A SURVEY OF NUMERICAL METHODS FOR LÉVY MARKETS

MICHAEL CHIU

AUGUST 7, 2014

Abstract. The modeling of financial markets by Lévy process has become an active area of research during recent years. This has motivated an equal amount of, if not more, research activity into the necessary numerical methods. Due to the large body of work in this area, we focus our survey on fast numerical methods for Lévy markets. Particular emphasis is placed on grid-based methods.

1. Introduction. Numerical methods for computational finance can be grouped into three main categories: Monte Carlo methods, Finite Difference/Element/Volume (FD/FE/FV) methods and Transform methods. Monte Carlo methods are by far the most prevalent of the three categories due to their being the simplest to implement and possibly the most widely applicable. Indeed, for some complex or high dimensional structures, Monte Carlo might be the only way possible to numerically value such derivatives. However, it is well known that the Monte Carlo method suffers from slow convergence, and consequently, is not applicable in many practical situations. This drawback motivates research into alternative methods. The second class of methods, FD/FE/FV methods, are natural applications of standard methods from computational mathematics originally designed for problems in science and engineering. That the solution of the option pricing problem, in the Black-Scholes (BS) case, reduces to the solution of the heat equation corroborates the applicability of such classes of methods. Transform methods, the third class of numerical methods, in contrast to the PDE-based framework of the FD/FE/FV methods, look to exploit properties of integration in order to efficiently arrive at a solution. Transform methods are extremely efficient when applicable and in many cases are simpler for Lévy Processes than Gaussian Processes. Gaussian Processes are also Lévy Processes, but in a trivial sense–see the quick review section on Lévy Processes. In all cases, due to the Lévy-Ito decomposition, Lévy processes can be decomposed into three main components: a linear Brownian Motion, a compound Poisson Process, and a square integrable pure jump martingale [77]. In many cases, this translates to the fact that existing methods only need to be extended such that the additional jump terms are taken into account. As such, it will be worthwhile expounding on the numerical methods for some products in the Gaussian case (i.e. American options), as the Lévy case follows via additional consideration of the jump term.

1.1. Lévy Processes. It is well known that financial returns deviate from normality and there has been much research into more realistic processes for modelling asset returns. Naturally, incorporating jumps is an intuitive step forward from the BS model; this can clearly be accomplished with models driven by Lévy Processes. We begin by reviewing some basic facts regarding Lévy Processes.

DEFINITION 1.1. *A Stochastic process $X = \{X_t : t \geq 0\}$ defined on the standard probability space with triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is a Lévy Processes if it satisfies the following:*

(i) $\mathbb{P}(X_0 = 0) = 1$.

(ii) *X has càdlàg paths, i.e. trajectories are $\mathbb{P}$-a.s. right continuous with left limits.*

(iii) *X has independent increments, i.e. $\forall t, h > 0$, $X_{t+h} - X_t$ is independent of X_s for all $s \leq t$.*

(iv) *X has stationary increments, i.e. $\forall h > 0$, $X_{t+h} - X_t$ has the same probability law as X_h.*

Clearly, from the definitions above, we see that Lévy Processes are also Markov processes. We defer to [83] for details into the theory of Markov Processes. The following theorem describes explicitly and completely, a Lévy Process in terms of its Fourier Transform. This representation also links a Lévy Process with its characteristic function.

THEOREM 1.2. *Lévy-Khintchine representation. Let X be a Lévy process, then*

$$\mathbb{E}\big[e^{iuX_t}\big] = e^{it\Psi(u)}$$

where the characteristic exponent Ψ has the form

$$\Psi(u) := i\gamma u + \frac{1}{2}\sigma^2 u^2 + \int_\Omega \big(1 - e^{iuz} + iuz1_{|z|\leq 1}\big)\nu(dz)$$

and $\gamma \in \mathbb{R}$, $\sigma \geq 0$ (constants) and ν is a σ-finite measure on $\mathcal{B}(\mathbb{R} \setminus 0)$ satisfying the conditions

$$\int_{\mathbb{R}\setminus 0} \min(1, z^2)\nu(dz) < \infty \tag{1.1}$$

the triplet (σ, γ, ν) is known as the characteristic triplet and ν is known as the Lévy measure.

The advantages of being able to characterize processes in this manner should be clear. However, we highlight the simple and well known fact that linear Brownian motion has the characteristic exponent $\Psi(u) := \sigma^2 u^2/2 + i\gamma u$ [15]. The fact that the Lévy measure is missing in this characteristic exponent is the reason that the Brownian motion is also known as a trivial Lévy process.

From this representation, Lévy processes are usually decomposed as $X_t = \sigma B_t + Y_t$, where B_t is a Brownian motion and Y_t is a quadratic pure jump Lévy process independent of B_t. Lévy process with $\sigma = 0$ are known as *pure jump processes*. It is through various parameterizations of the Lévy measure ν, that proposed exponential Lévy models in the literature differ. We show, in a later section, a few examples of popular choices for the Lévy measure. As mentioned above, Lévy processes are Markov processes and, consequently, they posses an infinitesimal generator $\mathcal{L}^X$ of the form

$$\begin{aligned}\mathcal{L}^X f(x) &= \lim_{t\to 0} \frac{\mathbb{E}[f(x + X_t)] - f(x)}{t} \\ &= \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} + \gamma\frac{\partial f}{\partial x} + \int_{\mathbb{R}} \nu(dy)\left[f(x+y) - f(x) - y1_{|y|\leq 1}\frac{\partial f}{\partial x}\right]\end{aligned} \tag{1.2}$$

which is the basic and fundamental connection between a (jump) diffusion and a partial integro differential operator.

We refer the reader to the treatises [7, 15, 59, 77] for more standard and rigorous expositions on Lévy Processes.

1.2. Exponential Lévy Models. Let $(S_t)_{t\in[0,T]}$ denote the price of a financial asset. This is usually modelled as a stochastic process on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. By no arbitrage arguments, there exists a martingale measure equivalent to $\mathbb{P}$ such that S_t is a martingale. This is usually known as the risk-neutral or pricing measure which we denote as $\mathbb{Q}$. Analogous to the Gaussian case, we assume that the price process follows, under $\mathbb{Q}$, exponential Lévy dynamics

$$S_t = S_0 e^{rt+X_t} \tag{1.3}$$

where X_t is a Lévy process with characteristic triplet (σ, γ, ν). Particular Lévy Models, which we introduce in the next section, correspond to different parameterizations of the Lévy measure ν. No arbitrage imposes the following conditions on the corresponding characteristic triplet (σ, γ, ν)

$$\int_{|y|>1} \nu_{\mathbb{Q}}(dy)e^y < \infty \tag{1.4}$$

$$\gamma = -\frac{\sigma^2}{2} - \int_{\mathbb{R}} (e^y - 1 - y1_{|y|\leq 1})\nu_{\mathbb{Q}}(dy) \tag{1.5}$$

The infinitesimal generator associated with the exponential Lévy model becomes

$$\mathcal{L}^X f = \frac{\sigma^2}{2}\left[\frac{\partial^2 f}{\partial x^2} - \frac{\partial f}{\partial x}\right] + \int_{\mathbb{R}} \nu_{\mathbb{Q}}(dy)\left[f(x+y) - f(x) - (e^y - 1)\frac{\partial f}{\partial x}\right] \tag{1.6}$$

This is the general form of the infinitesimal generator for the exponential Lévy model and at this point it is customary to assume that $\nu_{\mathbb{Q}}(dy)$ has a Lévy density, i.e., $\nu_{\mathbb{Q}}(dy) = k_{\mathbb{Q}}(y)dy$ for (practical) computational considerations.

1.2.1. Particular Lévy Models. As mentioned above, various jump diffusion models parameterize differently the Lévy measure $\nu_{\mathbb{Q}}(dy)$. These models include: the Variance Gamma model [63], based on a pure jump process of finite variation and its generalization, the CGMY model [22], extended to allow for infinite variation; the Hyperbolic Process [38] and its generalization the Generalized Hyperbolic Model [39]; Normal Inverse Gaussian (NIG) process [10] and its generalization, the Normal Tempered Stable Levy Processes [11, 12]. There are also models based on non-infinitely divisible processes such as the Truncated Levy Distributions [64] and their infinitely divisible analogue given in [69]. Empirical studies support the claim that infinite activity finite variation Lévy processes better represent stock price dynamics [48]. In many cases, these Lévy based models have the BS model as their limiting case. A very readable introduction into the current literature of proposed Lévy Models is [28].

1.2.2. Assumptions on the Lévy density. It is important to list standard assumptions on the Lévy measure. These assumptions are basic and are not always all satisfied by all proposed Lévy models. However, some assumptions, such as (A3), are necessary for more advanced analyses, such as the wavelet compression of the moment matrix of the associated Lévy density. Assumptions (A1-A3) are satisfied by virtually every proposed Lévy model.

(A1) *Activity of small jumps.*

$$\forall u \in \mathbb{R} \quad |\psi_0(u) - icu| \leq C(1 + |u|^2)^{Y/2} \tag{1.7}$$

where $\psi_0(u)$ denotes the pure jump portion, Y_t, of X_t. $C > 0$ and $Y < 2$ are constants.

(A2) *Semiheavy tails.*

$$\forall |z| > 1, \quad k(z) \leq C \begin{cases} e^{-G|z|} & \text{if} \quad z < 0 \\ e^{-M|z|} & \text{if} \quad z > 0 \end{cases} \tag{1.8}$$

where $C > 0$, $G > 0$, $M > 1$ are constants.

