Journal of Computational and Applied Mathematics 344 (2018) 259–274

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

High dimensional integration of kinks and jumps—Smoothing by preintegration

Andreas Griewank [a], Frances Y. Kuo [b,*], Hernan Leövey [c], Ian H. Sloan [b]

[a] School of Mathematical Sciences and Information Technology, Yachay Tech, Urcuqui 100119, Imbabura, Ecuador
[b] School of Mathematics and Statistics, The University of New South Wales, Sydney NSW 2052, Australia
[c] Structured Energy Management Team, Axpo AG, Parkstrasse 23, CH-5401 Baden, Switzerland

ARTICLE INFO

Article history:
Received 17 August 2017
Received in revised form 26 February 2018

Keywords:
High dimensional integration
Smoothing
Preintegration
ANOVA decomposition
Quasi Monte Carlo
Conditional sampling

ABSTRACT

We show how simple kinks and jumps of otherwise smooth integrands over $\mathbb{R}^d$ can be dealt with by a preliminary integration with respect to a single well chosen variable. It is assumed that this preintegration, or conditional sampling, can be carried out with negligible error, which is the case in particular for option pricing problems. It is proven that under appropriate conditions the preintegrated function of $d-1$ variables belongs to appropriate mixed Sobolev spaces, so potentially allowing high efficiency of Quasi Monte Carlo and Sparse Grid Methods applied to the preintegrated problem. The efficiency of applying Quasi Monte Carlo to the preintegrated function are demonstrated on a digital Asian option using the Principal Component Analysis factorization of the covariance matrix.

1. Introduction

In the present paper we analyze a natural method for numerical integration over $\mathbb{R}^d$, where d may be large, in the presence of "kinks" (i.e. discontinuities in the gradients) or "jumps" (i.e. discontinuities in the function). In this method one of the variables is integrated out in a "preintegration" step, with the aim of creating a smooth integrand over $\mathbb{R}^{d-1}$.

Integrands with kinks and jumps arise in option pricing, because an option is normally considered worthless if the value falls below a predetermined strike price. In the case of a continuous payoff function this introduces a kink, while in the case of a binary or other digital option it introduces a jump.

A simple strategy is to ignore the kinks and jumps, and apply directly a method for integration over $\mathbb{R}^d$. While there has been very significant recent progress in *Quasi Monte Carlo (QMC) methods* [1] and *Sparse Grid (SG) methods* [2] for high dimensional integration when the integrand is somewhat smooth, there has been little progress in understanding their performance when the integrand has kinks or jumps.

The performance of QMC and SG methods is degraded in the presence of kinks and jumps, but perhaps not as much as might have been expected, given that in both cases the standard error analysis fails in general for kinks and jumps: the standard assumption in both cases is that the integrand has mixed first partial derivatives for all variables, or at least that it has bounded Hardy and Krause variation over the unit cube $[0, 1]^d$, whereas even a straight non-aligned kink (one that is not orthogonal to one of the axes) lacks mixed first partial derivatives even for $d = 2$, and generally exhibits unbounded Hardy and Krause variation on $[0, 1]^d$ for $d \geq 3$ [3].

* Corresponding author.
E-mail addresses: griewank@yachaytech.edu.ec (A. Griewank), f.kuo@unsw.edu.au (F.Y. Kuo), hernaneugenio.leoevey@axpo.com (H. Leövey), i.sloan@unsw.edu.au (I.H. Sloan).

https://doi.org/10.1016/j.cam.2018.04.009
0377-0427/

A possible path towards understanding the performance of QMC and SG methods in the presence of kinks and jumps was developed in [4]. That paper studied the terms of the "ANOVA decomposition" of functions with kinks defined on d-dimensional Euclidean space $\mathbb{R}^d$, and showed that under suitable circumstances all but one of the 2^d ANOVA terms can be smooth, with the single exception of the highest order ANOVA term, the one depending on all d of the variables. If the "effective dimension" of the function is small, as is commonly thought to be the case in applications, then that single non-smooth term can be expected to make a very small contribution to both supremum and $\mathcal{L}_2$ norms. In a subsequent paper [5] the same authors showed, by strengthening the theorems and correcting a mis-statement in [4], that the smoothing of all but the highest order ANOVA term is a reality for the case of an arithmetic Asian option with the Brownian bridge construction.

More precisely, the papers [4] and [5] showed, for a function of the form $f(\boldsymbol{x}) = \max(\phi(\boldsymbol{x}), 0)$ with ϕ smooth (so that f generically has a kink along the manifold $\phi(\boldsymbol{x}) = 0$), that if the d-dimensional function ϕ has a positive partial derivative with respect to x_j for some $j \in \{1, \ldots, d\}$, and if certain growth conditions at infinity are satisfied, then all the ANOVA terms of f that do not depend on the variable x_j are smooth. The underlying reason, as explained in [4], is that integration of f with respect to x_j, under the stated conditions, results in a $(d-1)$-dimensional function that no longer has a kink, and indeed is as often differentiable as the function ϕ.

Going beyond kinks, we prove in this paper that Theorem 1 in [5] can be extended from kinks to jumps—thus jumps are smoothed under almost the same conditions as kinks. The smoothing occurs even in situations (for example in option pricing) where the location of the kink or jump treated as a function of the other $d-1$ variables moves off to infinity for some values of the other variables.

In this paper we pay particular attention to proving that the presmoothed integrand belongs to an appropriate mixed-derivative function space.

The preintegration method studied in the present paper has appeared as a practical tool under other names in many other papers, including those related to "conditional sampling" (see [6]; the paragraph leading up to and including Lemma 7.2 in [7]; the remark at the end of Section 3 in [8]), and other root-finding strategies for identifying where the payoff is positive (see [9,10]), as well as those under the name "smoothing" (see [11,12]). In contrast to the cited papers, the emphasis in this paper is on rigorous analysis. Also, we here prefer the description "preintegration" because to us "conditional sampling" suggests a probabilistic setting, which is not necessarily relevant here.

Even for the classical *Monte Carlo (MC) method* the preintegration step can be useful: to the extent that the preintegration can be considered exact, there is a reduction in the variance of the integrand, by the sum of the variances of all ANOVA terms that involve the preintegration variable x_j (since the ANOVA terms are eliminated because their exact integrals with respect to x_j are all zero). In our numerical experiments that reduction proves to be quite significant.

The problem class and the method are stated in Section 2. Section 3 gives numerical examples in the context of an option pricing problem with 256 time steps, treated as a problem of integration in 256 dimensions. Section 4 briefly discusses the variance reduction by preintegration for $\mathcal{L}_2$ functions. Section 5 focuses on the smoothing effect of preintegration. It gives mathematical background on needed function spaces and states two new smoothing theorems, extended here in a non-trivial way from [5, Theorem 1]. Section 6 applies our theoretical results to the option pricing example. Technical proofs are given in Section 7.

2. The problem and the method

The problem is the approximate evaluation of

$$I_d f := \int_{\mathbb{R}^d} f(\boldsymbol{x})\rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_d)\,\rho_d(\boldsymbol{x})\,\mathrm{d}x_1 \cdots \mathrm{d}x_d, \tag{1}$$

with

$$\rho_d(\boldsymbol{x}) := \prod_{k=1}^{d} \rho(x_k),$$

where ρ is a continuous and strictly positive probability density function on $\mathbb{R}$ with some smoothness, and f is a real-valued function integrable with respect to ρ_d.

To allow for both kinks and jumps we assume that the integrand is of the form

$$f(\boldsymbol{x}) = \theta(\boldsymbol{x})\,\mathrm{ind}(\phi(\boldsymbol{x})), \tag{2}$$

where θ and ϕ are somewhat smooth functions, and $\mathrm{ind}(\cdot)$ is the indicator function which gives the value 1 if the input is positive and 0 otherwise. When $\theta = \phi$ we have $f(\boldsymbol{x}) = \max(\phi(\boldsymbol{x}), 0)$ and thus we have the familiar kink seen in option pricing through the occurrence of a strike price. When θ and ϕ are different (for example, when $\theta(\boldsymbol{x}) = 1$) we have a structure that includes binary digital options.

Our key assumption on $\phi(\boldsymbol{x})$ is that it has a positive partial derivative (and so is an increasing function) with respect to some variable x_j, that is, we assume that for some $j \in \{1, \ldots, d\}$ we have

$$\frac{\partial \phi}{\partial x_j}(\boldsymbol{x}) > 0 \qquad \text{for all} \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{3}$$

In other words ϕ is monotone with respect to x_j.

We also make an assumption about the behavior as $x_j \to +\infty$. To state this it is convenient, given $j \in \{1, \ldots, d\}$, to write the general point $\boldsymbol{x} \in \mathbb{R}^d$ as $\boldsymbol{x} = (x_j, \boldsymbol{x}_{-j})$, where $\boldsymbol{x}_{-j}$ denotes the vector of length $d-1$ denoting all the variables other than x_j. With this notation, a second assumption is that

$$\lim_{x_j \to +\infty} \phi(\boldsymbol{x}) = \lim_{x_j \to +\infty} \phi(x_j, \boldsymbol{x}_{-j}) = +\infty \qquad \text{for fixed} \quad \boldsymbol{x}_{-j} \in \mathbb{R}^{d-1}. \tag{4}$$

The latter growth property follows automatically if we assume in addition to (3) that $(\partial^2\phi/\partial x_j^2)(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^d$. Additional properties at infinity will be assumed in Theorems 2 and 3.

Assuming that the properties (3) and (4) both hold for some $j \in \{1, \ldots, d\}$, the method is easily described: we write (1) as the repeated integral using Fubini's theorem

$$I_d f = \int_{\mathbb{R}^{d-1}} \left(\int_{-\infty}^{\infty} f(x_j, \boldsymbol{x}_{-j})\, \rho(x_j)\, \mathrm{d}x_j \right) \rho_{d-1}(\boldsymbol{x}_{-j})\, \mathrm{d}\boldsymbol{x}_{-j},$$

and first evaluate the inner integral for each needed value of $\boldsymbol{x}_{-j}$. This is the "*preintegration*" step. The essential point of the method is that the outer integral can then be evaluated by a standard QMC or SG method, in the knowledge that the integrand for this $(d-1)$-dimensional integral is smooth.

Looking more closely at the preintegration step, we write

$$I_d f = \int_{\mathbb{R}^{d-1}} (P_j f)(\boldsymbol{x}_{-j})\, \rho_{d-1}(\boldsymbol{x}_{-j})\, \mathrm{d}\boldsymbol{x}_{-j}, \tag{5}$$

where P_j is the operation of integration with respect to x_j,

$$(P_j f)(\boldsymbol{x}_{-j}) := \int_{-\infty}^{\infty} f(x_j, \boldsymbol{x}_{-j})\, \rho(x_j)\, \mathrm{d}x_j. \tag{6}$$

It follows from (2) and (3) that the integrand in this integral has generically a jump at the (unique) point at which $\phi(x_j, \boldsymbol{x}_{-j})$ passes through zero. By the implicit function theorem (see Theorem 6) for each $\boldsymbol{x}_{-j}$ there is a unique value $\psi(\boldsymbol{x}_{-j})$ of x_j at which $\phi(x_j, \boldsymbol{x}_{-j})$ passes from negative to positive values with increasing x_j. The preintegration step may then be written as

$$(P_j f)(\boldsymbol{x}_{-j}) = \int_{\psi(\boldsymbol{x}_{-j})}^{\infty} f(x_j, \boldsymbol{x}_{-j})\, \rho(x_j)\, \mathrm{d}x_j.$$

An essential ingredient in any implementation of the method is the accurate evaluation of $\psi(\boldsymbol{x}_{-j})$, for each point $\boldsymbol{x}_{-j}$ of the outer integration rule. The semi-infinite integral $P_j f$ may then be evaluated, for each needed point $\boldsymbol{x}_{-j}$, by a standard method for 1-dimensional integrals, for example by a formula of Gauss type. On the other hand, in certain important applications such as option pricing, the integration can be performed in more or less closed form. Note that numerical errors in the evaluation of $P_j f$ are benign, in the sense that if the error in $(P_j f)(\boldsymbol{x}_{-j})$ is bounded uniformly by ε then, by (5), the absolute error in $I_d f$ is also bounded by ε. The same conclusion also holds under the weaker condition that the error in $(P_j f)(\boldsymbol{x}_{-j})$, as a function of $\boldsymbol{x}_{-j}$, is integrable with respect to the weight function $\rho_{d-1}(\boldsymbol{x}_{-j})$ and the integral is bounded by ϵ.

