Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn

Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

M. Griebel, F.Y. Kuo and I.H. Sloan

The ANOVA decomposition of a non-smooth function of infinitely many variables can have every term smooth

INS Preprint No. 1403

2014

MATHEMATICS OF COMPUTATION
Volume 00, Number 0, Pages 000–000
S 0025-5718(XX)0000-0

THE ANOVA DECOMPOSITION OF A NON-SMOOTH FUNCTION OF INFINITELY MANY VARIABLES CAN HAVE EVERY TERM SMOOTH

MICHAEL GRIEBEL, FRANCES Y. KUO, AND IAN H. SLOAN

Abstract. The pricing problem for a continuous path-dependent option results in a path integral which can be recast into an infinite-dimensional integration problem. We study ANOVA decomposition of a function of infinitely many variables arising from the Brownian bridge formulation of the continuous option pricing problem. We show that all resulting ANOVA terms can be smooth in this infinite-dimensional case, despite the non-smoothness of the underlying payoff function. This result may explain why quasi-Monte Carlo methods or sparse grid quadrature techniques work for such option pricing problems.

1. Introduction

In this paper we are concerned with the smoothness of the terms in the ANOVA decomposition of the integrand for a continuous (and hence infinite-dimensional) problem motivated by option pricing. Option pricing problems have presented a particular challenge to modern cubature methods for high-dimensional integration (we think especially of quasi-Monte Carlo [10] and Smolyak or sparse grid [5] methods) because of the simple fact that Asian call options, for example, are considered to be worthless if the average asset value falls below a predetermined strike price K. For this reason the integrand in the expected value of the payoff then takes the form

$$\max\left(\text{average value} - K, 0\right).$$

The max function introduces a kink into the integrand, and precludes it from belonging to any of the mixed derivative function spaces for which the cubature theories are built. Yet empirical evidence suggests that the max function is not a barrier to the practical success of the methods.

This paper attempts to throw light on the apparent success of high-dimensional cubature methods in the particular case of the Brownian bridge (or more precisely, Lévy-Ciesielski) construction, by showing, in the infinite-dimensional setting, that while the integrand might not be smooth, *every term in its ANOVA decomposition is smooth.* This expansion is in the classical sense where the terms are obtained by integration rather than by "anchoring". ANOVA decomposition in other infinite-dimensional settings have previously been considered in [2, 9, 15].

The paper builds on our previous results [12, 13] for the discrete-time option pricing problem. In the second of those papers we showed that all terms in the

Received by the editor April 2016.
2010 *Mathematics Subject Classification.* Primary 41A63, 41A99. Secondary 65D30.

(finite) ANOVA decomposition of the integrand for the option pricing problem, for both the Brownian bridge and standard constructions (see below) are smooth except for the very last one, the one that depends simultaneously on all the integration variables.[1]

That result led us to speculate that for the corresponding *infinite-dimensional* problem there are *no* non-smooth ANOVA terms. We shall show in this paper that the speculation is, in a certain precise sense, true. The smoothness result for the ANOVA decomposition of a function of infinitely many variables in this paper opens a possible path to the error analysis of cubature methods for the finite-dimensional discrete-time option pricing problem. The idea is that the analysis might be applied not to the discrete problem itself, but rather to a truncated sum of the ANOVA decomposition of the option pricing integrand of infinitely many variables. This truncated sum is smooth, unlike the original integrand; yet we prove in this paper, see Corollary 5.5, that the expected value of the difference converges to zero as the dimensionality goes to infinity, or equivalently, as the time interval of the discretization goes to zero. The next step, not covered here, would be to quantify the rate of convergence to zero.

The outline of the paper is as follows. In Section 2 we introduce the option pricing problem, in both its continuous and discrete versions. In Section 3 we review the ANOVA decomposition in the finite-dimensional setting and summarize the results from [13]. In Section 4 we establish the theory for the integration problem over the sequence space. In Section 5 we develop an ANOVA decomposition for a function of infinitely many variables. In Section 6 we prove that all ANOVA terms are smooth under conditions that apply in particular to our Brownian bridge formulation of the finance problem.

2. The option pricing problem

In this paper we consider a continuous version of a path-dependent call option with strike price K in a Black-Scholes model with risk-free interest rate $r > 0$ and constant volatility $\sigma > 0$. The price $S(t)$ at time t then satisfies the stochastic differential equation

$$\mathrm{d}S(t) \;=\; S(t)\,(r\,\mathrm{d}t + \sigma\,\mathrm{d}B(t))\,, \tag{2.1}$$

where $B(t) = B(t,\omega)$ denotes standard Brownian motion on some probability space $(\Omega, \mathcal{F}, P)$, that is, for each $t \in [0,1]$, $B(t,\cdot)$ is a zero-mean Gaussian random variable, and for each pair $t,s \in [0,1]$ the covariance of $B(t,\cdot)$ and $B(s,\cdot)$ is

$$\mathbb{E}[B(t,\cdot)B(s,\cdot)] \;=\; \min(t,s). \tag{2.2}$$

The solution to (2.1) is given explicitly by

$$S(t) \;=\; S(0)\exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right), \quad 0 \le t \le T.$$

[1]Note that in [13] the main result, Theorem 3.1, though correct, does not as claimed apply to the Asian option pricing problem because assumption (3.3) in the theorem is not satisfied by the Asian option pricing problem. A strengthened theorem is given in [14], without this assumption, as a result of which the conclusions of [13] stand.

The discounted payoff (taking for definiteness the case of a continuous arithmetic Asian option with terminal time T) is

$$\mathcal{P} := e^{-rT} \max\left(\frac{1}{T}\int_0^T S(t)\,\mathrm{d}t - K, 0\right). \tag{2.3}$$

The pricing problem is then to compute the expected value $\mathbb{E}(\mathcal{P})$, which in this setting is an (infinite-dimensional) path integral, see Section 4.

It is instructive to consider first the finite-dimensional problem, in which the time interval T is divided into intervals of size T/d, and the integral in (2.3) is replaced by a finite sum,

$$\mathcal{P}_d := e^{-rT} \max\left(\frac{1}{d}\sum_{i=1}^d S(t_i) - K, 0\right),$$

where $t_i = iT/d$. The expected value of the discounted payoff $\mathcal{P}_d$ is then a d–dimensional Gaussian integral

$$\begin{aligned}\mathbb{E}(\mathcal{P}_d) \;=\; & e^{-rT}\int_{\mathbb{R}^d} \max\left(\frac{S(0)}{d}\sum_{i=1}^d \exp\left(\left(r-\frac{\sigma^2}{2}\right)t_i + \sigma y_i\right) - K, 0\right) \\ & \times \frac{\exp(-\frac{1}{2}\boldsymbol{y}^{\mathtt{T}}\Sigma_d^{-1}\boldsymbol{y})}{\sqrt{(2\pi)^d\det(\Sigma_d)}}\,\mathrm{d}\boldsymbol{y},\end{aligned} \tag{2.4}$$

where $\Sigma_d \in \mathbb{R}^d \times \mathbb{R}^d$ is the covariance matrix for the Brownian motion, with the (i,j)-th entry given by $\min(t_i, t_j)$. Using a factorization

$$\Sigma_d = A_d A_d^{\mathtt{T}},$$

followed by the substitution $\boldsymbol{y} = A_d\boldsymbol{x}$, the integral (2.4) then reduces to

$$\begin{aligned}\mathbb{E}(\mathcal{P}_d) \;=\; & e^{-rT}\int_{\mathbb{R}^d} \max\left(\frac{S(0)}{d}\sum_{i=1}^d \exp\left(\left(r-\frac{\sigma^2}{2}\right)t_i + \sigma\sum_{j=1}^d A_{d,i,j}x_j\right) - K, 0\right) \\ & \times \frac{\exp(-\frac{1}{2}\boldsymbol{x}^{\mathtt{T}}\boldsymbol{x})}{\sqrt{(2\pi)^d}}\,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{2.5}$$

As is well known, see for example [11], different factorizations of Σ_d correspond to different methods of construction of the Brownian motion, the most common trio being the *standard construction*, the *Brownian bridge construction* [6], and the *principal components construction* [1]. (There are other methods including the so-called *linear transformation* [16].) The standard construction of the Brownian motion corresponds to the Cholesky factorisation, with A_d lower-triangular, and all non-zero values equal to $\sqrt{T/d}$. The principal components construction has

$$A_d = [\sqrt{\lambda_1}\,\boldsymbol{\xi}_1; \cdots; \sqrt{\lambda_d}\,\boldsymbol{\xi}_d],$$

where $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of Σ_d in non-increasing order and $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_d$ are the corresponding column eigenvectors, normalised to be of unit length in the ℓ_2 sense.

Assuming that $d = 2^N$, the Brownian bridge construction generates the Brownian motion $\boldsymbol{y} = (b(\frac{T}{d}), b(\frac{2T}{d}), \ldots, b(T))^{\mathtt{T}}$ in the order of $T, T/2, T/4, 3T/4, \ldots$ as follows

$$\begin{aligned} b(0) &= 0, \\ b(T) &= \sqrt{T}\,x_1, \\ b(\tfrac{T}{2}) &= \tfrac{1}{2}(b(0)+b(T)) + \sqrt{\tfrac{T}{4}}\,x_2, \\ b(\tfrac{T}{4}) &= \tfrac{1}{2}(b(0)+b(\tfrac{T}{2})) + \sqrt{\tfrac{T}{8}}\,x_3, \\ b(\tfrac{3T}{4}) &= \tfrac{1}{2}(b(\tfrac{T}{2})+b(T)) + \sqrt{\tfrac{T}{8}}\,x_4, \\ &\vdots \\ b(\tfrac{(d-1)T}{d}) &= \tfrac{1}{2}(b(\tfrac{(d-2)T}{d})+b(T)) + \sqrt{\tfrac{T}{2d}}\,x_d. \end{aligned} \tag{2.6}$$

This yields a different matrix A_d in the substitution $A_d\boldsymbol{x} = \boldsymbol{y}$.