(A3) *Smoothness.*

$$\forall z \neq 0, \quad |k^{(\alpha)}(z)| \leq C(\alpha)|z|^{-(1+Y+\alpha)_+} \tag{1.9}$$

where $\alpha \in \mathbb{N}_0$ and $C(\alpha)$ is a constant depending on α.

(A4) *Boundedness from below.*

$$\forall 0 < |z| < 1, \quad \frac{1}{2}(k(-z) + k(z)) \geq \frac{C_-}{|z|^{1+Y}} \tag{1.10}$$

If X_t is an infinite activity Lévy process, then k is not integrable; k is integrable if X_t is a finite activity process. Furthermore, infinite processes satisfy (A4) above with $Y < 0$, whereas finite processes do not (with $0 < Y < 2$).

1.2.3. CGMY. A very popular model proposed by Carr et al. [22], the CGMY process is typically considered part of the infinite activity family of models and subsumes the Variance Gamma model (CGMY with $Y = 0$). The Lévy density takes the functional form

$$k(z) = C \begin{cases} \frac{e^{-G|z|}}{|z|^{1+Y}} & \text{if} \quad z < 0 \\ \frac{e^{-M|z|}}{|z|^{1+Y}} & \text{if} \quad z > 0 \end{cases} \tag{1.11}$$

with $C > 0$, $G, M \geq 0$ and $Y < 2$. These four parameters give the CGMY model great flexibility and allow the model to be either of finite or infinite activity and variation.

1.2.4. Pricing and the (Parabolic) Partial Integro-Differential Equation. The mathematical formulation of the pricing of an European option leads to a PDE in the Gaussian case and a PIDE in the Lévy case. As our main focus is on Lévy markets and the fact that the Gaussian case is well expounded upon, we review how a PIDE arises from the pricing of an European option. The value of an option, C_t, is defined to be the conditional expectation of its payoff, discounted under the pricing measure:

$$C_t = \mathbb{E}[e^{-r(T-t)} H_T | \mathcal{F}_t]$$

where H_T is the payoff of the option at maturity and $\mathcal{F}_t$ is the associated filtration of the underlying probability space. We can rewrite this in a more discrete form using the Markov property of Lévy processes:

$$C(t, S) = \mathbb{E}[e^{-r(T-t)} H_T(S_T) | S_t = S]$$

If C is sufficiently smooth (i.e, $u \in C^{1,2}$ with bounded derivatives), it is well known that we can obtain a Black-Scholes-like PIDE for the option price by applying standard no-arbitrage arguments [30]:

$$\frac{\partial C}{\partial t}(t, S) + rS\frac{\partial C}{\partial S}(t, S) + \frac{\sigma^2 S^2}{2}\frac{\partial^2 C}{\partial S^2}(t, S) - rC(t, S) \tag{1.12}$$

$$+ \int v(\mathrm{d}x)\Big[C(t, Se^x) - C(t, S) - S(e^x - 1)\frac{\partial C}{\partial S}(t, S)\Big] = 0$$

Equation (1.12) reduces to the Black-Scholes equation in the Gaussian case for which the last integral in (1.12) is zero.

2. Grid-based Methods. Mesh based methods (the Trinity being the well known Finite Difference/Element/Volume Methods), the solution method of choice for many numerical problems in other computational fields, have also taken root in computational finance. Although both Finite Difference (FD) and Finite Element (FE) methods can be applied to a broad swath of problems, FD methods are most commonly applied to simple problems (payoffs) in low dimensions or as the time discretization scheme in mixed schemes. FE methods, on the other hand, are more commonly applied to more complex problems in high dimensions, or for problems where a high accuracy is needed. The cost of the computational benefits of FE schemes is a more involved implementation. Finite volume (FV) methods have also been applied in computational finance but are less popular and widespread. This may be because most problems in finance are inherently parabolic, whereas FV methods are usually applied to hyperbolic problems (conservation laws). In this section, attention is largely devoted to the pricing of American options in order to highlight methods that are centered around a PDE approach. As explained later, the transform approach is ideally suited to (pricing) problems of European exercise, whereas the PDE approach is better suited to problems with American exercise features or path dependency. Indeed, for options of European exercise, PDE-based methods are, in general, not competitive with transform based methods. Furthermore, American options are ubiquitous on exchanges and as a contract feature (consider convertible bonds and other derivative contracts with embedded optionality, exercisable up to a certain date).

Outline of current and related research. Before further exposition, we provide a brief overview of current research in grid based methods for Lévy models in hopes of illuminating the current research landscape. Finite Difference methods have been proposed in [6, 29, 35, 58, 95]. Andersen and Andreasen [6] were one of the first to solve, numerically, classic pricing problems in Lévy models and suggested an operator splitting method based on the Crank-Nicholson method, treating the jump integral with an explicit method. However, this method does not extend to problems having singular kernels (infinite activity Lévy models). d'Halluin et al. [35] tackle the LCP (arising from the formulation of the American option problem in a jump-diffusion setting, which is discussed later) with an implicit discretization, treating the jump integral via a FFT. We remark that, in the literature, it is popular to handle the jump integral via the FFT due to its numerical efficiency. Zhang [95] treats the PIDE in a variational setting, devising a FD scheme for jump diffusion with finite intensity possessing all exponential moments. Although treated in a variational framework, the exponential moment condition severely limits the proposed method to the Merton model. Cont et al. [29] propose a more general FD scheme under the framework of viscosity solutions, which allows for rigorous analysis of stability, convergence, and estimates of the rate of convergence to be undertaken. Arguably, when applicable, analysis of the PIDE should be done in a viscosity solution framework as *all* viscosity solutions are stable. Viscosity solutions can also be seen to extend the notion of weak solutions for nonlinear settings in which it provides a nice regularity theory [32]. Whole families of methods have been developed for solving the basic Linear Complementarity Problem arising from the American option pricing problem and as we devote a section to discussing this, we defer to the following section for an overview of proposed solution methods. Penalty methods are known to be applicable to LCPs

[8] and are popular as they allow for a single homogeneous method to be applied to single and multidimensional problems [35, 46]. Furthermore, they are also able to easily include various generalizations (stochastic volatility, transactions costs, etc) in one framework.

Another exciting approach to pricing problems is that of the HJB/stochastic optimal control perspective [45, 52, 73, 92]. This new class of methods arise from traditional pricing problems when more realistic features are modelled, such as jumps in the underlying or market frictions such as transactions costs. Though as we have seen there are far simpler ways to deal with jumps, the control theoretic framework lends itself well when modelling various market frictions. At this point, it is pertinent to mention that, although market frictions such as transaction costs are of great practical relevance, in both academia and industry, they are rarely dealt with rigorously, if at all, because of the added complexity required to model them. Of course, this is changing as the field matures. One of the main attractive features of the control theoretic perspective is that it can exploit the fact that many problems in such settings can be phrased as optimization problems, thereby allowing one to take advantage of the advances in the field of (mathematical) optimization. We note however, that it is not rare for specialized solution methods to be devised for a particular re-framed optimization problem [45].

Finally, we must mention the class of methods based on the Galerkin framework, ranging from finite element methods [3, 4], to spectral element methods [25, 47], to the exciting new Wavelet methods [65, 66, 67, 85]. Galerkin based methods are more sophisticated and more difficult to implement compared to other grid based methods such as Finite Differences, however, they are more amenable to problems of higher dimensions and, in general, can be more efficient when requiring greater accuracy. Furthermore, as is well known, the Galerkin method also provides an a priori framework for the error, which is of great importance in computational finance. In more practical considerations, Galerkin methods can lead to a more homogeneous code base which is of great importance when maintaining production codes. We devote a section to the more detailed exposition of Wavelet methods as they are an exciting subclass of Galerkin methods, owing to such properties as certain integral operators being sparse when represented in the wavelet basis.

2.1. American Options and Linear Complementarity Problems. As briefly mentioned above, in the study of numerical methods for option pricing problems in Lévy markets, grid based methods in most cases, for European options, are not competitive with transform based methods. As most of the methods reviewed in this section are based on the American option problem, we first review the standard variational formulation of the American option problem. In general, grid based methods rely on an equivalent variational formulation of a PDE. In the case of the American put problem, the price satisfies a parabolic variational inequality. When discretized,

this results in a Linear Complementarity Problem (LCP) [91]:

$$\mathcal{L}v \geq 0 \qquad s > 0, \quad t \in [0,T] \tag{2.1a}$$
$$(v - H) \geq 0 \qquad s > 0, \quad t \in [0,T] \tag{2.1b}$$
$$(\mathcal{L}v)(v - H) = 0 \qquad s > 0, \quad t \in [0,T] \tag{2.1c}$$
$$v_t - (\frac{1}{2}\sigma^2 s^2 v_{ss} + rsv_s - rv) = \mathcal{L}v \qquad s > 0, \quad t \in [0,T] \tag{2.1d}$$
$$v = K \qquad s = 0, \quad t \in [0,T] \tag{2.1e}$$
$$v = H \qquad s > 0, \quad t = T \tag{2.1f}$$
$$v \to 0 \qquad s \to \infty, \quad t \in [0,T] \tag{2.1g}$$

where s is the spot price, $H = \max(K - s, 0)$ is the payoff, with K being the strike, v is the value of the American option, and $\mathcal{L}$ is the BS (linear parabolic differential) operator. Though in (2.1) above we have defined $\mathcal{L}$ to be the BS operator, the general method of numerically solving the American Option problem can be extended to stochastic volatility models by defining $\mathcal{L}$ to be the Heston operator. The same holds in a Lévy setting, where $\mathcal{L}$ then represents the corresponding Lévy parabolic integro-differential operator [29]. Many existing numerical methods in the literature explore different ways to solve the LCP, some of which include Projected Successive Over Relaxation (PSOR) [91], Operator Splitting [54], Projected Multigrid Methods [70], Penalty Methods [46], Policy Iteration [73], and Component Wise Splitting Methods [55]. Other approaches worth mentioning for solving the American option problem include solving for the Free Boundary (via front-fixing) [68] and relatively new techniques such as the use of Malliavin calculus [1]. LCPs are important classes of numerical problems in the computational sciences, in particular, in computational mechanics. LCPs have been found in the symmetric case to be equivalent to bound-constrained quadratic programming (BQPs) problems. The LCP specific numerical methods mentioned above are long established methods, though, as of late, there has been a rekindling of interest in numerical methods for LCPs. This stems from new techniques of subspace acceleration that allow for more complex splittings and improved identification of the active set [44, 74] which includes application to the American option problem.