The monotonicity condition (3) and the infinite growth condition (4) imply that for fixed $\boldsymbol{x}_{-j}$ the function $\phi(x_j, \boldsymbol{x}_{-j})$ either has a simple root $x_j = \psi(\boldsymbol{x}_{-j})$ or is positive for all $x_j \in \mathbb{R}$. The zero set of ϕ, denoted by

$$\phi^{-1}(0) := \{\boldsymbol{x} \in \mathbb{R}^d : \phi(\boldsymbol{x}) = 0\},$$

is then a hypersurface, i.e., a continuous manifold of dimension $d-1$. However, its projection onto $\mathbb{R}^{d-1}$ obtained by ignoring the component x_j can be very complicated, even if ϕ is highly differentiable.

An example with $d = 2$ and $j = 1$ illustrating the complications that can arise is given by

$$\phi(x_1, x_2) := \begin{cases} \exp(x_1) - x_2^m \sin(1/x_2) & \text{for } x_2 > 0, \\ \exp(x_1) & \text{for } x_2 \leq 0, \end{cases} \tag{7}$$

for some large m. Since $\phi(x_1, x_2)$ is monotonically increasing in x_1, the explicit solution of $\phi(x_1, x_2) = 0$ is

$$x_1 = \psi(x_2) := m \log(x_2) + \log((\sin(1/x_2))_+) \quad \text{for } x_2 \in U_1,$$

where $z_+ := \max(0, z)$, and

$$U_1 := \{x_2 \in \mathbb{R} : \phi(x_1, x_2) = 0 \text{ for some } x_1 \in \mathbb{R}\} = \left(\tfrac{1}{\pi}, \infty\right) \cup \bigcup_{k \in \mathbb{N}} \left(\tfrac{1}{(2k+1)\pi}, \tfrac{1}{2k\pi}\right),$$

while $\phi(x_1, x_2) = 0$ has no solution for x_2 in the complicated complementary set

$$U_1^+ := (-\infty, 0] \cup \bigcup_{k \in \mathbb{N}} \left[\tfrac{1}{2k\pi}, \tfrac{1}{(2k-1)\pi}\right].$$

The graph of the zero set $\phi^{-1}(0)$ for $m = 2$ is shown in Fig. 1. The solid lines represent the zero set, while the broken lines parallel to the horizontal axis define the boundaries on the x_2 axis between subsets of U_1^+ for which there is no solution of

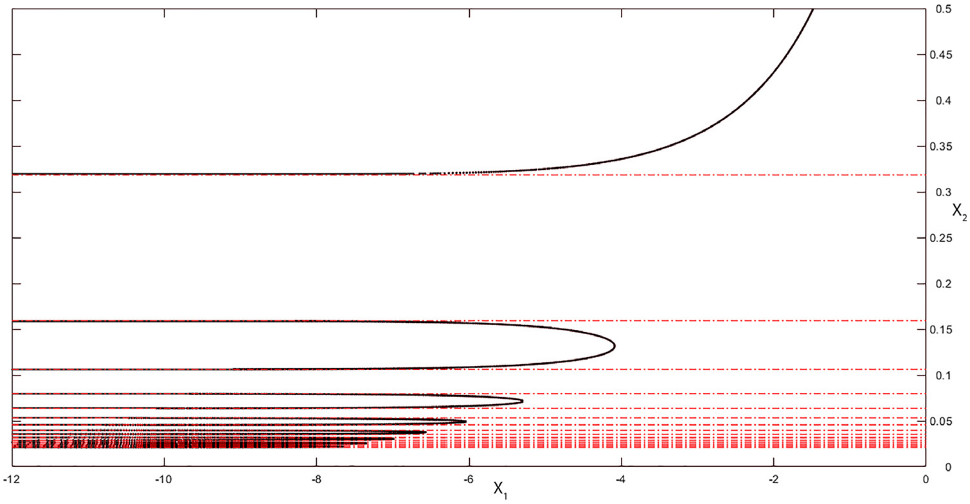

Fig. 1. The zero set $\phi^{-1}(0)$ for the example in (7) with $m = 2$.

$\phi(x_1, x_2) = 0$ and subsets of U_1 for which there is a solution $x_1 = \psi(x_2)$. The preintegrated version of f given by (2) for any smooth θ will rather clearly be differentiable on both U_1 and U_1^+, but it is not obvious that this is the case on the complicated boundary between the two sets. To ensure the necessary differentiability properties it turns out in Section 5 to be necessary to assume that the functions θ and ϕ and their derivatives, when multiplied by the appropriate weight functions, decay sufficiently rapidly as $x_1 = \psi(x_2)$ runs to $-\infty$.

3. Numerical experiments: application to option pricing

In this section we apply the preintegration method to an option pricing example, for which the payoff function is discontinuous.

An important aspect of the method presented in this paper is that the user needs to choose a variable x_j such that the condition (3) is satisfied. In the paper [4] it is shown that for the standard and Brownian bridge constructions for path simulation of Brownian motions every choice of the variable x_j will be suitable. More interesting for the present paper is the popular Principal Component Analysis (or PCA) method of constructing the Brownian motion [13]: for this case the only result known to us, from [14, Section 5], is that the property (3) is guaranteed if x_j is the variable that corresponds to the largest eigenvalue of the Brownian motion covariance matrix. For this reason it is of particular interest to apply the present theory to the PCA case, as we do below.

For our tests, we consider now the example of an arithmetic average digital Asian option. We assume that the underlying asset S_t follows the *geometric Brownian motion* model based on the stochastic differential equation

$$dS_t = rS_t\, dt + \sigma S_t\, dW_t\,, \tag{8}$$

where r is the risk-free interest rate, σ is the (constant) volatility and W_t is the standard Brownian Motion. The solution of this stochastic equation can be given as

$$S_t = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right) t + \sigma W_t\right). \tag{9}$$

The problem of simulating asset prices can be reduced to the problem of simulating discretized Brownian motion paths taking values $W_{t_1}, \ldots, W_{t_d}$, where d is the number of time steps taken in the discretization of the continuous time period $[0, T]$. In our tests, the asset prices are assumed to be sampled at equally spaced times $t_\ell := \ell\Delta t$, $\ell = 1, \ldots, d$, where $\Delta t := T/d$. The Brownian motion is a Gaussian process, therefore the vector $(W_{t_1}, \ldots, W_{t_d})$ is normally distributed, and in this particular case is a vector with mean zero and covariance matrix C given by

$$C = (\min(t_\ell, t_k))_{\ell,k=1}^d.$$

The value of an arithmetic average digital Asian call option is

$$V = \frac{e^{-rT}}{(2\pi)^{d/2}\sqrt{\det(C)}} \int_{\mathbb{R}^d} \mathrm{ind}\left(\frac{1}{d}\sum_{\ell=1}^{d} S_{t_\ell}(w_\ell) - K\right) e^{-\frac{1}{2}\boldsymbol{w}^\top C^{-1}\boldsymbol{w}}\, d\boldsymbol{w}\,,$$

with $\boldsymbol{w} = (W_{t_1}, \ldots, W_{t_d})^\top$. After a factorization $C = AA^\top$ of the covariance matrix is chosen (for the choice of A is not unique), we can rewrite the integration problem using the substitution $\boldsymbol{w} = A\boldsymbol{x}$ as

$$V = \frac{e^{-rT}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{ind}\left(\frac{1}{d} \sum_{\ell=1}^{d} S_{t_\ell}((A\boldsymbol{x})_\ell) - K \right) e^{-\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{x}} \,\mathrm{d}\boldsymbol{x}\,.$$

The new variable vector $\boldsymbol{x} = (x_1, \ldots, x_d)^\top$ can be assumed to consist of independent standard normally distributed random variables. Then the identity $\boldsymbol{w} = A\boldsymbol{x}$ defines a construction method for Brownian paths. We therefore have an integral of the form (1)–(2) with $\rho(x) = e^{-x^2}/\sqrt{2\pi}$, $\theta(\boldsymbol{x}) = e^{-rT}$, and

$$\phi(\boldsymbol{x}) = \frac{S_0}{d} \sum_{\ell=1}^{d} \exp\left(\left(r - \tfrac{\sigma^2}{2} \right) \ell \Delta t + \sigma \sum_{k=1}^{d} A_{\ell k}\, x_k \right) - K. \tag{10}$$

We use in our experiments the PCA factorization of C, which is based on the orthogonal factorization

$$C = (\boldsymbol{u}_1; \ldots; \boldsymbol{u}_d)\, \mathrm{diag}(\lambda_1, \ldots, \lambda_d)\, (\boldsymbol{u}_1; \ldots; \boldsymbol{u}_d)^\top,$$

where the eigenvalues $\lambda_1, \ldots, \lambda_d$ (all positive) are given in non-increasing order, with corresponding unit-length column eigenvectors $\boldsymbol{u}_1 \ldots, \boldsymbol{u}_d$, and as a result

$$A = (\sqrt{\lambda_1}\boldsymbol{u}_1; \ldots; \sqrt{\lambda_d}\boldsymbol{u}_d).$$

Note that we have $A_{\ell 1} > 0$ for $1 \le \ell \le d$ because the elements of the eigenvector $\boldsymbol{u}_1$ are all positive.

For approximate integration with quadrature, we generate randomized QMC or MC samples $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}$ over $\mathbb{R}^d$ by first generating classical randomized QMC or MC samples over unit cube $(0, 1)^d$, and then transforming them to $\mathbb{R}^d$ using in each coordinate the univariate inverse normal cumulative distribution function $\Phi^{-1}(\cdot)$. The randomized QMC points over $(0, 1)^d$ are obtained by first generating Sobol′ points over $[0, 1]^d$ with direction numbers taken from [15], and then applying the random linear-affine scrambling method as proposed by Matousek [16] (as implemented in the statistics toolbox of MATLAB). Note that taking randomly scrambled Sobol′ points not only allows us to generate statistically independent QMC samples, but also allows us to avoid in practice having points lying on the boundary of $(0, 1)^d$ (which is usually the case for non-randomized QMC points), since the boundary is sampled with zero probability. The MC points were taken from the Mersenne Twister PRNG. For the function $\Phi^{-1}(\cdot)$, we have used Moro's algorithm [13]. The matrix A can be given explicitly [14], but more importantly, each matrix–vector multiplication $A\boldsymbol{x}^{(i)}$, $1 \le i \le N$ can be done with $O(d \log d)$ cost by means of the fast-sine transform [17] (as long as time steps for discretization are taken of equal size).