3. Smoothness of finite ANOVA decompositions

Let $\mathcal{L}_2(\mathbb{R}^d) = \mathcal{L}_{2,\rho}(\mathbb{R}^d)$ be the class of square-integrable functions $f : \mathbb{R}^d \to \mathbb{R}$ with respect to the Gaussian weight

$$\rho_d(\boldsymbol{x}) \;=\; \prod_{j=1}^{d} \rho(x_j), \qquad \rho(x) \;=\; \frac{\exp(-\frac{1}{2}x^2)}{\sqrt{2\pi}}. \tag{3.1}$$

For a function $f \in \mathcal{L}_2(\mathbb{R}^d)$, the ANOVA decomposition takes the form (see e.g., [18, 20, 21])

$$f \;=\; \sum_{\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}, \tag{3.2}$$

where the sum is over all subsets of $\{1:d\} := \{1,2,\ldots,d\}$, and where $f_{\mathfrak{u}}(\boldsymbol{x})$ depends only on $\boldsymbol{x}_{\mathfrak{u}} = (x_j)_{j\in\mathfrak{u}}$; we write for convenience

$$f_{\mathfrak{u}}(\boldsymbol{x}) \;=\; f_{\mathfrak{u}}(\boldsymbol{x}_{\mathfrak{u}}), \quad \mathfrak{u} \subseteq \{1:d\}.$$

Moreover, each component $f_{\mathfrak{u}}$ with $\mathfrak{u} \neq \emptyset$ has the property that

$$\int_{-\infty}^{\infty} f_{\mathfrak{u}}(\boldsymbol{x})\rho(x_j)\,\mathrm{d}x_j \;=\; 0 \quad \text{for} \quad j \in \mathfrak{u},$$

which with (3.2) implies

$$\int_{\mathbb{R}^d} f(\boldsymbol{x})\rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \;=\; f_{\emptyset},$$

and is also easily seen to imply the $\mathcal{L}_2(\mathbb{R}^d)$-orthogonality of the terms in (3.2),

$$\int_{\mathbb{R}^d} f_{\mathfrak{u}}(\boldsymbol{x}) f_{\mathfrak{v}}(\boldsymbol{x})\,\rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \;=\; 0 \quad \text{if} \quad \mathfrak{u} \neq \mathfrak{v}, \quad \mathfrak{u},\mathfrak{v} \subseteq \{1:d\}.$$

This in turn ensures that

$$\sigma^2(f) \;=\; \sum_{\mathfrak{u}\subseteq\{1:d\}} \sigma^2(f_{\mathfrak{u}}),$$

where

$$\sigma^2(f) \;:=\; \int_{\mathbb{R}^d} f^2(\boldsymbol{x})\,\rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \left(\int_{\mathbb{R}^d} f(\boldsymbol{x})\,\rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)^2,$$

and

$$\sigma^2(f_\emptyset) \;=\; 0 \qquad \text{and} \qquad \sigma^2(f_{\mathfrak{u}}) \;=\; \int_{\mathbb{R}^d} f_{\mathfrak{u}}^2(\boldsymbol{x})\,\rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \quad \text{for} \quad \mathfrak{u} \neq \emptyset.$$

The terms $f_{\mathfrak{u}}$ in the ANOVA decomposition are in a sense obtained by integrating out from f the variables x_j for $j \notin \mathfrak{u}$. More precisely, following [13] we define P_j to be the projection obtained by integrating out the jth variable,

$$(P_j f)(\boldsymbol{x}) = \int_{-\infty}^{\infty} f(x_1, \ldots, x_{j-1}, t_j, x_{j+1}, \ldots, x_d)\,\rho(t_j)\,\mathrm{d}t_j,$$

and then defining the operator product

$$P_{\mathfrak{u}} \;=\; \prod_{j \in \mathfrak{u}} P_j, \qquad \mathfrak{u} \subseteq \{1:d\}.$$

Then it is known that $f_{\mathfrak{u}}$ is given recursively by

$$f_{\mathfrak{u}} \;=\; P_{\{1:d\}\setminus\mathfrak{u}} f - \sum_{\mathfrak{v} \subsetneq \mathfrak{u}} f_{\mathfrak{v}}, \qquad f_\emptyset \;=\; P_{\{1:d\}} f,$$

and that (see [18]) $f_{\mathfrak{u}}$ is given explicitly by

$$f_{\mathfrak{u}} \;=\; \sum_{\mathfrak{v} \subseteq \mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} P_{\{1:d\}\setminus\mathfrak{v}} f. \tag{3.3}$$

It is the integrating out of the inactive variables (those with labels not in $\mathfrak{u}$) that has the smoothing effect. The essence of the matter is that for a function of the form $f(\boldsymbol{x}) = \max(\phi(\boldsymbol{x}), 0)$ with ϕ a smooth function on $\mathbb{R}^d$ and having suitable behavior at infinity, $P_j f$ is a smooth function provided the partial derivative $\partial\phi/\partial x_j$ does not change sign; see in particular [13, Theorem 3.1] and the more general [14, Theorem 1].

For the particular case of option pricing, this property is used to show, in [14], that for the standard construction and the Brownian bridge construction, *all terms in the ANOVA decomposition are smooth except the last term* $f_{\{1:d\}}$. That term $f_{\{1:d\}}$ does not benefit from any integration, thus this highest term retains all of the non-smoothness existing in f itself. That led us to speculate that in the infinite-dimensional case, where there is no highest term, perhaps all terms of the Brownian bridge construction are smooth.

Later in the paper we shall need, in addition to the finite ANOVA decomposition (3.2), the so-called "anchored" decomposition: with $\boldsymbol{c} \in \mathbb{R}^d$ an arbitrarily prescribed "anchor", the decomposition is (see [18])

$$f \;=\; \sum_{\mathfrak{u} \subseteq \{1:d\}} \tilde{f}_{\mathfrak{u}}, \tag{3.4}$$

where instead of (3.3) we have

$$\tilde{f}_{\mathfrak{u}}(\boldsymbol{x}) \;=\; \sum_{\mathfrak{v} \subseteq \mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} f(\boldsymbol{x}_{\mathfrak{v}}, \boldsymbol{c}_{\{1:d\}\setminus\mathfrak{v}}), \tag{3.5}$$

with $(\boldsymbol{x}_{\mathfrak{v}}, \boldsymbol{c}_{\{1:d\}\setminus\mathfrak{v}})$ being the d-vector whose jth component is x_j if $j \in \mathfrak{v}$ and is c_j if $j \in \{1:d\} \setminus \mathfrak{v}$. Together (3.4) and (3.5) give the identity

$$f(\boldsymbol{x}) \;=\; f(x_1, \ldots, x_d) \;=\; \sum_{\mathfrak{u} \subseteq \{1:d\}} \sum_{\mathfrak{v} \subseteq \mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} f(\boldsymbol{x}_{\mathfrak{v}}, \boldsymbol{c}_{\{1:d\}\setminus\mathfrak{v}}). \tag{3.6}$$

Unlike the ANOVA case, there is no smoothing property associated with the anchored decomposition, because while integration can have a smoothing effect, point evaluation cannot.

4. The continuous problem

In this section we deal with the continuous version of the problem in Section 2. This means that instead of the d-dimensional integral we now need to work with path integrals (see [25]), which are essentially integrals over Banach spaces. For simplicity, from now on we take $T = 1$ for the terminating time of the Brownian motion.

4.1. **A path integral formulation.** In this subsection our path integrals are integrals over the different Brownian paths, and the integration domain is the space $\mathcal{C}[0,1]$ of continuous functions on $[0,1]$, with the uniform norm $\|\cdot\|_\infty$, and equipped with the Wiener measure μ. More formally, we consider integration with respect to the measure space $(\mathcal{C}[0,1], \mathcal{A}, \mu)$ where $\mathcal{A}$ is the σ-algebra associated with $\mathcal{C}[0,1]$, and the Wiener measure μ (see [4, page 34]) is a Gaussian measure. The path integral is defined for real-valued functions g belonging to the class of measurable functions with respect to the Wiener measure. Denoting a particular realization of $B(\cdot,\omega)$ by $b \in \mathcal{C}[0,1]$, we write the path integral of g as (see [3, Section 15])

$$\int_{\mathcal{C}[0,1]} g(b)\,\mu(\mathrm{d}b). \tag{4.1}$$

The covariance (2.2) can then be written as

$$\mathbb{E}[B(t,\cdot)\,B(s,\cdot)] \;=\; \int_{\mathcal{C}[0,1]} b(t)b(s)\,\mu(\mathrm{d}b) \;=\; \min(t,s), \quad t,s \in [0,1].$$

In our finance applications we need path integrals for which g is non-negative and satisfies

$$g(b) \;\le\; c\exp(\alpha\|b\|_\infty) \qquad \text{for some} \quad c > 0 \quad \text{and} \quad \alpha > 0.$$

As a concrete example, we have from (2.3), with $T = 1$,

$$\begin{aligned} 0 \le g(b) \;&:= e^{-r}\max\left(S(0)\int_0^1 \exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma b(t)\right)\mathrm{d}t - K, 0\right) \\ &\le e^{-r}S(0)\exp\left(\left|r - \frac{\sigma^2}{2}\right| + \sigma\|b\|_\infty\right). \end{aligned} \tag{4.2}$$

Fernique's theorem [4, Theorem 2.8.5] states that there exists a $\beta > 0$ such that

$$\int_{\mathcal{C}[0,1]} \exp(\beta\,\|b\|_\infty^2)\,\mu(\mathrm{d}b) < \infty.$$

It follows immediately that a function g with merely exponential growth in $\|b\|_\infty$ is integrable.

4.2. **Lévy-Ciesielski construction of Brownian paths.** The Lévy-Ciesielski (or Brownian bridge) construction expresses the Brownian path $B(t,\omega)$ in terms of a Faber-Schauder basis $\{\eta_0, \eta_{n,i} : n \in \mathbb{N}, i = 1, \ldots, 2^{n-1}\}$ of continuous functions on $[0,1]$, where $\eta_0(t) := t$ and

$$\eta_{n,i}(t) := \begin{cases} 2^{(n-1)/2}\left(t - \dfrac{2i-2}{2^n}\right), & t \in \left[\dfrac{2i-2}{2^n}, \dfrac{2i-1}{2^n}\right], \\ 2^{(n-1)/2}\left(\dfrac{2i}{2^n} - t\right), & t \in \left[\dfrac{2i-1}{2^n}, \dfrac{2i}{2^n}\right], \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

For a proof that this is a basis in $\mathcal{C}[0,1]$, see [23, Theorem 2.1(iii)] or [24]. The Brownian path corresponding to the sample point $\omega \in \Omega$ is in this construction given by

$$B(t,\omega) = X_0(\omega)\,\eta_0(t) + \sum_{n=1}^{\infty} \sum_{i=1}^{2^{n-1}} X_{n,i}(\omega)\,\eta_{n,i}(t), \tag{4.4}$$

where X_0 and all the $X_{n,i}, i = 1, \ldots, 2^{n-1}, n \in \mathbb{N}$ are independent standard normal random variables.

The Lévy-Ciesielski construction has the important property that it converges almost surely to a continuous Brownian path. For the following classical theorem see the original works by [7, 19], and for example [22] for the more general exposition.

Theorem 4.1. *The Lévy-Ciesielski expansion* (4.4) *converges uniformly in t, almost surely, to a continuous function, and the limit is a Brownian motion.*

For future reference we sketch a proof, following the argument of [8]. For $N \in \mathbb{N}$ we define

$$B_N(t,\omega) := X_0(\omega)\,\eta_0(t) + \sum_{n=1}^{N} \sum_{i=1}^{2^{n-1}} X_{n,i}(\omega)\,\eta_{n,i}(t),$$

so that $B_N(t,\omega)$ for each ω is a piecewise-linear function. Note that $B_N(t,\omega)$ is equal to $B(t,\omega)$ at special values of t: we easily see that

$$B(0,\omega) = B_N(0,\omega) = 0, \quad B(1,\omega) = B_N(1,\omega) = X_0(\omega),$$

and with $t = (2\ell-1)/2^N$ we have

$$B\left(\frac{2\ell-1}{2^N}, \omega\right) = B_N\left(\frac{2\ell-1}{2^N}, \omega\right), \quad \ell = 1, \ldots, 2^{N-1},$$

because the terms in (4.4) with $n > N$ vanish. The successive values returned by the usual (discrete) Brownian bridge construction, see (2.6), are the values of $B(t,\omega)$ at the corresponding special t values $0, 1, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \ldots$. Between the values at which the series has already terminated, $B_N(t,\omega)$ is a piecewise-linear interpolant. It is clear that

$$\begin{aligned} |B_N(t,\omega)| &\le |X_0(\omega)| + \sum_{n=1}^{N} \left(\max_{1 \le i \le 2^{n-1}} |X_{n,i}(\omega)|\right) \left(\sum_{i=1}^{2^{n-1}} \eta_{n,i}(t)\right) \\ &\le |X_0(\omega)| + \sum_{n=1}^{N} \max_{1 \le i \le 2^{n-1}} |X_{n,i}(\omega)|\, 2^{-(n+1)/2}, \end{aligned} \tag{4.5}$$

where in the last step we used the fact that for a given $n \geq 1$ the disjoint nature of the Faber-Schauder functions ensures that at most one value of i contributes to the sum over i, and also that the $\eta_{n,i}$ have the same maximum value $2^{-(n+1)/2}$. Thus if we define

$$H_n(\omega) := \begin{cases} |X_0(\omega)| & \text{for } n = 0, \\ \max_{1 \leq i \leq 2^{n-1}} |X_{n,i}(\omega)| \, 2^{-(n+1)/2} & \text{for } n \geq 1, \end{cases} \tag{4.6}$$

then we have

$$|B_N(t,\omega)| \leq \sum_{n=0}^{N} H_n(\omega),$$

in which the right-hand side is independent of t. Now it is known that (see [8, Proof of Theorem 3]), as a consequence of (4.6) and the Borel-Cantelli lemma, one can construct a sequence $(\beta_n)_{n\geq 1}$ of positive numbers such that

$$\sum_{n=1}^{\infty} \beta_n < \infty, \quad \text{and} \quad P\left(H_n(\cdot) > \beta_n \text{ infinitely often}\right) = 0.$$