2.2. Operator Splitting. Introduced by Ikonen and Toivanen [54] to finance, the operator splitting method originated in fluid mechanics where it is used to handle the incompressibility constraint [2]. The main idea is in splitting the operators arising in the LCP into fractional time steps as opposed to solving them simultaneously in one time step. This results in a series of simpler problems during each time step, as opposed to directly solving the LCP. Applied to the option pricing problem, this amounts to solving for the equation associated with the BS operator in a fractional time step, and then imposing the constraint (2.2) in a second fractional time step. An equivalent way of looking at this is that the equation associated with the BS operator is solved first and then the solution is adjusted in a second fractional time step so that the constraint is satisfied. To begin, we first rewrite the LCP (2.1) in a slightly different but equivalent form where we introduce an auxiliary variable to represent the BS operator and make slight notional changes more appropriate for this method:

$$\lambda = v_t + \frac{1}{2}\sigma^2 s^2 v_{ss} + rsv_s - rv \qquad s > 0, \quad t \in [0,T] \tag{2.2a}$$

$$[v - (K-s)^+] \cdot \lambda = 0 \qquad s > 0, \quad t \in [0,T] \tag{2.2b}$$

$$v - (K-s)^+ \geq 0, \quad \lambda \geq 0 \qquad s > 0, \quad t \in [0,T] \tag{2.2c}$$

$$v = \max(K-s, 0) \qquad s > 0, \quad t = T \tag{2.2d}$$

$$v = K \qquad s = 0, \quad t \in [0,T] \tag{2.2e}$$

$$v \to 0 \qquad s \to \infty, \quad t \in [0,T] \tag{2.2f}$$

2.2.1. Space Discretization. To begin, we introduce the standard finite difference discretization schemes for the space variable - the standard central difference scheme. Note that the subscript i refers to the ith point in the spatial discretization

$$\frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} \approx \sigma^2 s_i^2 \frac{(v_{i-1} - 2v_i + v_{i+1})}{2(\Delta s)^2} \tag{2.3}$$

$$rs\frac{\partial v}{\partial s} \approx rs_i \frac{(v_{i+1} - v_{i-1})}{2\Delta s} \tag{2.4}$$

Substituting the above discretizations into (2.2a), we arrive at the semidiscrete equation

$$\frac{\partial v}{\partial t} + Av - \lambda = 0 \tag{2.5}$$

where the matrix A contains the above space discretizations (and $-rv$).

2.2.2. Time Discretization. Next, we discretize (2.5) in time. In the literature, it is common to discretize the time variable via a θ-scheme given as follows

$$\frac{1}{\Delta t}\big(v^{(k+1)} - v^{(k)}\big) + A\big((1-\theta)v^{(k+1)} + \theta v^{(k)}\big) - \lambda^{(k)} = 0 \tag{2.6}$$

where superscript k denotes the kth time level. Note that we are integrating backward in time and that v^k, λ^k are implicit. Another popular time discretization method is the two-step backward differentiation formula (BDF) [70]:

$$\frac{1}{\Delta t}\left(\frac{4}{3}v^{(k+1)} - v^{(k)} - \frac{1}{3}v^{(k+2)}\right) + \frac{2}{3}\left(Av^{(k)} - \lambda^{(k)}\right) = 0 \tag{2.7}$$

The remaining constraints from the LCP are discretized in a straightforward and consistent manner with the discretizations introduced above and are as follows:

$$\left[v_i^{(k)} - (K - s_i)^+\right] \cdot \lambda_i^{(k)} = 0, v_i^{(k)} \geq (K - s_i)^+, \lambda_i^{(k)} \geq 0. \tag{2.8}$$

For the reminder of this section, we do not explicitly mention (2.8), the remaining discretized LCP conditions and assume that they are satisfied.

2.2.3. Operator Splitting. Having to solve (2.6) or (2.7) simultaneously with (2.8) is difficult; indeed this comes from the simple fact that constructing solutions to linear problems is easier than constructing solutions to LCPs. As previously mentioned, the general strategy of operator splitting methods is to decouple the LCP into simpler linear problems via splitting the time discretization into two fractional time steps. The first fractional time step for both methods (the θ-scheme and the BDF, respectively) are, respectively,

$$\frac{1}{\Delta t}(v^{(k+1)} - \tilde{v}^{(k)}) + A\bigg((1-\theta)v^{(k+1)} + \theta\tilde{v}^{(k)}\bigg) - \lambda^{(k+1)} = 0 \tag{2.9}$$

$$\frac{1}{\Delta t}(\frac{4}{3}v^{(k+1)} - \tilde{v}^{(k)} - \frac{1}{3}v^{(k+2)}) + \frac{2}{3}(A\tilde{v}^{(k)} - \lambda^{(k+1)}) = 0 \tag{2.10}$$

The $\tilde{v}^{(k)}$ denotes the intermediate value that links the two fractional time steps and $\lambda^{(k+1)}$ is known from the previous time step. Note that in the first fractional time step, all the terms at the $k+1$ and $k+2$ time steps are known and that only $\tilde{v}^{(k)}$ is solved for (since we are going backwards in time). Thus, $\tilde{v}^{(k)}$ can be computed via an LU-decomposition. The second fractional time steps are

$$\frac{1}{\Delta t}(\tilde{v}^{(k)} - v^{(k)}) + \lambda^{(k+1)} - \lambda^{(k)} = 0 \tag{2.11}$$

$$\frac{1}{\Delta t}(\tilde{v}^{(k)} - v^{(k)}) + \frac{2}{3}(\lambda^{(k+1)} - \lambda^{(k)}) = 0 \tag{2.12}$$

for the θ-scheme and the BDF, respectively. It is in the second fractional time step that the additional LCP conditions (2.8) are incorporated into the method. Indeed it is those equations together with (2.11) or (2.12) that allows us to compute $v^{(k)}$ and $\lambda^{(k)}$. Note that $v^{(k)}$ and $\lambda^{(k)}$ are decoupled from the other components and can be solved for algebraically, as we would already have obtained the values of $\tilde{v}^{(k)}$ from the first fractional time step, $v^{(k)}$ and $\lambda^{(k)}$ are decoupled because of the fact that they must satisfy a complementarity condition. One of the two must be known at each iteration. For example, if it were the case that $\lambda^{(k)} = 0$, the complementarity condition implies then that $v_i^{(k)} \geq (K - s_i)^+$. More importantly, the consequence of the case that $\lambda^{(k)} = 0$ results in the second fractional time steps being unknown in only one variable. In the BDF-scheme, this allows us to solve directly for $v^{(k)}$:

$$v^{(k)} = \tilde{v}^{(k)} + \frac{2}{3}\lambda^{(k+1)}\Delta t \tag{2.13}$$

Similarly, if it is the case that $v_i^{(k)} = (K - s_i)^+$ (i.e., $\lambda^{(k)} \geq 0$), we can use the complementarity condition to arrive at a solution for $\lambda^{(k)}$. However, this has only shown that $\lambda^{(k)}$ and $v^{(k)}$ are decoupled and that it is possible to use additional information from the complementarity conditions in order to arrive at a solution to the second fractional time step. The original paper does not discuss why the method works. It can be shown, although out of the scope of this survey, that the method proposed is equivalent to the Douglas-Rachford algorithm. In order to show that this method does indeed work, one would need to first reformulate the LCP into a form consistent with the Douglas-Rachford scheme. The resulting form has two operators which must be shown to be maximal monotone operators in order to

use the standard result to prove that the scheme converges [71]. Note that maximal monotone operators stem from the theory on proximal algorithms, specifically, the underlying proximal operators. Some theory on the relation between the Douglas-Rachford splitting algorithm and proximal algorithms can be found in [40].

2.2.4. Numerical Results and Convergence. Numerical experiments show that the operator splitting method is approximately twice as efficient as the PSOR method; the CPU time for PSOR was shown to increase by a factor of 8 when the discretization (both space and time) was made twice as fine, whereas the Operator Splitting method only increased by a factor of 4 [54]. With regards to accuracy, it is shown that the operator splitting method is essentially the same as the PSOR method. Although operator splitting methods are known to be efficient, one disadvantage of this class of methods is a lack of convergence results for each possible splitting scheme: convergence analysis has to be done from scratch. Moreover, as the operator splitting method, fundamentally, looks *not* to solve the discretized LCP, any convergence results for the continuous LCP cannot be exploited.

2.2.5. Extensions. The Operator Splitting method, like many other methods, has been extended to more general cases such as stochastic volatility [55], where the A matrix in (2.5) is decomposed into three simpler matrices denoted A_x, A_{xy}, A_y. This component-wise method is also known as the Marchuk-Yanekno scheme. Similar to the method of Policy Iteration, to be discussed in the next section, one of the most important considerations in the analysis of the method is the properties of the A matrix. For example, if the A matrix has properties, such as being a Z-matrix, then the method may have nice convergence results. Following this line of thought, Huang et al. [53] discuss some properties of the matrices resulting from the discretization of the LCP arising from models such as the Cox's CEV and a jump-diffusion model. A more general and rigorous discussion of the properties of the matrices arising from LCPs can be found in [31].