For the preintegration approach, we generate randomized QMC or MC points over $[0, 1]^{d-1}$, following the procedure for the d-dimensional case, and so obtain N sample points over $\mathbb{R}^{d-1}$. We then evaluate the paths without using the first variable x_1, i.e., we sample over the coordinates $x_2, \ldots, x_d$. Once a sample point on these coordinates is fixed, the resulting problem is a one-dimensional integral on the variable x_1. We then take the approximation

$$V \;\approx\; Q_{N,d-1}\left(P_1(f)\right) = P_1\left(Q_{N,d-1}(f)\right) = \frac{1}{N} \sum_{i=1}^{N} \int_{-\infty}^{\infty} f(\xi, \boldsymbol{x}^{(i)})\, \rho(\xi)\, \mathrm{d}\xi\,,$$

where a high-precision quadrature rule such as Gauss quadrature should be applied to the univariate integral with respect to ξ for each of the N sample points $\boldsymbol{x}^{(i)}$ in $\mathbb{R}^{d-1}$. In the PCA case in this problem, the resulting N univariate integrals can be calculated in terms of the normal cumulative distribution function by completing squares and identifying the points $\xi_\star^{(i)}$, $1 \le i \le N$ (if they exist), where we have $\frac{1}{d}\sum_{\ell=1}^{d} S_{t_\ell}((A(\xi_\star^{(i)}, \boldsymbol{x}^{(i)})^\top)_\ell) = K$. Finding the points $\xi_\star^{(i)}$, $1 \le i \le N$, is not a difficult numerical task since each of them can be obtained as the root of an equation defined by a univariate convex function, for which Newton's method converges in few steps to a satisfactorily accurate solution. The computational cost of univariate root-finding and quadrature, though relatively low compared to the $(d-1)$-dimensional quadrature, scales like N.

The parameters in our tests were fixed to $K = 100$, $S_0 = 100$, $r = 0.1$, $\sigma = 0.1$, $T = 1$. We summarize our numerical experiments in Fig. 2. In the figure we show the box-plots of the $\log_{10}$ of relative root mean square error (RMSE), obtained from 10 independent blocks of 10 random replications, with PCA factorization of covariance matrix for the arithmetic average digital Asian option. We took the reference solution to be the empirical mean of the 100 replications. Results are shown in four groups containing three box-plots each. Each group corresponds to one of the following method: in order, MC, QMC, MC with preintegration and QMC with preintegration. In each group we have three box-plots to characterize the error convergence, each box-plot containing RMSE sampled with a given sample size. For all integration methods we chose the sample sizes $N = 2^{12}, 2^{14}, 2^{16}$. Note that for MC and QMC we generate samples over $\mathbb{R}^d$, while for the preintegration MC and QMC we generate samples only over $\mathbb{R}^{d-1}$. For the cases considered in our computations, we observe that the amount of time taken with preintegration is roughly two to three times that without preintegration.

The results show that randomized QMC exhibits higher convergence than MC, but the convergence rate is still not optimal $(\sim N^{-0.6})$. When we combine the preintegration method with MC, we observe an improvement in the implied error constant, as predicted, but the convergence rate remains the same as MC $(= N^{-0.5})$, as of course it should. Combining the preintegration method with randomized QMC reduces the error satisfactorily, and improves the convergence rate to close to the best possible rate $N^{-1.0}$.

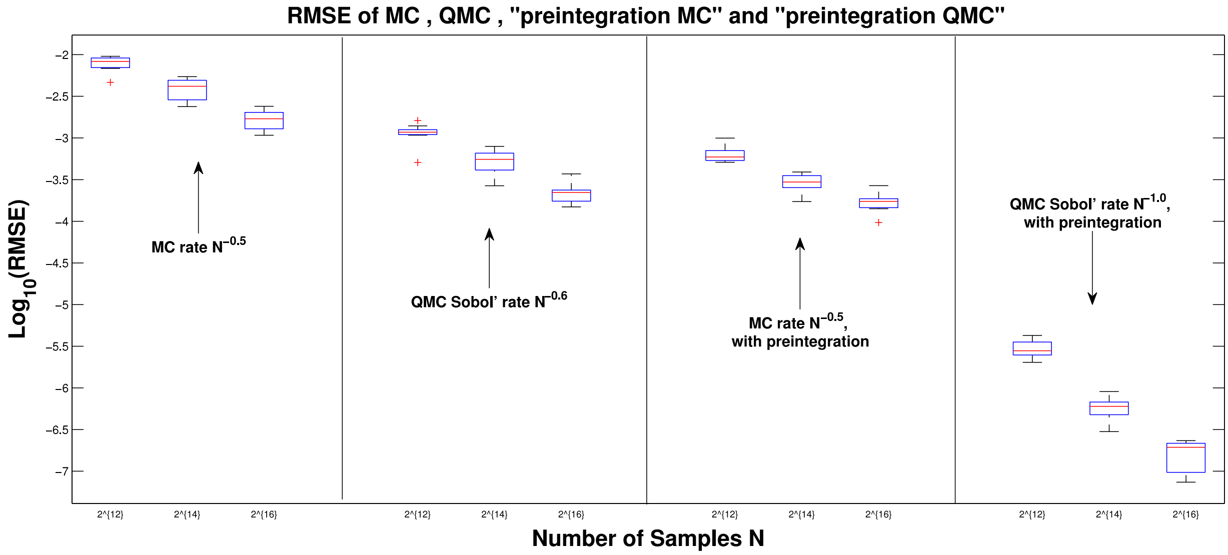

Fig. 2. Root mean square errors for (from left to right) Monte Carlo, Quasi Monte Carlo, preintegrated Monte Carlo and preintegrated Quasi Monte Carlo.

4. Variance reduction by preintegration

In this section we consider the space $\mathcal{L}_{2,\rho_d}$ of square-integrable functions on $\mathbb{R}^d$, with ρ_d-weighted $\mathcal{L}_2$ inner product and norm.

The preintegration step (6) can be viewed more generally as a projection, which is the key operation underlying the well-known ANOVA decomposition. For a general function $g \in \mathcal{L}_{2,\rho_d}$ the ANOVA decomposition takes the form [18]

$$g = \sum_{\mathfrak{u} \subseteq \mathfrak{D}} g_{\mathfrak{u}}, \tag{11}$$

where the sum is over all the 2^d subsets of $\mathfrak{D} := \{1, \ldots, d\}$, and each term $g_{\mathfrak{u}}$ depends only on the variables x_k with $k \in \mathfrak{u}$, and with the additional property that the projection operator P_k defined by (as in (6))

$$(P_k g)(\boldsymbol{x}_{-k}) := \int_{-\infty}^{\infty} g(x_k, \boldsymbol{x}_{-k})\, \rho(x_k)\, \mathrm{d}x_k$$

annihilates all ANOVA terms $g_{\mathfrak{u}}$ with $k \in \mathfrak{u}$:

$$P_k g_{\mathfrak{u}} = 0 \quad \text{for} \quad k \in \mathfrak{u}, \quad \text{whereas} \quad P_k g_{\mathfrak{u}} = g_{\mathfrak{u}} \quad \text{for} \quad k \notin \mathfrak{u}. \tag{12}$$

The ANOVA terms can be written explicitly as [19]

$$g_{\mathfrak{u}} = \sum_{\mathfrak{v} \subseteq \mathfrak{u}} (-1)^{|\mathfrak{u}|-|\mathfrak{v}|} \Big(\prod_{k \notin \mathfrak{v}} P_k \Big) g.$$

It follows from (12), since I_d involves integration with respect to every variable x_k for $k \in \mathfrak{D}$, that

$$I_d g = g_{\emptyset}.$$

Another consequence is that the ANOVA terms are orthogonal in $\mathcal{L}_{2,\rho_d}$,

$$\int_{\mathbb{R}^d} g_{\mathfrak{u}}(\boldsymbol{x})\, g_{\mathfrak{v}}(\boldsymbol{x})\, \rho_d(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = 0 \quad \text{for} \quad \mathfrak{u} \neq \mathfrak{v}.$$

As a result, the variance of g has the well known property that it is a sum of the variances of the separate ANOVA terms,

$$\sigma^2(g) = \sum_{\emptyset \neq \mathfrak{u} \subseteq \mathfrak{D}} \sigma^2(g_{\mathfrak{u}}), \tag{13}$$

where

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

With these preparations, we are now ready to make a simple observation that preintegration is a variance reduction strategy for any general $\mathcal{L}_2$ function, not specific to our functions with kinks or jumps. This explains why the preintegration strategy improves the performance of MC methods.

Lemma 1. *The projection P_k reduces the variance of g for all $g \in \mathcal{L}_{2,\rho_d}$ and all $k \in \mathfrak{D}$.*

Proof. For any $g \in \mathcal{L}_{2,\rho_d}$ and any $k \in \mathfrak{D}$, it follows from (11) and (12) that

$$P_k g = \sum_{k \notin \mathfrak{u} \subseteq \mathfrak{D}} g_{\mathfrak{u}},$$

that is, the operation P_k applied to g has the effect of annihilating those ANOVA terms $g_{\mathfrak{u}}$ of g with $k \in \mathfrak{u}$. As a result, the ANOVA terms of the resulting function $P_k g$ are precisely the ANOVA terms $g_{\mathfrak{u}}$ of g for which $k \notin \mathfrak{u}$. Hence we have

$$\sigma^2(P_k g) = \sum_{k \notin \mathfrak{u} \subseteq \mathfrak{D},\, \mathfrak{u} \neq \emptyset} \sigma^2(g_{\mathfrak{u}}). \tag{14}$$

The result follows by comparing (14) with (13). □

5. Smoothing by preintegration

In this section we first slightly generalize the mathematical setting from [4], providing some details on Sobolev spaces and weak derivatives which are needed for the formulation of our main smoothing theorems. Then we establish two new smoothing theorems for these Sobolev spaces, extending [5, Theorem 1] from kinks to jumps.

5.1. Sobolev spaces with generalized weight functions

Following [4, Section 2.2], for $j \in \mathfrak{D}$ let D_j denote the partial derivative operator

$$(D_j g)(\boldsymbol{x}) = \frac{\partial g}{\partial x_j}(\boldsymbol{x}).$$

Throughout this paper, the term *multi-index* refers to a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ whose components are nonnegative integers, and we use the notation $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_d$ to denote the sum of its components. For any multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$, we define

$$D^{\boldsymbol{\alpha}} = \prod_{j=1}^{d} D_j^{\alpha_j} = \prod_{j=1}^{d} \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j} = \frac{\partial^{|\boldsymbol{\alpha}|}}{\prod_{j=1}^{d} \partial x_j^{\alpha_j}}, \tag{15}$$

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 $g(\boldsymbol{x}) = \exp(\sum_{j=1}^d x_j^2)$ belongs to $\mathcal{C}^r(\mathbb{R}^d)$ for all values of r. For convenience we write $\mathcal{C}^\infty(\mathbb{R}^d) = \cap_{r\geq 0}\mathcal{C}^r(\mathbb{R}^d)$.

In addition to classical derivatives, we shall need also *weak* derivatives. By definition, the weak derivative $D^{\boldsymbol{\alpha}} g$ is a measurable function on $\mathbb{R}^d$ which satisfies

$$\int_{\mathbb{R}^d} (D^{\boldsymbol{\alpha}} g)(\boldsymbol{x})\, v(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = (-1)^{|\boldsymbol{\alpha}|} \int_{\mathbb{R}^d} g(\boldsymbol{x})\, (D^{\boldsymbol{\alpha}} v)(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \quad \text{for all} \quad v \in \mathcal{C}_0^\infty(\mathbb{R}^d), \tag{16}$$

where $\mathcal{C}_0^\infty(\mathbb{R}^d)$ denotes the space of infinitely differentiable functions with compact support in $\mathbb{R}^d$, and where the derivatives on the right-hand side of (16) are classical partial derivatives. It can be shown, using the definition (16), that $D_j D_k = D_k D_j$ for all $j, k \in \mathfrak{D}$, thus the ordering of the weak first derivatives that make up $D^{\boldsymbol{\alpha}}$ in (15) is irrelevant.

If g has classical continuous derivatives up to order $|\boldsymbol{\alpha}|$, then they satisfy (16), which in the classical sense is just the integration by parts formula on $\mathbb{R}^d$. Unless stated otherwise, the derivatives in this paper are weak derivatives, which in principle allows the possibility that they are defined only "almost everywhere". However, a recurring theme is that our weak derivatives are shown to be continuous (or strictly, can be represented by continuous functions), in which case the weak derivatives are at the same time classical derivatives.

