Quantization error under high-order convolution smoothing in parabolic PDE models for finance

Ruining (Ray) Wu, Christina C. Christara
Department of Computer Science
University of Toronto
Toronto, Ontario M5S 3G4, Canada
{rwu, ccc}@cs.toronto.edu

Abstract

Numerical solution of parabolic PDEs with nonsmooth initial conditions often exhibits reduced and/or unstable convergence orders due to the so-called quantization error. For a class of nonsmooth initial conditions, including representative payoffs of financial derivatives, we study how high-order convolution-based smoothings [13] modify the quantization error. For each initial condition and smoothing operator, we derive closed-form expressions for the leading order terms of the quantization error. Examining the dependence of the quantization error leading coefficient on the relative location of the discontinuity with respect to the discretization grid, we identify the minimum smoothing order required to guarantee stable fourth-order convergence (with or without grid alignment), when using fourth-order discretizations in space and time. These results explain when extrapolation-based acceleration methods, including Richardson extrapolation and the sparse grid combination approach [8], succeed or fail.

Key words: quantization error; high-order methods; nonsmooth initial conditions; Black-Scholes PDE; sparse grids; Richardson extrapolation

1 Introduction

Option pricing problems are associated with nonsmooth initial conditions, and special difficulties arise when applying standard PDE discretization methods to such problems [2], even for second-order methods. In addition to the usual truncation error in space and time arising from discretization, two other sources of error are introduced by the nonsmooth initial condition. They are high-frequency error [1, 7], and quantization error [20, 2].

High-frequency error is propagated by the diagonal Padé schemes [18, 12], a class of methods the popular Crank-Nicolson method belongs to, and lead to spurious oscillations in the solution and its spatial derivatives, even away from initial time. It can be damped by either initializing with subdiagonal Padé timestepping methods, analyzed in [18], or by using BDF methods, analyzed in [21]. The technique of replacing a fixed number of Crank-Nicolson timesteps at initial time by fully implicit timesteps is referred to as *Rannacher timestepping*, and is analyzed in detail in [1, 7]. In this paper, we consider the fourth-order diagonal (2, 2)-Padé scheme, as well as the BDF4 timestepping method to damp high-frequency error. In both cases, a subdiagonal (1, 2)-Padé initialization is applied. The stability of BDF4, with a less stable Runge-Kutta and BDF3 initializer, was analyzed in [21].

Quantization error is loosely defined as error arising from grid resolution of the point of nonsmoothness. It may depend on the positioning of the discrete grid relative to the point of nonsmoothness, and therefore, it may cause fluctuations in the order of convergence. Additionally, the quantization error may be of a

	order of smoothing operator			
	1	2	3	4
1-dim	9.01e-04	8.07e-04	4.27e-03	4.95e-03
2-dim	1.43e-01	1.87e-01	1.68e+00	1.10e+01
3-dim	5.20e+01	1.48e+02	6.73e+03	7.17e+04

Table 1: Computation time in seconds for convolution smoothings of various orders, using MATLAB's `integral`, `integral2`, and `integral3` of a max-call payoff, with $N = 16$, the number of intervals in one dimension, held fixed.

lower order than that of the discretization method, in which case it causes a reduction in the order of convergence.

Although there are many techniques for restoring the order of convergence, especially in one dimension, in this work, we consider the family of convolution smoothing operators in [13]. Compared to other Fourier-based smoothings, such as those developed in [21], the application of the smoothings in [13] does not require the knowledge of the Fourier transform of the initial conditions, thus, the convolution-based smoothings can be applied in a systematic manner to any initial condition. This makes the convolution smoothings particularly attractive for multidimensional problems, used in [4, 9], among other works.

Other methods used in the computational finance literature to restore the optimal order of convergence of numerical methods are projection methods [17, 18], grid alignment [20], grid stretching methods [16], and, for a discretization technique that scales especially well with dimension, known as sparse grids, coordinate rotation methods [14, 20], combined with one-dimensional smoothing techniques.

High-order methods in numerical PDEs, including computational finance, have gained interest [4, 9, 16, 21] due to their ability to reduce the error at a faster rate compared to their low-order counterparts. This advantage is even more pronounced in multidimensional settings, due to the curse of dimensionality. While precise requirements for stable convergence of second-order methods have been well-studied in computational finance [20, 17, 1, 7, 2, 22], there is not as much detail in the literature for fourth-order methods. This paper aims to bridge this gap. We consider a fourth-order finite difference scheme, along with a timestepping method of the same order.

Compared to low-order methods, high-order methods have even more stringent requirements on the smoothness of initial conditions. These requirements can be satisfied with high-order smoothing methods [13], that restore the optimal order of convergence of such numerical methods.

Our goal is ultimately the study of multidimensional problems, on highly efficient sparse grid discretizations [8] which reduce the impact of the curse of dimensionality. In such settings, based on the error expansion requirements given in [8], a truncation error that is independent of the grid alignment is key to stable convergence. Hence, although we study the one-dimensional problem, where grid alignment can easily be achieved, whenever possible, we are interested in a numerical scheme that exhibits a stable order of convergence independently of the grid alignment.

For multidimensional problems, smoothing can be computationally expensive, and its cost scales with both the order of the operator and the dimension d, as shown in Table 1. Hence, we are interested in using the minimum order smoothing possible while maintaining stable and optimal order of convergence.

The contributions of this paper are

- Considering initial conditions of various orders of continuity, we derive explicit formulae for computing the respective smoothed initial conditions under any order smoothing operators of the convolution type [13] in one dimension.
- Under the general framework for analysis of quantization error for a finite difference scheme in relation to the relative position $\alpha \in (0, 1]$ of the point of nonsmoothness on the grid, introduced in [2], we derive explicit formulae for the leading order terms of the quantization error for various types of initial conditions and under various smoothing operators.
- For each smoothing operator and smoothed initial condition considered, we examine the coefficient of the leading term of the quantization error as a function of α, and identify two cases of behaviour: (a) the coefficient is independent of α, or (b) the coefficient exhibits weak dependence on α.
- The distinction between these two cases allows us to suggest the minimum smoothing requirements, with or without alignment of the discontinuity with the grid, that suffice to obtain stable and unreduced order of convergence.
- Furthermore, the distinction between these two cases allows us to suggest whether the discretization method can be effectively used within the context of Richardson extrapolation (to obtain higher order of convergence), and/or within the context of sparse grids (which are also an extrapolation-based technique).

The rest of the paper is organized as follows. Section 2 presents numerical experiments that serve as motivation, defines the simplified model problem that we study, and recalls some results from [21] that we use. Section 3 presents a formula for computing convolution smoothings of any order applied to any of the one-dimensional initial conditions considered. In the same section, the formula is applied to derive explicit smoothed initial conditions values, to be used in fourth-order methods. In Section 4, we study the error in Fourier space of the smoothed initial conditions and derive formulae for the leading terms. Section 4.2 discusses the convergence behavior of the error, if the leading coefficient depends on α. Section 5 studies Richardson extrapolation, and shows that α-independence is strictly necessary for a successful extrapolation. Furthermore, Section 5 considers the sparse grid combination method [8], and shows that, for some weak α-dependence, stable order of convergence is obtained. Section 6 presents numerical results verifying our derivations. Section 7 concludes.

2 Model problem

The one-dimensional Black-Scholes PDE, in the basic form, is given by

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV, \tag{1}$$

where V denotes the option price, τ the reverse time counted from expiry T, S the asset price on which the option is written, r the risk-free interest rate, and σ the volatility of the asset price. Although the asset price S is on a semi-infinite domain $[0, \infty)$, for computational purposes, we truncate the domain to $[0, S_{\max}]$.

When an option is written on $d > 1$ underlying assets, the multidimensional Black-Scholes PDE is given

by

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\sum_{i,j=1}^{d} \sigma_i \sigma_j \rho_{i,j} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{d} r S_i \frac{\partial V}{\partial S_i} - rV, \tag{2}$$

where V and S_i are the option and asset prices as before, σ_i is the volatility of the price of the i-th asset, and $\rho_{i,j}$ is the correlation between S_i and S_j, satisfying $|\rho_{i,j}| \leq 1$, with $\rho_{i,j} = 1$ if and only if i is equal to j.

Many financial derivatives have nonsmooth, but piecewise smooth payoff functions. The most commonly occurring are calls and puts, which in one dimension are given by $\max(S - K, 0)$ and $\max(K - S, 0)$ respectively, for some positive constant K known as the strike price. The first derivative of this initial condition is not continuous at the strike $S = K$, which causes difficulty in high-order convergence of numerical methods. A similar payoff is the digital call option, which has a payoff taking on the value 1 when $S \geq K$, and 0 otherwise. The payoff itself is not continuous at the strike. Analytical solutions for the Black-Scholes PDE (1) with call/put/digital initial conditions are known, and given in standard mathematical finance textbooks, such as [10].

Two multidimensional examples are digital basket call options and basket call options. These are digital call and call options written on a basket of assets $\sum w_i S_i$, for constants $w_i \geq 0$, $\sum w_i = 1$. For two assets, the initial conditions for digital basket call options are given by

$$V^*_{\text{dcall}}(S_1, S_2) = H((wS_1 + (1-w)S_2) - K), \tag{3}$$

where H is the Heaviside function, defined later, and for basket call options, by

$$V^*_{\text{call}}(S_1, S_2) = \max((wS_1 + (1-w)S_2) - K, 0), \tag{4}$$

where K is a constant known as the strike price, and w is the proportion of S_1 in the basket of two assets. Analytical solutions of the basket call and digital basket call options are not known, hence, in our numerical experiments to test for convergence, we measure the difference between successive refinements instead of the error.

In the following, we present some preliminary experiments that motivate our work. We consider solving (1) with a digital call payoff. We use fourth-order finite differences and BDF4 timestepping initialized by three subdiagonal (1, 2)-Padé steps.

Let the *level* l of grid refinement be defined such that, for some fixed constant N_0 independent of l, the grid at level l has $N_0 2^l$ gridpoints. Let the grid alignment parameter $\alpha \in (0, 1]$ be defined such that αh is the offset of the first gridpoint to the left of the discontinuity, which may not necessarily be on a gridpoint. For a general value of $\alpha \in (0, 1]$, as the grid refines from level l to $l+1$, by inserting midpoints, α changes as

$$\alpha_{l+1} = \begin{cases} 2\alpha_l - 1 & \text{if } \alpha_l > 0.5 \\ 2\alpha_l & \text{if } \alpha_l \leq 0.5. \end{cases} \tag{5}$$

Hence, if the initial discrete grid is not aligned to the discontinuity ($\alpha_1 \neq 1$), refining the grid by inserting midpoints may cause α_l to fluctuate. In the following numerical experiment we will see how this impacts the order of convergence.

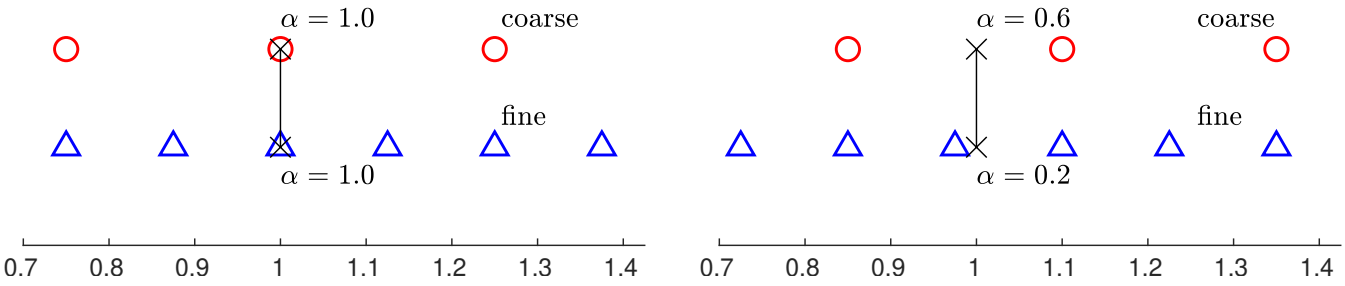

Figure 1: The discrete grids near the strike $K = 1$ for the two coarsest refinements used in Tables 2 (left) and 3 (right). As the grid is refined, if there is no grid alignment ($\alpha \neq 1$), then the grid alignment parameter α may fluctuate.