2.3. Policy Iteration. Most of the methods discussed in this paper are traditional numerical methods, i.e. grid based methods or numerical integration methods. In this section, we describe a numerical method following the recent trend of framing and solving problems from an optimization standpoint. In a recent paper [73], the authors demonstrate the use of policy iteration as a means of pricing American options. It is shown that the Policy Iteration can be seen as a standard Penalty Method [92] with the regularization parameter approaching infinity. The key insight that permits tackling the problem from an optimization standpoint is the recognization that the LCP can be expressed as a discrete HJB problem. Policy Iteration is a numerical method designed for solving HJB equations that was first introduced to finance in [45]. Policy Iteration was originally devised by Bellman [13, 14] and Howard [51] for the solution of stationary infinite horizon Markovian Dynamic Programming (MDP) problems. Central to each MDP is the Bellman equation in the discrete case or the Hamilton-Jacobi-Bellman equation in the continuous case which is the dynamic programming equation describing the cost of the dynamical system. The dynamic programming equation, also known as a value function, essentially measures the cost of each stage (payoff) so that one may choose, usually, the path that results in the lowest cost. We refer to the standard treatises [13, 14, 51] for theoretical background and [57] for a more numerical treatment, as even a quick review is beyond the scope of this paper. It is worthwhile to mention that it has been shown, under specific regularity conditions, that Policy Iteration is mathematically equivalent to Newton's

Method [72].

The policy iteration algorithm used in this paper is adapted from the original [45] to the case of the discretized LCP arising from the American option problem. The resulting discretized linear complementarity problem takes the form

$$\begin{aligned} Ax &\geq b \\ x &\geq c \\ (Ax-b)_i \cdot (x-c)_i &= 0 \qquad 1 \leq i \leq N \end{aligned} \tag{2.14}$$

where $A \in \mathbb{R}^{N\times N}$ is an M-matrix and $b, c \in \mathbb{R}^N$. Recall that an M-matrix is a nonsingular matrix A such that $A^{-1} \geq 0$ [50]. It is clear that (2.14) is equivalent to the following optimization problem

$$\min\{Ax-b, x-c\} = 0 \tag{2.15}$$

also known as the *obstacle problem*, a discrete HJB equation. The obstacle problem is a classic problem in variational inequalities and free boundary problems, where the goal is to find the equilibrium position of an elastic membrane constrained to lie above a certain obstacle [8]. Note that the optimization problem (2.15) is computed in a component-wise fashion.

2.3.1. Application to American Options. The application of Policy Iteration to American options in [73] is predicated on the equivalence of (2.14) and (2.15). The LCP arising from the American option problem is reformulated as a discrete HJB problem that can be solved via the Policy Iteration adapted to this specific case. The component-wise form of (2.15) is

$$\min_{\phi\in\{0,1\}} \{\phi(Ax-b)_i + (1-\phi)(x-c)_i\} = 0 \tag{2.16}$$

Equivalently, this amounts to considering

$$(\phi_i(Ax)_i + (1-\phi_i)(x)_i) - (\phi_i(b)_i + (1-\phi_i)(c)_i) = 0 \qquad 1 \leq i \leq N$$

where $\phi_i \in \arg\min_{\phi\in\{0,1\}}\{\phi(Ax-b)_i + (1-\phi)(x-c)_i\}$ represents the optimal policy in state i. The algorithm amounts to point-wise checks and solving the corresponding equations which violate the equality most. Formally, the problem specific Policy Iteration algorithm is

Policy Iteration Algorithm applied to the American Option Problem. Let $x^0 \in \mathbb{R}^N$. Given x^n, let $\phi^n \in \mathbb{R}^N$, $A \in \mathbb{R}^{N\times N}$, $b \in \mathbb{R}^N$ be such that for $1 \leq i \leq N$

$$\begin{aligned} (\phi^n)_i \in \arg\min_{\phi\in\{0,1\}}\{\phi(Ax^n-b)_i + (1-\phi)(x^n-c)_i\} \\ (A^n)_i = (\phi^n)_i(A)_i + (1-(\phi^n)_i)(I_N)_i \\ (b^n)_i = (\phi^n)_i(b)_i + (1-(\phi^n)_i)(c)_i \end{aligned}$$

where $(A^n)_i$ and $(I_N)_i$ stand for the ith row of A^n or I_N, respectively, and I_N is the $N \times N$ identity matrix.

It then follows that

$$(A^n x^n - b^n)_i = \min_{\phi\in\{0,1\}}\{\phi(Ax^n-b)_i + (1-\phi)(x^n-c)_i\}$$

Each step of the algorithm consists of correcting the row that violates the inequality the most. It is shown that this algorithm terminates in finitely many steps. In particular, the policy iteration converges in at most $N+1$ steps for the vanilla American option payoff [73]. In practice, however, the algorithm often converges in much fewer iterations. Recall that the policy iteration steps through via finding improved policies; if the policy cannot be improved further, then the algorithm terminates.

2.3.2. Policy Iteration in the Lévy Case. Since the method mainly relies on the M-matrix property of the operator matrix, Policy Iteration can be applied to Lévy (and other) models *provided* that their discretizations satisfy the M-matrix property. At the time of this writing, there have not been explicit results published on this; it could be a possible research topic.

2.4. Wavelet Galerkin Methods. Galerkin methods are a mainstay tool of the numerical analyst and one of the main solution methods of choice for the numerical solution of PDEs in computational mathematics due to such properties as their a posteriori error analysis and mesh adaptivity. The Galerkin method has been applied to finance in the BS [3] and the stochastic volatility cases [4]. As mentioned earlier, Galerkin methods are, in general, more complex than simpler discretization methods such as Finite Differences. However, Galerkin methods are more amenable to more exotic payoff features and higher dimensions when compared to less sophisticated discretizations. Analogous to the finite element literature in other fields, other flavours of Galerkin methods have also been applied to finance [25, 47]. Our focus is on the wavelet Galerkin method (known simply in the literature as wavelet methods) due to properties such as wavelet compression and optimal preconditioning[1], which allow it to handle features characteristic of PIDEs, like integral operators. The particular method that we discuss discretizes the spatial dimension via a Wavelet-based Galerkin method, whereas the time dimension is discretized via a discontinuous Galerkin (dG) approach. The resulting linear systems are then solved by the incomplete GMRES.

2.4.1. Background. Wavelets originated in signal processing and image analysis and have traditionally been mainly applied to those fields. Recently, interest in their applicability to the numerical solution of PDEs has surged. One major reason for this is that PDEs describing physical phenomena exhibit multiscale behaviour. Naturally, numerical methods such as multigrid methods which are able to exploit this are quite successful. It is well known that multigrid methods take advantage of mesh refinements, and in that sense can be considered asymptotically optimal. Wavelets inherently describe components living in different scales, thereby naturally possessing an important property which make multigrid methods so effective. Another attractive property of wavelets is that many operators (not just integral operators) and their inverses are sparse when represented in wavelet coordinates. Wavelet analysis in many cases parallels Fourier analysis and indeed many of the results obtained in the Fourier case have analogous results in the wavelet case. Due to the considerable depth of the analysis underlying wavelet methods, we highlight the basics of wavelets so that our review of wavelet based Galerkin methods will be easier to understand.

2.4.2. Preliminaries. The analysis of wavelets is a broad field and, as previously mentioned, many results parallel that of Fourier analysis which is an even older and broader field, with many ties to Harmonic analysis. Wavelets are functions that are inherently multiscale and consequently hierarchical. This is the basis

[1] in a sense to be discussed

for many of their important properties which we summarize below. To this end, we define a space V_h as the space of piecewise polynomials of total degree $p \geq 0$ vanishing on the boundary $\partial\Omega$. V_h is the finite element space that we will be working with. The hierarchical basis functions ψ_j^l (wavelets) possess the following properties

1. $V_h = \text{span}\,\{\psi_j^l \,|\, 0 \leq l \leq L, 1 \leq j \leq M^l\}$
2. The wavelets have support $S_j^l = \text{supp}\ \psi_j^l$ with $\text{diam}(S_j^l) \leq C2^{-l}$
3. Wavelets such that $\bar{S}_j^l \cap \partial\Omega = \emptyset$ have moments that vanish up to order p; $(\psi_j^l, \text{q}) = 0$ for all polynomials q of total degree p or less.
4. ψ_j^l, where $l \geq l_0$ are scalar transformations and scalings of the wavelet $\psi_j^{l_0}$
5. A function $v \in V_h$ has the decomposition

$$v = \sum_{l=0}^{L} \sum_{j=1}^{M^l} v_j^l \psi_j^l$$

   where $v_j^l = (v, \bar{\psi}_j^l)$. $\bar{\psi}_j^l$ are known as dual wavelets.
6. The following norm equivalence holds. There exist constants $c_1, c_2 > 0$ such that

$$c_1 \|v\|^2_{\tilde{H}^s(\Omega)} \leq \sum_{l=0}^{\infty} \sum_{j=0}^{M^l} |v_j^l|^2 \, 2^{2ls} \leq c_2 \|v\|^2_{\tilde{H}^s(\Omega)}$$

The aforementioned properties are the basic machinery behind the compression property and efficiency of wavelets. Further properties and analysis can be found in the standard texts for wavelets for numerical analysis: Cohen [26] and Daubechie [34]. Standard references in the literature include [27, 33, 80, 81, 82].