Letting Ω be the set of all sample points ω, we may define $\tilde{\Omega}$ to be the subset of the sample points ω for which $H_n(\omega) > \beta_n$ for only finitely many values of n. Then $\tilde{\Omega}$ is of full Gaussian measure, and for each $\omega \in \tilde{\Omega}$ there exists $N(\omega) \in \mathbb{N}$ such that

$$H_n(\omega) \leq \beta_n \quad \text{for} \quad n > N(\omega),$$

leading for $\omega \in \tilde{\Omega}$ to

$$|B_N(t,\omega)| \leq \sum_{n=1}^{\infty} H_n(\omega) \leq \sum_{n=1}^{N(\omega)} H_n(\omega) + \sum_{n=N(\omega)+1}^{\infty} \beta_n < \infty.$$

Thus almost surely the series $B_N(t,\omega)$ converges uniformly in t, and hence $B(t,\omega)$ is continuous in t.

That $B(t,\omega)$ has the correct covariance structure (2.2) follows (see [7, pp. 406–407]) from the fact that the value of each Faber-Schauder basis function at t is the integral from 0 to t of a Haar basis function, where the Haar basis forms a complete orthonormal set. The last step of the argument uses the Parseval identity.

4.3. **The path integral as an integral over a sequence space.** The results in the preceding subsection will allow us to express the path integral in (4.1) as an integral over a sequence space. We remark that from this point on we will find it convenient to use generally the language of measure and integration rather than of probability and expectation.

Recall that the Lévy-Ciesielski expansion (4.4) expresses the Brownian path $B(t,\omega)$ in terms of an infinite sequence $\boldsymbol{X}(\omega)$ of independent standard normal random variables X_0, $(X_{n,i})_{n\in\mathbb{N},i=1,\ldots,2^{n-1}}$. In the following we will denote a particular realization of this sequence $\boldsymbol{X}$ by

$$\boldsymbol{x} = \left(x_0, (x_{n,i})_{n\in\mathbb{N},i=1,\ldots,2^{n-1}}\right) = (x_1, x_2, \ldots) \in \mathbb{R}^\infty.$$

For convenience we will switch between the single-index labeling $(x_1, x_2, \ldots)$ and the double-index labeling $(x_0, x_{1,1}, x_{2,1}, x_{2,2}, \ldots)$ as appropriate. Our convention with indexing is that $x_1 \equiv x_0$, and $x_j \equiv x_{n,i}$ with $j = 2^{n-1} + i$ for $n \geq 1$ and $1 \leq i \leq 2^{n-1}$.

Motivated by the bound (4.5), we define a norm of this sequence $\boldsymbol{x}$ by

$$\|\boldsymbol{x}\|_{\mathcal{X}} := |x_0| + \sum_{n=1}^{\infty} \max_{1\le i\le 2^{n-1}} |x_{n,i}|\, 2^{-(n+1)/2}, \tag{4.7}$$

and we define a corresponding normed space by

$$\mathcal{X} := \{\boldsymbol{x} \in \mathbb{R}^{\infty} \;:\; \|\boldsymbol{x}\|_{\mathcal{X}} < \infty\}.$$

It is easily seen that $\mathcal{X}$ is a Banach space.

Each choice of $\boldsymbol{x} \in \mathcal{X}$ corresponds to a continuous function $b(\boldsymbol{x}) \in \mathcal{C}[0,1]$, defined by

$$b(\boldsymbol{x})(t) = b(\boldsymbol{x}, t) := x_0\, \eta_0(t) + \sum_{n=1}^{\infty} \sum_{i=1}^{2^{n-1}} x_{n,i}\, \eta_{n,i}(t), \quad t \in [0,1]. \tag{4.8}$$

Note that $|b(\boldsymbol{x}, t)| \le \|\boldsymbol{x}\|_{\mathcal{X}}$, implying

$$\|b(\boldsymbol{x})\|_{\infty} \le \|\boldsymbol{x}\|_{\mathcal{X}} < \infty \qquad \text{for} \quad \boldsymbol{x} \in \mathcal{X}, \tag{4.9}$$

so that (4.8) does indeed define a continuous function.

We define $\mathcal{A}_{\mathbb{R}^{\infty}}$ to be the σ-algebra generated by products of Borel sets of $\mathbb{R}$, see [4, p. 372]. On the Banach space $\mathcal{X}$, we now define a product Gaussian measure (see [4, p. 392 and Example 2.35])

$$\rho(\mathrm{d}\boldsymbol{x}) := \bigotimes_{j=1}^{\infty} \rho_j(\mathrm{d}x_j), \tag{4.10}$$

where

$$\rho_j(\mathrm{d}x_j) := \rho(x_j)\, \mathrm{d}x_j = \frac{\exp(-\frac{1}{2}x_j^2)}{\sqrt{2\pi}}\, \mathrm{d}x_j.$$

The space $\mathcal{X}$ has full Gaussian measure: indeed, the norm $\|\boldsymbol{X}(\omega)\|_{\mathcal{X}}$ corresponds precisely to the sum $\sum_{j=1}^{\infty} H_n(\omega)$ in the sketched proof of Theorem 4.1, thus we have

$$P\left(\|\boldsymbol{X}(\cdot)\|_{\mathcal{X}} < \infty\right) = P\left(\sum_{j=1}^{\infty} H_n(\cdot) < \infty\right) = 1.$$

We now study integration on the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^{\infty}}, \rho)$, and we denote the integral of a measurable function f by

$$\int_{\mathcal{X}} f(\boldsymbol{x})\, \rho(\mathrm{d}\boldsymbol{x}). \tag{4.11}$$

In particular, in connection with the Lévy-Ciesielski construction of the Brownian path, we consider functions f of the special form

$$f(\boldsymbol{x}) := g(b(\boldsymbol{x})) = g\left(x_0\eta_0 + \sum_{n=1}^{\infty} \sum_{i=1}^{2^{n-1}} x_{n,i}\, \eta_{n,i}\right), \quad \boldsymbol{x} \in \mathcal{X}, \tag{4.12}$$

where $g : \mathcal{C}[0,1] \to \mathbb{R}$ is non-negative and continuous, and satisfies the conditions in the following theorem. The theorem ensures that the particular functions g from our finance problem, see (4.2), are integrable in the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^{\infty}}, \rho)$, and that the integral of $f = g(b)$ is equal to the path integral given by (4.1).

Theorem 4.2. *Let f be given by (4.12), where $g : \mathcal{C}[0,1] \to \mathbb{R}$ is non-negative and continuous, and satisfies*

$$0 \le f(\boldsymbol{x}) = g(b(\boldsymbol{x})) \le G(\|\boldsymbol{x}\|_{\mathcal{X}}) \quad \textit{for} \quad b \in \mathcal{C}[0,1],$$

for some function $G : \mathbb{R} \to \mathbb{R}^+$ which is monotone increasing and having the property that $G(\|\cdot\|_{\mathcal{X}})$ is integrable in the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^\infty}, \rho)$. Then f is integrable in the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^\infty}, \rho)$, and

$$\int_{\mathcal{X}} f(\boldsymbol{x})\, \rho(\mathrm{d}\boldsymbol{x}) \;=\; \int_{\mathcal{C}[0,1]} g(b)\, \mu(\mathrm{d}b). \tag{4.13}$$

In the theorem above, the integrability of the measurable function f is ensured by f being upper-bounded by an integrable function $G(\|\cdot\|_{\mathcal{X}})$. Equation (4.13) expresses the fact that the integral on the left is merely a concrete representation of the expected value of g over the Brownian paths on the right. In effect, following [3, Theorem 37.1], in the representation on the left-hand side we have redefined the probability space to exclude discontinuous Brownian paths, which occur with probability zero.

In our finance application we see from (4.2) that g is non-negative and continuous, and that the upper bound on g in the theorem is satisfied using (4.9) with G given by

$$G(\alpha) \;:=\; e^{-r} S(0) \exp\left(\left| r - \frac{\sigma^2}{2} \right| + \sigma\alpha \right), \tag{4.14}$$

which is monotone increasing, and which by another application of Fernique's theorem $G(\|\cdot\|_{\mathcal{X}})$ is integrable in the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^\infty}, \rho)$. Thus the conditions of the theorem are satisfied.

4.4. **Approximation by a sequence of finite-dimensional integrals.** Given a function f integrable in the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^\infty}, \rho)$, we define

$$f^{[d]}(\boldsymbol{x}) \;:=\; f^{[d]}(\boldsymbol{x}_{\{1:d\}}) \;:=\; f(\boldsymbol{x}_{\{1:d\}}, \boldsymbol{0}_{\{d+1:\infty\}}), \quad \boldsymbol{x} \in \mathcal{X}. \tag{4.15}$$

Then, because $f^{[d]}$ depends only on the components x_j of $\boldsymbol{x}$ for which $j \in \{1:d\}$, it follows from the definition of the product measure ρ, see (4.10) and (3.1), that

$$\int_{\mathcal{X}} f^{[d]}(\boldsymbol{x})\, \rho(\mathrm{d}\boldsymbol{x}) \;=\; \int_{\mathbb{R}^d} f^{[d]}(\boldsymbol{x}_{\{1:d\}})\, \rho_d(\boldsymbol{x}_{\{1:d\}})\, \mathrm{d}\boldsymbol{x}_{\{1:d\}}. \tag{4.16}$$

Note that for f of the form (4.12) and

$$d \;=\; 1 + \sum_{n=1}^{N} 2^{n-1} = 2^N, \tag{4.17}$$

we have

$$f^{[d]}(\boldsymbol{x}) \;=\; g\biggl(x_0 \eta_0 + \sum_{n=1}^{N} \sum_{i=1}^{2^{n-1}} x_{n,i}\, \eta_{n,i} \biggr),$$

which is just the Brownian bridge approximation to the Brownian path. Thus in this case the integral (4.16) is the standard Brownian bridge approximation to the Wiener path integral (4.1). (The integral (4.16) is, however, not exactly the same as the discrete problem (2.5) in that a rectangle rule is used there to approximate the inner integral over t.)

For $p \in [1,\infty)$, let $\mathcal{L}_p(\mathcal{X})$ denote the class of measurable functions in the measure space $(\mathcal{X}, \mathcal{A}_{\mathbb{R}^\infty}, \rho)$ with finite norm

$$\|f\|_{\mathcal{L}_p(\mathcal{X})} := \left(\int_{\mathcal{X}} |f(\boldsymbol{x})|^p \, \rho(\mathrm{d}\boldsymbol{x})\right)^{1/p}. \tag{4.18}$$

By the Hölder inequality we have $\|f\|_{\mathcal{L}_1(\mathcal{X})} \le \|f\|_{\mathcal{L}_p(\mathcal{X})}$ for $p \in (1,\infty)$.

