we assume $P\tilde{\theta} = 0$ (i.e., $\theta_{1j} = \theta_{i1} = 0$ for all i, j). To see what $(d\theta)_{norm}$ means, we compute

$$\begin{aligned} d\theta &= E^k \wedge \nabla_{E_k}(\theta_{ij} E^i \wedge E^j) \\ &= E_k \theta_{ij} E^k \wedge E^i \wedge E^j + \theta_{ij} E^k \wedge \nabla_{E_k}(E^i \wedge E^j) \\ &= I_1 + I_2. \end{aligned}$$

Apparently

$$(I_1)_{norm} = E_1 \theta_{ij} E^1 \wedge E^i \wedge E^j, \tag{3.1}$$

since $\theta_{1j} = \theta_{i1} = 0$. On the other hand, we have

$$\begin{aligned} I_2 &= \theta_{ij} E^k \wedge (\nabla_{E_k} E^i \wedge E^j) + \theta_{ij} E^k \wedge (E^i \wedge \nabla_{E_k} E^j) \\ &= J_1 + J_2. \end{aligned}$$

Note that at x,

$$(\nabla_{E_k} E^i)(E_l) = -E^i(\nabla_{E_k} E_l) = -\langle \nabla_{E_k} E_l, E_i \rangle,$$

we therefore have

$$\nabla_{E_k} E^i = -\langle \nabla_{E_k} E_l, E_i \rangle E^l.$$

Hence at x,

$$J_1 = -\langle \nabla_{E_k} E_l, E_i \rangle \theta_{ij} E^k \wedge E^l \wedge E^j.$$

Keeping in mind that $\theta_{1j} = \theta_{i1} = 0$ and $\nabla_{E_1} E_i = 0$ for $i \neq 1$, we obtain

$$(J_1)_{norm} = -\langle \nabla_{E_k} E_1, E_i \rangle \theta_{ij} E^k \wedge E^1 \wedge E^j.$$

Re-indexing it we have

$$(J_1)_{norm} = \langle \nabla_{E_i} E_1, E_k \rangle \theta_{kj} E^1 \wedge E^i \wedge E^j. \tag{3.2}$$

Similarly

$$(J_2)_{norm} = \langle \nabla_{E_j} E_1, E_k \rangle \theta_{ik} E^1 \wedge E^i \wedge E^j. \tag{3.3}$$

Note that here $-\langle \nabla_{E_i} E_1, E_j \rangle$ is the matrix of second fundamental form on 1-forms. So we conclude that, by (3.1)–(3.3), when $\theta_{norm} = 0$, $(d\theta)_{norm} = 0$ is equivalent to

$$(E_1\theta_{ij} + \langle \nabla_{E_i} E_1, E_k \rangle \theta_{kj} + \langle \nabla_{E_j} E_1, E_k \rangle \theta_{ik}) E^1 \wedge E^i \wedge E^j = 0,$$

i.e., $Q(N - H)\tilde{\theta} = 0$. The proof is completed. □

Remark 3.2 Lemma 3.1 gives us a clear picture of the role the second fundamental form plays in the absolute boundary condition. Together with the discussion in Sect. 2, the initial boundary valued problem (1.1) can be lifted onto $\mathscr{O}(M)$ as

$$\begin{cases} \frac{\partial \tilde{\theta}}{\partial t} = \frac{1}{2}[\triangle_{\mathscr{O}(M)} + D^*\Omega]\tilde{\theta}, \\ \tilde{\theta}(\cdot, 0) = \tilde{\theta}_0, \\ Q[N - H]\tilde{\theta} - P\tilde{\theta} = 0. \end{cases} \tag{3.4}$$

Finally, we state an easy corollary of Lemma 3.1, which will be needed later. For each $x \in \partial M$, by the way we extended H to a linear map on $T_x M$, $\gamma_1 = 0$ is an eigenvalue of H associated to the eigenvector $n(x)$. Suppose that $\gamma_2(x), \ldots, \gamma_n(x)$ are other eigenvalues of H on $T_x \partial M$. We may define a real-valued function σ_k on ∂M by (see Donnelly-Li [2]),

$$\sigma_k(x) = \min_I \big(\gamma_{i_1}(x) + \gamma_{i_2}(x) + \ldots + \gamma_{i_k}(x)\big), \tag{3.5}$$

where $I = \{(i_1, \ldots, i_k)\}$ is the collection of multi-indices $(i_1, \ldots, i_k)$ such that $i_s \neq i_l$ if $s \neq l$; $s, l = 2, 3, \ldots, k$. Apparently, $\sigma_k(x)$ is a combination of eigenvalues of the second fundamental form H on $T_x \partial M$.

Corollary 3.3 *For any $x \in \partial M$ we have*

$$\sigma_k(x) = \inf_{\theta \in \wedge^k \partial M, |\theta| = 1} \langle H(x)\theta, \theta \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the canonical inner product on forms and $|\theta|^2 := \langle \theta, \theta \rangle$.

Proof Fix $x \in \partial M$, let $\{E_2, \ldots, E_n\}$ be a the set of orthonormal eigenvectors corresponding to the eigenvalues $\{\gamma_2, \ldots, \gamma_n\}$, and $\{E^i\}$ its dual. We first prove for any k-form θ with $|\theta| = 1$ we have

$$\sigma_k(x) \leq \langle H(x)\theta, \theta \rangle. \tag{3.6}$$

Let $\theta = \theta_{i_1,\ldots,i_k} E^{i_1} \wedge \ldots \wedge E^{i_k}$ with $|\theta|^2 = \sum \theta^2_{i_1,\ldots,i_k} = 1$. By the previous lemma we have

$$H(x)\theta = (\gamma_{i_1} + \ldots + \gamma_{i_k})\theta_{i_1,\ldots,i_k} E^{i_1} \wedge \ldots \wedge E^{i_k}.$$

Hence

$$\langle H(x)\theta, \theta\rangle = \sum(\gamma_{i_1} + \ldots + \gamma_{i_k})\theta^2_{i_1,\ldots,i_k} \geq \sigma_k(x)\sum \theta^2_{i_1,\ldots,i_k} = \sigma_k(x),$$

which proves (3.6). On the other hand, it's not hard to see that the equality can be achieved. The proof is completed. □

4 Reflecting Brownian Motion

Let $\omega = \{\omega_t\}$ be an n-dimensional Euclidean Brownian motion. Recall the definition of N in the previous section, and consider the following stochastic differential equation on the fame bundle $\mathscr{O}(M)$

$$du_t = \sum_{i=1}^{n} H_i(u_t) \circ d\omega_t^i + N(u_t)dl_t. \tag{4.1}$$

The solution $\{u_t\}$ is a horizontal reflecting Brownian motion starting at an initial frame u_0. Let $x_t = \pi u_t$. Then $\{x_t\}$ is a reflecting Brownian motion on M, with its transition density the Neumann heat kernel $p_M(t, x, y)$. The nondecreasing process l_t is the boundary local time, which increases only when $x_t \in \partial M$.

Now suppose that we have two smooth functions

$$R : \mathscr{O}(M) \to \mathrm{End}(\wedge^*\mathbb{R}^n), \quad A : \partial\mathscr{O}(M) \to \mathrm{End}(\wedge^*\mathbb{R}^n).$$

Define the $\mathrm{End}(\wedge^*\mathbb{R}^n)$-valued, continuous multiplicative functional $\{M_t\}$ by

$$dM_t + M_t\{-\frac{1}{2}R(u_t)dt + A(u_t)dl_t\} = 0, \;\; M_0 = I.$$

Since M_t takes values in $\mathrm{End}(\wedge^k\mathbb{R}^n)$, it is also helpful to think $\{M_t\}$ as a matrix-valued process.

Lemma 4.1 *Let* $\mathcal{L} = \frac{\partial}{\partial s} - \frac{1}{2}[\triangle_{\mathscr{O}(M)} + R]$ *and* $F : \mathscr{O}(M) \times \mathbb{R}_+ \to \wedge^*\mathbb{R}^n$ *be a solution to*

$$\begin{cases} \mathcal{L}F = 0 & u \in \mathscr{O}(M)/\partial\mathscr{O}(M) \\ (N - A)F = 0 & u \in \partial\mathscr{O}(M), \end{cases} \tag{4.2}$$

we have

$$M_t F(u_t, T - t) = F(u_0, T) + \int_0^t \langle M_s \nabla^H F(u_s, T - s), d\omega\rangle,$$

where $\nabla^H F = \{H_1 F, H_2 F, \dots, H_n F\}$ *is the horizontal gradient of a function* F *on* $\mathcal{O}(M)$. *In this case, we say that* $\{M_t\}$ *is the multiplicative functional associated with the operator* $\mathcal{L}$ *with the boundary condition* $(N - A)F = 0$.

Proof Apply Itô's formula to $M_t F(u_t, T-t)$. □

5 Discontinuous Multiplicative Functional

We have shown that the heat equation on k-forms with absolute boundary condition is equivalent to the following heat equation on $O(n)$-invariant functions F: $\mathcal{O}(M) \times \mathbb{R}_+ \to \wedge^k \mathbb{R}^n$:

$$\begin{cases} \frac{\partial F}{\partial t} = \frac{1}{2}[\triangle_{\mathcal{O}(M)} + D^* \Omega] F, \\ F(\cdot, 0) = f, \\ QNF - (H + P)F = 0. \end{cases} \tag{5.1}$$

Compared with the boundary condition in (4.2), $QN - (H+P)$ is degenerate, because Q is a projection (hence is not of full rank as a linear map). Thus Lemma 4.1 cannot be applied directly. In this section we follow closely the idea of Hsu [3] to construct the End($\wedge^k \mathbb{R}^n$)-valued multiplicative functional associated to (5.1).

Observe that the boundary condition in (5.1) consists of two orthogonal components:

$$Q[N - H]F = 0, \quad PF = 0. \tag{5.2}$$

We replace PF above by $(-\varepsilon PN + P)F$ and rewrite the boundary condition as

$$\left[N - H - \frac{P}{\varepsilon}\right] F = 0.$$

According to Lemma 4.1, the multiplicative functional for this approximate boundary condition is given by

$$dM_t^\varepsilon + M_t^\varepsilon \left\{ -\frac{1}{2} D^* \Omega(u_t) dt + \left[\frac{1}{\varepsilon} P(u_t) + H(u_t) \right] dl_t \right\} = 0. \tag{5.3}$$

In the rest of this section, we show that $\{M_t^\varepsilon\}$ converges to a discontinuous multiplicative functional $\{M_t\}$ which turns out to be the right one for the boundary condition (5.2).

Recall the definition of σ_k in (3.5) and let

$$\lambda(x) = \sup_{\theta \in \wedge_x^k M, \langle \theta, \theta \rangle = 1} \langle D^* R(x) \theta, \theta \rangle. \tag{5.4}$$

When $k = 1$, it is well known that $D^*R(x) = -Ric(x)$, where $Ric(x)$ is the Ricci transformation at x (see Hsu [4], for example), hence $\lambda(x)$ is the negative lower bound of the Ricci transform at x.

Proposition 5.1 *Let $|\cdot|_{2,2}$ be the norm of a linear transform on $\wedge^k\mathbb{R}^n$ with the standard Euclidean norm. Then for all positive ε such that $\varepsilon^{-1} \geq \min_{x\in\partial M}\sigma_k(x)$, we have*

$$|M_t^\varepsilon|_{2,2} \leq \exp\left[\frac{1}{2}\int_0^t \lambda(x_s)ds - \int_0^t \sigma_k(x_s)dl_s\right].$$

Proof We only outline the proof here, the technical details being mostly the same as that in [3]. Instead of considering M_t^ε, we prove for the adjoint (transpose, if we think M_t^ε as a matrix-valued process) of M_t^ε , namely $(M_t^\varepsilon)^T$. Let $f(t) = |(M_t^\varepsilon)^T\tilde{\theta}|^2 = \langle (M_t^\varepsilon)^T\tilde{\theta}, (M_t^\varepsilon)^T\tilde{\theta}\rangle$. Differentiate f with respect to t. By (5.3), our assumption on ε and standard estimate we have

$$df(t) \leq f(t)\{\lambda(x_t)dt - 2\sigma_k(x_t)dl_t\},$$

which gives us the desired result. □

The integrability of M_t^ε is given by the following lemma.

Lemma 5.2 *For any positive constant C, there is a constant C_1 depending on C but independent of x such that*

$$\mathbb{E}_x e^{Cl_t} \leq C_1 e^{C_1 t}.$$

Proof This can be obtained by a heat kernel upper bound and the strong Markov property of reflecting Brownian motion. See [3, Lemma 3.2] for a detailed proof. □

If we formally let $\varepsilon \downarrow 0$ in (5.3), one should expect $M_t^\varepsilon P(u_t) \to 0$ for all t such that $u_t \in \partial\mathscr{O}(M)$. The next lemma shows it is indeed the case. Define

$$T_{\partial M} = \inf\{s \geq 0 : x_s \in \partial M\} = \text{the first hitting time of } \partial M.$$

A point $t \geq T_{\partial M}$ such that $l_t - l_{t-\delta} > 0$ for all positive $\delta \leq t$ is called a *left support point* of the boundary local time l.

Proposition 5.3 *When $\varepsilon \downarrow 0$, $M_t^\varepsilon P(u_t) \to 0$ for all left support points $t \geq T_{\partial M}$.*

Proof The proof is almost identical to the one for 1-forms in [3]. For the convenience of the reader, we still provide some details here. We drop the superscript ε for simplicity. Let $\theta \in \wedge^k M$ be a k-form and define

$$f(s) = \langle M_s^T\tilde{\theta}, P(u_t)M_s^T\tilde{\theta}\rangle = \langle\tilde{\theta}, M_sP(u_t)M_s^T\tilde{\theta}\rangle.$$

Differentiating f with respect to s, by (5.3) we have $df(s) = -\frac{2}{\varepsilon}f(s) + dN_s$, which gives us

$$f(t) = e^{-2(l_t - l_{t-\delta})/\varepsilon} f(t-\delta) + \int_{t-\delta}^{t} e^{-2(l_t - l_s)/\varepsilon} dN_s. \tag{5.5}$$

Here dN_s is equal to

$$\begin{aligned}
&\frac{1}{\varepsilon}\langle \tilde{\theta}, M_s(2P(u_t) - P(u_s)P(u_t) - P(u_t)P(u_s))M_s^T\tilde{\theta}\rangle dl_s \\
&\quad + \langle \tilde{\theta}, \frac{1}{2}M_s(D^*\Omega(u_s)P(u_t) + P(u_t)(D^*\Omega(u_s))^T)M_s^T\tilde{\theta}\rangle ds \\
&\quad - \langle \tilde{\theta}, M_s(H(u_s)P(u_t) + P(u_t)H(u_s))M_s^T\tilde{\theta}\rangle dl_s.
\end{aligned}$$

In the above we used the fact that $H^T = H$ and $P^T = P$. By continuity of P and Proposition 5.1, for any $\eta > 0$ there exists a $\delta > 0$ such that, for all $s \in [t-\delta, t]$ with $x_s \in \partial M$,

$$\langle \tilde{\theta}, M_s(2P(u_t) - P(u_s)P(u_t) - P(u_t)P(u_s))M_s^T\tilde{\theta}\rangle \leq \eta|\tilde{\theta}|^2.$$

Also by Proposition 5.1, there is a constant C such that, for all $s \in [t-\delta, t]$ with $x_s \in \partial M$,

$$\langle \tilde{\theta}, \frac{1}{2}M_s(D^*\Omega(u_s)P(u_t) + P(u_t)(D^*\Omega(u_s))^T)M_s^T\tilde{\theta}\rangle \leq C|\tilde{\theta}|^2$$

and

$$\langle \tilde{\theta}, M_s(H(u_s)P(u_t) + P(u_t)H(u_s))M_s^T\tilde{\theta}\rangle \leq C|\tilde{\theta}|^2.$$

It follows that

$$|dN_s| \leq |\tilde{\theta}|^2\left[\left(\frac{\eta}{\varepsilon} + C\right)dl_s + Cds\right].$$

Substituting in (5.5), we obtain

$$\begin{aligned}
|M_tP(u_t)|_{2,2}^2 \leq & e^{-2(l_t - l_{t-\delta})/\varepsilon}|M_{t-\delta}|_{2,2}^2 + \frac{\eta + C\varepsilon}{2}\{1 - e^{-2(l_t - l_{t-\delta})/\varepsilon}\} \\
& + C\int_{t-\delta}^{t} e^{-2(l_t - l_{t-\delta})/\varepsilon} ds.
\end{aligned} \tag{5.6}$$

Because t is a left support point, $l_t - l_s > 0$ for all $s < t$. We first let $\varepsilon \downarrow 0$ and then $\eta \to 0$ in (5.6), we have $M_tP(u_t) \to 0$. □

We now come to the main result of this section, namely, the limit $\lim_{\varepsilon\to 0} M_t^\varepsilon = M_t$ exists. From the definition of M_t^ε, if t is such that $x_t \notin \partial M$ we have

$$dM_t^\varepsilon - \frac{1}{2} M_t^\varepsilon D^* \Omega(u_t) dt = 0.$$

Let $\{e(s,t), t \geq s\}$ be the solution of

$$\frac{d}{dt} e(s,t) - \frac{1}{2} e(s,t) D^* \Omega(u_t) = 0, \quad e(s,s) = I.$$

Then, for $t \geq T_{\partial M}$ we have $M_t^\varepsilon = M_{t*}^\varepsilon e(t_*, t)$. Here for each $t \geq T_{\partial M}$, t_* is defined to be the last exit time from ∂M, more precisely, $t_* = \sup\{s \leq t : x_s \in \partial M\}$.

Define

$$Y_t^\varepsilon = M_t^\varepsilon P(u_t), \quad Z_t^\varepsilon = M_t^\varepsilon Q(u_t).$$

Since when $t \leq T_{\partial M}$ we have $M_t^\varepsilon = e(0,t)$; and when $t \geq T_{\partial M}$ we have

$$M_t^\varepsilon = M_{t*}^\varepsilon e(t_*, t) = \{Z_{t*}^\varepsilon + Y_{t*}^\varepsilon\} e(t_*, t),$$

we can write

$$\begin{aligned} Y_t^\varepsilon &= I_{\{t \leq T_{\partial M}\}} M_t^\varepsilon P(u_t) + I_{\{t > T_{\partial M}\}} M_t^\varepsilon P(u_t) \\ &= I_{\{t \leq T_{\partial M}\}} e(0,t) P(u_t) + I_{\{t > T_{\partial M}\}} Z_{t*}^\varepsilon e(t_*, t) P(u_t) + \alpha_t^\varepsilon, \end{aligned} \tag{5.7}$$

where

$$\alpha_t^\varepsilon = I_{\{t > T_{\partial M}\}} Y_{t*}^\varepsilon e(t_*, t) P(u_t). \tag{5.8}$$

If $t > T_{\partial M}$, then t_* is a left support point of l. By Proposition 5.3, $Y_{t*}^\varepsilon \to 0$ as $\varepsilon \downarrow 0$; hence $\alpha_t^\varepsilon \to 0$. On the other hand, by Eq. (5.3) for M_t^ε we have

$$\begin{aligned} Z_t^\varepsilon &= Q(u_0) + \int_0^t dM_s^\varepsilon Q(u_s) + \int_0^t M_s^\varepsilon dQ(u_s) \\ &= Q(u_0) + \int_0^t [Y_s^\varepsilon + Z_s^\varepsilon] d\chi_s, \end{aligned} \tag{5.9}$$

where

$$d\chi_s = -H(u_s) dl_s + \frac{1}{2} D^* \Omega(u_s) Q(u_s) ds + dQ(u_s).$$

Formally letting $\varepsilon \downarrow 0$ in (5.7) and (5.9) above, we expect that the limit (Y_t, Z_t) satisfies following equations:

$$\begin{cases} Y_t = I_{\{t \le T_{\partial M}\}} e(0,t) P(u_t) + I_{\{t > T_{\partial M}\}} Z_{t*} e(t_*, t) P(u_t), \\ Z_t = Q(u_0) + \int_0^t (Y_s + Z_s) d\chi_s. \end{cases} \tag{5.10}$$

Substituting the first equation into the second, we obtain an equation for Z itself in the form

$$Z_t = Q(u_0) + \int_0^t \Phi(Z)_s d\chi_s, \tag{5.11}$$

where

$$\Phi(Z)_s = Z_s + I_{\{s \le T_{\partial M}\}} e(0,s) P(u_s) + I_{\{s > T_{\partial M}\}} Z_{s*} e(s_*, s) P(u_s).$$

Now we can state the main result in this section. For an End($\wedge^k \mathbb{R}^n$)-valued stochastic process $M = \{M_t\}$, we define

$$|M|_t = \sup_{0 \le s \le t} |M_s|_{2,2}.$$

Theorem 5.4 *We have*

(1) *Equation (5.11) has a unique solution Z. Define Y by the first equation in (5.10) and let $M_t = Y_t + Z_t$. Then $\{M_t\}$ is right continuous with left limits and $M_t P(u_t) = 0$ whenever $x_t \in \partial M$.*
(2) *For each fixed t,*

$$\mathbb{E}|Z^\varepsilon - Z|_t \to 0, \quad \mathbb{E}|Y_t^\varepsilon - Y_t|_{2,2}^2 \to 0, \text{ as } \varepsilon \downarrow 0.$$

Hence $\mathbb{E}|M_t^\varepsilon - M_t|_{2,2}^2 \to 0$ as $\varepsilon \downarrow 0$.

Proof The proof of the stated results follow the lines of proofs of Theorem 3.4 and Theorem 3.5 of [3]. □

Corollary 5.5 *For the limit process $\{M_t\}$ we have*

$$|M_t|_{2,2} \le \exp\left[\frac{1}{2}\int_0^t \lambda(x_s) ds - \int_0^t \sigma_k(x_s) dl_s\right].$$

Proof Letting $\varepsilon \downarrow 0$ in Lemma 5.1, the result follows immediately. □

Corollary 5.6 *The End($\wedge^k \mathbb{R}^n$)-valued process M_t is the multiplicative functional associated to Eq. (5.1).*

Proof Since F is a solution to (5.1), from Lemma 4.1 with $\mathcal{L} = \frac{\partial}{\partial s} - \frac{1}{2}[\triangle_{\mathscr{O}(M)} + D^*\Omega]$, we have

$$M_t^\varepsilon F(u_t, T-t) = F(u_0, T) + \int_0^t \langle M_s^\varepsilon \nabla^H F(u_s, T-s), d\omega_s\rangle + \int_0^t M_s^\varepsilon \left[N - \frac{1}{\varepsilon}P - H\right] F(u_s, T-s) dl_s.$$

The terms involving $1/\varepsilon$ vanish because, by the assumption on F, $P(u_s)F(u_s, T-s) = 0$ for $u_s \in \partial\mathscr{O}(M)$. Using the previous theorem, we let $\varepsilon \downarrow 0$ and note that $Q[N-H]F = [N-H]F$ and $M_s = MQ(u_s)$ when $u_s \in \partial\mathscr{O}(M)$ (by Theorem 5.4), we obtain the desired equality. □

6 Heart Kernel Representation and Applications

With the multiplicative functional M_t constructed in the previous section, we have the following probabilistic representation of the solution to (1.1).

Theorem 6.1 *Let $\theta \in \wedge^k M$ be the solution of the initial boundary value problem (1.1). Then*

$$\tilde{\theta}(u,t) = \mathbb{E}_u\{M_t\tilde{\theta}_0(u_t)\}. \tag{6.1}$$

Hence θ is given by

$$\theta(x,t) = u\mathbb{E}_x\{M_t u_t^{-1}\theta_0(x_t)\} \tag{6.2}$$

for any $u \in \mathscr{O}(M)$ such that $\pi u = x$.

Proof By Corollary 5.6, $\{M_s\tilde{\theta}(u_s, t-s), 0 \le s \le t\}$ is a martingale. Equating the expected values at $s = 0$ and $s = t$ gives us (6.1). The second equality is a restatement of the first one on the manifold M. □

There are several application with the above representation. We will examine two of them below. Let

$$p_M^*(t,x,y) : \wedge_y^* M \to \wedge_x^* M$$

be the heat kernel on differential forms with absolute boundary condition. By the above theorem we have

$$u\mathbb{E}_x\{M_t u_t^{-1}\theta(x_t)\} = \int_M p_M^*(t,x,y)\theta(y)dy, \quad \pi u = x. \tag{6.3}$$

On the other hand, we can also write

$$u\mathbb{E}_x\{M_t u_t^{-1}\theta(x_t)\} = u\mathbb{E}_x\mathbb{E}_x\{M_t u_t^{-1}\theta(x_t)|x_t = y\} \\ = \int_M p_M(t,x,y)u\mathbb{E}_x\{M_t u_t^{-1}\theta(x_t)|x_t = y\}dy. \tag{6.4}$$

Here $p_M(t,x,y)$ is the heat kernel on functions with Neumann boundary condition, i.e., the transition probability of $\{x_t\}$. From (6.3) and (6.4) the heat kernel on differential forms can be written as

$$p_M^*(t,x,y) = p_M(t,x,y)u\mathbb{E}_x\{M_t u_t^{-1}|x_t = y\}. \tag{6.5}$$

Recall that

$$\sigma_k = \min_I \gamma_{i_1} + \gamma_{i_2} + \ldots + \gamma_{i_k},$$

where $\gamma_2, \ldots, \gamma_n$ are eigenvalues of the second fundamental form of ∂M, and $I = \{(i_1, \ldots, i_k)\}$ is the collection of multi-indices $(i_1, \ldots, i_k)$ with $i_s = 2, 3, \ldots, k$ and $i_s = i_l$ if $s \neq l$; and that

$$\lambda(x) = \sup_{\theta\in\wedge_x^k M, \langle\theta,\theta\rangle=1} \langle D^* R(x)\theta, \theta\rangle. \tag{6.6}$$

We have the following heat kernel domination.

Theorem 6.2 *Let $p_M^k(t,x,y)$ be the heat kernel on k-forms. Define*

$$\bar{\sigma}_k = \inf_{x\in\partial M} \sigma_k \text{ and } \bar{\lambda} = \sup_{x\in\partial M} \lambda(x).$$

We have

$$|p_M^k(t,x,y)|_{2,2} \leq e^{\frac{1}{2}\bar{\lambda}t - \bar{\sigma}_k l_t} p_M(t,x,y),$$

where l_t is the Brownian motion boundary local time.

Proof This is a direct application of representation (6.5) and Proposition 5.1. □

Remark 6.3 When $\bar{\sigma}_k \geq 0$ then we have

$$|p_M^k(t,x,y)|_{2,2} \leq e^{\frac{1}{2}\bar{\lambda}t} p_M(t,x,y).$$

This special case was proved by Donnelly and Li [2], and Shigekawa [8].

For $\theta \in \wedge^k M$, let $P_t\theta(x) = \int_M p^*(t,x,y)\theta(y)dy$. Then we have the following generalized gradient inequality.

Theorem 6.4 *Keep all the notation above, we have*

$$|dP_t\theta(x)| \le \mathbb{E}_x \left\{ |d\theta| \exp\left[\frac{1}{2}\int_0^t \lambda(x_s)ds - \int_0^t \sigma_{k+1}(x_s)dl_s \right] \right\} .$$

Proof Let $\eta(x,t) = dP_t\theta(x)$. Then η is a $k+1$-form satisfying the absolute boundary condition, since $d\eta = d(dP_t\theta) = 0$ and $(\eta)_{norm} = (dP_t\theta)_{norm} = 0$. On the other hand, because d commute with the Hodge-de Rham Laplacian, we have

$$\frac{\partial\eta}{\partial t} = d\left(\frac{\partial P_t\theta}{\partial t}\right) = \frac{1}{2} d\,\Box P_t\theta = \frac{1}{2}\Box\, dP_t\theta = \frac{1}{2}\Box\eta.$$

So η is a solution to the heat equation (1.1). The rest of the proof is thus again an easy application of (6.2) and Proposition 5.1. ☐

Remark 6.5 When θ is a 0-form, i.e., a function on M, denoted as f. Then the above inequality reduces to

$$|\nabla P_t f(x)| \le \mathbb{E}_x \left\{ |\nabla f(x_t)| \exp\left[\frac{1}{2}\int_0^t \lambda(x_t)ds - \int_0^t \sigma_1(x_s)dl_s \right] \right\} ,$$

where σ_1 is just the smallest eigenvalue of the second fundamental form at x and $-\lambda$ is the low bound of Ricci curvature(since in one dimension $D^*R = -Ricci$). This special case was proved by Hsu [3]. In the case when M is a non-compact manifold, similar bound was obtained in [9, Chap. 3].

References

1. H. Airault, Pertubations singuliéres et solutions stochastiques de problémes de D.Neumann-Spencer. J. Math. Pures Appl. **55**(9), 233–267 (1976)
2. H. Donnelly, P. Li, Lower bounds for the eigenvalues of Riemannian manifolds. Mich. Math. J. **29**, 149–161 (1982)
3. E.P. Hsu, Multiplicative functional for the heat equation on manifolds with boundary. Mich. Math. J. **55**, 351–367 (2002)
4. E.P. Hsu, *Stochastic Analysis on Manifolds*. Graduate Series in Mathematics, vol. 38 (American Mathematical Society, Providence, 2002)
5. N. Ikeda, S. Watanabe, Heat equation and diffusion on Riemannian manifold with boundary. in *Proceedings of the International Symposium SDE (Kyoto, 1976)* (Wiley, New York, 1978), pp. 77–94
6. N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. (North-Holland, Amsterdam, 1989)
7. D.B. Ray, I.M. Singer, R-torsion and the Laplacian on Riemannian manifolds. Adv. Math. **7**, 145–210 (1971)
8. I. Shigekawa, Semigroup domination on a Riemannian manifold with boundary. Acta Applicandae Math. **63**, 385–410 (2000)
9. F.-Y. Wang, *Analysis for Diffusion Processes on Riemannian Manifolds*. Advanced Series on Statistical Science & Applied Probability, vol. 18 (World Scientific, Singapore, 2013)

Connections Between the Dirichlet and the Neumann Problem for Continuous and Integrable Boundary Data

Lucian Beznea, Mihai N. Pascu, and Nicolae R. Pascu

Dedicated to Rodrigo Banuelos on the occasion of his sixtieth birthday

Abstract We present results concerning the representation of the solution of the Neumann problem for the Laplace operator on the n-dimensional unit ball in terms of the solution of an associated Dirichlet problem. We show that the representation holds in the case of integrable boundary data, thus providing an explicit solution of the generalized solution of the Neumann problem.

Keywords Dirichlet problem • Dirichlet-to-Neumann operator • Infinite-dimensional Laplace operator • Laplace operator • Neumann problem

1991 Mathematics Subject Classification 31B05, 31B10, 42B37, 35J05, 35J25

L. Beznea (✉)
Simion Stoilow Institute of Mathematics of the Romanian Academy, Research Unit No. 2, P.O. Box 1-764, RO-014700 Bucharest, Romania

Faculty of Mathematics and Computer Science, University of Bucharest, Bucharest, Romania

Centre Francophone en Mathématique de Bucarest, Bucarest, Romania
e-mail: lucian.beznea@imar.ro

M.N. Pascu
Department of Mathematics and Computer Science, Transilvania University of Braşov, Str. Iuliu Maniu Nr. 50, 500091 Braşov, Romania
e-mail: mihai.pascu@unitbv.ro

N.R. Pascu
Department of Mathematics, Kennesaw State University, 1100 S. Marietta Parkway, Marietta, GA 30060-2896, USA
e-mail: npascu@kennesaw.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_4

1 Introduction

The classical Dirichlet and Neumann problems on a smooth bounded domain $D \subset \mathbb{R}^n$ $(n \geq 1)$ are the problem of finding $u \in C^2(D) \cap C(\overline{D})$ which solves

$$\begin{cases} \Delta u = 0 \text{ in } D \\ u = \varphi \text{ on } \partial D \end{cases}, \tag{1}$$

respectively $U \in C^2(D) \cap C^1(\overline{D})$ which solves

$$\begin{cases} \Delta U = 0 \text{ in } D \\ \frac{\partial U}{\partial \nu} = \phi \text{ on } \partial D \end{cases}, \tag{2}$$

where ν is the outward unit normal to the boundary of D.

As it is known, for continuous boundary data, the Dirichlet problem (1) has a unique solution and the Neumann problem (2) has a solution, unique up to additive constants, if we require in addition the condition $\int_{\partial D} \phi(z)\,\sigma(dz) = 0$. Note that this is a necessary condition for the existence of a solution, since by Green's first identity we have

$$\int_{\partial D} \phi(z)\,\sigma(dz) = \int_{\partial D} 1 \frac{\partial U}{\partial \nu}(z)\,\sigma(dz) = \int_{D} 1 \Delta U(z) + \nabla 1 \cdot \nabla U(z)\,dz = 0.$$

In this paper, we present explicit relations between the solutions of (1) and (2), which appeared recently in [4]. This shows that the Dirichlet and Neumann problems are "equally hard", in the sense that solving one of them leads to the solution of the other one. The central results for continuous boundary data (Theorem 1, and its extensions given in Theorems 2 and 5) provide an explicit relation between the solution(s) of (2) and (1), in the sense that the normalized solution of (2) can be found as a weighted average of the solution of (1).

The link between the solution of the Dirichlet problem and the Neumann problem is provided by the operator defined by (3). What is interesting here is that the same operator also provides a relationship between the solution of Dirichlet and Neumann problem in the infinite-dimensional setting of generalized Laplacian on an abstract Wiener space (see [4], Sect. 3). In Sect. 3 we show that the same operator can be used in order to construct a generalized solution of the Neumann boundary problem in the case of the unit ball in $\mathbb{R}^n$ $(n \geq 1)$ for integrable boundary data. While the existence of such a generalized solution for the Dirichlet boundary problem for integrable boundary data is known (the Perron-Wiener-Brelot theory [1, 9], or alternately the method of controlled convergence introduced by Cornea [5, 6]), in the case of the Neumann problem this is a new result, and it is the main result of the present paper, given in Theorem 12.

In Sect. 2, we consider the case of continuous boundary data for the Dirichlet and Neumann problems. This section is based on the recent results on the subject from

[4]. The main result giving the connection between the Dirichlet and the Neumann problem in the case of the unit ball is given in Theorem 1. The result can be extended to other operators besides the Laplacian, and in Theorem 2 we present such an extension.

As an application, in Theorem 4 we give an explicit representation of the inverse of the Dirichlet-to-Neumann operator (a particular case of the Poincaré-Steklov operator, which encapsulates the boundary response of a system modeled by a certain partial differential equation).

By using conformal mapping arguments (in the 2-dimensional case), the main result obtained in the case of the unit disk is extended (Theorem 5) to the general case of smooth bounded simply connected domains.

In what follows, we will identify as usual the complex plane $\mathbb{C}$ with $\mathbb{R}^2$, that is we identify the vector $(x, y) \in \mathbb{R}^2$ with the complex number $z = x + iy \in \mathbb{C}$. In particular, the dot product of two vectors $a, b \in \mathbb{R}^2$ will be written in terms of multiplication of complex numbers as $a \cdot b = \operatorname{Re}\left(a\overline{b}\right)$, and for a complex number $z \in \mathbb{C}$ we denote the real part and the imaginary part of z by $\operatorname{Re}(z)$, respectively $\operatorname{Im}(z)$. Also, for a function u defined on a subset D of $\mathbb{R}^2$ (or $\mathbb{C}$), we will write equivalently $u(x, y)$ or $u(z)$, where $z = x + iy \in D$.

For a smooth bounded domain we will be denote by $\sigma(\cdot)$ and $\sigma_0(\cdot)$ the surface measure on its boundary, respectively the surface measure normalized to have total mass 1.

2 The Case of Continuous Boundary Data

We start by recalling some recent results [4] concerning the equivalence between the Dirichlet and the Neumann problem for the Laplace operator in the case of continuous boundary data.

Heuristic arguments from Complex analysis (in the 2-dimensional case) led us to consider the operator which associates to a continuous function $u : D\subset \mathbb{R}^n \to \mathbb{R}$ with $u(0) = 0$ the function $U : D\subset \mathbb{R}^n \to \mathbb{R}$ defined by

$$U(z) = \int_0^1 \frac{u(\rho z)}{\rho} d\rho, \qquad z \in D, \tag{3}$$

where $D \subset \mathbb{R}^n$ is a smooth bounded subset, starlike with respect to the origin (i.e. $\rho z \in D$ for any $z \in D$ and $\rho \in [0, 1]$).

A first result concerning the operator defined above is that in the case of the n-dimensional unit ball $D = \mathbb{U} = \{z \in \mathbb{R}^n : |z| < 1\}$, the relation (3) provides an explicit solution of the Neumann problem (2) in terms of the Dirichlet problem (1) with the boundary condition $\varphi = \phi$. Conversely, since for a harmonic function the Laplacian and the partial derivatives commute, one can see that it is possible to solve the Dirichlet problem by solving an appropriate Neumann problem. The result is the following.

Theorem 1 ([4]) *The following assertions hold.*

(i) Assume $\phi : \partial\mathbb{U} \to \mathbb{R}$ *is continuous and satisfies* $\int_{\partial\mathbb{U}} \phi(z)\,\sigma_0(dz) = 0$*. If* u *is the solution of the Dirichlet problem (1) with boundary condition* $\varphi = \phi$ *on* $\partial\mathbb{U}$*, then*

$$U(z) = \int_0^1 \frac{u(\rho z)}{\rho} d\rho, \qquad z \in \overline{\mathbb{U}}, \tag{4}$$

is the solution to the Neumann problem (2) with $U(0) = 0$*.*

(ii) Assume $\varphi : \partial\mathbb{U} \to \mathbb{R}$ *is continuous. If* U *is the solution of the Neumann problem (2) with boundary condition* $\phi = \varphi - \int_{\partial\mathbb{U}} \varphi(\xi)\,\sigma_0(d\xi)$*, then*

$$u(z) = z \cdot \nabla U(z) + \int_{\partial\mathbb{U}} \varphi(\xi)\,\sigma_0(d\xi), \qquad z \in \overline{\mathbb{U}}, \tag{5}$$

is the solution to the Dirichlet problem (1).

As shown in [4], the previous result can also be applied to other operators besides the Laplacian. For example, considering the operator $\mathscr{L}$ defined by

$$\mathscr{L}f(z) = \sum_{i,j=1}^{n} a_{ij}(z) \frac{\partial^2 f}{\partial z_i \partial z_j}(z) + \sum_{i=1}^{n} a_i(z) \frac{\partial f}{\partial z_i}(z), \tag{6}$$

where the coefficients a_{ij} are smooth and homogeneous of degree $k \in [0, 1]$, i.e.

$$a_{ij}(\rho z) = \rho^k a_{ij}(z), \qquad 0 \le \rho \le 1, z \in \mathbb{U}, 1 \le i, j \le n, \tag{7}$$

and the coefficients a_i are also smooth and homogeneous of degree $k - 1$, i.e.

$$a_i(\rho z) = \rho^{k-1} a_{ij}(z), \qquad 0 \le \rho \le 1, z \in \mathbb{U}, 1 \le i \le n, \tag{8}$$

if u (with $u(0) = 0$) and U are related by (4), then

$$\mathscr{L}U(z) = \int_0^1 \rho^{1-k} \mathscr{L}u(\rho z)\,d\rho, \qquad z \in \mathbb{U},$$

and

$$\frac{\partial U}{\partial \nu}(z) = u(z), \qquad z \in \partial\mathbb{U}.$$

The previous observation leads to the following more general result.

Theorem 2 ([4]) *Assume $\phi : \partial\mathbb{U} \to \mathbb{R}$ is continuous. If u is the solution of the Dirichlet problem*

$$\begin{cases} \mathscr{L}u = 0 \text{ in } \mathbb{U} \\ u = \phi \text{ on } \partial\mathbb{U} \end{cases} \tag{9}$$

where $\mathscr{L}$ is the operator given by (6) which satisfies (7) and (8), and if $u(0) = 0$, then

$$U(z) = \int_0^1 \frac{u(\rho z)}{\rho} d\rho, \qquad z \in \overline{\mathbb{U}}, \tag{10}$$

is the solution to the Neumann problem

$$\begin{cases} \mathscr{L}U = 0 \ \text{ in } \mathbb{U} \\ \frac{\partial U}{\partial \nu} = \phi \ \text{ on } \partial\mathbb{U} \end{cases}, \tag{11}$$

with $U(0) = 0$.

Remark 3 The above result was stated in [4], Theorem 2, under the condition $\int_{\partial\mathbb{U}} \phi(z)\, \sigma_0(dz) = 0$ instead of $u(0) = 0$. If $\mathscr{L} = \Delta$, then these two conditions are equivalent, due to the Poisson formula.

As an application of the correspondence between the solutions of the Dirichlet and Neumann problems given above, we obtained an explicit representation of the inverse of the Dirichlet-to-Neumann operator Λ_n in the case of the unit ball $\mathbb{U} \subset \mathbb{R}^n$, $n \geq 2$. See for example [10, Sect. 5.0], or [4] for details on the Dirichlet-to-Neumann operator Λ_n and its inverse.

Theorem 4 *Assume $\phi : \partial\mathbb{U} \to \mathbb{R}$ is continuous and satisfies $\int_{\partial\mathbb{U}} \phi(\xi)\, \sigma(d\xi) = 0$. We have*

$$\Lambda_n^{-1}(\phi)(z) = \int_{\partial\mathbb{U}} \phi(\xi)\, k_n(z, \xi) \sigma_0(d\xi), \qquad z \in \partial\mathbb{U}, \tag{12}$$

where $k_n(z, \zeta) = \displaystyle\int_0^1 \frac{1}{\rho}\left(\frac{1-\rho^2}{|\rho z - \xi|^n} - 1\right) d\rho$, $z, \xi \in \partial\mathbb{U}$.

Explicitly, $k_2(z, \xi) = -2\ln|z - \xi|$, $k_3(z, \xi) = \frac{2}{|z-\xi|} - 2 + \ln 2 - \ln\left(\frac{|z-\xi|^2}{2} + |z - \xi|\right)$, and for $n > 4$ the kernel $k_n(z, \xi)$ can be computed using the recurrence formulae

$$k_n(z, \xi) = k_{n-2}(z, \xi) + \frac{2\left(1 - |z - \xi|^{n-2}\right)}{(n-2)\,|z - \xi|^{n-2}} - \frac{1 - |z - \xi|^{n-4}}{(n-4)\,|z - \xi|^{n-4}} + \left(1 - \frac{|z - \xi|^2}{2}\right) J_{n-2}(z, \xi), \tag{13}$$

where $J_n(z,\xi) = \int_0^1 \frac{1}{|\rho z - \xi|^n} d\rho$ *satisfies*

$$J_n(z,\xi) = \frac{4(n-3)J_{n-2}(z,\xi)}{(n-2)\left(4-|z-\xi|^2\right)|z-\xi|^2} + \frac{2\left(1+4|z-\xi|^{n-4}-|z-\xi|^{n-2}\right)}{(n-2)\left(4-|z-\xi|^2\right)|z-\xi|^{n-2}}. \tag{14}$$

Using conformal mapping arguments (in the 2-dimensional case), the result in Theorem 1 can be extended to the general case of a smooth bounded simply connected domain $D \subset \mathbb{C}$ ($C^{1,\alpha}$ boundary with $0 < \alpha < 1$ will suffice). The result is the following.

Theorem 5 ([4]) *Let* $D \subset \mathbb{C}$ *be a smooth bounded simply connected domain (*$C^{1,\alpha}$ *boundary with* $0 < \alpha < 1$ *will suffice), and for an arbitrarily fixed* $w_0 \in D$ *let* $f : \mathbb{U} \to D$ *be the conformal map of the unit disk* $\mathbb{U}$ *onto* D *with* $f(0) = w_0$, $\arg f'(0) = 0$, *and let* $g = f^{-1} : D \to \mathbb{U}$ *be its inverse.*

Assume $\phi : \partial D \to \mathbb{R}$ *is continuous and satisfies* $\int_{\partial D} \phi(w)\,\sigma(dw) = 0$. *If* u *is the solution of the Dirichlet problem (1) with boundary condition*

$$\varphi(w) = \frac{1}{|g'(w)|}\phi(w), \qquad w \in \partial D, \tag{15}$$

then

$$U(w) = \int_0^1 \frac{u(f(\rho g(w)))}{\rho} d\rho, \qquad w \in D, \tag{16}$$

is the solution to the Neumann problem (2) with $U(w_0) = 0$.

The result in Theorem 1 can also be extended to the case of Dirichlet and Neumann problems for the infinite-dimensional ball on an abstract Wiener space, in the setup stated in [7, 8], and [3]; for details see Sect. 3 from [4].

3 The Case of Integrable Boundary Data

In order to extend the result in Theorem 1 to a correspondence between the solutions of the Dirichlet problem and the Neumann problem for the unit ball in the general case of integrable boundary data, we will use Cornea's notion of *controlled convergence* [5, 6]. Even in the case of the unit ball $\mathbb{U} \subset \mathbb{R}^n$ ($n \geq 2$) which we consider here this is a new result, and it provides an explicit solution to the general Neumann problem for the Laplace operator.

It can be shown that in the case of the unit ball Cornea's approach is equivalent to the Perron-Wiener-Brelot approach for the generalized solution of the Dirichlet problem. More precisely, it can be shown that for integrable boundary data, both methods indicate that the generalized solution of the Dirichlet problem is given by the stochastic solution $H_{\mathbb{U}}^{f}$ defined by (18) below (see [5], Corollary 2, [6], Corollary 2.13, [1], Theorem 6.4.6, and [3], Theorem 4.5).

We will first recall the notion of *controlled convergence* introduced in [5, 6].

Definition 6 (Controlled convergence (A. Cornea, [5, 6])) Let $D \subset \mathbb{R}^d$ be a bounded open set, $\partial D \subset \Delta \subset \overline{D}$, $f : \partial D \to \mathbb{R}$ and $h, k : D \to \overline{\mathbb{R}}$, $k \geq 0$. The function h converges to f controlled by k (we write $h \xrightarrow{k} f$) if the following conditions hold:

For any set $A \subset D$ and any point $z_0 \in \overline{A} \cap \Delta$ we have

(*) If $\limsup_{A \ni z \to z_0} k(z) < +\infty$, then $f(z_0) \in \mathbb{R}$ and $\lim_{A \ni z \to z_0} h(z) = f(z)$.
(**) If $\lim_{A \ni z \to z_0} k(z) = +\infty$, then $\lim_{A \ni z \to z_0} \frac{h(z)}{1+k(z)} = 0$.

The function k will be called a *control function* for f.

Remark 7 It can be shown (see [6], Theorem 1.5, or [5], Theorem 1) that h converges to f controlled by k, in the sense of the above definition if and only if for any $z_0 \in \partial D$ the following equivalent conditions are satisfied:

(a) If $\liminf_{D \ni z \to z_0} k(z) < +\infty$, then $f(z_0) \in \mathbb{R}$ and $\lim_{D \ni z \to z_0} \frac{h(z) - f(z)}{1+k(z)} = 0$.
(b) If $\lim_{D \ni z \to z_0} k(z) = +\infty$, then $\lim_{D \ni z \to z_0} \frac{h(z)}{1+k(z)} = 0$.

Using the above definition, Cornea [5, 6] introduced the notion of generalized solution of the Dirichlet problem (1) as follows.

Definition 8 ([5, 6]) A *generalized solution* of the Dirichlet problem (1) is a harmonic function $u : D \to \mathbb{R}$ which satisfies

$$\lim_{z \to z_0} u(z) = \varphi(z), \qquad z_0 \in \partial D, \tag{17}$$

controlled by a continuous, non-negative (super)harmonic function $k : D \to \mathbb{R}_+$.

A function $\varphi : \partial D \to \overline{\mathbb{R}}$ for which the Dirichlet problem has a generalized solution is called *resolutive*. We denote by $\mathscr{R}(D)$ the set of resolutive functions $\varphi : \partial D \to \overline{\mathbb{R}}$.

In the same spirit, we propose the following definition for the generalized solution of the Neumann problem (2).

Definition 9 Let $D \subset \mathbb{R}^d$ be a bounded open set, $\partial D \subset \Delta \subset \overline{D}$, $h, k : D \to \overline{\mathbb{R}}$, $k \geq 0$. We say that the function h has a continuous extension to $\overline{D}$ controlled by k if the following conditions hold:

For any set $A \subset D$ and any point $z_0 \in \overline{A} \cap \Delta$ we have:

(i) If $\limsup_{A \ni z \to z_0} k(z) < +\infty$, we have $h(z_0) := \lim_{A \ni z \to z_0} h(z) \in \mathbb{R}$.
(ii) If $\lim_{A \ni z \to z_0} k(z) = +\infty$, then $\lim_{A \ni z \to z_0} \frac{h(z)}{1+k(z)} = 0$.

Remark 10 The previous remark shows that h has a continuous extension to $\overline{D}$ iff the equivalent conditions (a)–(b) above are satisfied.

If h has a continuous extension to $\overline{D}$ controlled by k, then the function h can be extended by continuity at the set of points z_0 belonging to the set $\left\{z_0 \in \partial D : \limsup_{D \ni z \to z_0} k(z) < +\infty\right\}$. On the set, $E = \left\{z_0 \in \partial D : \limsup_{D \ni z \to z_0} k(z) = +\infty\right\}$, the limit $\lim_{D \ni z \to z_0} h(z)$ may not exist, and the function h may fail to be continuous (this set of points is "controlled" by the function k).

Definition 11 A *generalized solution* of the Neumann problem (2) is a harmonic function $U : D \to \mathbb{R}$ which has a continuous extension to ∂D, controlled by a non-negative harmonic function $k : D \to \mathbb{R}_+$, and for any $z_0 \in \partial D$ for which $\limsup_{[0,z_0] \ni z \to z_0} k(z) < +\infty$ we have

$$\lim_{\varepsilon \nearrow 0} \frac{U(z_0 + \varepsilon \nu(z_0)) - U(z_0)}{\varepsilon} = \phi(z_0),$$

where $\nu(z)$ denotes the outward unit normal to the boundary of D at $z \in \partial D$.

In [5], the author showed that in the case of the unit ball $D = \mathbb{U} \subset \mathbb{R}^n$, every function $f \in L^1(\partial \mathbb{U}, \sigma_0)$ is resolutive for the Dirichlet problem. Moreover, by Beznea [2], the generalized solution coincides in fact with the stochastic solution, that is

$$u(z) = H^f_{\mathbb{U}}(z) = E^z f(B_\tau), \tag{18}$$

where $(B_t)_{t \geq 0}$ is a n-dimensional Brownian motion starting at $z \in \mathbb{U}$ and $\tau = \tau_{\partial \mathbb{U}} = \inf\{t \geq 0 : B_t \in \partial \mathbb{U}\}$ is the hitting time of the boundary of $\mathbb{U}$, and the controlled convergence to the boundary data f holds outside an exceptional (polar) set. It is also known (see [2], Corollary 4.3) that the generalized solution of the Dirichlet problem is unique.

With this preparation, we can now prove the main result, as follows.

Theorem 12 *Assume $\phi : \partial \mathbb{U} \to \mathbb{R}$ is integrable and satisfies $\int_{\partial \mathbb{U}} \phi(z)\, \sigma_0(dz) = 0$. If u is the generalized solution of the Dirichlet problem (1) with boundary condition $\varphi = \phi$ on $\partial \mathbb{U}$, then*

$$U(z) = \int_0^1 \frac{u(\rho z)}{\rho} d\rho, \qquad z \in \overline{\mathbb{U}}, \tag{19}$$

is a generalized solution to the Neumann problem (2) with $U(0) = 0$.

Proof Before proceeding with the proof, note that by symmetry, the exit distribution from $\mathbb{U}$ of the Brownian motion starting at the origin is the (normalized) surface measure σ_0 on $\partial \mathbb{U}$, and using the hypothesis we obtain

$$u(0) = E^0 \phi(B_{\tau_{\mathbb{U}}}) = \int_{\partial \mathbb{U}} \phi(z)\, \sigma_0(dz) = 0. \tag{20}$$

Using this, we have

$$\lim_{\rho\searrow 0}\frac{u(\rho z)}{\rho}=\lim_{\rho\searrow 0} z\frac{u(\rho z)-u(0)}{\rho z-0}=z\cdot\nabla u(0), \qquad z\in\mathbb{U}, \tag{21}$$

which shows that the integrand in (19) can be extended by continuity at $\rho = 0$, so U is well defined for all $z \in \mathbb{U}$. Note that the relation (20) also shows that $U(0) = 0$.