2.4.3. Variational Formulation. Due to constraints on the scope of this report, we focus the bulk of the discussion on the European case. We discuss the American case, albeit in lesser detail, afterwards. It is, by now, well known that the European option pricing problem results in the following PIDE:

$$\frac{\partial u}{\partial \tau}(\tau, x) + \mathcal{L}^X(u)(\tau, x) = 0 \quad \text{in } (0, T) \times \mathbb{R} \tag{2.17}$$

$$u(0, x) = g(x) \quad \text{in } \mathbb{R} \tag{2.18}$$

where $g(x) = f(e^x)$ is the payoff function in log-space. Note that the PIDE above is in log-space $x = \log(S)$, and time to maturity, τ. The operator $\mathcal{L}^X$ is, more specifically, $\mathcal{L}^X = \mathcal{L}^{\text{BS}} + \mathcal{L}^{\text{jump}}$ [66] with

$$\mathcal{L}^{BS}[v] = -\frac{\sigma^2}{2} v'' + \left(\frac{\sigma^2}{2} - r\right) v' + rv$$

$$\mathcal{L}^{jump}[v] = -\int_{\mathbb{R}} \big(v(x - y) - v(x) - (e^y - 1)v'(x)\big) k(y) dy$$

The infinitesimal generator $\mathcal{L}^X$ of the Lévy process X_t is split into the BS component $\mathcal{L}^{BS}$ and the jump component $\mathcal{L}^{jump}$. The payoff function may grow exponentially as $|x| \to \infty$ (i.e, $(e^x - K)^+, (K - e^x)^+ \notin L^2(\mathbb{R})$). As such, instead of classical Sobolev spaces, as in the Brownian case, exponentially weighted Sobolev spaces are most

often used. To this end, we introduce the exponentially weighted Sobolev spaces with exponent w:

$$H^1_{\pm w}(\mathbb{R}) := \{v \in L^1_{\text{loc}}(\mathbb{R}) |\, e^{\pm w} v, e^{\pm w} v' \in L^2(\mathbb{R})\} \tag{2.19}$$

where

$$w(x) = \begin{cases} w_- |x| & \text{if} \quad x < 0 \\ w_+ |x| & \text{if} \quad x > 0 \end{cases}$$

and $w \in L^1_{\text{loc}}(\mathbb{R})$, $w' \in L^\infty(\mathbb{R})$. As in the standard variational analysis needed for Galerkin methods, we introduce a bilinear form $a^\eta(\cdot, \cdot)$ corresponding to our operator $\mathcal{L}^X$:

$$a^{\pm w}(\varphi, \psi) := \int_{\mathbb{R}} \mathcal{L}^X[\varphi](x)\psi(x)e^{\pm 2w(x)}dx \qquad \varphi, \psi \in C_0^\infty(\mathbb{R}) \tag{2.20}$$

Under some conditions, the bilinear form $a^w()$ can be shown to satisfy a Gårding inequality. Further assuming that the exponent w has $0 \leq w_- < M$ and $0 < w_+ < G$, where M and G are the constants from the Lévy density (1.8), the solution $u \in L^2((0,T); H^1_{-w}(\mathbb{R})) \cap H^1((0,T); H^1_{-w}(\mathbb{R}))$ of

$$\frac{\partial}{\partial \tau}(u(\tau), v)_{L^2_{-w}(\mathbb{R})} + a^{-w}(u(t), v) = 0 \quad \forall v \in H^1_{-w}(\mathbb{R}) \tag{2.21}$$

$$u(0) = g \tag{2.22}$$

exists and is unique [67].

2.4.4. Localization. The variational formulation has domain on the whole real line. Clearly, this is not feasible for numerical solutions. As such, the variational formulation must be localized to a suitable computational domain $\Omega_R = (-R, R)$, so that a computer solution is feasible. A step towards obtaining an equivalent localized problem is by the introduction of what is referred to as the *excess to payoff* function.

$$U = u - g \tag{2.23}$$

This excess to payoff function is introduced because it decays exponentially. Hence, the corresponding PIDE for the excess to payoff function can be localized as in more classical cases and we recover again a local property. Restricting the solution of the PIDE for U to our computational domain Ω_R, we obtain

$$\frac{\partial U_R}{\partial \tau} + \mathcal{L}^X_R[U_R] = -\mathcal{L}^X_R[g] \quad \text{in} \quad (0,T) \times \Omega_R \tag{2.24}$$

$$U_R(\tau, x) = 0 \quad \text{in} \quad (0,T) \times \mathbb{R} \setminus \Omega_R \tag{2.25}$$

$$U_R(0, x) = 0 \quad \text{in} \quad \Omega_R \tag{2.26}$$

It can be shown that there does exist a solution for the PIDE above and that it is unique. We refer to [66, 67] for details of the verification.

2.4.5. Discretization and Numerical Solution. The exposition in the previous section dealt mostly with the solution of the PIDE in a variational setting. In this section we discuss the particulars of the discretization. We mention that, in general, the discretization of nonlocal operators such as $\mathcal{L}^{\text{jump}}$ leads to dense and

ill-conditioned stiffness matrices. This undoubtedly increases the computational complexity of the numerical solution. It is for this reason that wavelet bases are so attractive - their properties such as matrix compressions result in the matrix being *sparse* and in most cases, better conditioned.

We begin by introducing the general functional space necessary for the analysis of the discretization. For $0 \leq \rho \leq 2$ the spaces

$$V = \tilde{H}^{\rho/2}(\Omega) \begin{cases} H^{\rho/2}(\Omega) & \text{if} \quad 0 \leq \rho < 1 \\ H_{00}^{1/2}(\Omega) & \text{if} \quad \rho = 1 \\ H_0^{\rho/2}(\Omega) & \text{if} \quad 1 < \rho \leq 2 \end{cases} \tag{2.27}$$

defined on a bounded interval $\Omega = (a, b)$ denote the standard Sobolev Spaces [5]. Let V^* denote the dual of V. We remark that the analysis of the discretization and the problem form required for numerical solution is carried out on standard Sobolev Spaces. This is because the bilinear form $a(\cdot, \cdot)$ can be extended continuously into $V \times V$. We do not specify which particular space V refers to as the analysis can proceed without specifics. The specific cases of V exists to deal with certain cases of the PIDE. This is due to the fact that as the volatility of the process changes, the order of the operator changes accordingly. Details of the continuous extension and the cases of the PIDE in different volatility regimes are presented in [66].

Basis Functions. The wavelet functions chosen as a basis are the biorthgonal wavelets

$$\{\varphi_j^l\}_{j,l} \qquad l = 0, 1, .., L \quad j = 1, 2, ..., M \tag{2.28}$$

whose general properties have been discussed above.

Matrix Compression is a deep topic in itself (see [78] for more advanced compression methods), but given that the problem is semidiscretized in space via wavelets, the matrix compression for our bilinear form a on $V_h \times V_h$ is

$$\tilde{A}_{(j,l),(j',l')} := \begin{cases} A_{(j,l),(j',l')} & \text{if } \operatorname{dist}(S_j^l, S_{j'}^{l'}) \leq \delta_{l,l'} \text{ or } S_j^l \cap \partial\Omega \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{2.29}$$

where the truncation parameter is $\delta_{l,l'} = c \max\{2^{-L+\hat{\alpha}(2L-l-l')}, 2^{-l}, 2^{-l'}\}$ with parameters $c > 0$, $0 < \hat{\alpha} \leq 1$, and $S_j^l = \operatorname{supp} \psi_j^l$.

Domain Discretization. The problem domain is discretized as follows. For the spatial dimension, $\mathcal{T}^0$ is introduced as a fixed coarse partition of Ω_R. Then, for $l > 0$, $\mathcal{T}^l$ is defined to be a bisection of each of the intervals in $\mathcal{T}^{l-1}$. Clearly, we have discretized the spatial domain in a manner amenable to the inherent multiscale properties of the wavelet basis functions. Let $\mathcal{T}_h = \mathcal{T}^L$ for some $L > 0$ such that $h = C2^{-L}$. V_h is then defined to be the space of all continuous piecewise polynomial functions vanishing on the $\partial\Omega$ of degree $p \geq 0$ ($p \geq 1$ in the case where the operator is of order $1 \leq \rho \leq 2$). We can then arrive at a hierarchy of spaces based on the discretization (triangulation) $V^0 \subset V^1 \subset \cdots V^L = V_h$ with each space V^l corresponding to the partition $\mathcal{T}^l$. Let $N^l = \dim V^l$ and $N = \dim V_h = N^L = C2^L$.

Problem Discretization. The resulting spatial semidiscrete problem is

$$\left(\frac{d}{dt}\tilde{u}_h, v_h\right) + \tilde{a}(\tilde{u}_h, v_h) = 0 \quad \forall v_h \in V_h \tag{2.30}$$

$$\tilde{u}_h(0) = P_{L^2}^{V_h} u_0 \tag{2.31}$$

where $P_{L^2}^{V_h}$ denotes the (L^2) projection of onto V_h

The time dimension is discretized via $\mathcal{M} = \{I_m\}_{m=1}^M$, a partition into M subintervals $I_m = (t_{m-1}, t_m)$ of the time domain $(0,T)$ where $m = 1, 2, .., M$ and $0 = t_0 < t_1 < \cdots < t_M = T$. The functional spaces used for the dG method are

$$\mathcal{S}^{\underline{r}}(\mathcal{M}, V_h) := \{u \in L^2((0,T), V_h) : u|_{I_m} \in \mathcal{P}_{r_m}(I_m, V_h),\ m = 1, 2, ., , M\} \tag{2.32}$$

where $\mathcal{P}_{r_m}(I_m, V_h)$ denotes the space of polynomials of degree $\underline{r}_m$, at most, on I_m. The time domain is discretized via the dG method, which is fairly non-standard; see [36, 49] for standard references on the dG method. The fully discretized problem is that of finding $\tilde{U}_h^{dG} \in \mathcal{P}_{r_n}(I_n, V_h)$, $1 \le n \le M$

$$\int_{I_n} \big((\partial_t \tilde{U}_h^{dG}, W) + \tilde{a}(\tilde{U}_h^{dG}, W)\big)dt + (\tilde{U}_{n-1}^{dG+}, W_{n-1}^{+}) = (U_{n-1}^{dG-}, W_{n-1}^{+}) \tag{2.33}$$

for all $W \in \mathcal{P}_{r_n}(I_n, V_h)$, with $U_0^{dG-} = P_{L^2}^{V_h} u_0$. Note that the '-' superscript denotes a one-sided limit (i.e., $u_m^- := \lim_{s \to 0^+} u(t_m - s)$). This can be interpreted as a time stepping scheme of size $t_m - t_{m-1}$ of order r_m. The resulting linear systems that must be solved at each time step $m = 1, 2, ..., M$ are of size $(r_m + 1)N^L$.