In the next theorem we assume continuity of f. Note that continuity is well defined on the space $\mathcal{X}$ with norm (4.7): f is continuous if for each $\boldsymbol{x}_0 \in \mathcal{X}$

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_{\mathcal{X}} \to 0 \quad \Longrightarrow \quad |f(\boldsymbol{x}) - f(\boldsymbol{x}_0)| \to 0. \tag{4.19}$$

The theorem holds for a more general function f than (4.12).

Theorem 4.3. *Let $f \in \mathcal{L}_1(\mathcal{X})$ be continuous and satisfy $|f(\boldsymbol{x})| \le G(\|\boldsymbol{x}\|_{\mathcal{X}})$ for all $\boldsymbol{x} \in \mathcal{X}$ and for some function $G : \mathbb{R} \to \mathbb{R}^+$ which is monotone increasing and with the property that $G(\|\cdot\|_{\mathcal{X}}) \in \mathcal{L}_1(\mathcal{X})$. Define $f^{[d]}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathcal{X}$ as in (4.15). Then*

$$\lim_{d\to\infty} f^{[d]}(\boldsymbol{x}) \;=\; f(\boldsymbol{x}) \qquad \textit{for all} \qquad \boldsymbol{x} \in \mathcal{X},$$

and

$$\lim_{d\to\infty} \int_{\mathcal{X}} f^{[d]}(\boldsymbol{x})\, \rho(\mathrm{d}\boldsymbol{x}) \;=\; \int_{\mathcal{X}} f(\boldsymbol{x})\, \rho(\mathrm{d}\boldsymbol{x}). \tag{4.20}$$

Proof. Let $N = \lfloor \log_2 d \rfloor$. Then $2^N \le d < 2^{N+1}$, and from (4.7)

$$\begin{aligned}\|\boldsymbol{x} - (\boldsymbol{x}_{\{1:d\}}, \boldsymbol{0}_{\{d+1:\infty\}})\|_{\mathcal{X}} \;&\le\; \sum_{n=N}^{\infty} \max_{1\le i\le 2^{n-1}} |x_{n,i}|\, 2^{-(n+1)/2} \\ &\to 0 \quad \text{as} \quad d\to\infty \quad \text{for} \quad \boldsymbol{x}\in\mathcal{X}.\end{aligned}$$

Since f is continuous, we conclude that $f^{[d]}(\boldsymbol{x}) = f(\boldsymbol{x}_{\{1:d\}}, \boldsymbol{0}_{\{d+1:\infty\}})$ converges to $f(\boldsymbol{x})$ pointwise. Also we have

$$|f^{[d]}(\boldsymbol{x})| \;=\; |f(\boldsymbol{x}_{\{1:d\}}, \boldsymbol{0}_{\{d+1:\infty\}})| \;\le\; G(\|\boldsymbol{x}_{\{1:d\}}, \boldsymbol{0}_{\{d+1:\infty\}}\|_{\mathcal{X}}) \;\le\; G(\|\boldsymbol{x}\|_{\mathcal{X}})$$

which is integrable. The dominated convergence theorem (see [3, Theorem 16.4]) then ensures that (4.20) holds. □

5. ANOVA decomposition of a function of infinitely many variables

In this section we seek to develop an ANOVA decomposition for a function of the form (4.12), where $g : \mathcal{C}[0,1] \to \mathbb{R}$ satisfies the conditions in Theorem 4.2. Note that $\boldsymbol{x} \in \mathcal{X}$ has an infinite number of components.

5.1. Defining the projections and the ANOVA terms.

As in the finite ANOVA decomposition in Section 3, the principal tool for ANOVA decomposition is integration with respect to a restricted subset of the variables. For $f \in \mathcal{L}_1(\mathcal{X})$ and $\mathfrak{u} \subseteq \mathbb{N}$ (with $\mathfrak{u}$ either finite or infinite), we can write the product Gaussian measure (4.10) as

$$\rho(\mathrm{d}\boldsymbol{x}) \;=\; \rho_{\mathfrak{u}}(\mathrm{d}\boldsymbol{x}_{\mathfrak{u}}) \otimes \rho_{\mathbb{N}\setminus\mathfrak{u}}(\mathrm{d}\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}),$$

where $\rho_{\mathfrak{u}}(\mathrm{d}\boldsymbol{x}_{\mathfrak{u}}) := \otimes_{j\in\mathfrak{u}} \rho_j(\mathrm{d}x_j)$. Fubini's theorem [3, Section 18] allows us to interchange the order of integration. We define $P_{\mathfrak{u}} f$ to be the function obtained by integrating out all the variables labeled by the indices in $\mathfrak{u}$; that is, for $\boldsymbol{x} \in \mathcal{X}$,

$$(P_{\mathfrak{u}} f)(\boldsymbol{x}) \;:=\; (P_{\mathfrak{u}} f)(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}) \;=\; \int_{\mathcal{X}_{\mathfrak{u}}} f(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})\, \rho_{\mathfrak{u}}(\mathrm{d}\boldsymbol{y}_{\mathfrak{u}}) \;=\; \int_{\mathcal{X}} f(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})\, \rho(\mathrm{d}\boldsymbol{y}).$$

Here $\boldsymbol{y}_{\mathfrak{u}}$ denotes the restriction of $\boldsymbol{y}$ to the components with indices in the set $\mathfrak{u}$, $\mathcal{X}_{\mathfrak{u}}$ denotes the corresponding restriction from $\mathcal{X}$, while $(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})$ is the sequence whose jth component is x_j if $j \notin \mathfrak{u}$ and is y_j if $j \in \mathfrak{u}$. In particular,

$$(P_j f)(\boldsymbol{x}) \,:=\, (P_{\{j\}} f)(\boldsymbol{x}) \,=\, \int_{-\infty}^{\infty} f(\boldsymbol{x}_{\mathbb{N}\setminus\{j\}}, y_j)\, \rho(y_j)\, \mathrm{d}y_j, \qquad j \in \mathbb{N},$$

while $P_{\mathfrak{u}} f = (\prod_{j\in\mathfrak{u}} P_j) f$ for $\mathfrak{u}$ a finite subset of $\mathbb{N}$, and $P_{\mathbb{N}} f$ is precisely the integral (4.11). It follows from Fubini's theorem that for any $\mathfrak{u}, \mathfrak{v} \subseteq \mathbb{N}$ we have $P_{\mathfrak{u}} P_{\mathfrak{v}} = P_{\mathfrak{u}\cup\mathfrak{v}}$.

Lemma 5.1. *For $p \in [1,\infty)$, if $f \in \mathcal{L}_p(\mathcal{X})$ then $P_{\mathfrak{u}} f \in \mathcal{L}_p(\mathcal{X})$ for all $\mathfrak{u} \subseteq \mathbb{N}$.*

Proof. We note first that for $f \in \mathcal{L}_p(\mathcal{X})$ and arbitrary $\mathfrak{u} \subseteq \mathbb{N}$ we can write

$$\begin{aligned} \|f\|_{\mathcal{L}_p(\mathcal{X})}^p \,&=\, \int_{\mathcal{X}} |f(\boldsymbol{x})|^p \rho(\mathrm{d}\boldsymbol{x}) \,=\, \int_{\mathcal{X}} |f(\boldsymbol{x})|^p \left(\rho_{\mathfrak{u}}(\mathrm{d}\boldsymbol{x}_{\mathfrak{u}}) \otimes \rho_{\mathbb{N}\setminus\mathfrak{u}}(\mathrm{d}\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}})\right) \\ &=\, \int_{\mathcal{X}} \left(\int_{\mathcal{X}} |f(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})|^p \rho(\mathrm{d}\boldsymbol{y}) \right) \rho(\mathrm{d}\boldsymbol{x}), \end{aligned}$$

where the last step follows from Fubini's theorem. We have

$$\begin{aligned} \|P_{\mathfrak{u}} f\|_{\mathcal{L}_p(\mathcal{X})}^p \,&=\, \int_{\mathcal{X}} \left| \int_{\mathcal{X}} f(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})\, \rho(\mathrm{d}\boldsymbol{y}) \right|^p \rho(\mathrm{d}\boldsymbol{x}) \\ &\le\, \int_{\mathcal{X}} \left(\int_{\mathcal{X}} |f(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})|\, \rho(\mathrm{d}\boldsymbol{y}) \right)^p \rho(\mathrm{d}\boldsymbol{x}) \\ &\le\, \int_{\mathcal{X}} \left(\int_{\mathcal{X}} |f(\boldsymbol{x}_{\mathbb{N}\setminus\mathfrak{u}}, \boldsymbol{y}_{\mathfrak{u}})|^p\, \rho(\mathrm{d}\boldsymbol{y}) \right) \rho(\mathrm{d}\boldsymbol{x}) \,=\, \|f\|_{\mathcal{L}_p(\mathcal{X})}^p \,<\, \infty, \end{aligned}$$

where the second inequality follows from $\|\cdot\|_{\mathcal{L}_1(\mathcal{X})} \le \|\cdot\|_{\mathcal{L}_p(\mathcal{X})}$. This completes the proof. □

For f belonging to the least restrictive space $\mathcal{L}_1(\mathcal{X})$ and for $\mathfrak{u}$ a finite subset of $\mathbb{N}$, we define the ANOVA term $f_{\mathfrak{u}}$ by (cf. (3.3))

$$f_{\mathfrak{u}} \,:=\, \sum_{\mathfrak{v}\subseteq\mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} P_{\mathbb{N}\setminus\mathfrak{v}} f. \tag{5.1}$$

From this and Lemma 5.1 we conclude that if $f \in \mathcal{L}_p(\mathcal{X})$ then $f_{\mathfrak{u}} \in \mathcal{L}_p(\mathcal{X})$.

Lemma 5.2. *Let $f \in \mathcal{L}_1(\mathcal{X})$ and let $\mathfrak{u}$ be a non-empty finite subset of $\mathbb{N}$. Then the ANOVA term $f_{\mathfrak{u}}$ from* (5.1) *satisfies*

$$P_j f_{\mathfrak{u}} = 0 \quad \textit{for } j \in \mathfrak{u}.$$

Proof. We have for $j \in \mathfrak{u}$,

$$P_j f_{\mathfrak{u}} \,=\, \sum_{\mathfrak{v}\subseteq\mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} P_j P_{\mathbb{N}\setminus\mathfrak{v}} f.$$

On splitting the terms on the right-hand side into those for which $j \notin \mathfrak{v}$ and those for which $j \in \mathfrak{v}$ (in which case we write $\mathfrak{v} = \mathfrak{w} \cup \{j\}$, with $j \notin \mathfrak{w}$), we obtain

$$P_j f_{\mathfrak{u}} \,=\, \sum_{j\notin\mathfrak{v}\subseteq\mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} P_{\mathbb{N}\setminus\mathfrak{v}} f + \sum_{j\notin\mathfrak{w}\subseteq\mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{w}|+1} P_{\mathbb{N}\setminus\mathfrak{w}} f \,=\, 0,$$

as required. □

Lemma 5.3. *Let $f \in \mathcal{L}_1(\mathcal{X})$ and let $\mathfrak{u}$ be a finite subset of $\mathbb{N}$. Then the ANOVA terms $f_{\mathfrak{v}}$ from* (5.1) *with $\mathfrak{v} \subseteq \mathfrak{u}$ satisfy*

$$\sum_{\mathfrak{v}\subseteq\mathfrak{u}} f_{\mathfrak{v}} \;=\; P_{\mathbb{N}\setminus\mathfrak{u}} f, \quad \textit{or equivalently,} \quad f_{\mathfrak{u}} \;=\; P_{\mathbb{N}\setminus\mathfrak{u}} f - \sum_{\mathfrak{v}\subsetneq\mathfrak{u}} f_{\mathfrak{v}}.$$