When solving (1) with a digital call payoff, we pick as initial discrete grid a uniform grid on the truncated interval $[h_1(1-\alpha_1), 16K+h_1(1-\alpha_1)]$ with initial stepsize $h_1 = 1/4$, and initial grid alignment parameter α_1 set to 1 and to 0.6, respectively, in two examples. The discrete grids are shown in Figure 1.

Dirichlet conditions with the exact solution are imposed on the two endpoints. From a coarse mesh indexed by $l-1$ to a finer mesh indexed by l, we define the convergence order estimate

$$\Upsilon_l \equiv \log_2 \left(\frac{|\text{error}_{l-1}|}{|\text{error}_l|} \right), \tag{6}$$

where the error is defined to be the difference between the numerical approximation and the exact solution at terminal time T, at the point of evaluation. For the fourth-order numerical methods considered, we expect Υ_l's to be close to 4.

The left set of numerical results in Table 2 show that only first-order convergence is observed experimentally. Additionally, as shown in the left set of results in Table 3, if the discontinuity is not aligned with a grid point, then an erratic/fluctuating order of convergence may be observed, while the average order is about one, still lower than the one the numerical methods employed should exhibit. We defer the discussion on the right set of the results in these tables until after our analysis. Here, we only mention that lower errors, as well as optimal and stable order of convergence can be obtained if an appropriate smoothing is employed.

Reduced and/or erratic order of convergence, such as that in the left set of results in Tables 2 and 3 is generally problematic. In the case of reduced (even if stable) convergence, one needs to use very fine grids to reach a certain tolerance resulting in inefficient computations. In the case of erratic (and possibly reduced) convergence, the error is not guaranteed to improve when the mesh size is refined. Furthermore, erratic convergence does not allow the application of certain techniques that improve accuracy and/or efficiency. For example, Richardson extrapolation is a popular technique to obtain a higher-order convergence and a more accurate solution from a coarse grid and a finer grid. However, extrapolation is only possible when the convergence is stable, which is not the case in the left set of results of Table 3.

It is worth noting, that for the same grid size, the errors in the left set of results in Table 3 are smaller than those in the left set of Table 2, and this is not by chance. As was derived in [2], the *quantization error* that arises from sampling the digital call payoff on the discrete grid has a leading term of $(\alpha - \frac{1}{2})h$. As a result, when we have grid alignment ($\alpha = 1$), the leading term of the quantization error is maximized.

			no smoothing			Φ_3 smoothing		
N_S	N_τ	α	value	error	Υ_l	value	error	Υ_l
64	32	1.00	0.4075306052	6.07e-02		0.3470279278	2.04e-04	
128	64	1.00	0.3761787308	2.94e-02	1.05	0.3468288942	4.97e-06	5.36
256	128	1.00	0.3612801114	1.45e-02	1.02	0.3468241684	2.43e-07	4.35
512	256	1.00	0.3539973890	7.17e-03	1.01	0.3468239400	1.47e-08	4.05
1024	512	1.00	0.3503970322	3.57e-03	1.01	0.3468239262	8.83e-10	4.05
2048	1024	1.00	0.3486070745	1.78e-03	1.00	0.3468239253	4.87e-11	4.18

Table 2: Results of solving (1) with unsmoothed (left) and Φ_3-smoothed (right) digital call payoff, with parameters $\sigma = 0.8$, $r = 0.02$, $K = 1$, and $T = 1$. Computed values of V at $S = K$. Fourth-order finite differences with BDF4 timestepping used. Discontinuity aligned with a gridpoint. Exact solution to 14 decimal places is `0.34682392528322`.

			no smoothing			Φ_4 smoothing		
N_S	N_τ	α	value	error	Υ_l	value	error	Υ_l
64	32	0.60	0.3568100083	9.99e-03		0.3471733218	3.49e-04	
128	64	0.20	0.3298054876	-1.70e-02	-0.77	0.3468414857	1.76e-05	4.31
256	128	0.40	0.3438821684	-2.94e-03	2.53	0.3468245129	5.88e-07	4.90
512	256	0.80	0.3510972981	4.27e-03	-0.54	0.3468239472	2.19e-08	4.75
1024	512	0.60	0.3475298205	7.06e-04	2.60	0.3468239263	1.06e-09	4.36
2048	1024	0.20	0.3457562179	-1.07e-03	-0.60	0.3468239253	5.47e-11	4.28

Table 3: Results of solving (1) with unsmoothed (left) and Φ_4-smoothed (right) digital call payoff, with parameters $\sigma = 0.8$, $r = 0.02$, $K = 1$, and $T = 1$. Computed values of V at $S = K$. Fourth-order finite differences with BDF4 timestepping used. Discontinuity not aligned with a gridpoint. Exact solution to 14 decimal places is `0.34682392528322`.

Sparse grids [8] reduce the curse of dimensionality of multidimensional problems, making them a highly efficient methodology. In computational finance, some works that use sparse grids are [9, 14, 19]. Sparse grids rely on error-cancellation of coarse error terms, making them similar to extrapolation. They also depend on stable convergence of the underlying method. In Table 4, we show results from solving a basket call option on sparse grids. The convergence is both reduced and unstable. Note that, in the sparse grid setting, elaborated further in Section 5.2, due to anisotropic discretizations, it is generally much more difficult to keep α consistent across refinements.

For the problem of Table 4, we defer the presentation of results with optimal order of convergence until Section 6.4, where detailed discussion and extended results are presented.

2.1 The convection-diffusion equation

In order to characterize the effect of nonsmooth initial conditions on the discretization error, we analyze the initial value problem given by the convection-diffusion equation (7),

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + a\frac{\partial u}{\partial x}, \quad (t, x) \in (0, T] \times (-\infty, \infty) \tag{7}$$

l	N_S	N_τ	value	difference	Υ_l
1	961	16	5.0699669318		
2	4867	96	5.2036951733	1.34e-01	
3	15749	384	5.1714756971	-3.22e-02	2.05
4	43655	1280	5.1793056116	7.83e-03	2.04
5	111753	3840	5.1654453384	-1.39e-02	-0.82
6	272523	10752	5.1659331648	4.88e-04	4.83

Table 4: Results of solving (2) with (unsmoothed) basket call payoff (4) on sparse grids, and parameters are $\sigma_1 = 0.4$, $\sigma_2 = 0.2$, $\rho = 0.2$, $w = 1/2$, $r = 0.03$, $K = 100$, $T = 1/4$. Computed values of V at $(S_1, S_2) = (K, K)$. Fourth-order finite differences with BDF4 timestepping used.

where a and T are constants, and some initial condition $u^{(0)}$, which is of limited continuity $\mathcal{C}^k$, $k < 2$. In particular, it may contain jump discontinuities and therefore does not necessarily belong in H^1. The spatial variable x is over the entire real line $\mathbb{R}$. Subsequent numerical experiments are performed on a sufficiently large truncated interval, such that boundary effects are negligible over the simulated time horizon $[0, T]$.

This analysis is not very restrictive as it is known that a change to logarithmic variables ($x = \log(S)$) transforms the one-dimensional Black-Scholes PDE (1) to a constant-coefficient convection-diffusion PDE, and further transformations convert it to the heat equation.

The solution u of (7) at a point (t, x), for any initial condition $u^{(0)}$, is given by

$$u(t, x) = \int_{-\infty}^{\infty} G(t, \tilde{x}) u^{(0)}(x - \tilde{x})\, d\tilde{x}, \tag{8}$$

where G is the Green's function or fundamental solution of (7), given by

$$G(t, x) = \frac{1}{\sqrt{4\pi t}} \exp\left(- \frac{(x + at)^2}{4t} \right). \tag{9}$$

The derivative $\partial u/\partial t$ of (8) at a point (t, x) is given by

$$\frac{\partial}{\partial t} u(t, x) = \int_{-\infty}^{\infty} \frac{\partial}{\partial t} G(t, \tilde{x}) u^{(0)}(x - \tilde{x})\, d\tilde{x}, \tag{10}$$

where the derivative of G at time $t \in (0, T]$ is

$$\frac{\partial}{\partial t} G(t, x) = \left(\frac{x^2 - 2t}{4t^2} - \frac{a^2}{4} \right) G(t, x). \tag{11}$$

We use (8)-(9) to compute, to a very high precision, pointwise solutions for various initial conditions described below. Additionally, (8)-(9) and (10)-(11) are also used to compute the values of u and $\partial u/\partial t$ respectively at the boundaries, where applicable.

We are interested in a family of initial conditions that are not sufficiently continuous, which, if directly sampled, could impact the order of convergence of a high-order method.

These four initial conditions are

1. Dirac delta, defined formally as a distribution satisfying the following properties:
   - $D(x) = 0$ for all $x \neq 0$, and
   - $\int_{\mathbb{R}} D(x)\, dx = 1$.
2. The Heaviside function, defined by

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

   Note that we define $H(0) = 1$ in order to be consistent with previous works [2, 21].

3. The ramp (in machine learning, also known as ReLU) function, given by

$$R(x) = \max(x, 0), \tag{13}$$

   and finally,

4. The quadratic ramp function, given by

$$Q(x) = \frac{1}{2} \max(x, 0)^2, \tag{14}$$

   where we define Q with a coefficient of $\frac{1}{2}$, again, to be consistent with [21].

The Dirac delta initial condition must be smoothed in order to be meaningfully represented on a discrete grid. The Heaviside, ramp, and quadratic ramp functions have $\mathcal{C}^{-1}$, $\mathcal{C}^0$, and $\mathcal{C}^1$ continuity respectively. These are not typically continuous enough for optimal convergence of fourth-order finite difference methods.

In general, for $p > 0$, we define p-th power ramp functions as

$$R_p(x) = \frac{1}{p!} \max(x, 0)^p, \tag{15}$$

which has continuity $\mathcal{C}^{p-1}$. For $p > 2$, fourth-order finite differences for problems with these initial conditions converge stably at fourth order.

We would like to mention that the Heaviside and ramp functions have particular interest in the computational finance community, corresponding to the payoffs of digital call and call options, respectively. For $p > 1$, p-th power ramp functions also find use in the testing of numerical methods, in order to control the continuity of initial conditions [4]. Furthermore, the respective put versions can be handled similarly.

It is also worth noting that, although we consider four particular initial conditions, our work extends naturally to any linear combination of these four initial conditions, including the respective puts, the combination involving possibly multiple discontinuities. See the results on capped call options in Section 3.5.

2.2 Numerical methods

We consider a fourth-order finite difference method in space. Let $\mathbf{u}$ denote the discrete vector of values u at the gridpoints. Then, with grid spacing h, the fourth-order finite difference approximations to the first and second derivatives at the ith gridpoint are given by

$$[\mathbf{u}_x]_i \equiv \frac{1}{12h}(\mathbf{u}_{i-2} - 8\mathbf{u}_{i-1} + 8\mathbf{u}_{i+1} - \mathbf{u}_{i+2}), \tag{16a}$$

$$[\mathbf{u}_{xx}]_i \equiv \frac{1}{12h^2}(-\mathbf{u}_{i-2} + 16\mathbf{u}_{i-1} - 30\mathbf{u}_i + 16\mathbf{u}_{i+1} - \mathbf{u}_{i+2}), \tag{16b}$$

respectively. We use uniform gridpoints throughout this analysis, however, these results generalize well to nonuniform grids which are popular in computational finance [3]. The coefficients in the finite difference approximations on nonuniform grids can be computationally generated with well-known algorithms [5] and related software [15]. For multidimensional problems, a natural (in 2D, bottom-up, then left-right) ordering of the gridpoints is used.

Near the boundary, we use second-order finite differences, which do not affect the overall fourth-order accuracy of the solution. They are given by

$$[\mathbf{u}_x]_i \equiv \frac{1}{2h}(-\mathbf{u}_{i-1} + \mathbf{u}_{i+1}), \tag{17a}$$

$$[\mathbf{u}_{xx}]_i \equiv \frac{1}{h^2}(\mathbf{u}_{i-1} - 2\mathbf{u}_i + \mathbf{u}_{i+1}). \tag{17b}$$

We impose Dirichlet boundary conditions. Where the closed-form solution is known, such as the digital call option, capped call option, or any of the initial conditions associated with the convection-diffusion equation, we use the known values and derivatives at the boundaries. Where the closed-form solution is not known, we use the time-discounted initial conditions.