Linear System. In matrix form, (2.33) becomes

$$\mathbf{R}u = f \quad \text{with} \quad \mathbf{R} = \mathbf{C} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \tilde{\mathbf{A}} \tag{2.34}$$

$$C_{ij} = \sigma_{ij}\sqrt{(i+0.5)(j+0.5)} \quad \sigma_{ij} = \begin{cases} (-1)^{i+j} & j > 1 \\ 1 & \text{otherwise} \end{cases} \tag{2.35}$$

where $\mathbf{M}$ is the mass matrix corresponding to the simple inner product $(\cdot, \cdot)$ and $\tilde{\mathbf{A}}$ is the compressed stiffness matrix corresponding to our bilinear form $\tilde{a}(\cdot, \cdot)$.

Decoupling. Although (2.34) can be solved as a single linear system of size $(r_m + 1)N^L$, it is decoupled into r+1 linear systems of size N^L with the help of a Schur decomposition. Though not discussed nor mentioned in the original paper, the decoupled systems are amenable to solutions via parallel methods. Indeed, such an investigation could be basis for future work. A Schur decomposition of $\mathbf{C} \in \mathbb{C}^{(r+1)\times(r+1)}$ results in $\mathbf{C} = \mathbf{Q}\mathbf{T}\mathbf{Q}^H$ where $\mathbf{T}$ is an upper triangular matrix having the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{r+1}$ of $\mathbf{C}$ on its diagonal and $\mathbf{Q}$ is a unitary matrix. Left multiplying (2.34) by $(\mathbf{Q}^H \otimes \mathbf{I})$ results in

$$\left(\mathbf{T} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \tilde{\mathbf{A}}\right) w = g, \quad w = (\mathbf{Q}^H \otimes \mathbf{I})u \quad g = (\mathbf{Q}^H \otimes \mathbf{I})f \tag{2.36}$$

Complexity. The most evident advantage of working in a wavelet basis is that the representation of the operator can be compressed via matrix compression. Complexity-wise, this means that an otherwise dense matrix representation of a nonlocal operator in the standard hat functions having non zero entries on the order of $\mathcal{O}(N^2)$ (where N denotes the size of the matrix) has non zero entries of order $\mathcal{O}(N \log N)$ only in the wavelet basis. The complexity for the full solution of the discrete scheme is shown to be bounded by $\mathcal{O}(N(\log N)^{4+2\alpha})$ where α is some constant, which has been shown numerically to be around 1 [85].

2.4.6. Extensions and variations. We note that the method discussed above with the wavelet discretization in space and dG discretization in time is valid for Markov processes more general than Lévy processes such as time-inhomogeneous Markov processes (i.e. where the law of $X_{t+h} - X_t$ can depend on t). This is mainly due to working in the weak solution (Galerkin) framework and is one of the advantages over using simpler discretizations such as Finite Differences. In other variations of a wavelet-Galerkin based method, the time discretization is a θ-scheme, thereby resulting in slightly different bounds for the total complexity [67, 85]. Higher dimensional extensions for the abstract parabolic evolution problem have been explored in [86]. For such a problem, the space dimension was discretized via sparse wavelet grids whereas the time dimension was discretized via an *hp* discontinuous Galerkin scheme. The associated linear systems were solved with the GMRES algorithm. It is pertinent to note that for such a high dimensional case, applications to high dimensional computational finance problems such as the pricing of basket options were not considered; this could be the basis of future work.

3. Transform Methods. In contrast to PDE-based methods, transform methods take an alternative view rooted in applied probability and signal processing. As opposed to deriving a PIDE to solve, the problem is tackled as a pure integration problem. Although mathematically pleasing, the pure integration problem would have not gained as much attention if not for the existence of the Fast Fourier Transform (FFT); the reader is referred to [84] for a more complete treatment. The FFT allows the integration to be computed in an extremely efficient manner. Before proceeding further we give an overview of this section. We begin this section with a brief description of more traditional transform methods proposed by Carr and Madan [23]. However, we do not delve too deeply into this as it is widely documented in literature and the underlying idea is straightforward. We then provide an overview of the current literature of various related transform methods followed by some preliminaries on the basics of the Fourier transform. Then, we review the COS method and the Fourier time-stepping method, which is the focus of this section due to its novelty and effectiveness.

Overview of the literature. Among the first proponents of the Fourier Transform approach are Carr and Madan [23]. There existed Fourier analytic methods in the option pricing literature before Carr and Madan's paper, although the underlying analysis was unable to harness the power of the FFT. Related works on this theme include [9, 37, 61]. Shortly following the intense research into FFT methods, it was discovered that pricing via transform methods could incur a large error for out of the money options. Lee [60] extended the line of reasoning from the authors above but was also one of the first to address the control of the error in such methods. Levendorskii also did important work in this regard and illuminated the complex analytic machinery behind the efficacy of Fourier Transform methods [20, 21]. Furthermore, the complex analytic framework of the transform methods shows that both the COS method [41] and the CONV method [62] are variations of the general transform methods devised in [20]. Having deduced the important drivers of the method, they also make recommendations for improving other transform methods, namely the Hilbert transform method of Feng and Linetsky [43].

Preliminaries. Recall that for a sufficiently integrable function, the Fourier transform of a function is

$$\phi(\omega) = \int_{\mathbb{R}} e^{ix\omega} f(x) dx \tag{3.1}$$

and that the inverse Fourier transform is

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ix\omega} \phi(\omega) d\omega \tag{3.2}$$

Note that our choice of notation conveniently coincides with that of a probability density, f, and its characteristic function, ϕ.

3.1. COS Method. The COS method is an extremely efficient option pricing method first introduced by Fang and Oosterlee [41] utilizing the analytic techniques of Sinc methods [79]. Specifically, the density function in the risk neutral valuation formula is expanded via a Fourier-Cosine series, which is intimately related to its characteristic function. It is due to this expansion that the method exhibits exponential convergence for continuous payoffs and algebraic convergence for discontinuous payoffs. The risk neutral valuation formula is canonically given as

$$v(x, t_0) = e^{-r\tau} \mathbb{E}^{\mathbb{Q}}[v(y, T)|x] = e^{-r\tau} \int_{\mathbb{R}} v(y, T) f(y|x) dy \tag{3.3}$$

where τ denotes time to maturity, $f(y|x)$ is the probability density of y given x at time t_0, and $v(y, T)$ is the payoff function. This is an inherently different approach from PIDE methods, where, from the risk neutral valuation formula, no-arbitrage arguments are applied in order to derive a PIDE which the derivative instrument must satisfy. In most cases, this PIDE is then solved numerically. The COS method, on the other hand, treats the evaluation of the contingent claim as a pure numerical integration problem. It is important to acknowledge this subtle yet important point – we realize that the COS method is not natural for contingent claims with American features as those problems are fundamentally (stochastic) optimal control problems.[2] The effectiveness of the COS method comes from exploiting the exponential decay of the Fourier-Cosine series coefficients that stems from the relationship between the density and the characteristic function (Fourier-cosine expansion). The COS method not only applies to vanilla European options and has been extended to barrier options [62], European style Asian options [94], and swing options [93].

3.1.1. The Fourier-Cosine Expansion. The Fourier-Cosine expansion of a general function f supported on $[0, \pi]$ is given as

$$f(\theta) = {\sum_{k=0}^{\infty}}' A_k \cdot \cos(k\theta) \tag{3.4}$$

$$A_k = \frac{2}{\pi} \int_0^{\pi} f(\theta) \cos(k\theta) d\theta \tag{3.5}$$

where ${\sum}'$ denotes the otherwise usual summation with the first term weighted by $1/2$. Note that by the change of variables $\theta = \frac{x-a}{b-a}\pi$ and $x = \frac{b-a}{\pi}\theta + a$ we are able to expand functions supported on $[a, b]$ by generalizing the formula for A_k to

$$A_k = \frac{2}{b-a} \int_a^b f(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx$$

[2]However, the COS method is applicable to Bermudan Options from which we can derive approximate American Option prices via Richardson extrapolation [37].