We now turn to the definition of the function spaces. For $p \in [1, \infty]$, we first define weighted $\mathcal{L}_p$ norms:

$$\|g\|_{\mathcal{L}_{p,\widetilde{\rho}_d}} = \begin{cases} \left(\displaystyle\int_{\mathbb{R}^d} |g(\boldsymbol{x})|^p\, \widetilde{\rho}_d(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}\right)^{1/p} & \text{if } p \in [1, \infty), \\ \operatorname{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^d} |g(\boldsymbol{x})| & \text{if } p = \infty \end{cases}$$

where $\widetilde{\rho}_d$ is a positive integrable function on $\mathbb{R}^d$.

When dealing with function spaces of derivatives of a function g, it turns out to be convenient to allow flexibility in the choice of weight function $\widetilde{\rho}_d$. We therefore generalize the setting in [4] and introduce a family $\zeta_{d,\boldsymbol{\alpha}}$ of such weight functions, one for each derivative $D^{\boldsymbol{\alpha}}$, given by

$$\zeta_{d,\boldsymbol{\alpha}}(\boldsymbol{x}) := \prod_{k=1}^{d} \zeta_{\alpha_k}(x_k), \tag{17}$$

where $\{\zeta_i\}_{i\geq 0}$ is a sequence of continuous integrable functions on $\mathbb{R}$, satisfying

$$\rho(x) = \zeta_0(x) \leq \zeta_1(x) \leq \zeta_2(x) \leq \dots \qquad \text{for all} \quad x \in \mathbb{R}. \tag{18}$$

The intuitive idea is that higher derivatives with respect to every coordinate need to be limited in their growth towards infinity by making ζ_i decay more slowly for larger order of derivatives i.

With these generalized weight functions $\zeta_{d,\boldsymbol{\alpha}}$, denoted collectively by $\boldsymbol{\zeta}$, we consider two kinds of Sobolev space: the *isotropic Sobolev space* with smoothness parameter $r \geq 0$, for r a nonnegative integer,

$$\mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} = \left\{ g \, : \, D^{\boldsymbol{\alpha}} g \in \mathcal{L}_{p,\zeta_{d,\boldsymbol{\alpha}}} \quad \text{for all} \quad |\boldsymbol{\alpha}| \leq r \right\},$$

and the *Sobolev space of dominating mixed smoothness* with smoothness multi-index $\boldsymbol{r} = (r_1, \dots, r_d)$,

$$\mathcal{W}^{\boldsymbol{r}}_{d,p,\boldsymbol{\zeta},\mathrm{mix}} = \left\{ g \, : \, D^{\boldsymbol{\alpha}} g \in \mathcal{L}_{p,\zeta_{d,\boldsymbol{\alpha}}} \quad \text{for all} \quad \boldsymbol{\alpha} \leq \boldsymbol{r} \right\},$$

where $\boldsymbol{\alpha} \leq \boldsymbol{r}$ is to be understood componentwise, and the derivatives are weak derivatives. For convenience we also write $\mathcal{W}^{0}_{d,p,\boldsymbol{\zeta}} = \mathcal{L}_{p,\rho_d}$ and $\mathcal{W}^{\infty}_{d,p,\boldsymbol{\zeta}} = \cap_{r\geq 0} \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}}$. Analogously, we define $\mathcal{C}^{\boldsymbol{r}}_{\mathrm{mix}}(\mathbb{R}^d)$ to be the space of functions g whose classical derivatives $D^{\boldsymbol{\alpha}} g$ with $\boldsymbol{\alpha} \leq \boldsymbol{r}$ are all continuous at every point in $\mathbb{R}^d$, with no limitation on their behavior at infinity.

The norms corresponding to the two kinds of Sobolev space can be defined, for example, as in the classical sense, by

$$\|g\|_{\mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}}} = \left(\sum_{|\boldsymbol{\alpha}|\leq r} \|D^{\boldsymbol{\alpha}} g\|^2_{\mathcal{L}_{p,\zeta_{d,\boldsymbol{\alpha}}}} \right)^{1/2} \quad \text{and} \quad \|g\|_{\mathcal{W}^{\boldsymbol{r}}_{d,p,\boldsymbol{\zeta},\mathrm{mix}}} = \left(\sum_{\boldsymbol{\alpha}\leq \boldsymbol{r}} \|D^{\boldsymbol{\alpha}} g\|^2_{\mathcal{L}_{p,\zeta_{d,\boldsymbol{\alpha}}}} \right)^{1/2}.$$

We have the following relationships between the spaces:

(i) $\mathcal{W}^{r}_{d,p',\boldsymbol{\zeta}} \subseteq \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}}$ and $\mathcal{W}^{\boldsymbol{r}}_{d,p',\boldsymbol{\zeta},\mathrm{mix}} \subseteq \mathcal{W}^{\boldsymbol{r}}_{d,p,\boldsymbol{\zeta},\mathrm{mix}}$ for $1 \leq p \leq p' \leq \infty$.
(ii) $\mathcal{W}^{\boldsymbol{r}}_{d,p,\boldsymbol{\zeta},\mathrm{mix}} \subseteq \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} \iff \min_{j\in\mathfrak{D}} r_j \geq r$ and $\mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} = \cap_{|\boldsymbol{r}|=r} \mathcal{W}^{\boldsymbol{r}}_{d,p,\boldsymbol{\zeta},\mathrm{mix}}$.
(iii) $\mathcal{W}^{s,\dots,s}_{d,p,\boldsymbol{\zeta},\mathrm{mix}} \subseteq \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} \iff s \geq r$ and $\mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} \subseteq \mathcal{W}^{s,\dots,s}_{d,p,\boldsymbol{\zeta},\mathrm{mix}} \iff r \geq s\,d$.
(iv) $\mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} \subseteq \mathcal{C}^{k}(\mathbb{R}^d)$ if $r > k + d/p$ (Sobolev embedding theorem).
(v) For $p \in [1,\infty)$ and $r \geq 1$, if $g \in \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}}$ then $D^{\boldsymbol{\alpha}} g \in \mathcal{W}^{r-|\boldsymbol{\alpha}|}_{d,p,\boldsymbol{\zeta}}$ for all $|\boldsymbol{\alpha}| \leq r$.
(vi) For $p \in [1,\infty)$ and $\boldsymbol{r} \geq \boldsymbol{1}$, if $g \in \mathcal{W}^{\boldsymbol{r}}_{d,p,\boldsymbol{\zeta}}$ then $D^{\boldsymbol{\alpha}} g \in \mathcal{W}^{\boldsymbol{r}-\boldsymbol{\alpha}}_{d,p,\boldsymbol{\zeta}}$ for all $\boldsymbol{\alpha} \leq \boldsymbol{r}$.

Properties (i)–(iv) are straightforward. Properties (v) and (vi) are a bit more involved due to the varying generalized weight functions considered here. Indeed, when $\overline{\boldsymbol{\alpha}}$ is a multi-index satisfying $0 \leq |\overline{\boldsymbol{\alpha}}| \leq r - |\boldsymbol{\alpha}|$, we have

$$\begin{aligned} \|D^{\overline{\boldsymbol{\alpha}}}(D^{\boldsymbol{\alpha}} g)\|_{\mathcal{L}_{p,\zeta_{d,\overline{\boldsymbol{\alpha}}}}} &= \left(\int_{\mathbb{R}^d} |(D^{\overline{\boldsymbol{\alpha}}}(D^{\boldsymbol{\alpha}} g))(\boldsymbol{x})|^p \, \zeta_{d,\overline{\boldsymbol{\alpha}}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right)^{1/p} \\ &\leq \left(\int_{\mathbb{R}^d} |(D^{\widehat{\boldsymbol{\alpha}}} g)(\boldsymbol{x})|^p \, \zeta_{d,\widehat{\boldsymbol{\alpha}}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right)^{1/p} = \|D^{\widehat{\boldsymbol{\alpha}}} g\|_{\mathcal{L}_{p,\zeta_{d,\widehat{\boldsymbol{\alpha}}}}} < \infty, \end{aligned}$$

where we introduced $\widehat{\boldsymbol{\alpha}} := \overline{\boldsymbol{\alpha}} + \boldsymbol{\alpha}$ and we used $\zeta_{d,\overline{\boldsymbol{\alpha}}}(\boldsymbol{x}) \leq \zeta_{d,\widehat{\boldsymbol{\alpha}}}(\boldsymbol{x})$ since $\overline{\boldsymbol{\alpha}} \leq \widehat{\boldsymbol{\alpha}}$. The finiteness in the final step follows from $g \in \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}}$ and $|\widehat{\boldsymbol{\alpha}}| \leq r$. This justifies (v). The argument can easily be modified to justify (vi).

5.2. New smoothing theorems

In this subsection we establish two smoothing theorems: one for the isotropic Sobolev space, the other for the mixed Sobolev space. The proofs are modeled on the proof of [5, Theorem 1], but are extended here to cover discontinuous integrands.

Theorem 2 (*Result for the Isotropic Sobolev Space with Weight Functions $\zeta_{d,\boldsymbol{\alpha}}$*). *Let $d \geq 2$, $r \geq 1$, $p \in [1,\infty)$, and let $\rho \in \mathcal{C}^{r-1}(\mathbb{R})$ be a strictly positive probability density function. Let*

$$f(\boldsymbol{x}) := \theta(\boldsymbol{x}) \, \mathrm{ind}(\phi(\boldsymbol{x})), \quad \textit{where} \quad \theta, \phi \in \mathcal{W}^{r}_{d,p,\boldsymbol{\zeta}} \cap \mathcal{C}^{r}(\mathbb{R}^d),$$

with generalized weight functions $\zeta_{d,\boldsymbol{\alpha}}$ satisfying (17) and (18), and with $\mathrm{ind}(\cdot)$ denoting the indicator function. Let $j \in \mathfrak{D} := \{1, \dots, d\}$ be fixed, and suppose that

$$(D_j\phi)(\boldsymbol{x}) > 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^d, \qquad \textit{and} \qquad \phi(\boldsymbol{x}) \to \infty \quad \textit{as} \quad x_j \to \infty \quad \forall \boldsymbol{x}_{-j} \in \mathbb{R}^{d-1}. \tag{19}$$

Writing $\boldsymbol{y} := \boldsymbol{x}_{-j}$ *so that* $\boldsymbol{x} = (x_j, \boldsymbol{y})$*, let*

$$U_j := \{\boldsymbol{y} \in \mathbb{R}^{d-1} : \phi(x_j, \boldsymbol{y}) = 0 \text{ for some } x_j \in \mathbb{R}\} \qquad \text{and} \qquad U_j^+ := \mathbb{R}^{d-1} \setminus U_j.$$

If U_j *is empty, then* $f = \theta$*. If* U_j *is not empty, then* U_j *is open, and there exists a unique function* $\psi \equiv \psi_j \in \mathcal{C}^r(U_j)$ *such that* $\phi(x_j, \boldsymbol{y}) = 0$ *if and only if* $x_j = \psi(\boldsymbol{y})$ *for* $\boldsymbol{y} \in U_j$*. In the latter case we assume that every function of the form*

$$\begin{cases} h(\boldsymbol{y}) = \dfrac{(D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y}) \prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^b} \rho^{(c)}(\psi(\boldsymbol{y})), \quad \boldsymbol{y} \in U_j, \\ \text{where } a, b, c \text{ are integers and } \boldsymbol{\gamma}^{(i)}, \boldsymbol{\eta} \text{ are multi-indices with the constraints} \\ 1 \le a, b \le 2r - 1, \quad 1 \le |\boldsymbol{\gamma}^{(i)}| \le r, \quad 0 \le |\boldsymbol{\eta}|, c \le r - 1, \quad 1 \le |\boldsymbol{\gamma}^{(i)}| + |\boldsymbol{\eta}| + c \le r, \end{cases} \tag{20}$$

satisfies both

$$h(\boldsymbol{y}) \to 0 \qquad \text{as} \quad \boldsymbol{y} \text{ approaches a boundary point of } U_j \text{ lying in } U_j^+, \tag{21}$$

and

$$\int_{U_j} |h(\boldsymbol{y})|^p \, \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} < \infty, \qquad \text{for all } |\boldsymbol{\alpha}| \le r \text{ with } \alpha_j = 0, \tag{22}$$

where $\boldsymbol{\alpha}_{-j}$ *denotes the multi-index with* $d-1$ *components obtained from* $\boldsymbol{\alpha}$ *by leaving out* α_j*. Then*

$$P_j f \in \mathcal{W}^r_{d-1,p,\zeta} \cap \mathcal{C}^r(\mathbb{R}^{d-1}).$$

Proof. We defer the proof of this theorem to Section 7. □

In effect, under the conditions in the theorem, the single integration with respect to x_j is sufficient to ensure that $P_j f$ inherits the full smoothness of θ and ϕ.