Let $\mathbf{I}$ be an identity matrix of the appropriate size. Assembling the finite difference coefficients of the generic parabolic PDE $\partial u/\partial t = \mathcal{L}u$ into a matrix $\mathbf{L}$, and discretizing the term introduced by nonzero boundary conditions as a vector $\mathbf{g}_b$, the PDE semi-discretized in space becomes

$$\mathbf{u}_t = \mathbf{L}\mathbf{u} + \mathbf{g}_b. \tag{18}$$

We consider two fourth-order timestepping methods, (2, 2)-Padé and BDF4. The diagonal (2, 2)-Padé scheme is A-stable. It is also known [18] that, if the first four timesteps are replaced with third-order (1, 2)-Padé, the high-frequency components are damped.

The (1, 2)-subdiagonal Padé timestepping gives the linear system

$$\left(\mathbf{I} - \frac{2}{3}\Delta t\mathbf{L} + \frac{1}{6}\Delta t^2\mathbf{L}^2\right)\mathbf{u}^n = \left(\mathbf{I} + \frac{1}{3}\Delta t\mathbf{L}\right)\mathbf{u}^{n-1} + \Delta t\left(\frac{2}{3}\mathbf{g}_b^n + \frac{1}{3}\mathbf{g}_b^{n-1}\right) - \frac{1}{6}\Delta t^2(\mathbf{L}\mathbf{g}_b^n + \mathbf{g}_{bt}^n). \tag{19}$$

The (2, 2)-diagonal Padé timestepping gives the linear system

$$\begin{aligned}\left(\mathbf{I} - \frac{1}{2}\Delta t\mathbf{L} + \frac{1}{12}\Delta t^2\mathbf{L}^2\right)\mathbf{u}^n = \left(\mathbf{I} + \frac{1}{2}\Delta t\mathbf{L} + \frac{1}{12}\Delta t^2\mathbf{L}^2\right)\mathbf{u}^{n-1} \\ + \frac{1}{2}\Delta t(\mathbf{g}_b^n + \mathbf{g}_b^{n-1}) + \frac{1}{12}\Delta t^2(\mathbf{L}(\mathbf{g}_b^n - \mathbf{g}_b^{n-1}) + \mathbf{g}_{bt}^n - \mathbf{g}_{bt}^{n-1}).\end{aligned} \tag{20}$$

In (19) and (20), $\mathbf{g}_b^n$ and $\mathbf{g}_{bt}^n$ are $\mathbf{g}_b$ and the time derivative of $\mathbf{g}_b$ sampled at the n-th timestep, respectively.

A possible disadvantage of the (2, 2)-Padé timestepping is that it is computationally expensive, since a matrix quadratic needs to be solved per timestep. Although it is possible to reduce computation time for multidimensional problems by factoring the matrices in (19) and (20) into complex components [6], we also consider BDF4 timestepping, which is popular and efficient.

For any timestep $n > 3$, BDF4 timestepping is given by

$$\left(\mathbf{I} - \frac{12}{25}\Delta t\mathbf{L}\right)\mathbf{u}^n = \frac{48}{25}\mathbf{u}^{n-1} - \frac{36}{25}\mathbf{u}^{n-2} + \frac{16}{25}\mathbf{u}^{n-3} - \frac{3}{25}\mathbf{u}^{n-4} + \frac{12}{25}\Delta t\mathbf{g}_b^n. \tag{21}$$

Recent work in [21] shows the damping of high-frequency components by BDF4, assuming the first two timesteps are third-order Runge-Kutta, and the third step is BDF3. In the current work, when using BDF4, we initialize the first three timesteps by subdiagonal (1, 2)-Padé.

In general, we use (2, 2)-Padé for the convection-diffusion equation and BDF4 for all the option pricing problems, initialized by four and three steps, respectively, of (1, 2)-Padé.

We emphasize that although our numerical experiments have these discretizations, von Neumann (Fourier) analysis assumes an infinite discrete grid, which makes the boundary stencils (17a) and (17b) unnecessary. Our numerical experiments on finite domains replicate the behavior shown by von Neumann analysis well.

As is standard practice, von Neumann stability analysis is employed on an infinite discrete grid, which we define by

$$x_j = (j + (1 - \alpha))h, \tag{22}$$

where j takes on the range of all integers. Recall that $\alpha \in (0, 1]$ is the grid alignment parameter. When $\alpha = 1$, a gridpoint is placed on the discontinuity at $x = 0$, and α remains 1 at each grid refinement. Otherwise, α follows (5) in each refinement.

2.3 Convergence of parabolic PDEs with nonsmooth initial conditions

We summarize essential results from the convergence analysis of parabolic PDEs with nonsmooth initial conditions [13, 18]. Recall that the nonsmoothness of the initial condition, when directly sampled, introduces both high-frequency error components in the solution and its spatial derivatives, as well as quantization error, due to low-order terms in Fourier space. In a robust PDE solver, both issues need to be addressed.

In [13], considering parabolic PDEs with initial data of certain weak regularity, discretized by finite differences and one-step timestepping schemes, and under some restrictions for the time stepsize, it is shown that if the initial data is smoothed via a smoothing operator with $\mathcal{O}(h^p)$ accuracy in Fourier space for the low frequencies, and damping properties for the high frequencies, then the error of the computed solution $\mathcal{U}$ in the L_2 norm at any time $t > 0$ for a pth-order finite difference scheme satisfies

$$\|\mathcal{U}(t) - u(t)\|_{L_2} \leq Ch^p t^{-p/2}, t > 0. \tag{23}$$

Since $t = T$ is fixed away from zero, this guarantees the optimal convergence rate of $\mathcal{O}(h^p)$. Thus, in our analysis of the initial conditions under smoothing, we aim to have $\mathcal{O}(h^4)$ accuracy in Fourier space.

Furthermore, Rannacher [18], considering convection-diffusion equation with nonsmooth initial conditions, discretized in space by order p finite elements, and in time by the family of diagonal Padé schemes of order 2μ (e.g., Crank-Nicolson for $\mu = 1$), showed that the necessary damping of high-frequency components and the optimal order of convergence $\mathcal{O}(h^p + \Delta t^{2\mu})$ can be achieved without restrictions on the size of the time step Δt, by projecting the initial conditions (in the L_2 norm) onto the space of order p finite elements, *and* replacing the first 2μ timesteps of the A-stable (μ, μ)-diagonal scheme with the strongly A-stable $(\mu - 1, \mu)$-subdiagonal scheme. This timestepping technique is known as *Rannacher timestepping*, and lifts the restriction on the time stepsize imposed by the analysis in [13]. In our numerical experiments, we adopt the strongly stable (1, 2)-Padé scheme for the initialization of all problems, which corresponds to $\mu = 2$ in [18]. Although Rannacher's analysis is formulated for finite elements, the same damping mechanism applies to finite-difference discretizations, and is widely used in practice, including in option pricing [12]. Compared to the convolution-based smoothing used in this paper, the projection smoothing in [18] is of nonlocal nature, and requires the solution of a linear system.

2.4 Fourier transform of initial conditions

We review some results from [21]. Throughout this work, we define the discrete-time Fourier transform as

$$\hat{U}(\omega) = h \sum_{j=-\infty}^{\infty} U_j \exp(-i\omega x_j), \tag{24}$$

where i is the canonical imaginary unit such that $i^2 = -1$, and ω is the variable in frequency space. The inverse operator of the discrete-time Fourier transform (24) is defined as

$$U_j = \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} \hat{U}(\omega) \exp(i\omega x_j)\, d\omega. \tag{25}$$

However, the Fourier transforms of the Heaviside, ramp, and quadratic ramp functions are not convergent. Hence, when the initial condition is one of these cases, we consider some real, positive constant η, and apply the (scaled) discrete-time Fourier transform

$$\hat{U}(\omega) = h \sum_{j=-\infty}^{\infty} \exp(-\eta j h) U_j \exp(-i\omega x_j), \tag{26}$$

which has inverse operator

$$U_j = \exp(\eta j h) \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} \hat{U}(\omega) \exp(i\omega x_j)\, d\omega. \tag{27}$$

We note that the continuous Fourier transforms of D, H, R, and Q, are $\tilde{v}_D^{(0)} = 1$, $\tilde{v}_H^{(0)} = (i\kappa)^{-1}$, $\tilde{v}_R^{(0)} = (i\kappa)^{-2}$, and $\tilde{v}_Q^{(0)} = (i\kappa)^{-3}$, respectively, where $\kappa = \omega - i\eta$. Next, we recall some results from [21] indicating the accuracy of the discrete Fourier transform of the unsmoothed initial conditions as compared to their continuous counterparts.

1. For the Dirac delta distribution, since there are only a finite number of nonzero values, we will work directly with the smoothed version, and do not need the discretized unsmoothed function (which doesn't exist).

2. For the Heaviside function, given any α, the Taylor series expansion of the Fourier transform of the discretized Heaviside initial condition is given in [21] by

$$\hat{v}^{(0)}_{H,h,\alpha} = \tilde{v}^{(0)}_H + (\alpha - \frac{1}{2})h + \frac{i}{2}(\alpha^2 - \alpha + \frac{1}{6})\kappa h^2 + \frac{1}{12}\alpha(1-\alpha)(2\alpha-1)\kappa^2 h^3 \\ - \frac{i}{24}(\alpha^2(1-\alpha)^2 - \frac{1}{30})\kappa^3 h^4 + \mathcal{O}(\kappa^4 h^5). \tag{28}$$

3. For the ramp function, given any α, the Taylor series expansion of the Fourier transform of the discretized ramp initial condition is given in [21] by

$$\hat{v}^{(0)}_{R,h,\alpha} = \tilde{v}^{(0)}_R - \frac{1}{2}(\alpha^2 - \alpha + \frac{1}{6})h^2 + \frac{i}{6}\alpha(1-\alpha)(2\alpha-1)\kappa h^3 \\ + \frac{1}{8}(\alpha^2(1-\alpha)^2 - \frac{1}{30})\kappa^2 h^4 + \mathcal{O}(\kappa^3 h^5). \tag{29}$$

4. For the quadratic ramp function, given any α, the Taylor series expansion of the Fourier transform of the discretized quadratic ramp initial condition is given in [21] by

$$\hat{v}^{(0)}_{Q,h,\alpha} = \tilde{v}^{(0)}_Q - \frac{1}{12}\alpha(1-\alpha)(2\alpha-1)h^3 + \frac{i}{8}(\alpha^2(1-\alpha)^2 - \frac{1}{30})\kappa h^4 + \mathcal{O}(\kappa^2 h^5). \tag{30}$$

We will compute the discrete-time Fourier transform of the smoothed, discretized versions of D, H, R, and Q, and compare their Taylor series to their continuous counterparts.

3 Computing convolution smoothings

We introduce a general method to compute the convolution smoothings [13] of any order μ, for any initial condition among those considered. The smoothing operator in Fourier space $\hat{\Phi}_\mu$ is designed as a polynomial in $\sin(\omega)$, giving the smoothing operator a particularly convenient structure. Polynomials in $\sin(\omega)$ correspond, after inverse Fourier transformation, to linear combinations of spline basis functions, which have compact or near-compact support. In physical space, convolution with such a spline means an integration over a bounded domain $[-\nu(\mu)h, \nu(\mu)h]$, which is computationally attractive, especially for multidimensional initial conditions. Here, $\nu(\mu) = \mu - 1$ if μ even, and $\nu(\mu) = \mu - 1/2$ if μ odd.