Now, suppose that (3.1) can be sufficiently well approximated on a finite interval. That is,

$$\phi_1(\omega) = \int_a^b e^{i\omega x} f(x)\mathrm{d}x \approx \int_{\mathbb{R}} e^{i\omega x} f(x)\mathrm{d}x = \phi(\omega) \tag{3.6}$$

where ϕ_1 denotes (3.1) restricted to a finite interval $[a, b]$. Then A_k can be rewritten as

$$A_k = \frac{2}{b-a}\,\mathrm{Re}\left\{\phi_1\left(\frac{k\pi}{b-a}\right)\cdot\exp\left(-i\frac{ka\pi}{b-a}\right)\right\} \tag{3.7}$$

Replacing ϕ_1 by ϕ in (3.7), we obtain the approximation

$$F_k = \frac{2}{b-a}\,\mathrm{Re}\left\{\phi\left(\frac{k\pi}{b-a}\right)\cdot\exp\left(-i\frac{ka\pi}{b-a}\right)\right\} \tag{3.8}$$

to A_k. Expanding the density via the Fourier-Cosine method with F_k as the Fourier coefficient, we obtain a first approximation

$$f_1(x) = \sum_{k=0}^{\infty}{}' F_k \cos(k\pi\frac{x-a}{b-a}) \tag{3.9}$$

to $f(x)$. Truncating the infinite series in (3.9), we obtain a second approximation

$$f_2(x) = \sum_{k=0}^{N-1}{}' F_k \cos(k\pi\frac{x-a}{b-a}) \tag{3.10}$$

to $f(x)$, which is the approximation of the Fourier-Cosine expansion of $f(x)$ that we exploit for the probability density function.

3.1.2. European Options. We apply this to European options (on a finite interval) by expanding the density function $f(y\,|\,x)$ in (3.3) in a Fourier-Cosine series, bringing us to

$$v(x,t_0) = e^{-r\tau}\int_a^b v(y,T)\sum_{k=0}^{\infty}{}' A_k(x)\cos(k\pi\frac{y-a}{b-a})dy \tag{3.11}$$

where

$$A_k(x) = \frac{2}{b-a}\int_a^b f(y|x)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \tag{3.12}$$

Now define

$$V_k = \frac{2}{b-a}\int_a^b v(y,T)\cos(k\pi\frac{y-a}{b-a})dy \tag{3.13}$$

the cosine series coefficients of the payoff function. Then, (3.11) can be approximated by

$$v_1(x,t_0) = \frac{1}{2}(b-a)e^{-r\tau}\sum_{k=0}^{N-1}{}' A_k(x)V_k \tag{3.14}$$

Approximating A_k by F_k as above, one obtains the approximation

$$v_2(x,t_0) = e^{-r\tau} \sum_{k=0}^{N-1}{}' \operatorname{Re}\left\{ \phi\left(\frac{k\pi}{b-a}, x\right) e^{-ik\pi\frac{a}{b-a}} \right\} V_k \tag{3.15}$$

to $v(x,t_0)$, where $\phi(\omega,x)$ is the characteristic function of $f(y|x)$, thereby arriving at a semi-discrete formula with a fixed number, N, of terms from a pure integration problem on an infinite interval. It can be shown that V_k can be determined analytically for the case of vanilla and digital options, whereby we obtain a fast and efficient analytical formula for the evaluation of options. We note in passing that, for more complicated contracts, V_k can be determined numerically in the cases where it cannot be determined analytically.

3.1.3. The Lévy Case. Clearly, the COS method is ideally suited to models where the characteristic function is given analytically, as is the case for many Lévy processes. What may seem counterintuitive is that the method when applied to a more complicated stochastic processes is, in fact, greatly simplified. This is due to the fact that many Lévy characteristic functions are trivially dependent on the (log) space variable in the sense that it can essentially be factored out. Such characteristic functions are usually of the form

$$\phi(\omega, \mathbf{x}) = \varphi_{levy}(\omega) \cdot e^{i\omega\mathbf{x}} \tag{3.16}$$

where φ_{levy} represents the part of the characteristic function dependent on the model parameters. It is due to this that multiple strikes can be computed simultaneously. This leads to the pricing formula

$$v_2(x,t_0) = e^{-r\tau} \sum_{k=0}^{N-1}{}' \operatorname{Re}\left\{ \varphi_{levy}\left(\frac{k\pi}{b-a}\right) e^{-ik\pi\frac{\mathbf{x}-a}{b-a}} \right\} \mathbf{V}_k \tag{3.17}$$

It is shown in [41] that $\mathbf{V}_k = U_k\mathbf{K}$ where U_k corresponds to the cosine series coefficients of the payoff function, introduced above in (3.13) (which can be determined analytically for European and digital options) and $\mathbf{K}$ is a vector of strikes. It is due to this V_k decomposition that we are able to price multiple strikes with one evaluation in the Lévy case.

Remark 1. The Heston model also enjoys simplifications similar to the Lévy case for the same reason that the (log) space variable can essentially be factored out.

Remark 2. The COS method (and other numerical methods that require the evaluation of the characteristic function) can still be applied when an analytic characteristic function is not available as it is possible to numerically evaluate the characteristic function for affine jump diffusion models via a system of ODEs [37].

3.1.4. Greeks. As the pricing formula (3.15) is analytic, we can do a straightforward partial differentiation to obtain another formula to compute the sensitivities (i.e., Greeks). As the COS method is highly efficient, the additional computational cost of another COS-pricing formula is small compared to the cost of computing Greeks in other numerical methods. For example, the delta for the vanilla European

option case for a general underlying stochastic process is

$$\begin{aligned}\Delta =&\frac{\partial v}{\partial S_o} = \frac{\partial v}{\partial x}\frac{\partial x}{\partial S_o} = \frac{1}{S_o}\frac{\partial v}{\partial x} \\ \approx \quad & e^{-r\tau}\sum_{k=0}^{N-1}{}' \operatorname{Re}\left\{\partial\phi\left(\frac{k\pi}{b-a}, x\right)\Big/\partial x \cdot e^{-ik\pi\frac{a}{b-a}}\right\}\frac{V_k}{S_0}\end{aligned} \tag{3.18}$$

The other Greeks of interest are obtained in a similar fashion.

3.1.5. Bermudan Options. A Bermudan option is a contract which allows for exercise at multiple (pre-determined) dates before maturity. It is clear that the level of exercisability is between that of an European and an American option. Indeed, it is intuitive that in the limit of the exercise dates approaching infinity, the Bermudan option price should approach that of the American option. It can be easily seen that (due to the Markov property) between exercise dates, the value of the option can be viewed as a simple European option. As such, the valuation of a Bermudan option amounts to an induction backwards in time, taking the maximum of either the payoff or the continuation value at that exercise point. The continuation value is the value of continuing to hold the option (forgoing exercise). Formally, this means that given a collection of exercise times $\mathcal{T} = \{t_1, ..., t_M\}$ and t_0 the initial time, for $m = M, M-1, ..., 2$

$$\begin{cases} c(x, t_{m-1}) = e^{-r(t_m - t_{m-1})}\int_{\mathbb{R}} v(y, t_m) f(y|x) dy \\ v(x, t_{m-1}) = \max\left(g(x, t_{m-1}), c(x, t_{m-1})\right) \end{cases} \tag{3.19}$$

$$v(x, t_0) = e^{-r(t_1 - t_0)}\int_{\mathbb{R}} v(y, t_1) f(y|x) dy$$

where x and y are state variables (log-space), with v, c, g denoting the option value, continuation value, and payoff, respectively. It is also important to define, for what follows, the free boundary point , x_m^*, which separates the continuation region, where the option should be held, from the exercise region, where the option should be exercised. As the valuation of a Bermudan option primarily relies on the valuation of European options it can be solved by the COS method. This is explored in [42]. We expound on the basic methodology in pricing such options within the COS method and avoid detailed discussion of the numerical intricacies required in implementing the method in an efficient and practical manner. Utilizing the COS method described above, we can calculate the continuation value and the option value as

$$\begin{aligned}\hat{c}(x, t_{m-1}) &:= e^{-r\Delta t}\sum_{k=0}^{N-1}{}' \operatorname{Re}\left\{\phi(\frac{k\pi}{b-a}; x)e^{-ik\pi\frac{a}{b-a}}\right\}V_k(t_m) \\ V_k(t_m) &:= \frac{2}{b-a}\int_a^b v(y, t_m)\cos(k\pi\frac{y-a}{b-a})dy\end{aligned} \tag{3.20}$$

From (3.20), it can be observed that the majority of the work in computing the Bermudan option price is in determining the V_k. Since it is much easier in determining the early-exercise point in Bermudan options as opposed to American options, we can determine the early exercise point via a root finder (Newton's method) for $\hat{c}(x, t_m) -$

$g(x, t_m) = 0$. This is valid because we have a semi-analytic formula for $\hat{c}$ as a function of x, whereby we can obtain a derivative necessary for Newton's method. Note that we also have a good initial guess of x_m^* from the previous step if $t_{m+1} - t_m$ is small. Given knowledge of the x_m^*, the Fourier cosine coefficients V_k can thus be split into parts corresponding to the intervals $[a, x_m^*]$ and $(x_m^*, b]$:

$$V_k(t_m) = \begin{cases} C_k(a, x_m^*, t_m) + G_k(x_m^*, b, t_m) & \text{call} \\ G_k(a, x_m^*, t_m) + C_k(x_m^*, b, t_m) & \text{put} \end{cases}$$

for $m = M - 1, M - 2, ..., 1$, and

$$V_k(t_M) = \begin{cases} G_k(0, b, t_m) & \text{call} \\ G_k(a, 0, t_M) & \text{put} \end{cases} \tag{3.21}$$

where we assume that

$$\begin{aligned} G_k(x_1, x_2, t_m) &:= \frac{2}{b-a} \int_{x_1}^{x_2} g(x, t_m) \cos(k\pi \frac{x-a}{b-a}) dx \\ C_k(x_1, x_2, t_m) &:= \frac{2}{b-a} \int_{x_1}^{x_2} c(x, t_m) \cos(k\pi \frac{x-a}{b-a}) dx \end{aligned} \tag{3.22}$$

where we recall that $g(x, t_m)$ is the payoff. Hence, similar to the COS method, G_k can be determined analytically.