We remark that when $\theta = \phi$ we are back at the same function $f(\boldsymbol{x}) = \max(\phi(\boldsymbol{x}, 0))$ as considered in [5, Theorem 1]. However, for this case we see that the new result is not as sharp as the old one in the sense that the upper bounds on the values of $a, b, c, |\boldsymbol{\gamma}^{(i)}|$ in the condition (20) are larger than those in [5, Theorem 1]. This is because the explicit prior knowledge of $\theta = \phi$ means that we know a certain term vanishes (precisely, the second term on the right-hand side of [5, Formula (11)]). This observation also indirectly explains how the new result for jumps requires stronger conditions on the functions θ, ϕ and ρ than the corresponding result for kinks.

The conditions (21) and (22) in the theorem are difficult to verify directly because the function h depends explicitly on the inverse function $\psi(\boldsymbol{y})$. Fortunately, a sufficient condition for (21) to hold is that

$$\left| \frac{(D^{\boldsymbol{\eta}}\theta)(x_j, \boldsymbol{y}) \prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(x_j, \boldsymbol{y})]}{[(D_j\phi)(x_j, \boldsymbol{y})]^b} \rho^{(c)}(x_j) \right| \le E_1(x_j) E_2(\boldsymbol{y}), \tag{23}$$

where E_1, E_2 are positive functions satisfying

- E_1 is bounded and $E_1(x_j) \to 0$ as $x_j \to -\infty$,
- E_2 is locally bounded (bounded over compact sets) and $\int_{\mathbb{R}^{d-1}} |E_2(\boldsymbol{y})|^p \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} < \infty$ for all $|\boldsymbol{\alpha}| \le r$.

Considering a point $\boldsymbol{y}^\star$ on the boundary $\Gamma(U_j) \subset U_j^+$, and a sequence $(\boldsymbol{y}_n)_{n\in\mathbb{N}} \subset U_j$ such that $\boldsymbol{y}_n \to \boldsymbol{y}^\star$, we see that $E_2(\boldsymbol{y}_n)$ is bounded and $E_1(\psi(\boldsymbol{y}_n)) \to 0$ since $\psi(\boldsymbol{y}_n) \to -\infty$. Therefore (23) is sufficient for (21). Moreover, for $|\boldsymbol{\alpha}| \le r$ we have

$$\begin{aligned} &\int_{U_j} \left| \frac{(D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y}) \prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^b} \rho^{(c)}(\psi(\boldsymbol{y})) \right|^p \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \\ &\le \int_{U_j} |E_1(\psi(\boldsymbol{y}))|^p |E_2(\boldsymbol{y})|^p \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \le B \int_{U_j} |E_2(\boldsymbol{y})|^p \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} < \infty, \end{aligned}$$

for some positive constant B. Therefore (23) is also sufficient for (22).

We can also deduce a result for Sobolev spaces of dominating mixed smoothness.

Theorem 3 (*Result for the Sobolev Space of Dominating Mixed Smoothness*)**.** *Let* $d \ge 2$*,* $p \in [1, \infty)$*,* $j \in \mathfrak{D}$*, and let* $\rho \in \mathcal{C}^{r-1}(\mathbb{R})$ *be a strictly positive probability density function. Let* $\boldsymbol{r} = (r_1, \ldots, r_d)$ *be a multi-index satisfying*

$$r_j \ge \sum_{1 \le k \le d,\, k \ne j} r_k \ge 1.$$

If we replace the conditions on θ, ϕ *in Theorem 2 by*

$$\theta, \phi \in \mathcal{W}^{\boldsymbol{r}}_{d,p,\zeta,\mathrm{mix}} \cap \mathcal{C}^{\boldsymbol{r}}_{\mathrm{mix}}(\mathbb{R}^d),$$

and further restrict (20)–(22) to functions h with multi-indices $\boldsymbol{\gamma}^{(i)} \le \boldsymbol{r}$, $\boldsymbol{\eta} < \boldsymbol{r}$, *and* $\boldsymbol{\alpha} \le \boldsymbol{r}$, *then the conclusion becomes:*

$$P_j f \in \mathcal{W}^{\boldsymbol{r}_{-j}}_{d-1,p,\zeta} \cap \mathcal{C}^{\boldsymbol{r}_{-j}}_{\mathrm{mix}}(\mathbb{R}^{d-1}).$$

Proof. The proof is obtained from minor modifications of the proof of Theorem 2 in Section 7. In particular, the requirement that r_j is greater than or equal to the sum of the remaining r_k for $k \neq j$ is needed because, for any multi-index $\boldsymbol{\alpha} \le \boldsymbol{r}$ with $\alpha_j = 0$, it is clear from (34) and a generalization of (35) that the expression for $D^{\boldsymbol{\alpha}} P_j f$ includes some terms that depend on $D_j^{|\boldsymbol{\alpha}|}\phi$ and some terms that depend on $D_j^{|\boldsymbol{\alpha}|-1}\theta$. □

6. Applying the theory to option pricing

We now apply our results to the option pricing example.

Recall from Section 3 that after PCA factorization the function f from the digital option pricing example takes the form (2), with θ a constant function and ϕ given by (10). It follows that

$$(D_j\phi)(\boldsymbol{x}) = \frac{\sigma S_0}{d} \sum_{\ell=1}^{d} \exp\left(\left(\mu - \tfrac{\sigma^2}{2}\right)\ell \Delta t + \sigma \sum_{k=1}^{d} A_{\ell k} x_k\right) A_{\ell j}.$$

In particular, we see that $(D_1\phi)(\boldsymbol{x}) > 0$ because, as explained in Section 3, $A_{\ell 1} > 0$ for all ℓ, thus in this case it is appropriate to take $j = 1$ in Theorem 2. It is also clear that ϕ is in $\mathcal{C}^r(\mathbb{R}^d)$ for all $r \in \mathbb{Z}^+$. Additionally, we may take all the weight functions ζ_i in (18) equal to the standard normal density ρ. It is then clear that the sufficient condition (23) is satisfied, and moreover that all the integrability and decay conditions in Theorem 2 are satisfied, because all derivatives of ϕ are "killed" at infinity by the Gaussian weight functions and their derivatives. It then follows from Theorem 2 that

$$P_1 f \in \mathcal{W}^r_{d-1,p,\boldsymbol{\rho}} \cap \mathcal{C}^r(\mathbb{R}^{d-1}) \quad \forall\, r \in \mathbb{Z}^+, \quad \forall\, p \in [1,\infty).$$

7. Proof of the main smoothing theorem

Before we proceed to prove Theorem 2, we quote three theorems from [4, Section 2.4], but state them with respect to the Sobolev spaces defined with generalized weight functions $\zeta_{d,\boldsymbol{\alpha}}$. We outline the subtle additional steps needed in the proofs of [4] to allow for this generalization.

The classical Leibniz theorem allows us to swap the order of differentiation and integration. In this paper we need a more general form of the Leibniz theorem as given below.

Theorem 4 (*The Leibniz Theorem [4, Theorem 2.1]*)**.** *Let* $p \in [1,\infty)$. *For* $g \in \mathcal{W}^1_{d,p,\zeta}$ *with generalized weight functions* $\zeta_{d,\boldsymbol{\alpha}}$ *satisfying (17) and (18), we have*

$$D_k P_j g = P_j D_k g \qquad \text{for all} \quad j,k \in \mathfrak{D} \quad \text{with} \quad j \neq k.$$

Proof. We follow the proof of [4, Theorem 2.1] to the last paragraph where Fubini's theorem was applied a second time. This application of Fubini's Theorem is valid because

$$\begin{aligned} &\left| \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} (D_k g)(t_j, \boldsymbol{x}_{-j})\, \rho(t_j)\, \mathrm{d}t_j\, v(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \right| \\ &\le \int_{\mathbb{R}^{d-1}} \int_{-\infty}^{\infty} |(D_k g)(t_j, \boldsymbol{x}_{-j})|\, \zeta_1(t_j)\, \mathrm{d}t_j\, \rho_{d-1}(\boldsymbol{x}_{-j}) \frac{\int_{-\infty}^{\infty} |v(x_j, \boldsymbol{x}_{-j})|\, \mathrm{d}x_j}{\rho_{d-1}(\boldsymbol{x}_{-j})}\, \mathrm{d}\boldsymbol{x}_{-j} \\ &\le \|D_k g\|_{\mathcal{L}_{1,\zeta_{d,\boldsymbol{e}_k}}} \frac{\sup_{\boldsymbol{x}_{-j}\in V} \int_{-\infty}^{\infty} |v(x_j, \boldsymbol{x}_{-j})|\, \mathrm{d}x_j}{\inf_{\boldsymbol{x}_{-j}\in V} \rho_{d-1}(\boldsymbol{x}_{-j})} < \infty, \end{aligned}$$

where $\boldsymbol{e}_k$ is the multi-index consisting of 1 in the position k, and 0 elsewhere, and where we made use of $\rho(t_j) \le \zeta_1(t_j)$ and $g \in \mathcal{W}^1_{d,1,\zeta}$, and that the set V defined in the proof of [4, Theorem 2.1] is a compact set because of the compactness of $\mathrm{supp}(v)$. The remainder of that proof then stands. □

The next theorem is an application of the Leibniz theorem; it establishes that $P_j f$ inherits the smoothness of g.

Theorem 5 (*The Inheritance Theorem [4, Theorem 2.2]*)**.** *Let* $r \ge 0$ *and* $p \in [1,\infty)$. *For* $g \in \mathcal{W}^r_{d,p,\zeta}$ *with generalized weight functions* $\zeta_{d,\boldsymbol{\alpha}}$ *satisfying (17) and (18), we have*

$$P_j g \in \mathcal{W}^r_{d-1,p,\zeta} \qquad \text{for all} \quad j \in \mathfrak{D}.$$