We are interested in computing the smoothed initial condition

$$u^{(0),\Phi_\mu}_{h,\alpha}(x) = \frac{1}{h}\int_{-\nu(\mu)h}^{\nu(\mu)h} \Phi_\mu(\tilde{x}/h) u^{(0)}(x - \tilde{x})\, d\tilde{x}, \tag{31}$$

with the initial condition $u^{(0)}(x) = \max(x, 0)^p$ for various values of p, and orders of smoothing μ. For fourth-order finite differences in space, we are specifically interested in the third, fourth, and fifth order convolution operators, which have Φ_μ defined by

$$\Phi_3(x) = \frac{1}{16}\Big(\max(x-2.5,0)^2 - 13\max(x-1.5,0)^2 + 34\max(x-0.5,0)^2 \\ - 34\max(x+0.5,0)^2 + 13\max(x+1.5,0)^2 - \max(x+2.5,0)^2 \Big), \tag{32}$$

$$\Phi_4(x) = \frac{1}{36}\Big(-\max(x-3,0)^3 + 12\max(x-2,0)^3 \\ - 39\max(x-1,0)^3 + 56\max(x,0)^3 - 39\max(x+1,0)^3 \\ + 12\max(x+2,0)^3 - \max(x+3,0)^3\Big), \tag{33}$$

and

$$\Phi_5(x) = \frac{1}{27648}\Big(-47\max(x-4.5,0)^4 + 663\max(x-3.5,0)^4 - 4524\max(x-2.5,0)^4 \\ + 14748\max(x-1.5,0)^4 - 25842\max(x-0.5,0)^4 \\ + 25842\max(x+0.5,0)^4 - 14748\max(x+1.5,0)^4 \\ + 4524\max(x+2.5,0)^4 - 663\max(x+3.5,0)^4 + 47\max(x+4.5,0)^4\Big). \tag{34}$$

As mentioned above, Φ_μ is a spline, provided in the form

$$\Phi_\mu(x) = \sum_j c_j \max(x - v_j, 0)^{\mu-1} \tag{35}$$

for some given constants c_j and knots v_j.

Since Φ_μ is in the form (35), its antiderivative $F_{p,\mu}$ of any order p is given by

$$F_{p,\mu}(x) = \left(\prod_{k=0}^{p-1} \frac{1}{\mu+k}\right) \sum_j c_j \max(x - v_j, 0)^{\mu+p-1}. \tag{36}$$

Substituting (35) in (31), applying integration by parts repeatedly to (31), and using (36), reduces the power p in $u^{(0)}$ to 0, and allows us to directly evaluate (31). The smoothed values are given by

$$u_{h,\alpha}^{(0),\Phi_\mu}(x) = h^p F_{p+1,\mu}(x/h). \tag{37}$$

Despite the Dirac delta and the Heaviside initial conditions not technically belonging to the family of initial conditions $\max(x,0)^p$, their smoothed values can still be computed with (37). For the Dirac delta and Heaviside initial conditions, setting $p = -1$ and $p = 0$, respectively, in (37) gives the correct smoothed values.

Next, we use (37) to compute explicit formulas for the smoothing operator applied to the initial conditions of interest, and derive approximation properties in Fourier space.

3.1 Dirac delta

We study the Φ_3, Φ_4, and Φ_5 smoothings applied to the Dirac delta initial condition. Given any h and α, assuming that the discontinuity is located between $j = 0$ and $j = -1$, the Φ_3-smoothed values are in Table 5.

j	$\alpha > 1/2$	$\alpha \leq 1/2$
-3	unmodified	$-(2\alpha-1)^2/(64h)$
-2	$-(2\alpha-3)^2/(64h)$	$(12\alpha^2-10\alpha+1)/(16h)$
-1	$(12\alpha^2-34\alpha+23)/(16h)$	$-(44\alpha^2-29)/(32h)$
0	$-(44\alpha^2-88\alpha+15)/(32h)$	$(12\alpha^2+10\alpha+1)/(16h)$
1	$(12\alpha^2-14\alpha+3)/(16h)$	$-(2\alpha+1)^2/(64h)$
2	$-(2\alpha-1)^2/(64h)$	unmodified

Table 5: Values of Φ_3-smoothed Dirac delta.

j	$\alpha > 1/2$	$\alpha \leq 1/2$
-3	unmodified	$\frac{1}{384}(8\alpha^3-12\alpha^2+6\alpha-1)$
-2	$\frac{1}{384}(8\alpha^3-36\alpha^2+54\alpha-27)$	$\frac{1}{192}(-48\alpha^3+60\alpha^2-12\alpha-7)$
-1	$\frac{1}{192}(-48\alpha^3+204\alpha^2-276\alpha+113)$	$\frac{1}{96}(44\alpha^3-87\alpha+48)$
0	$\frac{1}{96}(44\alpha^3-132\alpha^2+45\alpha+91)$	$\frac{1}{192}(-48\alpha^3-60\alpha^2-12\alpha+199)$
1	$\frac{1}{192}(-48\alpha^3+84\alpha^2-36\alpha+199)$	$\frac{1}{384}(8\alpha^3+12\alpha^2+6\alpha+385)$
2	$\frac{1}{384}(8\alpha^3-12\alpha^2+6\alpha+383)$	unmodified

Table 6: Values of Φ_3-smoothed Heaviside function.

For values of x_j not included in the table, the smoothed values are simply the function values sampled at x_j. Note that, for the Φ_3 smoothing and for any odd μ in general, the support of Φ_μ ends at the midpoints of the discrete grid, hence, which values get smoothed depend on whether $\alpha > 1/2$ or $\alpha \leq 1/2$.

For the analysis of the Φ_4 smoothing applied to the Dirac delta initial condition, we use the values provided in Table 3.1 of [21].

We also study the Φ_5 smoothings, with the values given in Table 24 of Appendix A.

3.2 Heaviside function

Given any h and α, assuming again that the discontinuity is located between $j = 0$ and $j = -1$, the Φ_3-smoothed Heaviside function values at x_j are given in Table 6.

We will also study the Φ_4-smoothed Heaviside function with values provided in Table 3.1 of [21].

3.3 Ramp function

Given any h and α, assuming again that the discontinuity is located between $j = 0$ and $j = -1$, the Φ_3-smoothed ramp function values at x_j are given in Table 7.

We will also study the Φ_4-smoothed ramp function with values provided in Table 3.1 of [21].

3.4 Quadratic ramp function

For the quadratic ramp function, [21] does not provide analytical formulas for either Φ_3 smoothing or Φ_4 smoothing. For Φ_3 smoothing, the smoothed values are given in Table 8. For Φ_4 smoothing, the smoothed values are given in Table 9.

j	$\alpha > 1/2$	$\alpha \leq 1/2$
-3	unmodified	$\frac{1}{3072}(-16\alpha^4+32\alpha^3-24\alpha^2+8\alpha-1)h$
-2	$\frac{1}{3072}(-16\alpha^4+96\alpha^3-216\alpha^2+216\alpha-81)h$	$\frac{1}{768}(48\alpha^4-80\alpha^3+24\alpha^2+28\alpha-17)h$
-1	$\frac{1}{768}(48\alpha^4-272\alpha^3+552\alpha^2-452\alpha+107)h$	$\frac{1}{1536}(-176\alpha^4+696\alpha^2-768\alpha+197)h$
0	$\frac{1}{1536}(-176\alpha^4+704\alpha^3-360\alpha^2-1456\alpha+1485)h$	$\frac{1}{3072}(48\alpha^4+80\alpha^3+24\alpha^2-796\alpha+751)h$
1	$\frac{1}{768}(48\alpha^4-112\alpha^3+72\alpha^2-796\alpha+1539)h$	$\frac{1}{3072}(-16\alpha^4-32\alpha^3-24\alpha^2-3080\alpha+6143)h$
2	$\frac{1}{3072}(-16\alpha^4+32\alpha^3-24\alpha^2-3064\alpha+9215)h$	unmodified

Table 7: Values of Φ_3-smoothed ramp function.

j	Smoothed values ($\alpha > 1/2$)
-2	$\frac{1}{30720}(32\alpha^5-240\alpha^4+720\alpha^3-1080\alpha^2+810\alpha-243)h^2$
-1	$\frac{1}{15360}(-192\alpha^5+1360\alpha^4-3680\alpha^3+4520\alpha^2-2140\alpha+17)h^2$
0	$\frac{1}{453120}(10384\alpha^5-51920\alpha^4+35400\alpha^3+214760\alpha^2-438075\alpha+229451)h^2$
1	$\frac{1}{15360}(-192\alpha^5+560\alpha^4-480\alpha^3+7960\alpha^2-30780\alpha+30727)h^2$
2	$\frac{1}{30720}(32\alpha^5-80\alpha^4+80\alpha^3+15320\alpha^2-92150\alpha+138239)h^2$

j	Smoothed values ($\alpha \leq 1/2$)
-3	$\frac{1}{30720}(32\alpha^5-80\alpha^4+80\alpha^3-40\alpha^2+10\alpha-1)h^2$
-2	$\frac{1}{15360}(-192\alpha^5+400\alpha^4-160\alpha^3-280\alpha^2+340\alpha-115)h^2$
-1	$\frac{1}{453120}(10384\alpha^5-68440\alpha^3+113280\alpha^2-58115\alpha)h^2$
0	$\frac{1}{15360}(-192\alpha^5-400\alpha^4-160\alpha^3+7960\alpha^2-15020\alpha+7795)h^2$
1	$\frac{1}{30720}(32\alpha^5+80\alpha^4+80\alpha^3+15400\alpha^2-61430\alpha+61441)h^2$

Table 8: Values of Φ_3-smoothed quadratic ramp function.

j	Smoothed values
-3	$\frac{1}{4320}(-\alpha^6+6\alpha^5-15\alpha^4+20\alpha^3-15\alpha^2+6\alpha-1)h^2$
-2	$\frac{1}{4320}(11\alpha^6-60\alpha^5+120\alpha^4-80\alpha^3-60\alpha^2+120\alpha-52)h^2$
-1	$\frac{1}{1080}(-7\alpha^6+27\alpha^5-150\alpha^3+270\alpha^2-153\alpha)h^2$
0	$\frac{1}{1080}(7\alpha^6-15\alpha^5-30\alpha^4-20\alpha^3+555\alpha^2-1050\alpha+553)h^2$
1	$\frac{1}{4320}(-11\alpha^6+6\alpha^5+15\alpha^4+20\alpha^3+2175\alpha^2-8634\alpha+8641)h^2$
2	$\frac{1}{4320}(\alpha^6+2160\alpha^2-12960\alpha+19440)h^2$

Table 9: Values of Φ_4-smoothed quadratic ramp function.

3.5 Extension to multiple discontinuities

The analysis above assumes a single discontinuity at $x = 0$. In practice, some financial initial conditions may have multiple points of nonsmoothness. One such example is the capped call option (bull spread option) [10], which is written as the difference between two call payoffs

$$V^*_{\text{capped-call}}(S) = \max(S - K_1, 0) - \max(S - K_2, 0), \tag{38}$$

where K_1 and K_2 are the two discontinuities, with $K_1 < K_2$. More generally, any piecewise smooth payoff can be written as a linear combination of the basic initial conditions.

Due to linearity of the smoothing operator, applying the smoothing operator to the capped call initial condition is equivalent to applying the smoothing operator to the two constituent call initial conditions and taking the difference.

4 Fourier analysis of smoothed initial conditions and stable convergence

In this section, we derive explicit formulae for the leading terms of the quantization error, for the initial conditions and smoothing operators considered. More specifically, we derive formulae for the difference between the smoothed values of the initial conditions in Fourier space, and the respective continuous Fourier transform. We also study the leading term coefficient of the difference. Our target is to get Taylor expansions of the smoothed initial conditions in Fourier space, similar to those in (28), (29), (30). These Taylor expansions indicate the accuracy, in Fourier space, of the smoothed initial conditions as compared to the original ones.

4.1 Analyzing the accuracy of the smoothings in Fourier space

We are interested in computing the Fourier transforms (24) and (26) for the smoothed initial conditions. To avoid computing by hand the Fourier transforms of infinite series, we make use of the discrete-time Fourier transforms of the discretized, unsmoothed initial conditions, and compute the Fourier transform of the difference between the unsmoothed and smoothed initial conditions.

For each of the initial conditions considered, we denote the difference between the smoothed and unsmoothed initial conditions (in physical space) to be $d^{(0),\Phi_\mu}_{\cdot,h,\alpha}$, where $\cdot$ corresponds to D, H, R or Q. In other words,

$$d^{(0),\Phi_\mu}_{\cdot,h,\alpha} \equiv u^{(0),\Phi_\mu}_{\cdot,h,\alpha} - u^{(0)}_{\cdot,h,\alpha}. \tag{39}$$

The nonzero values of $d^{(0),\Phi_\mu}_{\cdot,h,\alpha}$ correspond to the points which smoothing changes the value of the function, and can be easily calculated from the unsmoothed (sampled) and the smoothed values, the latter computed by the technique of Section 3.