Next, we show that under the given hypotheses the function U has a continuous extension (controlled by k) to the boundary $\partial\mathbb{U}$, and it has the appropriate normal derivative. To be precise, for an arbitrary $z_0 \in \partial\mathbb{U}$ we'll show the following:

a) if $\liminf_{\mathbb{U}\ni z\to z_0} k(z) < \infty$ then there exists $U(z_0) \in \mathbb{R}$ such that $\lim_{\mathbb{U}\ni z\to z_0}\frac{U(z)-U(z_0)}{1+k(z)} = 0$. Moreover, if $\limsup_{[0,z_0]\ni z\to z_0} k(z) < \infty$, then $\lim_{\varepsilon\nearrow 0}\frac{U(z_0+\varepsilon\nu(z_0))-U(z_0)}{\varepsilon} = \phi(z_0)$.

b) if $\lim_{\mathbb{U}\ni z\to z_0} k(z) = \infty$, then $\lim_{\mathbb{U}\ni z\to z_0}\frac{U(z)}{1+k(z)} = 0$

Consider $z_0 \in \partial\mathbb{U}$ and assume that $\liminf_{z\to z_0} k(z) < \infty$. Since $u \to \phi$ controlled by k, we have $\lim_{z\to z_0}\frac{u(z)-u(z_0)}{1+k(z)} = 0$. Since u is continuous in $\mathbb{U}$, it follows that the function $(\rho, z) \mapsto \frac{u(\rho z)-u(\rho z_0)}{1+k(z)}$ is bounded on the set

$$\{(\rho,z)\in[0,1]\times\mathbb{U} : |\rho z - z_0| < \delta\},$$

for some $\delta > 0$. Since u and k are also bounded (being continuous) on the compact cone

$$C_\delta=\left\{\rho z:\rho\in[0,1], z\in\overline{\mathbb{U}} \text{ s.t. } |z-z_0|=\frac{\delta}{2}\right\}\cap\{z\in\mathbb{U}:|z-z_0|\ge\delta\}\subset\mathbb{U},$$

(see Fig. 1), it follows that $\frac{u(\rho z)-u(\rho z_0)}{1+k(z)}$ is bounded on $[0,1]\times\{z\in\mathbb{U}:|z-z_0|<\delta/2\}$.

Using the bounded convergence theorem and the above, we obtain

$$\lim_{z\to z_0}\frac{U(z)-U(z_0)}{1+k(z)}=\lim_{z\to z_0}\int_0^1\frac{u(\rho z)-u(\rho z_0)}{1+k(z)}d\rho=\int_0^1\lim_{z\to z_0}\frac{u(\rho z)-u(\rho z_0)}{1+k(z)}d\rho=0.$$

Suppose now that $z_0 \in \partial\mathbb{U}$ is such that $\lim_{z\to z_0} k(z) = \infty$. In order to show that $\lim_{z\to z_0}\frac{U(z)}{1+k(z)} = 0$, we will first show that for $\rho_0 \in (0, 1)$ arbitrarily fixed we have

$$\lim_{z\to z_0}\frac{1}{1+k(z)}\int_0^{\rho_0}\frac{u(\rho z)}{\rho}d\rho=0.$$

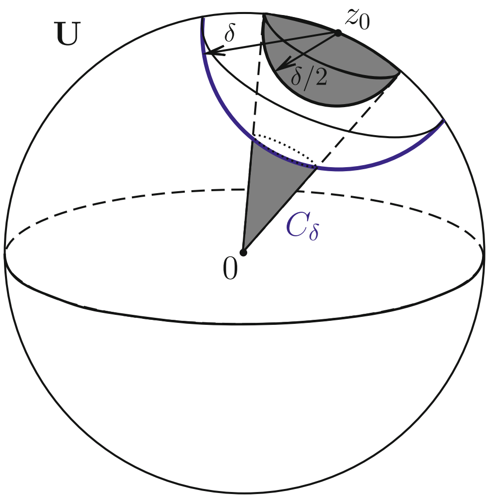

Fig. 1 The cone C_δ in the proof of Theorem 12

Since $u \in C^1(\mathbb{U})$, and using the substitution $w = \rho z$ we obtain

$$\lim_{\rho \searrow 0} \frac{u(\rho z)}{\rho z} = \lim_{w \to 0} \frac{u(w)}{w} = \nabla u(0),$$

uniformly with respect to $z \in \mathbb{U}$. It follows that

$$\left| \frac{u(\rho z)}{(1 + k(z))\rho} \right| \le 1 + |z \cdot \nabla u(0)| \le 1 + |\nabla u(0)|$$

is bounded for $\rho < \rho_1$ sufficiently small, uniformly with respect to $z \in \mathbb{U}$. For $\rho \in [\rho_1, \rho_0]$, we have

$$\left| \frac{u(\rho z)}{(1 + k(z))\rho} \right| \le \frac{1}{\rho_1} \max_{|w| \le \rho_0} |u(w)|,$$

and combining with the above we conclude that $\frac{u(\rho z)}{(1+k(z))\rho}$ is bounded for $\rho \in [0, \rho_0]$, uniformly with respect to $z \in \mathbb{U}$. Using the bounded convergence theorem and $\lim_{z \to z_0} k(z) = \infty$, we conclude

$$\lim_{z \to z_0} \frac{1}{1 + k(z)} \int_0^{\rho_0} \frac{u(\rho z)}{\rho} d\rho = \int_0^{\rho_0} \lim_{z \to z_0} \frac{u(\rho z)}{(1 + k(z))\rho} d\rho = 0, \tag{22}$$

thus proving the claim. In order to prove that $\lim_{z\to z_0}\frac{U(z)}{1+k(z)}=0$, it remains to show that

$$\lim_{z\to z_0}\frac{1}{1+k(z)}\int_{\rho_0}^{1}\frac{u(\rho z)}{\rho}d\rho=0,$$

for an arbitrarily fixed $\rho_0\in(0,1)$.

For $\varepsilon>0$ arbitrarily fixed, consider $n_0\in\mathbb{N}$ such that $n_0\geq\frac{1}{\varepsilon}$, and let $\phi_0=\frac{1}{n_0}\sum_{i=1}^{n_0}f_i$. We have

$$\begin{aligned}u(z)+\varepsilon k(z) &\geq u(z)+\frac{1}{n_0}k(z)=H_{\phi}^{\mathbb{U}}(z)+\frac{1}{n_0}H^{\mathbb{U}}_{\sum_{n\geq1}(f_n-g_n)}(z)\\ &\geq H_{\phi}^{\mathbb{U}}(z)+\frac{1}{n_0}H^{\mathbb{U}}_{\sum_{n=1}^{n_0}(f_n-g_n)}(z)\\ &=H_{\phi_0}^{\mathbb{U}}(z)+\frac{1}{n_0}H^{\mathbb{U}}_{\sum_{n=1}^{n_0}(f-g_n)}(z)\\ &\geq H_{\phi_0}^{\mathbb{U}}(z),\end{aligned}$$

for any $z\in\mathbb{U}$.

Since by construction the functions f_n are lower bounded, there exists $M>0$ such $\phi_0\geq M$, and therefore $H_{\phi_0}^{\mathbb{U}}(z)\geq M$ for any $z\in\mathbb{U}$. We obtain

$$\begin{aligned}\frac{1}{1+k(z)}\int_{\rho_0}^{1}\frac{u(\rho z)}{\rho}d\rho &= \frac{1}{1+k(z)}\int_{\rho_0}^{1}\left(\frac{u(\rho z)+\varepsilon k(\rho z)}{\rho}-\varepsilon\frac{k(\rho z)}{\rho}\right)d\rho\\ &\geq \frac{1}{1+k(z)}\int_{\rho_0}^{1}\left(\frac{M}{\rho}-\varepsilon\frac{k(\rho z)}{\rho}\right)d\rho\\ &\geq \frac{-M\ln\rho_0}{1+k(z)}-\frac{\varepsilon}{\rho_0(1+k(z))}\int_{\rho_0}^{1}k(\rho z)\,d\rho\end{aligned}$$

An argument similar to the one in the beginning of the proof shows that $\frac{k(\rho z)}{1+k(z)}$ is bounded for $\rho\in[\rho_0,1]$ and z in a neighborhood of z_0. Passing to the limit in the above inequality, and using $\lim_{z\to z_0}k(z)=\infty$ (which in particular implies $\lim_{z\to z_0}\frac{k(\rho z)}{1+k(z)}=0$ for any $\rho\in[\rho_0,1)$, and $\lim_{z\to z_0}\frac{k(z)}{1+k(z)}=1$), we obtain

$$\begin{aligned}\liminf_{z\to z_0}\frac{1}{1+k(z)}\int_{\rho_0}^{1}\frac{u(\rho z)}{\rho}d\rho &\geq \liminf_{z\to z_0}\frac{-M\ln\rho_0}{1+k(z)}-\frac{\varepsilon}{\rho_0}\int_{\rho_0}^{1}\limsup_{z\to z_0}\frac{k(\rho z)}{1+k(z)}\\ &\times d\rho\geq-\frac{\varepsilon}{\rho_0}(1-\rho_0).\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary chosen and $\rho_0 \in (0, 1)$, the above shows that

$$\liminf_{z \to z_0} \frac{1}{1 + k(z)} \int_{\rho_0}^{1} \frac{u(\rho z)}{\rho} d\rho \geq 0.$$

Repeating the proof above with $\tilde{\phi} = -\phi$ in place of ϕ (for which the corresponding functions are $\tilde{u} = -u$, $\tilde{k} = k$, and $\tilde{U} = -U$), we also have

$$\limsup_{z \to z_0} \frac{1}{1 + k(z)} \int_{\rho_0}^{1} \frac{u(\rho z)}{\rho} d\rho \leq 0,$$

and therefore

$$\lim_{z \to z_0} \frac{1}{1 + k(z)} \int_{\rho_0}^{1} \frac{u(\rho z)}{\rho} d\rho = 0.$$

This, combined with (22) shows that

$$\lim_{z \to z_0} \frac{1}{1 + k(z)} \int_{0}^{1} \frac{u(\rho z)}{\rho} d\rho = 0,$$

concluding the proof of part b) of claim.

To see that U has the prescribed normal derivative on $\partial\mathbb{U}$ (recall that we are using the outward normal $\nu(z_0) = z_0$ to the boundary of $\partial\mathbb{U}$), fix $z_0 \in \partial\mathbb{U}$ such that $\limsup_{[0,z_0]\ni z \to z_0} k(z) < \infty$. Since $u \to \phi$ controlled by k, choosing the particular set $A = [0, z_0]$ in the Definition 6 of controlled convergence, we have that $\lim_{\rho \nearrow 1} u(\rho z_0) = \phi(z_0) \in \mathbb{R}$.

Using a change of variables and the mean value theorem, we obtain

$$\begin{aligned}
\lim_{\varepsilon \nearrow 0} \frac{U(z_0 + \varepsilon \nu(z_0)) - U(z_0)}{\varepsilon} &= \lim_{\varepsilon \nearrow 0} \frac{1}{\varepsilon} \Big(\int_0^{1+\varepsilon} \frac{u(\rho z_0)}{\rho} d\rho - \int_0^1 \frac{u(\rho z_0)}{\rho} d\rho \Big) \\
&= \lim_{\varepsilon \nearrow 0} \frac{1}{\varepsilon} \int_1^{1+\varepsilon} \frac{u(\rho z_0)}{\rho} d\rho = \lim_{\varepsilon \nearrow 0} \frac{u(\rho^* z_0)}{\rho^*} = \phi(z_0),
\end{aligned}$$

where we denoted by $\rho^* \in (1 + \varepsilon, 1)$ the intermediate point given by the mean value theorem. This shows that the directional derivative of the function U in the direction of the normal to the boundary of $\mathbb{U}$ has the appropriate value $\frac{\partial U}{\partial \nu}(z_0) = \phi(z_0)$ at z_0, thus concluding the proof. □

Acknowledgements The first author acknowledges support from the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-III-P4-ID-PCE-2016-0372. The second author kindly acknowledges the support by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PNII-ID-PCCE-2011-2-0015.

References

1. D.H. Armitage, S.J. Gardiner, *Classical Potential Theory*. Springer Monographs in Mathematics (Springer, London, 2001)
2. L. Beznea, The stochastic solution of the Dirichlet problem and controlled convergence. Lect. Notes Semin. Interdisciplinare Mat. **10**, 115–136 (2011)
3. L. Beznea, A. Cornea, M. Röckner, Potential theory of infinite dimensional Lévy processes. J. Funct. Anal. **261**(10), 2845–2876 (2011)
4. L. Beznea, M.N. Pascu, N.R. Pascu, An equivalence between the Dirichlet and the Neumann problem for the Laplace operator. Potential Anal. **44**(4), 655–672 (2016)
5. A. Cornea, Résolution du probléme de Dirichlet et comportement des solutions à la frontière à l'aide des fonctions de contrôle. C. R. Acad. Sci. Paris Ser. I Math. **320**, 159–164 (1995)
6. A. Cornea, Applications of controlled convergence in analysis, in *Analysis and Topology* (World Science, Singapore, 1998)
7. V. Goodman, Harmonic functions on Hilbert space. J. Funct. Anal. **10**, 451–470 (1972)
8. L. Gross, Potential theory on Hilbert space. J. Funct. Anal. **1**, 123–181 (1967)
9. L.L. Helms, *Introduction to Potential Theory, Pure and Applied Mathematics*, vol. XXII. (Robert E. Krieger Publishing Co., Huntington, 1975)
10. V. Isakov, *Inverse Problems for Partial Differential Equations*. Applied Mathematical Sciences, vol. 127, 2nd edn. (Springer, New York, 2006)

Decomposition and Limit Theorems for a Class of Self-Similar Gaussian Processes

Daniel Harnett and David Nualart

Abstract We introduce a new class of self-similar Gaussian stochastic processes, where the covariance is defined in terms of a fractional Brownian motion and another Gaussian process. A special case is the solution in time to the fractional-colored stochastic heat equation described in Tudor (Analysis of variations for self-similar processes: a stochastic calculus approach. Springer, Berlin, 2013). We prove that the process can be decomposed into a fractional Brownian motion (with a different parameter than the one that defines the covariance), and a Gaussian process first described in Lei and Nualart (Stat Probab Lett 79:619–624, 2009). The component processes can be expressed as stochastic integrals with respect to the Brownian sheet. We then prove a central limit theorem about the Hermite variations of the process.

Keywords Fractional Brownian motion • Hermite variations • Self-similar processes • Stochastic heat equation

AMS 2010 Classification 60F05, 60G18, 60H07

1 Introduction

The purpose of this paper is to introduce a new class of Gaussian self-similar stochastic processes related to stochastic partial differential equations, and to establish a decomposition in law and a central limit theorem for the Hermite variations of the increments of such processes.

D. Harnett
Department of Mathematical Sciences, University of Wisconsin Stevens Point, Stevens Point, WI 54481, USA
e-mail: dharnett@uwsp.edu

D. Nualart (✉)
Department of Mathematics, University of Kansas, 405 Snow Hall, Lawrence, KS 66045-2142, USA
e-mail: nualart@ku.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_5

Consider the d-dimensional stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \dot{W}, \ t \geq 0, \ x \in \mathbb{R}^d, \tag{1.1}$$

with zero initial condition, where $\dot{W}$ is a zero mean Gaussian field with a covariance of the form

$$\mathbb{E}\left[\dot{W}^H(t,x)\dot{W}^H(s,y)\right] = \gamma_0(t-s)\Lambda(x-y), \quad s,t \geq 0, \ x,y \in \mathbb{R}^d.$$

We are interested in the process $U = \{U_t, t \geq 0\}$, where $U_t = u(t,0)$.

Suppose that $\dot{W}$ is white in time, that is, $\gamma_0 = \delta_0$ and the spatial covariance is the Riesz kernel, that is, $\Lambda(x) = c_{d,\beta}|x|^{-\beta}$, with $\beta < \min(d,2)$ and $c_{d,\beta} = \pi^{-d/2}2^{\beta-d}\Gamma(\beta/2)/\Gamma((d-\beta)/2)$. Then U has the covariance (see [14])

$$\mathbb{E}[U_tU_s] = D\left((t+s)^{1-\frac{\beta}{2}} - |t-s|^{1-\frac{\beta}{2}}\right), \quad s,t \geq 0, \tag{1.2}$$

for some constant

$$D = (2\pi)^{-d}(1-\beta/2)^{-1}\int_{\mathbb{R}^d} e^{-\frac{|\xi|^2}{2}}\frac{d\xi}{|\xi|^{d-\beta}}. \tag{1.3}$$

Up to a constant, the covariance (1.2) is the covariance of the *bifractional Brownian motion* with parameters $H = \frac{1}{2}$ and $K = 1 - \frac{\beta}{2}$. We recall that, given constants $H \in (0,1)$ and $K \in (0,1)$, the bifractional Brownian motion $B^{H,K} = \{B_t^{H,K}, t \geq 0\}$, introduced in [4], is a centered Gaussian process with covariance

$$R_{H,K}(s,t) = \frac{1}{2^K}\left((t^{2H}+s^{2H})^K - |t-s|^{2HK}\right), \quad s,t \geq 0.$$

When $K = 1$, the process $B^H = B^{H,1}$ is simply the fractional Brownian motion (fBm) with Hurst parameter $H \in (0,1)$, with covariance $R_H(s,t) = R_{H,1}(s,t)$. In [5], Lei and Nualart obtained the following decomposition in law for the bifractional Brownian motion

$$B^{H,K} = C_1 B^{HK} + C_2 Y^K_{t^{2H}},$$

where B^{HK} is a fBm with Hurst parameter HK, the process Y^K is given by

$$Y_t^K = \int_0^\infty y^{-\frac{1+K}{2}}(1-e^{-yt})dW_y, \tag{1.4}$$

with $W = \{W_y, y \geq 0\}$ a standard Brownian motion independent of $B^{H,K}$, and C_1, C_2 are constants given by $C_1 = 2^{\frac{1-K}{2}}$ and $C_2 = \sqrt{\frac{2-K}{\Gamma(1-K)}}$. The process Y^K

has trajectories which are infinitely differentiable on $(0,\infty)$ and Hölder continuous of order $HK-\epsilon$ in any interval $[0,T]$ for any $\epsilon>0$. In particular, this leads to a decomposition in law of the process U with covariance (1.2) as the sum of a fractional Brownian motion with Hurst parameter $\frac{1}{2}-\frac{\beta}{4}$ plus a regular process.

The classical one-dimensional space-time white noise can also be considered as an extension of the covariance (1.2) if we take $\beta=1$. In this case the covariance corresponds, up to a constant, to that of a bifractional Brownian motion with parameters $H=K=\frac{1}{2}$.

The case where the noise term $\dot{W}$ is a fractional Brownian motion with Hurst parameter $H\in(\frac{1}{2},1)$ in time and a spatial covariance given by the Riesz kernel, that is,

$$\mathbb{E}\left[\dot{W}^H(t,x)\dot{W}^H(s,y)\right]=\alpha_H c_{d,\beta}|s-t|^{2H-2}|x-y|^{-\beta},$$

where $0<\beta<\min(d,2)$ and $\alpha_H=H(2H-1)$, has been considered by Tudor and Xiao in [14]. In this case the corresponding process U has the covariance

$$\mathbb{E}[U_tU_s]=D\alpha_H\int_0^t\int_0^s|u-v|^{2H-2}(t+s-u-v)^{-\gamma}dudv, \tag{1.5}$$

where D is given in (1.3) and $\gamma=\frac{d-\beta}{2}$. This process is self-similar with parameter $H-\frac{\gamma}{2}$ and it has been studied in a series of papers [1, 8, 12–14]. In particular, in [14] it is proved that the process U can be decomposed into the sum of a scaled fBm with parameter $H-\frac{\gamma}{2}$, and a Gaussian process V with continuously differentiable trajectories. This decomposition is based on the stochastic heat equation. As a consequence, one can derive the exact uniform and local moduli of continuity and Chung-type laws of the iterated logarithm for this process. In [12], assuming that $d=1,2$ or 3, a central limit theorem is obtained for the renormalized quadratic variation

$$V_n=n^{2H-\gamma-\frac{1}{2}}\sum_{j=0}^{n-1}\left\{(U_{(j+1)T/n}-U_{jT/n})^2-\mathbb{E}\left[(U_{(j+1)t/n}-U_{jT/n})^2\right]\right\},$$

assuming $\frac{1}{2}<H<\frac{3}{4}$, extending well-known results for fBm (see for example [6, Theorem 7.4.1]).

The purpose of this paper is to establish a decomposition in law, similar to that obtained by Lei and Nualart in [5] for the bifractional Brownian motion, and a central limit theorem for the Hermite variations of the increments, for a class of self-similar processes that includes the covariance (1.5). Consider a centered Gaussian process $\{X_t,t\geq 0\}$ with covariance

$$R(s,t)=\mathbb{E}[X_sX_t]=\mathbb{E}\left[\left(\int_0^t Z_{t-r}dB_r^H\right)\left(\int_0^s Z_{s-r}dB_r^H\right)\right], \tag{1.6}$$

where

(i) $B^H = \{B_t^H, t \geq 0\}$ is a fBm with Hurst parameter $H \in (0, 1)$.
(ii) $Z = \{Z_t, t > 0\}$ is a zero-mean Gaussian process, independent of B^H, with covariance

$$\mathbb{E}[Z_s Z_t] = (s + t)^{-\gamma}, \tag{1.7}$$

where $0 < \gamma < 2H$.

In other words, X is a Gaussian process with the same covariance as the process $\{\int_0^t Z_{t-r} dB_r^H, t \geq 0\}$, which is not Gaussian.

When $H \in (\frac{1}{2}, 1)$, the covariance (1.6) coincides with (1.5) with $D = 1$. However, we allow the range of parameters $0 < H < 1$ and $0 < \gamma < 2H$. In other words, up to a constant, X has the law of the solution in time of the stochastic heat equation (1.1), when $H \in (0, 1)$, $d \geq 1$ and $\beta = d - 2\gamma$. Also of interest is that X can be constructed as a sum of stochastic integrals with respect to the Brownian sheet (see the proof of Theorem 1).

1.1 Decomposition of the Process X

Our first result is the following decomposition in law of the process X as the sum of a fractional Brownian motion with Hurst parameter $\frac{\alpha}{2} = H - \frac{\gamma}{2}$ plus a process with regular trajectories.

Theorem 1 *The process X has the same law as $\{\sqrt{\kappa} B_t^{\frac{\alpha}{2}} + Y_t, t \geq 0\}$, where here and in what follows, $\alpha = 2H - \gamma$,*

$$\kappa = \frac{1}{\Gamma(\gamma)} \int_0^\infty \frac{z^{\gamma-1}}{1+z^2}\, dz, \tag{1.8}$$

$B^{\frac{\alpha}{2}}$ is a fBm with Hurst parameter $\frac{\alpha}{2}$, and Y (up to a constant) has the same law as the process Y^K defined in (1.4), with $K = 2\alpha + 1$, that is, Y is a centered Gaussian process with covariance given by

$$\mathbb{E}\,[Y_t Y_s] = \lambda_1 \int_0^\infty y^{-\alpha-1}(1 - e^{-yt})(1 - e^{-ys})\, dy,$$

where

$$\lambda_1 = \frac{4\pi}{\Gamma(\gamma)\Gamma(2H+1)\sin(\pi H)} \int_0^\infty \frac{\eta^{1-2H}}{1+\eta^2} d\eta.$$

The proof of this theorem is given in Sect. 3.

1.2 Hermite Variations of the Process

For each integer $q \geq 0$, the qth Hermite polynomial is given by

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q} e^{-\frac{x^2}{2}}.$$

See [6, Sect. 1.4] for a discussion of properties of these polynomials. In particular, it is well known that the family $\{\frac{1}{\sqrt{q!}} H_q, q \geq 0\}$ constitutes an orthonormal basis of the space $L^2(\mathbb{R}, \gamma)$, where γ is the $N(0, 1)$ measure.

Suppose $\{Z_n, n \geq 1\}$ is a stationary, Gaussian sequence, where each Z_n follows the $N(0, 1)$ distribution with covariance function $\rho(k) = \mathbb{E}[Z_n Z_{n+k}]$. If $\sum_{k=1}^{\infty} |\rho(k)|^q < \infty$, it is well known that as n tends to infinity, the Hermite variation

$$V_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} H_q(Z_j) \tag{1.9}$$

converges in distribution to a Gaussian random variable with mean zero and variance given by $\sigma^2 = q! \sum_{k=1}^{\infty} \rho(k)^q$. This result was proved by Breuer and Major in [3]. In particular, if B^H is a fBm, then the sequence $\{Z_{j,n}, 0 \leq j \leq n-1\}$ defined by

$$Z_{j,n} = n^H \left(B^H_{\frac{j+1}{n}} - B^H_{\frac{j}{n}} \right)$$

is a stationary sequence with unit variance. As a consequence, if $H < 1 - \frac{1}{q}$, we have that

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} H_q \left(n^H \left(B^H_{\frac{j+1}{n}} - B^H_{\frac{j}{n}} \right) \right)$$

converges to a normal law with variance given by

$$\sigma_q^2 = \frac{q!}{2^q} \sum_{m \in \mathbb{Z}} \left(|m+1|^{2H} - 2|m|^{2H} + |m-1|^{2H} \right)^q . \tag{1.10}$$

See [3] and Theorem 7.4.1 of [6].

The above Breuer-Major theorem can not be applied to our process because X is not necessarily stationary. However, we have a comparable result.

Theorem 2 *Let $q \geq 2$ be an integer and fix a real $T > 0$. Suppose that $\alpha < 2 - \frac{1}{q}$, where α is defined in Theorem 1. For $t \in [0, T]$, define,*

$$F_n(t) = n^{-\frac{1}{2}} \sum_{j=0}^{\lfloor nt \rfloor - 1} H_q \left(\frac{\Delta X_{\frac{j}{n}}}{\left\| \Delta X_{\frac{j}{n}} \right\|_{L^2(\Omega)}} \right),$$

where $H_q(x)$ denotes the qth Hermite polynomial. Then as $n \to \infty$, the stochastic process $\{F_n(t), t \in [0,T]\}$ converges in law in the Skorohod space $D([0,T])$, to a scaled Brownian motion $\{\sigma B_t, t \in [0,T]\}$, where $\{B_t, t \in [0,T]\}$ is a standard Brownian motion and $\sigma = \sqrt{\sigma^2}$ is given by

$$\sigma^2 = \frac{q!}{2^q} \sum_{m \in \mathbb{Z}} (|m+1|^\alpha - 2|m|^\alpha + |m-1|^\alpha)^q . \tag{1.11}$$

The proof of this theorem is given in Sect. 4.

2 Preliminaries

2.1 Analysis on the Wiener Space

The reader may refer to [6, 7] for a detailed coverage of this topic. Let $Z = \{Z(h), h \in \mathcal{H}\}$ be an *isonormal Gaussian process* on a probability space $(\Omega, \mathcal{F}, P)$, indexed by a real separable Hilbert space $\mathcal{H}$. This means that Z is a family of Gaussian random variables such that $\mathbb{E}[Z(h)] = 0$ and $\mathbb{E}[Z(h)Z(g)] = \langle h, g \rangle_{\mathcal{H}}$ for all $h, g \in \mathcal{H}$.

For integers $q \geq 1$, let $\mathcal{H}^{\otimes q}$ denote the qth tensor product of $\mathcal{H}$, and $\mathcal{H}^{\odot q}$ denote the subspace of symmetric elements of $\mathcal{H}^{\otimes q}$.

Let $\{e_n, n \geq 1\}$ be a complete orthonormal system in $\mathcal{H}$. For elements $f, g \in \mathcal{H}^{\odot q}$ and $p \in \{0, \dots, q\}$, we define the pth-order contraction of f and g as that element of $\mathcal{H}^{\otimes 2(q-p)}$ given by

$$f \otimes_p g = \sum_{i_1, \dots, i_p = 1}^{\infty} \left\langle f, e_{i_1} \otimes \cdots \otimes e_{i_p} \right\rangle_{\mathcal{H}^{\otimes p}} \otimes \left\langle g, e_{i_1} \otimes \cdots \otimes e_{i_p} \right\rangle_{\mathcal{H}^{\otimes p}} , \tag{2.1}$$

where $f \otimes_0 g = f \otimes g$. Note that, if $f, g \in \mathcal{H}^{\odot q}$, then $f \otimes_q g = \langle f, g \rangle_{\mathcal{H}^{\odot q}}$. In particular, if f, g are real-valued functions in $\mathcal{H}^{\otimes 2} = L^2(\mathbb{R}^2, \mathcal{B}^2, \mu^2)$ for a non-atomic measure μ, then we have

$$f \otimes_1 g = \int_{\mathbb{R}} f(s, t_1) g(s, t_2) \, \mu(ds). \tag{2.2}$$

Let $\mathcal{H}_q$ be the qth Wiener chaos of Z, that is, the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_q(Z(h)), h \in \mathcal{H}, \|h\|_{\mathcal{H}} = 1\}$, where $H_q(x)$ is the qth Hermite polynomial. It can be shown (see [6, Proposition 2.2.1]) that if $Z, Y \sim N(0,1)$ are jointly Gaussian, then

$$\mathbb{E}\left[H_p(Z) H_q(Y)\right] = \begin{cases} p! \, (\mathbb{E}[ZY])^p & \text{if } p = q \\ 0 & \text{otherwise} \end{cases} . \tag{2.3}$$

For $q \geq 1$, it is known that the map

$$I_q(h^{\otimes q}) = H_q(Z(h)) \tag{2.4}$$

provides a linear isometry between $\mathcal{H}^{\odot q}$ (equipped with the modified norm $\sqrt{q!}\|\cdot\|_{\mathcal{H}^{\otimes q}}$) and $\mathcal{H}_q$, where $I_q(\cdot)$ is the generalized Wiener-Itô stochastic integral (see [6, Theorem 2.7.7]). By convention, $\mathcal{H}_0 = \mathbb{R}$ and $I_0(x) = x$.

We use the following integral multiplication theorem from [7, Proposition 1.1.3]. Suppose $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$. Then

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q} r! \binom{p}{r}\binom{q}{r} I_{p+q-2r}(f\widetilde{\otimes}_r g), \tag{2.5}$$

where $f\widetilde{\otimes}_r g$ denotes the symmetrization of $f \otimes_r g$. For a product of more than two integrals, see Peccati and Taqqu [9].

2.2 Stochastic Integration and fBm

We refer to the 'time domain' and 'spectral domain' representations of fBm. The reader may refer to [10, 11] for details. Let $\mathcal{E}$ denote the set of real-valued step functions on $\mathbb{R}$. Let B^H denote fBm with Hurst parameter H. For this case, we view B^H as an isonormal Gaussian process on the Hilbert space $\mathfrak{H}$, which is the closure of $\mathcal{E}$ with respect to the inner product $\langle f, g\rangle_{\mathfrak{H}} = \mathbb{E}\left[I(f)I(g)\right]$. Consider also the inner product space

$$\tilde{\Lambda}_H = \left\{ f : f \in L^2(\mathbb{R}), \int_{\mathbb{R}} |\mathcal{F}f(\xi)|^2 |\xi|^{1-2H} d\xi < \infty \right\},$$

where $\mathcal{F}f = \int_{\mathbb{R}} f(x)e^{i\xi x}dx$ is the Fourier transform, and the inner product of $\tilde{\Lambda}_H$ is given by

$$\langle f, g\rangle_{\tilde{\Lambda}_H} = \frac{1}{C_H^2}\int_{\mathbb{R}} \mathcal{F}f(\xi)\overline{\mathcal{F}g(\xi)}|\xi|^{1-2H}d\xi, \tag{2.6}$$

where $C_H = \left(\frac{2\pi}{\Gamma(2H+1)\sin(\pi H)}\right)^{\frac{1}{2}}$. It is known (see [10, Theorem 3.1]) that the space $\tilde{\Lambda}_H$ is isometric to a subspace of $\mathfrak{H}$, and $\tilde{\Lambda}_H$ contains $\mathcal{E}$ as a dense subset. This inner product (2.6) is known as the 'spectral measure' of fBm. In the case $H \in (\frac{1}{2}, 1)$, there is another isometry from the space

$$|\Lambda_H| = \left\{ f : \int_0^\infty \int_0^\infty |f(u)||f(v)||u-v|^{2H-2}du\, dv < \infty \right\}$$

to a subspace of $\mathfrak{H}$, where the inner product is defined as

$$\langle f, g\rangle_{|\Lambda_H|} = H(2H-1)\int_0^\infty \int_0^\infty f(u)g(v)|u-v|^{2H-2}du\, dv,$$

see [10] or [7, Sect. 5.1].

3 Proof of Theorem 1

For any $\gamma > 0$ and $\lambda > 0$, we can write

$$\lambda^{-\gamma} = \frac{1}{\Gamma(\gamma)}\int_0^\infty y^{\gamma-1}e^{-\lambda y}dy,$$

where Γ is the Gamma function defined by $\Gamma(\gamma) = \int_0^\infty y^{\gamma-1}e^{-y}dy$. As a consequence, the covariance (1.7) can be written as

$$\mathbb{E}[Z_s Z_t] = \frac{1}{\Gamma(\gamma)}\int_0^\infty y^{\gamma-1}e^{-(t+s)y}dy. \tag{3.1}$$

Notice that this representation implies the covariance (1.7) is positive definite. Taking first the expectation with respect to the process Z, and using formula (3.1), we obtain

$$\begin{aligned} R(s,t) &= \frac{1}{\Gamma(\gamma)}\int_0^\infty \mathbb{E}\left[\left(\int_0^t e^{yu}dB_u^H\right)\left(\int_0^s e^{yu}dB_u^H\right)\right] y^{\gamma-1}e^{-(t+s)y}dy \\ &= \frac{1}{\Gamma(\gamma)}\int_0^\infty \left\langle e^{yu}\mathbf{1}_{[0,t]}(u), e^{yv}\mathbf{1}_{[0,s]}(v)\right\rangle_{\mathfrak{H}} y^{\gamma-1}e^{-(t+s)y}dy. \end{aligned}$$

Using the isometry between $\tilde{\Lambda}_H$ and a subspace of $\mathfrak{H}$ (see Sect. 2.2), we can write

$$\begin{aligned} \left\langle e^{yu}\mathbf{1}_{[0,t]}(u), e^{yv}\mathbf{1}_{[0,s]}(v)\right\rangle_{\mathfrak{H}} &= C_H^{-2}\int_{\mathbb{R}} |\xi|^{1-2H}(\mathcal{F}\mathbf{1}_{[0,t]}e^{y\cdot})(\overline{\mathcal{F}\mathbf{1}_{[0,s]}e^{y\cdot}})\, d\xi \\ &= C_H^{-2}\int_{\mathbb{R}} \frac{|\xi|^{1-2H}}{y^2+\xi^2}\left(e^{yt+i\xi t}-1\right)\left(e^{ys-i\xi s}-1\right)\, d\xi, \end{aligned}$$

where $(\mathcal{F}\mathbf{1}_{[0,t]}e^{x\cdot})$ denotes the Fourier transform and $C_H = \left(\frac{2\pi}{\Gamma(2H+1)\sin(\pi H)}\right)^{\frac{1}{2}}$. This allows us to write, making the change of variable $\xi = \eta y$,

$$R(s,t) = \frac{1}{\Gamma(\gamma)C_H^2}\int_0^\infty \int_{\mathbb{R}} y^{\gamma-1}\frac{|\xi|^{1-2H}}{y^2+\xi^2}\left(e^{i\xi t}-e^{-yt}\right)\left(e^{-i\xi s}-e^{-ys}\right)\, d\xi\, dy$$

$$= \frac{1}{\Gamma(\gamma)C_H^2} \int_0^\infty \int_{\mathbb{R}} y^{-\alpha-1} \frac{|\eta|^{1-2H}}{1+\eta^2} \left(e^{i\eta yt} - e^{-yt}\right) \left(e^{-i\eta ys} - e^{-ys}\right) \, d\eta \, dy, \tag{3.2}$$

where $\alpha = 2H - \gamma$. By Euler's identity, adding and subtracting 1 to compensate the singularity of $y^{-\alpha-1}$ at the origin, we can write

$$e^{i\eta yt} - e^{-yt} = (\cos(\eta yt) - 1 + i \sin(\eta yt)) + (1 - e^{-yt}). \tag{3.3}$$

Substituting (3.3) into (3.2) and taking into account that the integral of the imaginary part vanishes because it is an odd function, we obtain

$$\begin{aligned} R(s,t) = {} & \frac{2}{\Gamma(\gamma)C_H^2} \int_0^\infty \int_0^\infty y^{-\alpha-1} \frac{\eta^{1-2H}}{1+\eta^2} \Big((1-\cos(\eta yt))(1-\cos(\eta ys)) \\ & + \sin(\eta yt)\sin(\eta ys) + (\cos(\eta ys) - 1)(1 - e^{-yt}) + (\cos(\eta yt) - 1)(1 - e^{-ys}) \\ & + (1 - e^{-yt})(1 - e^{-ys}) \Big) \, d\eta \, dy. \end{aligned}$$

Let $B^{(j)} = \{B^{(j)}(\eta, t), \eta \geq 0, t \geq 0\}, j = 1, 2$ denote two independent Brownian sheets. That is, for $j = 1, 2$, $B^{(j)}$ is a continuous Gaussian field with mean zero and covariance given by

$$\mathbb{E}\left[B^{(j)}(\eta, t) B^{(j)}(\xi, s)\right] = \min(\eta, \xi) \times \min(t, s).$$

We define the following stochastic processes:

$$U_t = \frac{\sqrt{2}}{\sqrt{\Gamma(\gamma)}C_H} \int_0^\infty \int_0^\infty y^{-\frac{\alpha}{2}-\frac{1}{2}} \sqrt{\frac{\eta^{1-2H}}{1+\eta^2}} \, (\cos(\eta yt) - 1) \, B^{(1)}(d\eta, dy), \tag{3.4}$$

$$V_t = \frac{\sqrt{2}}{\sqrt{\Gamma(\gamma)}C_H} \int_0^\infty \int_0^\infty y^{-\frac{\alpha}{2}-\frac{1}{2}} \sqrt{\frac{\eta^{1-2H}}{1+\eta^2}} \, (\sin(\eta yt)) \, B^{(2)}(d\eta, dy), \tag{3.5}$$

$$Y_t = \frac{\sqrt{2}}{\sqrt{\Gamma(\gamma)}C_H} \int_0^\infty \int_0^\infty y^{-\frac{\alpha}{2}-\frac{1}{2}} \sqrt{\frac{\eta^{1-2H}}{1+\eta^2}} \left(1 - e^{-yt}\right) B^{(1)}(d\eta, dy), \tag{3.6}$$

where the integrals are Wiener-Itô integrals with respect to the Brownian sheet. We then define the stochastic process $X = \{X_t, t \geq 0\}$ by $X_t = U_t + V_t + Y_t$, and we have $\mathbb{E}[X_s X_t] = R(s,t)$ as given in (3.2). These processes have the following properties:

(I) The process $W_t = U_t + V_t$ is a fractional Brownian motion with Hurst parameter $\frac{\alpha}{2}$ scaled with the constant $\sqrt{\kappa}$. In fact, the covariance of this process is

$$\mathbb{E}[W_t W_s] = \frac{2}{\Gamma(\gamma)C_H^2} \int_0^\infty \int_0^\infty y^{-\alpha-1} \frac{\eta^{1-2H}}{1+\eta^2} \Big((\cos(\eta yt) - 1)(\cos(\eta ys) - 1)$$
$$+ \sin(\eta yt)\sin(\eta ys) \Big) d\eta dy$$
$$= \frac{1}{\Gamma(\gamma)C_H^2} \int_0^\infty \int_{\mathbb{R}} y^{\gamma-1} \frac{|\xi|^{1-2H}}{y^2+\xi^2} (e^{i\xi t} - 1)(e^{-i\xi s} - 1) d\xi dy.$$

Integrating in the variable y we finally obtain

$$\mathbb{E}[W_t W_s] = \frac{c_1}{\Gamma(\gamma)C_H^2} \int_{\mathbb{R}} \frac{(e^{i\xi t} - 1)(e^{-i\xi s} - 1)}{|\xi|^{\alpha+1}} d\xi,$$

where $c_1 = \int_0^\infty \frac{z^{\gamma-1}}{1+z^2} dz = \kappa\Gamma(\gamma)$. Taking into account the Fourier transform representation of fBm (see [11, p. 328]), this implies $\kappa^{-\frac{1}{2}} W$ is a fractional Brownian motion with Hurst parameter $\frac{\alpha}{2}$.

(II) The process Y coincides, up to a constant, with the process Y^K introduced in (1.4) with $K = 2\alpha + 1$. In fact, the covariance of this process is given by

$$\mathbb{E}[Y_t Y_s] = \frac{2c_2}{\Gamma(\gamma)C_H^2} \int_0^\infty y^{-\alpha-1} (1 - e^{-yt})(1 - e^{-ys}) dy, \tag{3.7}$$

where

$$c_2 = \int_0^\infty \frac{\eta^{1-2H}}{1+\eta^2} d\eta.$$

Notice that the process X is self-similar with exponent $\frac{\alpha}{2}$. This concludes the proof of Theorem 1.

4 Proof of Theorem 2

Along the proof, the symbol C denotes a generic, positive constant, which may change from line to line. The value of C will depend on parameters of the process and on T, but not on the increment width n^{-1}.

For integers $n \geq 1$, define a partition of $[0, \infty)$ composed of the intervals $\{[\frac{j}{n}, \frac{j+1}{n}), j \geq 0\}$. For the process X and related processes U, V, W, Y defined in Sect. 3, we introduce the notation

$$\Delta X_{\frac{j}{n}} = X_{\frac{j+1}{n}} - X_{\frac{j}{n}} \text{ and } \Delta X_0 = X_{\frac{1}{n}},$$

with corresponding notation for U, V, W, Y. We start the proof of Theorem 2 with two technical results about the components of the increments.

4.1 Preliminary Lemmas

Lemma 3 *Using above notation with integers $n \geq 2$ and $j, k \geq 0$, we have*

(a) $\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right] = \frac{\kappa}{2} n^{-\alpha} \left(|j-k-1|^{\alpha} - 2|j-k|^{\alpha} + |j-k-1|^{\alpha}\right)$, *where κ is defined in (1.8).*

(b) For $j + k \geq 1$,

$$\left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right| \leq C n^{-\alpha} (j+k)^{\alpha-2}$$

for a constant $C > 0$ that is independent of j, k and n.

Proof Property (a) is well-known for fractional Brownian motion. For (b), we have from (3.7):

$$\begin{aligned}\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right] &= \frac{2c_2}{\Gamma(\gamma) C_H^2 n^{\alpha}} \int_0^{\infty} y^{-\alpha-1} \left(e^{-yj} - e^{-y(j+1)}\right) \left(e^{-yk} - e^{-y(k+1)}\right) dy \\ &= \frac{2c_2}{\Gamma(\gamma) C_H^2 n^{\alpha}} \int_0^{\infty} y^{-\alpha+1} \int_0^1 \int_0^1 e^{-y(j+k+u+v)} du\, dv\, dy.\end{aligned}$$

Note that the above integral is nonnegative, and we can bound this with

$$\begin{aligned}\left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right| &\leq C n^{-\alpha} \int_0^{\infty} y^{-\alpha+1} e^{-y(j+k)}\, dy \\ &= C n^{-\alpha} (j+k)^{\alpha-2} \int_0^{\infty} u^{-\alpha+1} e^{-u} du \\ &\leq C n^{-\alpha} (j+k)^{\alpha-2}.\end{aligned}$$

□

Lemma 4 *For $n \geq 2$ fixed and integers $j, k \geq 1$,*

$$\left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right| \leq C n^{-\alpha} j^{2H-2} k^{-\gamma}$$

for a constant $C > 0$ that is independent of j, k and n.

Proof From (3.4)–(3.6) in the proof of Theorem 1, observe that

$$\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right] = \mathbb{E}\left[(\Delta U_{\frac{j}{n}} + \Delta V_{\frac{j}{n}}) \Delta Y_{\frac{k}{n}}\right] = \mathbb{E}\left[\Delta U_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right].$$

Assume $s, t > 0$. By self-similarity we can define the covariance function ψ by $\mathbb{E}[U_t Y_s] = s^\alpha \mathbb{E}\left[U_{t/s} Y_1\right] = s^\alpha \psi(t/s)$, where, using the change-of-variable $\theta = \eta x$,

$$\psi(x) = \int_0^\infty \int_0^\infty y^{-\alpha-1} \frac{\eta^{1-2H}}{1+\eta^2} (\cos(y\eta x) - 1)(1 - e^{-y})\, d\eta\, dy$$
$$= \int_0^\infty y^{-\alpha-1}(1-e^{-y}) \int_0^\infty \frac{\theta^{1-2H} x^{2H}}{x^2+\theta^2} (\cos(y\theta) - 1)\ d\theta\ dy.$$

Then using the fact that

$$\left| \frac{\theta^{1-2H} x^{2H}}{x^2+\theta^2} \right| \le |\theta^{-2H}|\ |x|^{2H-1}, \tag{4.1}$$

we see that $|\psi(x)| \le Cx^{2H-1}$, and

$$\psi'(x) = 2H \int_0^\infty y^{-\alpha-1}(1-e^{-y}) \int_0^\infty \frac{\theta^{1-2H} x^{2H-1}}{x^2+\theta^2} (\cos(y\theta) - 1)\ d\theta\ dy$$
$$- 2\int_0^\infty y^{-\alpha-1}(1-e^{-y}) \int_0^\infty \frac{\theta^{1-2H} x^{2H+1}}{(x^2+\theta^2)^2} (\cos(y\theta) - 1)\ d\theta\ dy.$$

Using (4.1) and similarly

$$\left| \frac{\theta^{1-2H} x^{2H+1}}{(x^2+\theta^2)^2} \right| \le |\theta^{-2H}|\ |x|^{2H-2}, \tag{4.2}$$

we can write

$$\left|\psi'(x)\right| \le x^{2H-2}|2H-2| \int_0^\infty y^{-\alpha-1}(1-e^{-y}) \int_0^\infty \theta^{-2H} (\cos(y\theta) - 1)\ d\theta\ dy \le Cx^{2H-2}.$$

By continuing the computation, we can find that $|\psi''(x)| \le Cx^{2H-3}$. We have for $j, k \ge 1$,

$$\mathbb{E}\left[\Delta U_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right] = n^{-\alpha}(k+1)^\alpha \left(\psi\left(\frac{j+1}{k+1}\right) - \psi\left(\frac{j}{k+1}\right)\right)$$
$$- n^{-\alpha} k^\alpha \left(\psi\left(\frac{j+1}{k}\right) - \psi\left(\frac{j}{k}\right)\right)$$
$$= n^{-\alpha}\left((k+1)^\alpha - k^\alpha\right)\left(\psi\left(\frac{j+1}{k+1}\right) - \psi\left(\frac{j}{k+1}\right)\right)$$
$$+ n^{-\alpha} k^\alpha \left(\psi\left(\frac{j+1}{k+1}\right) - \psi\left(\frac{j}{k+1}\right) - \psi\left(\frac{j+1}{k}\right) + \psi\left(\frac{j}{k}\right)\right).$$

With the above bounds on ψ and its derivatives, the first term is bounded by

$$
\begin{aligned}
n^{-\alpha}\left|(k+1)^{\alpha}-k^{\alpha}\right|&\left|\psi\left(\frac{j+1}{k+1}\right)-\psi\left(\frac{j}{k+1}\right)\right| \\
&\le \alpha n^{-\alpha}\int_0^1 (k+u)^{\alpha-1}du\int_0^{\frac{1}{k+1}}\left|\psi'\left(\frac{j}{k+1}+v\right)\right|dv \\
&\le Cn^{-\alpha}k^{\alpha-2}\left(\frac{j}{k}\right)^{2H-2}\le Cn^{-\alpha}k^{-\gamma}j^{2H-2},
\end{aligned}
$$

and

$$
\begin{aligned}
n^{-\alpha}k^{\alpha}&\left|\psi\left(\frac{j+1}{k+1}\right)-\psi\left(\frac{j}{k+1}\right)-\psi\left(\frac{j+1}{k}\right)+\psi\left(\frac{j}{k}\right)\right| \\
&= n^{-\alpha}k^{\alpha}\left|\int_0^{\frac{1}{k+1}}\psi'\left(\frac{j}{k+1}+u\right)du-\int_0^{\frac{1}{k}}\psi'\left(\frac{j}{k}+u\right)du\right| \\
&\le n^{-\alpha}k^{\alpha}\int_{\frac{1}{k+1}}^{\frac{1}{k}}\left|\psi'\left(\frac{j}{k}+u\right)\right|du+\int_0^{\frac{1}{k+1}}\int_{\frac{j}{k+1}}^{\frac{j}{k}}\left|\psi''(u+v)\right|dv\,du \\
&\le Cn^{-\alpha}k^{\alpha-2}\left(\frac{j}{k}\right)^{2H-2}+Cn^{-\alpha}k^{\alpha-3}j\left(\frac{j}{k}\right)^{2H-3}\le Cn^{-\alpha}k^{-\gamma}j^{2H-2}.
\end{aligned}
$$

This concludes the proof of the lemma. □

4.2 Proof of Theorem 2

We will make use of the notation $\beta_{j,n}=\left\|\Delta X_{\frac{j}{n}}\right\|_{L^2(\Omega)}$. We have for integer $j\ge 1$,

$$
\beta_{j,n}^2=\mathbb{E}\left[\Delta W_{\frac{j}{n}}^2\right]+\mathbb{E}\left[\Delta Y_{\frac{j}{n}}^2\right]+2\mathbb{E}\left[\Delta W_{\frac{j}{n}}\Delta Y_{\frac{j}{n}}\right]=\kappa n^{-\alpha}(1+\theta_{j,n}),
$$

where

$$
\kappa n^{-\alpha}\theta_{j,n}=\mathbb{E}\left[\Delta Y_{\frac{j}{n}}^2\right]+2\mathbb{E}\left[\Delta W_{\frac{j}{n}}\Delta Y_{\frac{j}{n}}\right].
$$

It follows from Lemmas 3 and 4 that $|\theta_{j,n}|\le Cj^{\alpha-2}$ for some constant $C>0$. Notice that, in the definition of $F_n(t)$, it suffices to consider the sum for $j\ge n_0$ for a fixed n_0. Then, we can choose n_0 in such a way that $Cn_0^{\alpha-2}\le\frac{1}{2}$, which implies

$$
\beta_{j,n}^2\ge\kappa n^{-\alpha}(1-Cj^{\alpha-2}) \tag{4.3}
$$

for any $j \geq n_0$.

By (2.4),

$$\beta_{j,n}^{q} H_q\left(\beta_{j,n}^{-1} \Delta X_{\frac{j}{n}}\right) = I_q^X\left(\left(\mathbf{1}_{[\frac{j}{n},\frac{j+1}{n})}\right)^{\otimes q}\right),$$

where I_q^X denotes the multiple stochastic integral of order q with respect to the process X. Thus, we can write

$$F_n(t) = n^{-\frac{1}{2}} \sum_{j=n_0}^{\lfloor nt\rfloor -1} \beta_{j,n}^{-q} I_q^X\left(\mathbf{1}_{\left[\frac{j}{n},\frac{j+1}{n}\right)}^{\otimes q}\right).$$

The decomposition $X = W + Y$ leads to

$$I_q^X\left(\mathbf{1}_{\left[\frac{j}{n},\frac{j+1}{n}\right)}^{\otimes q}\right) = \sum_{r=0}^{q} \binom{q}{r} I_r^W\left(\mathbf{1}_{\left[\frac{j}{n},\frac{j+1}{n}\right)}^{\otimes r}\right) I_{q-r}^Y\left(\mathbf{1}_{\left[\frac{j}{n},\frac{j+1}{n}\right)}^{\otimes q-r}\right).$$

We are going to show that the terms with $r = 0, \ldots, q-1$ do not contribute to the limit. Define

$$G_n(t) = n^{-\frac{1}{2}} \sum_{j=n_0}^{\lfloor nt\rfloor -1} \beta_{j,n}^{-q} I_q^W\left(\mathbf{1}_{\left[\frac{j}{n},\frac{j+1}{n}\right)}^{\otimes q}\right)$$

and

$$\widetilde{G}_n(t) = n^{-\frac{1}{2}} \sum_{j=n_0}^{\lfloor nt\rfloor -1} \left\| \Delta W_{j/n} \right\|_{L^2(\Omega)}^{-q} I_q^W\left(\mathbf{1}_{\left[\frac{j}{n},\frac{j+1}{n}\right)}^{\otimes q}\right).$$

Consider the decomposition

$$F_n(t) = (F_n(t) - G_n(t)) + (G_n(t) - \widetilde{G}_n(t)) + \widetilde{G}_n(t).$$

Notice that all these processes vanish at $t = 0$. We claim that for any $0 \leq s < t \leq T$, we have

$$\mathbb{E}[|F_n(t) - G_n(t) - (F_n(s) - G_n(s))|^2] \leq \frac{(\lfloor nt\rfloor - \lfloor ns\rfloor)^{\delta}}{n} \tag{4.4}$$

and

$$\mathbb{E}[|G_n(t) - \widetilde{G}_n(t) - (G_n(s) - \widetilde{G}_n(s))|^2] \leq \frac{(\lfloor nt\rfloor - \lfloor ns\rfloor)^{\delta}}{n}, \tag{4.5}$$

where $0 \le \delta < 1$. By Lemma 3, $\left\|\Delta W_{j/n}\right\|^2_{L^2(\Omega)} = \kappa n^{-\alpha}$ for every j. As a consequence, using (2.4) we can also write

$$\widetilde{G}_n(t) = n^{-\frac{1}{2}} \sum_{j=n_0}^{\lfloor nt \rfloor -1} H_q\left(\kappa^{-\frac{1}{2}} n^{\frac{\alpha}{2}} \Delta W_{\frac{j}{n}}\right).$$

Since $\kappa^{-\frac{1}{2}} W$ is a fractional Brownian motion, the Breuer-Major theorem implies that the process $\widetilde{G}$ converges in $D([0,T])$ to a scaled Brownian motion $\{\sigma B_t, t \in [0,T]\}$, where σ^2 is given in (1.11). By the fact that all the p-norms are equivalent on a fixed Wiener chaos, the estimates (4.4) and (4.5) lead to

$$\mathbb{E}[|F_n(t) - G_n(t) - (F_n(s) - G_n(s))|^{2p}] \le \frac{(\lfloor nt \rfloor - \lfloor ns \rfloor)^{\delta p}}{n^p} \tag{4.6}$$

and

$$\mathbb{E}[|G_n(t) - \widetilde{G}_n(t) - (G_n(s) - \widetilde{G}_n(s))|^{2p}] \le \frac{(\lfloor nt \rfloor - \lfloor ns \rfloor)^{\delta p}}{n^p}, \tag{4.7}$$

for all $p \ge 1$. Letting n tend to infinity, we deduce from (4.6) and (4.7) that for any $t \in [0,T]$ the sequences $F_n(t) - G_n(t)$ and $G_n(t) - \widetilde{G}_n(t)$ converge to zero in $L^{2p}(\Omega)$ for any $p \ge 1$. This implies that the finite dimensional distributions of the processes $F_n - G_n$ and $G_n - \widetilde{G}_n$ converge to zero in law. Moreover, by Billingsley [2, Theorem 13.5], (4.6) and (4.7) also imply that the sequences $F_n - G_n$ and $G_n - \widetilde{G}_n$ are tight in $D([0,T])$. Therefore, these sequences converge to zero in the topology of $D([0,T])$.