Proof. For the case $r = 0$ the proof is exactly the same as the proof of [4, Theorem 2.2]. Consider now $r \geq 1$. Let $j \in \mathfrak{D}$ and let $\boldsymbol{\alpha}$ be any multi-index with $|\boldsymbol{\alpha}| \leq r$ and $\alpha_j = 0$. Since now $g \in \mathcal{W}^r_{d,p,\zeta}$ with generalized weight functions, we have $\|D^{\boldsymbol{\alpha}} g\|_{\mathcal{L}_{p,\zeta_{d,\boldsymbol{\alpha}}}} < \infty$. To show that $P_j g \in \mathcal{W}^r_{d-1,p,\zeta}$ we need to show that $\|D^{\boldsymbol{\alpha}} P_j g\|_{\mathcal{L}_{p,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}}} < \infty$. Mimicking the proof of [4, Theorem 2.2], we write successively

$$\begin{aligned} D^{\boldsymbol{\alpha}} P_j g &= \left(\prod_{i=1}^{|\boldsymbol{\alpha}|} D_{k_i}\right) P_j g = \left(\prod_{i=2}^{|\boldsymbol{\alpha}|} D_{k_i}\right) P_j D_{k_1} g \\ &= \ldots = D_{k_{|\boldsymbol{\alpha}|}} P_j \left(\prod_{i=1}^{|\boldsymbol{\alpha}|-1} D_{k_i}\right) g = P_j \left(\prod_{i=1}^{|\boldsymbol{\alpha}|} D_{k_i}\right) g = P_j D^{\boldsymbol{\alpha}} g, \end{aligned} \tag{24}$$

where $k_i \in \mathfrak{D} \setminus \{j\}$ and $k_1, \ldots, k_{|\boldsymbol{\alpha}|}$ need not be distinct. Each step in (24) involves a single differentiation under the integral sign, and is justified by the Leibniz theorem (Theorem 4) because we know from the property (v) in Section 5.1 that $(\prod_{i=1}^{\ell} D_{k_i}) g \in \mathcal{W}^{r-\ell}_{d,p,\zeta} \subseteq \mathcal{W}^1_{d,p,\zeta}$ for all $\ell \leq |\boldsymbol{\alpha}| - 1 \leq r - 1$. We have therefore

$$\begin{aligned} &\|D^{\boldsymbol{\alpha}} P_j g\|_{\mathcal{L}_{p,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}}} = \|P_j D^{\boldsymbol{\alpha}} g\|_{\mathcal{L}_{p,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}}} \\ &= \left(\int_{\mathbb{R}^{\mathfrak{D}\setminus\{j\}}} \left|\int_{-\infty}^{\infty} (D^{\boldsymbol{\alpha}} g)(\boldsymbol{x})\, \rho(x_j)\, \mathrm{d}x_j\right|^p \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{x}_{-j})\, \mathrm{d}\boldsymbol{x}_{-j}\right)^{1/p} \\ &\leq \left(\int_{\mathbb{R}^{\mathfrak{D}\setminus\{j\}}} \left(\int_{-\infty}^{\infty} |(D^{\boldsymbol{\alpha}} g)(\boldsymbol{x})|^p\, \rho(x_j)\, \mathrm{d}x_j\right) \zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{x}_{-j})\, \mathrm{d}\boldsymbol{x}_{-j}\right)^{1/p} = \|D^{\boldsymbol{\alpha}} g\|_{\mathcal{L}_{p,\zeta_{d,\boldsymbol{\alpha}}}} < \infty, \end{aligned}$$

where we applied Hölder's inequality to the inner integral as in [4, Equation (2.11)] and used $\zeta_{\alpha_j}(x_j) = \zeta_0(x_j) = \rho(x_j)$. This completes the proof. □

The implicit function theorem stated below is crucial for the main results of this paper. In the rest of the paper, for any $r \geq 0$, $k \geq 1$, and an open set $U \subset \mathbb{R}^k$, we define $\mathcal{C}^r(U)$ to be the space of functions whose classical derivatives of order $\leq r$ are all continuous at every point in U.

Theorem 6 (*The Implicit Function Theorem [4, Theorem 2.3]*)**.** *Let $j \in \mathfrak{D}$. Suppose $\phi \in \mathcal{C}^1(\mathbb{R}^d)$ satisfies*

$$(D_j\phi)(\boldsymbol{x}) > 0 \qquad \text{for all} \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{25}$$

Let

$$U_j := \{\boldsymbol{x}_{-j} \in \mathbb{R}^{d-1} \,:\, \phi(x_j, \boldsymbol{x}_{-j}) = 0 \text{ for some (unique) } x_j \in \mathbb{R}\}.$$

If U_j is not empty then there exists a unique function $\psi_j \in \mathcal{C}^1(U_j)$ such that

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

and for all $k \neq j$ we have

$$(D_k\psi_j)(\boldsymbol{x}_{\mathfrak{D}\setminus\{j\}}) = -\frac{(D_k\phi)(\boldsymbol{x})}{(D_j\phi)(\boldsymbol{x})}\bigg|_{x_j=\psi_j(\boldsymbol{x}_{\mathfrak{D}\setminus\{j\}})} \qquad \text{for all} \quad \boldsymbol{x}_{\mathfrak{D}\setminus\{j\}} \in U_j. \tag{26}$$

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

Note that the derivatives in the implicit function theorem are classical derivatives, and the condition (25) needs to hold for *all* $\boldsymbol{x} \in \mathbb{R}^d$.

We are almost ready to prove Theorem 2. But first we give a remark and a couple of auxiliary results.

Remark 7. It is easily seen that U_j and U_j^+ in Theorem 2 can also be defined by

$$\begin{aligned} U_j &= \left\{\boldsymbol{y} \in \mathbb{R}^{d-1} : \lim_{x_j\to-\infty} \phi(x_j, \boldsymbol{y}) < 0\right\}, \\ U_j^+ &= \left\{\boldsymbol{y} \in \mathbb{R}^{d-1} : \lim_{x_j\to-\infty} \phi(x_j, \boldsymbol{y}) \geq 0\right\}. \end{aligned}$$

In the proof of the theorem we make essential use of the following lemma. This result is needed to ensure that all the derivatives we encounter are continuous across the boundary between U_j and U_j^+.

Lemma 8. *Under the condition (19), the function $\psi_j : \mathbb{R} \to \mathbb{R}$ has the following property*

$$\psi_j(\boldsymbol{y}) \to -\infty \tag{27}$$

as $\boldsymbol{y}$ approaches a point on the boundary of U_j.

Proof. Consider a point $\boldsymbol{y}^\star$ a point on the boundary of U_j, and hence (because U_j is open) lying in U_j^+. Consider also a sequence $(\boldsymbol{y}_n)_{n\in\mathbb{N}} \subset U_j$ with $\boldsymbol{y}_n \to \boldsymbol{y}^\star$ as $n \to \infty$. We assert that the sequence $(\psi(\boldsymbol{y}_n))_{n\in\mathbb{N}}$ has no accumulation points in $\mathbb{R}$. This is true because if we assume otherwise then there would exist a convergent subsequence $(\psi(\boldsymbol{y}_{n_k}))_{k\in\mathbb{N}}$, with $\psi(\boldsymbol{y}_{n_k}) \to x_j^\star$ as $k \to \infty$ for $x_j^\star \in \mathbb{R}$. But because of the continuity of ϕ we must have

$$\phi(x_j^\star, \boldsymbol{y}^\star) = \lim_{k\to\infty} \phi(\psi(\boldsymbol{y}_{n_k}), \boldsymbol{y}_{n_k}) = 0,$$

since by definition $\phi(\psi(\boldsymbol{y}_{n_k}), \boldsymbol{y}_{n_k}) = 0$, $\forall k \in \mathbb{N}$. This implies that $\boldsymbol{y}^\star \in U_j$, which is a contradiction. Therefore the sequence $(\psi(\boldsymbol{y}_n))_{n\in\mathbb{N}}$ has no accumulation points in $\mathbb{R}$. This implies (due to the Bolzano–Weierstrass Theorem) that $(\psi(\boldsymbol{y}_n))_{n\in\mathbb{N}} \cap [a, b]$ is a finite set, for each interval $[a, b]$, $a, b \in \mathbb{R}$. Thus,

$$\lim_{n\to\infty} |\psi(\boldsymbol{y}_n)| = \infty.$$

To eliminate the possibility that $+\infty$ is an accumulation point of $\psi(\boldsymbol{y}_n)$ we observe that, due to condition (4), and with $\boldsymbol{y}^\star$ as above, there exists an $x_j^\star$ such that $\phi(x_j^\star, \boldsymbol{y}^\star) > 0$. Because of the continuity of ϕ, there is a ball around the point $(x_j^\star, \boldsymbol{y}^\star)$, denoted by $B(x_j^\star, \boldsymbol{y}^\star)$ such that ϕ is positive for each point in $B(x_j^\star, \boldsymbol{y}^\star)$. Assume now that we have a subsequence $(\psi(\boldsymbol{y}_{n_m}))_{m\in\mathbb{N}}$ such that $\lim_{m\to\infty} \psi(\boldsymbol{y}_{n_m}) = +\infty$. Because $\boldsymbol{y}_{n_m}$ converges to $\boldsymbol{y}^\star$ as $m \to \infty$, it follows that $(x_j^\star, \boldsymbol{y}_{n_m}) \in B(x_j^\star, \boldsymbol{y}^\star)$ for all m sufficiently large. But assumption $\psi(\boldsymbol{y}_{n_m}) \to +\infty$, and the monotonicity condition in (19) implies that for all m sufficiently large we have $\psi(\boldsymbol{y}_{n_m}) > x_j^\star$, and therefore $0 < \phi(x_j^\star, \boldsymbol{y}_{n_m}) < \phi(\psi(\boldsymbol{y}_{n_m}), \boldsymbol{y}_{n_m}) = 0$, which is clearly a contradiction. Therefore we conclude that

$$\lim_{n\to\infty} \psi(\boldsymbol{y}_n) = -\infty. \quad \square$$

Another auxiliary result is needed to show that the assumption (21) implies continuous differentiability of $P_j f$ at boundary points of U_j.

Lemma 9. *Let $r \geq 0$ and $k \geq 1$. Suppose $g \in \mathcal{C}^r(U)$ for some open domain $U \subset \mathbb{R}^k$ and $g(\boldsymbol{y}) = 0$ for all $\boldsymbol{y} \in U^c$. Suppose that for any multi-index $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}| \leq r$ and any sequence $(\boldsymbol{y}_n)_{n\in\mathbb{N}} \subset U$ we have*

$$\lim_{n\to\infty} \boldsymbol{y}_n = \boldsymbol{y}^\star \text{ with } \boldsymbol{y}^\star \in U^c \quad \Rightarrow \quad \lim_{n\to\infty} (D^{\boldsymbol{\alpha}} g)(\boldsymbol{y}_n) = 0. \tag{28}$$

Then we have $g \in \mathcal{C}^r(\mathbb{R}^k)$, with $(D^{\boldsymbol{\alpha}} g)(\boldsymbol{y}) = 0$ for all $\boldsymbol{y} \in U^c$.

Proof. The statement is obviously true for $r = 0$ where mere continuity of g is asserted. Now suppose it holds for a natural number $r_0 \geq 0$ and consider any multi-index $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{e}_i$, with $\boldsymbol{e}_i$ denoting a canonical basis vector and $|\boldsymbol{\alpha}_0| = r_0$, and hence $|\boldsymbol{\alpha}| = r_0 + 1 \leq r$. Then we have to show that $(D^{\boldsymbol{\alpha}} g)(\boldsymbol{y})$ exists at all $\boldsymbol{y} \in \mathbb{R}^k$ and is continuous at every point of $\mathbb{R}^k$. For points in U the derivative $D^{\boldsymbol{\alpha}} g$ exists and is continuous by assumption, and for points in the interior of U^c the derivative $D^{\boldsymbol{\alpha}} g$ exists and is continuous because g is zero there. So it remains to consider the existence and continuity of $D^{\boldsymbol{\alpha}} g$ at any boundary point $\boldsymbol{y}^\star$, i.e., at any limit point $\boldsymbol{y}^\star$ of a sequence $(\boldsymbol{y}_n)_{n\in\mathbb{N}} \subset U$.