Proof. We have

$$\begin{aligned}
\sum_{\mathfrak{v}\subseteq\mathfrak{u}} f_{\mathfrak{v}}(\boldsymbol{x}) &= \sum_{\mathfrak{v}\subseteq\mathfrak{u}}\sum_{\mathfrak{w}\subseteq\mathfrak{v}} (-1)^{|\mathfrak{v}|-|\mathfrak{w}|}(P_{\mathbb{N}\setminus\mathfrak{w}} f)(\boldsymbol{x}) \\
&= \sum_{\mathfrak{v}\subseteq\mathfrak{u}}\sum_{\mathfrak{w}\subseteq\mathfrak{v}} (-1)^{|\mathfrak{v}|-|\mathfrak{w}|}\int_{\mathcal{X}} f(\boldsymbol{x}_{\mathfrak{w}}, \boldsymbol{y}_{\mathbb{N}\setminus\mathfrak{w}})\,\rho(\mathrm{d}\boldsymbol{y}) \\
&= \int_{\mathcal{X}} \sum_{\mathfrak{v}\subseteq\mathfrak{u}}\sum_{\mathfrak{w}\subseteq\mathfrak{v}} (-1)^{|\mathfrak{v}|-|\mathfrak{w}|} f(\boldsymbol{x}_{\mathfrak{w}}, \boldsymbol{y}_{\mathbb{N}\setminus\mathfrak{w}})\,\rho(\mathrm{d}\boldsymbol{y}) \\
&= \int_{\mathcal{X}} \sum_{\mathfrak{v}\subseteq\mathfrak{u}}\sum_{\mathfrak{w}\subseteq\mathfrak{v}} (-1)^{|\mathfrak{v}|-|\mathfrak{w}|} f(\boldsymbol{x}_{\mathfrak{w}}, \boldsymbol{y}_{\mathfrak{u}\setminus\mathfrak{w}}, \boldsymbol{y}_{\mathbb{N}\setminus\mathfrak{u}})\,\rho(\mathrm{d}\boldsymbol{y}) \\
&= \int_{\mathcal{X}} f(\boldsymbol{x}_{\mathfrak{u}}, \boldsymbol{y}_{\mathbb{N}\setminus\mathfrak{u}})\,\rho(\mathrm{d}\boldsymbol{y}) \\
&= (P_{\mathbb{N}\setminus\mathfrak{u}} f)(\boldsymbol{x}),
\end{aligned}$$

where for the second last equality we used, for a given $\boldsymbol{y}_{\mathbb{N}\setminus\mathfrak{u}}$, the identity (3.6) with $\{1:d\}$ replaced by $\mathfrak{u}$, $\mathfrak{u}$ replaced by $\mathfrak{v}$, $\mathfrak{v}$ replaced by $\mathfrak{w}$, and with anchor $\boldsymbol{c}_{\{1:d\}}$ replaced by $\boldsymbol{y}_{\mathfrak{u}}$. □

5.2. **Infinite sum of the ANOVA terms.** To establish the ANOVA decomposition of a function of infinitely many variables, we need to show that, in a certain precise sense, $f(\boldsymbol{x})$ is expressible as an infinite sum of ANOVA terms $f_{\mathfrak{u}}(\boldsymbol{x})$.

Theorem 5.4. *Let $f \in \mathcal{L}_1(\mathcal{X})$ be continuous and satisfy $|f(\boldsymbol{x})| \le G(\|\boldsymbol{x}\|_{\mathcal{X}})$ for all $\boldsymbol{x} \in \mathcal{X}$ for some function $G : \mathbb{R} \to \mathbb{R}^+$ which is monotone increasing and with the property that $G(\|\cdot\|_{\mathcal{X}} + a) \in \mathcal{L}_1(\mathcal{X})$ for arbitrary $a \in \mathbb{R}$. Then the truncated sum of the ANOVA terms $f_{\mathfrak{u}}(\boldsymbol{x})$ from* (5.1) *converges pointwise to $f(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathcal{X}$: precisely,*

$$\lim_{d\to\infty} \sum_{\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \;=\; f(\boldsymbol{x}) \qquad \textit{for all} \quad \boldsymbol{x} \in \mathcal{X}. \tag{5.2}$$

Proof. We have from Lemma 5.3, with $\mathfrak{u}$ replaced by $\{1:d\}$ and $\mathfrak{v}$ replaced by $\mathfrak{u}$, that for all $\boldsymbol{x} \in \mathcal{X}$,

$$\sum_{\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \;=\; \int_{\mathcal{X}} f(\boldsymbol{x}_{\{1:d\}}, \boldsymbol{y}_{\{d+1:\infty\}})\,\rho(\mathrm{d}\boldsymbol{y}). \tag{5.3}$$

Thus

$$f(\boldsymbol{x}) - \sum_{\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \;=\; \int_{\mathcal{X}} \Big(f(\boldsymbol{x}) - f(\boldsymbol{x}_{\{1:d\}}, \boldsymbol{y}_{\{d+1:\infty\}}) \Big)\,\rho(\mathrm{d}\boldsymbol{y}).$$

Now let $N = \lfloor \log_2 d \rfloor$. Because f is continuous, and because

$$\begin{aligned}\|\boldsymbol{x} - (\boldsymbol{x}_{\{1:d\}}, \boldsymbol{y}_{\{d+1:\infty\}})\|_{\mathcal{X}} &= \|(\mathbf{0}_{\{1:d\}}, \boldsymbol{x}_{\{d+1:\infty\}} - \boldsymbol{y}_{\{d+1:\infty\}})\|_{\mathcal{X}} \\ &\le \|(\mathbf{0}_{\{1:d\}}, \boldsymbol{x}_{\{d+1:\infty\}})\|_{\mathcal{X}} + \|(\mathbf{0}_{\{1:d\}}, \boldsymbol{y}_{\{d+1:\infty\}})\|_{\mathcal{X}} \\ &\le \sum_{n=N}^{\infty} \Big(\max_{1\le i\le 2^{n-1}} |x_{n,i}| + \max_{1\le i\le 2^{n-1}} |y_{n,i}| \Big) 2^{-(n+1)/2} \\ &\to 0 \quad \text{as} \quad d \to \infty \quad \text{for} \quad \boldsymbol{y}, \boldsymbol{x} \in \mathcal{X},\end{aligned}$$

it follows that $f(\boldsymbol{x}) - f(\boldsymbol{x}_{\{1:d\}}, \boldsymbol{y}_{\{d+1:\infty\}})$ converges to 0 pointwise as a function of $\boldsymbol{y}$ as $d \to \infty$. Since we also have

$$|f(\boldsymbol{x}) - f(\boldsymbol{x}_{\{1:d\}}; \boldsymbol{y}_{\{d+1:\infty\}})| \le G(\|\boldsymbol{x}\|_{\mathcal{X}}) + G(\|\boldsymbol{x}\|_{\mathcal{X}} + \|\boldsymbol{y}\|_{\mathcal{X}}),$$

it follows from the dominated convergence theorem that

$$\int_{\mathcal{X}} \Big(f(\boldsymbol{x}) - f(\boldsymbol{x}_{\{1:d\}}; \boldsymbol{y}_{\{d+1:\infty\}}) \Big)\, \rho(\mathrm{d}\boldsymbol{y}) \to 0 \quad \text{as} \quad d \to \infty.$$

This completes the proof. □

Note that the f for our finance problem satisfies the condition in the theorem, with G defined by (4.14).

The integrated form of (5.2) is trivial, since for all d we already have from Lemma 5.2 that

$$\int_{\mathcal{X}} \Big(\sum_{\mathfrak{u} \subseteq \{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \Big) \rho(\mathrm{d}\boldsymbol{x}) \;=\; f_{\emptyset} \;=\; \int_{\mathcal{X}} f(\boldsymbol{x})\, \rho(\mathrm{d}\boldsymbol{x}). \tag{5.4}$$

Additionally, from (5.2) we conclude that

$$\Big| \sum_{\mathfrak{u} \subseteq \{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) - \sum_{\mathfrak{u} \subseteq \{1:d-1\}} f_{\mathfrak{u}}(\boldsymbol{x}) \Big| = \Big| \sum_{d \in \mathfrak{u} \subseteq \{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \Big| \to 0 \qquad \text{as} \quad d \to \infty,$$

but this does *not* allow us to conclude that $f_{\mathfrak{u}}(\boldsymbol{x}) \to 0$ as $|\mathfrak{u}| \to \infty$, since the signs of the terms in the latter sum are uncontrolled.

The following result links the dth partial sum of the ANOVA decomposition of a function f of infinitely many variables to the finite-dimensional version $f^{[d]}$. Both functions depend only on the variables $x_1, \ldots, x_d$, so we state the result in a form in which this is apparent. The corollary follows immediately from Theorems 4.3 and 5.4 and Equation (5.4).

Corollary 5.5. *Under the conditions in Theorems 4.3 and 5.4, we define $f^{[d]}$ as in* (4.15)*. Then the truncated sum of the ANOVA decomposition of f satisfies*

$$\lim_{d\to\infty} \Big| \sum_{\mathfrak{u} \subseteq \{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}_{\{1:d\}}) - f^{[d]}(\boldsymbol{x}_{\{1:d\}}) \Big| \;=\; 0 \qquad \textit{for all} \quad \boldsymbol{x} \in \mathcal{X},$$

and

$$\begin{aligned}\lim_{d\to\infty} \Big| \int_{\mathbb{R}^d} \Big(\sum_{\mathfrak{u} \subseteq \{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}_{\{1:d\}}) \Big) \rho_d(\boldsymbol{x}_{\{1:d\}}) \mathrm{d}\boldsymbol{x}_{\{1:d\}} \\ - \int_{\mathbb{R}^d} f^{[d]}(\boldsymbol{x}_{\{1:d\}})\, \rho_d(\boldsymbol{x}_{\{1:d\}})\, \mathrm{d}\boldsymbol{x}_{\{1:d\}} \Big| \;=\; 0.\end{aligned}$$

5.3. **Analysis of variance.** As the name ANOVA (analysis of variance) suggests, the decomposition even in the finite-dimensional setting has a strong connection with variance, and so requires an $\mathcal{L}_2$ theory. For the next two results we need a stronger assumption than $f \in \mathcal{L}_1(\mathcal{X})$, namely, that $f \in \mathcal{L}_2(\mathcal{X})$, so that the variance

$$\begin{aligned}\sigma^2(f) \;:=\; P_{\mathbb{N}}(f - P_{\mathbb{N}} f)^2 \;&=\; P_{\mathbb{N}}(f^2) - (P_{\mathbb{N}} f)^2 \\ &=\; \int_{\mathcal{X}} f^2(\boldsymbol{x})\,\rho(\mathrm{d}\boldsymbol{x}) - \left(\int_{\mathcal{X}} f(\boldsymbol{x})\,\rho(\mathrm{d}\boldsymbol{x})\right)^2\end{aligned}$$

is well defined.