Proof of (4.4) We can write

$$\mathbb{E}\left[|F_n(t) - G_n(t) - (F_n(s) - G_n(s))|^2\right] \le C \sum_{r=0}^{q-1} \mathbb{E}[\Phi^2_{r,n}],$$

where

$$\Phi_{r,n} = n^{-\frac{1}{2}} \sum_{j=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} \beta_{j,n}^{-q} I_r^W\left(\mathbf{1}^{\otimes r}_{\left[\frac{j}{n}, \frac{j+1}{n}\right)}\right) I_{q-r}^Y\left(\mathbf{1}^{\otimes q-r}_{\left[\frac{j}{n}, \frac{j+1}{n}\right)}\right).$$

We have, using (4.3),

$$\mathbb{E}[\Phi^2_{r,n}] \le n^{-1+q\alpha} \times \sum_{j,k=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} \left|\mathbb{E}\left[I_r^W\left(\mathbf{1}^{\otimes r}_{\left[\frac{j}{n}, \frac{j+1}{n}\right)}\right) I_{q-r}^Y\left(\mathbf{1}^{\otimes q-r}_{\left[\frac{j}{n}, \frac{j+1}{n}\right)}\right) I_r^W\left(\mathbf{1}^{\otimes r}_{\left[\frac{k}{n}, \frac{k+1}{n}\right)}\right) I_{q-r}^Y\left(\mathbf{1}^{\otimes q-r}_{\left[\frac{k}{n}, \frac{k+1}{n}\right)}\right)\right]\right|.$$

Using a diagram method for the expectation of four stochastic integrals (see [9]), we find that, for any j, k, the above expectation consists of a sum of terms of the form

$$\left(\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right)^{a_1} \left(\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right)^{a_2} \left(\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right)^{a_3} \left(\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right)^{a_4},$$

where the a_i are nonnegative integers such that $a_1+a_2+a_3+a_4 = q$, $a_1 \leq r \leq q-1$, and $a_2 \leq q-r$. First, consider the case with $a_3 = a_4 = 0$, so that we have the sum

$$n^{-1+q\alpha} \sum_{j,k=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} \left(\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right)^{a_1} \left(\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right)^{q-a_1},$$

where $0 \leq a_1 \leq q-1$. Applying Lemma 3, we can control each of the terms in the above sum by

$$n^{-q\alpha}(|j-k+1|^{\alpha} - 2|j-k|^{\alpha} + |j-k-1|^{\alpha})^{a_1}(j+k)^{(q-a_1)(\alpha-2)},$$

which gives

$$\begin{aligned}
&n^{-1+q\alpha} \sum_{j,k=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{a_1} \left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right|^{q-a_1} \\
&\quad \leq Cn^{-1} \left(\sum_{j=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} j^{\alpha-2} + \sum_{j,k=\lfloor ns \rfloor \vee n_0, j\neq k}^{\lfloor nt \rfloor -1} |j-k|^{(q-1)(\alpha-2)}(j+k)^{\alpha-2} \right) \\
&\quad \leq Cn^{-1} \sum_{j=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} \left(j^{\alpha-2} + j^{q(\alpha-2)+1}\right) \\
&\quad \leq Cn^{-1}\left((\lfloor nt \rfloor - \lfloor ns \rfloor)^{(\alpha-1)\vee 0} + (\lfloor nt \rfloor - \lfloor ns \rfloor)^{[q(\alpha-2)+2]\vee 0}\right). \qquad (4.8)
\end{aligned}$$

Next, we consider the case where $a_3 + a_4 \geq 1$. By Lemma 3, we have that, up to a constant C,

$$\left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right| \leq C \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|,$$

so we may assume $a_2 = 0$, and have to handle the term

$$n^{-1+q\alpha} \sum_{j,k=\lfloor ns \rfloor \vee n_0}^{\lfloor nt \rfloor -1} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{q-a_3-a_4} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right|^{a_3} \left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{a_4} \qquad (4.9)$$

for all allowable values of a_3, a_4 with $a_3 + a_4 \geq 1$. Consider the decomposition

$$n^{-1+q\alpha} \sum_{j,k=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{q-a_3-a_4} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right|^{a_3} \left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{a_4}$$

$$= n^{q\alpha-1} \sum_{j=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} \left|\mathbb{E}\left[\Delta W^2_{\frac{j}{n}}\right]\right|^{q-a_3-a_4} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{j}{n}}\right]\right|^{a_3+a_4}$$

$$+ n^{q\alpha-1} \sum_{j=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} \sum_{k=\lfloor ns\rfloor \vee n_0}^{j-1} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{q-a_3-a_4} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right|^{a_3} \left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{a_4}$$

$$+ n^{q\alpha-1} \sum_{k=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} \sum_{j=\lfloor ns\rfloor \vee n_0}^{k-1} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{q-a_3-a_4} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right|^{a_3} \left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{a_4}.$$

We have,by Lemmas 3 and 4,

$$n^{-1+q\alpha} \sum_{j,k=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{q-a_3-a_4} \left|\mathbb{E}\left[\Delta W_{\frac{j}{n}} \Delta Y_{\frac{k}{n}}\right]\right|^{a_3} \left|\mathbb{E}\left[\Delta Y_{\frac{j}{n}} \Delta W_{\frac{k}{n}}\right]\right|^{a_4}$$

$$\leq Cn^{-1} \sum_{j=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} j^{(a_3+a_4)(\alpha-2)}$$

$$+ Cn^{-1} \sum_{j=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} j^{a_3(2H-2)-a_4\gamma} \sum_{k=\lfloor ns\rfloor \vee n_0}^{j-1} k^{-a_3\gamma+a_4(2H-2)} |j-k|^{(q-a_3-a_4)(\alpha-2)}$$

$$+ Cn^{-1} \sum_{k=\lfloor ns\rfloor \vee n_0}^{\lfloor nt\rfloor -1} k^{-a_3\gamma+a_4(2H-2)} \sum_{j=\lfloor ns\rfloor \vee n_0}^{k-1} j^{a_3(2H-2)-a_4\gamma} |k-j|^{(q-a_3-a_4)(\alpha-2)}$$

$$\leq Cn^{-1} \Big((\lfloor nt\rfloor - \lfloor ns\rfloor)^{[(a_3+a_4)(\alpha-2)+1]\vee 0} + (\lfloor nt\rfloor - \lfloor ns\rfloor)^{[q(\alpha-2)+2]\vee 0}$$
$$+ (\lfloor nt\rfloor - \lfloor ns\rfloor)^{[a_3(2H-2)-a_4\gamma+1]\vee 0} + (\lfloor nt\rfloor - \lfloor ns\rfloor)^{[a_4(2H-2)-a_3\gamma+1]\vee 0} \Big). \tag{4.10}$$

Then (4.8) and (4.10) imply (4.4) because $\alpha < 2 - \frac{1}{q}$.

Proof of (4.5) We have

$$G_n(t) - \widetilde{G}_n(t) = n^{-\frac{1}{2}} \sum_{j=n_0}^{\lfloor nt\rfloor -1} \left(\beta_{j,n}^{-q} - \left\| \Delta W_{\frac{j}{n}} \right\|_{L^2(\Omega)}^{-q} \right) I_q^W \left(\mathbf{1}^{\otimes q}_{[\frac{j}{n}, \frac{j+1}{n})} \right)$$

and we can write, using (4.3) for any $j \geq n_0$,

$$\left| \beta_{j,n}^{-q} - \left\| \Delta W_{\frac{j}{n}} \right\|_{L^2(\Omega)}^{-q} \right| = (\kappa^{-1} n^{\alpha})^{\frac{q}{2}} \left| (1+\theta_{j,n})^{-\frac{q}{2}} - 1 \right| \leq C \left(\kappa^{-1} n^{\alpha} j^{\alpha-2} \right)^{\frac{q}{2}}.$$

This leads to the estimate

$$\mathbb{E}\left[\left|G_n(t)-\widetilde{G}_n(t)-(G_n(s)-\widetilde{G}_n(s))\right|^2\right] \le Cn^{-1}$$
$$\times\left(\sum_{j=\lfloor ns\rfloor\vee n_0}^{\lfloor nt\rfloor-1} j^{\alpha-2}+\sum_{j,k=\lfloor ns\rfloor\vee n_0, j\neq k}^{\lfloor nt\rfloor-1}|j-k|^{q(\alpha-2)}\right)$$
$$\le Cn^{-1}\left((\lfloor nt\rfloor-\lfloor ns\rfloor)^{(\alpha-1)\vee 0}+(\lfloor nt\rfloor-\lfloor ns\rfloor)^{[q(\alpha-2)+2]\vee 0}\right),$$

which implies (4.5).

This concludes the proof of Theorem 2.

Acknowledgements D. Nualart is supported by NSF grant DMS1512891 and the ARO grant FED0070445

References

1. R.M. Balan, C.A. Tudor, The stochastic heat equation with fractional-colored noise: existence of the solution. Latin Am. J. Probab. Math. Stat. **4**, 57–87 (2008)
2. P. Billingsley, *Convergence of Probability Measures*, 2nd edn. (Wiley, New York, 1999)
3. P. Breuer, P. Major, Central limit theorems for nonlinear functionals of Gaussian fields. J. Multivar. Anal. **13**(3), 425–441 (1983)
4. C. Houdré, J. Villa, An example of infinite dimensional quasi-helix. Stoch. Models Contemp. Math. **366**, 195–201 (2003)
5. P. Lei, D. Nualart, A decomposition of the bifractional Brownian motion and some applications. Stat. Probab. Lett. **79**, 619–624 (2009)
6. I. Nourdin, G. Peccati, *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality* (Cambridge University Press, Cambridge, 2012)
7. D. Nualart, *The Malliavin Calculus and Related Topics*, 2nd edn. (Springer, Berlin, 2006)
8. H. Ouahhabi, C.A. Tudor, Additive functionals of the solution to the fractional stochastic heat equation. J. Fourier Anal. Appl. **19**(4), 777–791 (2012)
9. G. Peccati, M.S. Taqqu, *Wiener Chaos: Moments, Cumulants and Diagrams* (Springer, Berlin, 2010)
10. V. Pipiras, M.S. Taqqu, Integration questions related to fractional Brownian motion. Probab. Theory Relat. Fields **118** 251–291 (2000)
11. G. Samorodnitsky, M.S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman & Hall, New York, 1994)
12. S. Torres, C.A. Tudor, F. Viens, Quadratic variations for the fractional-colored stochastic heat equation. Electron. J. Probab. **19**(76), 1–51 (2014)
13. C.A. Tudor, *Analysis of Variations for Self-Similar Processes: A Stochastic Calculus Approach* (Springer, Berlin, 2013)
14. C.A. Tudor, Y. Xiao, Sample paths of the solution to the fractional-colored stochastic heat equation. Stochastics Dyn. **17**(1) (2017)

On A Priori Estimates for Rough PDEs

Qi Feng and Samy Tindel

Dedicated to Rodrigo Bañuelos on occasion of his 60th birthday

Abstract In this note, we present a new and simple method which allows to get a priori bounds on rough partial differential equations. The technique is based on a weak formulation of the equation and a rough version of Gronwall's lemma. The method is presented on a simple linear example, but might be generalized to a wide number of situations.

Keywords A priori estimate • Rough Gonwall lemma • Rough paths • Stochastic PDEs

1 Introduction

This paper proposes to review a recent method allowing to get a priori estimates for rough partial differential equations, taken from [6]. Our aim here is to show how to implement the technique on a simple example. Namely, we shall consider the following noisy heat equation on an interval $[0, \tau] \times \mathbb{R}^d$ for $\tau > 0$ and a spatial dimension $d \geq 1$:

$$\partial_t u_t(x) = \frac{\Delta}{2} u_t(x) + \sum_{i=1}^{\infty} \beta_i u_t(x) e_i(x) dw_t^i, \tag{1.1}$$

where Δ stands for the Laplace operator, $\{e_i; \geq 1\}$ is an orthonormal basis of $L^2(\mathbb{R}^d)$ and $\{\beta_i; \geq 1\}$ is a family of coefficients satisfying some summability conditions (see Hypothesis 2.4 below). In Eq. (1.1), $\{w_i; \geq 1\}$ is also a family of noises,

Q. Feng • S. Tindel (✉)
Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907, USA
e-mail: stindel@purdue.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_6

interpreted as p-variation paths with $p < 3$, which can be lifted to a rough path $\mathbf{w}$ (see Hypothesis 2.3 for a more complete definition).

The recent activity on existence and uniqueness results for rough PDEs has been thriving. A lot of this activity concerns situations which require renormalization techniques and a way to handle pathwise products of distributions [10, 12, 13]. Here we are concerned with a different context, for which the noise is smooth enough in space, so that the solution of (1.1) is directly expected to be a function and the integrals with respect to w are usual rough paths integrals. This situation does not require the whole regularity structure machinery, and one advantage of this reduced setting is that more information on the solution is available. We are concerned in this paper about a priori estimates, which can be either seen as a crucial step in the proof of existence of solutions, or as a first piece of valuable information about the solution. Furthermore, we believe that a priori estimates exhibit the core of the pathwise methods for stochastic PDEs, even though many more technical steps have to be performed in order to get existence and uniqueness results.

Let us summarize some of the (unrelated) approaches leading to estimates of equations like (1.1).

1. The references [2, 11] handle stochastic PDEs by considering random flows (induced by a finite dimensional rough path) which change the stochastic PDE into a deterministic PDE with random coefficients. A priori bounds are then potentially obtained by composing bounds on deterministic PDEs and estimates on rough flows. This possibility has not been fully exploited yet, and might lead to nontrivial considerations.
2. In [5, 9], a variant of the rough paths theory is introduced in order to cope with PDEs of the form (1.1), considered in the mild sense. This involves some lengthy and intricate considerations on twisted increments of the form $\hat{\delta} f_{ts} = f_t - S_{t-s} f_s$, where S designates the heat semi-group and f is a generic $L^2(\mathbb{R}^d)$-valued function. However, this formalism yields a priori estimates for (1.1), especially when one considers related numerical schemes as in [4].
3. For linear equations like (1.1), Feynman-Kac representations for the solution are available. This gives raise to explicit moment computations for $u_t(x)$, for a fixed couple $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Many cases of Gaussian noises have been examined in this context, and we refer to [3] for a situation which is close to ours, namely a rough noise in time which is smooth in space.

Let us highlight again the fact that we only recall here results concerning smooth noises in space. In cases like [10, 12, 13] where renormalization is needed, the mere existence of moments for the renormalized solution is still an open problem (to the best of our knowledge).

With these preliminary considerations in mind, the main point of the current paper is to show that the variational approach to rough PDEs, introduced in [1, 6], provides a handy way to obtain $L^2(\mathbb{R}^d)$ (and more generally $L^\alpha(\mathbb{R}^d)$) estimates on the solution. The main advantages of this new setting are the following:

1. The variational formulation is convenient at an algebraic and analytic level, when compared with the other methods mentioned above.

2. Unlike Feynman-Kac representations, the variational approach is not restricted to linear equations (though generalizations require a nontrivial extra work).

We shall illustrate this point of view with the simple model (1.1), for which we shall deduce L^α-estimates in a detailed way. It should be noticed that variational methods have been considered previously in [15] for pathwise PDEs driven by a fractional Brownian motion. With respect to this reference, our computations are restricted to linear cases. However, [15] only considers fBm's with a Hurst parameter $H > \frac{1}{2}$, while we are concerned with a true rough case (corresponding to $\frac{1}{3} < H \leq \frac{1}{2}$ for fBm).

Our article is structured as follows: in Sect. 2 we introduce some notations and the variational method framework, and we also present our first a priori estimate in Proposition 2.8. This estimate (adapted from [6, Theorem 2.5]) is valid for general linear equations, and will be suitable for our stochastic heat equation with multidimensional noise. Then in Sect. 3 we prove our main a priori bounds, namely Theorems 3.5 and 3.9 for the solution of Eq. (1.1), both in $L^2(\mathbb{R}^d)$ and $L^\alpha(\mathbb{R}^d)$ norms. Finally, Sect. 4 is devoted to the application of our abstract results to equations driven by fractional Brownian motion. A first example concerns a bounded domain, which enables us to compare our result with those of [15], while a second example deals with the whole space $\mathbb{R}^d$.

2 Rough Variational Framework

As mentioned above, our framework relies on a variational formulation of the heat equation, which is algebraically quite convenient. In this section we first recall some basic vocabulary about algebraic integration, then we give the main general results needed for the rough heat equation (1.1).

2.1 Notions of Algebraic Integration

First of all, let us recall the definition of the increment operator, denoted by δ. If g is a path defined on $[0, T]$ and $s, t \in [0, T]$ then we set $\delta g_{st} := g_t - g_s$. Whenever g is a 2-index map defined on $[0, T]^2$, we define $\delta g_{sut} := g_{st} - g_{su} - g_{ut}$. The norm of the element g in the Banach space E will be written as $\mathcal{N}[g; E]$. For two quantities a and b the relation $a \lesssim_x b$ means $a \leq c_x b$, for a constant c_x depending on a multidimensional parameter x.

In the sequel, given an interval I we call *control on I* (and denote it by ω) any continuous superadditive map on $\Delta_I := \{(s, t) \in I^2 : s \leq t\}$, that is, any continuous map $\omega : \Delta_I \to [0,\infty)$ such that, for all $s \le u \le t$,

$$\omega(s,u) + \omega(u,t) \le \omega(s,t).$$

Given a control ω on an interval $I = [a,b]$, we will use the notation $\omega(I) := \omega(a,b)$. For a fixed time interval I, a parameter $p > 0$, a Banach space E and any continuous function $g : I \to E$ we define the norm

$$\mathcal{N}[g; V_1^p(I;E)] := \sup_{(t_i)\in\mathcal{P}(I)} \left(\sum_i |\delta g_{t_i t_{i+1}}|^p \right)^{\frac{1}{p}},$$

where $\mathcal{P}(I)$ denotes the set of all partitions of the interval I. In this case,

$$\omega_g(s,t) = \mathcal{N}[g; V_1^p([s,t];E)]^p$$

defines a control on I. We denote by $V_2^p(I;E)$ the set of continuous two-index maps $g : I \times I \to E$ for which there exists a control ω such that

$$|g_{st}| \le \omega(s,t)^{\frac{1}{p}}$$

for all $s,t \in I$. We also define the space $V_{2,\text{loc}}^p(I;E)$ of maps $g : I \times I \to E$ such that there exists a countable covering $\{I_k\}_k$ of I satisfying $g \in V_2^p(I_k;E)$ for any k.

The following result is often referred to as *sewing lemma* in the literature, and is at the core of our approach to generalized integration.

Lemma 2.1 *Fix an interval I, a Banach space E and a parameter $\zeta > 1$. Consider a function $h : I^3 \to E$ such that $h \in \operatorname{Im}\delta$ and for every $s < u < t \in I$,*

$$|h_{sut}| \le \omega(s,t)^{\zeta}, \tag{2.1}$$

for some control ω on I. Then there exists a unique element $\Lambda h \in V_2^{\frac{1}{\zeta}}(I;E)$ such that $\delta(\Lambda h) = h$ and for every $s < t \in I$,

$$|(\Lambda h)_{st}| \le C_\zeta\, \omega(s,t)^{\zeta}, \tag{2.2}$$

for some universal constant C_ζ.

Our computations also hinge on the following rough version of Gronwall's lemma, borrowed from [6, Lemma 2.7].

Lemma 2.2 *Fix a time horizon $T > 0$ and let $Q : [0,T] \to [0,\infty)$ be a path such that for some constants $C, L > 0$, $\kappa \ge 1$ and some controls ω_1, ω_2 on $[0,T]$, one has*

$$\delta Q_{st} \le C\Big(\sup_{0\le r\le t} Q_r \Big)\, \omega_1(s,t)^{\frac{1}{\kappa}} + \omega_2(s,t), \tag{2.3}$$

for every $s < t \in [0, T]$ satisfying $\omega_1(s,t) \le L$. Then it holds

$$\sup_{0\le t\le T} Q_t \le 2\exp\left(c_{\kappa,L}\,\omega_1(0,T)\right)\cdot\left\{Q_0 + \sup_{0\le t\le T}\Big(\omega_2(0,t)\,\exp\left(-c_{\kappa,L}\,\omega_1(0,t)\right)\Big)\right\},$$

for a strictly positive constant $c_{\kappa,L}$.

2.2 Linear Equations with Distributional Drifts

In this section we shall first generalize Eq. (1.1), and consider the following:

$$dg_t = \mu(dt) + \sum_{i=1}^{\infty} \beta_i g_t e_i dw_t^i, \tag{2.4}$$

where μ is a distributional-valued measure. Before we give a rigorous meaning to this equation, let us label our hypothesis on the coefficients. We start by a rough path assumption for each couple of components of the driving noise w:

Hypothesis 2.3 *Let $p \in [2, 3)$ be given. We assume that the family $\{w^i;\ i \ge 1\}$ is such that there exist increments $\mathbf{w}^{\mathbf{1},i}, \mathbf{w}^{\mathbf{2},ij}$ satisfying the two following properties:*

(i) Algebraic condition: *For each $i, j \ge 1$ and $0 \le s \le u \le t \le \tau$, Chen's relation holds true:*

$$\delta\mathbf{w}_{st}^{\mathbf{1},i} = 0, \quad \text{and} \quad \delta\mathbf{w}_{sut}^{\mathbf{2},ij} = \mathbf{w}_{su}^{\mathbf{1},i}\,\mathbf{w}_{ut}^{\mathbf{1},j}. \tag{2.5}$$

(ii) Analytic condition: *For all $i, j \ge 1$, we have*

$$\mathcal{N}[\mathbf{w}^{\mathbf{1},i}; V_2^p([s,t])] < \infty, \quad \text{and} \quad \mathcal{N}[\mathbf{w}^{\mathbf{2},ij}; V_2^{p/2}([s,t])] < \infty.$$

The rough variational setting introduced in [1, 6] uses the concept of scale. A scale is defined as a sequence $\left(E_n, \|\cdot\|_n\right)_{n\in\mathbb{N}_0}$ of Banach spaces such that E_{n+1} is continuously embedded into E_n. Besides, for $n \in \mathbb{N}_0$ we denote by E_{-n} the topological dual of E_n. For the heat equation (1.1), we will consider the scale $E_n = W^{n,\infty}$.

Having the concept of scale in mind, the noise w should also fulfill the following hypothesis as an infinite dimensional object:

Hypothesis 2.4 *Recall that the scale E_n is given by $E_n = W^{n,\infty}$. We assume that $\{\beta_i;\ \ge 1\}$ is a family of positive coefficients satisfying $\sum_{i\ge 1}\beta_i < \infty$. Consider an orthonormal basis $\{e_i;\ \ge 1\}$ of $L^2(\mathbb{R}^d)$, composed of bounded functions. The noise w is such that $\{w_i;\ \ge 1\}$ is a family of p-variation paths with $p < 3$, whose first and* *second order increments* $\mathbf{w}^{\mathbf{1},i}$, $\mathbf{w}^{\mathbf{2},ij}$ *are such that* $\omega_{\mathbf{w}^1}$ *and* $\omega_{\mathbf{w}^2}$ *below are two controls on* $[0,\tau]$*:*

$$\omega_{\mathbf{w}^1}(s,t) \equiv \left(\sum_{i=1}^{\infty} \beta_i \left(1 + |e_i|_{E_1}\right) \mathcal{N}[\mathbf{w}^{\mathbf{1},i}; V_2^p([s,t])] \right)^p \tag{2.6}$$

$$\omega_{\mathbf{w}^2}(s,t) \equiv \left(\sum_{i,j=1}^{\infty} \beta_i \beta_j |e_i|_{E_1} |e_j|_{E_1} \mathcal{N}[\mathbf{w}^{\mathbf{2},ij}; V_2^{p/2}([s,t])] \right)^{p/2} . \tag{2.7}$$

We can now give a more formal definition of solution to our Eq. (2.4), in terms of expansions of the increments up to a regularity order greater than 1:

Definition 2.5 Let $p \in [2,3)$ and fix an interval $I \subseteq [0,\tau]$. Let μ be a distributional-valued measure lying in $V_1^1(I; E_{-1})$. A path $g : I \to E_{-0}$ is called solution (on I) of Eq. (2.4) provided there exists $q < 3$ and $g^\natural \in V_{2,\mathrm{loc}}^{\frac{q}{3}}(I, E_{-1})$ such that we have:

$$\delta g_{st}(\varphi) = \sum_{i=1}^{\infty} \beta_i g_s(e_i\varphi)\mathbf{w}_{st}^{\mathbf{1},i} + \delta\mu_{st}(\varphi) + \sum_{i,j=1}^{\infty} \beta_i\beta_j g_s(e_ie_j\varphi)\mathbf{w}_{st}^{\mathbf{2},ij} + g_{st}^\natural(\varphi), \tag{2.8}$$

for every $s,t \in I$ satisfying $s < t$ and every $\varphi \in E_1$.

Remark 2.6 On top of (2.5), we will use the following expressions for δg_{st}:

$$\delta g_{st}(\varphi) = \sum_{i=1}^{\infty} \beta_i g_s(e_i\varphi)\mathbf{w}_{st}^{\mathbf{1},i} + g^\sharp(\varphi), \tag{2.9}$$

where $g^\sharp$ is a $V_2^{\frac{p}{2}}(E_{-1})$ increment satisfying:

$$g_{st}^\sharp(\varphi) = \delta g_{st}(\varphi) - \sum_{i=1}^{\infty} \beta_i g_s(e_i\varphi)\mathbf{w}_{st}^{\mathbf{1},i} = \delta\mu_{st}(\varphi) + \sum_{i,j=1}^{\infty} \beta_i\beta_j g_s(e_ie_j\varphi)\mathbf{w}_{st}^{\mathbf{2},ij} + g_{st}^\natural(\varphi). \tag{2.10}$$

Remark 2.7 Equation (2.8) is expressed as an expansion along the increments of w^i. However, according to [7, Theorem 4.10], a solution u of (2.8) also solves the following integral equation (which has to be interpreted in the rough paths sense in time and weak sense in space):

$$\delta g_{st} = \mu\left([s,t)\right) + \sum_{i=1}^{\infty} \beta_i e_i \int_s^t g_r dw_r^i. \tag{2.11}$$

Furthermore, a change of variable formula (see [7, Proposition 5.6]) holds for g verifying (2.11). Namely, for $h \in \mathcal{C}^3(\mathbb{R})$ we have (still in the weak rough paths sense):

$$\delta h(g)_{st} = \int_s^t h'(g_r)\,\mu(dr) + \sum_{i=1}^{\infty} \beta_i e_i \int_s^t h'(g_r)\,g_r\,dw_r^i. \tag{2.12}$$

2.3 A General Estimate for Linear Equations

The following proposition gives our first a priori estimate for the solution to Eq. (2.4). It should be seen as an adaptation of [6, Theorem 2.5] to our current context.

Proposition 2.8 *Let $p \in [2,3)$ and fix an interval $I \subseteq [0,T]$. Let $\mathbf{w}$ be a rough path verifying Hypothesis 2.3 and 2.4. Consider a path $\mu \in V_1^1(I; E_{-1})$ such that for every $\varphi \in E_1$, there exists a control ω_μ verifying*

$$|\delta\mu_{st}(\varphi)| \leq \omega_\mu(s,t)\,\|\varphi\|_{E_1}. \tag{2.13}$$

Let g be a solution on I of Eq. (2.4), with the following additional hypothesis: g is controlled over the whole interval I, that is we have $g^\natural \in V_2^{\frac{q}{3}}(I; E_{-1})$ for $q < 3$. Moreover let $S_t^g = \sup_{s \leq t} \|g_s\|_{E_{-0}}$, and consider the following control:

$$\omega_I(s,t) \equiv \omega_\mu(s,t)\left(\omega_{\mathbf{w}^1}^{1/p}(s,t) + \omega_{\mathbf{w}^2}^{2/p}(s,t)\right) + S_t^g\left(2\omega_{\mathbf{w}^1}^{1/p}(s,t)\omega_{\mathbf{w}^2}^{2/p}(s,t) + \omega_{\mathbf{w}^2}^{4/p}(s,t)\right). \tag{2.14}$$

Then there exists a constant $L = L_p > 0$ (independent of I) such that if

$$\omega_{\mathbf{w}^1}(s,t) + \omega_{\mathbf{w}^2}^2(s,t) \leq L,$$

then for all $s,t \in I$ such that $s < t$, we have:

$$\|g_{st}^\natural\|_{E_{-1}} \lesssim_p \omega_I(s,t). \tag{2.15}$$

Proof Let $\omega_\natural(s,t)$ be a regular control such that $\|g_{st}^\natural\|_{E_{-1}} \leq \omega_\natural(s,t)^{\frac{3}{q}}$ for any $s,t \in I$ such that $s < t$. We divide this proof in several steps.

Step 1: An Algebraic Identity Let $\varphi \in E_1$ be such that $\|\varphi\|_{E_3} \leq 1$. We first show that

$$\delta g_{sut}^\natural(\varphi) = \sum_{i=1}^{\infty} \beta_i g_{su}^\sharp(e_i\varphi)\mathbf{w}_{ut}^{\mathbf{1},i} + \sum_{i,j=1}^{\infty} \beta_i\beta_j \delta g_{su}(e_ie_j\varphi)\mathbf{w}_{ut}^{\mathbf{2},ij} \equiv K_{sut}^1 + K_{sut}^2, \tag{2.16}$$

where $g^\sharp$ was defined in (2.10). Indeed, owing to (2.8), we have

$$g^\natural_{st} = \delta g_{st}(\varphi) - \sum_{i=1}^{\infty} \beta_i g_s(e_i\varphi)\mathbf{w}^{\mathbf{1},i}_{st} - \delta\mu_{st}(\varphi) - \sum_{i,j=1}^{\infty} \beta_i\beta_j g_s(e_ie_j\varphi)\mathbf{w}^{\mathbf{2},ij}_{st}.$$

Applying δ on both sides of this identity and recalling Chen's relations (2.5) as well as the fact that $\delta\delta = 0$ we thus get

$$\delta g^\natural_{sut}(\varphi) = \sum_{i=1}^{\infty} \beta_i \delta g_{su}(e_i\varphi)\mathbf{w}^{\mathbf{1},i}_{ut} + \sum_{i,j=1}^{\infty} \beta_i\beta_j \delta g_{su}(e_ie_j\varphi)\mathbf{w}^{\mathbf{2},ij}_{ut} - \sum_{i,j=1}^{\infty} \beta_i\beta_j g_s(e_ie_j\varphi)\mathbf{w}^{\mathbf{1},i}_{su}\mathbf{w}^{\mathbf{1},j}_{ut}.$$

Plugging relation (2.10) again into this identity, we end up with our claim (2.16).

Step 2: Bound for K^1 In order to bound the term $g^\sharp_{su}(e_i\varphi)$ in K^1, we invoke decomposition (2.10), which yields:

$$g^\sharp_{su}(e_i\varphi) = \delta\mu_{su}(e_i\varphi) + \sum_{j,k=1}^{\infty} \beta_j\beta_k g_s(e_ie_je_k\varphi)\mathbf{w}^{\mathbf{2},kl}_{su} + g^\natural_{su}(e_i\varphi),$$

and hence:

$$\begin{aligned} &|g^\sharp_{su}(e_i\varphi)| \\ &\le \left[\omega_\mu(s,t)|e_i|_{E_1} + S^g_u \sum_{j,k=1}^{\infty} \beta_j\beta_k |e_i|_{E_0}|e_j|_{E_0}|e_k|_{E_0}\omega^{2/p}_{\mathbf{w},jk}(s,u) + \omega^{3/p}_\natural(s,u)|e_i|_{E_1}\right]|\varphi|_{E_1}. \end{aligned}$$

Therefore, thanks to our assumption (2.7), we have:

$$|g^\sharp_{su}(e_i\varphi)| \le \left[\omega_\mu(s,u) + S^g_u\, \omega^{2/p}_{\mathbf{w}^2}(s,u) + \omega^{3/p}_\natural(s,u)\right] |e_i|_{E_1}|\varphi|_{E_1}. \tag{2.17}$$

Plugging this identity into the definition of K^1, we have thus obtained:

$$\begin{aligned} |K^1_{sut}| &\le |\varphi|_{E_1}\left[\omega_\mu(s,u) + S^g_u\, \omega^{2/p}_{\mathbf{w}^2}(s,u) + \omega^{3/p}_\natural(s,u)\right] \sum_{i=1}^{\infty} \beta_i |e_i|_{E_1}\omega^{1/p}_{\mathbf{w}^{1,i}}(u,t) \\ &\le |\varphi|_{E_1}\left[\omega_\mu(s,u) + S^g_u\, \omega^{2/p}_{\mathbf{w}^2}(s,u) + \omega^{3/p}_\natural(s,u)\right] \omega^{1/p}_{\mathbf{w}^1}(u,t). \end{aligned} \tag{2.18}$$

Step 3: Bound for K^2 and $\delta g^\natural$ The main term to treat for K^2 is the increment δg_{su}. To this aim, we resort to decomposition (2.9). This yields:

$$K^2_{sut} = \sum_{i,j,k=1}^{\infty} \beta_i\beta_j\beta_k\, g_s(e_ie_je_k\varphi)\, \mathbf{w}^{\mathbf{1},k}_{su}\, \mathbf{w}^{\mathbf{2},ij}_{ut} + \sum_{i,j=1}^{\infty} \beta_i\beta_j\, g^\sharp(e_ie_j\varphi)\, \mathbf{w}^{\mathbf{2},ij}_{ut} \equiv K^{21}_{sut} + K^{22}_{sut}.$$

Furthermore, we have:

$$\begin{aligned} |K^{21}_{sut}| &\le S^g_t |\varphi|_{E_0} \left(\sum_{k=1}^{\infty} \beta_k |e_k|_{E_0} \omega^{1/p}_{\mathbf{w}^{\mathbf{1},k}}(s,u)\right) \left(\sum_{i,j=1}^{\infty} \beta_i\beta_j |e_i|_{E_0} |e_j|_{E_0} \omega^{2/p}_{\mathbf{w}^{\mathbf{2},ij}}(u,t)\right) \\ &\le S^g_t |\varphi|_{E_0} \omega^{1/p}_{\mathbf{w}^1}(s,u)\, \omega^{2/p}_{\mathbf{w}^2}(u,t). \end{aligned}$$

In order to handle K^{22}, we elaborate slightly on our estimate (2.17) in order to get:

$$\begin{aligned} |K^{22}_{sut}| &\le |\varphi|_{E_1} \left[\omega_\mu(s,u) + S^g_u\, \omega^{2/p}_{\mathbf{w}^2}(s,u) + \omega^{3/p}_\natural(s,u)\right] \left(\sum_{i,j=1}^{\infty} \beta_i\beta_j |e_i|_{E_1} |e_j|_{E_1} \omega^{2/p}_{\mathbf{w}^{\mathbf{2},ij}}(u,t)\right) \\ &\le |\varphi|_{E_1} \left[\omega_\mu(s,u) + S^g_u\, \omega^{2/p}_{\mathbf{w}^2}(s,u) + \omega^{3/p}_\natural(s,u)\right] \omega^{2/p}_{\mathbf{w}^2}(u,t). \end{aligned}$$

Hence, gathering our estimates on K^{21} and K^{22} we end up with:

$$|K^2_{sut}| \le |\varphi|_{E_1} \left[S^g_t \left(\omega^{1/p}_{\mathbf{w}^1}(s,u) + \omega^{2/p}_{\mathbf{w}^2}(s,u)\right) + \omega_\mu(s,u) + \omega^{3/p}_\natural(s,u)\right] \omega^{2/p}_{\mathbf{w}^2}(u,t). \tag{2.19}$$

We can now easily conclude for the increment $\delta g^\natural$: plugging (2.18) and (2.19) into (2.16), we get:

$$\begin{aligned} \left|\delta g^\natural_{sut}(\varphi)\right| \le |\varphi|_{E_1} \Big\{ &\left(\omega_\mu(s,u) + S^g_u\, \omega^{2/p}_{\mathbf{w}^2}(s,u)\right)\, \omega^{1/p}_{\mathbf{w}^1}(u,t) \\ &+ \left(\omega_\mu(s,u) + S^g_t \left(\omega^{1/p}_{\mathbf{w}^1}(s,u) + \omega^{2/p}_{\mathbf{w}^2}(s,u)\right)\right) \omega^{2/p}_{\mathbf{w}^2}(u,t) \\ &+ \omega^{3/p}_\natural(s,u) \left(\omega^{1/p}_{\mathbf{w}^1}(u,t) + \omega^{2/p}_{\mathbf{w}^2}(u,t)\right) \Big\}. \end{aligned}$$

Otherwise stated, with our definition (2.14) in mind, we have obtained:

$$\left|\delta g^\natural_{sut}(\varphi)\right| \le |\varphi|_{E_1} \left\{\omega_I(s,t) + \omega^{3/p}_\natural(s,t) \left(\omega^{1/p}_{\mathbf{w}^1}(s,t) + \omega^{2/p}_{\mathbf{w}^2}(s,t)\right)\right\}. \tag{2.20}$$

Step 4: Conclusion It is readily checked, thanks to the fact that ω_μ, $\omega_{\mathbf{w}^1}$, $\omega_{\mathbf{w}^2}$ and $\omega_\natural$ are controls, plus [8, Exercise 1.9], that ω_I is a control as well as $\omega^{3/p}_\natural(s,t)(\omega^{1/p}_{\mathbf{w}^1}(s,t) + \omega^{2/p}_{\mathbf{w}^2}(s,t))$. One can thus apply Lemma 2.1 to relation (2.20) and get:

$$\left|g^{\natural}_{st}(\varphi)\right| \le c_p\,|\varphi|_{E_1}\left\{\omega_I(s,t) + \omega_{\natural}^{3/p}(s,t)\left(\omega_{\mathbf{w}^1}^{1/p}(s,t) + \omega_{\mathbf{w}^2}^{2/p}(s,t)\right)\right\}.$$

We now take I such that $c_p\,(\omega_{\mathbf{w}^1}^{1/p}(s,t) + \omega_{\mathbf{w}^2}^{2/p}(s,t)) \le \frac{1}{2}$. We obtain:

$$\|g^{\natural}_{st}\|_{E_{-1}} \le 2c_p\omega_I(s,t),$$

which ends our proof. □

Remark 2.9 In order to apply Proposition 2.8 to the heat equation (1.1), we shall consider a measure μ defined by $\mu([0,t]) = \int_0^t \Delta u_s\,ds$. It is worth noting that for a noisy equation like (1.1), we cannot assume that Δu_s is properly defined. This is why we consider $\mu([0,t])$ as an element of E_{-1} and perform our computations with distributional increments.

3 L^2 and L^α Type Estimates

Let us now go back to Eq. (1.1), for which we will derive some a priori estimates in $L^2(\mathbb{R}^d)$ and $L^\alpha(\mathbb{R}^d)$. We start by giving some basic properties of our linear heat equation.

3.1 Preliminary Considerations

Let us begin by giving a precise meaning to Eq. (1.1), as a particular case of rough PDE in the weak sense.

Definition 3.1 Let $\mathbf{w}$ be a rough path satisfying Hypothesis 2.3 and 2.4. Consider the following equation:

$$du_t(x) = \frac{1}{2}\Delta u_t(x) + \sum_{i=1}^{\infty} \beta_i u_t(x) e_i dw_t^i. \tag{3.1}$$

We interpret this system as in Definition 2.5, with a measure μ given by

$$\mu([s,t)) = \int_s^t \Delta u_r\,dr.$$

As mentioned in the introduction, we are only focusing here on a priori estimates for the heat equation, which are representative of the methods at stake without being too technical. To this aim, we label the following assumption, which prevails until the end of the article:

Hypothesis 3.2 *One can construct a path u on $[0,\tau]$ which solves (3.1) according to Definition 3.1. In addition, u can be obtained as a limit of a sequence of functions u^ε, where u^ε solves:*

$$du_t^\varepsilon(x) = \frac{1}{2}\Delta u_t^\varepsilon(x) + \sum_{i=1}^{\infty} \beta_i u_t^\varepsilon(x) e_i dw_t^{\varepsilon,i}. \tag{3.2}$$

In (3.2), the family $\{w_t^{\varepsilon,i};\ \varepsilon > 0, i \geq 1\}$ is a sequence of smooth functions converging to w. Recalling our notations (2.6) and (2.7), we also assume that:

$$\lim_{\varepsilon\to 0} \omega_{\mathbf{w}^1-\mathbf{w}^{1,\varepsilon}}(0,\tau) + \omega_{\mathbf{w}^2-\mathbf{w}^{2,\varepsilon}}(0,\tau) = 0.$$

Remark 3.3 Since we assume that u is obtained as a limit of smoothed paths u^ε (see Hypothesis 3.2), all the remaining computations have to be understood as follows: we first derive our relations for u^ε, and we then take limits as $\varepsilon \to 0$. This step will often be implicit for sake of conciseness.

With Hypothesis 3.2 in hand, we now derive the equation followed by the path u^2 as a first step towards L^2 estimates.

Proposition 3.4 *Let u be the solution of Eq. (3.1) alluded to in Hypothesis 3.2. We also set*

$$f_t = \|u_t\|_{L^2}^2 + \int_0^t \|\nabla u_r\|_{L^2} dr, \quad \text{and} \quad S_t^f = \sup_{s\leq t} f_s. \tag{3.3}$$

Then the following holds true:

(i) *Let μ^2 be the E_{-1}-valued measure defined as:*

$$\delta\mu_{st}^2(\psi) = -\int_s^t |\nabla u|^2(\psi)dr - \int_s^t (u_r\nabla u_r)(\nabla\psi)dr. \tag{3.4}$$

Then we have:

$$\omega_{\mu^2}(s,t) \leq \frac{3}{2}\int_s^t \|\nabla u\|_{L^2}^2 dr + \frac{1}{2}\int_s^t \|u\|_{L^2}^2 dr \leq \frac{3}{2}\int_s^t \|\nabla u\|_{L^2}^2 dr + \frac{(t-s)S_t^f}{2}, \tag{3.5}$$

provided the quantity above is finite.

(ii) *The squared path u^2 admits the following representation:*

$$\delta u_{st}^2(\psi) = \delta\mu_{st}^2(\psi) + \sum_{i=1}^{\infty} 2\beta_i u_s^2(e_i\psi)\mathbf{w}_{st}^{\mathbf{1},i} + \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} 4u_s^2(\psi e_i e_j)\beta_i\beta_j \mathbf{w}_{st}^{\mathbf{2},ij} + u_{st}^{2,\natural}(\psi), \tag{3.6}$$

where ψ is a generic test function, and where $u^{2,\natural}$ is an element of $V_2^{\frac{q}{3}}$ for a certain $q < 3$.

(iii) *The increment f satisfies the following relation: for $0 \le s < t \le \tau$ we have*

$$\delta f_{st} = 2\sum_{i=1}^{\infty} u_s^2(e_i)\beta_i \mathbf{w}_{st}^{\mathbf{1},i} + 4\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} u_s^2(e_i e_j)\beta_i\beta_j \mathbf{w}_{st}^{\mathbf{2},ij} + u_{st}^{2,\natural}(\mathbf{1}), \tag{3.7}$$

where $\mathbf{1}$ designates the function defined on $\mathbb{R}^d$ and identically equal to 1.

Proof With Remark 3.3 in mind, let us divide our proof in several steps.

Proof of (i) Similarly to [6, Remark 2.6], and working in the scale $E_n = W^{n,\infty}(\mathbb{R}^d)$, we have

$$|(\delta\mu^2)_{st}(\psi)| \le \int_s^t \|\nabla u\|_{L^2}^2 dr \|\psi\|_{L^\infty} + \left(\int_s^t \|\nabla u\|_{L^2}^2 dr\right)^{\frac{1}{2}} \left(\int_s^t \|u\|_{L^2}^2 dr\right)^{\frac{1}{2}} \|\psi\|_{W^{1,\infty}}, \tag{3.8}$$

Invoking now Young's inequality (namely $AB \le \frac{A^\alpha}{\alpha} + \frac{B^\beta}{\beta}$ for two positive numbers A, B with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$) we get our claim (3.5).

Proof of (ii) According to Definitions 2.5 and 3.1, the solution of Eq. (3.1) can be decomposed as:

$$\delta u_{st}(\psi) = \sum_{i=1}^{\infty} \beta_i u_s(e_i\psi)\mathbf{w}_{st}^{\mathbf{1},i} + \sum_{i,j=1}^{\infty} \beta_i\beta_i u_s(e_i e_j\psi)\mathbf{w}_{st}^{\mathbf{2},ij} + \delta\mu_{st}(\psi) + u_{st}^{\natural}(\psi). \tag{3.9}$$

As mentioned in Remark 2.7, u can also be seen as a solution to the integral equation (2.11), for which the change of variable formula (2.12) holds true. Applying this relation (written in its weak form) to $h(z) = z^2$, we obtain:

$$\delta u_{st}^2(\psi) = 2\int_s^t \Delta u_r(u_r\psi)\, dr + 2\sum_{i=1}^{\infty} \beta_i \int_s^t u_r^2(e_i\psi) dw_r^i,$$

so that an integration by parts in the first integral above yields:

$$\delta u_{st}^2(\psi) = -2\int_s^t |\nabla u|^2(\psi)\, dr - 2\int_s^t (u_r\nabla u_r)(\nabla\psi)\, dr + 2\sum_{i=1}^{\infty} \beta_i \int_s^t u_r^2(e_i\psi)\, dw_r^i. \tag{3.10}$$

We now expand the rough integral in (3.10) along the increments of w. We end up with relation (3.6), for a certain remainder $u^{2,\natural} \in V_2^{\frac{q}{3}}(E_{-1})$.

Proof of (ii) Relation (3.7) is simply obtained from (3.6) by considering a sequence of test functions $\{\psi_n; n \geq 1\}$ such that $\lim_{n\to\infty} \psi_n = \mathbf{1}$ and $\lim_{n\to\infty} \nabla\psi_n = 0$. □

3.2 A Priori Estimate in L^2

With Proposition 3.4 in hand, we can now derive the main estimate of this section.

Theorem 3.5 *Suppose w fulfills Hypothesis 2.3 and 2.4, and let u be the solution of Eq. (3.1) given in Hypothesis 3.2. For* $0 \leq s < t \leq \tau$, *set:*

$$\omega_1(s,t) = \omega_{\mathbf{w}^1}(s,t) + \omega^2_{\mathbf{w}^2}(s,t) + \omega_{\mathbf{w}^1}(s,t)\,\omega^2_{\mathbf{w}^2}(s,t) + \omega^4_{\mathbf{w}^2}(s,t). \tag{3.11}$$

Then the following L^2 *norm estimate for the solution u holds true:*

$$S^f_\tau = \sup_{0\leq t\leq\tau}\left(\|u_r\|^2_{L^2} + \int_0^t \|\nabla u_r\|^2_{L^2} dr\right) \leq 2\exp\left(c_p\omega_1(0,\tau)\right)\|u_0\|^2_{L^2}, \tag{3.12}$$

where c_p *is a strictly positive constant.*

Remark 3.6 Notice that $\|u_r\|^2_{L^2}$ and $\int_0^t \|\nabla u_r\|^2_{L^2} dr$ are positive. Therefore relation (3.12) implies that both terms are bounded from above.

Proof of Theorem 3.5 Recall that we have obtained the following decomposition in Proposition 3.4:

$$\delta u^2_{st}(\psi) = \delta\mu^2_{st}(\psi) + \sum_{i=1}^{\infty} 2\beta_i u^2_s(e_i\psi)\mathbf{w}^{\mathbf{1},i}_{st} + \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} 4u^2_s(\psi e_i e_j)\beta_i\beta_j \mathbf{w}^{\mathbf{2},ij}_{st} + u^{2,\natural}_{st}(\psi), \tag{3.13}$$

If we now set $g = u^2$ and $\mu^g = \mu^2$, we can recast (3.13) as:

$$\delta g_{st}(\psi) = \delta\mu^g_{st}(\psi) + \sum_{i=1}^{\infty} 2\beta_i\, g_s(e_i\psi)\mathbf{w}^{\mathbf{1},i}_{st} + \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} 4\, g_s(\psi e_i e_j)\beta_i\beta_j \mathbf{w}^{\mathbf{2},ij}_{st} + g^\natural_{st}(\psi).$$

This equation is of the same form as (2.8), and thus we can apply Proposition 2.8 directly. We get the following bound for $g^\natural_{st}$, which is valid whenever $\omega_1(s,t) + \omega^2_2(s,t) \leq L_p$ (recall that p is the regularity index of $\mathbf{w}$):

$$\|g^\natural_{st}\|_{E_{-1}} \leq c_p\omega_I(s,t), \quad \text{or equivalently} \quad \|u^{2,\natural}_{st}\|_{E_{-1}} \leq c_p\,\omega_I(s,t), \tag{3.14}$$

where the control ω_I is defined by:

$$\omega_I(s,t) \equiv \omega_{\mu^2}(s,t)\left(\omega_{\mathbf{w}^1}^{1/p}(s,t)+\omega_{\mathbf{w}^2}^{2/p}(s,t)\right)+S_t^{u^2}\left(2\omega_{\mathbf{w}^1}^{1/p}(s,t)\omega_{\mathbf{w}^2}^{2/p}(s,t)+\omega_{\mathbf{w}^2}^{4/p}(s,t)\right), \tag{3.15}$$

and where we recall that we have set:

$$S_t^{u^2}=\sup_{s\le t}|u_s^2|_{E_{-0}}=\sup_{s\le t}|u_s|_{L^2}^2.$$

Let us now go back to (3.13), and apply this relation to $\psi=\mathbf{1}$ (notice that the function $\mathbf{1}$ obviously sits in E_1). It is readily checked from (3.4) that:

$$\delta\mu_{st}^2(\mathbf{1})=-\int_s^t\|\nabla u\|_{L^2}^2dr,$$

and thus, with our notation (3.3) in mind, relation (3.13) becomes:

$$\delta f_{st}=\sum_{i=1}^{\infty}2\beta_i u_s^2(e_i)\mathbf{w}_{st}^{\mathbf{1},i}+\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}4u_s^2(e_ie_j)\beta_i\beta_j\mathbf{w}_{st}^{\mathbf{2},ij}+u_{st}^{2,\natural}(\mathbf{1})$$

Therefore, bounding $\|u_s^2\|_{E_{-0}}$ by S_t^f and invoking (3.14) in order to estimate $u_{st}^{2,\natural}(\mathbf{1})$, we obtain:

$$|\delta f_{st}|\le\left[2\omega_{\mathbf{w}^1}^{1/p}(s,t)+4\omega_{\mathbf{w}^2}^{2/p}(s,t)\right]S_t^f+c_p\,\omega_I(s,t), \tag{3.16}$$

where ω_I is given by (3.15). In order to close this expression, let us further bound the term ω_{μ^2} in the definition of ω_I. Namely, according to (3.5), we have

$$\omega_{\mu^2}(s,t)\le\frac{3}{2}\int_s^t\|\nabla u\|_{L^2}^2dr+\frac{(t-s)S_t^f}{2}\le c_\tau S_t^f, \tag{3.17}$$

where we recall that we are working on a time interval $[0,\tau]$. Plugging this inequality into the definition of ω_I, we end up with:

$$\omega_I(s,t)\le c_\tau S_t^f\left(\omega_{\mathbf{w}^1}^{1/p}(s,t)+\omega_{\mathbf{w}^2}^{2/p}(s,t)+\omega_{\mathbf{w}^1}^{1/p}(s,t)\omega_{\mathbf{w}^2}^{2/p}(s,t)+\omega_{\mathbf{w}^2}^{4/p}(s,t)\right).$$

Reporting the relation above into (3.16), we get

$$\begin{aligned}|\delta f_{st}|&\le c_\tau S_t^f\left(\omega_{\mathbf{w}^1}^{1/p}(s,t)+\omega_{\mathbf{w}^2}^{2/p}(s,t)+\omega_{\mathbf{w}^1}^{1/p}(s,t)\,\omega_{\mathbf{w}^2}^{2/p}(s,t)+\omega_{\mathbf{w}^2}^{4/p}(s,t)\right)\\&\le c_{\tau,p}S_t^f\,\omega_1(s,t),\end{aligned} \tag{3.18}$$

where ω_1 is the control introduced in (3.11). Recall again that inequality (3.18) is valid when $\omega_1(s,t)+\omega_2^2(s,t)\le L_p$. It is thus also satisfied when $\omega_1(s,t)\le L_p$.