To show the existence, consider the scalar valued function $D^{\boldsymbol{\alpha}_0} g$ which is by the induction hypothesis continuous on all of $\mathbb{R}^k$ and vanishes at any boundary point $\boldsymbol{y}^\star$. Consider first the case $h > 0$ and a point $\boldsymbol{y}^\star + h\boldsymbol{e}_i$. If $\boldsymbol{y}^\star + h\boldsymbol{e}_i \in U$, then because U is open and $\boldsymbol{y}^\star \in U^c$, we have for

$$\bar{h} := \sup\{h' \ : \ 0 \leq h' < h, \ \boldsymbol{y}^* + h'\boldsymbol{e}_i \in U^c\}$$

that $\boldsymbol{y}^\star + h'\boldsymbol{e}_i \in U$ for all $\bar{h} < h' \leq h$. Furthermore, because U^c is closed it follows that $(D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star + \bar{h}\boldsymbol{e}_i) = 0 = (D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star)$, and thus we conclude from the mean value theorem that

$$(D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star + h\boldsymbol{e}_i) - (D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star) = (D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star + h\boldsymbol{e}_i) - (D^{\boldsymbol{\alpha}_0} g)(y^\star + \bar{h}\boldsymbol{e}_i) = (h - \bar{h})(D_i D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^* + h^\star \boldsymbol{e}_i)$$

for some $h^\star$ satisfying $\bar{h} < h^\star < h$. Hence we have for the quotient

$$\frac{(D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star + h\boldsymbol{e}_i) - (D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star)}{h} = \begin{cases} \dfrac{h - \bar{h}}{h} (D_i D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star + h^\star \boldsymbol{e}_i), \ \ \bar{h} < h^\star < h, & \text{if } \boldsymbol{y}^* + h\boldsymbol{e}_i \in U, \\ 0 & \text{if } \boldsymbol{y}^\star + h\boldsymbol{e}_i \in U^c. \end{cases}$$

Then using the assumption (28), letting h be arbitrarily small, using that $|\bar{h}| \leq |h|$, and considering the analogous situation for $h < 0$, we obtain the existence of $D_i D^{\boldsymbol{\alpha}_0} g$ at $\boldsymbol{y}^\star$, with $(D_i D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}^\star) = 0$.

To show the derivative continuity at a boundary point $\boldsymbol{y}^\star$, consider any sequence $(\boldsymbol{y}_n)_{n\in\mathbb{N}} \subset \mathbb{R}^k$ with $\lim_{n\to\infty} \boldsymbol{y}_n = \boldsymbol{y}^\star$. For a given n either $\boldsymbol{y}_n \in U$, in which case (28) applies, or $\boldsymbol{y}_n \in U^c$, in which case $(D_i D^{\boldsymbol{\alpha}_0} g)(\boldsymbol{y}_n) = 0$, as above, so that both subsequences converge to $0 = (D_i D^{\alpha_0} g)(\boldsymbol{y}^\star)$. Finally, since all partial derivatives of order $r_0 + 1$ are now proved continuous in $\mathbb{R}^k$, those mixed partial derivatives are symmetric, and we can write $D_i D^{\alpha_0} g = D^{\alpha} g$.

Hence $D^{\alpha} g$ exists and is continuous on all of $\mathbb{R}^k$, i.e., the induction step is proved. It follows then that $g \in \mathcal{C}^r(\mathbb{R}^k)$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. We focus on the non-trivial case when U_j is not empty. Given that $\phi \in \mathcal{C}^r(\mathbb{R}^d)$, that $(D_j\phi)(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \mathbb{R}^d$, and that U_j is not empty, it follows from the implicit function theorem (see Theorem 6) that the set U_j is open, and that there exists a unique function $\psi \equiv \psi_j \in \mathcal{C}^r(U_j)$ for which

$$\phi(x_j, \boldsymbol{y}) = 0 \iff \psi(\boldsymbol{y}) = x_j \quad \text{for all } \boldsymbol{y} \in U_j. \tag{29}$$

This justifies the existence of the function ψ as stated in the theorem.

For the function $f(\boldsymbol{x}) = \theta(x_j, \boldsymbol{y})\,\mathrm{ind}(\phi(x_j, \boldsymbol{y}))$ we can write $P_j f$ defined by (6) as

$$(P_j f)(\boldsymbol{y}) = \int_{x_j \in \mathbb{R}\,:\,\phi(x_j,\boldsymbol{y}) \ge 0} \theta(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j. \tag{30}$$

It follows from the condition $(D_j\phi)(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in \mathbb{R}^d$ and the continuity of $D_j\phi$ that, for fixed $\boldsymbol{y}$, $\phi(x_j, \boldsymbol{y})$ is a strictly increasing function of x_j.

We now determine the limits of integration in (30). If $\boldsymbol{y} \in U_j^+$, then $\phi(x_j, \boldsymbol{y}) \neq 0$ for all $x_j \in \mathbb{R}$. Since ϕ is continuous, strictly increasing in x_j, and tends to $+\infty$ as $x_j \to +\infty$, we conclude that $\phi(x_j, \boldsymbol{y}) > 0$ for all $x_j \in \mathbb{R}$, and thus we integrate x_j from $-\infty$ to ∞. On the other hand, if $\boldsymbol{y} \in U_j$, in which case $\phi(x_j, \boldsymbol{y})$ changes sign once as x_j goes from $-\infty$ to ∞, then there exists a unique $x_j^* = \psi(\boldsymbol{y}) \in \mathbb{R}$ for which $\phi(x_j^*, \boldsymbol{y}) = 0$, and in this case we integrate x_j from $\psi(\boldsymbol{y})$ to ∞. Hence we can write (30) as

$$(P_j f)(\boldsymbol{y}) = \begin{cases} \displaystyle\int_{-\infty}^{\infty} \theta(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j & \text{if } \boldsymbol{y} \in U_j^+, \\ \displaystyle\int_{\psi(\boldsymbol{y})}^{\infty} \theta(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j & \text{if } \boldsymbol{y} \in U_j. \end{cases}$$

Note that $P_j f$ is continuous across the boundary between U_j and U_j^+, since from Lemma 8 it follows that $\psi(\boldsymbol{y}) \to -\infty$ as $\boldsymbol{y} \in U_j$ approaches a boundary point of U_j.

By the Leibniz Theorem and the Inheritance Theorem, we know that the function $(P_j\theta)(\boldsymbol{y}) = \int_{-\infty}^{\infty} \theta(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j$ for $\boldsymbol{y} \in \mathbb{R}^{d-1}$ is as smooth as θ, i.e., $P_j\theta \in \mathcal{W}^r_{d-1,p,\xi} \cap \mathcal{C}^r(\mathbb{R}^{d-1})$. Therefore, to obtain the same smoothness property for $P_j f$ it suffices that we consider in the remainder of this proof the difference

$$g(\boldsymbol{y}) := (P_j f)(\boldsymbol{y}) - (P_j\theta)(\boldsymbol{y}) = \begin{cases} \displaystyle -\int_{-\infty}^{\psi(\boldsymbol{y})} \theta(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j & \text{if } \boldsymbol{y} \in U_j, \\ 0 & \text{if } \boldsymbol{y} \in U_j^+. \end{cases} \tag{31}$$

First we differentiate g with respect to the kth coordinate for any $k \neq j$. For $\boldsymbol{y} \in U_j$ we obtain, using the fundamental theorem of calculus,

$$\begin{aligned} (D_k g)(\boldsymbol{y}) &= -\int_{-\infty}^{\psi(\boldsymbol{y})} (D_k\theta)(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j - \theta(\psi(\boldsymbol{y}), \boldsymbol{y})\,\rho(\psi(\boldsymbol{y}))\,(D_k\psi)(\boldsymbol{y}) \\ &= -\int_{-\infty}^{\psi(\boldsymbol{y})} (D_k\theta)(x_j, \boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j + \theta(\psi(\boldsymbol{y}), \boldsymbol{y})\,\frac{(D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}{(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}\,\rho(\psi(\boldsymbol{y})), \end{aligned} \tag{32}$$

where we substituted using (26)

$$(D_k\psi)(\boldsymbol{y}) = -\frac{(D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}{(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}. \tag{33}$$

It follows from (21) and Lemma 8 that both terms in (32) go to 0 as $\boldsymbol{y} \in U_j$ approaches a boundary point $\boldsymbol{y}^\star$ of U_j lying in U_j^+. Hence the condition (28) in Lemma 9 holds with $r = 1$, and we conclude that $g \in \mathcal{C}^1(\mathbb{R}^{d-1})$.

Next we differentiate with respect to the ℓth coordinate for any $\ell \neq j$ (allowing the possibility that $\ell = k$). For $\boldsymbol{y} \in U_j$ it is useful to note that for any sufficiently smooth d-variate function ξ the rule for partial differentiation and the chain rule give

$$D_\ell(\xi(\psi(\boldsymbol{y}), \boldsymbol{y})) = (D_\ell\xi)(\psi(\boldsymbol{y}), \boldsymbol{y}) + (D_j\xi)(\psi(\boldsymbol{y}), \boldsymbol{y})(D_\ell\psi)(\boldsymbol{y}). \tag{34}$$

Thus we find for $\boldsymbol{y} \in U_j$

$$
\begin{aligned}
(D_\ell D_k g)(\boldsymbol{y}) = &- \int_{-\infty}^{\psi(\boldsymbol{y})} (D_\ell D_k \theta)(x_j, \boldsymbol{y})\, \rho(x_j)\, \mathrm{d}x_j \;-\; (D_k\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, \rho(\psi(\boldsymbol{y}))\, (D_\ell \psi)(\boldsymbol{y}) \\
&+ [(D_\ell\, \theta)(\psi(\boldsymbol{y}), \boldsymbol{y}) + (D_j\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, (D_\ell \psi)(\boldsymbol{y})]\, \frac{(D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}{(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}\, \rho(\psi(\boldsymbol{y})) \\
&+ \theta(\psi(\boldsymbol{y}), \boldsymbol{y})\, \frac{[(D_\ell D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y}) + (D_j D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})\, (D_\ell \psi)(\boldsymbol{y})]}{(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}\, \rho(\psi(\boldsymbol{y})) \\
&- \theta(\psi(\boldsymbol{y}), \boldsymbol{y})\, \frac{(D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})\, [(D_\ell D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y}) + (D_j D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})\, (D_\ell \psi)(\boldsymbol{y})]}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^2}\, \rho(\psi(\boldsymbol{y})) \\
&+ \theta(\psi(\boldsymbol{y}), \boldsymbol{y})\, \frac{(D_k\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}{(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})}\, \rho'(\psi(\boldsymbol{y}))\, (D_\ell \psi)(\boldsymbol{y}),
\end{aligned}
\tag{35}
$$

where we used again (33). We have from (21) and Lemma 8 that all terms in (35) go to 0 as $\boldsymbol{y} \in U_j$ approaches a boundary point $\boldsymbol{y}^\star$ of U_j lying in U_j^+. Hence the condition (28) in Lemma 9 holds with $r = 2$, and we conclude that $g \in \mathcal{C}^2(\mathbb{R}^{d-1})$.

In general, for every non-zero multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ with $|\boldsymbol{\alpha}| \le r$ and $\alpha_j = 0$, we claim that for $\boldsymbol{y} \in U_j$

$$
(D^{\boldsymbol{\alpha}} g)(\boldsymbol{y}) = - \int_{-\infty}^{\psi(\boldsymbol{y})} (D^{\boldsymbol{\alpha}}\, \theta)(x_j, \boldsymbol{y})\, \rho(x_j)\, \mathrm{d}x_j \;+\; \sum_{m=1}^{M_{|\boldsymbol{\alpha}|}} h_{\boldsymbol{\alpha},m}(\boldsymbol{y}),
\tag{36}
$$

where $M_{|\boldsymbol{\alpha}|}$ is a nonnegative integer, and each function $h_{\boldsymbol{\alpha},m}$ is of the form (20), with integers β, a, b, c and multi-indices $\boldsymbol{\gamma}^{(i)}$ and $\boldsymbol{\eta}$ satisfying

$$
1 \le a, b \le 2|\boldsymbol{\alpha}| - 1, \quad 1 \le |\boldsymbol{\gamma}^{(i)}| \le |\boldsymbol{\alpha}|, \quad 0 \le |\boldsymbol{\eta}|, c \le |\boldsymbol{\alpha}| - 1, \quad 1 \le |\boldsymbol{\gamma}^{(i)}| + |\boldsymbol{\eta}| + c \le |\boldsymbol{\alpha}|.
\tag{37}
$$

We have from (21) and Lemma 8 that all terms in (36) go to 0 as $\boldsymbol{y} \in U_j$ approaches a boundary point $\boldsymbol{y}^\star$ of U_j lying in U_j^+. Hence the condition (28) in Lemma 9 holds for a general r, and we conclude that $g \in \mathcal{C}^r(\mathbb{R}^{d-1})$.