Lemma 5.6. *Let $f \in \mathcal{L}_2(\mathcal{X})$, and assume that $\mathfrak{u}, \mathfrak{v} \subseteq \mathbb{N}$ with $\mathfrak{u} \neq \mathfrak{v}$. Then*

$$P_{\mathbb{N}}(f_{\mathfrak{u}} f_{\mathfrak{v}}) \;=\; 0.$$

Proof. Since $\mathfrak{u} \neq \mathfrak{v}$, there exists $j \in \mathbb{N}$ such that *either* $j \in \mathfrak{u}$, $j \notin \mathfrak{v}$, *or* $j \notin \mathfrak{u}$, $j \in \mathfrak{v}$. In either case by Lemma 5.2 above we have $P_j(f_{\mathfrak{u}} f_{\mathfrak{v}}) = 0$, and hence $P_{\mathbb{N}}(f_{\mathfrak{u}} f_{\mathfrak{v}}) = P_{\mathbb{N}\setminus\{j\}} P_j(f_{\mathfrak{u}} f_{\mathfrak{v}}) = 0$. □

Theorem 5.7. *Let $f \in \mathcal{L}_2(\mathcal{X})$ be continuous and satisfy $|f(\boldsymbol{x})| \leq G(\|\boldsymbol{x}\|_{\mathcal{X}})$ for all $\boldsymbol{x} \in \mathcal{X}$ for some function $G : \mathbb{R} \to \mathbb{R}^+$ which is monotone increasing and having the properties that $G(\|\cdot\|_{\mathcal{X}} + a) \in \mathcal{L}_2(\mathcal{X})$ for arbitrary $a \in \mathbb{R}$ and also that*

$$\int_{\mathcal{X}} \int_{\mathcal{X}} (G(\|\boldsymbol{x}\|_{\mathcal{X}} + \|\boldsymbol{y}\|_{\mathcal{X}}))^2\,\rho(\mathrm{d}\boldsymbol{x})\,\rho(\mathrm{d}\boldsymbol{y}) \;<\; \infty.$$

Then the variance of the ANOVA terms of f satisfies

$$\lim_{d\to\infty} \sum_{\mathfrak{u}\subseteq\{1:d\}} \sigma^2(f_{\mathfrak{u}}) \;=\; \sigma^2(f).$$

Proof. Clearly $f_\emptyset = P_{\mathbb{N}} f$, $\sigma^2(f_\emptyset) = 0$, and for $\mathfrak{u}$ a non-empty finite subset of $\mathbb{N}$ we have

$$\sigma^2(f_{\mathfrak{u}}) \;=\; P_{\mathbb{N}}(f_{\mathfrak{u}}^2) - (P_{\mathbb{N}} f_{\mathfrak{u}})^2 \;=\; P_{\mathbb{N}}(f_{\mathfrak{u}}^2),$$

where we used Lemma 5.2. Thus

$$\sum_{\mathfrak{u}\subseteq\{1:d\}} \sigma^2(f_{\mathfrak{u}}) \;=\; \sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} P_{\mathbb{N}}(f_{\mathfrak{u}}^2) \;=\; P_{\mathbb{N}}\Bigg(\sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}^2\Bigg) \;=\; P_{\mathbb{N}}\Bigg(\sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}\Bigg)^2,$$

where we used Lemma 5.6. From Theorem 5.4 we know that for $\boldsymbol{x} \in \mathcal{X}$,

$$\sum_{\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \quad \text{converges to} \quad f(\boldsymbol{x}) \quad \text{as} \quad d \to \infty,$$

which is equivalent to

$$\sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \quad \text{converges to} \quad f(\boldsymbol{x}) - P_{\mathbb{N}} f \quad \text{as} \quad d \to \infty,$$

and in turn this implies that

$$\Bigg(\sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x})\Bigg)^2 \quad \text{converges to} \quad (f(\boldsymbol{x}) - P_{\mathbb{N}} f)^2 \quad \text{as} \quad d \to \infty.$$

It remains to prove that the squared sum above is dominated by a suitable integrable function. We can write as in (5.3) that

$$\begin{aligned}\Big(\sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}(\boldsymbol{x})\Big)^2 &= \Big(\int_{\mathcal{X}} f(\boldsymbol{x}_{\{1:d\}},\boldsymbol{y}_{\{d+1:\infty\}})\,\rho(\mathrm{d}\boldsymbol{y}) - f_\emptyset\Big)^2\\ &\le \Big(\int_{\mathcal{X}} |f(\boldsymbol{x}_{\{1:d\}},\boldsymbol{y}_{\{d+1:\infty\}})|\,\rho(\mathrm{d}\boldsymbol{y}) + |f_\emptyset|\Big)^2\\ &\le 2\Big(\int_{\mathcal{X}} |f(\boldsymbol{x}_{\{1:d\}},\boldsymbol{y}_{\{d+1:\infty\}})|^2\,\rho(\mathrm{d}\boldsymbol{y}) + \int_{\mathcal{X}} |f(\boldsymbol{y})|^2\rho(\mathrm{d}\boldsymbol{y})\Big)\\ &\le 4\int_{\mathcal{X}} (G(\|\boldsymbol{x}\|_{\mathcal{X}}+\|\boldsymbol{y}\|_{\mathcal{X}}))^2\,\rho(\mathrm{d}\boldsymbol{y}),\end{aligned}$$

which by assumption is integrable with respect to $\boldsymbol{x}$. The dominated convergence theorem then yields

$$P_{\mathbb{N}}\Big(\sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:d\}} f_{\mathfrak{u}}\Big)^2 \quad\text{converges to}\quad P_{\mathbb{N}}\,(f-P_{\mathbb{N}}f)^2 = \sigma^2(f) \quad\text{as}\quad d\to\infty.$$

This completes the proof. □

6. Smoothing for functions with kinks

In our finance application, we consider a function f of the form (4.12) with g given, for example, by (4.2). This motivates us to consider a general function f of the form

$$f(\boldsymbol{x}) \;=\; \max(\phi(\boldsymbol{x}),0), \qquad \boldsymbol{x}\in\mathcal{X}, \tag{6.1}$$

with a smooth function ϕ. In the particular example from (4.2) we have

$$\phi(\boldsymbol{x}) \;:=\; e^{-r}\Big[S(0)\int_0^1 \exp\Big(\Big(r-\frac{\sigma^2}{2}\Big)t+\sigma\sum_{j=1}^\infty x_j\,\eta_j(t)\Big)\mathrm{d}t - K\Big]. \tag{6.2}$$

Note that for convenience we here use the single-index labeling for x_j and $\eta_j(t)$, instead of double-index labeling $x_{n,i}$ and $\eta_{n,i}(t)$. For this example, the function f is continuous but has a kink along the manifold $\phi(\boldsymbol{x})=0$. For any $j\in\mathbb{N}$ we have

$$\frac{\partial\phi}{\partial x_j}(\boldsymbol{x}) \;=\; e^{-r}S(0)\int_0^1 \sigma\eta_j(t)\exp\Big(\Big(r-\frac{\sigma^2}{2}\Big)t+\sigma\sum_{k=1}^\infty x_k\,\eta_k(t)\Big)\mathrm{d}t > 0.$$

Clearly the function ϕ can be differentiated repeatedly with respect to any variable, and the derivative will remain nonnegative.

In this section we will establish that *all the ANOVA terms* $f_{\mathfrak{u}}$*,* $|\mathfrak{u}|<\infty$*, of* f *are smooth*. The result holds for all functions ϕ sharing the same characteristics as (6.2), and these will be made precise in Theorem 6.5 below. We will follow the argument from [14] closely, but generalize it to the infinite-dimensional setting. First we need to define what we mean by *smooth* in the infinite-dimensional setting, and for this we need to generalize the definition of *weak derivative*.

6.1. **The smoothness class $\mathcal{W}_p^r(\mathcal{X})$.** For $j \in \mathbb{N}$, let D_j denote the partial derivative operator

$$(D_j f)(\boldsymbol{x}) = \frac{\partial f}{\partial x_j}(\boldsymbol{x}) \qquad \text{for} \quad \boldsymbol{x} \in \mathcal{X}.$$

For any multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots)$ with each $\alpha_j \in \mathbb{N} \cup \{0\}$ and $|\boldsymbol{\alpha}| := \sum_{j=1}^{\infty} \alpha_j < \infty$, we define $a(\boldsymbol{\alpha}) := \{j : \alpha_j > 0\}$ to be the set of *active* indices, and

$$D^{\boldsymbol{\alpha}} = \prod_{j \in a(\boldsymbol{\alpha})} D_j^{\alpha_j} = \prod_{j \in a(\boldsymbol{\alpha})} \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j} = \frac{\partial^{|\boldsymbol{\alpha}|}}{\prod_{j \in a(\boldsymbol{\alpha})} \partial x_j^{\alpha_j}}, \tag{6.3}$$

and we say that the derivative $D^{\boldsymbol{\alpha}} f$ is of order $|\boldsymbol{\alpha}|$.

Let $\mathcal{C}(\mathbb{R}^d) = \mathcal{C}^0(\mathbb{R}^d)$ denote the linear space of continuous functions defined on $\mathbb{R}^d$. For a nonnegative integer $r \geq 0$, we define $\mathcal{C}^r(\mathbb{R}^d)$ to be the space of functions whose classical derivatives of order $\leq r$ are all continuous at every point in $\mathbb{R}^d$, with no limitation on their behavior at infinity. For example, the function $\exp(\sum_{j=1}^d x_j^2)$ belongs to $\mathcal{C}^r(\mathbb{R}^d)$ for all values of r. For convenience we write $\mathcal{C}^\infty(\mathbb{R}^d) = \bigcap_{r \geq 0} \mathcal{C}^r(\mathbb{R}^d)$. We generalize these definitions for any subset of indices $\mathfrak{u} \subset \mathbb{N}$, to have $\mathcal{C}^\infty(\mathbb{R}^{\mathfrak{u}}) = \bigcap_{r \geq 0} \mathcal{C}^r(\mathbb{R}^{\mathfrak{u}})$. We also extend these definitions to the sequence space $\mathcal{X}$ (with continuity defined as in (4.19)), to give

$$\mathcal{C}^\infty(\mathcal{U}) := \bigcap_{r \geq 0} \mathcal{C}^r(\mathcal{U}),$$

where $\mathcal{U} = \mathcal{X}$, or $\mathcal{U}$ is a projection of $\mathcal{X}$ obtained by removing finitely many variables.

In addition to classical derivatives, we shall also need a generalization of the notion of *weak* derivative.

Definition 6.1. For $f \in \mathcal{L}_1(\mathcal{X})$ and $\boldsymbol{\alpha}$ a multi-index with finitely many nonzero entries, i.e., $|\boldsymbol{\alpha}| < \infty$, we say that $w \in \mathcal{L}_1(\mathcal{X})$ is a weak derivative of f with multi-index $\boldsymbol{\alpha}$, and we denote it by $w = D^{\boldsymbol{\alpha}} f$, if it satisfies

$$\begin{aligned} &\int_{\mathbb{R}^{a(\boldsymbol{\alpha})}} w(\boldsymbol{x}_{a(\boldsymbol{\alpha})}, \boldsymbol{y}_{\mathbb{N} \setminus a(\boldsymbol{\alpha})})\, v(\boldsymbol{x}_{a(\boldsymbol{\alpha})})\, \mathrm{d}\boldsymbol{x}_{a(\boldsymbol{\alpha})} \\ &= (-1)^{|\boldsymbol{\alpha}|} \int_{\mathbb{R}^{a(\boldsymbol{\alpha})}} f(\boldsymbol{x}_{a(\boldsymbol{\alpha})}, \boldsymbol{y}_{\mathbb{N} \setminus a(\boldsymbol{\alpha})})\, (D^{\boldsymbol{\alpha}} v)(\boldsymbol{x}_{a(\boldsymbol{\alpha})})\, \mathrm{d}\boldsymbol{x}_{a(\boldsymbol{\alpha})} \end{aligned} \tag{6.4}$$

for all $v \in \mathcal{C}_0^\infty(\mathbb{R}^{a(\boldsymbol{\alpha})})$ and for almost all $\boldsymbol{y} \in \mathcal{X}$, where $\mathcal{C}_0^\infty(\mathbb{R}^{a(\boldsymbol{\alpha})})$ denotes the subset of $\mathcal{C}^\infty(\mathbb{R}^{a(\boldsymbol{\alpha})})$ whose members have compact support in $\mathbb{R}^{a(\boldsymbol{\alpha})}$, and where the derivatives $D^{\boldsymbol{\alpha}} v$ on the right-hand side of (6.4) are classical partial derivatives.