Using the linearity of the Fourier transform, and some re-arrangement, we get

$$\hat{u}^{(0),\Phi_\mu}_{\cdot,h,\alpha} = \hat{u}^{(0)}_{\cdot,h,\alpha} + \hat{d}^{(0),\Phi_\mu}_{\cdot,h,\alpha}. \tag{40}$$

The Taylor series expansions of the discrete, unsmoothed initial conditions $\hat{u}^{(0)}_{\cdot,h,\alpha}$ are already computed in [21], and reproduced in Section 2.4. To compute $\hat{u}^{(0),\Phi_\mu}_{\cdot,h,\alpha}$, for each of the initial conditions considered, we compute the respective Taylor series expansions of $\hat{d}^{(0),\Phi_\mu}_{\cdot,h,\alpha}$. For illustration, we provide one example for

IC	Φ_3-smoothed	Φ_4-smoothed	Φ_5-smoothed
Dirac delta	not fourth-order in general	$-\frac{1}{72}(3\alpha^2(\alpha-1)^2+2)$ range $\approx [0.0278, 0.0304]$	0
Heaviside	$\frac{1}{2880}(120\alpha^4-60\alpha^2+59)$ if $\alpha<\frac{1}{2}$, $\frac{1}{2880}(120(1-\alpha)^4-60(1-\alpha)^2+59)$ else range $\approx [0.0179, 0.0205]$	$7/240 \approx 0.0292$	not studied
Ramp	$37/1920 \approx 0.0192$	$7/240 \approx 0.0292$	not studied
Quad. ramp	$-37/1920 \approx -0.0192$	$-7/240 \approx -0.0292$	not studied

Table 10: Coefficients $C_4^{(\mu)}(\cdot,\alpha)$ of the fourth-order term of quantization error, for various initial conditions, and orders of convolution-based smoothing operators.

the ramp function under Φ_3 smoothing ($\alpha > 1/2$ branch) in Appendix B; the other cases are very similar. Once we have $\hat{d}^{(0),\Phi_\mu}_{\cdot,h,\alpha}$, through (40), we get $\hat{u}^{(0),\Phi_\mu}_{\cdot,h,\alpha}$, and this gives us the accuracy of the smoothed initial conditions as compared to the respective continuous Fourier transforms.

More specifically, for the initial conditions we considered, and for most smoothings we considered, we get

$$\hat{u}^{(0),\Phi_\mu}_{D,h,\alpha} = \tilde{u}^{(0)}_D + C_4^{(\mu)}(D,\alpha)\omega^4 h^4 + O(h^5), \tag{41}$$
$$\hat{u}^{(0),\Phi_\mu}_{H,h,\alpha} = \tilde{u}^{(0)}_H + C_4^{(\mu)}(H,\alpha)i\kappa^3 h^4 + O(h^5), \tag{42}$$
$$\hat{u}^{(0),\Phi_\mu}_{R,h,\alpha} = \tilde{u}^{(0)}_R + C_4^{(\mu)}(R,\alpha)\kappa^2 h^4 + O(h^5), \tag{43}$$
$$\hat{u}^{(0),\Phi_\mu}_{Q,h,\alpha} = \tilde{u}^{(0)}_Q + C_4^{(\mu)}(Q,\alpha)i\kappa h^4 + O(h^5), \tag{44}$$

where the coefficients $C_4^{(\mu)}(\cdot,\alpha)$ of the h^4 term are shown in Table 10.

However, the one exception is the Φ_3-smoothed Dirac delta initial condition, which satisfies

$$\hat{u}^{(0),\Phi_3}_{D,h,\alpha} = \tilde{u}^{(0)}_D + C_3^{(3)}(D,\alpha)i\omega^3 h^3 + \mathcal{O}(h^4), \tag{45}$$

where $C_3^{(3)}(D,\alpha)$ is given by

$$C_3^{(3)}(D,\alpha) = \begin{cases} -\frac{1}{24}(\alpha-1)(2\alpha-1)(2\alpha-3) & \text{if } \alpha > 1/2 \\ -\frac{1}{24}\alpha(2\alpha-1)(2\alpha+1) & \text{otherwise.} \end{cases} \tag{46}$$

This means that for the Dirac delta initial condition, applying third-order smoothing to it does not result in a fourth-order approximation, unless if $\alpha = 1/2$ or $\alpha = 1$, which both cancel out the h^3 error term. Applying fourth-order smoothing results in the coefficient associated with the h^4 term in the quantization error to be dependent on α, hence, we also study the fifth-order smoothing, which results in a fifth-order approximation.

For the Heaviside, ramp, and quadratic ramp initial conditions, we study the application of both the third and fourth-order smoothings. All combinations result in fourth-order approximation, however, the Heaviside initial condition with third-order smoothing has a varying coefficient dependent on α. Note that, in general, for the Heaviside, ramp, and quadratic ramp initial conditions, applying third-order smoothing instead of fourth-order smoothing results in a smaller coefficient for the quantization error.

4.2 On stable convergence

In [2], α-independence of the leading order coefficient of the quantization error is linked to stable convergence. While this is true, in this paper we go deeper, explain in more detail the underlying reasons, and show why, in some cases, despite not having a coefficient that is independent of α, we still observe some stability in the order of convergence.

Consider the case of a varying α, with a varying coefficient α_1 and α_2 between two successive refinements. Then, for a discretization method of theoretical order Υ, the estimated order of convergence Υ_2 calculated from these two refinements is

$$\Upsilon_2 = \log_2\left(\frac{C_4^{(\mu)}(\cdot,\alpha_1)(2h)^\Upsilon}{C_4^{(\mu)}(\cdot,\alpha_2)(h)^\Upsilon}\right) = \Upsilon + \log_2\left(\frac{C_4^{(\mu)}(\cdot,\alpha_1)}{C_4^{(\mu)}(\cdot,\alpha_2)}\right). \tag{47}$$

Therefore, the order of accuracy is perturbed from the theoretical one by at most $\log_2\left(\frac{C_4^{(\mu)}(\cdot,\alpha_1)}{C_4^{(\mu)}(\cdot,\alpha_2)}\right)$ in either direction. This value is maximized (minimized) when the ratio $\frac{C_4^{(\mu)}(\cdot,\alpha_1)}{C_4^{(\mu)}(\cdot,\alpha_2)}$ itself is maximized (minimized).

Examining the coefficients shows that for the Φ_4-smoothed Dirac delta initial condition, the difference between the estimated convergence rate Υ_2 and the ideal convergence rate Υ is at most

$$\log_2\left(\frac{1/36}{35/1152}\right) \approx -0.13. \tag{48}$$

In other words, for the fourth-order methods used in this work, the convergence rate will be between 3.87 and 4.13.

For the Φ_3-smoothed Heaviside function, with a similar calculation, the estimated order of convergence is between 3.80 and 4.20.

The qualitative difference is apparent, when we contrast with the leading term of the quantization error under no smoothing for the ramp function, which is of optimal convergence for second-order methods, if grid alignment is given. Without smoothing, the quantization error for the ramp function is second-order, and its coefficient [2] is

$$C_2^{(0)}(R,\alpha) = \frac{1}{2}(\alpha^2 - \alpha + \frac{1}{6}). \tag{49}$$

Compared with $C_4^{(4)}(D,\alpha)$ for the Φ_4-smoothed Dirac delta or $C_4^{(3)}(H,\alpha)$ for the Φ_3-smoothed Heaviside function, the leading coefficient $C_2^{(0)}(R,\alpha)$ not only changes sign throughout the interval $(0,1]$, possibly causing nonmonotone convergence, but can also become arbitrarily small as α approaches the roots of $C_2^{(0)}(R,\alpha)$, given by

$$\alpha^* = 1/2 \pm \sqrt{1/4 - 1/6} \approx 0.2113, 0.7887, \tag{50}$$

which results in an arbitrarily large change to the quantization error from one refinement to the next, causing unstable convergence. In our numerical examples, we will see that despite having a varying $C_4^{(\mu)}(\cdot,\alpha)$ coefficient from one refinement to the next, we still observe a reasonably stable order of convergence that mostly remains within the given bounds.

In addition, there are the following special cases of stable convergence. Note that choosing $\alpha = 1/3$ or $\alpha = 2/3$ causes the value of α in successive refinements to alternate between $1/3$ and $2/3$. When $C_4^{(4)}(D, 1/3) = C_4^{(4)}(D, 2/3)$, as is the case in Φ_4-smoothed Dirac delta, or $C_4^{(3)}(H, 1/3) = C_4^{(3)}(H, 2/3)$, for the case of Φ_3-smoothed Heaviside, this results in stable fourth-order convergence. When $C_2^{(0)}(R, 1/3) = C_2^{(0)}(R, 2/3)$, such as for the unsmoothed ramp function, this results in stable second-order convergence. We summarize the special cases of values of α that may have desirable properties.

- For the Dirac delta distribution, choosing $\alpha = 1$ for Φ_1-smoothing gives stable fourth-order convergence. Additionally, for Φ_3-smoothing, choosing $\alpha = 1/2$ or $\alpha = 1$ results in stable fourth-order convergence. Similarly, for Φ_4-smoothing, choosing $\alpha = 1/3$, $\alpha = 2/3$, or $\alpha = 1$, results in stable fourth-order convergence.
- For the Heaviside function, with no smoothing, choosing $\alpha = 1/2$ cancels out the first-order term, resulting in second-order convergence. With Φ_3-smoothing, choosing $\alpha = 1/3$, $\alpha = 2/3$, or $\alpha = 1$, results in stable fourth-order convergence.
- For the ramp function, with no smoothing, choosing $\alpha = 1/3$, $\alpha = 2/3$, or $\alpha = 1$, results in stable second-order convergence.
- For the quadratic ramp function, with no smoothing, choosing $\alpha = 1/2$ or $\alpha = 1$ cancels out the third-order term, resulting in fourth-order convergence.

5 Extrapolation methods

Having characterized how convolution smoothing affects the leading coefficient of the error and its possible dependence on the grid alignment parameter α, we now examine Richardson extrapolation and the sparse grid combination method to illustrate how α-dependence determines their effectiveness.

5.1 Richardson extrapolation

Richardson extrapolation increases the order of convergence of a numerical method, by combining low order approximations of different refinements, in a way that cancels out the low order terms of the error, and reveals the higher order error terms. Compared with convergence of the original numerical method, Richardson extrapolation is far more sensitive to non-stable convergence. For extrapolation of fourth-order methods, the appropriate combination to cancel out the leading error term is

$$u_{\text{extrap}} = \frac{16u_h - u_{2h}}{15}. \tag{51}$$

It is clear that upon substitution of the error expressions $C_4^{(\mu)}(\cdot, \alpha_1)(2h)^4$ and $C_4^{(\mu)}(\cdot, \alpha_2)h^4$ into (51), the fourth-order error terms cancel out only if $C_4^{(\mu)}(\cdot, \alpha_1) = C_4^{(\mu)}(\cdot, \alpha_2)$. We will observe that, in contrast to the convergence of the original numerical scheme, a higher order of convergence is observed only when $C_4^{(\mu)}(\cdot, \alpha_1) = C_4^{(\mu)}(\cdot, \alpha_2)$ holds strictly.

5.2 Sparse grid combination method

The sparse grid combination method gives fine grid accuracy by combining approximations of same order from a set of possibly anisotropic and (partly) coarse grids, in a way that cancels out the coarse grid error

terms and reveals the fine grid error terms. We give only the briefest introduction to the combination method, and only in the two-dimensional case. We index the grids with (i, j), where incrementing i or j by one doubles the number of intervals in that dimension. That is, the indices i and j denote the level of refinement in the x and y dimensions, respectively. We use the notation $u_{i,j}$ to denote a pointwise discrete solution computed on the grid (i, j). The combination method solution $u^c_{l,l}$ is given by

$$u^c_{l,l} = \sum_{i+j=l+1} u_{i,j} - \sum_{i+j=l} u_{i,j}, \tag{52}$$

which approximates the (l, l) grid. Compared to the (l, l) grid, which, for fourth-order methods, has an accuracy of $\mathcal{O}(h_l^4)$, the combination method approximation has reduced accuracy $\mathcal{O}(h_l^4 \log_2(h_l^{-1}))$. However, due to its ability to use $\mathcal{O}(h \log(h^{-1}))$ gridpoints, the combination method is more efficient per-unknown, with computational savings over full grids increasing as dimension increases. For any p-th order method, its accuracy depends on a (signed) error expansion of the form

$$u - u_{i,j} = D_x h_i^p + D_y h_j^p + D_{xy} h_i^p h_j^p, \tag{53}$$

with stable coefficients D_x, D_y, and D_{xy}. When the leading order terms of the error do not depend on the relative positioning α of the discontinuity to grid, the above coefficients are stable.