We will prove (36)–(37) by induction on $|\boldsymbol{\alpha}|$. The case $|\boldsymbol{\alpha}| = 1$ is shown in (32); there we have $M_1 = 1$, $a = 1$, $b = 1$, $c = 0$, $\beta = 1$, $|\boldsymbol{\gamma}^{(1)}| = 1$, $|\boldsymbol{\eta}| = 0$, and $|\boldsymbol{\gamma}^{(i)}| + |\boldsymbol{\eta}| + c = 1$. The case $|\boldsymbol{\alpha}| = 2$ is shown in (35); there we have $M_2 = 8$, $1 \le a, b \le 3$, $0 \le c \le 1$, $\beta = \pm 1$, $1 \le |\boldsymbol{\gamma}^{(i)}| \le 2$, $0 \le |\boldsymbol{\eta}| \le 1$, and $1 \le |\boldsymbol{\gamma}^{(i)}| + |\boldsymbol{\eta}| + c \le 2$. To establish the inductive step we now differentiate $D^{\boldsymbol{\alpha}} g$ once more: for $\ell \ne j$ we have from (36)

$$
\begin{aligned}
(D_\ell D^{\boldsymbol{\alpha}} g)(\boldsymbol{y}) = &- \int_{-\infty}^{\psi(\boldsymbol{y})} (D_\ell D^{\boldsymbol{\alpha}}\theta)(x_j, \boldsymbol{y})\, \rho(x_j)\, \mathrm{d}x_j \;-\; (D^{\boldsymbol{\alpha}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, \rho(\psi(\boldsymbol{y}))\, (D_\ell \psi)(\boldsymbol{y}) \\
&+ \sum_{m=1}^{M_{|\boldsymbol{\alpha}|}} (D_\ell\, h_{\boldsymbol{\alpha},m})(\boldsymbol{y}).
\end{aligned}
\tag{38}
$$

For a typical term in (38), we have from (20)

$$
\begin{aligned}
&(D_\ell\, h)(\boldsymbol{y}) \\
&= \beta\, \frac{[(D_\ell D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y}) + (D_j D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, (D_\ell \psi)(\boldsymbol{y})]\, \prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^b}\, \rho^{(c)}(\psi(\boldsymbol{y})) \\
&\quad + \beta\, \frac{(D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, D_\ell \big(\prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]\big)}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^b}\, \rho^{(c)}(\psi(\boldsymbol{y})) \\
&\quad + \beta\, \frac{(D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, \prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^b}\, \rho^{(c+1)}(\psi(\boldsymbol{y}))\, (D_\ell \psi)(\boldsymbol{y}) \\
&\quad - \beta b\, \frac{(D^{\boldsymbol{\eta}}\theta)(\psi(\boldsymbol{y}), \boldsymbol{y})\, \prod_{i=1}^{a} [(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]}{[(D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})]^{b+1}}\, \rho^{(c)}(\psi(\boldsymbol{y})) \\
&\qquad \Big[(D_\ell D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y}) + (D_j D_j\phi)(\psi(\boldsymbol{y}), \boldsymbol{y})\, (D_\ell \psi)(\boldsymbol{y})\Big],
\end{aligned}
$$

where

$$D_\ell\left(\prod_{i=1}^{a}[(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}),\boldsymbol{y})]\right)$$
$$=\sum_{t=1}^{a}\left(\left[(D_\ell D^{\boldsymbol{\gamma}^{(t)}}\phi)(\psi(\boldsymbol{y}),\boldsymbol{y})+(D_j D^{\boldsymbol{\gamma}^{(t)}}\phi)(\psi(\boldsymbol{y}),\boldsymbol{y})\,(D_\ell\psi)(\boldsymbol{y})\right]\prod_{\substack{i=1\\ i\neq t}}^{a}(D^{\boldsymbol{\gamma}^{(i)}}\phi)(\psi(\boldsymbol{y}),\boldsymbol{y})\right).$$

Thus we conclude that $D_\ell\, h$ is a sum of functions of the form (20), but with a and b increased by at most 2, c increased by at most 1, $|\beta|$ multiplied by a factor of at most b, $|\boldsymbol{\gamma}^{(i)}|$ and $|\boldsymbol{\eta}|$ increased by at most 1, and with $|\boldsymbol{\gamma}^{(i)}|+|\boldsymbol{\eta}|+c$ increased by at most 1. Hence, $D_\ell D^{\boldsymbol{\alpha}} g$ consists of a sum of functions of the form (20) satisfying the constraints in (37). This completes the induction proof for (36)–(37).

We now turn to the task of showing that $D^{\boldsymbol{\alpha}} g\in\mathcal{L}_{p,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}}$ for $p\in[1,\infty)$ and all $\boldsymbol{\alpha}$ satisfying $|\boldsymbol{\alpha}|\le r$ and $\alpha_j=0$. We need to consider

$$\int_{\mathbb{R}^{d-1}}|(D^{\boldsymbol{\alpha}}g)(\boldsymbol{y})|^p\,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}=\int_{U_j}|(D^{\boldsymbol{\alpha}}g)(\boldsymbol{y})|^p\,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

where we have split the integral noting that U_j is open and its complement U_j^+ is closed, as they are both Borel measurable, and that $D^{\boldsymbol{\alpha}}g$ is zero on U_j^+.

For $\boldsymbol{y}\in U_j$, it follows from Hölder's inequality and the special form of $D^{\boldsymbol{\alpha}}g$ in (36) that

$$\begin{aligned}|(D^{\boldsymbol{\alpha}}g)(\boldsymbol{y})|^p&=\left|-\int_{-\infty}^{\psi(\boldsymbol{y})}(D^{\boldsymbol{\alpha}}\theta)(x_j,\boldsymbol{y})\,\rho(x_j)\,\mathrm{d}x_j+\sum_{m=1}^{M_{|\boldsymbol{\alpha}|}}h_{\boldsymbol{\alpha},m}(\boldsymbol{y})\right|^p\\
&\le\left(\int_{-\infty}^{\psi(\boldsymbol{y})}|(D^{\boldsymbol{\alpha}}\theta)(x_j,\boldsymbol{y})|\,\rho(x_j)\,\mathrm{d}x_j+\sum_{m=1}^{M_{|\boldsymbol{\alpha}|}}|h_{\boldsymbol{\alpha},m}(\boldsymbol{y})|\right)^p\\
&\le(M_{|\boldsymbol{\alpha}|}+1)^{p-1}\left(\left(\int_{-\infty}^{\psi(\boldsymbol{y})}|D^{\boldsymbol{\alpha}}\theta(x_j,\boldsymbol{y})|\,\rho(x_j)\,\mathrm{d}x_j\right)^p+\sum_{m=1}^{M_{|\boldsymbol{\alpha}|}}|h_{\boldsymbol{\alpha},m}(\boldsymbol{y})|^p\right)\\
&\le(M_{|\boldsymbol{\alpha}|}+1)^{p-1}\left(\int_{-\infty}^{\psi(\boldsymbol{y})}|(D^{\boldsymbol{\alpha}}\theta)(x_j,\boldsymbol{y})|^p\,\rho(x_j)\,\mathrm{d}x_j+\sum_{m=1}^{M_{|\boldsymbol{\alpha}|}}|h_{\boldsymbol{\alpha},m}(\boldsymbol{y})|^p\right),\end{aligned}$$

and thus

$$\int_{U_j}|(D^{\boldsymbol{\alpha}}g)(\boldsymbol{y})|^p\,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$
$$\le(M_{|\boldsymbol{\alpha}|}+1)^{p-1}\left(\int_{\mathbb{R}^d}|(D^{\boldsymbol{\alpha}}\theta)(\boldsymbol{x})|^p\,\zeta_{d,\boldsymbol{\alpha}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}+\sum_{m=1}^{M_{|\boldsymbol{\alpha}|}}\int_{U_j}|h_{\boldsymbol{\alpha},m}(\boldsymbol{y})|^p\,\zeta_{d-1,\boldsymbol{\alpha}_{-j}}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)<\infty,$$

with the finiteness coming because $\rho(x_j)=\zeta_0(x_j)=\zeta_{\alpha_j}(x_j)$ and $\theta\in\mathcal{W}^r_{d,p,\zeta}$, and because each integral involving $h_{\boldsymbol{\alpha},m}$ is finite due to the condition (22). This proves that $g\in\mathcal{W}^r_{d-1,p,\zeta}$ as claimed. □

Acknowledgment

The authors acknowledge the support of the Australian Research Council under the projects FT130100655 and DP150101770.

References

[1] J. Dick, F.Y. Kuo, I.H. Sloan, High dimensional integration–the quasi-Monte Carlo way, Acta Numer. 22 (2013) 133–288.
[2] H. Bungartz, M. Griebel, Sparse grids, Acta Numer. 13 (2004) 147–269.
[3] A.B. Owen, Multidimensional variation for quasi-Monte Carlo, in: International Conference on Statistics in honour of Professor Kai-Tai Fang's 65th birthday. Jianqing Fan and Gang Li (Eds.), 2005, pp. 49–74.
[4] M. Griebel, F.Y. Kuo, I.H. Sloan, The smoothing effect of integration in $\mathbb{R}^d$ and the ANOVA decomposition, Math. Comp. 82 (2013) 383–400.
[5] M. Griebel, F.Y. Kuo, I.H. Sloan, Note on the smoothing effect of integration in $\mathbb{R}^d$ and the ANOVA decomposition, Math. Comp. 86 (2017) 1847–1854.
[6] P. Glasserman, J. Staum, Conditioning on one-step survival for barrier option simulations, Oper. Res. 49 (2001) 923–937.
[7] N. Achtsis, R. Cools, D. Nuyens, Conditional sampling for barrier option pricing under the LT method, SIAM J. Financ. Math. 4 (2013) 327–352.

[8] N. Achtsis, R. Cools, D. Nuyens, Conditional sampling for barrier option pricing under the Heston model, in: J. Dick, F.Y. Kuo, G.W. Peters, I.H. Sloan (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2012, Springer-Verlag, Berlin/Heidelberg, 2013, pp. 253–269.
[9] M. Holtz, Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance (Ph.D. thesis), Springer-Verlag, Berlin, 2011.
[10] D. Nuyens, B.J. Waterhouse, Global adaptive quasi-Monte Carlo algorithm for functions of low truncation dimension applied to problems from finance, in: L. Plaskota, H. Woźniakowski (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2010, Springer-Verlag, Berlin/Heidelberg, 2012, pp. 589–607.
[11] C. Bayer, M. Siebenmorgen, R. Tempone, Smoothing the payoff for efficient computation of Basket option prices, Quant. Finance (2018) 491–505.
[12] C. Weng, X. Wang, Z. He, Efficient computation of option prices and greeks by quasi-Monte Carlo method with smoothing and dimension reduction, SIAM J. Sci. Comput. 39 (2017) B298–B322.
[13] P. Glasserman, Monte Carlo Methods in Financial Engineering, Springer-Verlag, 2003.
[14] M. Griebel, F.Y. Kuo, I.H. Sloan, The smoothing effect of the ANOVA decomposition, J. Complexity 26 (2010) 523–551.
[15] S. Joe, F.Y. Kuo, Constructing Sobol' sequences with better two-dimensional projection, SIAM J. Sci. Comput. 30 (2008) 2635–2654.
[16] J. Matousek, On the L2-discrepancy for anchored boxes, J. Complexity 14 (1998) 527–556.
[17] K. Scheicher, Complexity and effective dimension of discrete Lévy areas, J. Complexity 23 (2007) 152–168.
[18] I.M. Sobol', Sensitivity estimates for nonlinear mathematical models, Mat. Model. 2 (1) (1990) 112–118 (in Russian). English translation in Mathematical Modeling and Computatinoal Experiment, 407–414 (1993).
[19] F.Y. Kuo, I.H. Sloan, G.W. Wasilkowski, H. Wózniakoski, On decompositions of multivariate functions, Math. Comp. 79 (2010) 953–966.