We are now in a position to directly apply our rough Gronwall Lemma 2.2 to (3.18), with $Q=f$, $\kappa=1/p$ and $\omega_2=0$. It is readily checked that ω_1 is a control, and hence:

$$S_t^f \le 2\exp\Big(c_p\,\omega_1(0,\tau)\Big)f_0 = 2\exp\Big(c_p\,\omega_1(0,\tau)\Big)\|u_0\|_{L^2}^2, \tag{3.19}$$

which ends our proof. □

3.3 L^α Type Estimates

In this part, we are going to derive some L^α estimates for the solution of Eq. (3.1), generalizing the case $\alpha=2$. As the reader will notice, the method is the same as for the L^2 case, but we include some computational details for convenience.

Remark 3.7 We will handle the case of L^α estimates for an even integer α, in order to have $u^\alpha\ge 0$ and $u^{\alpha-2}\ge 0$ in the computations below. However, notice that other values of α can then be reached by simple interpolation methods.

We start this section with an analogue of Proposition 3.4.

Proposition 3.8 *Let u be the solution of Eq.* (3.1) *alluded to in Hypothesis 3.2, and consider an even integer α. We also set*

$$\ell_t = \|u_t\|_{L^\alpha}^\alpha + \int_0^t u_r^{\alpha-2}\|\nabla u_r\|^2 dr, \quad \textit{and} \quad S_t^\ell = \sup_{s\le t}\ell_s.$$

Then the following holds true:

(i) *Let μ^α be the E_{-1}-valued measure defined as:*

$$\delta\mu_{st}^\alpha(\psi) = -\frac{\alpha(\alpha-1)}{2}\int_s^t u_r^{\alpha-2}|\nabla u|^2(\psi)dr - \frac{\alpha}{2}\int_s^t (u_r^{\alpha-1}\nabla u_r)(\nabla\psi)dr \tag{3.20}$$

Then we have:

$$\omega_{\mu^\alpha}(s,t) \le \frac{\alpha(\alpha-1)}{4}\int_s^t u_r^{\alpha-2}|\nabla u_r|^2 dr + \frac{\alpha(t-s)S_t^\ell}{4}, \tag{3.21}$$

provided the quantity above is finite.

(ii) *The path u^α admits the following representation :*

$$\delta u_{st}^\alpha(\psi) = \delta\mu_{st}^\alpha(\psi) + \sum_{i=1}^\infty \alpha\beta_i u_s^\alpha(e_i\psi)\mathbf{w}_{st}^{\mathbf{1},i} + \sum_{j=1}^\infty\sum_{i=1}^\infty \alpha^2 u_s^\alpha(\psi e_i e_j)\beta_i\beta_j\mathbf{w}_{st}^{\mathbf{2},ij} + u_{st}^{\alpha,\natural}(\psi) \tag{3.22}$$

where ψ is a generic test function, and where $u^{\alpha,\natural}$ is an element of $V_2^{\frac{q}{3}}$ for a certain $q < 3$.

(iii) *The increment ℓ satisfies the following relation: for $0 \le s < t \le \tau$ we have*

$$\delta\ell_{st} = \alpha \sum_{i=1}^{\infty} u_s^{\alpha}(e_i)\beta_i \mathbf{w}_{st}^{\mathbf{1},i} + \alpha^2 \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} u_s^{\alpha}(e_i e_j)\beta_i\beta_j \mathbf{w}_{st}^{\mathbf{2},ij} + u_{st}^{\alpha,\natural}(\mathbf{1}), \tag{3.23}$$

where $\mathbf{1}$ designates the function defined on $\mathbb{R}^d$ and identically equal to 1.

Proof With Remark 3.3 in mind and μ^{α} defined in (3.21), it is readily checked that:

$$|(\delta\mu^{\alpha})_{st}(\psi)| \le \frac{\alpha(\alpha-1)}{2} \int_s^t u_r^{\alpha-2}|\nabla u_r|^2 dr \|\psi\|_{L^\infty} + \frac{\alpha}{2}\left(\int_s^t u_r^{\alpha-2}|\nabla u_r|^2 dr\right)^{1/2}\left(\int_s^t \|u_r\|_{L^\alpha}^{\alpha} dr\right)^{1/2} \|\psi\|_{W^{1,\infty}}. \tag{3.24}$$

Invoking now Young's inequality as we did in the previous L^2 case, we get our claim (3.21).

The proof of (3.22) is similar to the L^2 case, except that we apply the change of variable formula and relation (3.9) to $h(z) = z^{\alpha}$. We obtain:

$$\delta u_{st}^{\alpha}(\psi) = \alpha \int_s^t \Delta u_r(u_r^{\alpha-1}\psi)\, dr + \alpha \sum_{i=1}^{\infty} \beta_i \int_s^t u_r^{\alpha}(e_i\psi) dw_r^i,$$

so that an integration by parts in the first integral above yields:

$$\delta u_{st}^{\alpha}(\psi) = -\alpha(\alpha-1)\int_s^t u_r^{\alpha-2}|\nabla u|^2(\psi)\, dr - \alpha \int_s^t (u_r^{\alpha-1}\nabla u_r)(\nabla\psi)\, dr + \alpha \sum_{i=1}^{\infty} \beta_i \int_s^t u_r^{\alpha}(e_i\psi)\, dw_r^i. \tag{3.25}$$

We now expand the rough integral in (3.10) along the increments of w. We end up with relation (3.22), for a certain remainder $u^{\alpha,\natural} \in V_2^{q/3}(E_{-1})$.

As in the L^2 case, relation (3.23) is simply obtained from (3.22) by considering a sequence of test functions $\{\psi_n; n \ge 1\}$ such that $\lim_{n\to\infty} \psi_n = 1$. □

With Proposition 3.8 in hand, we can now derive the announced estimate in L^{α} type spaces.

Theorem 3.9 *Suppose w fulfills Hypothesis 2.3 and 2.4, and let u be the solution of Eq. (3.1) given in Hypothesis 3.2. For $0 \le s < t \le \tau$, set:*

$$\omega_1(s,t) = \omega_{\mathbf{w}^1}(s,t) + \omega_{\mathbf{w}^2}^2(s,t) + \omega_{\mathbf{w}^1}(s,t)\,\omega_{\mathbf{w}^2}^2(s,t) + \omega_{\mathbf{w}^2}^4(s,t). \tag{3.26}$$

Then for any even integer α, the following L^α norm estimate for the solution u holds true:

$$\sup_{0\le t\le\tau}\left(\|u_r\|_{L^\alpha}^\alpha+\int_0^t u_r^{\alpha-2}|\nabla u_r|^2dr\right)\le 2\exp\left(c_p\omega_1(0,\tau)\right)\|u_0\|_{L^\alpha}^\alpha, \tag{3.27}$$

where c_p is a strictly positive constant.

Proof Recall that we have obtained the following decomposition in Proposition 3.8:

$$\delta u_{st}^\alpha(\psi)=\delta\mu_{st}^\alpha(\psi)+\sum_{i=1}^\infty\alpha\beta_i u_s^\alpha(e_i\psi)\mathbf{w}_{st}^{\mathbf{1},i}+\sum_{j=1}^\infty\sum_{i=1}^\infty\alpha^2u_s^\alpha(\psi e_ie_j)\beta_i\beta_j\mathbf{w}_{st}^{\mathbf{2},ij}+u_{st}^{\alpha,\natural}(\psi). \tag{3.28}$$

If we now set $g=u^\alpha$ and $\mu^g=\mu^\alpha$, we can proceed as in Theorem 3.5 and recast (3.13) as:

$$\delta g_{st}(\psi)=\sum_{i=1}^\infty\alpha\beta_i g_s(e_i\varphi)\mathbf{w}_{st}^{\mathbf{1},i}+\delta\mu_{st}^g(\varphi)+\sum_{i,j=1}^\infty\alpha^2\beta_i\beta_j g_s(e_ie_j\varphi)\mathbf{w}_{st}^{\mathbf{2},ij}+g_{st}^\natural(\varphi),$$

This equation is of the same form as (2.8), and thus we can apply Proposition 2.8 directly. We get the following bound for $g_{st}^\natural$, which is valid whenever $\omega_1(s,t)+\omega_2^2(s,t)\le L_{p,\alpha}$:

$$\|g_{st}^\natural\|_{E_{-1}}\le c_p\omega_I(s,t),\quad\text{or equivalently}\quad \|u_{st}^{\alpha,\natural}\|_{E_{-1}}\le c_p\omega_I(s,t), \tag{3.29}$$

where the control ω_I is defined by:

$$\omega_I(s,t)\equiv\omega_{\mu^\alpha}(s,t)\left(\omega_{\mathbf{w}^1}^{1/p}(s,t)+\omega_{\mathbf{w}^2}^{2/p}(s,t)\right)+S_t^{u^\alpha}\left(2\omega_{\mathbf{w}^1}^{1/p}(s,t)\omega_{\mathbf{w}^2}^{2/p}(s,t)+\omega_{\mathbf{w}^2}^{4/p}(s,t)\right). \tag{3.30}$$

and where we recall that we have

$$S_t^{u^\alpha}=\sup_{s\le t}|u_s^\alpha|_{E_{-0}}=\sup_{s\le t}|u_s|_{L^\alpha}^\alpha.$$

Let us now go back to (3.13), and apply this relation to $\psi=\mathbf{1}$ (notice that the function $\mathbf{1}$ obviously sits in E_1). It is readily checked from (3.20) that:

$$\delta\mu_{st}^\alpha(\mathbf{1})=-\frac{\alpha(\alpha-1)}{2}\int_s^t u_r^{\alpha-2}|\nabla u|^2dr,$$

and thus (3.13) becomes:

$$\delta\ell_{st} = \sum_{i=1}^{\infty} \alpha\beta_i u_s^{\alpha}(e_i)\mathbf{w}_{st}^{\mathbf{1},i} + \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} \alpha^2 u_s^{\alpha}(e_i e_j)\beta_i\beta_j \mathbf{w}_{st}^{\mathbf{2},ij} + u_{st}^{\alpha,\natural}(\mathbf{1})$$

Therefore, bounding $\|u_s^{\alpha}\|_{E_{-0}}$ by S_t^{ℓ} and invoking (3.29) in order to estimate $u_{st}^{\alpha,\natural}(\mathbf{1})$, we obtain:

$$|\delta\ell_{st}| \le \left[2\omega_{\mathbf{w^1}}^{1/p}(s,t) + 4\omega_{\mathbf{w^2}}^{2/p}(s,t)\right] S_t^{\ell} + \omega_I(s,t), \tag{3.31}$$

where ω_I is given by (3.15). In order to close this expression, let us further bound the term $\omega_{\mu^{\alpha}}$ in the definition of ω_I. Namely, according to (3.21), we have

$$\omega_{\mu^{\alpha}}(s,t) \le \frac{\alpha(\alpha-1)}{4}\int_s^t u_r^{\alpha-2}|\nabla u_r|^2 dr + \frac{\alpha(t-s)S_t^{\ell}}{4} \le c_{\tau,p} S_t^{\ell},$$

which is the equivalent of relation (3.17) in our context. Starting from this point, we can conclude exactly as in Theorem 3.5. □

4 Application to Fractional Brownian Motion

This section is devoted to the application of our abstract results of Sect. 3 to some more concrete examples of heat equations driven by an infinite dimensional fractional Brownian motion. Though our general analysis was focused on equations in $\mathbb{R}^d$, we shall treat the case of both bounded and unbounded domains.

4.1 Equations in Bounded Domains

We first consider the case of an equation in a bounded domain D. This will enable us to compare our hypothesis with the assumptions contained in [15] for similar situations. Let us first label the conditions on our domain.

Hypothesis 4.1 *In this section, we consider an open, bounded domain D with smooth boundary ∂D, and satisfying the cone property.*

On such a domain D, we wish to give conditions which are close enough to the ones produced in [15]. This is why we consider an operator C given as follows:

Hypothesis 4.2 *In the remainder of the section, C will stand for a linear, self-adjoint, positive trace-class operator on $L^2(D)$. This operator admits an orthonormal basis $(e_i)_{i\in\mathbb{N}_+}$ of eigenfunctions, with corresponding eigenvalues $(\lambda_i)_{i\in\mathbb{N}_+}$. It* *also admits an integral representation, whose generating kernel is denoted as* κ*. Summarizing, for all* $i \geq 0$ *and for almost every* $x \in D$ *we have:*

$$Ce_i(x) = \int_D \kappa(x,y)e_i(y)\,dy = \lambda_i e_i(x). \tag{4.1}$$

We can now formulate our a priori estimate in this context:

Proposition 4.3 *Let* $D \subset \mathbb{R}^d$ *be a domain fulfilling Hypothesis 4.1, together with an operator* C *as in Hypothesis 4.2. On D, we consider the following equation:*

$$du_t(x) = \frac{1}{2}\Delta u_t(x) + \sum_{i=1}^{\infty} \lambda_i^{\nu} u_t(x)e_i(x)dB_t^i, \tag{4.2}$$

where $(B_t^i)_{t\in\mathbb{R}^+})_{i\in N^+}$ *is a sequence of one-dimensional, independent, identically distributed fractional Brownian motions with Hurst parameter* $H \in (\frac{1}{3}, 1)$*, and* $\nu \geq 0$ *is a positive parameter. For the definition of* e_i *and* λ_i*, we refer to Hypothesis 4.2. In addition, we suppose that our operator* C *and its kernel* κ *satisfy the following conditions:*

$$A_\kappa \equiv \sup_{x\in D} \|\kappa(x,\cdot)\|_{L^2(D)} + \|\nabla\kappa(x,\cdot)\|_{L^2(D)} < \infty, \quad \textit{and} \quad \sum_{i\geq 0} \lambda_i^{\nu-1} < \infty. \tag{4.3}$$

Then the results from Theorems 3.5 and 3.9 apply.

Proof It is well known (see e.g. [8, Chap. 15]) that any finite dimensional fractional Brownian motion $(B^i)_{i\leq N}$ can be lifted as a rough path. It is thus enough to prove conditions (2.6) and (2.7). We shall focus on condition (2.6), the other one being checked with the same kind of arguments.

In order to verify (2.6), similarly to [15], we start by recasting (4.1) as:

$$e_i(x) = \lambda_i^{-1} \int_D \kappa(x,y)e_i(y)\,dy, \quad \text{and} \quad \nabla e_i(x) = \lambda_i^{-1} \int_D \nabla\kappa(x,y)e_i(y)\,dy.$$

Hence, invoking Cauchy-Schwarz' inequality and relation (4.3), we obtain:

$$|e_i|_{E_1} \leq \lambda_i^{-1} A_\kappa \|e_i\|_{L^2(D)} = \lambda_i^{-1} A_\kappa. \tag{4.4}$$

Now notice that (2.6) is ensured by the condition $\mathbb{E}[\omega_{\mathbf{w}^1}^{1/p}(0,\tau)] < \infty$, where τ is our time horizon. Furthermore,

$$\mathbb{E}\left[\omega_{\mathbf{w}^1}^{1/p}(0,\tau)\right] = \sum_{i=1}^{\infty} \lambda_i^{\nu}\,(1 + |e_i|_{E_1})\,\mathbb{E}\left[\mathcal{N}[\mathbf{w}^{1,i}; V_2^p([s,t])]\right],$$

and since $\mathbb{E}[\mathcal{N}[\mathbf{w}^{\mathbf{1},i}; V_2^p([s,t])]]$ is uniformly bounded in i, we end up with

$$\mathbb{E}\left[\omega_{\mathbf{w}^{\mathbf{1}}}^{1/p}(0,\tau)\right] \leq c_{\kappa,\mathbf{w}} \sum_{i=1}^{\infty} \lambda_i^{\nu-1},$$

which is a finite quantity according to our assumption (4.3). In conclusion, Hypothesis 2.3 and 2.4 are satisfied, and Theorems 3.5 and 3.9 hold true. □

Remark 4.4 With respect to [15], we have added here the assumption

$$\sup_{x \in D} \|\nabla \kappa(x, \cdot)\|_{L^2(D)} < \infty,$$

which is an artifact of our variational approach. This being said, let us recall that our method applies to rough situations (compared to the case $H > 1/2$ dealt with in [15]). We also believe that our method extends to non linear equations, with a noisy term of the form $\sum_{i=1}^{\infty} \lambda_i^{\nu} \sigma(u_t(x)) e_i(x) dB_t^i$ for a smooth coefficient σ.

4.2 *Equations in* $\mathbb{R}^d$

On the whole space $\mathbb{R}^d$, choices of orthonormal basis of L^2 are wide. For sake of concreteness, we will stick here to a wavelet basis based on Shannon's wavelet, though a much more general setting can be found e.g. in [14].

Let us start by defining the L^2 basis alluded to above (we refer again to [14] for proofs of general facts on wavelets).

Lemma 4.5 *Let* $\psi : \mathbb{R} \to \mathbb{R}$ *be defined as*

$$\psi(x) = \frac{\sin 2\pi(x - 1/2)}{2\pi(x - 1/2)} - \frac{\sin \pi(x - 1/2)}{\pi(x - 1/2)}.$$

Then $\psi \in L^2(\mathbb{R})$, *and the following holds true:*

(i) *Let us introduce a family of scaled functions* $\{\psi_{j,k};\, j \geq 0,\, k \in \mathbb{Z}\}$ *by:*

$$\psi_{j,k}(x) = 2^{-\frac{j}{2}} \psi\left(\frac{x - 2^j k}{2^j}\right). \tag{4.5}$$

This family is an orthonormal basis of $L^2(\mathbb{R})$.

(ii) *One can obtain an orthonormal basis of* $L^2(\mathbb{R}^d)$ *by tensorizing the previous basis of* $L^2(\mathbb{R})$. *Namely, for all* $j \geq 0$ *and for* $n = (n_1, \cdots, n_d)$, *we denote*

$$\psi_{j,n}(x) = 2^{-dj/2} \psi\left(\frac{x_1 - 2^j n_1}{2^j}, \cdots, \frac{x_d - 2^j n_d}{2^j}\right).$$

Then $\{\psi_{j,n}(x)\}_{(j,n)\in\mathbb{Z}^{d+1}}$ *is an orthonormal basis of* $L^2(\mathbb{R}^d)$*. In addition, it is readily checked that:*

$$|\psi_{j,k}|_{E_1} \le 2^{\frac{jd}{2}}, \tag{4.6}$$

where we recall that we work in the scale $E_n = W^{n,\infty}(\mathbb{R})$*.*

Remark 4.6 A completely correct version of Lemma 4.5 should include a so-called father wavelet ϕ. We omit this step for notational sake.

Under the setting of Lemma 4.5, here is our example of stochastic heat equation on $\mathbb{R}^d$:

Proposition 4.7 *Consider the equation*

$$du_t(x) = \frac{1}{2}\Delta u_t(x) + \sum_{j=0}^{\infty}\sum_{n\in\mathbb{Z}^d} \beta_{j,n} u_t(x)\psi_{j,n}(x)dB_t^{j,n},$$

where $\{B^{j,n}; j \ge 0, n \in \mathbb{Z}^d\}$ *is a sequence of one-dimensional, independent, identically distributed fractional Brownian motions with Hurst parameter* $H \in (\frac{1}{3}, 1)$*, and* $\{\beta_{j,n}; j \ge 0, n \in \mathbb{Z}^d\}$ *is a family of positive coefficients. We assume that*

$$A_\beta \equiv \sum_{j=0}^{\infty}\sum_{n\in\mathbb{Z}^d} 2^{\frac{dj}{2}} \beta_{j,n} < \infty. \tag{4.7}$$

Then the results of Theorems 3.5 and 3.9 apply.

Proof We proceed as for Proposition 4.3, and we are easily reduced to show that $\mathbb{E}[\omega_{\mathbf{w}^1}^{1/p}(0,\tau)]$ is a finite quantity. In our case, we have

$$\mathbb{E}\left[\omega_{\mathbf{w}^1}^{1/p}(0,\tau)\right] = \sum_{j=0}^{\infty}\sum_{n\in\mathbb{Z}^d} \beta_{j,n}\left(1 + |\psi_{j,n}|_{E_1}\right)\mathbb{E}\left[\mathcal{N}[\mathbf{w}^{\mathbf{1}j,n}; V_2^p([s,t])]\right].$$

Moreover, the coefficients $\mathbb{E}[\mathcal{N}[\mathbf{w}^{\mathbf{1}j,n}; V_2^p([s,t])]]$ are uniformly bounded in j, n. Hence, owing to relation (4.6), we get:

$$\mathbb{E}\left[\omega_{\mathbf{w}^1}^{1/p}(0,\tau)\right] \le c_{\psi,\mathbf{w}} \sum_{j=0}^{\infty}\sum_{n\in\mathbb{Z}^d} 2^{\frac{dj}{2}} \beta_{j,n} = c_{\psi,\mathbf{w}} A_\beta,$$

where A_β is introduced in condition (4.7). This concludes our proof in a straightforward way. □

References

1. I. Bailleul, M. Gubinelli, Unbounded rough drivers (2015). arXiv:1501.02074
2. M. Caruana, P. Friz, H. Oberhauser, A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. Ann. Inst. H. Poincaré Anal. Non Linéaire **28**(1), 27–46 (2011)
3. L. Chen, Y. Hu, K. Kalbasi, D. Nualart, Intermittency for the stochastic heat equation driven by a rough time fractional Gaussian noise (2016). arXiv:1602.05617
4. A. Deya, A discrete approach to rough parabolic equations. Electron. J. Probab. **16**(54), 1489–1518 (2011)
5. A. Deya, M. Gubinelli, S. Tindel, Non-linear rough heat equations. Probab. Theory Relat. Fields **153**, 97–147 (2012)
6. A. Deya, M. Gubinelli, M. Hofmanova, S. Tindel, General a priori estimates for rough PDEs with applications to rough conservation laws (2016). arXiv:1604.00437
7. P. Friz, M. Hairer, *A Course on Rough Paths* (Springer, Berlin, 2014)
8. P. Friz, N. Victoir, *Multidimensional Dimensional Processes Seen as Rough Paths* (Cambridge University Press, Cambridge, 2010)
9. M. Gubinelli, S. Tindel, Rough evolution equations. Ann. Probab. **38**(1), 1–75 (2010)
10. M. Gubinelli, P. Imkeller, N. Perkowski, Paracontrolled distributions and singular PDEs. Forum Math. Pi **3**(e6), 75 (2015)
11. M. Gubinelli, S. Tindel, I. Torrecilla, Controlled viscosity solutions of fully nonlinear rough PDEs (2014). arXiv:1403.2832
12. M. Hairer, Solving the KPZ equation. Ann. Math. **178**(2), 559–664 (2013)
13. M. Hairer, A theory of regularity structures. Invent. Math. **198**(2), 269–504 (2014)
14. S. Mallat, *A Wavelet Tour of Signal Processing* (Academic, London, 1998)
15. D. Nualart, P.-A. Vuillermot, Variational solutions for partial differential equations driven by a fractional noise. J. Funct. Anal. **232**, 390–454 (2006)

Lévy Systems and Moment Formulas for Mixed Poisson Integrals

Krzysztof Bogdan, Jan Rosiński, Grzegorz Serafin, and Łukasz Wojciechowski

Abstract We propose Mecke-Palm formula for multiple integrals with respect to the Poisson random measure and its intensity measure performed, or mixed, in an arbitrary order. We apply the formulas to mixed Lévy systems of Lévy processes and obtain moment formulas for mixed Poisson integrals.

Keywords Lévy system • Poisson-Skorochod integral

2010 Mathematics Subject Classification 60G51, 60G57

1 Introduction

The Mecke-Palm formula is an important identity in stochastic analysis of Poisson random measures. In this work we propose its generalization named the (multiple) mixed-type Mecke-Palm formula. We show that the generalization is useful and has a considerable scope of applications.

Part of our motivation comes from recent results on moments of stochastic integrals [7, 21]. These were obtained for 1-processes in [21] by using combinatorics of the binomial convolution to undo the usual compensation in stochastic integration against Poisson random measures [11, 12]; and they were extended in [7, Theorem 3.1] to ensembles of integrals of 1-processes.

K. Bogdan (✉) • G. Serafin
Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, 50-370 Wrocław, Poland
e-mail: krzysztof.bogdan@pwr.edu.pl; grzegorz.serafin@pwr.edu.pl

J. Rosiński
Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA
e-mail: rosinski@math.utk.edu

Ł. Wojciechowski
Mathematical Institute, University of Wrocław, 50-384 Wrocław, Poland

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_7

By compensation in the previous paragraph we mean integration against the difference of the random Poisson measure and its intensity, or control, measure. It is well-known that such integration fits well into the framework of L^2 Hilbert spaces [19]. In opposition, the results of this paper mainly concern iterated integrations against the (uncompensated) Poisson random measure mixed with integrations against the control measure. Such integrations preserve nonnegativity and are performed under nonnegativity or absolute integrability conditions, rather than the square-integrability conditions (for which see Lemma 4.8 below or [19]). In both settings, however, the main feature of the iterated stochastic integration is the impact of the diagonals in the corresponding Cartesian products of the state space, which cannot be ignored because the random measure has atoms. The impact is accounted for by using partitions of the set of coordinates. We shall see below that in the setting of the uncompensated stochastic integration the description is simpler than in the compensated, or L^2, setting, for which we refer the reader to [19, Chap. 5]. In fact, the integrals against the compensated Poisson measure can be considered as (limits of) linear combinations of mixed integrals with respect to the Poisson random measure and its control measure, which explains the added complexity. Moreover, we may consider the results obtained in both settings as consequences of the mixed Mecke-Palm formula and the structure of the family of partitions. Our presentation is essentially self-contained in that it relies on the mixed Mecke-Palm formula, which we explain from the first principles. We should also remark that the integrands we consider are random, and in this respect they are more general than those in [19]. A complete survey of results on integration with respect to random measures is beyond the scope of this paper, but for more information we like to refer the reader to [13, 14, 17].

Below we first prove the mixed Mecke-Palm formula and use its along with the so-called linearization to obtain moments of stochastic integrals in more generality than known before: we consider moments of k-processes with arbitrary integer $k \geq 1$, and we allow Poisson stochastic integrations to be mixed, up to arbitrary multiplicity and order, with integrations against the intensity measure of the Poisson random measure. Our proofs are more direct as compared to [21] and [7], because they easily follow from the mixed Mecke-Palm formula.

When the random measure is given by the jumps of a Lévy process, the mixed Mecke-Palm formula translates into multiple Lévy systems of mixed type, which is our second main application. By the multiple Lévy systems of mixed type we mean identities for expectations of functionals defined by accumulated summations indexed by the jumps of the Lévy process and integrations against the product of the linear Lebesgue measure on the time scale and the Lévy measure of the process in space. They generalize the classical (single) Lévy system [3, 4, 8], which is an important tool in the study of jump-type Markov processes. The multiple variants have interesting applications and we indicate some of them.

The structure of the paper is as follows. In Theorem 2.4 of Sect. 2 we give the mixed Mecke-Palm formula for k-processes. In Theorem 3.1 of Sect. 3 we derive general moment formulas for ensembles of k-processes. These are illustrated by the moments formulas for 1-processes and 2-processes in Sects. 3.2 and 3.3, respectively. In Theorem 4.3 of Sect. 4 we present the multiple mixed Lévy systems for Lévy processes in $\mathbb{R}^d$. In Sect. 4.2 we present several applications of the Lévy systems including applications that merge the topics and techniques from Sects. 3 and 4. Some of the results are known, but even then the presentation may be of interest. In Sect. 4.2 we give a proof of the simple Mecke-Palm formula, to make the paper more self-contained.

2 Mixed Mecke-Palm Formulas

A direct approach to calculus of Poisson random measures is based on the configuration space: Given a locally compact separable metric space $\mathbb{X}$, any locally finite subset of $\mathbb{X}$ is called a configuration on $\mathbb{X}$. The configuration space is defined as $\Omega = \Omega_{\mathbb{X}} = \{\omega \subset \mathbb{X} : \omega \text{ is a configuration on } \mathbb{X}\}$ [20]. The elements of Ω can be identified with the class of locally finite, nonnegative-integer valued measures: if $\omega = \{y_1, y_2, \ldots\}$, where $y_i \in \mathbb{X}$ are all different, then we also write

$$\omega = \sum_i \delta_{y_i},$$

where δ_y is the probability measure concentrated at $y \in \mathbb{X}$. According to this identification, ω will have two meanings depending on the context: a configuration on $\mathbb{X}$ or a measure on $\mathbb{X}$. We equip Ω with a σ-algebra $\mathcal{F}$, which is the smallest sigma-algebra of subsets of Ω making the maps $\omega \mapsto \omega(A)$ measurable for each Borel set $A \subset \mathbb{X}$ cf. [12, Chap. 10]. A jointly measurable map

$$f : (\mathbb{X})^k \times \Omega \ni (x_1, \ldots, x_k; \omega) \ \mapsto f(x_1, \ldots, x_k; \omega) \in \bar{\mathbb{R}}$$

is called a process or, more specifically, a k-process. Here $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, $k \in \mathbb{N}_0 = \{0, 1, \ldots\}$, and when $k = 0$, i.e., $f : \ \Omega \ni \omega \mapsto f(\omega) \in \bar{\mathbb{R}}$, we call f a random variable. We also note that for every Borel function $\phi \geq 0$ on $\mathbb{X}$, the map

$$\omega \mapsto \int \phi(x)\omega(dx)$$

is well-defined and measurable, hence a random variable. We say that a k-process f depends only on $\mathcal{X} \subseteq \mathbb{X}$, if $f(x_1, \ldots, x_k; \omega) = f(x_1, \ldots, x_k; \omega \cap \mathcal{X})$ for all $\omega \in \Omega$ and $x_1, \ldots, x_k \in \mathbb{X}$.

We let $\mathbb{X}^n_{diag} = \{x = (x_1, \cdots, x_n) \in \mathbb{X}^n : x_i = x_j \text{ for some } i \neq j\}$ and $\mathbb{X}^n_{\neq} = \mathbb{X}^n \setminus \mathbb{X}^n_{diag}$, where $\mathbb{X}^1_{\neq} = \mathbb{X}$. Given a k-process f and $n \in \mathbb{N}$ we define the n-th coefficient $f_{(n)}$ of f as a function $f_{(n)} : \mathbb{X}^k \times \mathbb{X}^n_{\neq} \mapsto \bar{\mathbb{R}}$ such that

$$f_{(n)}(x_1, \ldots, x_k; y_1, \ldots, y_n) = f(x_1, \ldots, x_k; \omega), \quad \text{where } \omega = \{y_1, \ldots, y_n\}.$$

We also let $f_{(0)}(x_1,\dots,x_k) = f(x_1,\dots,x_k;\emptyset)$. Thus, coefficients $f_{(n)}$ are Borel functions on $\mathbb{X}^k \times \mathbb{X}^n_{\neq}$ invariant upon permutations of the last n coordinates. In particular, for random variables (0-processes) f we simply have $f_{(n)}(y_1,\dots,y_n) = f(\{y_1,\dots,y_n\})$, where $y_1,\dots,y_n$ are all different, and $f_{(0)} = f(\emptyset)$. Of course, if f is a k-process, then $\omega \mapsto f(x_1,\dots,x_k;\omega)$ is a random variable for every choice of $x_1,\dots,x_k \in \mathbb{X}$, and the n-th coefficient of this random variable is $f_{(n)}(x_1,\dots,x_k;y_1,\dots,y_n)$, provided $(y_1,\dots,y_n) \in \mathbb{X}^n_{\neq}$.

Now we define a Poisson probability measure $\mathbb{P}$ on $(\Omega,\mathcal{F})$ and the corresponding expectation $\mathbb{E}$. Notice that $N(A,\omega) := \omega(A)$ is a random measure on $\mathbb{X}$ under any probability measure on Ω, but we will consider the probability $\mathbb{P}$ which makes N a Poisson random measure with intensity measure $\sigma(A) = \mathbb{E}N(A)$. Here is a construction of $\mathbb{P}$. The main analytic datum is a non-atomic measure σ finite on compact subsets of $\mathbb{X}$. If $\mathcal{X}$ is a Borel subset of $\mathbb{X}$ and $\sigma(\mathcal{X}) < \infty$, then the corresponding probability, say $\mathbb{P}_{\mathcal{X}}$, is concentrated on *finite* configurations $\Omega_{\mathcal{X}}$ on $\mathcal{X}$ and defined by

$$\begin{aligned}\mathbb{E}_{\mathcal{X}}f &= \int_{\Omega_{\mathcal{X}}} f(\omega)\mathbb{P}_{\mathcal{X}}(d\omega)\\ &= e^{-\sigma(\mathcal{X})}\sum_{n=0}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n} f_{(n)}(y_1,\dots,y_n)\sigma(dy_n)\cdots\sigma(dy_1),\end{aligned} \tag{2.1}$$

cf. [20, p. 196]. Here the first term on the rightmost of (2.1) is $e^{-\sigma(\mathcal{X})}f_{(0)}$, according to a general convention.

Further, let Borel sets $\mathcal{X}_1,\mathcal{X}_2,\ldots \subset \mathbb{X}$ be such that $\bigcup_m \mathcal{X}_m = \mathbb{X}$, $\mathcal{X}_m \cap \mathcal{X}_n = \emptyset$ for $m \neq n$, and $\sigma(\mathcal{X}_m) < \infty$ for every m. We identify $\Omega_{\mathbb{X}}$ with $\otimes_m \Omega_{\mathcal{X}_m}$ by identifying ω with $(\omega \cap \mathcal{X}_m)_m$. Then $\mathbb{P}$ is unambiguously defined as the product measure,

$$\mathbb{P} = \otimes_m \mathbb{P}_{\mathcal{X}_m}.$$

For $\mathcal{X} \subset \mathbb{X}$, $\mathbb{P}_{\mathcal{X}}$ may be considered as a marginal distribution of $\mathbb{P}$, and for random variables f_1, f_2 depending only on disjoint $\mathcal{X}_1,\mathcal{X}_2 \subset \mathbb{X}$, respectively, we have

$$\mathbb{E}[f_1(\omega)f_2(\omega)] = \mathbb{E}_{\mathcal{X}_1}[f_1(\omega)]\,\mathbb{E}_{\mathcal{X}_2}[f_2(\omega)]. \tag{2.2}$$

Here the notions of the independence of a function from a set of arguments, and the probabilistic independence happily meet. In what follows $\mathbb{E}$ and $\mathbb{P}$ are always the expectation and distribution making ω a Poisson random measure with control measure σ (in Sect. 4 we make additional structure assumptions on $\mathbb{X}$ and σ).

In what follows we denote $\omega_1 = \omega$, $\omega_0 = \sigma$, for $\omega \in \Omega$. For a 1-process $f \geq 0$ and $\epsilon \in \{0,1\}$, we have

$$\mathbb{E}\int_{\mathbb{X}} f(x;\omega)\;\omega_\epsilon(dx) = \int_{\mathbb{X}} \mathbb{E}f\,(x;\omega+\epsilon\delta_x)\;\sigma(dx). \tag{2.3}$$

Indeed, for $\epsilon = 0$ the identity follows from Fubini-Tonelli, and if $\epsilon = 1$, then it is the celebrated Mecke-Palm formula, see also [15, (2.10)]. (For the readers's convenience a direct proof of the Mecke-Palm formula is given in Sect. 4.2.)

We say that a k-process f vanishes on the diagonals if for all $\omega \in \Omega = \Omega_{\mathbb{X}}$ we have $f(x_1, \ldots, x_k; \omega) = 0$ whenever $(x_1, \ldots, x_k) \in \mathbb{X}^k_{diag}$, i.e. whenever $x_i = x_j$ for some $1 \leq i < j \leq k$. This condition is restrictive only if $k \geq 2$. We propose the following mixed Mecke-Palm formula.

Lemma 2.1 *If $f \geq 0$ vanishes on the diagonals and $\epsilon_1, \ldots, \epsilon_k \in \{0, 1\}$, then*

$$\mathbb{E}\int_{\mathbb{X}^k} f(x_1, \ldots, x_k; \omega)\, \omega_{\epsilon_1}(dx_1) \cdots \omega_{\epsilon_k}(dx_k) \tag{2.4}$$

$$= \int_{\mathbb{X}^k} \mathbb{E} f\Big(x_1, \ldots, x_k; \omega + \sum_{i=1}^k \epsilon_i \delta_{x_i}\Big)\, \sigma(dx_1) \cdots \sigma(dx_k).$$

Proof Case $k = 0$ is trivial: $\mathbb{E} f(\omega) = \mathbb{E} f(\omega)$. Case $k = 1$ is precisely (2.3). For $k > 1$ we define

$$g(x; \omega; \epsilon_1, \ldots, \epsilon_{k-1}) = \int_{\mathbb{X}^{k-1}} f(x_1, \ldots, x_{k-1}, x; \omega)\, \omega_{\epsilon_1}(dx_1) \cdots \omega_{\epsilon_{k-1}}(dx_{k-1}).$$

Since $f(x_1, \ldots, x_{k-1}, x; \omega)$ vanishes on $\mathbb{X}^k_{diag}$, we get

$$\begin{aligned} &g(x; \omega + \delta_x; \epsilon_1, \ldots, \epsilon_{k-1}) \\ &= \int_{\mathbb{X}^{k-1}} f(x_1, \ldots, x_{k-1}, x; \omega + \delta_x)\, (\omega_{\epsilon_1} + \epsilon_1 \delta_x)(dx_1) \cdots (\omega_{\epsilon_{k-1}} + \epsilon_{k-1}\delta_x)(dx_{k-1}) \\ &= \int_{\mathbb{X}^{k-1}} f(x_1, \ldots, x_{k-1}, x; \omega + \delta_x)\, \omega_{\epsilon_1}(dx_1) \cdots \omega_{\epsilon_{k-1}}(dx_{k-1}). \end{aligned} \tag{2.5}$$

By (2.3), (2.5) and induction we obtain

$$\begin{aligned} &\mathbb{E}\int_{\mathbb{X}^k} f(x_1, \ldots, x_k; \omega)\, \omega_{\epsilon_1}(dx_1) \cdots \omega_{\epsilon_k}(dx_k) \\ &= \mathbb{E}\int_{\mathbb{X}} g(x_k; \omega; \epsilon_1, \ldots, \epsilon_{k-1})\, \omega_{\epsilon_k}(dx_k) = \int_{\mathbb{X}} \mathbb{E} g(x_k; \omega + \epsilon_k \delta_{x_k}; \epsilon_1, \ldots, \epsilon_{k-1})\, \sigma(dx_k) \\ &= \int_{\mathbb{X}} \mathbb{E}\int_{\mathbb{X}^{k-1}} f(x_1, \ldots, x_{k-1}, x_k; \omega + \epsilon_k \delta_{x_k})\, \omega_{\epsilon_1}(dx_1) \cdots \omega_{\epsilon_{k-1}}(dx_{k-1}) \sigma(dx_k) \\ &= \int_{\mathbb{X}^k} \mathbb{E} f\Big(x_1, \ldots, x_k; \omega + \sum_{i=1}^k \epsilon_i \delta_{x_i}\Big)\, \sigma(dx_1) \cdots \sigma(dx_k), \end{aligned}$$

which proves (2.4). □

Remark 2.2 Lemma 2.1 extends to signed processes f satisfying

$$\int_{\mathbb{X}^k} \mathbb{E}\left| f\Big(x_1, \ldots, x_k; \omega + \sum_{i=1}^{k} \epsilon_i \delta_{x_i}\Big)\right| \sigma(dx_1)\cdots\sigma(dx_k) < \infty,$$

because of the decomposition $f = f_+ - f_-$, where $f_+ = \max(f, 0)$ and $f_- = \max(-f, 0)$. In what follows we leave such extensions to the reader.

Remark 2.3 The assumption in Lemma 2.1 that f should vanish on the diagonals is essential. Indeed, take $k = 2$ and (deterministic) $f(x_1, x_2; \omega) = \mathbf{1}_{x_1=x_2}$ for $(x_1, x_2) \in \mathbb{X}^2$. Considering the atoms of ω we have

$$\int_{\mathbb{X}^2} f(x_1, x_2; \omega)\ \omega(dx_1)\omega(dx_2) = \sum_{x_1\in\omega}\sum_{x_2\in\omega} \mathbf{1}_{x_1=x_2} = \omega(\mathbb{X}),$$

hence

$$\mathbb{E}\int_{\mathbb{X}^2} f(x_1, x_2; \omega)\ \omega(dx_1)\omega(dx_2) = \sigma(\mathbb{X}).$$

On the other hand σ is non-atomic, therefore

$$\int_{\mathbb{X}^2} \mathbb{E}f(x_1, x_2; \omega)\ \sigma(dx_1)\sigma(dx_2) = \int_{\mathbb{X}^2} f(x_1, x_2; \omega)\sigma(dx_1)\sigma(dx_2) = 0.$$

Motivated by the above example we shall give a version of the multiple Mecke-Palm formula for processes which do not necessarily vanish on the diagonals. This calls for a notation that can handle *partitions*: For integers $k, n \geq 1$ we consider a family of pairwise disjoint nonempty sets (blocks) of integers $P = \{P_1, \ldots, P_k\}$, such that $\bigcup_{i=1}^{k} P_i = \{1, \ldots, n\}$. Thus, P is a partition of $\{1, \ldots, n\}$. We denote by $\mathcal{P}_n$ the set of all such partitions. We will use partitions to describe effects of mixed integrations with respect to the Poisson measure and the control measure on the diagonals of $\mathbb{X}^n$, in a manner which resembles the approach to multiple Itô integrals and compensated Poisson integrals in [19]. For $P \in \mathcal{P}_n$ we let

$$\mathbb{X}^n_P = \Big\{(x_1, \ldots, x_n) \in \mathbb{X}^n : x_i = x_j \text{ iff } i, j \in P_s \text{ for some } s \in \{1, \ldots, k\}\Big\}.$$

For $P = \{P_1, \ldots, P_k\} \in \mathcal{P}_n$ and $y \in \mathbb{X}^k_{\neq}$, we define $y^{[P]} = (y^{[P]}_1, \ldots, y^{[P]}_n)$ by letting $y^{[P]}_i = y_j$ if $i \in P_j$. We have, as in Remark 2.3,

$$\begin{aligned}
&\int_{\mathbb{X}^n} f(x_1, \ldots, x_n; \omega)\omega(dx_1)\cdots\omega(dx_n)\\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}_n} \int_{\mathbb{X}^n_P} f(x;\omega)\omega(dx_1)\cdots\omega(dx_n)\\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}_n} \int_{\mathbb{X}^k_{\neq}} f(y^{[P]};\omega)\omega(dy_1)\cdots\omega(dy_k). \qquad (2.6)
\end{aligned}$$

As in Remark 2.3 we also note that for $n > 1$ and all ω,

$$\int_{\mathbb{X}^n} \mathbf{1}_{x_1=x_2=\ldots=x_n}\sigma(dx_1)\omega(dx_2)\cdots\omega(dx_n) = 0, \qquad (2.7)$$

because the first marginal of the product measure is non-atomic. Therefore in view of generalizing (2.6) to mixed integrations against ω_1 and ω_0, we propose the following notation. For $\epsilon_1, \ldots, \epsilon_n \in \{0, 1\}$ we let $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$ and consider the family $\mathcal{P}^\epsilon_n$ of all the partitions $P = \{P_1, \ldots, P_k\}$ of $\{1, \ldots, n\}$ such that for every block $P_i \in P$ with $|P_i| > 1$ we have $\epsilon_j = 1$ for all $j \in P_i$. For $P \in \mathcal{P}^\epsilon_n$ we let $\epsilon^{[P]} = (\epsilon^{[P]}_1, \ldots, \epsilon^{[P]}_k)$, where $\epsilon^{[P]}_1 = \epsilon_{i_1}, \ldots, \epsilon^{[P]}_k = \epsilon_{i_k}$ and $i_1 \in P_1, \ldots, i_k \in P_k$. For $y = (y_1, \ldots, y_k) \in \mathbb{X}^k$ we then let $y_{\epsilon^{[P]}} = \{y_i : \epsilon^{[P]}_i = 1\}$. In the following extension of Lemma 2.1 we write x for $(x_1, \ldots, x_n) \in \mathbb{X}^n$ and $\sigma^k(dy) = \sigma(dy_1)\cdots\sigma(dy_k)$. The identity (2.8) below gives an algorithm to calculate expectations of Poisson integrals mixed with integrations against the control measure.

Theorem 2.4 *Let $\mathbb{E}$ be the expectation making configurations ω on $\mathbb{X}$ a Poisson random measure with control measure σ. For every n-process $f \geq 0$ and $\epsilon_1, \ldots, \epsilon_n \in \{0, 1\}$ we have*

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{X}^n} f(x;\omega)\,\omega_{\epsilon_1}(dx_1)\cdots\omega_{\epsilon_n}(dx_n) \qquad (2.8)\\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}^\epsilon_n} \int_{\mathbb{X}^k_{\neq}} \mathbb{E}f\left(y^{[P]};\omega\cup y_{\epsilon^{[P]}}\right)\sigma^k(dy).
\end{aligned}$$

Proof By similar reasons as in (2.6), and by Lemma 2.1,

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{X}^n} f(x;\omega)\,\omega_{\epsilon_1}(dx_1)\cdots\omega_{\epsilon_n}(dx_n) \\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}_n} \mathbb{E}\int_{\mathbb{X}^n_P} f(x;\omega)\ \omega_{\epsilon_1}(dx_1)\cdots\omega_{\epsilon_n}(dx_n) \\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}^\epsilon_n} \mathbb{E}\int_{\mathbb{X}^n_P} f(x;\omega)\ \omega_{\epsilon_1}(dx_1)\cdots\omega_{\epsilon_n}(dx_n) \qquad (2.9)\\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}^\epsilon_n} \mathbb{E}\int_{\mathbb{X}^k_{\neq}} f\left(y^{[P]};\omega\right)\ \omega_{\epsilon[P]}(dy) \\
&= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}^\epsilon_n} \int_{\mathbb{X}^k_{\neq}} \mathbb{E}f\left(y^{[P]};\omega\cup y_{\epsilon[P]}\right)\ \sigma^k(dy).
\end{aligned}$$

In (2.9) we use (2.7) to eliminate $P \notin \mathcal{P}^\epsilon_n$. □

3 Moments

In this section we give applications of the mixed Mecke-Palm formula to expectations of products of mixed stochastic integrals. As before, ω denotes the Poisson random measure with control measure σ on $\mathbb{X}$ and probability $\mathbb{P}$ and expectation $\mathbb{E}$.

3.1 General Moment Formulas

Theorem 3.1 below generalizes moment formulas of [7, 21]. As we see in the proof, the result is equivalent to the mixed Mecke-Palm formula (2.8) and is obtained after a simple *linearization* procedure. Let S be a finite set and $\mathbb{X}^S = \{x : S \to \mathbb{X}\}$. For $x \in \mathbb{X}^S$ and $s \in S$ we write $x_s = x(s)$. We consider $\mathcal{P}(S)$, the class of all the partitions $P = \{P_1, \ldots, P_k\}$ of S. Here (blocks) $P_1, \ldots, P_k$ are disjoint, and $\bigcup_{\alpha=1}^k P_\alpha = S$. Let $P \in \mathcal{P}(S)$ and consider the P-diagonal:

$$\mathbb{X}^S_P = \{x \in \mathbb{X}^S : x_s = x_t \text{ iff there is } \in \{1,\ldots,k\} \text{ such that } s,t \in P_\alpha\}.$$

For $\epsilon : S \to \{0,1\}$ we denote $\omega_\epsilon(dx) = \otimes_{s\in S}\omega_{\epsilon_s}(dx_s)$. We note that ω_ϵ vanishes on X^S_P if there is block $P_\alpha \in P$ with cardinality $|P_\alpha| > 1$ and such that $\epsilon = 0$ at some point of P_α. This is so because the product measure has a non-atomic marginal. The set of all the remaining partitions will be denoted $\mathcal{P}^\epsilon(S)$. In particular, if $P \in \mathcal{P}^\epsilon(S)$ then ϵ is constant on every block of P, and we may define $\epsilon^P_\alpha := \epsilon_s$ if $s \in P_\alpha$, $\alpha = 1, \ldots, k$. We denote $\epsilon^P = (\epsilon^P_1, \ldots, \epsilon^P_k)$. For $y = (y_1, \ldots, y_k) \in \mathbb{X}^k$ we let $y^P_s = y_\alpha$ if $s \in P_\alpha$. Thus, $\epsilon^P \in \{0,1\}^k$ and $y^P \in \mathbb{X}^S$. For measurable $f : \mathbb{X}^S \to \mathbb{R}_+$ we have

$$\int_{\mathbb{X}^P} f(x)\omega_\epsilon(dx) = \int_{\mathbb{X}^k} f(y^P)\omega_{\epsilon_1}(dy_1)\cdots\omega_{\epsilon_k}(dy_k),$$

which follows because ω is a sum of Dirac measures supported at different points.

Let $l \in \mathbb{N}$ and $r_1, n_1, \ldots, r_l, n_l \geq 1$. We define

$$S = \{(\alpha, \beta, \gamma) : 1 \leq \alpha \leq l, 1 \leq \beta \leq r_\alpha, 1 \leq \gamma \leq n_\alpha\}.$$

If $1 \leq \alpha \leq l$ and $1 \leq \gamma \leq n_\alpha$, then we let

$$S_{\alpha,\gamma} = \{(\alpha, \beta, \gamma) \in S : 1 \leq \beta \leq r_\alpha\}.$$

For $z \in \mathbb{X}^S$ we write, as usual, $z_{S_{\alpha,\gamma}}$ for the restriction of z to $S_{\alpha,\gamma}$. If $P = \{P_1, \ldots, P_k\} \in \mathcal{P}(S)$ and $y \in \mathbb{X}^k$, then $y^P_{S_{\alpha,\gamma}}$ denotes $(y^P)_{S_{\alpha,\gamma}}$. In particular, $y^P_{S_{\alpha,\gamma}} \in \mathbb{X}^{S_{\alpha,\gamma}}$.

Theorem 3.1 *Let $\mathbb{E}$ be the expectation making ω a Poisson random measure on $\mathbb{X}$ with control measure σ. Let $f_0, f_1, \ldots, f_l \geq 0$ be $0, r_1, \ldots, r_l$-processes, respectively. Let $\epsilon_{(1)} \in \{0,1\}^{r_1}, \ldots, \epsilon_{(l)} \in \{0,1\}^{r_l}$. For $s = (\alpha, \beta, \gamma) \in S$ we define $\epsilon_s = \epsilon_{(\alpha)}(\beta)$. Then,*

$$\mathbb{E}\left[f_0(\omega)\left(\int_{\mathbb{X}^{r_1}} f_1(y;\omega)\omega_{\epsilon_{(1)}}(dy)\right)^{n_1} \ldots \left(\int_{\mathbb{X}^{r_l}} f_l(y;\omega)\omega_{\epsilon_{(l)}}(dy)\right)^{n_l}\right] \tag{3.1}$$
$$= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}^\epsilon(S)} \mathbb{E}\left[\int_{\mathbb{X}^k_{\neq}} f_0(\omega \cup \bigcup_{\epsilon_s=1}\{y^P_s\}) \prod_{\alpha=1}^{l}\prod_{\gamma=1}^{n_\alpha} f_\alpha(y^P_{S_{\alpha,\gamma}}; \omega \cup \bigcup_{\epsilon_s=1}\{y^P_s\})\sigma^k(dy)\right].$$

Proof The first transformation in the calculation below we call linearization, and the last one follows from Theorem 2.4:

$$\mathbb{E}\left[f_0(\omega)\left(\int_{\mathbb{X}^{r_1}} f_1(y;\omega)\omega_{\epsilon_{(1)}}(dy)\right)^{n_1} \ldots \left(\int_{\mathbb{X}^{r_l}} f_l(y;\omega)\omega_{\epsilon_{(l)}}(dy)\right)^{n_l}\right]$$
$$= \mathbb{E}\Big[f_0(\omega)\int_{\mathbb{X}^{r_1}} \ldots \int_{\mathbb{X}^{r_1}} \prod_{\gamma=1}^{n_1} f_1(y_{S_{1,\gamma}};\omega)\omega_{\epsilon_{(1)}}(dy_{S_{1,1}})\ldots\omega_{\epsilon_{(1)}}(dy_{S_{1,n_1}})\times$$
$$\cdots \times \int_{\mathbb{X}^{r_l}} \ldots \int_{\mathbb{X}^{r_l}} \prod_{\gamma=1}^{n_l} f_l(y_{S_{l,\gamma}};\omega)\omega_{\epsilon_{(l)}}(dy_{S_{l,1}})\ldots\omega_{\epsilon_{(l)}}(dy_{S_{l,n_l}})\Big]$$

$$= \mathbb{E}\Big[\int_{\mathbb{X}^S} f_0(\omega)\prod_{\alpha=1}^{l}\prod_{\gamma=1}^{n_\alpha} f_\alpha(y_{S_{\alpha,\gamma}};\omega)\omega_{\epsilon(1)}(dy_{S_{1,1}})\dots\omega_{\epsilon(1)}(dy_{S_{1,n_1}})\dots$$

$$\dots\omega_{\epsilon(l)}(dy_{S_{l,1}})\dots\omega_{\epsilon(l)}(dy_{S_{l,n_l}})\Big]$$

$$= \mathbb{E}\left[\sum_{P=\{P_1,\dots,P_k\}\in\mathcal{P}(S)}\int_{\mathbb{X}^k_{\neq}} f_0(\omega)\prod_{\alpha=1}^{l}\prod_{\gamma=1}^{n_\alpha} f_\alpha(y^P_{S_{\alpha,\gamma}};\omega)\omega_{\epsilon^P}(dy)\right]$$

$$= \mathbb{E}\sum_{P=\{P_1,\dots,P_k\}\in\mathcal{P}^\epsilon(S)}\int_{\mathbb{X}^k_{\neq}} f_0(\omega\cup\bigcup_{\epsilon_s=1}\{y^P_s\})\prod_{\alpha=1}^{l}\prod_{\gamma=1}^{n_\alpha} f_\alpha(y^P_{S_{\alpha,\gamma}};\omega\cup\bigcup_{\epsilon_s=1}\{y^P_s\})\sigma^k(dy).$$

□

In concrete computations one may either use Theorem 3.1, along with its somewhat heavy notation, or just follow its proof, i.e. use linearization and the mixed Mecke-Palm formula. For instance in Lemma 3.3 below it is simpler to use the latter approach.