In the context of this paper, stability of D_x, D_y, and D_{xy} is achieved when the leading coefficient of $\hat{u}^{(0),\Phi_\mu}_{h,\alpha} - \tilde{u}^{(0)}$ is independent of α.

6 Numerical results

In all of our numerical experiments for the model problem (7), we truncate the domain, and our initial grid is given by

$$x_j = \{(j - \alpha_1)h_1\}_{j=-32}^{32}, \tag{54}$$

where $h_1 = 1/2$, and α_1 is the value of α at the first refinement. Grids are refined by inserting new gridpoints at midpoints of the coarser grid.

We set the remaining parameters as follows: the terminal time $T = 1/2$, the speed of convection $a = 2/5$, and $N_\tau = N_S/2$. We always evaluate the solution at $x = 0$. If x_0 is between two gridpoints, in order to avoid introducing fourth and sixth order interpolation error, which would impact the convergence study of the method and convergence study of Richardson extrapolation, respectively, we use a eighth-order polynomial interpolant using the nine nearest values to recover the solution. Although high-degree polynomial interpolation is generally not recommended due to stability reasons, especially on uniform gridpoints, in our specific application, all the higher-order derivatives away from the strike approach zero, so we are confident that it is stable.

For the numerical results of option pricing problems, following the recommendation in [11], we truncate the domain to $[0, 8S_{\text{spot}}]$, where S_{spot} is the point of evaluation, often the strike. In each dimension we use a nonuniform grid introduced in [3], computed by

$$S_j = \left(1 + \frac{\sinh(\beta - (z_i/S_{\max} - \xi))}{\sinh(\beta\xi)}\right) S_{\text{spot}}, \tag{55}$$

where z_j are uniform gridpoints on $[0, S_{\max}]$, β is numerically solved for to ensure that the last gridpoint is $S_{\max} = 8S_{\text{spot}}$, and ξ is the parameter that specifies the proportion of points in the interval $[0, S_{\text{spot}}]$. We

	Φ_1, aligned			Φ_4, not aligned			Φ_5, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l	α	error	Υ_l
64	1.00	2.15e-03		0.80	3.49e-04		0.80	1.60e-03	
128	1.00	1.37e-04	3.97	0.60	1.46e-05	4.58	0.60	1.28e-04	3.64
256	1.00	8.78e-06	3.97	0.20	1.23e-06	3.57	0.20	8.62e-06	3.89
512	1.00	5.53e-07	3.99	0.40	5.61e-08	4.45	0.40	5.50e-07	3.97
1024	1.00	3.46e-08	4.00	0.80	4.96e-09	3.50	0.80	3.47e-08	3.99
2048	1.00	2.16e-09	4.00	0.60	2.92e-10	4.09	0.60	4.02e-09	3.11

Table 11: Results of solving (7) with the Dirac delta initial condition, and parameters $a = 2/5$ and $T = 1/2$. Errors of u at $x = 0$. Various orders of smoothings considered. While we do not notice fluctuations for Φ_1 aligned and Φ_5 not aligned, minor fluctuations are present in Φ_4 not aligned. The solution accurate to 14 decimal places is `0.391042693975456`.

set $\xi = 1/4$ in our examples. Grids are refined by inserting new gridpoints of the underlying uniform grid at midpoints of the coarser grid. Nonuniform grids are commonly used in finance to reduce the error in the region of interest, usually around the strike. The parameters for the capped call option pricing problem are $K_1 = 100$, $K_2 = 115$, $S_{\text{spot}} = 110$, $T = 1/2$, $\sigma = 0.3$, $r = 0.02$, and the parameters for the digital basket call and basket call option pricing problems are identical, with $K = 100$, $S_{\text{spot}} = (K, K)$, $T = 1/4$, $\sigma_x = 0.4$, $\sigma_y = 0.2$, $r = 0.03$, $\rho = 0.2$, and $w = 1/2$.

When we have nonuniform grids, we apply *approximations* to the convolution smoothing operators [13] based on a uniform approximation of the local mesh size. For any gridpoint S_j where smoothing is to be applied, we use for h the average mesh with $\bar{h}_j = (h_j + h_{j+1})/2$. These allow us to efficiently compute convolution smoothings (31) in multiple dimensions while maintaining fourth-order convergence.

6.1 The convection-diffusion equation

We consider the solution of (7) with all four initial conditions.

Table 11 shows convergence results with the Dirac delta initial condition. We consider Φ_1 smoothing, with alignment, and Φ_4, Φ_5 smoothings without alignment. We observe that the convergence is very stable for Φ_1, as is expected from the theory in [21] which states that the low-order quantization error terms are cancelled out for the choice $\alpha = 1$. We also see stable convergence for the choice of Φ_5 smoothing. For Φ_4 smoothing, we see that while the order of convergence is approximately 4, there is slight fluctuation, due to the changing α and by extension, $C_4^{(4)}(\alpha)$. This agrees with our analysis for the Φ_4 and Φ_5 smoothings, as well as with the results in [21], where minor fluctuations are shown for Φ_4-smoothed Dirac delta, when no grid alignment is in effect.

Next, we present results for the Heaviside initial condition, in Table 12. We see that, as the theory suggests, Φ_3 smoothing results in an error that depends on the alignment α. This is shown in the first two sets of results, where keeping grid alignment gives stable fourth-order convergence, while not keeping alignment results in slight fluctuations in the order of convergence. The last set of results shows that using Φ_4 smoothing results in stable fourth-order convergence, even when we do not have grid alignment.

For the ramp initial condition, we present results with Φ_3- and Φ_4-smoothed initial conditions, both without grid alignment, in Table 13. Despite not having grid alignment, for both smoothings, stable fourth-order convergence is observed. Additionally, we observe that for the same grid, using Φ_3 smoothing

	Φ_3, aligned			Φ_3, not aligned			Φ_4, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l	α	error	Υ_l
64	1.00	-3.28e-04		0.80	-2.94e-04		0.80	-4.23e-04	
128	1.00	-2.13e-05	3.94	0.60	-1.86e-05	3.98	0.60	-2.85e-05	3.89
256	1.00	-1.36e-06	3.97	0.20	-1.34e-06	3.80	0.20	-1.84e-06	3.95
512	1.00	-8.53e-08	3.99	0.40	-7.77e-08	4.11	0.40	-1.16e-07	3.99
1024	1.00	-5.33e-09	4.00	0.80	-5.14e-09	3.92	0.80	-7.24e-09	4.00
2048	1.00	-3.33e-10	4.00	0.60	-3.02e-10	4.09	0.60	-4.53e-10	4.00

Table 12: Results of solving (7) with the Heaviside initial condition, and parameters $a = 2/5$ and $T = 1/2$. Errors of u at $x = 0$. Various orders of smoothings considered. While we do not notice fluctuations for Φ_3 aligned and Φ_4 not aligned, minor fluctuations are present in Φ_3 not aligned. The solution accurate to 14 decimal places is `0.579259709439142`.

	Φ_3, not aligned			Φ_4, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l
64	0.80	-9.53e-06		0.80	2.01e-04	
128	0.60	-9.71e-07	3.29	0.60	1.31e-05	3.94
256	0.20	-6.93e-08	3.81	0.20	8.28e-07	3.98
512	0.40	-4.54e-09	3.93	0.40	5.20e-08	3.99
1024	0.80	-2.90e-10	3.97	0.80	3.25e-09	4.00
2048	0.60	-1.80e-11	4.01	0.60	2.03e-10	4.00

Table 13: Results of solving (7) with the ramp initial condition, and parameters $a = 2/5$ and $T = 1/2$. Errors of u at $x = 0$. Various orders of smoothings considered. Neither Φ_3 not aligned nor Φ_4 not aligned have any fluctuations, and the error for Φ_3 smoothing is generally lower. The solution accurate to 14 decimal places is `0.506894635863370`.

generally results in a slightly lower error, which reflects the analysis in Section 4 showing a larger coefficient for Φ_4 smoothing compared to Φ_3 smoothing.

Finally, for the quadratic ramp initial condition, we present results with no smoothing, Φ_3, and Φ_4 smoothing applied to initial conditions, the latter both without grid alignment, in Table 14. We see that if grid alignment is present, specifically if $\alpha = 1$, then fourth-order convergence can be achieved without smoothing. However, if there is no grid alignment, then to achieve stable fourth-order convergence, Φ_3 smoothing is still required. This is exactly the same requirement as for the ramp function, and the extra continuity that the quadratic ramp function has is of no benefit, except in the aforementioned grid alignment case. Like the case for the ramp function, we observe that for the same grid, using Φ_3 smoothing generally results in a slightly lower error, which reflects the analysis in Section 4, showing a larger coefficient for Φ_4 smoothing compared to Φ_3 smoothing. Additionally, for the case with grid alignment and $\alpha = 1$, the error is lower than that of both Φ_3 and Φ_4 smoothing applied to the initial condition. This is because the quantization errors in h^3 and h^4 are both cancelled by the choice of $\alpha = 1$.

6.2 Extrapolation

For all of the cases in Section 6.1, presented in Tables 11, 12, 13, 14, we present the same results with Richardson extrapolation, in Tables 15, 16, 17, 18, respectively. We see that in Table 15, approximate

	no smoothing, $\alpha = 1$			Φ_3, not aligned			Φ_4, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l	α	error	Υ_l
64	1.00	9.47e-05		0.80	1.62e-04		0.80	2.04e-04	
128	1.00	5.98e-06	3.99	0.60	1.05e-05	3.95	0.60	1.34e-05	3.93
256	1.00	3.76e-07	3.99	0.20	6.62e-07	3.98	0.20	8.49e-07	3.98
512	1.00	2.35e-08	4.00	0.40	4.15e-08	4.00	0.40	5.33e-08	3.99
1024	1.00	1.47e-09	4.00	0.80	2.60e-09	4.00	0.80	3.33e-09	4.00
2048	1.00	9.22e-11	4.00	0.60	1.63e-10	4.00	0.60	2.09e-10	4.00

Table 14: Results of solving (7) with the quadratic ramp initial condition, and parameters $a = 2/5$ and $T = 1/2$. Errors of u at $x = 0$. Various orders of smoothings considered. Both Φ_3 and Φ_4 smoothings give stable fourth-order convergence, without alignment. However, alignment without smoothing exhibits a lower error, due to cancellation of h^3 and h^4 terms in quantization error. The solution accurate to 14 decimal places is `0.340319318305638`.

	Φ_1, aligned			Φ_4, not aligned			Φ_5, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l	α	error	Υ_l
128	1.00	3.07e-06		0.60	-7.73e-06		0.60	2.98e-05	
256	1.00	2.18e-07	3.81	0.20	3.42e-07	4.50	0.20	6.68e-07	5.48
512	1.00	4.18e-09	5.71	0.40	-2.21e-08	3.95	0.40	1.21e-08	5.79
1024	1.00	6.87e-11	5.93	0.80	1.55e-09	3.84	0.80	3.34e-10	5.18
2048	1.00	1.14e-12	5.91	0.60	-1.94e-11	6.31	0.60	1.98e-09	-2.57

Table 15: Richardson extrapolation of results in Table 11. While we achieve a higher order of convergence for Φ_1 aligned and Φ_5 not aligned, due to the α-independence of fourth-order convergence in the original numerical method, the minor fluctuations that are present in Φ_4 not aligned make Richardson extrapolation unsuccessful, that is, a higher order of convergence is not attained. The solution accurate to 14 decimal places is `0.391042693975456`.

sixth-order convergence is shown in both Φ_1 aligned and Φ_5 not aligned cases. This shows that despite approximate fourth-order convergence for the original method, because Richardson extrapolation is much more sensitive to a varying $C_4^{(\mu)}(\alpha)$, it requires the original method to be formally independent of α for the extrapolation to attain a higher order of convergence. The results in Table 16 indicate similar behavior, where the application of Richardson extrapolation to the errors arising from Φ_3-smoothing of initial conditions with no alignment fails to show a higher order of convergence, for the same reasons. The remaining cases attain a higher order of convergence. Tables 17 and 18 both show a higher order of convergence for all cases, due to the strict α-independent convergence of the original numerical method.