The weak derivative is unique, since a function $w \in \mathcal{L}_1(\mathcal{X})$ that satisfies

$$\int_{\mathbb{R}^{a(\boldsymbol{\alpha})}} w(\boldsymbol{x}_{a(\boldsymbol{\alpha})}, \boldsymbol{y}_{\mathbb{N} \setminus a(\boldsymbol{\alpha})})\, v(\boldsymbol{x}_{a(\boldsymbol{\alpha})})\, \mathrm{d}\boldsymbol{x}_{a(\boldsymbol{\alpha})} = 0$$

for all $v \in \mathcal{C}_0^\infty(\mathbb{R}^{a(\boldsymbol{\alpha})})$ and for almost all $\boldsymbol{y} \in \mathcal{X}$, necessarily vanishes almost everywhere.

It follows easily that $D_j D_k = D_k D_j$ for all $j, k \in \mathbb{N}$, that is, the ordering of the weak first derivatives that make up $D^{\boldsymbol{\alpha}}$ in (6.3) is irrelevant. We stress that the integrals on both sides of (6.4) are over finitely many variables and there is no Gaussian weight function.

For $r \geq 0$ and $p \in [1,\infty)$, we define a class of functions

$$\mathcal{W}_p^r(\mathcal{X}) := \left\{ f \in \mathcal{L}_1(\mathcal{X}) \;:\; \|D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_p(\mathcal{X})} < \infty \quad \text{for all} \quad |\boldsymbol{\alpha}| \leq r \right\},$$

where the $\mathcal{L}_p(\mathcal{X})$ norm is defined in (4.18). For convenience we also write

$$\mathcal{W}_p^0(\mathcal{X}) = \mathcal{L}_p(\mathcal{X}) \quad \text{and} \quad \mathcal{W}_p^\infty(\mathcal{X}) = \bigcap_{r\geq 0} \mathcal{W}_p^r(\mathcal{X}).$$

6.2. **Useful theorems.** In this subsection we generalize the three theorems from [13, Subsection 2.4] to the infinite-dimensional setting. These are needed to obtain our main result in the next subsection.

Theorem 6.2 (The Leibniz Theorem)**.** *Let $p \in [1,\infty)$. For $f \in \mathcal{W}_p^1(\mathcal{X})$ we have*

$$D_k P_j f \;=\; P_j D_k f \qquad \text{for all} \quad j,k \in \mathbb{N} \quad \text{with} \quad j \neq k.$$

Proof. The proof follows closely the steps from the proof of [13, Theorem 2.1]. The differences are due to the new definition of weak derivative. From the definition (6.4) we need to prove, for all $v \in \mathcal{C}_0^\infty(\mathbb{R})$ and all $\boldsymbol{y} \in \mathcal{X}$, that

$$-\int_{-\infty}^{\infty} (P_j f)(x_k, \boldsymbol{y}_{\mathbb{N}\setminus\{j,k\}})\,(D_k v)(x_k)\,\mathrm{d}x_k \;=\; \int_{-\infty}^{\infty} (P_j D_k f)(x_k, \boldsymbol{y}_{\mathbb{N}\setminus\{j,k\}})\,v(x_k)\,\mathrm{d}x_k. \tag{6.5}$$

Since the variables $\boldsymbol{y}_{\mathbb{N}\setminus\{j,k\}}$ are merely passengers in this desired result, we can simplify the notation by leaving them out temporarily (thereby making it clear that the infinite-dimensional setting plays here no role). The LHS of (6.5) becomes

$$\begin{aligned} &-\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(y_j, x_k)\,\rho(y_j)\,\mathrm{d}y_j \right) (D_k v)(x_k)\,\mathrm{d}x_k \\ &= \int_{-\infty}^{\infty} \left(-\int_{-\infty}^{\infty} f(y_j, x_k)\,(D_k v)(x_k)\,\mathrm{d}x_k \right) \rho(y_j)\,\mathrm{d}y_j, \end{aligned} \tag{6.6}$$

where in the last step we used Fubini's theorem to interchange the order of integration. Fubini's theorem is applicable because the last integral with the integrand replaced by its absolute value is finite for $v \in \mathcal{C}_0^\infty(\mathbb{R})$, being bounded by

$$\sup_{x\in\mathrm{supp}(v)} |D_k v(x)| \int_{-\infty}^{\infty} \left(\int_{\mathrm{supp}(v)} |f(y_j, x_k)|\,\mathrm{d}x_k \right) \rho(y_j)\,\mathrm{d}y_j.$$

Now we use again the definition of weak derivative (6.4), this time in the inner integral of (6.6), followed again by Fubini's theorem, to obtain

$$\begin{aligned} \text{RHS of (6.6)} \;&=\; \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} (D_k f)(y_j, x_k)\,v(x_k)\,\mathrm{d}x_k \right) \rho(y_j)\,\mathrm{d}y_j \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} (D_k f)(y_j, x_k)\,\rho(y_j)\,\mathrm{d}y_j \right) v(x_k)\,\mathrm{d}x_k \\ &= \int_{-\infty}^{\infty} (P_j D_k f)(x_k)\,v(x_k)\,\mathrm{d}x_k, \end{aligned}$$

which is equal to the RHS of (6.5) once the passenger variables are restored. This proves for $j \neq k$ that $D_k P_j f$ exists, and is equal to $P_j D_k f$. □

Theorem 6.3 (The Inheritance Theorem)**.** *Let $r \geq 0$ and $p \in [1,\infty)$. For $f \in \mathcal{W}_p^r(\mathcal{X})$ we have*

$$P_j f \;\in\; \mathcal{W}_p^r(\mathcal{X}) \qquad \text{for all} \quad j \in \mathbb{N}.$$

Proof. The case $r = 0$ is already established in Lemma 5.1. Consider now $r \geq 1$. Let $j \in \mathbb{N}$ and let $\boldsymbol{\alpha}$ be any multi-index with $|\boldsymbol{\alpha}| \leq r$ and $\alpha_j = 0$. We follow the proof of [13, Theorem 2.2] and write

$$\begin{aligned} D^{\boldsymbol{\alpha}} P_j f \;&=\; \left(\prod_{i=1}^{|\boldsymbol{\alpha}|} D_{k_i}\right) P_j f \;=\; \left(\prod_{i=2}^{|\boldsymbol{\alpha}|} D_{k_i}\right) P_j D_{k_1} f \\ &=\; \cdots \;=\; D_{k_{|\boldsymbol{\alpha}|}} P_j \left(\prod_{i=1}^{|\boldsymbol{\alpha}|-1} D_{k_i}\right) f \;=\; P_j \left(\prod_{i=1}^{|\boldsymbol{\alpha}|} D_{k_i}\right) f \;=\; P_j D^{\boldsymbol{\alpha}} f, \end{aligned}$$

where each step involves a single differentiation under the integral sign, and is justified by the new Leibniz theorem (Theorem 6.2) because $(\prod_{i=1}^{\ell} D_{k_i}) f \in \mathcal{W}_p^{r-\ell}(\mathcal{X}) \subseteq \mathcal{W}_p^1(\mathcal{X})$ for all $\ell \leq |\boldsymbol{\alpha}| - 1 \leq r - 1$. Note that the k_i need not be distinct. It then follows from Lemma 5.1 that $\|D^{\boldsymbol{\alpha}} P_j f\|_{\mathcal{L}_p(\mathcal{X})} = \|P_j D^{\boldsymbol{\alpha}} f\|_{\mathcal{L}_p(\mathcal{X})} < \infty$. □

The implicit function theorem stated below is crucial for the main results of this paper. For simplicity, we write

$$\boldsymbol{x} = (x_j, \boldsymbol{x}_{\mathbb{N}\setminus\{j\}}) = (x_j, \boldsymbol{x}_{\text{-}j})$$

to separate out the jth component of $\boldsymbol{x}$ in our notation, and we define

$$\mathcal{X}_{\text{-}j} \;:=\; \{\boldsymbol{x}_{\text{-}j} \;:\; (x_j, \boldsymbol{x}_{\text{-}j}) \in \mathcal{X} \text{ for some } x_j \in \mathbb{R}\}.$$

Theorem 6.4 (The Implicit Function Theorem)**.** *Let $j \in \mathbb{N}$. Suppose $\phi \in \mathcal{C}^1(\mathcal{X})$ satisfies*

$$(D_j\phi)(\boldsymbol{x}) \;\neq\; 0 \qquad \textit{for all} \quad \boldsymbol{x} \in \mathcal{X}. \tag{6.7}$$

Let

$$U_j \;:=\; \text{interior}\{\boldsymbol{x}_{\text{-}j} \in \mathcal{X}_{\text{-}j} \;:\; \phi(x_j, \boldsymbol{x}_{\text{-}j}) = 0 \textit{ for some } x_j \in \mathbb{R}\}. \tag{6.8}$$

(*The x_j in* (6.8) *if it exists is necessarily unique because of the condition* (6.7).) *If U_j is not empty then there exists a unique function $\psi_j \in \mathcal{C}^1(U_j)$ such that*

$$\phi(\,\psi_j(\boldsymbol{x}_{\text{-}j}), \boldsymbol{x}_{\text{-}j}) \;=\; 0 \qquad \textit{for all} \quad \boldsymbol{x}_{\text{-}j} \in U_j,$$

and for all $k \neq j$ we have

$$(D_k\psi_j)(\boldsymbol{x}_{\text{-}j}) \;=\; -\frac{(D_k\phi)(\boldsymbol{x})}{(D_j\phi)(\boldsymbol{x})}\bigg|_{x_j = \psi_j(\boldsymbol{x}_{\text{-}j})} \qquad \textit{for all} \quad \boldsymbol{x}_{\text{-}j} \in U_j.$$

If in addition $\phi \in \mathcal{C}^r(\mathcal{X})$ for some $r \geq 2$, then $\psi_j \in \mathcal{C}^r(U_j)$.

Proof. The proof is essentially the same as that of [13, Theorem 2.3], but with $\mathbb{R}^d$ replaced by the sequence space $\mathcal{X}$ which is a metric space, and with $\mathbb{R}^{\mathfrak{D}\setminus\{j\}}$ replaced by $\mathcal{X}_{\text{-}j}$. In particular, [17, Theorem 3.2.1] extends to the sequence space $\mathcal{X}$. □

Note that the derivatives in the implicit function theorem are classical derivatives.

6.3. **Main result.** We now return to our general function f given by (6.1). For any $k \in \mathbb{N}$, we have the weak derivative

$$(D_k f)(\boldsymbol{x}) \;=\; \begin{cases} (D_k\phi)(\boldsymbol{x}) & \text{if } \phi(\boldsymbol{x}) > 0, \\ 0 & \text{if } \phi(\boldsymbol{x}) < 0, \end{cases}$$

noting that the condition (6.7) ensures that the solution set of $\phi(\boldsymbol{x}) = 0$ is of ρ-measure zero.