3.2 *Moment Formulas for Stochastic Integrals of 1-Processes*

We first specialize to 1-processes. Let $k, l, n_1, \dots, n_l \in \mathbb{N} = \{1, 2, \dots\}$ and $n = n_1 + \dots + n_l$. For $j = 1, \dots, l$ and $P = \{P_1, \dots, P_k\} \in \mathcal{P}_n$ we denote

$$P_{i,j} = \{d \in P_i : \sum_{0<m<j} n_m < d \le \sum_{0<m\le j} n_m\}.$$

Let $|P_{i,j}|$ be the number of elements of $P_{i,j}$.

Corollary 3.2 *For a random variable* $f_0 \ge 0$ *and* 1*-processes* $f_1, \dots, f_l \ge 0$,

$$\mathbb{E}\Big[f_0(\omega)\Big(\int_{\mathbb{X}} f_1(x;\omega)\omega(dx)\Big)^{n_1}\cdots\Big(\int_{\mathbb{X}} f_l(x;\omega)\omega(dx)\Big)^{n_l}\Big] \tag{3.2}$$

$$= \sum_{P\in\mathcal{P}_n}\mathbb{E}\int_{\mathbb{X}^k} f_0(\omega+\sum_{i=1}^k\delta_{y_i})f_1^{|P_{1,1}|}(y_1;\omega+\sum_{i=1}^k\delta_{y_i})\cdots f_l^{|P_{1,l}|}(y_1;\omega+\sum_{i=1}^k\delta_{y_i})\times$$

$$\times f_1^{|P_{k,1}|}(y_k;\omega+\sum_{i=1}^k\delta_{y_i})\cdots f_l^{|P_{k,l}|}(y_k;\omega+\sum_{i=1}^k\delta_{y_i})\sigma(dy_1)\dots\sigma(dy_k).$$

Proof The result follows from Theorem 3.1. □
For $l = 1$ we recover [21, (1.2)]:

$$\mathbb{E}\left[v(\omega)\left(\int_{\mathbb{X}} u(x;\omega)\omega(dx)\right)^n\right]$$
$$= \sum_{P=\{P_1,\ldots,P_k\}\in\mathcal{P}_n} \mathbb{E}\left[\int_{\mathbb{X}^k} v(\omega\cup y)u(y_1;\omega\cup y)^{|P_1|}\ldots u(y_k;\omega\cup y)^{|P_k|}\sigma(dy_1)\ldots\sigma(dy_k)\right],$$

where $y = \{y_1,\ldots,y_k\}$ and $u \geq 0$ is a 1-process. With arbitrary l we obtain an alternative proof of [7, Theorem 3.1] for random Poisson measures. In passing we also refer the reader to recent papers [16] and [6].

3.3 The Second Moment of Stochastic Integrals of 2-Processes

Moments of arbitrary k-processes require formulas of increasing complexity, but they are entirely explicit. Here is a telling example.

Lemma 3.3 *If $f \geq 0$ is a 2-process, then*

$$\mathbb{E}\left(\int_{\mathbb{X}^2} f(x_1,x_2;\omega)\omega(dx_1)\omega(dx_2)\right)^2 = \int_{\mathbb{X}} \mathbb{E}f^2(x,x;\omega\cup\{x\})\sigma(dx) \tag{3.3}$$
$$+ 2\int_{\mathbb{X}^2_{\neq}} \mathbb{E}f(x,x;\omega\cup\{x,y\})f(x,y;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$
$$+ 2\int_{\mathbb{X}^2_{\neq}} \mathbb{E}f(x,x;\omega\cup\{x,y\})f(y,x;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$
$$+ \int_{\mathbb{X}^2_{\neq}} \mathbb{E}f(x,x;\omega\cup\{x,y\})f(y,y;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$
$$+ \int_{\mathbb{X}^2_{\neq}} \mathbb{E}f^2(x,y;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$
$$+ \int_{\mathbb{X}^2_{\neq}} \mathbb{E}f(x,y;\omega\cup\{x,y\})f(y,x;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$
$$+ 2\int_{\mathbb{X}^3_{\neq}} \mathbb{E}f(x,x;\omega\cup\{x,y,z\})f(y,z;\omega\cup\{x,y,z\})\sigma(dx)\sigma(dy)\sigma(dz)$$
$$+ 2\int_{\mathbb{X}^3_{\neq}} \mathbb{E}f(x,y;\omega\cup\{x,y,z\})f(z,x;\omega\cup\{x,y,z\})\sigma(dx)\sigma(dy)\sigma(dz)$$

$$+\int_{\mathbb{X}^3_{\neq}} \mathbb{E}f(x,y;\omega\cup\{x,y,z\})f(x,z;\omega\cup\{x,y,z\})\sigma(dx)\sigma(dy)\sigma(dz)$$

$$+\int_{\mathbb{X}^3_{\neq}} \mathbb{E}f(y,x;\omega\cup\{x,y,z\})f(z,x;\omega\cup\{x,y,z\})\sigma(dx)\sigma(dy)\sigma(dz)$$

$$+\int_{\mathbb{X}^4_{\neq}} \mathbb{E}f(x,y;\omega\cup\{x,y,z,t\})f(z,t;\omega\cup\{x,y,z,t\})\sigma(dx)\sigma(dy)\sigma(dz)\sigma(dt).$$

Proof By linearization,

$$\left(\int_{\mathbb{X}^2} f(x_1,x_2;\omega)\omega(dx_1)\omega(dx_2)\right)^2$$
$$=\int_{\mathbb{X}^4} g(x,y,z,t)\omega(dx)\omega(dy)\omega(dz)\omega(dt),$$

where $g(x,y,z,t;\omega)=f(x,y;\omega)f(z,t;\omega)$. We will use Theorem 2.4. The partitions involved have $k=1$, 2, 3 or 4 blocks, because the number 4 can be represented as the following sums: 4, $3+1$, $2+2$, $2+1+1$, $1+1+1+1$. In particular, the partition of $\{1,2,3,4\}$ with only one block ($k=1$), namely $\{\{1,2,3,4\}\}$, contributes

$$\mathbb{E}\int_{\mathbb{X}} g(x,x,x,x;\omega\cup\{x\})\sigma(dx)=\int_{\mathbb{X}}\mathbb{E}f^2(x,x;\omega\cup\{x\})\sigma(dx)$$

to (3.3). Then, partitions with $k=2$ blocks are of type $3+1$ and $2+2$. In the first case there are 4 different partitions as there are 4 different choices of the singleton. For instance, $P=\{\{1,2,3\},\{4\}\}$ contributes

$$\int_{\mathbb{X}^2_{\neq}} \mathbb{E}g(x,x,x,y;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$
$$=\int_{\mathbb{X}^2_{\neq}} \mathbb{E}f(x,x;\omega\cup\{x,y\})f(x,y;\omega\cup\{x,y\})\sigma(dx)\sigma(dy)$$

to (3.3). The contribution to (3.3) from all the partitions of type $3+1$ are the 2nd and the 3rd terms on the right-hand side of (3.3). In the case $2+2$, $P=\{\{1,2\},\{3,4\}\}$ contributes

$$\int_{\mathbb{X}^2_{\neq}} \mathbb{E}f(x,x;\omega\cup\{x,y\})f(y,y;\omega\cup\{x,y\})\sigma(dx)\sigma(dy),$$

to (3.3), and the contributions from all the partitions of type $2+2$ are precisely the 4th through 6th terms on the right-hand side of (3.3).

For $k = 3$ we have partitions of type $2+1+1$, e.g. $P = \{\{1,2\},\{3\},\{4\}\}$, which contributes

$$\int_{\mathbb{X}^3_{\neq}} \mathbb{E} f(x,x;\omega \cup \{x,y,z\}) f(y,z;\omega \cup \{x,y,z\}) \sigma(dx)\sigma(dy)\sigma(dz),$$

to (3.3), and all partitions of type $2+1+1$ result in the 7th through 10th terms on the right-hand side of (3.3). Finally, the partition into $k = 4$ singletons yields

$$\int_{\mathbb{X}^4_{\neq}} \mathbb{E} f(x,y;\omega \cup \{x,y,z,t\}) f(z,t;\omega \cup \{x,y,z,t\}) \sigma(dx)\sigma(dy)\sigma(dz)\sigma(dt).$$

This finishes the verification of (3.3). □

We now investigate the second moment of mixed double stochastic integrals, the ones with respect to the random measures $\omega \otimes \sigma$ and $\sigma \otimes \omega$.

Lemma 3.4 *If $f \geq 0$ is a 2-process, then*

$$\begin{aligned}
&\mathbb{E}\left(\int_{\mathbb{X}^2} f(x_1,x_2;\omega)\omega(dx_2)\sigma(dx_1)\right)^2 \\
&\quad = \mathbb{E}\left(\int_{\mathbb{X}^2} f(x_1,x_2;\omega)\sigma(dx_1)\omega(dx_2)\right)^2 \qquad (3.4)\\
&\quad = \mathbb{E}\int_{\mathbb{X}^3_{\neq}} f(x,y;\omega \cup \{y\}) f(z,y;\omega \cup \{y\}) \sigma(dx)\sigma(dy)\sigma(dz) \qquad (3.5)\\
&\quad + \mathbb{E}\int_{\mathbb{X}^4_{\neq}} f(x,y;\omega \cup \{y,t\}) f(z,t;\omega \cup \{y,t\}) \sigma(dx)\sigma(dy)\sigma(dz)\sigma(dt).
\end{aligned}$$

Proof Equation (3.4) follows from Fubini-Tonelli. Then the expectation in (3.4) is written as

$$\mathbb{E}\int_{\mathbb{X}^4} f(x,y;\omega) f(z,t;\omega) \sigma(dx)\omega(dy)\sigma(dz)\omega(dt),$$

and by Theorem 2.4 we get the equality (3.5), as in the proof of Lemma 3.3. □

4 Lévy Systems

An important motivation for this work is due to the so-called Lévy systems for Lévy processes. These are identities between expectations of sums taken with respect to the jumps of a Lévy process and expectations of integrals taken with respect to the corresponding intensity measure. There exists a considerable variety of (multiple) Lévy systems, which we discuss below.

4.1 General Result

We consider (time) $\mathbb{R}_+ = (0, \infty)$, (space) $\mathbb{R}^d$ and (space-time) $\mathbb{R}_+ \times \mathbb{R}^d$.

Let ν be a non-zero Lévy measure on $\mathbb{R}^d$, thus $\nu(\{0\}) = 0$ and

$$\int_{\mathbb{R}^d} \min\{1, z^2\}\nu(dz) < \infty.$$

Let $X = \{X_t\}_{t\geq 0}$ be a Lévy process in $\mathbb{R}^d$ with Lévy triplet (ν, A, b), where A is a symmetric, nonnegative-definite $d \times d$ matrix and $b \in \mathbb{R}^d$ [22]. Let $\mathbb{P}$ and $\mathbb{E}$ be the distribution and the expectation of the process and consider

$$p_t(A) = \mathbb{P}(X_t \in A),$$

the convolution semigroup of X. Let $\Delta X_u = X_u - X_{u-}$ and

$$\omega = \sum_{u>0,\, \Delta X_u \neq 0} \delta_{(u, \Delta X_u)}.$$

Then ω is a Poisson random measure with the intensity (control) measure $\sigma(dudz) = du\nu(dz)$ on

$$\mathbb{X} = \mathbb{R}_+ \times \mathbb{R}_0^d$$

[10, Sects. I.9, II.3, Example II.4.1] related to X by the Lévy-Itô decomposition [22, Chap. 4], [10, Example II.4.1]. We may and do identify ω, $\mathbb{P}$ and $\mathbb{E}$ with those from Sect. 2 given by $\sigma(dudz) = du\nu(dz)$. The following well-known identity is called the (simple) Lévy system (more comments are given after the proof).

Lemma 4.1 *If* $F : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \bar{\mathbb{R}}$ *is nonnegative, then*

$$\mathbb{E} \sum_{\substack{0<u<\infty \\ \Delta X_u \neq 0}} F(u, X_{u-}, X_u) = \mathbb{E} \int_0^\infty \int_{\mathbb{R}^d} F(u, X_u, X_u + z)\nu(dz)du\,. \tag{4.1}$$

Proof First, let X be a compound Poisson process, that is $\nu(\mathbb{R}^d) < \infty$, $X(t) = \sum_{i=1}^{N(t)} Z_i$, where $N(t)$ has Poisson distribution with expectation $t\nu(\mathbb{R}^d)$, and Z_i are i.i.d. random variables with distribution $\nu/\nu(\mathbb{R}^d)$. Therefore

$$p_t = e^{-|\nu|t} e^{*t\nu} = e^{-|\nu|t} \sum_{n=0}^{\infty} \frac{t^n \nu^{*n}}{n!}.$$

By Fubini-Tonelli theorem the right-hand side of (4.1) equals

$$\int_0^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(u, x, x+z) p_u(dx) \nu(dz) du. \tag{4.2}$$

Let $S_i = \inf\{t > 0 : N(t) = i\}$, the arrival time of the i-th jump of X. Recall that S_i has gamma distribution, and clearly X_{S_i} has distribution $\widetilde{\nu}^{*i}$. By Fubini-Tonelli the left-hand side of (4.1) equals

$$\begin{aligned}
&\mathbb{E} \sum_{i=1}^\infty F(S_i, X_{S_i-}, X_{S_i}) \\
&= \sum_{i=1}^\infty \int_0^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(u, x, x+z) \frac{|\nu|^i u^{i-1}}{(i-1)!} e^{-|\nu| u}\, \widetilde{\nu}^{*(i-1)}(dx) \widetilde{\nu}(dz) du \\
&= \int_0^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(u, x, x+z) \left(e^{-|\nu|u} \sum_{i=1}^\infty \frac{u^{i-1} \nu^{*(i-1)}}{(i-1)!}(dx) \right) \nu(dz) du \\
&= \int_0^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(u, x, x+z) p_u(dx) \nu(dz) du.
\end{aligned}$$

This yields (4.1) for compound Poisson process X. Now let X be a general Lévy process. We shall prove that for every $\epsilon > 0$,

$$\mathbb{E} \sum_{\substack{0<u<\infty \\ |\Delta X_u| \geq \epsilon}} F(u, X_{u-}, X_u) = \mathbb{E} \int_0^\infty \int_{|z|\geq\epsilon} F(u, X_u, X_u + z) \nu(dz) dv. \tag{4.3}$$

To this end we use the following decomposition,

$$X_t = V_t + Z_t.$$

The terms in the decomposition have the following properties. Process V_t is a Lévy process with the triplet $(A, \nu\big|_{|z|<\epsilon}, b)$, on a probability space $(\Omega^V, \mathcal{F}^V, \mathbb{P}^V)$. Here $\nu\big|_{|z|<\epsilon}$ is the measure ν restricted to $\{z \in \mathbb{R}^d : |z| < \epsilon\}$. Z_t is a compound Poisson process on an independent probability space $(\Omega^Z, \mathcal{F}^Z, \mathbb{P}^Z)$, and has the Lévy measure $\nu\big|_{|z|\geq\epsilon}$. We denote by $\mathbb{E}^V, \mathbb{E}^Z$ and $\mathbb{P}^V$, $\mathbb{P}^Z$ the corresponding expectations and probabilities. We may assume that $\Omega = \Omega^V \times \Omega^Z$ and $\mathbb{P} = \mathbb{P}^V \otimes \mathbb{P}^Z$, according to the fact that V and Z are independent. In what follows we consider

$$\widetilde{F}(v, x, y) = F(v, V_{v-} + x, V_v + y). \tag{4.4}$$

By Fubini-Tonelli theorem and by (4.1) for the compound Poisson process Z, the left hand side of (4.3) becomes

$$\mathbb{E}^V\mathbb{E}^Z \sum_{|\Delta(V_u+Z_u)|\geq\epsilon} F(u, V_{u-}+Z_{u-}, V_u+Z_u)$$

$$= \mathbb{E}^V\mathbb{E}^Z \sum_{|\Delta Z_u|\geq\epsilon} \widetilde{F}(u, Z_{u-}, Z_u) = \mathbb{E}^V\mathbb{E}^Z \int_0^\infty \int_{|z|\geq\epsilon} \widetilde{F}(u, Z_u, Z_u+z)\nu(dz)du$$

$$= \mathbb{E} \int_0^\infty \int_{|z|\geq\epsilon} F(u, X_u, X_u+z)\nu(dz)du.$$

We have proved (4.3). Let $\epsilon \downarrow 0$. By the monotone convergence theorem,

$$\mathbb{E} \sum_{|\Delta Y_u|\geq\epsilon} F(u, X_{u-}, X_u) \to \mathbb{E} \sum_{\Delta Y_u\neq 0} F(u, X_{u-}, X_u), \tag{4.5}$$

and

$$\mathbb{E} \int_0^\infty \int_{|z|\geq\epsilon} F(u, X_u, X_u+z)\nu(dz)du \to \mathbb{E} \int_0^\infty \int_{\mathbb{R}^d} F(u, X_u, X_u+z)\nu(dz)du. \tag{4.6}$$

By (4.6), (4.5) and (4.3) we obtain (4.1). □

Lemma 4.1 asserts that the expected sum over the jumps of the Lévy process X equals to the expectation of the integral with respect to the corresponding intensity measure. As we remarked, the result is well-known, see [3], [8, p. 375], [4, VII.2(d)], but the above direct proof seems original, and will be used below. We next present a reformulation of Lemma 4.1 followed by extensions of Lemma 4.1, which we call multiple mixed Lévy systems.

Lemma 4.2 *If $F : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \bar{\mathbb{R}}$ is nonnegative, then*

$$\mathbb{E} \sum_{\substack{0<u<\infty \\ \Delta X_u\neq 0}} F(u, X_{u-}, \Delta X_u) = \mathbb{E} \int_0^\infty \int_{\mathbb{R}^d} F(u, X_u, z)\nu(dz)du\,.$$

Here $\mathbb{R}_+ = (0, \infty)$. The multiple mixed Lévy systems can be described within the framework presented in the previous sections. We consider the "simplex"

$$\mathbb{X}^n_< = \{(u_1, z_1; \ldots ; u_n, z_n) \in \mathbb{X}^n : 0 < u_1 < \cdots < u_n\}.$$

The following are all the multiple mixed Lévy systems.

Theorem 4.3 *Let X be a Lévy process in $\mathbb{R}^d$ with the Lévy measure ν, the expectation $\mathbb{E}$ and the Poisson random measure of jumps ω. Let $\epsilon_1, \dots, \epsilon_n \in \{0,1\}$ and let $F : (\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d)^n \mapsto [0, \infty]$ be measurable. Then,*

$$\mathbb{E} \int_{\mathbb{X}^n_<} F(u_1, X_{u_1-}, z_1; \dots; u_n, X_{u_n-}, z_n) \omega_{\epsilon_1}(du_1 dz_1) \dots \omega_{\epsilon_n}(du_n dz_n) \tag{4.7}$$

$$= \int_{\mathbb{X}^n_<} \mathbb{E}F(u_1, X_{u_1-}, z_1; \dots; u_j, X_{u_j-} + \sum_{i=1}^{j-1} \epsilon_i z_i,\ z_j; \dots;$$
$$u_n, X_{u_n-} + \sum_{i=1}^{n-1} \epsilon_i z_i,\ z_n)\ du_1 \nu(dz_1) \cdots du_n \nu(dz_n)$$
$$= \int_{\mathbb{X}^n_<} \int_{(\mathbb{R}^d)^n} F(u_1, y_1, z_1; \dots; u_n, \sum_{i=1}^{n} y_i + \sum_{i=1}^{n-1} \epsilon_i z_i, z_n)$$
$$p_{u_1}(dy_1) \dots p_{u_n - u_{n-1}}(dy_n)\, du_1 \nu(dz_1) \cdots du_n \nu(dz_n). \tag{4.8}$$

Proof We first prove this result for compound Poisson process X. By the Lévy-Itô decomposition for $t \geq 0$ we have

$$X_{t-} = X_{t-}(\omega) = \int_{(0,t)} \int_{\mathbb{R}^d} z\omega(dudz) - t\nu(\mathbb{R}^d),$$

and

$$X_t = X_t(\omega) = \int_{(0,t]} \int_{\mathbb{R}^d} z\omega(dudz) - t\nu(\mathbb{R}^d).$$

We note that X_{t-} is a 1-process on $\mathbb{X}$, and

$$\mathbf{1}_{\mathbb{X}^n_<}(u_1, z_1; \dots; u_n, z_n) F(u_1, X_{u_1-}, z_1; \dots; u_n, X_{u_n-}, z_n)$$

is an n-process, which vanishes on the diagonals. Using the notation from the proof of Lemma 2.1, by Theorem 2.4 we see that the left-hand side of (4.7) equals

$$\int_{\mathbb{X}^n_<} \mathbb{E}F\big(u_1, X_{u_1-}(\omega + \sum_{i=1}^{n} \epsilon_i \delta_{(u_i, z_i)}), z_1; \dots; u_n, X_{u_n-}(\omega + \sum_{i=1}^{n} \epsilon_i \delta_{(u_i, z_i)}), z_n\big)$$
$$du_1 \nu(dz_1) \cdots du_n \nu(dz_n).$$

Since

$$X_{u_j-}(\omega + \sum_{i=1}^{j} \epsilon_i \delta_{(u_i,z_i)}) = X_{u_j-}(\omega) + \sum_{i=1}^{j-1} \epsilon_i z_i,$$

(4.7) follows. Then we note that the distribution of X_{u-} is the same as that of X_u, which is p_u, and we use Fubini-Tonelli to get (4.8). This resolves the case of compound Poisson processes. The case of the general Lévy processes follows as in the proof of Lemma 4.1. □

The next two results are direct consequences of Theorem 4.3.

Corollary 4.4 *If X is a Lévy process and F is nonnegative, then*

$$\mathbb{E} \sum_{\substack{0<u_1<\ldots<u_n\leq\infty \\ \Delta X_{u_1}\neq 0,\ldots,\Delta X_{u_n}\neq 0}} F(u_1, X_{u_1-}, X_{u_1}; \ldots; u_n, X_{u_n-}, X_{u_n}) \tag{4.9}$$

$$= \mathbb{E} \int_0^\infty \ldots \int_{u_{n-1}}^\infty \int_{\mathbb{R}^d} \ldots \int_{\mathbb{R}^d} F(u_1, X_{u_1}, X_{u_1} + z_1; \ldots;$$

$$u_n, X_{u_n} + z_1 + \ldots + z_{n-1}, X_{u_n} + z_1 + \ldots + z_n)\nu(dz_n)\ldots\nu(dz_1)du_n\ldots du_1.$$

Corollary 4.5 *If X is a Lévy process and F is nonnegative, then*

$$\mathbb{E} \sum_{\substack{0<s<\infty \\ \Delta Y_s\neq 0}} \int_s^\infty \int_{\mathbb{R}^d} F(s, X_{s-}, X_s; s_1, X_{s_1}, X_{s_1} + z_1)\nu(dz_1)ds_1 \tag{4.10}$$

$$= \mathbb{E} \int_0^\infty \int_{\mathbb{R}^d} \sum_{\substack{s<s_1<\infty \\ \Delta X_{s_1}\neq 0}} F(s, X_s, X_s + z; s_1, X_{s_1-} + z, X_{s_1} + z)\nu(dz)ds$$

$$= \mathbb{E} \int_0^\infty \int_s^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(s, X_s, X_s + z; s_1, X_{s_1} + z, X_{s_1} + z + z_1)\nu(dz_1)\nu(dz)ds_1 ds.$$

We note in passing that Corollaries 4.4 and 4.5 can also be proved without using Mecke-Palm formula, in a way similar to the first part of the proof of Lemma 4.1, see [24]. The proofs are quite involved and the proof of the general mixed Lévy systems is fraught with problems if similar approach is to be used. On the contrary, Theorem 4.3 offers a clear insight into the structure of multidimensional mixed-type Lévy systems. The structure is explained by accumulating z_i, the i-th variable of the integrations performed in (4.8), as a jump of the process X at the moment u_i, but only if z_i is integrated against the Poisson random measure, rather than it's control measure. By accumulation we mean that such jumps are indeed added to the trajectory of the process. We encourage the reader to consider the statement of Corollary 4.5 from this perspective.

Remark 4.6 We note that Theorem 4.3 may be generalized to allow for n-processes more complicated than $F(u_1, X_{u_1-}, z_1; \ldots; u_n, X_{u_n-}, z_n)$, with similar proofs based on the mixed Mecke-Palm formula. Such extensions may involve predictable factors, cf. [8, p. 375], [4, VII.2(d)], and integration of processes which are not adapted to the usual filtration associated with the Lévy process.
To illustrate Remark 4.6 we give the following classical result, cf. [4, VII.2(d)]. An additional discussion is given at the end of Sect. 4.2.

Lemma 4.7 *If $F \geq 0$ and $g_t \geq 0$ is predictable, then*

$$\mathbb{E} \sum_{\substack{0<u<\infty \\ \Delta X_u \neq 0}} g_u F(u, X_{u-}, X_u) = \mathbb{E} \int_0^\infty \int_{\mathbb{R}^d} g_u F(u, X_u, X_u + z) \nu(dz) du \,. \tag{4.11}$$

Proof $g_u F(u, X_{u-}, X_u) = g_u(\omega) F(u, X_{u-}(\omega), X_{u-}(\omega) + z)$ is a 1-process on $\mathbb{R}_+ \times \mathbb{R}^d$. By predictability, $g_u(\omega + \delta_{(u,z)}) = g_u(\omega)$ almost surely. The result then follows from the usual Mecke-Palm identity (1). □

4.2 Applications

The purpose of this section is to present some consequences of our formulas. One of them is the well-known Ikeda-Watanabe formula [9], given as (4.13) below. It concerns the situation of the Lévy process X in $\mathbb{R}^d$ at the moment of the first exit from an open set $D \subset \mathbb{R}^d$. To state the result we employ the usual Markovian notation: for $x \in \mathbb{R}^d$ we write $\mathbb{E}^x$ and $\mathbb{P}^x$ for the expectation and distribution of $x + X$, and the latter is simply denoted by X, cf. [22, Chap. 8]. We let $p_t(x, A) = p_t(A - x) = \mathbb{P}^x(X_t \in A)$, so that

$$\mathbb{E}^x \int_0^\infty f(t, X_t) dt = \int_0^\infty \int_{\mathbb{R}^d} f(t, y) p_t(x, dy)$$

for (Borel) functions $f \geq 0$ and $x \in \mathbb{R}^d$. The time of the first exit of X from D is

$$\tau_D = \inf\{t > 0 : X_t \notin D\}.$$

The Dirichlet heat kernel $p_t^D(x, dy)$ is then defined by

$$\int_{\mathbb{R}^d} f(y) p_t^D(x, dy) = \mathbb{E}^x[f(X_t); \tau_D > t],$$

and we have

$$\mathbb{E}^x \int_0^{\tau_D} f(t, X_t)dt = \int_0^\infty \int_{\mathbb{R}^d} f(t, y) p_t^D(x, dy).$$

We now consider function $F(u, y, w) = \mathbf{1}_I(u)\mathbf{1}_A(y)\mathbf{1}_B(w)$, where I is a bounded interval, and $A \subset D$, $B \subset (\overline{D})^c$ are Borel sets in $\mathbb{R}^d$. We let

$$M(t) = \sum_{\substack{0<u\le t \\ |\Delta X_u| \neq 0}} F(u, X_{u-}, X_u) - \int_0^t \int_{\mathbb{R}^d} F(u, X_u, X_u + z)\nu(dz)dv.$$

We note that

$$\mathbb{E} \int_0^t \int_{\mathbb{R}^d} F(u, X_u, X_u + z)\nu(dz)dv \le |I|\nu(\{|z| > \operatorname{dist}(A, B)\}) < \infty, \tag{4.12}$$

so by Lemma 4.1, $\mathbb{E}M(t) = 0$. Let $0 \le s \le t$. By considering the Lévy process $u \mapsto X_{s+u} - X_s$, independent of X_r, $0 \le r \le s$, we calculate the conditional expectation

$$\mathbb{E}[\sum_{\substack{s<u\le t \\ |\Delta X_u| \neq 0}} F(u, X_{u-}, X_u) - \int_s^t \int_{\mathbb{R}^d} F(u, X_u, X_u + z)\nu(dz)dv \big| X_r, 0 \le r \le s] = 0,$$

cf. (4.4). By the above, M is a uniformly integrable martingale. By stopping at τ_D, we obtain,

$$\mathbb{P}^x[\tau_D \in I,\ X_{\tau_D -} \in A,\ X_{\tau_D} \in B] = \int_I \int_{B-y} \int_A p_u^D(x, dy)\nu(dz)du. \tag{4.13}$$

This defines the joint distribution of $(\tau_D, X_{\tau_D -}, X_{\tau_D})$ restricted to the event $\{X_{\tau_D -} \in D\}$ and calculated under $\mathbb{P}^x$.

As another application we use the double mixed Lévy system to prove the following classical result [10, II (3.9)].

Lemma 4.8 *Let X be a Lévy process in $\mathbb{R}^d$ with Lévy measure ν. Let the function $F : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \bar{\mathbb{R}}$ satisfy*

$$\mathbb{E} \int_0^\infty \int_{\mathbb{R}^d} F^2(v, X_v, X_v + z)\nu(dz)dv < \infty. \tag{4.14}$$

For every $t \in [0, \infty)$ the following limit exists in L^2

$$M_t = \lim_{\epsilon \to 0} \Bigg(\sum_{\substack{0<v\le t \\ |\Delta X| \ge \epsilon}} F(v, X_{v-}, X_v) - \int_0^t \int_{|z|\ge\epsilon} F(v, X_v, X_v + z)\nu(dz)dv \Bigg),$$

$t \mapsto M_t$ is a martingale with respect to $(\mathcal{F}_t)$, $\mathbb{E}M_t = 0$ and

$$\mathbb{E}M_t^2 = \mathbb{E}\int_0^t \int_{\mathbb{R}^d} F^2(v, X_v, X_v + z)\nu(dz)dv.$$

Furthermore, the square bracket of M is

$$[M]_t = \sum_{\substack{0<v\leq t \\ \Delta X_v \neq 0}} F^2(v, X_{v-}, X_v), \tag{4.15}$$

and the predictable quadratic variation of M is

$$\langle M\rangle_t = \int_0^t \int_{|z|\geq\epsilon} F(v, X_v, X_v + z)^2\nu(dz)dv. \tag{4.16}$$

Recall that $[M]$ is defined as the unique adapted right-continuous non-decreasing process with jumps $\Delta[M]_t = |\Delta M_t|^2$, and such that $t \mapsto |M|_t^2 - [M]_t$ is a (continuous) martingale starting at 0 [8, VII.42]. We verify the martingale property of $|M|_t^2 - [M]_t$ by using Corollaries 4.4 and 4.5. Notice that $\mathbb{E}[M]_t = \mathbb{E}\langle M\rangle_t$ by the single Lévy system. More details and applications can be found in [24]. In particular, the square bracket $[M]$ is used in [5] to estimate the L^p norms of Fourier multipliers defined in terms of Lévy processes. We refer the reader to [8, VII-VIII] and [10, II] for further details and reading.

As the third application we will calculate moments of the Lévy integral. Let $X_t = (\eta_t, \xi_t)$, where $t \geq 0$, be a Lévy process in $\mathbb{R}^2$. To simplify the discussion we further assume that η and ξ are (possibly dependent) subordinators with no drift [22, 23]. Let ν be the Lévy measure of X. Of course, ν is concentrated on $\mathbb{R}^2_{++} := (0,\infty) \times (0,\infty)$. Let ϕ be the Laplace exponent of η:

$$\mathbb{E}\left[e^{-x\eta_t}\right] = e^{-t\phi(x)}, \quad x \geq 0.$$

The following expression is called the Lévy integral,

$$Z = \int_0^\infty e^{-\eta_{t-}} d\xi_t = \sum_{\Delta X_t \neq 0} e^{-\eta_{t-}} \Delta\xi_t.$$

Lévy integrals represent stationary distributions of generalized Ornstein-Uhlenbeck process (see [18] for details, applications and references). By Lemma 4.1,

$$\mathbb{E}\left[\xi_1\right] = \mathbb{E}\sum_{\substack{0<t\leq 1 \\ \Delta X_t \neq 0}} \Delta\xi_t = \int_{\mathbb{R}^2_{++}} y\, d\nu(x, y).$$

We can use the multiple Lévy systems to calculate the moments of Z. The first three moments of Z take on the following form

$$\mathbb{E}[Z] = \frac{\int y\, d\nu(x,y)}{\phi(1)},$$

$$\mathbb{E}[Z^2] = \frac{2\int e^{-x} y\, d\nu(x,y) \int y\, d\nu(x,y)}{\phi(1)\phi(2)} + \frac{\int y^2\, d\nu(x,y)}{\phi(2)},$$

$$\mathbb{E}[Z^3] = \frac{6\int y\, d\nu(x,y) \int e^{-x} y\, d\nu(x,y) \int e^{-2x} y\, d\nu(x,y)}{\phi(1)\phi(2)\phi(3)} + \frac{\int y^3\, d\nu(x,y)}{\phi(3)}$$
$$+ \frac{3\int y^2\, d\nu(x,y) \int e^{-2x} y\, d\nu(x,y)}{\phi(2)\phi(3)} + \frac{3\int y\, d\nu(x,y) \int e^{-x} y^2\, d\nu(x,y)}{\phi(1)\phi(3)}.$$

Indeed, by Lemma 4.1,

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{\Delta X_t \neq 0} e^{-\eta_{t-}} \Delta\xi_t\right] = \mathbb{E}\left[\int_0^\infty \int e^{-\eta_t} y\, d\nu(x,y) dt\right] = \frac{\int y\, d\nu(x,y)}{\phi(1)}.$$

For the higher moments we use linearization, as in Sect. 3, e.g., we obtain

$$\mathbb{E}[Z^2] = \mathbb{E}\left[\left(\sum_{\Delta X_t \neq 0} e^{-\eta_{t-}} \Delta\xi_t\right)^2\right] = \mathbb{E}\left[\left(\sum_{\Delta X_s \neq 0} e^{-\eta_{s-}} \Delta\xi_s\right)\left(\sum_{\Delta X_t \neq 0} e^{-\eta_{t-}} \Delta\xi_t\right)\right]$$
$$= \mathbb{E}\left[2 \sum_{\substack{s<t \\ \Delta X_s, \Delta X_t \neq 0}} e^{-\eta_{s-} - \eta_{t-}} \Delta\xi_s \Delta\xi_t\right] + \mathbb{E}\left[\sum_{\Delta X_t \neq 0} (e^{-\eta_{t-}} \Delta\xi_t)^2\right] = 2\mathrm{I} + \mathrm{II},$$

where, by Corollary 4.4,

$$\mathrm{I} = \mathbb{E}\left[\int_0^\infty \int_s^\infty \int\int e^{-\eta_s} y_1 e^{-\eta_t - x_1} y_2\, d\nu(x_1,y_1)\, d\nu(x_2,y_2) dt\, ds\right]$$
$$= \int\int e^{-x_1} y_1 y_2\, d\nu(x_1,y_1)\, d\nu(x_2,y_2)\, \mathbb{E}\left[\int_0^\infty \int_s^\infty e^{-(\eta_t - \eta_s) - 2\eta_s} dt\, ds\right]$$
$$= \int e^{-x} y\, d\nu(x,y) \int y\, d\nu(x,y) \int_0^\infty \int_s^\infty \mathbb{E}\left[e^{-\eta_{t-s}}\right] \mathbb{E}\left[e^{-2\eta_s}\right] dt\, ds$$
$$= \frac{\int e^{-x} y\, d\nu(x,y) \int y\, d\nu(x,y)}{\phi(1)\phi(2)},$$

and

$$\mathrm{II} = \mathbb{E}\left[\int_0^\infty \int y^2 e^{-2\eta_t}\, d\nu(x,y)dt\right] = \frac{\int y^2\, d\nu(x,y)}{\phi(2)}.$$

The third and the higher moments are obtained analogously. We note that [2, Theorem 3.1] gives the first and the second moments of Z, but not the higher moments, which are cumbersome to obtain by the methods of [2] (private communication). Our approach also gives moments of anticipating integrals like

$$Y := \int_0^\infty e^{-\eta_t} d\xi_t = \sum_{\Delta X_t \neq 0} e^{-\eta_{t-} - \Delta\eta_t} \Delta\xi_t.$$

Here, similar calculations as for Z yield

$$\mathbb{E}[Y] = \frac{\int e^{-x} y\, d\nu(x,y)}{\phi(1)},$$

and higher moments of Y can be obtained analogously. We notice the difference between the formulas for the expectations of Z and Y.

Acknowledgements The inspiration to study Lévy systems came from the joint work of the first named author with Rodrigo Bañuelos [1], and our interest in moment formulas is due to the work of Nicolas Privault [21]. The present paper is based in part on the PhD dissertation of the fourth named author [24]. We thank Mateusz Kwaśnicki for many discussions and suggestions, Jean Jacod and Alexandre Popier for references to literature, Anita Behme for discussions on the Lévy integral in Sect. 4.2, Nicolas Privault for a discussion of our results and Aleksander Janicki and Rodrigo Bañuelos for their inspirations. K. Bogdan and Ł. Wojciechowski were supported in part by NCN grant 2012/07/B/ST1/03356. J. Rosiński's research was partially supported by a grant # 281440 from the Simons Foundation.

Appendix

The following Mecke-Palm identity holds for 1-processes $f(x;\omega) \geq 0$ [20],

$$\mathbb{E}\int_{\mathcal{X}} f(x;\omega)\, \omega(dx) = \int_{\mathcal{X}} \mathbb{E} f(x;\omega \cup \{x\})\, \sigma(dx). \tag{1}$$

For the reader's convenience we give a direct proof of (1) in the setting of Sect. 2. We first consider $\sigma(\mathcal{X}) < \infty$ and nonnegative process $f(x;\omega) = f(x;\omega \cap \mathcal{X})$, i.e. depending only on $\mathcal{X}$. If $\omega = \{y_1, \dots, y_n\}$, a set with n elements, then

$$\int_{\mathcal{X}} f(x;\omega)\omega(dx) = \sum_{i=1}^n f_{(n)}(y_i; y_1, \dots, y_n).$$

The above quantity is invariant upon permutations of $y_1, \dots, y_n$, in fact it is the n-th coefficient of the random variable $\int_{\mathcal{X}} f(x;\omega)\omega(dx)$. By (2.1), the left-hand side of (1) equals

$$e^{-\sigma(\mathcal{X})} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{i=1}^{n} \int_{\mathcal{X}^n} f_{(n)}(y_i; y_1, \dots, y_n)\sigma(dy_1)\cdots\sigma(dy_n). \tag{2}$$

If $\omega = \{y_1, \dots, y_n\}$, a set with n elements, and $x \notin \omega$, then

$$\begin{aligned} f(x;\omega \cup \{x\}) &= f_{(n+1)}(x; x, y_1, \dots, y_n) = \dots = f_{(n+1)}(x; y_1, \dots, y_n, x) \\ &= \frac{1}{n+1} \sum_{i=1}^{n+1} f_{(n+1)}(x; y_1, \dots, y_{i-1}, x, y_i, \dots, y_n). \end{aligned}$$

Since σ is non-atomic, we have $\mathbb{P}(x \in \omega) = 0$ and so $\{x\} \cup \omega$ has $n+1$ elements if ω has n, for almost all $x \in \mathcal{X}$, cf. (2.1). Therefore, by (2.1), the right-hand side of (1) equals

$$\begin{aligned} &\mathbb{E} \int_{\mathcal{X}} f(x; \omega \cup \{x\})\sigma(dx) \\ &= e^{-\sigma(X)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{X}} \int_{\mathcal{X}^n} \frac{1}{n+1} \sum_{i=1}^{n+1} f_{(n+1)}(x; y_1, \dots, y_n, x)\sigma(dy_1)\cdots\sigma(dy_n)\sigma(dx). \end{aligned}$$

This verifies (1) when $\sigma(\mathcal{X}) < \infty$, e.g., if $\mathcal{X} \subset \mathbb{X}$ is compact; we note in passing that (2) is an explicit representation of either side of (1).

We next let $\mathbb{X} = \bigcup_m \mathcal{X}_m$ be a countable decomposition of $\mathbb{X}$ into disjoint Borel sets with $\sigma(\mathcal{X}_m) < \infty$. For arbitrary process $f(x;\omega) \geq 0$ we have

$$\int_{\mathbb{X}} f(x;\omega)\omega(dx) = \sum_m \int_{\mathcal{X}_m} f(x;\omega)\omega(dx). \tag{3}$$

For fixed m, we write $\omega_* = \omega \cap \mathcal{X}_m$, $\omega^* = \omega \setminus \mathcal{X}_m$, and denote by $\mathbb{E}_*$ and $\mathbb{E}^*$ the expectation $\mathbb{E}$ when restricted to random variables depending only on $\mathcal{X}_m$ and $\mathbb{X} \setminus \mathcal{X}_m$, respectively. By (2.2) and by (1) for $\mathcal{X}_m$,

$$\begin{aligned} &\mathbb{E} \int_{\mathcal{X}_m} f(x;\omega)\omega(dx) = \mathbb{E}^*\mathbb{E}_* \int_{\mathcal{X}_m} f(x; \omega_* \cup \omega^*)\omega_*(dx) \\ &= \mathbb{E}^* \int_{\mathcal{X}_m} \mathbb{E}_* f(x; \omega_* \cup \{x\} \cup \omega^*)\sigma(dx) = \int_{\mathcal{X}_m} \mathbb{E} f(x; \omega \cup \{x\})\sigma(dx). \end{aligned}$$

This yields (1) in the general case, cf. (3).

Needless to say, (1) also holds for signed processes f under the assumption of absolute integrability, because we can decompose both sides of (1) according to $f = f_+ - f_-$, where $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$.

References

1. R. Bañuelos, K. Bogdan, Lévy processes and Fourier multipliers. J. Funct. Anal. **250**(1), 197–213 (2007)
2. A.D. Behme, Distributional properties of solutions of $\mathrm{d}V_T = V_{T-}\mathrm{d}U_T + \mathrm{d}L_T$ with Lévy noise. Adv. Appl. Probab. **43**(3), 688–711 (2011)
3. A. Benveniste, J. Jacod, Systèmes de Lévy des processus de Markov. Invent. Math. **21**, 183–198 (1973)
4. R.M. Blumenthal, *Excursions of Markov Processes*. Probability and Its Applications (Birkhäuser, Boston, 1992)
5. K. Bogdan, L. Wojciechowski, Parabolic martingales and non-symmetric Fourier multipliers. Probab. Math. Stat. **32**(2), 241–253 (2012)
6. J.-C. Breton, N. Privault, Factorial moments of point processes. Stoch. Process. Appl. **124**(10), 3412–3428 (2014)
7. L. Decreusefond, I. Flint, Moment formulae for general point processes. J. Funct. Anal. **267**(2), 452–476 (2014)
8. C. Dellacherie, P.-A. Meyer, *Probabilities and Potential. B* (North-Holland, Amsterdam, 1982)
9. N. Ikeda, S. Watanabe, On some relations between the harmonic measure and the Lévy measure for a certain class of Markov processes. J. Math. Kyoto Univ. **2**, 79–95 (1962)
10. N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd edn. (North-Holland, Amsterdam, 1989)
11. J. Jacod, A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd edn. (Springer, Berlin, 2003)
12. O. Kallenberg, *Foundations of Modern Probability*. Probability and Its Applications, 2nd edn. (Springer, New York, 2002)
13. O. Kallenberg, J. Szulga, Multiple integration with respect to Poisson and Lévy processes. Probab. Theory Relat. Fields **83**(1–2), 101–134 (1989)
14. S. Kwapień, W.A. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*. Probability and Its Applications (Birkhäuser, Boston, 1992)
15. G. Last, M. Penrose, Poisson process fock space representation, chaos expansion and covariance inequalities. Probab. Theory Relat. Fields **150**, 663–690 (2011)
16. G. Last, M.D. Penrose, M. Schulte, C. Thäle, Moments and central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**(2), 348–364 (2014)
17. W. Linde, *Probability in Banach Spaces—Stable and Infinitely Divisible Distributions*. A Wiley-Interscience Publication, 2nd edn. (Wiley, Chichester, 1986)
18. A. Lindner, R. Maller, Lévy integrals and the stationarity of generalised Ornstein-Uhlenbeck processes. Stoch. Process. Appl. **115**(10), 1701–1722 (2005)
19. G. Peccati, M.S. Taqqu, *Wiener Chaos: Moments, Cumulants and Diagrams*. Bocconi & Springer Series, vol. 1 (Springer, Milan/Bocconi University Press, Milan, 2011). A survey with computer implementation, Supplementary material available online
20. N. Privault, *Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales*. Lecture Notes in Mathematics, vol. 1982 (Springer, Berlin, 2009)
21. N. Privault, Moments of Poisson stochastic integrals with random integrands. Probab. Math. Stat. **32**(2), 227–239 (2012)

22. K. Sato, *Lévy Processes and Infinitely Divisible Distributions* (Cambridge University Press, Cambridge, 1999)
23. R.L. Schilling, R. Song, Z. Vondraček, *Bernstein Functions: Theory and Applications*. de Gruyter Studies in Mathematics, vol. 37 (Walter de Gruyter, Berlin, 2010)
24. Ł. Wojciechowski, Stochastic analysis of Lévy processes. PhD dissertation, University of Wrocław, 2014

Conformal Transforms and Doob's *h*-Processes on Heisenberg Groups

Jing Wang

Abstract We study the stochastic processes that are images of Brownian motions on Heisenberg group $\mathbf{H}^{2n+1}$ under conformal maps. In particular, we obtain that Cayley transform maps Brownian paths in $\mathbf{H}^{2n+1}$ to a time changed Brownian motion on CR sphere $\mathbb{S}^{2n+1}$ conditioned to be at its south pole at a random time. We also obtain that the inversion of Brownian motion on $\mathbf{H}^{2n+1}$ started from $x \neq 0$, is up to time change, a Brownian bridge on $\mathbf{H}^{2n+1}$ conditioned to be at the origin.

Keywords Brownian bridge • Cayley transform • Doob's h-process • Heisenberg group • Kelvin transform

1 Introduction

The Brownian motions on sub-Riemannian model spaces has been widely studied in recent years. Due to strong symmetries of the model spaces, explicit computations analysis can be conducted (see [1–4, 7]). In this paper we focus on the relationships between Brownian motion on Heisenberg group and its images under certain conformal maps, namely Cayley transform and Kelvin transform.

Let $\mathbf{H}^{2n+1}$ be a $2n+1$ dimensional Heisenberg group that lives in $\mathbb{C}^n \times \mathbb{R}$ with coordinates $(z,t) = (z_1, \dots, z_n, t)$ where $z_j = x_j + iy_j$. It has the group law

$$(z,t)(z',t') = (z+z', t+t' + \mathbf{Im} z\bar{z'}).$$

It is a flat model space of sub-Riemannian manifolds. There is a canonical sub-Laplacian on $\mathbf{H}^{2n+1}$:

$$\bar{L}_{\mathbf{H}^{2n+1}} = \sum_{j=1}^{n} \left(\frac{\partial^2}{\partial x_j^2} + \frac{\partial^2}{\partial y_j^2} + 2y_j \frac{\partial^2}{\partial x_j \partial t} - 2x_j \frac{\partial^2}{\partial y_j \partial t} + |z_j|^2 \frac{\partial^2}{\partial t^2} \right)$$

J. Wang (✉)
Department of Mathematics, University of Illinois at Urbana Champaign, 1409 W. Green Street, Urbana, IL, USA
e-mail: wangjing@illinois.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_8

The Brownian motion on $\mathbf{H}^{2n+1}$ issued from $x' \in \mathbf{H}^{2n+1}$ is the strong Markov process that is generated by $\frac{1}{2}\bar{L}_{\mathbf{H}^{2n+1}}$.

Cayley transform is known to be a bi-holomorphic map between the Siegel domain Ω^{n+1} and a unit ball in $\mathbb{C}^{n+1}$. The restriction of Cayley transform on its boundary therefore provides a conformal map between $\mathbf{H}^{2n+1}$ and the unit sphere $\mathbb{S}^{2n+1}$ in $\mathbb{C}^{n+1}$. If we consider the image of a Brownian path on $\mathbf{H}^{2n+1}$ under Cayley transform, it then turns out to be a $\mathbb{S}^{2n+1}$-valued process. In particular, it is a time changed version of a Brownian path on $\mathbb{S}^{2n+1}$ conditioned to be at the south pole at a random time. Below we state our main result.

Theorem 1.1 *The Brownian motion on $\mathbf{H}^{2n+1}$ issued from x' is mapped by Cayley transform $\mathcal{C}_1$ to a time-changed Brownian motion on $\mathbb{S}^{2n+1}$ issued from $x = \mathcal{C}_1(x')$ and conditioned to be at the south pole $-e_n$ at time T, where T is an independent random variable with distribution*

$$\mathbb{P}_x^h[T > t] = \frac{\int_t^{+\infty} e^{-n^2 s} p_s(-e_n, x)ds}{\int_0^{+\infty} e^{-n^2 t} p_t(-e_n, x)dt}. \tag{1.1}$$

Here $p_t(x, y)$ denotes the subelliptic heat kernel on $\mathbb{S}^{2n+1}$.
This result extends the result by Carne in [5], where he proved that the Stereographic projection from $\mathbb{R}^n$ to S^n maps Brownian paths in $\mathbb{R}^n$ to the paths of conditioned Brownian motion on S^n.

Another object of our study is to probabilistically interpret the relation between the Brownian motion on $\mathbf{H}^{2n+1}$ started from any $x' \neq 0$ and its image under the inversion map, namely the Kelvin transform. This type of question was first posed by Schwartz (see [10]), who asked how Brownian motion in $\mathbb{R}^n$ can be interpreted as a Brownian bridge conditioned to be at the “ideal point at infinity”. A probabilistic approach was provided by Yor in [11]. In the present paper, we obtain the result in a setting of a flat sub-Riemannian manifold. The inversion of Brownian motion on $\mathbf{H}^{2n+1}$ issued from $x \neq 0$ turns out to be a Brownian bridge conditioned to be at the origin up to time change.

Theorem 1.2 *The Brownian motion on $\mathbf{H}^{2n+1}$ generated by $\frac{1}{2}L_{\mathbf{H}^{2n+1}}$ and issued from $x' \neq 0$ is mapped by Kelvin transform to a time-changed $\mathbf{H}^{2n+1}$-valued Brownian motion conditioned to be at the origin at $t = \infty$.*

The approaches to both results follow the idea of Carne. By analyzing the radial part of the corresponding conformal sub-Laplacians on $\mathbb{S}^{2n+1}$ and on $\mathbf{H}^{2n+1}$, we are able to obtain the relationship between Markov processes that are generated by $\frac{1}{2}L_{\mathbb{S}^{2n+1}}$ and $\frac{1}{2}L_{\mathbf{H}^{2n+1}}$ respectively through an argument of Doob’s h-processes.

In the next section, we deduce Theorem 1.1 after a detailed discussion of Cayley transform and radial process or Brownian motions on $\mathbb{S}^{2n+1}$ and $\mathbf{H}^{2n+1}$. In Sect. 3 we focus on the inverse transform on $\mathbf{H}^{2n+1}$ and the proof of Theorem 1.2.