6.3 One-dimensional options

We now come back to Tables 2 and 3 (digital call), and consider the respective right sets of results. We notice a stable fourth order of convergence on the digital call with Φ_3-smoothing aligned, and with Φ_4-smoothing regardless of alignment, as derived in Section 4.1. Furthermore, as derived in the analysis, the $C_4^{(3)}$ coefficient of Φ_3-smoothed Heaviside ranges from 0.0179 to 0.0205 (it is maximized at $59/2880 \approx 0.0205$ when $\alpha = 1$), while the $C_4^{(4)}$ coefficient of Φ_4-smoothed Heaviside is $7/240 \approx 0.0292$, regardless of α. Thus, we theoretically expect and experimentally obtain smaller errors with Φ_3-smoothing than with

	Φ_3, aligned			Φ_3, not aligned			Φ_4, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l	α	error	Υ_l
128	1.00	-9.23e-07		0.60	-2.83e-07		0.60	-2.19e-06	
256	1.00	-2.52e-08	5.20	0.20	-1.86e-07	0.61	0.20	-6.98e-08	4.97
512	1.00	-4.45e-10	5.82	0.40	6.32e-09	4.88	0.40	-5.96e-10	6.87
1024	1.00	-7.04e-12	5.98	0.80	-3.02e-10	4.39	0.80	-8.66e-13	9.43
2048	1.00	-5.31e-14	7.05	0.60	2.10e-11	3.84	0.60	-4.09e-13	1.08

Table 16: Richardson extrapolation of results in Table 12. While a higher order of convergence is observed for Φ_3 aligned and Φ_4 not aligned, Φ_3 not aligned shows the same order of convergence. The solution accurate to 14 decimal places is `0.579259709439142`.

	Φ_3, not aligned			Φ_4, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l
128	0.60	-4.01e-07		0.60	5.51e-07	
256	0.20	-9.24e-09	5.44	0.20	1.30e-08	5.40
512	0.40	-2.21e-10	5.38	0.40	2.23e-10	5.86
1024	0.80	-6.80e-12	5.02	0.80	3.41e-12	6.03
2048	0.60	1.11e-13	5.94	0.60	1.48e-13	4.52

Table 17: Richardson extrapolation of results in Table 13. In both Φ_3 and Φ_4 cases, fourth-order error terms are successfully cancelled by Richardson extrapolation, showing a higher order of convergence. The solution accurate to 14 decimal places is `0.506894635863370`.

	no smoothing, $\alpha = 1$			Φ_3, not aligned			Φ_4, not aligned		
N_S	α	error	Υ_l	α	error	Υ_l	α	error	Υ_l
128	1.00	6.54e-08		0.60	3.60e-07		0.60	6.62e-07	
256	1.00	2.17e-09	4.91	0.20	8.87e-09	5.34	0.20	1.41e-08	5.55
512	1.00	3.93e-11	5.79	0.40	1.51e-10	5.88	0.40	2.41e-10	5.87
1024	1.00	9.78e-13	5.33	0.80	2.79e-12	5.76	0.80	4.15e-12	5.86
2048	1.00	3.17e-13	1.62	0.60	4.26e-13	2.71	0.60	4.49e-13	3.21

Table 18: Richardson extrapolation of results in Table 14. In all cases, fourth-order error terms are successfully cancelled by Richardson extrapolation, showing a higher order of convergence. The solution accurate to 14 decimal places is `0.340319318305638`.

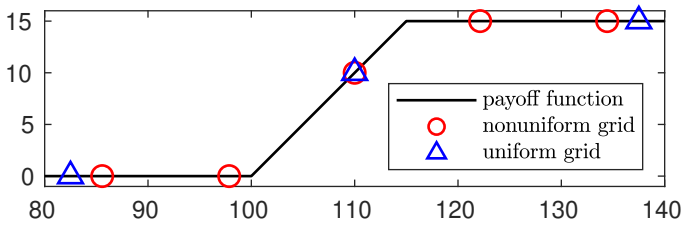

Figure 2: Capped call option payoff with nonsmoothness at $K_1 = 100$ and $K_2 = 115$, with uniform and nonuniform grids at the coarsest resolution overlaid. There is no grid alignment to either K_1 or K_2.

	uniform grid			nonuniform grid		
N_S	value	error	Υ_l	value	error	Υ_l
32	8.5891311890	8.54e-01		7.6326122535	-1.02e-01	
64	7.7172210507	-1.77e-02	5.59	7.7608554763	2.59e-02	1.98
128	7.7556698497	2.07e-02	-0.22	7.7337983347	-1.15e-03	4.49
256	7.7190189216	-1.59e-02	0.38	7.7356904705	7.39e-04	0.64
512	7.7364297939	1.48e-03	3.43	7.7345380206	-4.13e-04	0.84
1024	7.7354649037	5.14e-04	1.52	7.7350425979	9.15e-05	2.17
2048	7.7349077679	-4.33e-05	3.57	7.7349391235	-1.19e-05	2.94

Table 19: Results of solving (1) with (unsmoothed) capped call payoff (38), and parameters $\sigma = 0.3$, $r = 0.02$, $T = 1/2$, $K_1 = 100$, $K_2 = 115$. Computed values of V at $S = 110$. The solution accurate to 14 decimal places is `7.73495106401514`.

Φ_4-smoothing.

We next consider a payoff of $\mathcal{C}^0$ continuity, resembling the ramp function, but with two nonsmooth points. More specifically, in Tables 19 and 20, we show numerical experiments from solving the one-dimensional Black-Scholes PDE (1) with a capped call payoff (38). Figure 2 shows the initial condition on the coarsest grid refinement. Alignment becomes more challenging with two nonsmooth points (at K_1 and K_2); our experiments impose no grid alignment to either point.

Convergence is tested at $S = 110$ using both a uniform grid and the nonuniform grid (55), with both no smoothing and Φ_3 smoothing. The numerical results in Table 19 are, on average, second-order due to lack of smoothing, and unstable due to the lack of alignment to the nonsmooth points K_1 and K_2. With Φ_3 smoothing applied, stable fourth-order convergence is achieved for both grid types, even without alignment to either K_1 or K_2. In this example, the nonuniform grid reduces the error by a factor of roughly 20.

The analytical solution of the capped call option pricing problem can be computed as the difference of the prices of two call options with strikes at K_1 and K_2 at the point of evaluation.

6.4 Basket options on sparse grids

Finally, we show numerical experiments of multidimensional Black-Scholes PDEs on sparse grids. Sparse grids, in general, are an environment where alignment is not possible. Our analysis justifies the use of Φ_3-smoothing for call options, and suggests that Φ_3-smoothing may be attempted even for digital options. We see from the numerical results that Φ_3-smoothing indeed gives lower error than Φ_4-smoothing.

	uniform grid			nonuniform grid			
N_S	value	error	Υ_l	value	error	Υ_l	ratio
32	7.9604672787	2.26e-01		7.7500377691	1.51e-02		14.9
64	7.7575798609	2.26e-02	3.32	7.7359671913	1.02e-03	3.89	22.3
128	7.7364947199	1.54e-03	3.87	7.7350157602	6.47e-05	3.97	23.9
256	7.7350514580	1.00e-04	3.94	7.7349551634	4.10e-06	3.98	24.5
512	7.7349574031	6.34e-06	3.99	7.7349513184	2.54e-07	4.01	24.9
1024	7.7349514614	3.97e-07	4.00	7.7349510801	1.61e-08	3.99	24.8
2048	7.7349510889	2.49e-08	4.00	7.7349510650	1.01e-09	3.99	24.7

Table 20: Results of solving (1) with Φ_3-smoothed capped call payoff (38), and parameters $\sigma = 0.3$, $r = 0.02$, $T = 1/2$, $K_1 = 100$, $K_2 = 115$. Computed values of V at $S = 110$. The column "ratio" indicates the ratio of the error on the uniform grid over that on the corresponding nonuniform grid. The solution accurate to 14 decimal places is `7.73495106401514`.

			Φ_3 smoothing			Φ_4 smoothing		
l	N_S	N_τ	value	difference	Υ_l	value	difference	Υ_l
1	961	16	5.1625751813			5.1661475458		
2	4867	96	5.1609244254	-1.65e-03		5.1611183557	-5.03e-03	
3	15749	384	5.1608426513	-8.18e-05	4.34	5.1608519397	-2.66e-04	4.24
4	43655	1280	5.1608384970	-4.15e-06	4.30	5.1608386462	-1.33e-05	4.32
5	111753	3840	5.1608380642	-4.33e-07	3.26	5.1608380053	-6.41e-07	4.37
6	272523	10752	5.1608380356	-2.86e-08	3.92	5.1608379700	-3.53e-08	4.18

Table 21: Results of solving (2) with basket call payoff (4) on sparse grids, and parameters $\sigma_1 = 0.4$, $\sigma_2 = 0.2$, $\rho = 0.2$, $w = 1/2$, $r = 0.03$ and $T = 1/4$. Computed values of V at $(S_1, S_2) = (K, K)$. Various orders of smoothings considered. Note that Φ_3 smoothing has a lower error than Φ_4 smoothing, as our theory suggests.

More specifically, in Tables 21 and 22, we present results from solving (2) with basket call and digital basket call payoffs, on sparse grids. The Φ_3 smoothing indeed gives a lower error, and we see that even in the case of digital options, Φ_3 smoothing on sparse grids, corresponding to the case of no alignment, gives stable convergence. In these tables, l refers to the grid level, and N_S (N_τ) the total number of unknowns (timesteps) over *all* subproblems. We emphasize that about four orders of magnitude for the error are gained at level $l = 6$ by employing the appropriate smoothing: compare the results of Tables 4 and 21.

7 Conclusions

We considered the numerical solution of parabolic PDEs with nonsmooth initial conditions and provided a comprehensive analysis of the quantization error that arises when high-order convolution-based smoothing operators [13] are applied to the nonsmooth initial conditions. More specifically, we derived explicit closed-form expressions for the leading order quantization error terms as a function of the grid alignment parameter α. The derived formulae allow us to explicitly state the minimum smoothing requirements necessary to restore stable, optimal convergence, when fourth-order discretization methods are used.

			Φ_3 smoothing			Φ_4 smoothing		
l	N_S	N_τ	value	difference	Υ_l	value	difference	Υ_l
1	961	16	0.4915289911			0.4917791699		
2	4867	96	0.4908951947	-6.34e-04		0.4909175761	-8.62e-04	
3	15749	384	0.4908601300	-3.51e-05	4.18	0.4908616395	-5.59e-05	3.95
4	43655	1280	0.4908584295	-1.70e-06	4.37	0.4908585278	-3.11e-06	4.17
5	111753	3840	0.4908583620	-6.75e-08	4.65	0.4908583679	-1.60e-07	4.28
6	272523	10752	0.4908583602	-1.79e-09	5.24	0.4908583606	-7.31e-09	4.45

Table 22: Results of solving (2) with digital basket call payoff (3) on sparse grids, and parameters $\sigma_1 = 0.4$, $\sigma_2 = 0.2$, $\rho = 0.2$, $w = 1/2$, $r = 0.03$ and $T = 1/4$. Computed values of V at $(S_1, S_2) = (K, K)$. Various orders of smoothings considered. Note that Φ_3 smoothing has a lower error than Φ_4 smoothing, as our theory suggests. Additionally, the slight fluctuations in the convergence of the convection-diffusion equation (7) with the Heaviside initial condition shown in Table 12 do not appear here.

Our key findings establish practical guidance for high-order methods:

- For the Dirac delta initial condition, we showed that Φ_4 smoothing results in a quantization error term that depends weakly on the grid alignment parameter α. To achieve stable fourth-order convergence independent of grid alignment, Φ_5 smoothing must be used, as it is shown to remove the fourth-order term entirely.
- For the Heaviside function (modeling digital options), our analysis shows that Φ_3 smoothing results in a quantization error term that remains weakly dependent on the grid alignment parameter α. To achieve stable fourth-order convergence independent of grid alignment, the higher-order Φ_4 smoothing is required.
- For the ramp function (modeling call/put options), and for the quadratic ramp function, we demonstrated that Φ_3 smoothing is sufficient to achieve stable and unreduced fourth-order convergence even without grid alignment.

Additional findings for special cases of α values with desirable properties are given in Section 4.2.