If $\phi \in \mathcal{W}_p^r(\mathcal{X}) \cap \mathcal{C}^\infty(\mathcal{X})$ for some $r \geq 1$ and $p \in [1,\infty)$, then

$$\begin{aligned} \|D_k f\|_{\mathcal{L}_p(\mathcal{X})} \;&=\; \left(\int_{\mathcal{X}} |(D_k f)(\boldsymbol{x})|^p \,\rho(\mathrm{d}\boldsymbol{x})\right)^{1/p} \\ &=\; \left(\int_{\boldsymbol{x}\in\mathcal{X}\,:\,\phi(\boldsymbol{x})\geq 0} |(D_k\phi)(\boldsymbol{x})|^p \,\rho(\mathrm{d}\boldsymbol{x})\right)^{1/p} \;\leq\; \|D_k\phi\|_{\mathcal{L}_p(\mathcal{X})} \;<\; \infty, \end{aligned}$$

and we conclude that $f \in \mathcal{W}_p^1(\mathcal{X}) \cap \mathcal{C}(\mathcal{X})$. It is clear that in general $f \notin \mathcal{W}_p^r(\mathcal{X})$ for $r > 1$ because of the kink from the max function. However, as we prove in the following theorem, the main result of this paper, we have $P_j f \in \mathcal{W}_p^r(\mathcal{X})$ for arbitrary $r \geq 1$ and all $j \in \mathbb{N}$, provided that a number of conditions on ϕ are satisfied. In addition to the set U_j defined in (6.8), we define

$$\begin{aligned} U_j^+ \;&:=\; \{\boldsymbol{x}_{\text{-}j} \in \mathcal{X}_{\text{-}j} \;:\; \phi(x_j, \boldsymbol{x}_{\text{-}j}) > 0 \text{ for all } x_j \in \mathbb{R}\}, \\ U_j^- \;&:=\; \{\boldsymbol{x}_{\text{-}j} \in \mathcal{X}_{\text{-}j} \;:\; \phi(x_j, \boldsymbol{x}_{\text{-}j}) < 0 \text{ for all } x_j \in \mathbb{R}\}. \end{aligned}$$

Theorem 6.5. *Let $r \geq 1$, $p \in [1,\infty)$, and $j \in \mathbb{N}$. Let f be given by (6.1), where ϕ satisfies the following conditions:*

(i) $\phi \in \mathcal{W}_p^r(\mathcal{X}) \cap \mathcal{C}^\infty(\mathcal{X})$.
(ii) $(D_j\phi)(\boldsymbol{x}) \neq 0$ *for all* $\boldsymbol{x} \in \mathcal{X}$.
(iii) U_j *defined in* (6.8) *is not empty.*
(iv) *With $\psi_j \in \mathcal{C}^\infty(U_j)$ denoting the unique function such that $\phi(\boldsymbol{x}) = 0$ if and only if $x_j = \psi_j(\boldsymbol{x}_{\text{-}j})$ for $\boldsymbol{x} \in \mathcal{X}$, we assume that every function of the form*

$$\begin{cases} \mathcal{G}(\boldsymbol{x}_{\text{-}j}) \;=\; \dfrac{\prod_{i=1}^a [(D^{\boldsymbol{\alpha}^{(i)}}\phi)(\psi_j(\boldsymbol{x}_{\text{-}j}), \boldsymbol{x}_{\text{-}j})]}{[(D_j\phi)(\psi_j(\boldsymbol{x}_{\text{-}j}), \boldsymbol{x}_{\text{-}j})]^b}\, \rho^{(c)}(\psi_j(\boldsymbol{x}_{\text{-}j})), \quad \boldsymbol{x}_{\text{-}j} \in U_j, \\ \textit{where } a, b, c \textit{ are integers and } \boldsymbol{\alpha}^{(i)} \textit{ are multi-indices with the constraints} \\ 2 \leq a \leq 2r-2, \quad 1 \leq b \leq 2r-3, \quad 0 \leq c \leq r-2, \quad |\boldsymbol{\alpha}^{(i)}| \leq r-1, \end{cases}$$

satisfies both

$$\mathcal{G}(\boldsymbol{x}_{\text{-}j}) \to 0 \textit{ as } \boldsymbol{x}_{\text{-}j} \textit{ approaches a boundary point of } U_j \textit{ lying in } U_j^+ \textit{ or } U_j^-, \tag{6.9}$$

and

$$\int_{U_j} |\mathcal{G}(\boldsymbol{x}_{\text{-}j})|^p \,\rho_{\mathbb{N}\setminus\{j\}}(\mathrm{d}\boldsymbol{x}_{\text{-}j}) \;<\; \infty. \tag{6.10}$$

Then

$$P_j f \;\in\; \mathcal{W}_p^r(\mathcal{X}).$$

Moreover, if the above conditions hold for all $j \in \mathbb{N}$, then $P_{\mathfrak{u}} f \in \mathcal{W}_p^r(\mathcal{X})$ for all non-empty subsets $\mathfrak{u} \subseteq \mathbb{N}$, and

$$f_{\mathfrak{u}} \;\in\; \mathcal{W}_p^r(\mathcal{X}) \quad \textit{for all finite subsets} \quad \mathfrak{u} \subset \mathbb{N}.$$

Proof. The proof for $P_j f$ follows closely that of [14, Theorem 1], but makes use of the new theorems from the previous subsection. The result for $f_{\mathfrak{u}}$ then follows from the new inheritance theorem (Theorem 6.3). □

For the finance application in which ϕ is given by (6.2) we have verified already that $(D_j\phi)(\boldsymbol{x}) > 0$ for all $j \in \mathbb{N}$ and all $\boldsymbol{x} \in \mathcal{X}$. In our case, ϕ is essentially a sum of exponential functions involving only linear combinations of $x_1, x_2, \ldots$ in the exponents, and therefore (6.9) holds. The derivatives of ϕ will contain at worst exponential functions of the same form. Fernique's theorem then ensures that the Gaussian expected value (6.10) is finite. Since the result holds for all values of r and all finite values of p, we conclude that for our Brownian bridge continuous option pricing problem we have

$$f_{\mathfrak{u}} \in \mathcal{W}_p^{\infty}(\mathcal{X}) \quad \text{for all finite subset} \quad \mathfrak{u} \subset \mathbb{N}. \tag{6.11}$$

7. Concluding remarks

We have shown in (6.11) that all the ANOVA terms of our finance integrand of infinitely many variables are smooth. Corollary 5.5 indicates that the finite-dimensional (non-smooth) approximation differs from the smooth truncated sum of the ANOVA expansion of the continuous problem by a quantity that converges pointwise to zero. It remains a problem for the future to quantify this rate of convergence, and the rate of convergence to zero of the corresponding difference of Gaussian integrals in Corollary 5.5.

Acknowledgements

The authors thank Pierre Del Moral and Andreas Griewank for valuable discussions. The authors thank the referees for identifying a mistake in our previous manuscript and for providing very insightful suggestions. The authors acknowledge the support of the Australian Research Council under the projects DP110100442, FT130100655, and DP150101770. Michael Griebel was partially supported by the Sonderforschungsbereich 1060 *The Mathematics of Emergent Effects* funded by the Deutsche Forschungsgemeinschaft.

References

1. P. Acworth, M. Broadie, and P. Glasserman, *A comparison of some Monte Carlo and quasi-Monte Carlo techniques for option pricing*, in: Monte Carlo and quasi-Monte Carlo methods 1996 (P. Hellekalek, G. Larcher, H. Niederreiter, and P. Zinterhof, eds.), Springer Verlag, Berlin, 1–18 (1998).
2. J. Baldeaux and M. Gnewuch, *Optimal randomized multilevel algorithms for infinite-dimensional integration on funcion spaces with ANOVA-type decomposition*, SIAM J. Numer. Anal. **52**, 1128–1155 (2014).
3. P. Billingsley, *Probability and Measure*, Wiley, New York, 1979.
4. V. I. Bogachev, *Gaussian Measures*, American Mathematical Society, 1998.
5. H.-J. Bungartz and M. Griebel, *Sparse grids*, Acta Numer. **13**, 1–123 (2004).
6. R. E. Caflisch, W. Morokoff, and A. Owen, *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*, J. Comput. Finance **1**, 27–46 (1997).
7. Z. Ciesielski, *Hölder conditions for realizations of Gaussian processes*, Trans. Amer. Math. Soc. **99**, 403–413 (1961).
8. M. Davis, *Construction of Brownian motion*, Stochastic Processes I, 2004–2005. `http://www2.imperial.ac.uk/~mdavis/msc_home_page.htm`

9. J. Dick and M. Gnewuch, *Optimal randomized changing dimension algorithms for infinite-dimensional integration on funcion spaces with ANOVA-type decomposition*, J. Approx. Theory **184**, 111–145 (2014).
10. J. Dick, F. Y. Kuo, and I. H. Sloan, *High dimensional integration – the quasi-Monte Carlo way*, Acta Numer. **22**, 133–288 (2013).
11. P. Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer-Verlag. 2004.
12. M. Griebel, F. Y. Kuo, and I. H. Sloan, *The smoothing effect of the ANOVA decomposition*, J. Complexity **26**, 523–551 (2010).
13. M. Griebel, F. Y. Kuo, and I. H. Sloan, *The smoothing effect of integration in* $\mathbb{R}^d$ *and the ANOVA decomposition*, Math. Comp. **82**, 383–400 (2013).
14. M. Griebel, F. Y. Kuo, and I. H. Sloan, *Note on "the smoothing effect of integration in* $\mathbb{R}^d$ *and the ANOVA decomposition"*, Math. Comp., submitted together with the revision of the current manuscript.
15. F. J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter, *Multi-level Monte Carlo algorithms for infinite-dimensional integration on* $\mathbb{R}^{\mathbb{N}}$, J. Complexity **26**, 229–254 (2010).
16. J. Imai and K. S. Tan, *Minimizing effective dimension using linear transformation* in: Monte Carlo and Quasi-Monte Carlo Methods 2002 (H. Niederreiter, ed.), Springer Verlag, Berlin, 275–292 (2004).
17. S. G. Krantz and H. R. Parks, *The Implicit Function Theorem: History, Theory and Applications*, Birkhäuser, 2002.
18. F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski and H. Wozniakowski, *On decompositions of multivariate functions*, Math. Comp. **79**, 953–966 (2010).
19. P. Lévy, *Processus Stochastiques et Mouvement Brownien*, Suivi d'une note de M. Loève (French), Gauthier-Villars, Paris, 1948, 365pp.
20. R. Liu and A. Owen, *Estimating mean dimensionality of analysis of variance decompositions*, Journal of the American Statistical Association **101**(474), 712–720 (2006).
21. I. M. Sobol′, *Sensitivity estimates for nonlinear mathematical models*, Matem. Modelirovanie **2**(1), 112–118 (1990) (in Russian), English translation in: Mathematical Modeling and Computational Experiment **1**(4), 407–414 (1993).
22. J. M. Steele, *Stochastic calculus and financial applications*, Applications of Mathematics 45, Springer-Verlag, New York, 2001, 300pp.
23. H. Triebel, *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, EMS Tracts in Mathematics, Zürich, 2010.
24. H. Triebel, *Numerical integration and discrepancy, a new approach*, Math. Nachr. **283**, 139–159 (2010).
25. G. W. Wasilkowski and H. Wozniakowski, *On tractability of path integration*, J. Math. Phys. **37**, 2071–2088 (1996).

Universität Bonn, Institut für Numerische Simulation, Wegelerstr. 6, 53115, Bonn, Germany, and Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Schloss Birlinghoven, 53754 Sankt Augustin, Germany
E-mail address: griebel@ins.uni-bonn.de

School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
E-mail address: f.kuo@unsw.edu.au

School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
E-mail address: i.sloan@unsw.edu.au