2 Cayley Transformation and Doob's *h*-Process

2.1 Cayley Transform on CR Model Spaces

Cayley transforms on CR model spaces are natural analogues of stereographic projections on Riemannian models. Let $B^{n+1} = \{\zeta \in \mathbb{C}^{n+1} : |\zeta| < 1\}$ be the unit ball in $\mathbb{C}^{n+1}$ and $\Omega^{2n+1} = \{(z, w) \in \mathbb{C}^n \times \mathbb{C}, \mathbf{Im}(w) > |z|^2\}$ the Siegel domain. The Cayley transform $\mathcal{C} : B^{2n+1} \to \Omega^{n+1}$ is a biholomorphic map such that (see [6])

$$\mathcal{C} : (\zeta_1, \ldots, \zeta_{n+1}) \to \left(\frac{\zeta_1}{1+\zeta_{n+1}}, \ldots, \frac{\zeta_n}{1+\zeta_{n+1}}, i\frac{1-\zeta_{n+1}}{1+\zeta_{n+1}} \right), \quad \zeta^{n+1} \neq -1.$$

Let $\mathbb{S}^{2n+1} = \{\zeta \in \mathbb{C}^{n+1}, |\zeta| = 1\}$ be the unit sphere in $\mathbb{C}^{n+1}$. It also appears as a model space of CR manifolds. The restriction of $\mathcal{C}$ to the CR sphere $\mathbb{S}^{2n+1}$ minus a point gives a CR diffeomorphism to the boundary of the Siegel domain $\partial\Omega^{2n+1}$, which may be identified with the Heisenberg group $\mathbf{H}^{2n+1}$ through the CR isomorphism $\varphi : \mathbf{H}^{2n+1} \to \partial\Omega^{2n+1}$. For any $(z, t) \in \mathbf{H}^{2n+1}$,

$$\varphi(z, t) = (z, 2t + i|z|^2). \tag{2.2}$$

We denote the north pole of $\mathbb{S}^{2n+1}$ by $e_n = \{0, \ldots, 0, 1\}$ and denote the south pole by $-e_n$. Now we consider the CR equivalence between Heisenberg group and CR sphere minus the south pole $\mathcal{C}_1 : \mathbf{H}^{2n+1} \to \mathbb{S}^{2n+1} \backslash \{-e_n\}$. It is then given by $\mathcal{C}_1 = \mathcal{C}^{-1} \circ \varphi$. In local coordinates we have for any $(z, t) = (z_1, \ldots, z_n, t) \in \mathbf{H}^{2n+1}$,

$$\mathcal{C}_1 : (z, t) \to \left(\frac{2z_1}{(1+|z|^2) - 2it}, \ldots, \frac{2z_n}{(1+|z|^2) - 2it}, \frac{1-|z|^2+2it}{1+|z|^2-2it} \right). \tag{2.3}$$

It is a conformal map with inverse $\mathcal{C}_1^{-1} : \mathbb{S}^{2n+1} \backslash \{-e_n\} \to \mathbf{H}^{2n+1}$,

$$\mathcal{C}_1^{-1} : (\zeta_1, \cdots \zeta_{n+1}) \to \left(\frac{\zeta_1}{1+\zeta_{n+1}}, \ldots, \frac{\zeta_n}{1+\zeta_{n+1}}, \frac{i}{2} \frac{\overline{\zeta_{n+1}} - \zeta_{n+1}}{|1+\zeta_{n+1}|^2} \right). \tag{2.4}$$

Since $\mathbb{S}^{2n+1}$ is a model space of sub-Riemannian manifold with the Hopf fibration $\mathbb{S}^1 \to \mathbb{S}^{2n+1} \to \mathbb{CP}^n$, it is more convenient for us to use the so-called cylindrical coordinates that carries the structural information and are given by

$$(w, \theta) \to \frac{e^{i\theta}}{\sqrt{1+|w|^2}} (w, 1),$$

where $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, and $w = \zeta/\zeta_{n+1} \in \mathbb{CP}^n$. Here $w = (w_1, \cdots, w_n)$ parametrizes the complex lines passing through the origin, and θ determines a point on the line that is of unit distance from the north pole. Let $|w| = \tan r_S$, $r_S \in [0, \pi/2)$, then we have $\mathcal{C}_1^{-1}$ in cylindrical coordinates given by

$$\mathcal{C}_1^{-1} : \left(\frac{e^{i\theta}}{\sqrt{1 + |w|^2}} (w, 1) \right)$$
$$\rightarrow \left(\frac{e^{i\theta} \cos r_S + \cos^2 r_S}{1 + \cos^2 r_S + 2 \cos r_S \cos \theta} w, \frac{\cos r_S \sin \theta}{1 + \cos^2 r_S + 2 \cos r_S \cos \theta} \right).$$

Let $\psi_S : \mathbb{S}^{2n+1} \rightarrow [0, \pi/2) \times \mathbb{R}/2\pi\mathbb{Z}$ be such that

$$\psi_S \left(\frac{e^{i\theta}}{\sqrt{1 + |w|^2}} (w, 1) \right) = (r_S, \theta)$$

and $\psi_H : \mathbf{H}^{2n+1} \rightarrow \mathbb{R}_{\geq 0} \times \mathbb{R}$ be such that

$$\psi_H (z, t) = (r_H, t),$$

where $r_H = \sqrt{\sum_{j=1}^n |z_j|^2}$. We define a map $\mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow [0, \pi/2) \times \mathbb{R}/2\pi\mathbb{Z}$ by the chart below, and by abusing of notation we denote it by $\mathcal{C}_1$:

$$\begin{array}{ccc} \mathbf{H}^{2n+1} & \xrightarrow{\mathcal{C}_1} & \mathbb{S}^{2n+1} \\ \downarrow \psi_H & & \downarrow \psi_S \\ \mathbb{R}_{\geq 0} \times \mathbb{R} & \xrightarrow{\mathcal{C}_1} & [0, \pi/2) \times \mathbb{R}/2\pi\mathbb{Z} \end{array}$$

We easily compute that

$$\mathcal{C}_1 : (r_H, t)$$
$$\rightarrow \left(\arcsin \left(\frac{2r_H}{\sqrt{(1 + r_H^2)^2 + 4t^2}} \right), \arcsin \left(\frac{4t}{\sqrt{(1 + r_H^2)^2 + 4t^2} \sqrt{(1 - r_H^2)^2 + 4t^2}} \right) \right)$$

and

$$\mathcal{C}_1^{-1} : (r_S, \theta) \rightarrow \left(\frac{\sin r_S}{\sqrt{1 + \cos^2 r_S + 2 \cos r_S \cos \theta}}, \frac{\cos r_S \sin \theta}{1 + \cos^2 r_S + 2 \cos r_S \cos \theta} \right).$$

2.2 Brownian Motion and Doob's h-Process

Now we consider the Markov processes that are generated by sub-Laplacians $\bar{L}_{\mathbf{H}^{2n+1}}$ and $\bar{L}_{\mathbb{S}^{2n+1}}$, which are referred to as Brownian motions on $\mathbf{H}^{2n+1}$ and $\mathbb{S}^{2n+1}$ respectively throughout this paper. Due to the radial symmetries of these diffusion processes, it is sufficient for us to consider only the radial part of the sub-Laplacians.

We denote by $L_{\mathbf{H}^{2n+1}}$ the radial part of the sub-Laplacian on $\mathbf{H}^{2n+1}$ in coordinates (r_H, t), it is defined on the space $D_H = \{f \in C^\infty(\mathbb{R}_{\geq 0} \times \mathbb{R}, \mathbb{R}), \frac{\partial f}{\partial r_H}|_{r_H=0} = 0\}$. Let $L_{\mathbb{S}^{2n+1}}$ be the radial part of $\bar{L}_{\mathbb{S}^{2n+1}}$ in cylindric coordinates (r_S, θ), with domain $D_S = \{f \in C^\infty([0, \frac{\pi}{2}) \times \mathbb{R}/2\pi\mathbb{Z}, \mathbb{R}), \frac{\partial f}{\partial r_S}|_{r_S=0} = 0\}$. Then for any $f \in D_H$ and $g \in D_S$, we have

$$\bar{L}_{\mathbf{H}^{2n+1}}(f \circ \psi_H) = (L_{\mathbf{H}^{2n+1}} f) \circ \psi_H, \quad \bar{L}_{\mathbb{S}^{2n+1}}(g \circ \psi_S) = (L_{\mathbb{S}^{2n+1}} g) \circ \psi_S.$$

It is known that $L_{\mathbb{S}^{2n+1}}$ is essentially self-adjoint with respect to the volume measure $d\mu_{\mathbb{S}^{2n+1}} = \frac{2\pi^n}{\Gamma(n)}(\sin r_S)^{2n-1} \cos r_S dr_S d\theta$ on $\mathbb{S}^{2n+1}$, and $L_{\mathbf{H}^{2n+1}}$ is essentially self-adjoint with respect to the volume measure $d\mu_{\mathbf{H}^{2n+1}} = \frac{2\pi^n}{\Gamma(n)} r_H^{2n-1} dr_H dt$ on $\mathbf{H}^{2n+1}$. Moreover, we have explicitly

$$L_{\mathbf{H}^{2n+1}} = \frac{\partial^2}{\partial r_H^2} + \frac{2n-1}{r_H}\frac{\partial}{\partial r_H} + r_H^2 \frac{\partial^2}{\partial t^2} \tag{2.5}$$

and (see [1, 2], also [8])

$$L_{\mathbb{S}^{2n+1}} = \frac{\partial^2}{\partial r_S^2} + ((2n-1)\cot r_S - \tan r_S)\frac{\partial}{\partial r_S} + \tan^2 r_S \frac{\partial^2}{\partial \theta^2}. \tag{2.6}$$

Let us consider Green function of the conformal sub-Laplacian $-L_{\mathbb{S}^{2n+1}} + n^2$ with pole $(0, 0)$ (the north pole of $\mathbb{S}^{2n+1}$) and denote it by $G_{\mathbb{S}^{2n+1}}$. From [2] we have

$$G_{\mathbb{S}^{2n+1}}((0,0),(r_S,\theta)) = \frac{\Gamma\left(\frac{n}{2}\right)^2}{8\pi^{n+1}(1 - 2\cos r_S \cos\theta + \cos^2 r_S)^{n/2}}. \tag{2.7}$$

On the other hand the Green function of $-L_{\mathbf{H}^{2n+1}}$ with respect to $d\mu_{\mathbf{H}^{2n+1}}$ is given by

$$G_{\mathbf{H}^{2n+1}}((0,0),(r_H,t)) = \frac{\Gamma\left(\frac{n}{2}\right)^2}{8\pi^{n+1}(r_H^4 + 4t^2)^{n/2}} \tag{2.8}$$

We consider $h \in D_S$, such that for any $(r_S, \theta) \in [0, \frac{\pi}{2}) \times \mathbb{R}/2\pi\mathbb{Z}$,

$$h(r_S, \theta) = 1 + 2\cos r_S \cos\theta + \cos^2 r_S, \tag{2.9}$$

and $H \in D_H$, such that for any $(r_H, t) \in \mathbb{R}_{\geq 0} \times \mathbb{R}$,

$$H(r_H, t) = \frac{4}{(1 + r_H^2)^2 + 4t^2}. \tag{2.10}$$

It is an easy fact that h and H are harmonic functions with poles $(0, \pi)$ and $(0, 0)$ respectively. Moreover, we have

$$H = \mathcal{C}_1^* h = h \circ \mathcal{C}_1.$$

From (2.7) and (2.8) we can easily observe that

$$G_{\mathbb{S}^{2n+1}}((0,0),(r_S,\theta))(1+2\cos r_S \cos\theta + \cos^2 r_S)^{\frac{n}{2}} = (\mathcal{C}_1^{-1*} G_{\mathbf{H}^{2n+1}})((0,0),(r_S,\theta)).$$

In fact, for any $x, y \in [0, \frac{\pi}{2}) \times \mathbb{R}/2\pi\mathbb{Z}$ we have

$$G_{\mathbb{S}^{2n+1}}(x, y) = (\mathcal{C}_1^{-1*} G_{\mathbf{H}^{2n+1}})(x, y) h(x)^{-\frac{n}{2}} h(y)^{-\frac{n}{2}}. \tag{2.11}$$

From this we can then deduce the relation between $L_{\mathbf{H}^{2n+1}}$ and $L_{\mathbb{S}^{2n+1}} - n^2$.

Theorem 2.1 *For any function $f \in D_S$, the relation of $L_{\mathbf{H}^{2n+1}}$ and $L_{\mathbb{S}^{2n+1}} - n^2$ via Cayley transform is given by*

$$h^{(\frac{n}{2}+1)}(-L_{\mathbb{S}^{2n+1}} + n^2)\left(h^{-\frac{n}{2}} f\right) = -(\mathcal{C}_{1*} L_{\mathbf{H}^{2n+1}}) f \tag{2.12}$$

where h is as in (2.9).

Proof For any $f \in D_S$, let $F \in D_H$ be such that $F = (\mathcal{C}_1)^* f = f \circ \mathcal{C}_1$. We assume for some $\sigma_1, \sigma_2 \in D_S$ it holds that for any $x \in [0, \frac{\pi}{2}) \times \mathbb{R}/2\pi\mathbb{Z}$,

$$(-L_{\mathbb{S}^{2n+1}} + n^2)\,(\sigma_1 f)\,|_x = -\sigma_2\,(L_{\mathbf{H}^{2n+1}})\,(\mathcal{C}_1^* f)|_{\mathcal{C}_1^{-1}(x)}.$$

It then amounts to find σ_1, σ_2. Let $g = -L_{\mathbf{H}^{2n+1}} F$, then $F = (-L_{\mathbf{H}^{2n+1}})^{-1} g$. The above equation is equivalent to

$$\sigma_1 \cdot \left((-L_{\mathbf{H}^{2n+1}})^{-1} g\right) \circ \mathcal{C}_1^{-1} = (-L_{\mathbb{S}^{2n+1}} + n^2)^{-1}(\sigma_2(g \circ \mathcal{C}_1^{-1})). \tag{2.13}$$

Therefore, for all $x \in [0, \frac{\pi}{2}) \times \mathbb{R}/2\pi\mathbb{Z}$, we have

$$\int G_{\mathbf{H}^{2n+1}}(\mathcal{C}_1^{-1}(x), v) g(v) d\mu_{\mathbf{H}^{2n+1}} v = \sigma_1^{-1}(x) \int G_{\mathbb{S}^{2n+1}}(x, y) \sigma_2(y) g(\mathcal{C}_1^{-1}(y)) d\mu_{\mathbb{S}^{2n+1}} y \tag{2.14}$$

where $G_{\mathbb{S}^{2n+1}}$ and $G_{\mathbf{H}^{2n+1}}$ are Green functions as in (2.7) and (2.8). Moreover by changing variable $y = \mathcal{C}_1(v)$, the right hand side of the above equation writes

$$\sigma_1^{-1}(x) \int G_{\mathbb{S}^{2n+1}}(x, \mathcal{C}_1(v))\sigma_2(\mathcal{C}_1(v))g(v)|J_{\mathcal{C}_1}(v)|d\mu_{\mathbf{H}^{2n+1}}v, \tag{2.15}$$

where $|J_{\mathcal{C}_1}(v)|$ is the Jacobi determinant. We can easily compute that

$$|J_{\mathcal{C}_1}(v)| = H^{n+1}(v),$$

where H is given as in (2.10). Therefore (2.15) becomes

$$\sigma_1^{-1}(x) \int G_{\mathbb{S}^{2n+1}}(x, \mathcal{C}_1(v))\sigma_2(\mathcal{C}_1(v))g(v)H^{n+1}(v)d\mu_{\mathbf{H}^{2n+1}}v.$$

By plugging in (2.11) and comparing to (2.14), we obtain for all $x, y \in \mathbb{S}^{2n+1}$

$$\begin{cases} \sigma_1(x) = h^{-\frac{n}{2}}(x) \\ \sigma_2(y) = h^{-(1+\frac{n}{2})}(y), \end{cases}$$

□

hence the conclusion.

Corollary 2.2 *For any function $f \in D_S$, we have that*

$$(\mathcal{C}_{1*}L_{\mathbf{H}^{2n+1}})f = h\left(L_{\mathbb{S}^{2n+1}}f + \frac{2\Gamma_{\mathbb{S}^{2n+1}}(h^{-\frac{n}{2}}, f)}{h^{-\frac{n}{2}}}\right) \tag{2.16}$$

where $\Gamma_{\mathbb{S}^{2n+1}}(f, g) = \frac{1}{2}(L_{\mathbb{S}^{2n+1}}(fg) - fL_{\mathbb{S}^{2n+1}}g - gL_{\mathbb{S}^{2n+1}}f)$ *for any* $f, g \in D_S$.

Proof Notice that

$$(L_{\mathbb{S}^{2n+1}} - n^2)(h^{-\frac{n}{2}}) = 0.$$

hence

$$h^{\frac{n}{2}}(L_{\mathbb{S}^{2n+1}} - n^2)(h^{-\frac{n}{2}}f) = L_{\mathbb{S}^{2n+1}}f + 2h^{\frac{n}{2}}\Gamma_{\mathbb{S}^{2n+1}}(h^{-\frac{n}{2}}, f).$$

□

Now we are ready to prove the main result.

Proof of Theorem 1.1 The proof follows two steps.

Step 1 Notice that $h^{-\frac{n}{2}}$ is the Green function of the conformal sub-Laplacian $L_{\mathbb{S}^{2n+1}} - n^2$ with pole $(\pi/2, 0)$ (the south pole $-e_n$ of $\mathbb{S}^{2n+1}$). For any $f \in D_S$ we let

$$L^h f := L_{\mathbb{S}^{2n+1}} f + \frac{2\Gamma_{\mathbb{S}^{2n+1}}(h^{-\frac{n}{2}}, f)}{h^{-\frac{n}{2}}} = \frac{L_{\mathbb{S}^{2n+1}}(h^{-\frac{n}{2}} f)}{h^{-\frac{n}{2}}} - n^2 f. \tag{2.17}$$

Let X_t^h and X_t be Markov processes generated by $\frac{1}{2}L^h$ and $\frac{1}{2}L_{\mathbb{S}^{2n+1}}$, issued from $x \in \mathbb{S}^{2n+1}$. We first prove that X_t^h is X_t conditioned to be at the south pole $-e_n$ at time T, where T is a random time with distribution (1.1).

It is sufficient to prove that for any $f \in D_S$,

$$\mathbb{E}_x\left[f(X_t^h)\right] = \mathbb{E}_x\left[f(X_t)\mathbb{1}_{t<T} | X_T = -e_n\right] \tag{2.18}$$

Let P_t^h and P_t be the heat semigroups generated by L^h and $L_{\mathbb{S}^{2n+1}}$ respectively, then by iterating (2.17) it is not hard to obtain for any $x \in \mathbb{S}^{2n+1}$,

$$P_t^h(f(x)) = h(x)^{\frac{n}{2}} e^{-tn^2} P_t(h^{-\frac{n}{2}}(x) f(x)),$$

that is

$$\mathbb{E}_x\left[f(X_t^h)\right] = \frac{1}{h^{-\frac{n}{2}}(x)} e^{-tn^2} \mathbb{E}_x\left[h^{-\frac{n}{2}}(X_t) f(X_t)\right] = \mathbb{E}_x\left[\frac{e^{-tn^2} h^{-\frac{n}{2}}(X_t)}{h^{-\frac{n}{2}}(x)} f(X_t)\right].$$

Proving (2.18) is then equivalent to proving

$$\mathbb{E}_x\left[f(X_t)\mathbb{1}_{t<T} | X_T = -e_n\right] = \mathbb{E}_x\left[\frac{e^{-tn^2} h^{-\frac{n}{2}}(X_t)}{h^{-\frac{n}{2}}(x)} f(X_t)\right]. \tag{2.19}$$

Note that

$$\mathbb{E}_x\left[f(X_t)\mathbb{1}_{t<T} | X_T = -e_n\right] = \frac{\mathbb{E}_x\left[f(X_t)\mathbb{1}_{t<T}\mathbb{1}_{X_T=-e_n}\right]}{\mathbb{E}_x\left[X_T = -e_n\right]}.$$

Assume T is an exponential random variable with parameter $-n^2$ under the original probability measure, we have

$$\mathbb{E}_x\left[f(X_t)\mathbb{1}_{t<T}\mathbb{1}_{X_T=-e_n}\right] = \mathbb{E}_x\left[e^{-tn^2} h^{-\frac{n}{2}}(X_t) f(X_t)\right]$$

and

$$\mathbb{E}_x\left[X_T = -e_n\right] = \int_0^{+\infty} p_t(x, -e_n) e^{-n^2 t} dt = h^{-\frac{n}{2}}(x).$$

Thus (2.19) holds when T is an exponential random variable under the original probability measure. Switching to the conditioned probability measure, T then has the distribution

$$\mathbb{P}_x^h[T > t] = e^{-n^2 t}\frac{\mathbb{E}_x\left[h^{-\frac{n}{2}}(X_t)\right]}{h^{-\frac{n}{2}}(x)} = \frac{\int_t^{+\infty} e^{-n^2 s} p_s(-e_n, x)ds}{\int_0^{+\infty} e^{-n^2 t} p_t(-e_n, x)dt}.$$

Step 2 Next we prove the time change. Let Y_t be the Markov process generated by $\frac{1}{2}L_{\mathbf{H}^{2n+1}}$ and issued from $\mathcal{C}_1^{-1}(x)$, we claim that Y_t is mapped by Cayley transform to a time-changed version of X^h, i.e.,

$$X^h_{\mathcal{A}_t} = \mathcal{C}_1(Y_t) \tag{2.20}$$

where the time change is given by $\mathcal{A}_t = \int_0^t H(Y_s)^{-1} ds$. To see this, we consider for any $F = f \circ \mathcal{C}_1 \in D_H$, the associated martingale M_t^F that is given by

$$M_t^F = F(Y_t) - \frac{1}{2}\int_0^t L_{\mathbf{H}^{2n+1}} F(Y_s) ds.$$

By plugging in (2.20), (2.16) and (2.17) we have

$$M_t^F = f(X^h_{\mathcal{A}_t}) - \frac{1}{2}\int_0^t (L_{\mathbf{H}^{2n+1}} F) \circ \mathcal{C}_1^{-1}(X^h_{\mathcal{A}_s}) ds = f(X^h_{\mathcal{A}_t}) - \frac{1}{2}\int_0^t H(Y_s) L^h f(X^h_{\mathcal{A}_s}) ds.$$

Let σ_t be the hitting time such that $\sigma_t = \inf\{u, \mathcal{A}_u > t\}$, then clearly $\mathcal{A}_{\sigma_t} = t = \sigma_{\mathcal{A}_t}$. By changing variable $s = \sigma_u$ we obtain

$$M_t^F = f(X^h_{\mathcal{A}_t}) - \frac{1}{2}\int_0^{\sigma_{\mathcal{A}_t}} H(Y_s) L^h f(X^h_{\mathcal{A}_s}) ds = f(X^h_t) - \frac{1}{2}\int_0^{\mathcal{A}_t} H(Y_{\sigma_u}) L^h f(X^h_u) \sigma'_u du.$$

Note for any $u > 0$ we have $u = \mathcal{A}_{\sigma_u} = \int_0^{\sigma_u} H(Y_s) ds$. This implies that

$$1 = H(Y_{\sigma_u})\sigma'_u.$$

Therefore

$$M_t^F = f(X^h_t) - \frac{1}{2}\int_0^{\mathcal{A}_t} L^h f(X^h_u) du,$$

and it completes the proof.

3 Inversion of Brownian Motions on Heisenberg Group

In this section we consider the inversion of Brownian motion on Heisenberg group. First we construct the inverse map by composing two Cayley transforms $\mathcal{C}_1$ and $\mathcal{C}_2$, between $\mathbf{H}^{2n+1}$ and $\mathbb{S}^{2n+1}$ minus a point ($-e_n$ and e_n respectively). We have already discussed $\mathcal{C}_1$ in the previous section. Now let us consider $\mathcal{C}_2 : \mathbf{H}^{2n+1} \to \mathbb{S}^{2n+1} \backslash \{e_n\}$ where e_n is the north pole on $\mathbb{S}^{2n+1}$. We have

$$\mathcal{C}_2 : (z,t) \to \left(\frac{2z_1}{1+|z|^2+2it}, \dots, \frac{2z_n}{1+|z|^2+2it}, -\frac{1-|z|^2-2it}{1+|z|^2+2it} \right).$$

and

$$\mathcal{C}_2^{-1} : \{\zeta_1, \cdots \zeta_{n+1}\} \to \left\{ \frac{\zeta_1}{1-\zeta_{n+1}}, \dots, \frac{i}{2} \frac{\overline{\zeta_{n+1}} - \zeta_{n+1}}{|1-\zeta_{n+1}|^2} \right\}.$$

Let $\mathcal{K} : \mathbf{H}^{2n+1} \backslash \{0\} \longrightarrow \mathbf{H}^{2n+1} \backslash \{0\}$ be such that $\mathcal{K} = \mathcal{C}_2^{-1} \circ \mathcal{C}_1$, then

$$\mathcal{K} : (z_1, \cdots z_n, t) \to \left(\frac{z_1}{|z|^2 - 2it}, \dots, \frac{z_n}{|z|^2 - 2it}, \frac{t}{|z|^4 + 4t^2} \right).$$

Clearly $\mathcal{K}$ is an involution on $\mathbf{H}^{2n+1} \backslash \{0\}$ and preserve the Korányi ball $\{(z,t) \in \mathbf{H}^{2n+1}, |z|^4 + 4t^2 = 1\}$. Indeed it is the Kelvin transform generalized to Heisenberg group (see [9]).

For any $(r_H, t) \in \mathbb{R}_{\geq 0} \times \mathbb{R}$ and $(r_S, \theta) \in [0, \frac{\pi}{2}) \times \mathbb{R}/2\pi\mathbb{Z}$, we let $\tilde{h}(r_S, \theta) = 1 + \cos^2 r_S - 2\cos r_S \cos\theta$ and $\tilde{H}(r_H, t) = \frac{4(r_H^4 + 4t^2)}{(1+r_H^2)+4t^2}$, then $\mathcal{K}^* H = (\mathcal{C}_2 \circ \mathcal{C}_1^{-1})^* H = \tilde{H}$. Moreover, simple calculations show that

$$\tilde{h} = (\mathcal{C}_2^{-1})^* H, \quad h = (\mathcal{C}_2^{-1})^* \tilde{H}, \quad \tilde{h} = (\mathcal{C}_1^{-1})^* \tilde{H}.$$

Let $N(r_H, t) = r_H^4 + 4t^2$. By comparing the conformal Laplacians induced by $\mathcal{C}_1$ and $\mathcal{C}_2$, we obtain the following relation.

Theorem 3.1 *For any function $F \in D_H$,*

$$(\mathcal{K}_* L_{\mathbf{H}^{2n+1}})F = N^{\frac{n}{2}+1} L_{\mathbf{H}^{2n+1}} (N^{-n/2} F).$$

Proof First we notice that for all $f \in D_S$,

$$\tilde{h}^{\frac{n}{2}+1}(-L_{\mathbb{S}^{2n+1}} + n^2) \left(\tilde{h}^{-\frac{n}{2}} f \right) = -(\mathcal{C}_{2*} L_{\mathbf{H}^{2n+1}}) f.$$

Together with (2.12) we obtain

$$h^{-(\frac{n}{2}+1)} (\mathcal{C}_{1*} L_{\mathbf{H}^{2n+1}})(h^{\frac{n}{2}} f) = \tilde{h}^{-(\frac{n}{2}+1)} (\mathcal{C}_{2*} L_{\mathbf{H}^{2n+1}})(\tilde{h}^{\frac{n}{2}} f).$$

Thus

$$(\mathcal{C}_{1*}L_{\mathbf{H}^{2n+1}})f = n^{-(\frac{n}{2}+1)}(\mathcal{C}_{2*}L_{\mathbf{H}^{2n+1}})\left(n^{\frac{n}{2}}f\right),$$

where $n = \frac{\tilde{h}}{h}$. Note that $(\mathcal{C}_2)^*n = N^{-1}$, we have for any $F = \mathcal{C}_2^*f$,

$$(\mathcal{K}_*L_{\mathbf{H}^{2n+1}})F = N^{\frac{n}{2}+1}L_{\mathbf{H}^{2n+1}}(N^{-\frac{n}{2}}F).$$

□

Now we are ready to prove the relation between the inversion of Brownian motion on $\mathbf{H}^{2n+1}$ and the time changed Brownian bridge on $\mathbf{H}^{2n+1}$.

Proof of Theorem 1.2 Note that $N^{-\frac{n}{2}}$ is the Green function of the sub-Laplacian $L_{\mathbf{H}^{2n+1}}$ with pole $(0, 0)$. We let

$$L^N F := L_{\mathbf{H}^{2n+1}}F + 2N^{\frac{n}{2}}\Gamma_{\mathbf{H}^{2n+1}}(N^{-\frac{n}{2}}, F), \tag{3.21}$$

where $\Gamma_{\mathbf{H}^{2n+1}}(F, G) = \frac{1}{2}(L_{\mathbf{H}^{2n+1}}(FG) - fL_{\mathbf{H}^{2n+1}}G - GL_{\mathbf{H}^{2n+1}}F)$ for any $F, G \in D_H$. From the previous theorem we have

$$\mathcal{K}_*L_{\mathbf{H}^{2n+1}} = NL^N.$$

Let X_t^N and X_t be Markov processes generated by $\frac{1}{2}L^N$ and $\frac{1}{2}L_{\mathbf{H}^{2n+1}}$. We first prove that X_t^N is X_t conditioned to be at the origin.

It suffices to prove that for any $F \in D_H$,

$$\mathbb{E}_x\left[F(X_t^N)\right] = \mathbb{E}_x\left[F(X_t)\mathbb{1}_{t<T}|X_\infty = (0, 0)\right] \tag{3.22}$$

Let P_t^N and P_t be the heat semigroups generated by L^N and $L_{\mathbf{H}^{2n+1}}$ respectively, then by iterating (3.21) it is not hard to obtain

$$P_t^N(F(x)) = N(x)^{-\frac{n}{2}}P_t(N(x)^{-\frac{n}{2}}F(x)),$$

that is

$$\mathbb{E}_x\left[F(X_t^N)\right] = \frac{1}{N(x)^{-\frac{n}{2}}}\mathbb{E}_x\left[N(X_t)^{-\frac{n}{2}}F(X_t)\right] = \mathbb{E}_x\left[\frac{N(X_t)^{-\frac{n}{2}}}{N(x)^{-\frac{n}{2}}}F(X_t)\right].$$

From (3.22), we just need to show that

$$\mathbb{E}_x\left[F(X_t)|X_\infty = 0\right] = \mathbb{E}_x\left[\frac{N(X_t)^{-\frac{n}{2}}}{N^{-\frac{n}{2}}(x)}F(X_t)\right].$$

This is an easy consequence of $\mathbb{E}_x\left[X_\infty = 0\right] = N^{-\frac{n}{2}}(x)$ and

$$\mathbb{E}_x\left[F(X_t)\mathbb{1}_{X_\infty=0}\right] = \mathbb{E}_x\left[F(X_t)\mathbb{E}_x[\mathbb{1}_{X_\infty=0}|\mathcal{F}_t]\right] = \mathbb{E}_x\left[N(X_t)^{-\frac{n}{2}}F(X_t)\right].$$

Next we prove the time change. Consider the Markov process generated by $\frac{1}{2}\mathcal{K}_*(L_{\mathbf{H}^{2n+1}})$. It is the image of X_t under Kelvin transform, namely $\mathcal{K}(X_t)$. We claim

$$\mathcal{K}(X_t) = X^N_{\mathcal{A}_t} \tag{3.23}$$

where $\mathcal{A}_t = \int_0^t N(X_s)ds$ is the time-change of X^N. For any $F \in D_H$, we consider the associated martingale

$$M_t^F := F(X_t) - \frac{1}{2}\int_0^t L_{\mathbf{H}^{2n+1}}F(X_s)ds.$$

Denote $\tilde{F} = (\mathcal{K})^*F$. By plugging in (3.23), we obtain

$$M_t^F = \tilde{F}(X^N_{\mathcal{A}_t}) - \frac{1}{2}\int_0^t (L_{\mathbf{H}^{2n+1}}F)\circ\mathcal{K}^{-1}(X^h_{\mathcal{A}_s})ds = \tilde{F}(X^N_{\mathcal{A}_t}) - \frac{1}{2}\int_0^t N(X_s)L^N\tilde{F}(X^N_{\mathcal{A}_s})ds.$$

Let σ_t be the hitting time such that $\sigma_t = \inf\{u, \mathcal{A}_u > t\}$, then clearly $\mathcal{A}_{\sigma_t} = t = \sigma_{\mathcal{A}_t}$. By changing variable $s = \sigma_u$ we have

$$M_t^F = \tilde{F}(X^N_{\mathcal{A}_t}) - \frac{1}{2}\int_0^{\sigma_{\mathcal{A}_t}} N(X_s)L^N\tilde{F}(X^N_{\mathcal{A}_s})ds = \tilde{F}(X^N_t) - \frac{1}{2}\int_0^{\mathcal{A}_t} N(X_{\sigma_u})L^N\tilde{F}(X^N_u)\sigma'_u du.$$

Note for any $u > 0$ we have $u = \mathcal{A}_{\sigma_u} = \int_0^{\sigma_u} N(X_s)ds$. By differentiating both sides with respect to u we obtain

$$1 = N(X_{\sigma_u})\sigma'_u.$$

Hence

$$M_t^F = \tilde{F}(X^N_t) - \frac{1}{2}\int_0^{\mathcal{A}_t} L^N\tilde{F}(X^N_u)du,$$

and we have the conclusion.

References

1. F. Baudoin, M. Bonnefont, The subelliptic heat kernel on $SU(2)$: representations, asymptotics and gradient bounds. Math. Z. **263**, 647–672 (2009)
2. F. Baudoin, J. Wang, The subelliptic heat kernel on the CR sphere. Math. Z. **275**(1–2), 135–150 (2013)

3. F. Baudoin, J. Wang, Stochastic areas, winding numbers and Hopf fibrations. Probab. Theory Relat. Fields (2016). https://doi.org/10.1007/s00440-016-0745-x
4. R. Beals, B. Gaveau, P.C. Greiner, Hamilton-Jacobi theory and the heat kernel on Heisenberg groups. J. Math. Pures Appl. **79**(7), 633–689 (2000)
5. T.K. Carne, Brownian motion and stereographic projection. Ann. Inst. Henri Poincaré Probab. Stat. Sect. B **21**(2), 187–196 (1985)
6. S. Dragomir, G. Tomassini, *Differential Geometry and Analysis on CR Manifolds*, vol. 246 (Birkhäuser, Boston, 2006)
7. B. Gaveau, Principe de moindre action, propagation de la chaleur et estiméees sous elliptiques sur certains groupes nilpotents. Acta Math. **139**(1), 95–153 (1977)
8. D. Geller, The Laplacian and the Kohn Laplacian for the sphere. J. Differ. Geom. **15**, 417–435 (1980)
9. A. Korányi, Kelvin transforms and harmonic polynomials on the Heisenberg group. J. Funct. Anal. **49**(2), 177–185 (1982)
10. L. Schwartz, Le mouvement brownien sur $\mathbb{R}^N$, en tant que semi-martingale dans S_N. Ann. Inst. Henri Poincaré Probab. Stat. **21**(1), 15–25 (1985)
11. M. Yor, A prtopos de l'inverse du mouvement brownien dans $\mathbb{R}^n$ $(n \geq 3)$. Ann. Inst. Henri Poincaré Probab. Stat. **21**(1), 27–38 (1985)

On the Macroscopic Fractal Geometry of Some Random Sets

Davar Khoshnevisan and Yimin Xiao

Abstract This paper is concerned mainly with the macroscopic fractal behavior of various random sets that arise in modern and classical probability theory. Among other things, it is shown here that the macroscopic behavior of Boolean coverage processes is analogous to the microscopic structure of the Mandelbrot fractal percolation. Other, more technically challenging, results of this paper include:

(i) The computation of the macroscopic Minkowski dimension of the graph of a large family of Lévy processes; and
(ii) The determination of the macroscopic monofractality of the extreme values of symmetric stable processes.

As a consequence of (i), it will be shown that the macroscopic fractal dimension of the graph of Brownian motion differs from its microscopic fractal dimension. Thus, there can be no scaling argument that allows one to deduce the macroscopic geometry from the microscopic. Item (ii) extends the recent work of Khoshnevisan et al. (Ann Probab, to appear) on the extreme values of Brownian motion, using a different method.

Keywords Boolean models • Lévy processes • Macroscopic Minkowski dimension

AMS 2010 Subject Classification Primary 60G51; Secondary 28A80, 60G17, 60G52

D. Khoshnevisan (✉)
Department of Mathematics, University of Utah, 155 South 1400 East, JWB 102, Salt Lake City, UT 84112-0090, USA
e-mail: davar@math.utah.edu; davarkhoshnevisan@gmail.com

Y. Xiao
Department of Statistics & Probability, Michigan State University, East Lansing, MI 48824, USA
e-mail: xiao@stt.msu.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_9

1 Introduction

It has been known for some time that the curve of a Lévy process in $\mathbb{R}^d$ is typically an interesting "random fractal." For example, if $B = \{B_t\}_{t \geqslant 0}$ is a standard Brownian motion on $\mathbb{R}^d$, then the image and graph of B have Hausdorff dimension $d \wedge 2$ and $\max(d \wedge 2\,, 3/2)$ respectively. If in addition $d = 1$, then the level sets of B also have non-trivial Hausdorff dimension $1/2$. See the survey papers of Taylor [21] and Xiao [22] for historic accounts on these results and further developments.

The beginning student is often presented with some of these "random-fractal facts" via simulation. The well-versed reader will see in Fig. 1 a typical example. As a consequence of such a simulation, one is led to believe that one can deduce from a simulation, such as that in Fig. 1, the fractal nature of the graph $\cup_{0 \leqslant t \leqslant 1}\{(t\,, B_t)\}$ of Brownian motion up to time 1.

Figure 1, and other such simulations, are produced by running a random walk for a long time and then rescaling, using a central-limit scaling. The process is usually explained by appealing to Donsker's invariance principle. Unfortunately, the actual statement of Donsker's invariance principle is not sufficiently strong to ensure that we can "see" the various fractal properties of Brownian motion in simulations. Though Barlow and Taylor [1, 2] have introduced a theory of large-scale random fractals which, among other things, provides a more rigorous justification.

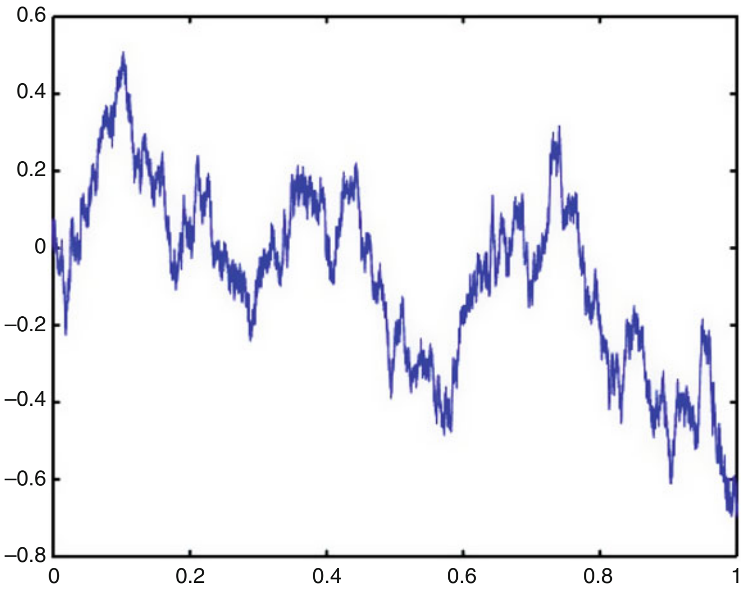

Fig. 1 The graph of one-dimensional Brownian motion

One of the goals of this paper is to test the extent to which one can experimentally deduce geometric facts about Brownian motion—and sometimes more general Lévy processes—from simulation analysis. This is achieved by presenting several examples in which one is able to compute the macroscopic fractal dimension of a macroscopic random fractal. One of the surprising lessons of this exercise is that our intuition is, at times, faulty. Yet, our instincts are correct at other times.

Here is an example in which our intuition is spot on: It is known that the level sets of Brownian motion have dimension $1/2$, both macroscopically and microscopically. This statement has the pleasant consequence that we can "see" the fractal structure of the level sets of Brownian motion from Fig. 1. As we shall soon see, however, the same cannot be said of the graph of Brownian motion: The microscopic and macroscopic fractal dimensions of the graph of Brownian motion do not agree!

In order to keep the technical level of the paper as low as possible, our choice of "fractal dimension" is the macroscopic Minkowski dimension, which we will present in the following section. There are more sophisticated notions which, we however, will not present here; see Barlow and Taylor [1, 2] for examples of these more sophisticated notions of macroscopic fractal dimension.

Throughout, we set $|x| := \max_{1\leqslant j\leqslant d} |x_j|$ and $\|x\| := (x_1^2 + \cdots + x_d^2)^{1/2}$ for all $x \in \mathbb{R}^d$. Whenever we write "$f(x) \lesssim g(x)$ [also $f(x) \gtrsim g(x)$] for all $x \in X$" we mean that there exists a finite constant K such that $f(x) \leqslant Kg(x)$ uniformly for all $x \in X$. If $f(x) \lesssim g(x)$ and $g(x) \lesssim f(x)$ for all $x \in X$, then we write "$f(x) \asymp g(x)$ for all $x \in X$."

2 Minkowski Dimension

The macroscopic Minkowski dimension is an easy-to-compute "fractal dimension number" that describes the large-scale fractal geometry of a set. In order to recall the Minkowski dimension, we first need to introduce some notation.

For all $x \in \mathbb{R}^d$ and $r > 0$ define

$$B(x;r) := [x_1 - r\,, x_1 + r) \times \cdots \times [x_d - r\,, x_d + r),$$

and

$$Q(x) := [x_1\,, x_1 + 1) \times \cdots \times [x_d\,, x_d + 1). \tag{1}$$

Of course, $Q(x) = B(y;\frac{1}{2})$ where $y_i := x_i + \frac{1}{2}$. But it is convenient for $Q(x)$ to have its own notation.

One can introduce a *pixelization map* which maps a set $F \subseteq \mathbb{R}^d$ to a set $\text{pix}(F) \subseteq \mathbb{Z}^d$ as follows:

$$\text{pix}(F) := \left\{x \in \mathbb{Z}^d : F \cap Q(x) \neq \varnothing\right\},$$

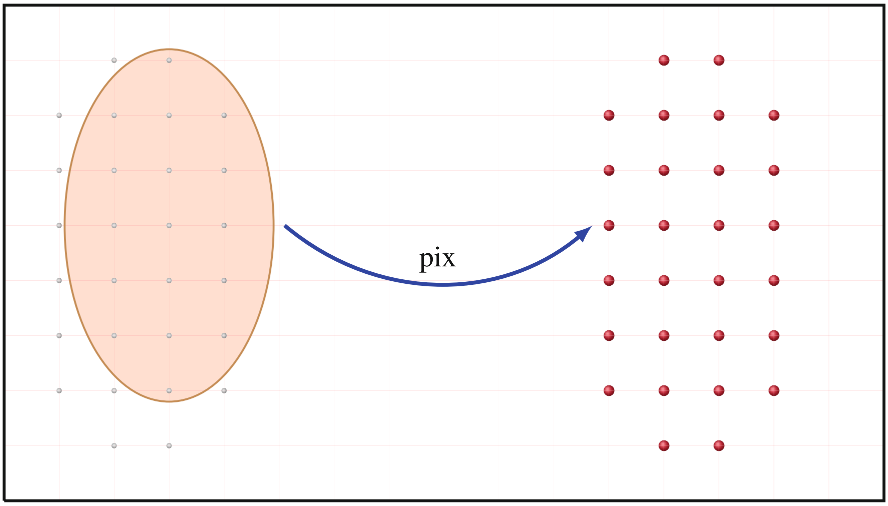

Fig. 2 The effect of the pixelization map on an ellipse

for all $F \subseteq \mathbb{R}^d$. It is clear that $F = \text{pix}(F)$ whenever F is a subset of the integer lattice $\mathbb{Z}^d$. For example, it should be clear that $\text{pix}(\mathbb{R}^d) = \mathbb{Z}^d$. Figure 2 below shows how the pixelization map works in a different simple case.

The following describes the role of the pixelization map in this paper.

Definition 2.1 The *macroscopic Minkowski dimension* of a set $F \subseteq \mathbb{R}^d$ is

$$\text{Dim}_{\text{M}}(F) := \limsup_{n\to\infty} n^{-1}\text{Log}_+ (|\text{pix}(F) \cap B(0; 2^n)|), \tag{2}$$

where $|\cdots|$ denotes cardinality and $\text{Log}_+(y) := \log_2(\max(y, 2))$.

Remark 2.2 The right-hand side of (2) coincides with the Barlow–Taylor [2] *upper mass dimension* of the discrete set $\text{pix}(F) \subseteq \mathbb{Z}^d$.

The proof of the following elementary result is left to the interested reader.

Lemma 2.3 *For every* $F \subseteq \mathbb{R}^d$,

$$\text{Dim}_{\text{M}}(F) = \limsup_{n\to\infty} n^{-1}\text{Log}_+ \left|\left\{x \in B(0; 2^n) \cap \mathbb{Z}^d : \ Q(x) \cap F \neq \varnothing\right\}\right|,$$

where $Q(x)$ *was defined in* (1).

Some of the elementary properties of Dim_{M} are listed below:

- If $A \subseteq B$ then $\text{Dim}_{\text{M}}(A) \leqslant \text{Dim}_{\text{M}}(B)$;
- If A is a bounded set, then $\text{Dim}_{\text{M}}(A) = 0$;
- $\text{Dim}_{\text{M}}(\mathbb{R}^d) = \text{Dim}_{\text{M}}(\mathbb{Z}^d) = d$.

The proof is omitted as it is easy to justify the preceding.

We end this section with a property of $\mathrm{Dim}_{\mathrm{M}}$ that is similar to the microscopic Minkowski dimension (compare with [6], for example), which will be used in the proof of Theorem 3.1 and in Example 3.16.

Lemma 2.4 $\mathrm{Dim}_{\mathrm{M}}(F) = \mathrm{Dim}_{\mathrm{M}}(\overline{F})$ *for every* $F \subseteq \mathbb{R}^d$*, where* $\overline{F}$ *denotes the closure of* F*.*

Proof Let $x_1, \ldots, x_{2^d}$ denote the corners of $B(0; r)$, where $r \in (0\,, 1)$, and let $x_j + \mathrm{pix}(F)$ denote the translate of $\mathrm{pix}(F)$ by x_j for all $1 \leqslant j \leqslant 2^d$. We may note that $\mathrm{pix}(\overline{F}) \subset \cup_{j=1}^{2^d}(x_j + \mathrm{pix}(F))$. Since $\|x_j\| = r\sqrt{d}$, it follows from the translation invariance of counting measure that

$$\begin{aligned} \left|\mathrm{pix}(\overline{F}) \cap B(0; 2^n)\right| &\leqslant \sum_{j=1}^{2^d} \left|\{x_j + \mathrm{pix}(F)\} \cap B\left(x_j; r\sqrt{d} + 2^n\right)\right| \\ &\leqslant 2^d \left|\mathrm{pix}(F) \cap B\left(0; r\sqrt{d} + 2^n\right)\right|. \end{aligned}$$

Let $r \downarrow 0$ to deduce the second inequality in the following, the first being a tautology:

$$|\mathrm{pix}(F) \cap B(0; 2^n)| \leqslant \left|\mathrm{pix}(\overline{F}) \cap B(0; 2^n)\right| \leqslant 2^d\, |\mathrm{pix}(F) \cap B(0; 2^n)|\,.$$

The lemma follows from the above and (2). □

2.1 Enumeration in Shells

There is a slightly different method of computing the macroscopic Minkowski dimension of a set. With this aim in mind, define

$$\mathcal{S}_0 := B(0; 1) \cap \mathbb{Z}^d, \qquad \mathcal{S}_{n+1} := \left(B(0; 2^{n+1}) \setminus B(0; 2^n)\right) \cap \mathbb{Z}^d \quad \text{for every integer } n \geqslant 0.$$

One can think of $\mathcal{S}_n$ as the nth *shell* in $\mathbb{Z}^d$.

The following provides an alternative description of $\mathrm{Dim}_{\mathrm{M}}(F)$.

Proposition 2.5 *For every* $F \subseteq \mathbb{R}^d$*,*

$$\mathrm{Dim}_{\mathrm{M}}(F) := \limsup_{n\to\infty} n^{-1} \mathrm{Log}_+ \left(|\mathrm{pix}(F) \cap \mathcal{S}_n|\right).$$

Proposition 2.5 tells us that we can replace $\mathrm{pix}(F) \cap B(0; 2^n)$, in Definition 2.1, by $\mathrm{pix}(F) \cap \mathcal{S}_n$ without altering the formula for $\mathrm{Dim}_{\mathrm{M}}(F)$.

Proof Our goal is to prove that $\mathrm{Dim}_{\mathrm{M}}(F) = \delta(F)$, where

$$\delta(F) := \limsup_{n\to\infty} n^{-1}\mathrm{Log}_+ \left(|\mathrm{pix}(F) \cap \mathcal{S}_n|\right).$$

Since $\mathcal{S}_n \subseteq B(0;2^n)$, the bound $\delta(F) \leqslant \mathrm{Dim}_{\mathrm{M}}(F)$ is immediate. We will establish the reverse inequality.

The definition of $\delta(F)$ ensures that for every $\varepsilon \in (0\,,1)$ there exists an integer $N(\varepsilon)$ such that

$$|\mathrm{pix}(F) \cap \mathcal{S}_k| \leqslant 2^{k\delta(F)(1+\varepsilon)} \qquad \text{for all } k \geqslant N(\varepsilon).$$

In particular, all $n \geqslant N(\varepsilon)$,

$$\begin{aligned} |\mathrm{pix}(F) \cap B(0;2^n)| &= \sum_{k=0}^{n} |\mathrm{pix}(F) \cap \mathcal{S}_k| \leqslant K(\varepsilon) + \sum_{k=N(\varepsilon)}^{n} 2^{k\delta(F)(1+\varepsilon)}, \\ &= 2^{n\delta(F)(1+o(1))} \qquad [n \to \infty], \end{aligned}$$

where $K(\varepsilon) := \sum_{0\leqslant k<N(\varepsilon)} |\mathcal{S}_k|$ is finite and depends only on $(d\,,\varepsilon)$. It follows from (2) that $\mathrm{Dim}_{\mathrm{M}}(F) \leqslant \delta(F)/(1-\varepsilon)$. This completes the proof since $\varepsilon \in (0\,,1)$ can be made to be as small as one would like. □

2.2 Boolean Models

In addition to the method of Proposition 2.5, there is at least one other useful method for computing the macroscopic Minkowski dimension of a set. In contrast with the enumerative method of Sect. 2.1, the method of this subsection is intrinsically probabilistic.

Let $\mathbf{p} := \{p(x)\}_{x\in\mathbb{Z}^d}$ denote a collection of numbers in $(0\,,1)$, and refer to the collection $\mathbf{p}$ as *coverage probabilities*, in keeping with the literature on Boolean coverage processes [7].

Let $\zeta := \{\zeta(x)\}_{x\in\mathbb{Z}^d}$ denote a field of totally independent random variables that satisfy the following for all $x \in \mathbb{Z}^d$:

$$\mathrm{P}\{\zeta(x) = 1\} = p(x) \quad \text{and} \quad \mathrm{P}\{\zeta(x) = 0\} = 1 - p(x).$$

By a *Boolean model* in $\mathbb{R}^d$ with *coverage probabilities* $\mathbf{p}$ we mean the random set

$$\mathbf{B}(\mathbf{p}) := \bigcup_{\substack{x\in\mathbb{Z}^d:\\ \zeta(x)=1}} Q(x),$$

where $Q(x)$ was defined earlier in (1). Figure 3 depicts simulations of two Booelan models.

If A and B are two subsets of $\mathbb{R}^d$, then we say that A is *recurrent* for B if $|\operatorname{pix}(A \cap B)| = \infty$. Equivalently, A is recurrent for B if $\operatorname{pix}(A \cap B) \cap \mathcal{S}_n \neq \varnothing$ for infinitely-many integers $n \geqslant 0$. Clearly, if A is recurrent for B, then B is also recurrent for A. Therefore, set recurrence is a symmetric relation.

As the following result shows, it is not hard to decide whether or not a nonrandom Borel set $A \subseteq \mathbb{R}^d$ is recurrent for $\mathbf{B}(\mathbf{p})$.

Lemma 2.6 *Let $A \subset \mathbb{R}^d$ be a nonrandom Borel set. Then,*

$$\mathrm{P}\left\{\left|\operatorname{pix}\left(A \cap \mathbf{B}(\mathbf{p})\right)\right| = \infty\right\} = \begin{cases} 1 & \text{if } \sum_{x \in \operatorname{pix}(A)} p(x) = \infty, \\ 0 & \text{if } \sum_{x \in \operatorname{pix}(A)} p(x) < \infty. \end{cases}$$

Lemma 2.6 is basically a reformulation of the Borel–Cantelli lemma for independent events. Therefore, we skip the proof. Instead, let us mention the following, more geometric, result which almost characterizes recurrent sets in terms of their macroscopic Minkowski dimension, in some cases.

Proposition 2.7 *Suppose* $\mathbf{p}$ *has an* index,

$$\operatorname{Ind}(\mathbf{p}) := -\lim_{|x| \to \infty} \frac{\log p(x)}{\log |x|}. \tag{3}$$

Then for every nonrandom Borel set $A \subseteq \mathbb{R}^d$,

$$\mathrm{P}\left\{\left|\operatorname{pix}\left(A \cap \mathbf{B}(\mathbf{p})\right)\right| = \infty\right\} = \begin{cases} 1 & \text{if } \operatorname{Dim}_{\mathrm{M}}(A) > \operatorname{Ind}(\mathbf{p}), \\ 0 & \text{if } \operatorname{Dim}_{\mathrm{M}}(A) < \operatorname{Ind}(\mathbf{p}). \end{cases}$$

We can compare this result to a similar result of Hawkes [8] about the hitting probabilities of the Mandelbrot fractal percolation. This comparison suggests that the Boolean models of this paper play an analogous role in the theory of macroscopic fractals as does fractal percolation in the better-studied theory of microscopic fractals.

Open Problem Is there a macroscopic analogue of the microscopic capacity theory of Peres [17, 18]?