Although our quantization error analysis focuses on initial conditions with a single discontinuity, the linearity of both the smoothing operators and the PDE extends the results directly to payoffs with multiple discontinuities, such as capped calls, by smoothing each constituent component.

The significance of the findings lies in the observation that the computational cost of smoothing increases with both the order and the dimension, making the minimum required smoothing order crucial for efficiency, especially in multidimensional problems.

Furthermore, our analysis explained the success and failure of extrapolation-based acceleration methods. We showed that Richardson extrapolation is extremely sensitive even to weak dependence of the leading term coefficient on α, and succeeds in achieving a higher order of convergence only when the convergence of the underlying numerical method is strictly α-independent. In contrast, the sparse grid combination method proves more robust, tolerating a weak dependence on α, while still providing stable convergence.

These explicit results provide the necessary theoretical foundation for reliably applying high-order finite

Initial condition		requirement for stable fourth-order convergence
Dirac delta	If alignment is possible,	set $\alpha = 1$ and any smoothing is sufficient.
	Otherwise,	Φ_3 smoothing with $\alpha = 1/2$ gives fourth-order convergence,
	and	Φ_4 smoothing gives fourth-order convergence, with a coefficient weakly depending on α, but remaining constant if $\alpha = 1/3$ or $\alpha = 2/3$,
	and	Φ_5 smoothing gives stable fourth-order convergence, with a coefficient independent of α.
Heaviside		Φ_3 smoothing gives fourth-order convergence, with a coefficient weakly depending on α, but remaining constant if $\alpha = 1/3$ or $\alpha = 2/3$,
	and	Φ_4 smoothing gives stable fourth-order convergence, with a coefficient independent of α, but slightly larger than the maximum Φ_3 coefficient value.
ramp		Φ_3 smoothing gives stable fourth-order convergence, with a smaller coefficient than that for Φ_4.
quadratic ramp	If alignment is possible,	set $\alpha = 1$ or $\alpha = 1/2$ and no smoothing is required.
	Otherwise,	Φ_3 smoothing gives stable fourth-order convergence, with a smaller coefficient than that for Φ_4.

Table 23: Summary of recommendations on the minimum order convolution-based smoothing required to obtain fourth-order error and stable convergence in the presence of insufficiently smooth initial conditions, when fourth-order finite differences and timestepping are used.

difference and timestepping methods to realistic financial PDE models, especially in the context of high-dimensional sparse grid discretizations, where grid alignment cannot be guaranteed.

In Table 23, we summarize our results for the minimum order of convolution smoothing required to preserve stable fourth-order convergence, including special cases of α.

8 Acknowledgements

This work is partly supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, through Discovery Grant RGPIN-2021-03502, and the Postgraduate Scholarship - Doctoral (PGS-D) program, and by the Ontario Graduate Scholarship (OGS).

A Fifth-order smoothing of Dirac delta

In Table 24, we provide the values that are produced by Φ_5-smoothing applied to Dirac delta.

j	Smoothed values ($\alpha > 1/2$)
-4	$(47(2\alpha-3)^4)/(442368h)$
-3	$(-1232\alpha^4+7016\alpha^3-14376\alpha^2+12026\alpha-3041)/(55296h)$
-2	$(15632\alpha^4-84688\alpha^3+158664\alpha^2-110788\alpha+16229)/(110592h)$
-1	$(-21680\alpha^4+103368\alpha^3-144936\alpha^2+25842\alpha+43677)/(55296h)$
0	$(120016\alpha^4-480064\alpha^3+450456\alpha^2+59216\alpha+38005)/(221184h)$
1	$(-21680\alpha^4+70072\alpha^3-45048\alpha^2+7246\alpha-4319)/(55296h)$
2	$(15632\alpha^4-40368\alpha^3+25704\alpha^2-7836\alpha+1917)/(110592h)$
3	$(-1232\alpha^4+2840\alpha^3-1848\alpha^2+710\alpha-77)/(55296h)$
4	$(47(2\alpha-1)^4)/(442368h)$

j	Smoothed values ($\alpha \leq 1/2$)
-5	$(47(2\alpha-1)^4)/(442368h)$
-4	$(-1232\alpha^4+2088\alpha^3-720\alpha^2-606\alpha+393)/(55296h)$
-3	$(15632\alpha^4-22160\alpha^3-1608\alpha^2+15004\alpha-4951)/(110592h)$
-2	$(-21680\alpha^4+16648\alpha^3+35088\alpha^2-40646\alpha+6271)/(55296h)$
-1	$(120016\alpha^4-269640\alpha^2+187629)/(221184h)$
0	$(-21680\alpha^4-16648\alpha^3+35088\alpha^2+40646\alpha+6271)/(55296h)$
1	$(15632\alpha^4+22160\alpha^3-1608\alpha^2-15004\alpha-4951)/(110592h)$
2	$(-1232\alpha^4-2088\alpha^3-720\alpha^2+606\alpha+393)/(55296h)$
3	$(47(2\alpha+1)^4)/(442368h)$

Table 24: Values of Φ_5-smoothed Dirac delta.

B Taylor expansion of ramp function under Φ_3 smoothing

For simplicity, consider the case $\alpha > 1/2$. We show the Taylor series expansion for $\hat{d}^{(0),\Phi_3}_{R,h,\alpha}$. Applying the discrete-time Fourier transform (26) to $u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)$, where $u^{(0),\Phi_3}_{R,h,\alpha}(x_j)$ are the function values in the right side of Table 7 and $R(x_j)$ are directly sampled values of the ramp function, gives

$$\hat{d}^{(0),\Phi_3}_{R,h,\alpha} = h\sum_{j=-2}^{2}\exp(-i\kappa x_j)(u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)). \tag{56}$$

The summation is over $j=-2$ to $j=2$, because these are the values modified by Φ_3 smoothing, when $\alpha > 1/2$. If $\alpha \leq 1/2$, the summation is over $j=-3$ to $j=1$.

We substitute the value of the discrete gridpoints $x_j = (j+(1-\alpha))h$ in, to get

$$\hat{d}^{(0),\Phi_3}_{R,h,\alpha} = h\sum_{j=-2}^{2}\exp(-i\kappa(j+(1-\alpha))h)(u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)), \tag{57}$$

Applying Taylor series expansion (in h) to the exponential $\exp(-hC)$, for $C = i\kappa(j+(1-\alpha))$, for terms up to $\mathcal{O}(h^2)$ (higher terms are not necessary since there is one factor of h before the summation, and for

all j, $u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j) = \mathcal{O}(h)$), gives

$$\exp(-hC) = 1 - hC + \frac{1}{2}(-hC)^2 + \mathcal{O}(h^3). \tag{58}$$

Next, we multiply this with $u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)$ and collect like terms in powers of h. Note that every term in $u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)$ is a constant multiplied by h. Then, the h^2 term of the expansion is simply given by the sum of the $u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)$ entries, which turns out to be $\frac{1}{2}(\alpha^2 - \alpha + \frac{1}{6})h^2$. Similarly, collecting the h^3 terms in the expansion, which is achieved by multiplying the $-hC$ term through the $u^{(0),\Phi_3}_{R,h,\alpha}(x_j) - R(x_j)$ values and summing, gives $-\frac{i}{6}\alpha(1-\alpha)(2\alpha-1)\kappa h^3$. And finally, for the h^4 term in the expansion, we get $-\frac{1}{8}((\alpha^2(1-\alpha)^2) - \frac{1}{30} + \frac{37}{240})\kappa^2 h^4$. Putting this all together gives

$$\hat{d}^{(0),\Phi_3}_{R,h,\alpha} = \frac{1}{2}(\alpha^2 - \alpha + \frac{1}{6})h^2 - \frac{i}{6}\alpha(1-\alpha)(2\alpha-1)\kappa h^3 - \frac{1}{8}((\alpha^2(1-\alpha)^2) - \frac{1}{30} + \frac{37}{240})\kappa^2 h^4 + \mathcal{O}(h^5), \tag{59}$$

which correctly cancels out the low-order terms in (29), leaving the term $\frac{37}{1920}\kappa^2 h^4$, which is independent of α, and higher order terms.

References

[1] R. CARTER AND M. B. GILES, *Sharp error estimates for discretizations of the 1D convection-diffusion equation with Dirac initial data*, IMA Journal of Numerical Analysis, 27 (2007), pp. 406–425.

[2] C. CHRISTARA AND N. C.-H. LEUNG, *Analysis of quantization error in financial pricing via finite difference methods*, SIAM Journal on Numerical Analysis, 56 (2018), pp. 1731–1757.

[3] N. CLARKE AND K. PARROTT, *Multigrid for American option pricing with stochastic volatility*, Applied Mathematical Finance, 6 (1999), pp. 177–195.

[4] B. DÜRING AND C. HEUER, *High-order compact schemes for parabolic problems with mixed derivatives in multiple space dimensions*, SIAM Journal on Numerical Analysis, 53 (2015), pp. 2113–2134.

[5] B. FORNBERG, *Classroom note: Calculation of weights in finite difference formulas*, SIAM Review, 40 (1998), pp. 685–691.

[6] J. D. GEORGE, S. Y. JUNG, AND N. M. MANGAN, *Walking into the complex plane to "order" better time integrators*, arXiv preprint arXiv:2110.04402, (2021).

[7] M. GILES AND R. CARTER, *Convergence analysis of Crank-Nicolson and Rannacher time-marching*, Journal of Computational Finance, 9 (2006), pp. 89–112.

[8] M. GRIEBEL, M. SCHNEIDER, AND C. ZENGER, *A combination technique for the solution of sparse grid problems*, Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra, (1992), pp. 263–292.

[9] C. HENDRICKS, C. HEUER, M. EHRHARDT, AND M. GÜNTHER, *High-order ADI finite difference schemes for parabolic equations in the combination technique with application in finance*, Journal of Computational and Applied Mathematics, 316 (2017), pp. 175–194.

[10] D. J. HIGHAM, *An introduction to Financial Option Valuation*, Cambridge University Press, 2004.

[11] R. KANGRO AND R. NICOLAIDES, *Far field boundary conditions for Black–Scholes equations*, SIAM Journal on Numerical Analysis, 38 (2000), pp. 1357–1368.

[12] A. KHALIQ, D. VOSS, AND M. YOUSUF, *Pricing exotic options with L-stable Padé schemes*, Journal of Banking and Finance, 31 (2007), pp. 3438–3461.

[13] H.-O. KREISS, V. THOMÉE, AND O. WIDLUND, *Smoothing of initial data and rates of convergence for parabolic difference equations*, Communications on Pure and Applied Mathematics, 23 (1970), pp. 241–259.

[14] C. LEENTVAAR AND C. W. OOSTERLEE, *On coordinate transformation and grid stretching for sparse grid pricing of basket options*, Journal of Computational and Applied Mathematics, 222 (2008), pp. 193–209.

[15] R. J. LEVEQUE, *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*, SIAM, 2007.

[16] C. W. OOSTERLEE, C. C. LEENTVAAR, AND X. HUANG, *Accurate American option pricing by grid stretching and high order finite differences*, Delft University of Technology, The Netherlands, Technical Report, (2005).

[17] D. POOLEY, K. VETZAL, AND P. FORSYTH, *Convergence remedies for non-smooth payoffs in option pricing*, Journal of Computational Finance, 6 (2003), pp. 25–40.

[18] R. RANNACHER, *Finite element solution of diffusion problems with irregular data*, Numerische Mathematik, 43 (1984), pp. 309–327.

[19] C. REISINGER AND G. WITTUM, *Efficient hierarchical approximation of high-dimensional option pricing problems*, SIAM Journal on Scientific Computing, 29 (2007), pp. 440–458.

[20] D. TAVELLA AND C. RANDALL, *Finite differences in pricing financial instruments*, Wiley, 2000.

[21] D. WANG, *High-order numerical methods for parabolic PDEs with nonsmooth data: a perspective from option pricing*, PhD thesis, University of Toronto, 2024. Available at: http://hdl.handle.net/1807/141099.

[22] R. WU AND C. C. CHRISTARA, *The sparse grid combination method for multidimensional Black-Scholes partial differential equations*, in Addressing Modern Challenges in the Mathematical, Statistical, and Computational Sciences, D. M. Kilgour, H. Kunze, R. N. Makarov, R. Melnik, and X. Wang, eds., Cham, 2025, Springer Nature Switzerland, pp. 399–409.