Proof of Proposition 2.7 Let us consider the process $N_0, N_1, N_2, \ldots$, defined as

$$N_n := \left|\operatorname{pix}\left(A \cap \mathbf{B}(\mathbf{p})\right) \cap \mathcal{S}_n\right| = \sum_{x \in \operatorname{pix}(A) \cap \mathcal{S}_n} \zeta(x) \qquad [n \geqslant 0].$$

Owing to (3) and the definition of $\operatorname{Dim}_{\mathrm{M}}$, we can verify that

$$\limsup_{n \to \infty} n^{-1} \operatorname{Log}_+ \mathrm{E}(N_n) = \operatorname{Dim}_{\mathrm{M}}(A) - \operatorname{Ind}(\mathbf{p}). \tag{4}$$

Suppose first that $\mathrm{Dim}_{\mathrm{M}}(A) < \mathrm{Ind}(\mathbf{p})$. We may combine (4) and Markov's inequality in order to see that $\sum_{n=1}^{\infty} \mathrm{P}\{N_n > 0\} \leqslant \sum_{n=1}^{\infty} \mathrm{E}(N_n) < \infty$. The Borel–Cantelli lemma then implies that with probability one $N_n = 0$ for all but finitely-many integers n. That is, $\left|\mathrm{pix}\left(A \cap \mathbf{B}(\mathbf{p})\right)\right| < \infty$ a.s. if $\mathrm{Dim}_{\mathrm{M}}(A) < \mathrm{Ind}(\mathbf{p})$. This proves half of the proposition.

For the remaining half let us assume that $\mathrm{Dim}_{\mathrm{M}}(A) > \mathrm{Ind}(\mathbf{p})$, and notice that $\mathrm{Var}(N_n) = \sum_{x \in \mathrm{pix}(A) \cap \mathcal{S}_n} p(x)(1-p(x)) \leqslant \mathrm{E}(N_n)$. Therefore,

$$\mathrm{P}\left\{N_n \leqslant \tfrac{1}{2}\mathrm{E}(N_n)\right\} \leqslant \mathrm{P}\left\{|N_n - \mathrm{E}N_n| \geqslant \tfrac{1}{2}\mathrm{E}(N_n)\right\} \leqslant \frac{4\,\mathrm{Var}(N_n)}{|\mathrm{E}(N_n)|^2} \leqslant \frac{4}{\mathrm{E}(N_n)}, \tag{5}$$

thanks to the Chebyshev's inequality. Because of (4) there exists an infinite collection $\mathcal{N}$ of positive integers such that

$$n^{-1}\mathrm{Log}_+\mathrm{E}(N_n) \to \mathrm{Dim}_{\mathrm{M}}(A) - \mathrm{Ind}(\mathbf{p}) > 0 \qquad \text{as } n \text{ approaches infinity in } \mathcal{N}.$$

This fact, and (5), together imply that $\sum_{n \in \mathcal{N}} \mathrm{P}\{N_n \leqslant \frac{1}{2}\mathrm{E}(N_n)\} < \infty$, and hence

$$\mathrm{Dim}_{\mathrm{M}}(\mathbf{B}(\mathbf{p}) \cap A) = \limsup_{n \to \infty} n^{-1}\mathrm{Log}_+ N_n \geqslant \lim_{\substack{n \to \infty:\\ n \in \mathcal{N}}} n^{-1}\mathrm{Log}_+ N_n \geqslant \mathrm{Dim}_{\mathrm{M}}(A) - \mathrm{Ind}(\mathbf{p}) > 0,$$

almost surely. This completes the proof. □

Remark 2.8 A quick glance at the proof shows that the independence of the ζ's was needed only to show that

$$\mathrm{Var}(N_n) = O(\mathrm{E}(N_n)) \qquad \text{as } n \to \infty. \tag{6}$$

Because $\mathrm{Var}(N_n) = \sum_{x,y \in \mathrm{pix}(A) \cap \mathcal{S}_n} \mathrm{P}\{\zeta(x) = \zeta(y) = 1\}$, (6) continues to hold if the independence of the ζ's is relaxed to a condition such as the following: There exists finite and positive constants c and K such that

$$\mathrm{P}[\zeta(x) = 1 \mid \zeta(y) = 1] \leqslant c\mathrm{P}\{\zeta(x) = 1\} \quad \text{whenever } \|x\| \wedge \|y\| \geqslant K.$$

We highlight the power of Proposition 2.7 by using it to give a quick computation of $\mathrm{Dim}_{\mathrm{M}}(A \cap \mathbf{B}(\mathbf{p}))$.

Corollary 2.9 *If $A \subseteq \mathbb{R}^d$ denotes a nonrandom Borel set, then*

$$\mathrm{Dim}_{\mathrm{M}}\left(A \cap \mathbf{B}(\mathbf{p})\right) = \mathrm{Dim}_{\mathrm{M}}(A) - \mathrm{Ind}(\mathbf{p}) \qquad a.s.$$

Because $\mathrm{Dim}_{\mathrm{M}}(\mathbb{R}^d) = d$, the following is an immediate consequence of Corollary 2.9.

Corollary 2.10 $\mathrm{Dim}_{\mathrm{M}}(\mathbf{B}(\mathbf{p})) = d - \mathrm{Ind}(\mathbf{p}) \qquad a.s.$

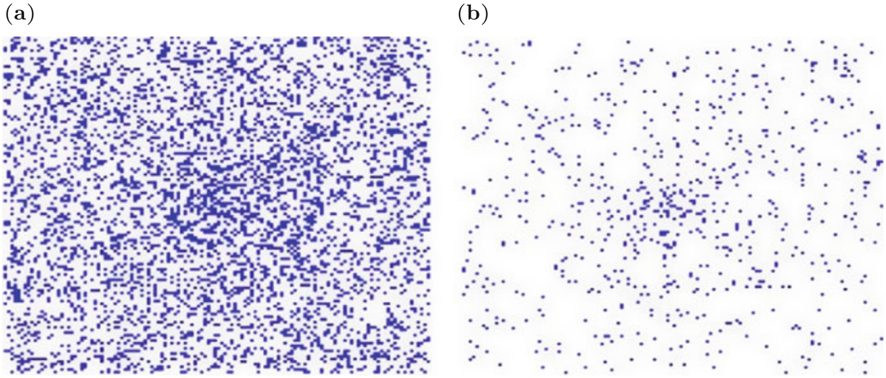

Fig. 3 A simulation of two Boolean models. Corollary 2.10 ensures that the Minkowski dimensions of the two figures are respectively 1.7 (**a**) and 1.3 (**b**). (**a**) $\text{Ind}(\mathbf{p}) = 0.3$, $\text{Dim}_{\text{M}}(\mathbf{B}(\mathbf{p})) = 1.7$. (**b**) $\text{Ind}(\mathbf{p}) = 0.7$, $\text{Dim}_{\text{M}}(\mathbf{B}(\mathbf{p})) = 1.3$

Therefore, it remains to establish Corollary 2.9. The proof uses a variation of an elegant "replica argument" that was introduced by Peres [18] in the context of [microscopic] Hausdorff dimension of fractal percolation processes.

Proof of Corollary 2.9 Let $\mathbf{B}'(\mathbf{p}')$ be an independent Boolean model with coverage probabilities $\mathbf{p}' = \{p'(x)\}_{x\in\mathbb{Z}^d}$ that have an index $\text{Ind}(\mathbf{p}')$. Define $q(x) := p(x)\times p'(x)$ for all $x \in \mathbb{Z}^d$. It is then easy to see that $\mathbf{C}(\mathbf{q}) := \mathbf{B}'(\mathbf{p}') \cap \mathbf{B}(\mathbf{p})$ is a Boolean model with coverage probabilities $\mathbf{q} = \{q(x)\}_{x\in\mathbb{Z}^d}$. Since $\text{Ind}(\mathbf{q}) = \text{Ind}(\mathbf{p}) + \text{Ind}(\mathbf{p}')$, Proposition 2.7 implies that

$$\mathrm{P}\left\{\left|\text{pix}\left(A\cap\mathbf{C}(\mathbf{q})\right)\right| = \infty\right\} = \begin{cases} 1 & \text{if } \text{Ind}(\mathbf{p}) + \text{Ind}(\mathbf{p}') < \text{Dim}_{\text{M}}(A),\\ 0 & \text{if } \text{Ind}(\mathbf{p}) + \text{Ind}(\mathbf{p}') > \text{Dim}_{\text{M}}(A).\end{cases}$$

At the same time, one can apply Proposition 2.7 conditionally in order to see that almost surely,

$$\begin{aligned}\mathrm{P}\left\{\left|\text{pix}\left(A\cap\mathbf{C}(\mathbf{q})\right)\right| = \infty \mid \mathbf{B}(\mathbf{p})\right\} &= \mathrm{P}\left\{\left|\text{pix}\left(A\cap\mathbf{B}(\mathbf{p})\cap\mathbf{B}'(\mathbf{p}')\right)\right| = \infty \mid \mathbf{B}(\mathbf{p})\right\}\\ &= \begin{cases} 1 & \text{if } \text{Dim}_{\text{M}}\left(A\cap\mathbf{B}(\mathbf{p})\right) > \text{Ind}(\mathbf{p}'),\\ 0 & \text{if } \text{Dim}_{\text{M}}\left(A\cap\mathbf{B}(\mathbf{p})\right) < \text{Ind}(\mathbf{p}').\end{cases}\end{aligned}$$

A comparison of the preceding two displays yields the following almost sure assertions:

1. If $\text{Ind}(\mathbf{p}) + \text{Ind}(\mathbf{p}') < \text{Dim}_{\text{M}}(A)$, then $\text{Dim}_{\text{M}}\left(A\cap\mathbf{B}(\mathbf{p})\right) \geqslant \text{Ind}(\mathbf{p}')$ a.s.; and
2. If $\text{Ind}(\mathbf{p}) + \text{Ind}(\mathbf{p}') > \text{Dim}_{\text{M}}(A)$, then $\text{Dim}_{\text{M}}\left(A\cap\mathbf{B}(\mathbf{p})\right) \leqslant \text{Ind}(\mathbf{p}')$ a.s.

Since $\mathbf{p}'$ can have any arbitrary index $\text{Ind}(\mathbf{p}') > 0$ that one wishes, the corollary follows. □

3 Transient Lévy Processes

Let $X := \{X_t\}_{t\geqslant 0}$ be a Lévy process on $\mathbb{R}^d$. That is, X is a strong Markov process that has càdlàg paths, takes values in $\mathbb{R}^d$, $X_0 = 0$, and X has stationary and independent increments. See, for example, Bertoin [3] for a pedagogic account. In this section we assume that X is transient and compute the macroscopic Minkowski dimension of the range $\mathcal{R}_X$ of X, where we recall the range is the following random set:

$$\mathcal{R}_X := \bigcup_{t\geqslant 0}\{X_t\}.$$

3.1 The Potential Measure

Let U_X denote the potential measure of X; that is,

$$\mathrm{U}_X(A) := \int_0^\infty \mathrm{P}\{X_t \in A\}\,\mathrm{d}t = \mathrm{E}\int_0^\infty \mathbb{1}_A(X_t)\,\mathrm{d}t. \tag{7}$$

Throughout we assume that X is transient; equivalently, U_X is a Radon measure. The following shows that the macroscopic Minkowski dimension of the range of X is linked intimately to the potential measure of X.

Theorem 3.1 *With probability one,*

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) = \inf\left\{\alpha > 0 : \int_{\mathbb{R}^d} \frac{\mathrm{U}_X(\mathrm{d}x)}{1+|x|^\alpha} < \infty\right\}.$$

Theorem 3.9 below contains an alternative formula for $\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X)$, in terms of the Lévy exponent of X, which is reminiscent of an old formula of Pruitt [20] for the [microscopic] Hausdorff dimension of $\mathcal{R}_X$. We refer to Ref.'s [11–13] for more recent developments on microscopic fractal properties of Lévy processes, based on potential theory of additive Lévy processes.

Example 3.2 Consider the case that $X := \{X_t\}_{t\geqslant 0}$ is a symmetric β-stable process on $\mathbb{R}^d$ for some $0 < \beta \leqslant 2$. Transience is equivalent to the condition $\beta < d$. This condition is known to imply that $\mathrm{U}_X(\mathrm{d}x)/\mathrm{d}x \propto \|x\|^{-d+\beta}$ for all $x \in \mathbb{R}^d \setminus \{0\}$ [3, 19]. Therefore, $\int_{\mathbb{R}^d}(1+|x|^\alpha)^{-1}\,\mathrm{U}_X(\mathrm{d}x) < \infty$ iff $\int_{|x|>1}|x|^{-\alpha-d+\beta}\,\mathrm{d}x < \infty$ iff $\alpha > \beta$. Theorem 3.1 then implies that $\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) = \beta$ a.s. This fact is essentially due to Barlow and Taylor [2].

Remark 3.3 Recall that the measure U_X is finite because X is transient. As a result, $\int_{\mathbb{R}^d}(1+|x|^{\alpha})^{-1}\,\mathrm{U}_X(\mathrm{d}x)$ converges iff $\int_{|x|>1}|x|^{-\alpha}\,\mathrm{U}_X(\mathrm{d}x)<\infty$. One can then deduce from this fact, from the definition (7) of U_X, and from Theorem 3.1 that

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X)=\inf\left\{\alpha>0:\ \int_0^\infty \mathrm{E}\left(|X_t|^{-\alpha};\,|X_t|>1\right)\mathrm{d}t<\infty\right\}\qquad \text{a.s.}$$

This is the macroscopic analogue of a result of Pruitt [20, p. 374].

Open Problem It is natural to ask if there is a nice formula for $\mathrm{Dim}_{\mathrm{M}}(A\cap\mathcal{R}_X)$ when $A\subseteq\mathbb{R}^d$ is Borel and nonrandom. We do not have an answer to this question when A is not "macroscopically self-similar."

The proof of Theorem 3.1 hinges on a few prefatory technical results. The first is a more-or-less well-known set of bounds on the potential measure of a ball.

Lemma 3.4 *For every $x\in\mathbb{R}^d$ and $r>0$,*

$$\mathrm{U}_X(B(x;r))\leqslant \mathrm{U}_X(B(0;2r))\cdot\mathrm{P}\left\{\overline{\mathcal{R}_X}\cap B(x;r)\neq\varnothing\right\}.$$

Proof Let $\inf\varnothing:=\infty$, and consider the stopping time

$$T(x;r):=\inf\{t\geqslant 0: X_t\in B(x;r)\}. \tag{8}$$

We can write $\mathrm{U}_X(B(x;r))$ in the following equivalent form:

$$\mathrm{E}\left(\int_0^\infty \mathbb{1}_{B(x-X_{T(x;r)},r)}\left(X_{t+T(x;r)}-X_{T(x;r)}\right)\mathrm{d}t\cdot\mathbb{1}_{\{T(x;r)<\infty\}}\right). \tag{9}$$

Since $|X_{T(x;r)}-x|\leqslant r$ a.s. on the event $\{T(x;r)<\infty\}$, the triangle inequality implies that $B(x-X_{T(x;r)}\,,r)\subseteq B(0;2r)$ a.s. on $\{T(x;r)<\infty\}$, and hence

$$\mathrm{U}_X(B(x;r))\leqslant\mathrm{U}_X(B(0;2r))\cdot\mathrm{P}\{T(x;r)<\infty\}.$$

This is another way to state the lemma. □

The next result is a standard upper bound on the hitting probability of a ball.

Lemma 3.5 *For every $x\in\mathbb{R}^d$ and $r>0$,*

$$\mathrm{U}_X(B(x;2r))\geqslant \mathrm{U}_X(B(0;r))\cdot\mathrm{P}\left\{\overline{\mathcal{R}_X}\cap B(x;r)\neq\varnothing\right\}.$$

Proof Similarly to (9), we see that $\mathrm{U}_X(B(x;2r))$ is bounded from below by

$$\mathrm{E}\left(\int_0^\infty \mathbb{1}_{B(x-X_{T(x;r)},2r)}\left(X_{t+T(x;r)}-X_{T(x;r)}\right)\mathrm{d}t\cdot\mathbb{1}_{\{T(x;r)<\infty\}}\right),$$

where $T(x;r)$ was defined in (8). By the triangle inequality, $B(x - X_{T(x;r)}, 2r) \supset B(0;r)$ almost surely on the event $\{T(x;r) < \infty\}$. Therefore, we apply the strong Markov property in order to see that

$$\mathrm{U}_X(B(x;2r)) \geqslant \mathrm{U}_X(B(0;r)) \cdot \mathrm{P}\{T(x;r) < \infty\}.$$

This is another way to write the lemma. □

The following is a "weak unimodality" result for the potential measure.

Lemma 3.6 $\mathrm{U}_X(B(x;r)) \leqslant 4^d \mathrm{U}_X(B(0;r))$ *for all* $x \in \mathbb{R}^d$ *and* $r > 0$.

Proof The proof will use the following elementary covering property of Euclidean spaces: For every $x \in \mathbb{R}^d$ and $r > 0$ there exist points $y_1, \ldots, y_{4^d} \in B(x;r)$ such that $B(x;r) = \cup_{1 \leqslant i \leqslant 4^d} B(y_i, r/2)$. This leads to the following "volume-doubling" bound: For all $r > 0$ and $x \in \mathbb{R}^d$,

$$\mathrm{U}_X(B(x;r)) \leqslant 4^d \sup_{y \in B(x,r)} \mathrm{U}_X(B(y;r/2)). \tag{10}$$

This inequality yields the lemma since $\mathrm{U}_X(B(y;r/2)) \leqslant \mathrm{U}_X(B(0;r))$ for all $y \in \mathbb{R}^d$ and $r > 0$, thanks to Lemma 3.4. □

The next result presents bounds for the probability that the pixelization of the range of X hits singletons. Naturally, both bounds are in terms of the potential measure of X.

Lemma 3.7 *There exist finite constants* $c_2 > 1 > c_1 > 0$ *such that, for all* $x \in \mathbb{Z}^d$,

$$c_1 \mathrm{U}_X(Q(x)) \leqslant \mathrm{P}\left\{x \in \mathrm{pix}\left(\overline{\mathcal{R}_X}\right)\right\} \leqslant c_2 \mathrm{U}_X(B(x;2)).$$

Proof For $x \in \mathbb{Z}^d$, let $y_i := x_i + \frac{1}{2}$ for $1 \leqslant i \leqslant d$ and recall that $Q(x) = B(y;1/2)$ in order to deduce from Lemmas 3.4 and 3.5 that

$$\frac{\mathrm{U}_X(Q(x))}{\mathrm{U}_X(B(0;1))} = \frac{\mathrm{U}_X(B(y;1/2))}{\mathrm{U}_X(B(0;1))} \leqslant \mathrm{P}\left\{x \in \mathrm{pix}\left(\overline{\mathcal{R}_X}\right)\right\} \leqslant \frac{\mathrm{U}_X(B(y;1))}{\mathrm{U}_X(B(0;1/2))}. \tag{11}$$

The denominators are strictly positive because X is càdlàg and $B(0;1/2)$ contains an open ball in $\mathbb{R}^d$; and they are finite because of the transience of X. Because $B(y;1) \subseteq B(x;2)$, (11) completes the proof. □

The following lemma is the final technical result of this section. It presents an upper bound for the probability that the range of X simultaneously intersects two given balls.

Lemma 3.8 *For all $x, y \in \mathbb{R}^d$ and $r > 0$,*

$$\mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(x;r) \neq \varnothing\,, \overline{\mathcal{R}_X} \cap B(y;r) \neq \varnothing\right\} \leqslant \frac{\mathrm{U}_X(B(x;2r))}{\mathrm{U}_X(B(0;r))} \cdot \frac{\mathrm{U}_X(B(y-x;4r))}{\mathrm{U}_X(B(0;2r))} + \frac{\mathrm{U}_X(B(y;2r))}{\mathrm{U}_X(B(0;r))} \cdot \frac{\mathrm{U}_X(B(x-y;4r))}{\mathrm{U}_X(B(0;2r))}.$$

Proof Let us recall the stopping time $T(x;r)$ from (8). First one notices that

$$\begin{aligned}\mathrm{P}\{T(x;r) \leqslant T(y;r) < \infty\} &= \mathrm{P}\left\{T(x;r) < \infty, \exists s \geqslant 0 : X_{s+T(x;r)} - X_{T(x;r)} \in B(y - X_{T(x;r)}; r)\right\}\\ &\leqslant \mathrm{P}\{T(x;r) < \infty\} \cdot \mathrm{P}\{T(y-x;2r) < \infty\}\\ &= \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(x;r) \neq \varnothing\right\} \cdot \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(y-x;2r) \neq \varnothing\right\},\end{aligned}$$

owing to the strong Markov property and the fact that $B(y - X_{T(x;r)}; r) \subseteq B(y - x; 2r)$ a.s. on $\{T(x;r) < \infty\}$ [the triangle inequality]. By exchanging the roles of x and y and appealing to the subadditivity of probabilities, one can deduce from the preceding that

$$\begin{aligned}&\mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(x;r) \neq \varnothing\,, \overline{\mathcal{R}_X} \cap B(y;r) \neq \varnothing\right\}\\ &\qquad \leqslant \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(x;r) \neq \varnothing\right\} \cdot \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(y-x;2r) \neq \varnothing\right\}\\ &\qquad\quad + \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(y;r) \neq \varnothing\right\} \cdot \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(x-y;2r) \neq \varnothing\right\}.\end{aligned}$$

An appeal to Lemma 3.5 completes the proof. □

With the requisite material for the proof of Theorem 3.1 under way, we are ready for the following.

Proof of Theorem 3.1 Because of Lemma 2.4, it is sufficient to to verify that $\mathrm{Dim}_{\mathrm{M}}(\overline{\mathcal{R}_X}) = \alpha_c$ a.s., where

$$\alpha_c := \inf\left\{\alpha > 0 : \int_{\mathbb{R}^d} \frac{\mathrm{U}_X(\mathrm{d}x)}{1 + |x|^\alpha} < \infty\right\}.$$

Let us begin by making some real-variable observations. First, let us note that because U_X is a finite measure [by transience],

$$\sum_{n=1}^{\infty} 2^{-n\alpha}\mathrm{U}_X(\mathcal{S}_n) = \sum_{n=1}^{\infty} 2^{-n\alpha}\int_{\mathcal{S}_n}\mathrm{U}_X(\mathrm{d}x) \asymp \int_{|x|>1}\frac{\mathrm{U}_X(\mathrm{d}x)}{|x|^\alpha} \asymp \int_{\mathbb{R}^d}\frac{\mathrm{U}_X(\mathrm{d}x)}{1+|x|^\alpha}.$$

Therefore,

$$\alpha_c = \inf\left\{\alpha > 0 : \sum_{n=1}^{\infty} 2^{-n\alpha}\mathrm{U}_X(\mathcal{S}_n) < \infty\right\}.$$

By the definition of α_c, if $0 < \alpha < \alpha_c$, then $\sum_n 2^{-n\alpha} \mathrm{U}_X(\mathcal{S}_n) = \infty$; as a result,

$$\limsup_{n\to\infty} 2^{-\beta n} \mathrm{U}_X(B(0;2^n)) \geqslant \limsup_{n\to\infty} 2^{-\beta n} \mathrm{U}_X(\mathcal{S}_n) = \infty,$$

whenever $0 < \beta < \alpha$. On the other hand, if $\beta > \alpha_c$, then $\lim_{n\to\infty} 2^{-\beta n} \mathrm{U}_X(\mathcal{S}_n) = 0$, and hence

$$\mathrm{U}_X(B(0;2^n)) = \sum_{k=0}^{n} \mathrm{U}_X(\mathcal{S}_k) = O(2^{\beta n}) \qquad \text{as } n \to \infty.$$

These remarks together show the following alternative representation of α_c:

$$\alpha_c = \limsup_{n\to\infty} n^{-1} \mathrm{Log}_+ \mathrm{U}_X(B(0;2^n)) = \limsup_{n\to\infty} n^{-1} \mathrm{Log}_+ \mathrm{U}_X(\mathcal{S}_n). \tag{12}$$

Now we begin the bulk of the proof. Lemma 3.7 and (12) together imply that for all $n \geqslant 2$,

$$\mathrm{E}\left|\mathrm{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right| \lesssim \sum_{x\in B(0;2^n)} \mathrm{U}_X(B(x;2)) \lesssim \mathrm{U}_X(B(0;2^{n+1})) \leqslant 2^{n(1+o(1))\alpha_c},$$

as $n \to \infty$. Therefore, the Chebyshev inequality implies that

$$\sum_{n=1}^{\infty} \mathrm{P}\left\{\left|\mathrm{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right| > 2^{n\theta}\right\} < \infty \qquad \text{for all } \theta > \alpha_c.$$

An application of the Borel–Cantelli lemma yields $\mathrm{Dim}_{\mathrm{M}}(\overline{\mathcal{R}_X}) \leqslant \alpha_c$ a.s., which implies a part of the assertion of the theorem.

For the next part, let us begin with the following consequence of Lemma 3.7:

$$\mathrm{E}\left|\mathrm{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right| \gtrsim \sum_{x\in B(0;2^n)} \mathrm{U}_X(Q(x)) \asymp \mathrm{U}_X(B(0;2^n)). \tag{13}$$

Next, we estimate the second moment of the same random variable as follows:

$$\begin{aligned}
\mathrm{E}\left(\left|\mathrm{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right|^2\right) &\leqslant \sum_{x,y\in B(0;2^n)} \mathrm{P}\left\{\overline{\mathcal{R}_X} \cap B(x;1) \neq \varnothing, \overline{\mathcal{R}_X} \cap B(y;1) \neq \varnothing\right\} \\
&\leqslant \sum_{x,y\in B(0;2^n)} \frac{\mathrm{U}_X(B(x;2))}{\mathrm{U}_X(B(0;1))} \cdot \frac{\mathrm{U}_X(B(y-x;4))}{\mathrm{U}_X(B(0;2))} \\
&\quad + \sum_{x,y\in B(0;2^n)} \frac{\mathrm{U}_X(B(y;2))}{\mathrm{U}_X(B(0;1))} \cdot \frac{\mathrm{U}_X(B(x-y;4))}{\mathrm{U}_X(B(0;2))};
\end{aligned}$$

see Lemma 3.8 for the final inequality. Since for all $x, y \in B(0; 2^n)$, we have $x - y, y - x \in B(0\,, 2^{n+1})$, it follows that

$$\begin{aligned} \mathrm{E}\left(\left|\operatorname{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0; 2^n)\right|^2\right) &\leqslant 2 \sum_{x \in B(0;2^n)} \frac{\mathrm{U}_X(B(x;2))}{\mathrm{U}_X(B(0;1))} \cdot \sum_{w \in B(0;2^{n+1})} \frac{\mathrm{U}_X(B(w;4))}{\mathrm{U}_X(B(0;2))} \\ &\leqslant K\mathrm{U}_X(B(0;2^{n+1})) \cdot \mathrm{U}_X(B(0;2^{n+2}) \\ &\leqslant 4^{3d} K[\mathrm{U}_X(B(0;2^n))]^2, \end{aligned}$$

where $K := 2[\mathrm{U}_X(B(0;1))\mathrm{U}_X(B(0;2))]^{-1}$ and the last line follows from (10). Therefore, the Paley–Zygmund inequality and (13) together imply that

$$\begin{aligned} \mathrm{P}\left\{\left|\operatorname{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right| > \frac{1}{2}\mathrm{U}_X(B(0;2^n))\right\} &\geqslant \frac{\left(\mathrm{E}\left|\operatorname{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right|\right)^2}{4\mathrm{E}\left(\left|\operatorname{pix}\left(\overline{\mathcal{R}_X}\right) \cap B(0;2^n)\right|^2\right)} \\ &\gtrsim 1, \end{aligned}$$

uniformly in n. The preceding and (12) together imply that $\mathrm{P}\{\mathrm{Dim}_{\mathrm{M}}(\overline{\mathcal{R}_X}) \geqslant \alpha_c\} > 0$ and hence $\mathrm{P}\{\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) \geqslant \alpha_c\} > 0$ thanks to Lemma 2.4. Since the event $\{\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) \geqslant \alpha_c\}$ is a tail event for the Lévy process X, the Kolmogorov 0–1 law implies that $\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) \geqslant \alpha_c$ a.s. This verifies the theorem since the other bound was verified earlier in the proof. □

3.2 *Fourier Analysis*

It is well-known that the law of X is determined by a socalled *characteristic exponent* $\Psi_X : \mathbb{R}^d \to \mathbb{C}$, which can be defined via $\mathrm{E}\exp(iz \cdot X_t) = \exp(-t\Psi_X(z))$ for all $t \geqslant 0$ and $z \in \mathbb{R}^d$. In particular, one can prove from this that $\Psi_X(z) \neq 0$ for almost all $z \in \mathbb{R}^d$. This fact is used tacitly in the sequel.

We frequently use the well-known fact that $\mathrm{Re}\Psi_X(z) \geqslant 0$ for all $z \in \mathbb{R}^d$. To see this fact, let X' be an independent copy of X and note that $t \mapsto X_t - X'_t$ is a Lévy process with characteristic exponent $2\mathrm{Re}\Psi_X$. Since $X_1 - X'_1$ is a symmetric random variable, one can conclude the mentioned fact that $\mathrm{Re}\Psi_X \geqslant 0$.

Port and Stone [19] have proved, among other things, that the transience of X is equivalent to the convergence of the integral

$$I(\Psi_X) := \int_{\|z\| \leqslant 1} \mathrm{Re}\left(\frac{1}{\Psi_X(z)}\right) \mathrm{d}z;$$

see also [3]. The following shows that the macroscopic dimension of the range of X is determined by the strength by which the Port–Stone integral $I(\Psi_X)$ converges.

Theorem 3.9 *With probability one,*

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) = \inf\left\{\alpha > 0 : \int_{\|z\|\leqslant 1} \mathrm{Re}\left(\frac{1}{\Psi_X(z)}\right)\frac{\mathrm{d}z}{\|z\|^{d-\alpha}} < \infty\right\}.$$

The proof of Theorem 3.9 hinges on a calculation from classical Fourier analysis. From now on, $\widehat{h}$ denotes the Fourier transform of a locally integrable function $h : \mathbb{R}^d \to \mathbb{R}$, normalized so that

$$\widehat{h}(z) = \int_{\mathbb{R}^d} \mathrm{e}^{iz\cdot x} h(x)\,\mathrm{d}x \qquad \text{for all } z \in \mathbb{R}^d \text{ and } h \in L^1(\mathbb{R}^d).$$

As is done customarily, we let K_ν denote the modified Bessel function [Macdonald function] of the second kind.

Lemma 3.10 *Choose and fix $\alpha > 0$ and define $f(x) := (1 + \|x\|^2)^{-\alpha/2}$ for all $x \in \mathbb{R}^d$. Then, the Fourier transform of f is*

$$\widehat{f}(z) = c_{d,\alpha} \cdot \frac{K_{(d-\alpha)/2}(\|z\|)}{\|z\|^{(d-\alpha)/2}} \qquad [z \in \mathbb{R}^d],$$

where $0 < c_{d,\alpha} < \infty$ depends only on $(d\,, \alpha)$.

Proof This is undoubtedly well known; the proof hinges on a simple abelian trick that can be included with little added effort.

For all $x \in \mathbb{R}^d$ and $\theta > 0$,

$$\int_0^\infty \mathrm{e}^{-t(1+\|x\|^2)} t^{\theta-1}\,\mathrm{d}t = \frac{\Gamma(\theta)}{(1 + \|x\|^2)^\theta}.$$

Therefore, for every rapidly decreasing test function $\varphi : \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned}\int_{\mathbb{R}^d} \frac{\varphi(x)}{(1 + \|x\|^2)^\theta}\,\mathrm{d}x &= \frac{1}{\Gamma(\theta)} \int_{\mathbb{R}^d} \varphi(x)\,\mathrm{d}x \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t(1+\|x\|^2)} t^{\theta-1}\\ &= \frac{1}{\Gamma(\theta)} \int_0^\infty \mathrm{e}^{-t} t^{\theta-1}\,\mathrm{d}t \int_{\mathbb{R}^d} \varphi(x)\mathrm{e}^{-t\|x\|^2}\,\mathrm{d}x.\end{aligned}$$

Since

$$\int_{\mathbb{R}^d} \varphi(x)\mathrm{e}^{-t\|x\|^2}\,\mathrm{d}x = (4\pi t)^{-d/2} \int_{\mathbb{R}^d} \overline{\widehat{\varphi}(z)} \exp\left(-\frac{\|z\|^2}{4t}\right)\mathrm{d}z,$$

it follows that

$$\int_{\mathbb{R}^d} \frac{\varphi(x)}{(1+\|x\|^2)^\theta}\,\mathrm{d}x = \frac{1}{b}\int_{\mathbb{R}^d} \overline{\widehat{\varphi}(z)}\,\mathrm{d}z \int_0^\infty \frac{\mathrm{d}t}{t^{(d/2)-\theta+1}} \exp\left(-t - \frac{\|z\|^2}{4t}\right)\mathrm{d}z$$
$$= \frac{1}{c}\int_{\mathbb{R}^d} \overline{\widehat{\varphi}(z)}\,\frac{K_{(d/2)-\theta}(\|z\|)}{\|z\|^{(d/2)-\theta}}\,\mathrm{d}z,$$

where $b := (4\pi)^{d/2}\Gamma(\theta)$ and $c := (4\pi)^{d/2}\Gamma(\theta)2^{-1-(d/2)+\theta}$. This proves the result, after we set $\theta := \alpha/2$. □

Proof of Theorem 3.9 It is not hard to check (see, for example, Port and Stone [19]) that $\widehat{\mathrm{U}}_X(z) = 1/\Psi_X(z)$ for almost all $z \in \mathbb{R}^d$. Because $\mathrm{Re}(1/\Psi_X(z)) = \mathrm{Re}\Psi_X(z)/|\Psi_X(z)|^2 > 0$ a.e., Lemma 3.10 and a suitable form of the Plancherel's theorem together imply that

$$\int_{\mathbb{R}^d} \frac{\mathrm{U}_X(\mathrm{d}x)}{1+|x|^\alpha} \asymp \int_{\mathbb{R}^d} \frac{\mathrm{U}_X(\mathrm{d}x)}{(1+|x|^2)^{\alpha/2}} \propto \int_{\mathbb{R}^d} \mathrm{Re}\left(\frac{1}{\Psi_X(z)}\right)\frac{K_{(d-\alpha)/2}(\|z\|)}{\|z\|^{(d-\alpha)/2}}\,\mathrm{d}z := T_1 + T_2,$$

where T_1 denotes the preceding integral with domain of integration restricted to $\{z \in \mathbb{R}^d : |\Psi_X(z)| < 1\}$ and T_2 is the same integral over $\{z \in \mathbb{R}^d : |\Psi_X(z)| \geqslant 1\}$.

A standard application of Laplace's method shows that for all $R > 0$ there exists a finite $A > 1$ such that

$$\frac{\mathrm{e}^{-w}}{A\sqrt{w}} \leqslant K_\nu(w) \leqslant \frac{A\mathrm{e}^{-w}}{\sqrt{w}},$$

whenever $w > R$. And one can check directly that for all $R > 0$ we can find a finite $B > 1$ such that

$$B^{-1}w^{-\nu} \leqslant K_\nu(w) \leqslant Bw^{-\nu} \qquad \text{whenever } 0 < w < R.$$

Since $\Psi_X : \mathbb{R}^d \to \mathbb{C}$ is a continuous function that vanishes at the origin, $\{z \in \mathbb{R}^d : |\Psi_X(z)| > 1\}$ does not intersect a certain ball about the origin of $\mathbb{R}^d$. Therefore, the inequality $\mathrm{Re}(1/\Psi_X(z)) \leqslant |\Psi_X(z)|^{-1}$, valid for all $z \in \mathbb{R}^d$, implies that

$$T_1 \asymp \int_{|\Psi_X(z)|<1} \mathrm{Re}\left(\frac{1}{\Psi_X(z)}\right)\frac{\mathrm{d}z}{\|z\|^{d-\alpha}},$$

and

$$T_2 \asymp \int_{|\Psi_X(z)|\geqslant 1} \mathrm{Re}\left(\frac{1}{\Psi_X(z)}\right)\frac{\mathrm{e}^{-\|z\|}}{\|z\|^{(d-\alpha+1)/2}}\,\mathrm{d}z \leqslant \int_{|\Psi_X(z)|\geqslant 1} \frac{\mathrm{e}^{-\|z\|}}{\|z\|^{(d-\alpha+1)/2}}\,\mathrm{d}z < \infty.$$

This verifies that

$$\int_{\mathbb{R}^d} \frac{\mathrm{U}_X(\mathrm{d}x)}{1+|x|^\alpha} < \infty \qquad \Longleftrightarrow \qquad T_1 < \infty,$$

which completes the theorem in light of Theorem 3.1 and a real-variable argument that implies that $T_1 < \infty$ iff $\int_{\|z\|\leqslant 1} \mathrm{Re}(1/\Psi_X(z))\|z\|^{-d+\alpha}\,\mathrm{d}z < \infty$. □

3.3 The Graph of a Lévy Process

Let $X := \{X_t\}_{t\geqslant 0}$ denote an arbitrary Lévy process on $\mathbb{R}^d$, not necessarily transient. It is easy to check that

$$Y_t := (t\,, X_t) \qquad [t \geqslant 0]$$

is a transient Lévy process in $\mathbb{R}^{d+1}$. Moreover,

$$\mathcal{G}_X := \mathcal{R}_Y$$

is the graph of the original Lévy process X. The literature on Lévy processes contains several results about the microscopic structure of $\mathcal{G}_X$. Perhaps the most noteworthy result of this type is the fact that

$$\dim_{\mathrm{H}}(\mathcal{G}_X) = 3/2 \qquad \text{a.s.,} \tag{14}$$

when X denotes a one-dimensional Brownian motion. In this section we compute the macroscopic Minkowski dimension of the same random set; in fact, we plan to compute the macroscopic Minkowski dimension of the graph of a large class of Lévy processes X.

The potential measure of the space-time process Y is, in general,

$$\mathrm{U}_Y(A\times B) := \mathrm{E}\left[\int_0^\infty \mathbb{1}_{A\times B}(s\,,X_s)\,\mathrm{d}s\right] = \int_A P_s(B)\,\mathrm{d}s,$$

for all Borel sets $A \subseteq \mathbb{R}_+$ and $B \subseteq \mathbb{R}^d$, where

$$P_s(B) := \mathrm{P}\{X_s \in B\}.$$

Therefore, Theorem 3.1 implies that

$$\mathrm{Dim}_{\mathrm{M}}\mathcal{G}_X = \inf\left\{\alpha > 0 : \int_0^\infty \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{1+s^\alpha+|x|^\alpha} < \infty\right\} \qquad \text{a.s.}$$

In order to understand what this formula says, let us first prove the following result.

Lemma 3.11 *If X is an arbitrary Lévy process on* $\mathbb{R}^d$*, then*

$$0 \leqslant \mathrm{Dim}_{\mathrm{M}}(\mathcal{G}_X) \leqslant 1 \qquad a.s.$$

Proof Since

$$\int_0^1 \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{1 + s^\alpha + |x|^\alpha} \leqslant \int_0^1 \mathrm{d}s \int_{\mathbb{R}^d} P_s(\mathrm{d}x) = 1,$$

it follows that

$$\mathrm{Dim}_{\mathrm{M}} \mathcal{G}_X = \inf\left\{\alpha > 0 : \int_1^\infty \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{s^\alpha + |x|^\alpha} < \infty\right\} \qquad \text{a.s.}$$

The proposition follows because

$$\int_1^\infty \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{s^\alpha + |x|^\alpha} \leqslant \int_1^\infty \frac{\mathrm{d}s}{s^\alpha} < \infty,$$

whenever $\alpha > 1$. □

It is possible to also show that, in a large number of cases, the graph of a Lévy process has macroscopic Minkowski dimension one, viz.,

Proposition 3.12 *Let X be a Lévy process on* $\mathbb{R}^d$ *such that* $X_1 \in L^1(\mathrm{P})$ *and* $\mathrm{E}(X_1) = 0$*. Then,* $\mathrm{Dim}_{\mathrm{M}}(\mathcal{G}_X) = 1$ *a.s.*

Therefore, we can see from Lemma 3.12 that the graph of one-dimensional Brownian motion has macroscopic Minkowski dimension 1, yet it has microscopic Hausdorff dimension 3/2; compare with (14).

Proof Lemma 3.11 implies that

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{G}_X) = \inf\left\{0 < \alpha \leqslant 1 : \int_1^\infty \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{s^\alpha + |x|^\alpha} < \infty\right\} \qquad \text{a.s.,} \tag{15}$$

where $\inf \varnothing := 1$. If $0 < \alpha < 1$, then

$$\int_1^\infty \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{s^\alpha + |x|^\alpha} \geqslant \int_1^\infty \mathrm{d}s \int_{|x| \leqslant s} \frac{P_s(\mathrm{d}x)}{s^\alpha + |x|^\alpha} \geqslant 2^{-\alpha} \int_1^\infty \mathrm{P}\{|X_s| \leqslant s\} \frac{\mathrm{d}s}{s^\alpha}.$$

Because $\mathrm{E}(X_1) = 0$, the law of large numbers for Lévy processes (see, for example, Bertoin [3, pp. 40–41]) implies that $\mathrm{P}\{|X_s| \leqslant s\} \to 1$ as $s \to \infty$. This shows that

$$\int_1^\infty \mathrm{P}\{|X_s| \leqslant s\} \frac{\mathrm{d}s}{s^\alpha} = \infty \quad \text{for every } \alpha \in (0\,,1),$$

and proves the lemma. □

Remark 3.13 The assumption $X_1 \in L^1(\mathrm{P})$ in Proposition 3.12 can be weakened. From the last part of the proof, we see that the conclusion of Proposition 3.12 still holds if there is a constant $c > 0$ such that $\mathrm{P}\{|X_s| \leqslant s\} \geqslant c$ for all $s \geqslant 1$. This is the case, for example, when X is a symmetric Cauchy process.

Finally, let us prove that the preceding result is unimprovable in the following sense: For every number $q \in (0\,,1)$, there exist a Lévy process X on $\mathbb{R}^d$ the macroscopic dimension of whose graph is q.

Theorem 3.14 *If X be a symmetric β-stable Lévy process on $\mathbb{R}^d$ for some $0 < \beta \leqslant 2$, then*

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{G}_X) = \beta \wedge 1 \qquad a.s.$$

The preceding is a large-scale analogue of a result due to McKean [15]. McKean's theorem asserts that with probability one, the (microscopic) Hausdorff dimension of the *range* (not graph!) of a *real-valued*, symmetric β-stable Lévy process is $\beta \wedge 1$.

Proof If $\beta > 1$, then the result follows from Proposition 3.12. When $\beta = 1$, the process X is a symmetric Cauchy process and the result follows from Remark 3.13. In the remainder of the proof we assume that $0 < \beta < 1$.

Let us observe the elementary estimate,

$$\begin{aligned}\int_1^\infty \mathrm{d}s \int_{\mathbb{R}^d} \frac{P_s(\mathrm{d}x)}{s^\alpha + |x|^\alpha} &\asymp \int_1^\infty \mathrm{d}s \int_{|x|<s} \frac{P_s(\mathrm{d}x)}{s^\alpha} + \int_1^\infty \mathrm{d}s \int_{|x|\geqslant s} \frac{P_s(\mathrm{d}x)}{|x|^\alpha} \\ &=: \mathcal{T}_1 + \mathcal{T}_2.\end{aligned} \tag{16}$$

For all $0 < \alpha < 1$,

$$\mathcal{T}_1 = \int_1^\infty \mathrm{P}\{|X_s| < s\} \frac{\mathrm{d}s}{s^\alpha} = \int_1^\infty \mathrm{P}\left\{|X_1| < s^{-(1-\beta)/\beta}\right\} \frac{\mathrm{d}s}{s^\alpha},$$

by scaling. It is well known that X_1 has a bounded, continuous, and strictly positive density function on $\mathbb{R}^d$. This shows that $\mathrm{P}\{|X_1| < s^{-(1-\beta)/\beta}\}$ is bounded above and below by constant multiples of $s^{-(1-\beta)/\beta}$, uniformly for all $s > 1$. In particular,

$$\mathcal{T}_1 < \infty \quad \text{iff} \quad 1 > \alpha > 2 - \beta^{-1}. \tag{17}$$

Next we note that if $0 < \alpha < 1$, then

$$\mathcal{T}_2 = \int_1^\infty \mathrm{E}\left(|X_1|^{-\alpha};\, |X_1| \geqslant s^{1-(1/\beta)}\right) \frac{\mathrm{d}s}{s^{\alpha/\beta}},$$

by scaling. Because X_1 has a strictly positive and bounded density in $\mathbb{R}^d$, the inequalities

$$\mathrm{E}\left(|X_1|^{-\alpha};\ |X_1| \geqslant 1\right) \leqslant \mathrm{E}\left(|X_1|^{-\alpha};\ |X_1| \geqslant s^{1-(1/\beta)}\right) \leqslant \mathrm{E}\left(|X_1|^{-\alpha}\right)$$

imply that

$$\mathcal{T}_2 < \infty \quad \text{iff} \quad \beta < \alpha. \tag{18}$$

Hence, we have shown that $\mathcal{T}_1 + \mathcal{T}_2 < \infty$ iff $\beta < \alpha$. The theorem follows from (15) to (18). □

3.4 Application to Subordinators

Let us now consider the special case that the Lévy process X is a subordinator. To be concrete, by the latter we mean that X is a Lévy process on $\mathbb{R}$ such that $X_0 = 0$ and the sample function $t \mapsto X_t$ is a.s. nondecreasing. If we assume further that $\mathrm{P}\{X_1 > 0\} > 0$, then it follows readily that $\lim_{t\to\infty} X_t = \infty$ a.s. and hence X is transient. As is customary, one prefers to study subordinators via their *Laplace exponent* $\Phi_X : \mathbb{R}_+ \to \mathbb{R}_+$. The Laplace exponent of X is defined via the identity

$$\mathrm{E}\exp(-\lambda X_t) = \exp(-t\Phi_X(\lambda)),$$

valid for all $t, \lambda \geqslant 0$. It is easy to see that $\Phi_X(\lambda) = \Psi_X(i\lambda)$, where Ψ_X now denotes [the analytic continuation, from $\mathbb{R}$ to $i\mathbb{R}$, of] the characteristic exponent of X.

Theorem 3.15 *If $\Phi_X : \mathbb{R}_+ \to \mathbb{R}_+$ denote the Laplace exponent of a subordinator X on $\mathbb{R}_+$, then*

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) = \inf\left\{0 < \alpha < 1 : \int_0^\infty \frac{\mathrm{d}y}{y^{1-\alpha}\,\Phi_X(y)} < \infty\right\} \qquad a.s.,$$

where $\inf \varnothing := 1$.

Theorem 3.15 is the macroscopic analogue of a theorem of Horowitz [9] (see also [4] for more results) which gave a formula for the microscopic Hausdorff dimension of the range of a subordinator. The following highlights a standard application of subordinators to the study of level sets of Markov process; see Bertoin [4] for much more on this connection.

Example 3.16 Let X be a symmetric, β-stable process on $\mathbb{R}$ where $1 < \beta \leqslant 2$. It is well known that $X^{-1}\{0\} := \{s > 0 : X_s = 0\}$ is a.s. nonempty, and coincides with the closure of the range of a stable subordinator $T := \{T_t\}_{t\geqslant 0}$ of index $1 - \beta^{-1}$.

It follows from Lemma 2.4 and Theorem 3.15 that

$$\mathrm{Dim}_{\mathrm{M}}\left(X^{-1}\{0\}\right) = \inf\left\{0<\alpha<1 : \int_0^1 \frac{\mathrm{d}t}{y^{1-\alpha+1-(1/\beta)}} < \infty\right\} = 1-\frac{1}{\beta} \qquad \text{a.s.} \tag{19}$$

Notice that (19) is analogous to the microscopic fractal dimension result for the zero set $X^{-1}\{0\}$. This is due to the fact that the Laplace exponent of the corresponding stable subordinator is a homogeneous function, which has the same asymptotic behavior at the origin and the infinity. For a Lévy process whose characteristic exponent has different asymptotic behaviors at the origin and the infinity, the macroscopic and microscopic fractal dimensions of the zero set may be different.

Proof of Theorem 3.15 The proof uses as its basis an old idea which is basically a "change of variables for subordinators," and is loosely connected to Bochner's method of subordination [5]. Before we get to that, let us observe first that Theorem 3.1 readily implies that

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{R}_X) = \inf\left\{0<\alpha<1 : \int_0^\infty x^{-\alpha}\,\mathrm{U}_X(\mathrm{d}x) < \infty\right\} \qquad \text{a.s.}$$

Now let us choose and fix some $\alpha \in (0\,,1)$, and let $Y := \{Y_s\}_{s\geqslant 0}$ be an independent α-stable subordinator, normalized to satisfy $\Phi_Y(x) = x^\alpha$ for every $x \geqslant 0$. Since $x^{-\alpha} = \int_0^\infty \exp(-sx^\alpha)\,\mathrm{d}s = \int_0^\infty \mathrm{E}\exp(-xY_s)\,\mathrm{d}s$, a few back-to-back appeals to the Tonelli theorem yield the following probabilistic change-of-variables formula[1]:

$$\begin{aligned}\int_0^\infty x^{-\alpha}\,\mathrm{U}_X(\mathrm{d}x) &= \mathrm{E}\left[\int_0^\infty \mathrm{U}_X(\mathrm{d}x)\int_0^\infty \mathrm{d}s\; \mathrm{e}^{-xY_s}\right] = \mathrm{E}\left[\int_0^\infty \mathrm{d}t \int_0^\infty \mathrm{d}s\; \mathrm{e}^{-X_tY_s}\right]\\ &= \int_0^\infty \mathrm{d}t\int_0^\infty \mathrm{d}s\;\mathrm{E}\left[\mathrm{e}^{-t\Phi_X(Y_s)}\right] = \mathrm{E}\left[\int_0^\infty \frac{\mathrm{d}s}{\Phi_X(Y_s)}\right] = \int_0^\infty \frac{\mathrm{U}_Y(\mathrm{d}y)}{\Phi_X(y)}.\end{aligned}$$

It is well-known that $\mathrm{U}_Y(\mathrm{d}y) \ll \mathrm{d}y$ (or one can verify this directly using transition density or characteristic function of Y), and the Radon–Nikodym density $u_Y(y) := \mathrm{U}_Y(\mathrm{d}y)/\mathrm{d}y$—this is the socalled *potential density of* Y—is given by $u_Y(y) = cy^{-1+\alpha}$ for all $y > 0$, where $c = c(\alpha)$ is a positive and finite constant [this follows from the scaling properties of Y]. Consequently, we see that $\int_0^\infty x^{-\alpha}\,\mathrm{U}_X(\mathrm{d}x) < \infty$ for some $0 < \alpha < 1$ if and only if $\int_0^\infty y^{-1+\alpha}\,\mathrm{d}y/\Phi_X(y) < \infty$ for the same α. The theorem follows from this. □

[1] The same argument shows that if X and Y are independent subordinators, then we have the change-of-variables formula,

$$\int_0^\infty \frac{\mathrm{U}_X(\mathrm{d}x)}{\Phi_Y(x)} = \int_0^\infty \frac{\mathrm{U}_Y(\mathrm{d}y)}{\Phi_X(y)}.$$

4 Tall Peaks of Symmetric Stable Processes

Let $B = \{B_t\}_{t \geqslant 0}$ be a standard Brownian motion. For every $\alpha > 0$, let us consider the set

$$\mathcal{H}_B(\alpha) := \left\{ t \geqslant \mathrm{e} : \; B_t \geqslant \alpha \sqrt{2 \log \log t} \right\}, \tag{20}$$

where "log" denotes the natural logarithm. In the terminology of Khoshnevisan et al. [14], the random set $\mathcal{H}_B(\alpha)$ denotes the collection of *the tall peaks of B in length scale α*.

Theorem 4.1 below follows from the law of the iterated logarithm for Brownian motion for $\alpha \neq 1$. The critical case of $\alpha = 1$ follows from Motoo [16, Example 2].

Theorem 4.1 *$\mathcal{H}_B(\alpha)$ is a.s. unbounded if $0 < \alpha \leqslant 1$ and is a.s. bounded if $\alpha > 1$.*

Recently, Khoshnevisan et al. [14] showed that the macroscopic Hausdorff dimension of $\mathcal{H}_B(\alpha)$ is 1 almost surely if $\alpha \leqslant 1$. Since the macroscopic Hausdorff dimension never exceeds the Minkowski dimension (see Barlow and Taylor [2]) Theorem 4.1 implies the following.

Theorem 4.2 $\mathrm{Dim}_{\mathrm{M}}(\mathcal{H}_B(\alpha)) = 1$ *a.s. for every* $0 < \alpha \leqslant 1$.

Together, Theorems 4.1 and 4.2 imply that the tall peaks of Brownian motion are macroscopic monofractals in the sense that either $\mathrm{Dim}_{\mathrm{M}}(\mathcal{H}_B(\alpha)) = 1$ or $\mathrm{Dim}_{\mathrm{M}}(\mathcal{H}_B(\alpha)) = 0$. In this section we extend the above results to facts about all symmetric stable Lévy processes. However, we are quick to point out that the proofs, in the stable case, are substantially more delicate than those in the Brownian case.

Let $X = \{X_t\}_{t \geqslant 0}$ be a real-valued, symmetric β-stable Lévy process for some $\beta \in (0\,, 2)$. We have ruled out the case $\beta = 2$ since X is Brownian motion in that case, and there is nothing new to be said about X in that case. To be concrete, the process X will be scaled so that it satisfies

$$\mathrm{E} \exp(izX_t) = \exp(-t|z|^\beta) \qquad \text{for every } t \geqslant 0 \text{ and } z \in \mathbb{R}. \tag{21}$$

In analogy with (20), for every $\alpha > 0$, let us consider the following set

$$\mathcal{H}_X(\alpha) := \left\{ t \geqslant \mathrm{e} : \; X_t \geqslant t^{1/\beta} (\log t)^\alpha \right\}$$

of tall peaks of X, parametrized by a "scale factor" $\alpha > 0$. The following is a re-interpretation of a classical result of Khintchine [10].

Theorem 4.3 *$\mathcal{H}_X(\alpha)$ is a.s. unbounded if $0 < \alpha \leqslant 1/\beta$, and it is a.s. bounded if $\alpha > 1/\beta$.*

We include a proof for the sake of completeness.

Proof It suffices to study only the case that $\alpha > 1/\beta$. The other case follows from the stronger Theorem 4.4 below.

Recall from [3, p. 221] that

$$\varrho := \lim_{\lambda\to\infty} \lambda^{\beta} \mathrm{P}\{X_1 > \lambda\} \tag{22}$$

exists and is in $(0\,,\infty)$. Consequently,

$$\mathrm{P}\{X_t > t^{1/\beta}\lambda\} \asymp \lambda^{-\beta} \qquad \text{for all } \lambda \geqslant 1 \text{ and } t > 0. \tag{23}$$

Let

$$X_t^* := \sup_{0\leqslant s\leqslant t} X_s \qquad \text{for all } t \geqslant 0.$$

The standard argument that yields the classical reflection principle also yields

$$\mathrm{P}\left\{X_t^* \geqslant \lambda\right\} \leqslant 2\mathrm{P}\left\{X_t \geqslant \lambda\right\} \qquad \text{for all } t, \lambda > 0.$$

Therefore, (23) implies that

$$\mathrm{P}\left\{X_t^* \geqslant \varepsilon t^{1/\beta}(\log t)^{\alpha}\right\} \leqslant 2\mathrm{P}\left\{X_t \geqslant \varepsilon t^{1/\beta}(\log t)^{\alpha}\right\} \asymp (\log t)^{-\alpha\beta},$$

for all $t \geqslant \mathrm{e}$ and $\varepsilon > 0$. This and the Borel–Cantelli lemma together show that, if $\alpha > 1/\beta$, then $X_t = o(t^{1/\beta}(\log t)^{\alpha})$ as $t \to \infty$, a.s. In other words, $\mathcal{H}_X(\alpha)$ is a.s. bounded if $\alpha > 1/\beta$. This completes the proof. □

Theorem 4.3 reduces the analysis of the peaks of X to the case where $\alpha \in (0\,, 1/\beta]$. That case is described by the following theorem, which is the promised extension of Theorem 4.2 to the stable case.

Theorem 4.4 *If* $0 < \alpha \leqslant 1/\beta$, *then* $\mathrm{Dim}_{\mathrm{M}}(\mathcal{H}_X(\alpha)) = 1$ *a.s.*

Proof It suffices to prove that

$$\mathrm{Dim}_{\mathrm{M}}(\mathcal{H}_X(\alpha)) \geqslant 1 \qquad \text{a.s.} \tag{24}$$

Throughout the proof, we choose and fix a constant $\gamma \in (0\,, 1)$. Let us define an increasing sequence $T_1, T_2, \ldots,$ where

$$T_j := 2^{2\beta j^{\gamma}/\gamma} = \exp\left(\frac{\beta \log(4) j^{\gamma}}{\gamma}\right).$$

Let us also introduce a collection of intervals $\mathcal{I}(1), \mathcal{I}(2), \ldots,$ defined as follows:

$$\mathcal{I}(j) := \left[T_j^{1/\beta}(\log T_j)^{\alpha}\ ,\ 2T_j^{1/\beta}(\log T_j)^{\alpha}\right).$$

Finally, let us introduce events $\mathscr{E}_1, \mathscr{E}_2, \ldots$, where

$$\mathscr{E}_j := \left\{\omega \in \Omega : X_{T_j}(\omega) \in \mathcal{I}(j)\right\}.$$

According to (22),

$$\begin{aligned} \mathrm{P}(\mathscr{E}_j) &= \mathrm{P}\left\{X_{T_j} \geqslant T_j^{1/\beta}(\log T_j)^{\alpha}\right\} - \mathrm{P}\left\{X_{T_j} \geqslant 2T_j^{1/\beta}(\log T_j)^{\alpha}\right\} \\ &\sim \frac{\varrho\left(1-2^{-\beta}\right)}{(\log T_j)^{\alpha\beta}} \qquad [\text{as } j \to \infty] \\ &= \frac{\varrho\left(1-2^{-\beta}\right)}{j^{\alpha\gamma\beta}}. \end{aligned}$$

For every integer $n \geqslant 1$, let us define

$$W_n := \sum_{j=2^{n-1}}^{2^n-1} \mathbb{1}_{\mathscr{E}_j}.$$

It follows from the preceding that there exists an integer $n_0 \geqslant 1$ such that

$$\mathrm{E}(W_n) \gtrsim 2^{n(1-\alpha\beta\gamma)} \qquad \text{uniformly for all } n \geqslant n_0. \tag{25}$$

Next, we estimate $\mathrm{E}(W_n^2)$, which may be written in the following form:

$$\mathrm{E}(W_n^2) = \mathrm{E}(W_n) + 2 \sum\sum_{2^{n-1} \leqslant j < k < 2^n} \mathrm{P}(\mathscr{E}_j \cap \mathscr{E}_k). \tag{26}$$

Henceforth, suppose $k > j$ are two integers between 2^{n-1} and $2^n - 1$.

Because X has stationary independent increments,

$$\mathrm{P}(\mathscr{E}_j \cap \mathscr{E}_k) \leqslant \mathcal{P}_j \times \mathcal{P}_{j,k}, \tag{27}$$

where

$$\begin{aligned} \mathcal{P}_j &= \mathrm{P}\left\{X_{T_j} \geqslant T_j^{1/\beta}(\log T_j)^{\alpha}\right\}, \\ \mathcal{P}_{j,k} &= \mathrm{P}\left\{X_{T_k - T_j} \geqslant T_k^{1/\beta}(\log T_k)^{\alpha} - 2T_j^{1/\beta}(\log T_j)^{\alpha}\right\}. \end{aligned}$$

In accord with (23),

$$\mathcal{P}_j = \mathrm{P}(\mathscr{E}_j) \asymp j^{-\alpha\beta\gamma}. \tag{28}$$

The analysis of $\mathcal{P}_{j,k}$ is somewhat mode complicated.

First, one might observe that

$$\begin{aligned}\mathcal{P}_{j,k} &= \mathrm{P}\left\{X_{T_k-T_j} \geqslant (\log T_k)^{\alpha}\left[T_k^{1/\beta} - 2T_j^{1/\beta}\right]\right\} \\ &= \mathrm{P}\left\{X_1 \geqslant k^{\alpha\gamma}\,\frac{T_k^{1/\beta} - 2T_j^{1/\beta}}{(T_k - T_j)^{1/\beta}}\right\} \quad \text{[by scaling]} \\ &\leqslant \mathrm{P}\left\{X_1 \geqslant k^{\alpha\gamma}\left[1 - 2\left(\frac{T_j}{T_k}\right)^{1/\beta}\right]\right\}.\end{aligned} \tag{29}$$

[The final inequality holds simply because $(T_k - T_j)^{1/\beta} \leqslant T_k^{1/\beta}$.]

If j and k are integers in $[2^{n-1}, 2^n)$ that satisfy $j \leqslant k - k^{1-\gamma}$, then

$$k^{\gamma} - j^{\gamma} = k^{\gamma}\left[1 - \left(\frac{j}{k}\right)^{\gamma}\right] \geqslant k^{\gamma}\left[1 - (1 - k^{-\gamma})^{\gamma}\right] \geqslant \gamma.$$

The preceding is justified by the following elementary inequality: $(1-x)^{\gamma} \leqslant 1 - \gamma x$ for all $x \in (0,1)$. As a result, we are led to the following bound:

$$1 - 2\left(\frac{T_j}{T_k}\right)^{1/\beta} = 1 - 2\exp\left(-\frac{2\log 2}{\gamma}\left[k^{\gamma} - j^{\gamma}\right]\right) \geqslant \frac{1}{2},$$

valid uniformly for all integers j and k that satisfy $k > j \geqslant k - k^{1-\gamma}$ and are between 2^{n-1} and $2^n - 1$. Therefore, (23) and (29) together imply that

$$\mathcal{P}_{j,k} \leqslant \mathrm{P}\{X_1 \geqslant k^{\alpha\gamma}\} \lesssim k^{-\alpha\beta\gamma},$$

uniformly for all integers $k > j$ that are in $[2^{n-1}, 2^n - 1)$ and satisfy $j \leqslant k - k^{1-\gamma}$, and uniformly for every integer $n \geqslant n_0$. It follows from this bound, (27), and (28) that

$$\sum\sum_{\substack{2^{n-1}\leqslant j<k<2^n \\ j\leqslant k-k^{1-\gamma}}} \mathrm{P}(\mathcal{E}_j \cap \mathcal{E}_k) \lesssim \sum\sum_{\substack{2^{n-1}\leqslant j<k<2^n \\ j\leqslant k-k^{1-\gamma}}} (jk)^{-\alpha\beta\gamma} \lesssim 4^{n(1-\alpha\beta\gamma)}, \tag{30}$$

uniformly for all integers $n \geqslant n_0$.

On the other hand,

$$\sum\sum_{\substack{2^{n-1}\leqslant j<k<2^n \\ j> k-k^{1-\gamma}}} \mathrm{P}(\mathcal{E}_j \cap \mathcal{E}_k) \leqslant \sum\sum_{\substack{2^{n-1}\leqslant j<k<2^n \\ j> k-k^{1-\gamma}}} \mathrm{P}(\mathcal{E}_k) \lesssim \sum\sum_{\substack{2^{n-1}\leqslant j<k<2^n \\ j> k-k^{1-\gamma}}} k^{-\alpha\beta\gamma} \qquad \text{[by (28)]}$$

$$\lesssim \sum_{k=2^{n-1}}^{2^n-1} k^{1-\gamma-\alpha\beta\gamma} \lesssim 2^{n(2-\gamma-\alpha\beta\gamma)}$$
$$\leqslant 4^{n(1-\alpha\beta\gamma)},$$

since $\alpha\beta \leqslant 1$. Therefore, (30) implies that

$$\sum\sum_{2^{n-1}\leqslant j<k<2^n} \mathrm{P}(\mathscr{E}_j \cap \mathscr{E}_k) \lesssim 4^{n(1-\alpha\beta\gamma)}.$$

This and (26) together imply that

$$\mathrm{E}(W_n^2) \leqslant \mathrm{E}(W_n) + (\mathrm{E}[W_n])^2\,, \tag{31}$$

uniformly for all $n \geqslant n_0$. Because of (25) and the condition $\alpha\beta \leqslant 1$, it follows that $\mathrm{E}(W_n) \gtrsim 1$, uniformly for all $n \geqslant 1$. Therefore, there exists a finite and positive constant c such that

$$\mathrm{E}(W_n^2) \leqslant c\,(\mathrm{E}[W_n])^2 \qquad \text{for all } n \geqslant n_0.$$

An appeal to the Paley–Zygmund inequality then yields the following: Uniformly for all integers $n \geqslant n_0$,

$$\inf_{n\geqslant n_0} \mathrm{P}\left\{W_n > \tfrac{1}{2}\mathrm{E}(W_n)\right\} \geqslant (4c)^{-1}.$$

From this and (25) it immediately follows that

$$\mathrm{P}\left\{\limsup_{n\to\infty} n^{-1}\mathrm{Log}_+ W_n \geqslant 1 - \alpha\beta\gamma\right\} \geqslant (4c)^{-1} > 0.$$

The event in the preceding event is a tail event for the Lévy process X. Therefore, the Kolmogorov 0–1 law implies that

$$\limsup_{n\to\infty} n^{-1}\mathrm{Log}_+ W_n \geqslant 1 - \alpha\beta\gamma \qquad \text{a.s.}$$

Because $\gamma \in (0\,,1)$ was arbitrary, this proves that $\limsup_{n\to\infty} n^{-1}\mathrm{Log}_+ W_n \geqslant 1$ a.s., and (24) follows since $\mathrm{Dim}_{\mathrm{M}}(\mathcal{H}_X(\alpha)) \geqslant \limsup_{n\to\infty} n^{-1}\mathrm{Log}_+ W_n$. This completes the proof of the theorem. □

Acknowledgements Research supported in part by the National Science Foundation grant DMS-1307470, DMS-1608575 and DMS-1607089.

References

1. M.T. Barlow, S.J. Taylor, Fractional dimension of sets in discrete spaces. J. Phys. A **22**, 2621–2626 (1989)
2. M.T. Barlow, S.J. Taylor, Defining fractal subsets of $\mathbb{Z}^d$. Proc. Lond. Math. Soc. **64**, 125–152 (1992)
3. J. Bertoin, *Lévy Processes* (Cambridge University Press, Cambridge, 1996)
4. J. Bertoin, Subordinators: examples and applications, in *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*. Lecture Notes in Mathematics, vol. 1717 (Springer, Berlin, 1999), pp. 1–91
5. S. Bochner, *Harmonic Analysis and the Theory of Probability* (University of California Press, Berkeley and Los Angeles, 1955)
6. K.J. Falconer, *Fractal Geometry–Mathematical Foundations and Applications*, 2nd edn. (Wiley, New York, 2003)
7. P. Hall, *Introduction to the Theory of Coverage Processes* (Wiley, New York, 1988)
8. J. Hawkes, Trees generated by a simple branching process. J. Lond. Math. Soc. (2) **24**, 373–384 (1981)
9. J. Horowitz, The Hausdorff dimension of the sample path of a subordinator. Isr. J. Math. **6**, 176–182 (1968)
10. A. Khintchine, Sur la croissance locale des processus stochastiques homogénes à accroissements indépendants (Russian). Bull. Acad. Sci. URSS. Sér. Math. [Izvestia Akad. Nauk SSSR] **3**, 487–508 (1939)
11. D. Khoshnevisan, Y. Xiao, Lévy processes: capacity and Hausdorff dimension. Ann. Probab. **33**, 841–878 (2005)
12. D. Khoshnevisan, Y. Xiao, Harmonic analysis of additive Lévy processes. Probab. Theory Relat. Fields **145**, 459–515 (2009)
13. D. Khoshnevisan, Y. Xiao, Y. Zhong, Measuring the range of an additive Lévy process. Ann. Probab. **31** 1097–1141 (2003)
14. D. Khoshnevisan, K. Kim, Y. Xiao, Intermittency and multifractality: a case study via parabolic stochastic PDEs. Ann. Probab. (to appear)
15. H.P. McKean Jr., Sample functions of stable processes. Ann. Math. **61**, 564–579 (1955)
16. M. Motoo, Proof of the law of the iterated logarithm through diffusion equation. Ann. Inst. Stat. Math. **10**(1), 21–28 (1959)
17. Y. Peres, Intersection equivalence of Brownian paths and certain branching processes. Commun. Math. Phys. **177**, 417–434 (1996)
18. Y. Peres, Remarks on intersection-equivalence and capacity-equivalence. Ann. Inst. Henri Poincaré (Physique théorique) **64**, 339–347 (1996)
19. S.C. Port, C.J. Stone, Infinitely divisible processes and their potential theory. I, II. Ann. Inst. Fourier (Grenoble) **21**(2), 157–275 (1971); Ann. Inst. Fourier (Grenoble) **21**(4), 179–265 (1971)
20. W.E. Pruitt, The Hausdorff dimension of the range of a process with stationary independent increments. J. Math. Mech. **19**, 371–378 (1969)
21. S.J. Taylor, The measure theory of random fractals. Math. Proc. Camb. Philos. Soc. **100**, 383–406 (1986)
22. Y. Xiao, Random fractals and Markov processes, in *Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot*, ed. by M.L. Lapidus, M. van Frankenhuijsen (American Mathematical Society, Providence, 2004), pp. 261–338

Critical Behavior of Mean-Field XY and Related Models

Kay Kirkpatrick and Tayyab Nawaz

Abstract We discuss spin models on complete graphs in the mean-field (infinite-vertex) limit, especially the classical XY model, the Toy model of the Higgs sector, and related generalizations. We present a number of results coming from the theory of large deviations and Stein's method, in particular, Cramér and Sanov-type results, limit theorems with rates of convergence, and phase transition behavior for these models.

Keywords Mean-field • Free energy • Density function • Gibbs measure

1 Introduction

We use mean-field theory to approximate a challenging problem and to study a many-body problem by converting it into a one-body problem. We survey a number of results obtained recently using the theory of large deviations along with Stein's method-type limit theorems to describe the asymptotic behavior of the $O(N)$ spin models such as the $N = 1$ Curie-Weiss model, the $N = 2$ model called the XY model, the $N = 3$ Heisenberg model, and the $N = 4$ Toy model of the Higgs sector [5, 12, 13]. We present these results mostly without proofs. In this section, we describe the mean-field XY model and the history, including the 2D XY model (which is currently intractable). In the next section we describe the asymptotic behavior of the XY model; in the last section, the behavior of its generalizations to N-dimensional spins.

The XY model on a complete graph K_n with n vertices in the absence of an external field is defined as follows: there is a circular spin $\sigma_i \in \mathbb{S}^1$ at each site $i \in 1, 2, \ldots, n$. The configuration space of the XY model is $\Omega_n = (\mathbb{S}^1)^n$ where each microstate is $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$. For the higher $O(N)$ spin models, we simply

K. Kirkpatrick • T. Nawaz (✉)
Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 W. Green Street, Urbana, IL 61801, USA

Department of Mathematics, COMSATS Institute of Information and Technology, Park Road, Islamabad 45550, Pakistan
e-mail: tnawaz2@illinois.edu

F. Baudoin, J. Peterson (eds.), *Stochastic Analysis and Related Topics*, Progress in Probability 72, DOI 10.1007/978-3-319-59671-6_10

replace $\mathbb{S}^1$ by $\mathbb{S}^{N-1}$, and in all cases the Hamiltonian energy is defined by

$$H_n(\sigma) = -\sum_{i,j} J_{i,j} \langle \sigma_i, \sigma_j \rangle .$$

The mean-field interaction for the XY and $O(N)$ models on the complete graph is defined by $J_{i,j} = \frac{1}{2n}$ for all i, j.

The simplest spin model is the Ising model, with one-dimensional ± 1 spins, a model that is used extensively in statistical mechanics, invented by Ernest Ising while working with his advisor Wilhelm Lenz [4, 11]. The one-dimensional Ising model has no phase transition, but there is a phase transition on an infinite two-dimensional lattice. The mean-field Ising model, or Curie-Weiss model, describes the Ising model well for higher dimensions, and the magnetization (average spin) in this model has a Gaussian law away from the critical temperature and a non-Gaussian law at the critical temperature [7]. Recently, Chatterjee and Shao [5] proved that the total spin in the Curie-Weiss model at the critical temperature satisfies a Berry–Esseen type error bound in this non-Gaussian limit.

The XY model, with two-dimensional circular spins, models superconductors and is interesting but challenging to study its phase transition rigorously [3]. On a lattice of two spatial dimensions, such a continuous circular symmetry cannot be broken at any finite temperature [16]. Thus the 2D XY model cannot have an ordered phase at low temperature quite like the Ising model, and it has a phase transition that is quite different from the Ising model [17, 18]. Instead, the 2D XY model exhibits the peculiar Kosterlitz-Thouless (KT) transition, a phase transition of infinite order and the subject of a Nobel prize. Above the transition temperature T_{KT} correlations between spins decay exponentially. At low temperatures, the system does not have any long-range order as the ground state is unstable, but there is a low-temperature quasi-ordered phase with a correlation function that decreases with the distance like a power, which depends on the temperature [14].

Because the 2D XY model is so challenging, we study the mean-field classical XY model instead, which can be viewed as the large-dimensional ($d \to \infty$) limit of the nearest-neighbor model on $\mathbb{Z}^d$, with spins in $\mathbb{S}^1$, and with critical inverse temperature $\beta_c = 2$ [1]. Furthermore, the large-dimensional limit approximates high-dimensional models nicely since below the critical temperature, the average spin goes to zero for all d, and above the critical temperature, the total spin has a non-zero limit as $d \to \infty$.

In the next section we will examine the XY model in detail, while Sect. 3 deals with extensions to higher spin dimensions.

2 The Mean-Field XY Model and Asymptotic Results

We consider the isotropic mean-field classical XY model on a finite complete graph K_n with n vertices. That is, at each site $i \in K_n$ of the graph is a spin living in $\Omega = \mathbb{S}^1$, so the state space is $\Omega_n = (\mathbb{S}^1)^n$. See Fig. 1 for a picture of the XY model

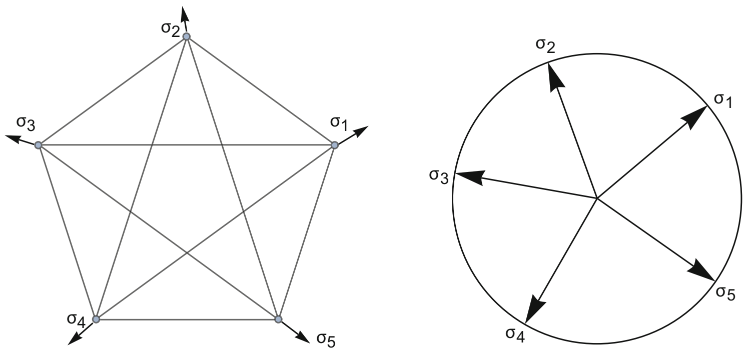

Fig. 1 *Left*: The classical mean-field XY Model on the complete graph K_5 with five sample spin vectors. *Right*: The projection of the same spin vectors from K_5 onto $\mathbb{S}^1$

on 5 vertices. The corresponding mean-field Hamiltonian energy $H_n : \Omega_n \to \mathbb{R}$ is given by:

$$H_n(\sigma) := -\frac{1}{2n} \sum_{i,j=1}^{n} \langle \sigma_i, \sigma_j \rangle = -\frac{1}{2n} \sum_{i,j} \cos(\theta_i - \theta_j),$$

where θ_i is the angle that the i-th spin makes with respect to some axis. The corresponding Gibbs measure is the probability measure $P_{n,\beta}$ on Ω_n with density function:

$$f(\sigma) := \frac{1}{Z(\beta)} e^{-\beta H_n(\sigma)}. \tag{1}$$

where Z is the normalizing constant, also known as the partition function, which encodes the statistical properties of the model such as free energy and magnetization. Note that Gibbs measure here is a normalization of the Boltzmann distribution, and that the inverse temperature β is equal to $(k_B T)^{-1}$, where k_B is the Boltzmann constant and T is the temperature of the system. We can understand the structural behavior of the spin vectors' distribution by studying extreme cases for the inverse temperature β as follows:

- At high temperature, from Eq. (1) we can predict that the Gibbs measure is uniform.
- At low temperature, again from Eq. (1) we can predict that the Gibbs measure decays quickly, and the spin vector distribution prefers the lowest-energy ground state.

The most likely physical system states corresponding to the Gibbs measure are called the canonical macrostates. We will consider the random measure of the spins $\{\sigma_i\}$, defined by $\mu_{n,\sigma} := \frac{1}{n} \sum_{i=1}^{n} \delta_{\sigma_i}$ on $\mathbb{S}^1$ and study the total empirical spin, defined by

$$S_n(\sigma) := \sum_{i=1}^{n} \sigma_i.$$

The relative entropy of a probability measure ν on $\mathbb{S}^1$, with respect to the uniform probability measure μ is defined by

$$H(\nu \mid \mu) := \begin{cases} \int_{\mathbb{S}^1} f \log(f) d\mu & \text{if } f := \frac{d\nu}{d\mu} \text{ exists;} \\ \infty & \text{otherwise.} \end{cases} \tag{2}$$

2.1 LDPs, Free Energy, and Macrostates for the XY Model

Let $M_1(\mathbb{S}^1)$ represent the probability measures on $\mathbb{S}^1$ with the weak-$*$ topology. We are interested in analyzing the total spin, $S_n := \sum_{i=1}^{n} \sigma_i$, as a function of the inverse temperature β in the Gibbs measures. This leads us to consider large deviation principles (LDPs) for the $\mu_{n,\sigma}$, and then we can rewrite the free energy more explicitly to describe the phase transition at $\beta = 2$. Part of Theorem 1 ($\beta = 0$) is a special case of Sanov's theorem, while the other cases ($\beta > 0$) follow from an abstract result of Ellis, Haven, and Turkington [10, Theorem 2.5].

Theorem 1 *If P_n is the n-fold product of uniform measure on $\mathbb{S}^1$ and $\mu_{n,\sigma}$ is the random measure as defined above. For $\Gamma \subset M_1(\mathbb{S}^1)$, the $\mu_{n,\sigma}$ satisfy an LDP with rate function*

$$I_\beta(\nu) := H(\nu \mid \mu) - \frac{\beta}{2} \left| \int_{\mathbb{S}^1} x d\nu(x) \right|^2 - \varphi(\beta), \tag{3}$$

where the free energy is given by

$$\varphi(\beta) = \inf_{\nu \in M_1(\mathbb{S}^1)} \left[H(\nu \mid \mu) - \frac{\beta}{2} \left| \int_{\mathbb{S}^1} x d\nu(x) \right|^2 \right]. \tag{4}$$

For fixed $\beta \geq 0$, every subsequence of $P_{n,\beta}[\mu_{n,\sigma} \in \cdot]$ converges weakly to a probability measure on $M_1(\mathbb{S}^1)$ concentrated on the canonical macrostates $\mathcal{E}_\beta := \{\nu : I_\beta(\nu) = 0\}$, i.e., the zeros of the rate function.

For $\beta = 0$, the relative entropy $H(\cdot \mid \mu)$ achieves its minima of 0 only for the uniform measure μ, implying that the canonical macrostate is disordered. For $\beta > 0$, canonical macrostates are defined abstractly through zeros of the rate function (3), and later Theorem 5 will describe the macrostates explicitly.

The free energy given by (4) can be transformed into the following more explicit form.

Theorem 2 (Kirkpatrick-Nawaz [13]) *The free energy φ has the formula:*

$$\varphi(\beta) = \begin{cases} 0, & \text{if } \beta < 2, \\ \Phi_\beta(g^{-1}(\beta)), & \text{if } \beta \geq 2, \end{cases}$$

where I_i below are modified Bessel functions of first kind and Φ_β is the functional defined by:

$$\Phi_\beta(r) := r\frac{I_1(r)}{I_0(r)} - \log[I_0(r)] - \frac{\beta}{2}\left(\frac{I_1(r)}{I_0(r)}\right)^2, \tag{5}$$

and

$$g(r) := r\frac{I_0(r)}{I_1(r)}.$$

Here the phase transition is continuous as the function φ and its derivative φ' are continuous at the critical threshold $\beta = \beta_c = 2$.

The magnetization for the classical XY model can be obtained by differentiating the partition function:

$$|m| = \left|\mathbb{E}\left[\frac{1}{n}\sum_i \sigma_i\right]\right| = \left|\mathbb{E}\left[\frac{1}{n}S_n\right]\right| = \frac{I_1(r)}{I_0(r)}$$

From the free energy we can precisely explain the phase transition at $\beta = 2$. For $0 \leq \beta \leq 2$, we have a unique global minimum for the free energy at the origin with a zero magnetization. For $\beta \geq 2$, we have a unique global minimum for a positive radius.

Let $\{\sigma_i\}_{i=1}^n$ be i.i.d. uniform random points on $\mathbb{S}^1 \subseteq \mathbb{R}^2$. We have the following Cramér-type LDP for the average spin.

Theorem 3 (Kirkpatrick-Nawaz [13]) *Let $P_{n,\beta}$ be the Gibbs measure defined above (1). Then for $\beta \geq 0$, the average spin $M_n = M_n(\sigma) := \frac{1}{n}\sum_{i=1}^n \sigma_i$ satisfies an LDP with rate function $I_\beta(x) = \Phi_\beta(r)$:*

$$P_{n,\beta}\left(M_n \simeq x\right) \simeq e^{-n\Phi_\beta(r)},$$

where Φ_β is given by (5) and $r = |x|$.

For an explicit representation of $\mathcal{E}_\beta$, we note from (2) that the relative entropy depends only on the distribution of f. By Fubini's theorem

$$\int f\log(f)d\mu = \int_0^\infty \mu\big[f\log(f) > t\big]dt - \int_0^\infty \mu\big[f\log(f) < -t\big]dt.$$

This implies that for a fixed f, the quantity $\left|\int x d\nu(x)\right|$ is maximized for corresponding densities which are symmetric about a fixed pole and decreasing as the distance from the pole increases. Using this reasoning, consider a density f that is symmetric about the north pole and decreasing away from the pole i.e., ν_f is the measure with density $f(x, y) = f(y)$ which is increasing in y. Then

$$\begin{aligned}
H(\nu_f \mid \mu) &= \frac{1}{2\pi}\int_{\mathbb{S}^1} f(x,y)\log[f(x,y)]dxdy \\
&= \frac{1}{\pi}\int_0^{2\pi}\int_0^{\pi} f(\cos(\theta))\log[f(\cos(\theta))]d\theta d\varphi \\
&= \frac{1}{\pi}\int_{-1}^{1}\frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy.
\end{aligned}$$

Similarly,

$$\int_{\mathbb{S}^1} x d\nu_f(x) = \frac{1}{\pi}\begin{bmatrix}0\\1\end{bmatrix}\int_{-1}^{1}\frac{yf(y)}{\sqrt{1-y^2}}dy.$$

Therefore, our minimization problem is reduced to minimizing the following functional

$$\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy - \frac{\beta}{2}\left(\frac{1}{\pi}\int_{-1}^{1}\frac{yf(y)}{\sqrt{1-y^2}}dy\right)^2$$

over $f : [-1, 1] \to \mathbb{R}_+$ such that f is increasing and $\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)}{\sqrt{1-y^2}}dy = 1$. We can rewrite the first term of the last expression to see that it involves the usual entropy $S(f) = \int f \log(f)$:

$$\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy = \frac{1}{\pi}\int_{-1}^{1}\frac{f(y)}{\sqrt{1-y^2}}\log\left[\frac{f(y)}{\pi}\right]dy + \log(\pi) = -S\left(\frac{f}{\pi}\right) + \log(\pi).$$

Now for $\left|\int x d\nu(x)\right| = c \in [0, 1]$, using constrained entropy maximization (see Theorem 12.1.1 from [6]), we will minimize $\frac{1}{\pi}\int_{-1}^{1}\frac{yf(y)}{\sqrt{1-y^2}}dy$, that is, maximize $S(f/\pi)$, over the $\nu \in M_1(\mathbb{S}^1)$ corresponding to this c.

Proposition 4 (Kirkpatrick-Nawaz [13]) *Consider a set of functions* $f : [-1, 1] \to \mathbb{R}_+$, *with weight function* $w(y) = \frac{1}{\sqrt{1-y^2}}$, *such that* $\int_{-1}^{1} f(y)w(y)dy = 1$, *and* $\left|\int_{-1}^{1} yf(y)w(y)dy\right| = c$. *i.e., weighted integral of* f *is 1 while first weighted moment is bounded. Then the exponential function* $f^*(y) = \pi a e^{by}$ *uniquely maximizes* $S(f/\pi)$ *over the densities satisfying these conditions.*

For $c \in [0,1]$, observe that f^* increasing gives all $b \in [0,\infty)$. Now for $b \in [0,\infty)$, our functional minimization reduces to the following one dimensional function:

$$\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy - \frac{\beta}{2}c^2 = b\frac{I_1(b)}{I_0(b)} - \log\left[I_0(b)\right] - \frac{\beta}{2}\left(\frac{I_1(b)}{I_0(b)}\right)^2$$
$$=: \Phi_\beta(b). \tag{6}$$

The following theorem, a special case proved using the calculus of variations in [13], describes the canonical macrostates:

Theorem 5 (Kirkpatrick-Nawaz [13])

(a) For $\beta \le 2$, $\inf_{b\ge 0}\left\{\Phi_\beta(b)\right\} = 0$ is achieved for $b = 0$ and the corresponding $a = 1$, so that the minimizing function $f^* = 1$ and therefore the only canonical macrostate is the uniform distribution μ.

(b) For $\beta > 2$, $\inf_{b\ge 0}\left\{\Phi_\beta(b)\right\} = \Phi_\beta(g^{-1}(\beta))$, where $b = g^{-1}(\beta)$ is the unique strictly positive solution to $g(b) = \beta$ where

$$g(b) = b\frac{I_0(b)}{I_1(b)},$$

$a = \frac{1}{\pi I_0(b)}$ and $\lim_{\beta\downarrow 2}\inf_{b\ge 0}\left\{\Phi_\beta(b)\right\} = 0$. In this case, the canonical macrostates are given by $\mathcal{E}_\beta = \{\nu_{f,x}\}_{x\in\mathbb{S}^1}$, where $\nu_{f,x}$ is the measure that is the rotation of ν_f from north pole to x-direction, which is symmetric about the north pole with density $f : [-1,1] \to \mathbb{R}$ given by $f(y) = \pi a e^{by}$ with a and b as above.

We can also visualize the Gibbs measure corresponding to subcritical or supercritical cases as shown in Fig. 2.

2.2 Limit Theorems for the Total Spin in the XY Model

Next we understand the asymptotics for the total spin of the mean-field XY model, in different regimes across the phase transition, describing the central and non-central limit theorems for each phase.

In the high temperature regime ($0 \le \beta < 2$), the average spin (magnetization) of the system goes to zero with increasing number of spins $n \to \infty$, and we have a multivariate central limit theorem with a rate of convergence in Theorem 6. The main idea is to use Stein's method [12, 15, 19] with the exchangeable pair (W_n, W_n') from the Gibbs sampling approach: our random variable representing the rescaled total spin of the original configuration is

$$W_n := \sqrt{\frac{2-\beta}{n}}\sum_{i=1}^{n}\sigma_i,$$

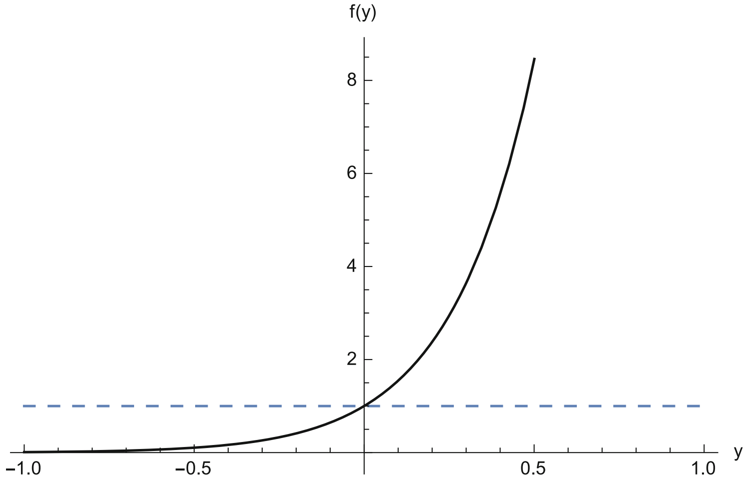

Fig. 2 Cross-sections of two canonical macrostates: For $\beta \le 2$ (the disordered regime), we have the uniform distribution $f(y) = 1$ as the *dotted line*; for $\beta = 5 > \beta_c = 2$ (the ordered regime), we have plotted the cross-section of the distribution ν_f, given by $f(x, y) = f(y) = \frac{e^{by}}{I_0(y)}$, showing that the spins point predominantly near the north-pole direction

while the random variable representing the rescaled total spin of the new configuration, with $I \in \{1, \ldots, n\}$ chosen uniformly at random, is

$$W_n' := W_n(\sigma') = W_n - \sqrt{\frac{2-\beta}{n}}\sigma_I + \sqrt{\frac{2-\beta}{n}}\sigma_I'.$$

Theorem 6 (Kirkpatrick-Nawaz [13]) *In the high temperature regime* $0 < \beta < 2$*, if* W_n *is defined as above,* Z *is a standard normal random variable in* $\mathbb{R}^2$*,* c_β *is a function depending on* β *only,* $L(g)$ *is the modulus of uniform continuity of* g*, and* $M(g)$ *is the maximum operator norm of the Hessian of* g*, then we have:*

$$\sup_{g:L(g),M(g)\le 1} |\mathbb{E}g(W_n) - \mathbb{E}g(Z)| \le \frac{c_\beta}{\sqrt{n}}$$

The proof of Theorem 6 proceeds in several steps, as a special case of [13]: first we use the fact that the density of the Gibbs measure is rotationally invariant to conclude that each spin has a uniform marginal distribution. We obtain the complete asymptotic behavior of the total spin using the rotational invariance of the total spin, a strategy adapted from [12]. We calculate the variance of the total spin to arrive at the proper scaling for defining the exchangeable pair and use the pair to derive expressions and bounds for the linear factor Λ appearing in the conditional expectation and the remainder terms R and R' [12, 13, 15]. The rest follows from a theorem of Meckes [15].

As the temperature decreases to zero, the spins start aligning. For smaller values of $\beta > 2$, the spins vectors are aligned weakly, while for larger β, this alignment is strong. For any $\beta > 2$, because of the large deviation principle in Theorem 3, we have that $|\sum \sigma_j|$ is close to bn/β with high probability, if b is the minimizer in Φ_β. And due to the circular symmetry, all points on the circle of radius bn/β are equally likely. With this reasoning, similar to [12], it is natural to consider the random variable representing the fluctuations of squared-length of total spin, i.e.,

$$W_n := \sqrt{n}\left[\frac{\beta^2}{n^2 b^2}\left|\sum_{j=1}^{n}\sigma_j\right|^2 - 1\right]. \tag{7}$$

Our multivariate central limit theorem in the low temperature (ordered) regime is as follows:

Theorem 7 (Kirkpatrick-Nawaz [13]) *If $\beta > 2$ and b is the solution of $b = \beta f(b) := \beta\frac{I_1(b)}{I_0(b)}$, and W_n is as defined above in (7), and if Z is a centered normal random variable with variance V, where*

$$V = \frac{4\beta^2}{(1-\beta f'(b))\,b^2}\left[1 - \frac{1}{b}\frac{I_1(b)}{I_0(b)} - \left(\frac{I_1(b)}{I_0(b)}\right)^2\right],$$

then there exists c_β, depending only on β, such that then

$$d_{BL}(W_n, Z) \leq c_\beta\left(\frac{\log(n)}{n}\right)^{1/4}.$$

where $d_{BL}(X,Y)$ is the bounded Lipschitz distance between random variables X and Y.

Again the proof of Theorem 7 follows from a univariate analogue of the abstract normal approximation of Stein [19], and relies on conditional moment bounds. The fact that the variance is positive was proved by Amos [2] while deriving the improved bounds on the ratio of Bessel functions.

At the critical temperature $\beta_c = 2$, we will consider the random variable

$$W_n := \frac{c}{n^{3/2}}\sum_{i,j=1}^{n}\langle\sigma_i, \sigma_j\rangle, \tag{8}$$

and make an exchangeable pair (W_n, W_n') using Glauber dynamics. Using symmetry of the total spin and Stein's method similar to [5, 12], we will obtain critical limiting density function p as defined below.

Theorem 8 (Kirkpatrick-Nawaz [13]) *For the critical inverse temperature $\beta = 2$, if W_n is as defined above in (8), and X is the random variable with the density*

$$p(t) = \begin{cases} \frac{1}{Z} e^{-t^2/64} & t \geq 0, \\ 0 & t < 0, \end{cases}$$

where Z is normalizing constant, then there exists a universal constant C such that

$$\sup_{\substack{\|h\|_\infty \leq 1, \|h'\|_\infty \leq 1 \\ \|h''\|_\infty \leq 1}} \left| \mathbb{E}h(W_n) - \mathbb{E}h(X) \right| \leq \frac{C \log(n)}{\sqrt{n}}.$$

The proof of the limit theorem for the critical temperature is essentially via the "density approach" to Stein's method introduced by Stein, Diaconis, Holmes, and Reinert [20]. Recently, also Chatterjee and Shao [5] have applied this approach to the total spin of the mean-field Ising model, i.e., the Curie-Weiss model.

We note that these limit theorems with explicit rates of convergence can be generalized to high-dimensional spins, but we will omit those technicalities in the following section.

3 High-Dimensional Spin $O(N)$ Models

We can use similar methods to extend our results for two-dimensional spin classical XY model to classical $O(N)$ models, or N-vector models. In this general case, with spins in $\mathbb{S}^{N-1} \subset \mathbb{R}^N$, the critical inverse temperature is $\beta_c = N$ [1, 13]. The N-vector models on a complete graph K_n have the Hamiltonian:

$$H_n(\sigma) := -\frac{1}{2n} \sum_{i,j=1}^{n} \left\langle \sigma_i, \sigma_j \right\rangle. \tag{9}$$

We present results about the magnetization, free energy, and critical behavior in the $O(N)$ models. It is important to note that we divide our asymptotic analysis into two cases: if N an even positive integer, we have modified Bessel functions of first kind with order $\nu = N/2$ and $\nu - 1$, while for N odd, we have hyperbolic functions arising from the half-integer order Bessel functions.

3.1 The Magnetization in $O(N)$ Models

Similar to the classical XY model, we can calculate the magnetization of the classical N-vector unit hyperspherical model using the conditional density, from the conditional expectations, and it turns out to be a ratio of modified Bessel function of first kind:

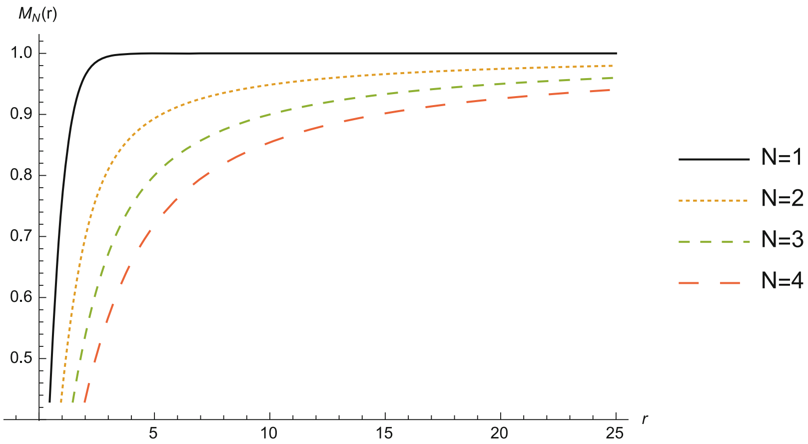

Fig. 3 Graph of magnetization limits $|M_N|$ for N-vector models, $1 \le N \le 4$. For the mean-field Ising model, $M_1 = tanh(x)$, for the mean-field XY model $|M_2| = \frac{I_1(r)}{I_0(r)}$, for the mean-field Heisenberg model $|M_3| = coth(r) - \frac{1}{r}$, and for the mean-field Toy model of the Higgs sector, $|M_4| = \frac{I_2(r)}{I_1(r)}$. Here r and β are related by the formula $g_N(r) := r\frac{I_{\frac{N}{2}-1}(r)}{I_{\frac{N}{2}}(r)} = \beta$

Theorem 9 (Kirkpatrick-Nawaz [13]) *Consider the $O(N)$ model with the above Hamiltonian (9), with N representing the dimension of the spin $\sigma_i \in \mathbb{S}^{N-1}$. Then on the complete graph K_n the $O(N)$ magnetization $M_{N,n} = \sum_{i=1}^{n} \sigma_i$ has the following mean-field limit:*

$$|M_N| = \frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}$$

From Fig. 3, we can observe that low-dimensional spin models can be magnetized easier in some sense, and as the spin gets higher dimensional, it takes more energy to magnetize the physical system.

3.2 *The Rate Function and Free Energy in $O(N)$ Models*

Next we will present rate functions for large deviation principles similar to Theorems 1 and 3, the first of which is the relative entropy for the N-vector model given by an abstract formula similar to before:

$$I_{\beta,N}(\nu) := H(\nu \mid \mu) - \frac{\beta}{2}\left|\int_{\mathbb{S}^{N-1}} x d\nu(x)\right|^2 - \varphi(\beta)$$

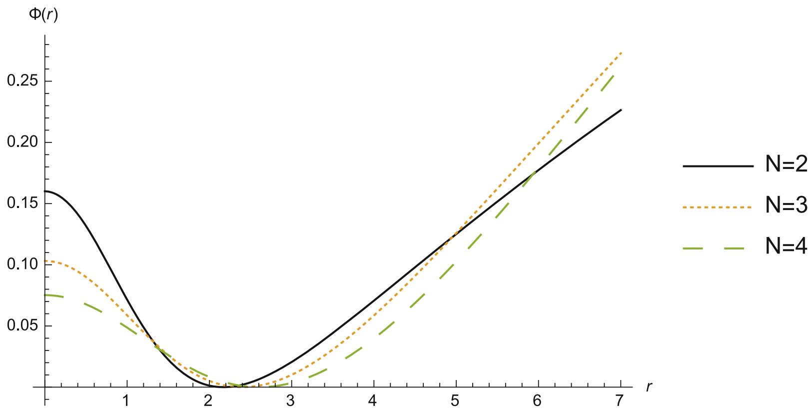

Fig. 4 Graph of the rate function $I_{\beta,N}(x) = \Phi_{\beta,N}(r)$ in the supercritical regime ($\beta = N + 1$) for $2 \le N \le 4$, which has minimum at radius $g_N^{-1}(\beta) = r$

where $H(\nu \mid \mu)$ is the relative entropy (2) and φ_N is the free energy defined abstractly as before:

$$\varphi_N(\beta) = \inf_{\nu \in M_1(\mathbb{S}^{N-1})} \left[H(\nu \mid \mu) - \frac{\beta}{2} \left| \int_{\mathbb{S}^{N-1}} x d\nu(x) \right|^2 \right]. \tag{10}$$

We can calculate the minima in the expression of this rate function and verify that in the subcritical regime ($\beta < N$) there is a unique minimum, while in the supercritical regime there is a family of minima parametrized by $\mathbb{S}^{N-1}$. The free energy given by (10) can be written in the following more explicit form using a method like the one in the previous section. In particular, we have a Cramér-type LDP for the average spin $M_n := \frac{1}{n} \sum_{i=1}^{n} \sigma_i \in \mathbb{R}^N$, with rate function $I_{\beta,N}(x) = \Phi_{\beta,N}(r)$, defined below for $\beta \ge 0$ and $r = |x|$ (Fig. 4).

Theorem 10 (Kirkpatrick-Nawaz [13]) *For dimension N, the free energy φ has the formula:*

$$\varphi_N(\beta) = \begin{cases} 0, & \text{if } \beta < N, \\ \Phi_{\beta,N}(g^{-1}(\beta)), & \text{if } \beta \ge N, \end{cases}$$

where $g^{-1}(\beta) = r$ with

$$g(r) = g_N(r) := r \frac{I_{\frac{N}{2}-1}(r)}{I_{\frac{N}{2}}(r)},$$

and

$$\Phi_{\beta,N}(r) = r\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)} + \log\left[\frac{A_N}{A_{N-1}}\frac{r^{\frac{N}{2}-1}}{B_N\pi I_{\frac{N}{2}-1}(r)}\right] - \frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}\right)^2,$$

with

$$A_N := \frac{2\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)}$$

and

$$B_N = \begin{cases} \prod_{k=0}^{\frac{N}{2}-1}|2k-1|, & \text{if } N \text{ even},\\ \frac{2^{\frac{N}{2}-1}\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}}, & \text{if } N \text{ odd}. \end{cases}$$

In particular, φ and φ' are continuous at the critical threshold $\beta = N$, implying that the phase transition is second-order or continuous.

3.3 *The Critical Density Function in $O(N)$ Models*

The limiting density for the critical case uses the (hyper-)spherical symmetry of the total spin for $O(N)$ models, giving the following non-normal limit theorem (Fig. 5).

Theorem 11 (Kirkpatrick-Nawaz [13]) *At the critical temperature $\beta = N$, the random variable $W_n = \frac{c_N|S_n|^2}{n^{3/2}}$ has as its limit as $n \to \infty$ the random variable X with density*

$$p_N(t) = \begin{cases} \frac{1}{Z}t^{\frac{N-2}{2}}e^{-kt^2}, & \text{if } t \geq 0,\\ 0, & \text{if } t < 0, \end{cases}$$

where $k = \frac{1}{N^2(4N+8)}$ and Z is the normalizing constant. To be precise about the rate of convergence, there exists a universal constant C such that

$$\sup_{\substack{\|h\|_\infty \leq 1,\ \|h'\|_\infty \leq 1\\ \|h''\|_\infty \leq 1}} \left|\mathbb{E}h(W_n) - \mathbb{E}h(X)\right| \leq \frac{C\log(n)}{\sqrt{n}}.$$

The proof of this theorem is in [13] and includes methods from [8, 9, 12].

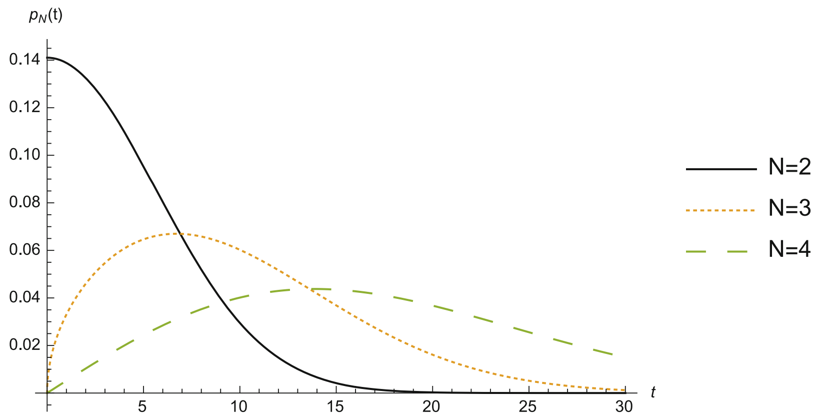

Fig. 5 Mean-field critical density functions p_N for $2 \leq N \leq 4$ and $t \geq 0$. For the XY model $p_2(t) = \frac{e^{-t^2/64}}{4\sqrt{\pi}}$, for the Heisenberg model $p_3(t) = \frac{\sqrt{t}e^{-t^2/180}}{5^{3/4}\sqrt{54}\Gamma[3/4]}$, and for the Toy model of the Higgs sector, $p_4(t) = \frac{te^{-t^2/384}}{192}$

Acknowledgements Both authors partially supported by NSF CAREER award DMS-1254791, and NSF grant 0932078 000 while in residence at MSRI during Fall 2015.

References

1. M. Aizenman, B. Simon, A comparison of plane rotor and Ising models. Phys. Lett. A **76** (1980). doi:10.1016/0375-9601(80)90493-4
2. D.E. Amos, Computation of modified Bessel functions and their ratios. Math. Comput. **28**(125), 239–251 (1974)
3. P.W. Anderson, Random-phase approximation in the theory of superconductivity. Phys. Rev. **112**(6), 1900–1915 (1958)
4. S.G. Brush, History of the Lenz-Ising Model. Rev. Mod. Phys. **39**, 883–893 (1967)
5. S. Chatterjee, Q.-M. Shao, Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. Ann. Appl. Probab. **21**(2), 464–483 (2011)
6. T.M. Cover, J.A. Thomas, *Elements of Information Theory*, 2nd edn. (Wiley-Interscience [Wiley], Hoboken, 2006)
7. R.S. Ellis, C.M. Newman, Limit theorems for sums of dependent random variables occurring in statistical mechanics. Z. Wahrsch. Verw. Gebiete **44**(2), 117–139 (1978)
8. R.S. Ellis, C.M. Newman, The statistics of Curie-Weiss models. J. Stat. Phys. **19**(2), 149–161 (1978)
9. R.S. Ellis, C.M. Newman, J.S. Rosen, Limit theorems for sums of dependent random variables occurring in statistical mechanics. II. Conditioning, multiple phases, and metastability. Z. Wahrsch. Verw. Gebiete **51**(2), 153 (1980)
10. R.S. Ellis, K. Haven, B. Turkington, Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. J. Stat. Phys. **101**(5–6), 999–1064 (2000)
11. E. Ising, Contribution to the theory of ferromagnetism. Z. Phys. **31**, 253–258 (1925)

12. K. Kirkpatrick, E. Meckes, Asymptotics of the mean-field Heisenberg model. J. Stat. Phys. **152**(1), 54–92 (2013)
13. K. Kirkpatrick, T. Nawaz, Asymptotics of mean-field O(N) models. J. Stat. Phys. **165**(6), 1114–1140 (2016) **165**(6), 1114–1140 (2016)
14. J.M. Kosterlitz, D.J. Thouless, Ordering, metastability and phase transitions in two-352 dimensional systems. J. Phys. C Solid State Phys. **6**, 1181–1203 (1973)
15. E. Meckes, On Stein's method for multivariate normal approximation, in *High Dimensional Probability V: The Luminy Volume* (Institute of Mathematical Statistics, Beachwood, 2009)
16. N.D. Mermin, H. Wagner, Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models. Phys. Rev. Lett. **17**, 1307 (1966)
17. M.A. Moore, Additional evidence for a phase transition in the plane-rotator and classical Heisenberg models for two-dimensional lattices. Phys. Rev. Lett. **23**, 861–863 (1969)
18. H.E. Stanley, Dependence of critical properties on dimensionality of spins. Phys. Rev. Lett. **20**, 589–592 (1968)
19. C. Stein, *Approximate Computation of Expectations*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 7 (Institute of Mathematical Statistics, Hayward, 1986)
20. C. Stein, P. Diaconis, S. Holmes, G. Reinert, Use of exchangeable pairs in the analysis of simulations, in *Stein's Method: Expository Lectures and Applications*. IMS Lecture Notes-Monograph Series, vol. 46 (Institute of Mathematical Statistics, Beachwood, 2004), pp. 1–26