3.1.6. American Options. As the COS method fundamentally treats the option pricing problem as an integration problem, it was not designed specifically to handle American options. This is due to the simple fact that the American option problem is inherently an Optimal Stopping problem. However, it is well known and intuitive that in the limit of the exercise dates of a Bermudan option approaching infinity, the value of a Bermudan option should approach that of an American option. Hence, via extrapolation techniques such as Richardson extrapolation, one can obtain the price of American options via the prices of Bermudan options [24] that are solved via the COS method.

3.1.7. 2-D Rainbow Options. The extension of the COS method to higher dimensions is introduced in [75]. The exposition here is focused on the 2D case; this can be easily extended to higher dimensions. In theory, this is a good and simple extension, however, practically, as with most non-Monte Carlo based methods, we eventually run into the curse of dimensionality. For this method, the total computational complexity is $\mathcal{O}(2^n N^n \log_2 N)$, where n refers to the dimensionality of the asset and N refers to the number of terms to include in the Fourier-cosine summation, (which is related to the desired accuracy). Empirical experiments show that up to 10 dimensions can be considered tractable for this method. Furthermore, memory requirements grow exponentially when the dimensions increase, also limiting the effectiveness of this method beyond a dimension of around 10. However, this is not to say that this mild multidimensional extension is nothing more than a theoretical exercise as the majority of real world multidimensional problems are of low dimensions. Indeed, it is common industry practice for classically intractable high dimensional problems ($d > 10$) in the equity world to be proxied by an index, thereby reducing the effective dimension.[3]

[3]The reader is referred to the papers [88, 89, 90] on the topic of effective dimension.

Intuitively, an option on two assets can be thought of as an option on one asset that has two dimensions. Recalling the fact that the COS method tackles the problem from a numerical integration perspective, the fundamental risk neutral valuation formula for the two dimensional problem can be written as

$$v(\mathbf{x}, t_o) = e^{-r\Delta t}\mathbb{E}^{t_0,\mathbf{x}}[g(\mathbf{X}_T)] = e^{-r\Delta t}\iint_{\mathbb{R}^2} g(\mathbf{y})f(\mathbf{y}|\mathbf{x})d\mathbf{y} \tag{3.23}$$

where $\Delta t = T - t_0$, $g(\cdot)$ denotes, as usual, the payoff function, $\mathbf{x} = (x_1, x_2)$ is the state variable, and $f(\cdot|\cdot)$ is the conditional density. Clearly, (3.23) takes the approach of treating two assets as a single asset in two dimensions, hence the double integral. As in the 1D COS method, we first discretize the integration domain, resulting in the approximation

$$v_1(\mathbf{x}, t_0) = e^{-r\Delta t}\int_{a_2}^{b_2}\int_{a_1}^{b_1} g(\mathbf{y})f(\mathbf{y}|\mathbf{x})d\mathbf{y} \tag{3.24}$$

to $v(\mathbf{x}, t_0)$. Expanding the density via a Fourier cosine expansion, we arrive at

$$v_1(\mathbf{x}, t_0) = e^{-r\Delta t}\int_{a_2}^{b_2}\int_{a_1}^{b_1} g(\mathbf{y}) {\sum_{k_1=0}^{\infty}}' {\sum_{k_2=0}^{\infty}}' A_{k_1,k_2}(\mathbf{x}) \cos\left(k_1\pi\frac{y_1 - a_1}{b_1 - a_1}\right)\cos\left(k_2\pi\frac{y_2 - a_2}{b_2 - a_2}\right)dy_1dy_2$$

with the series coefficients defined as

$$A_{k_1,k_2}(\mathbf{x}) := \frac{2}{b_1 - a_1}\frac{2}{b_2 - a_2}\int_{a_2}^{b_2}\int_{a_1}^{b_1} f(\mathbf{y}|\mathbf{x})\cos\left(k_1\pi\frac{y_1 - a_1}{b_1 - a_1}\right)\cos\left(k_2\pi\frac{y_2 - a_2}{b_2 - a_2}\right)dy_1dy_2$$

As in the 1D case, with some rearranging of the terms, we can then define the Fourier cosine coefficients of the payoff function as

$$V_{k_1,k_2}(T) := \frac{2}{b_1 - a_1}\frac{2}{b_2 - a_2}\int_{a_2}^{b_2}\int_{a_1}^{b_1} g(\mathbf{y})\cos\left(k_1\pi\frac{y_1 - a_1}{b_1 - a_1}\right)\cos\left(k_2\pi\frac{y_2 - a_2}{b_2 - a_2}\right)dy_1dy_2$$

Rearranging the terms in order to take advantage of the above definition and truncating the infinite series, we arrive at the second approximation

$$v_2(\mathbf{x}, t_0) = \frac{b_1 - a_1}{2}\frac{b_2 - a_2}{2}\, e^{-r\Delta t}\, {\sum_{k_1=0}^{N-1}}' {\sum_{k_2=0}^{N-1}}' A_{k_1,k_2}(\mathbf{x})V_{k_1,k_2}(T) \tag{3.25}$$

to $v(\mathbf{x}, t_0)$. Continuing analogously as in the 1D case, an approximation of $A_{k_1,k_2}(\mathbf{x})$ on the support of the whole line is taken

$$F_{k_1,k_2}(x) := \frac{2}{b_1 - a_1}\frac{2}{b_2 - a_2}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\cos\left(k_1\pi\frac{y_1 - a_1}{b_1 - a_1}\right)\cos\left(k_2\pi\frac{y_2 - a_2}{b_2 - a_2}\right)dy_1dy_2 \tag{3.26}$$

Before rewriting the final form of the discretization, we rewrite (3.26) by noticing that the product of the cosines satisfies a trigonometric identity and can be rewritten as

$$2F_{k_1,k_2}(\mathbf{x}) = F^{+}_{k_1,k_2}(\mathbf{x}) + F^{-}_{k_1,k_2}(\mathbf{x}) \tag{3.27}$$

where

$$\begin{aligned}
F^{\pm}_{k_1,k_2}(x) :=& \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\cos\left(k_1\pi\frac{y_1-a_1}{b_1-a_1} \pm k_2\pi\frac{y_2-a_2}{b_2-a_2}\right)dy_1dy_2 \\
=& \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\mathrm{Re}\Bigg\{\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\exp\left(ik_1\pi\frac{y_1}{b_1-a_1} \pm ik_2\pi\frac{y_2}{b_2-a_2}\right) \\
& \exp\left(-ik_1\pi\frac{a_1}{b_1-a_1} \mp ik_2\pi\frac{b_2}{b_2-a_2}\right)\Bigg\}d\mathbf{y} \\
=& \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\mathrm{Re}\Bigg\{\varphi(\frac{k_1\pi}{b_1-a_1} \pm \frac{k_2\pi}{b_2-a_2}|\mathbf{x})\cdot \\
& \exp\left(-ik_1\pi\frac{a_1}{b_1-a_1} \mp ik_2\pi\frac{a_2}{b_2-a_2}\right)\Bigg\} \qquad (3.28)
\end{aligned}$$

where $\varphi(\omega|x)$ is the bivariate characteristic function of $f(y|x)$. With this, we arrive at the 2D COS formula

$$v_3(\mathbf{x},t_0) := \frac{b_1-a_1}{2}\frac{b_2-a_2}{2}\, e^{-r\Delta t}\sum_{k_1=0}^{N-1}{}'\sum_{k_2=0}^{N-1}{}'\frac{1}{2}[F^{+}_{k_1,k_2}(\mathbf{x}) + F^{-}_{k_1,k_2}(\mathbf{x})]V(T) \qquad (3.29)$$

3.2. Fourier Time Stepping. First introduced by Jackson, Jaimungal, and Surkov [56], the Fourier Time Stepping (FTS) method, markedly different from the COS method above, is actually a framework for efficient options pricing. Options of different payoffs (vanilla/exotic) and observations (European/Barrier/American) can be priced by the FTS method in a consistent manner. In contrast to [41], where the problem is approached from a pure numerical integration standpoint, the FTS method approaches the option pricing problem instead from a PIDE standpoint. Namely, the risk neutral valuation equation (3.3), an integration problem, is transformed into PIDE form by applying the no-arbitrage condition (zero-drift), resulting in

$$\begin{cases}(\partial_t + \mathcal{L})v(x,t) = 0 \\ v(T,x) = \varphi(\mathbf{S}(0)e^{\mathrm{x}})\end{cases} \qquad (3.30)$$

where $\mathcal{L}$ is the infinitesimal generator of the Lévy process, in multidimensional form,

$$\mathcal{L}g(\mathbf{x}) = (\gamma'\partial_{\mathbf{x}} + \frac{1}{2}\partial'_{\mathbf{x}}\Sigma\partial_{\mathbf{x}})g(\mathbf{x}) + \int_{\mathbb{R}\setminus 0}(g(x+y)-g(x)-\mathbf{1}_{\{|\mathbf{y}|<1\}}y'\partial_{\mathbf{x}}g(x))\nu(\mathrm{d}y) \qquad (3.31)$$

An important piece of insight is that, when the Fourier Transform is applied to the infinitesimal generator, the characteristic exponent can be factored out

$$\begin{aligned}
\mathcal{F}[\mathcal{L}v](t,\omega) =& \left\{i\gamma'\omega - \frac{1}{2}\omega'\Sigma\omega + \int_{\mathbb{R}^n}\left(e^{i\omega'\mathbf{y}} - 1 - i\mathbf{1}_{\{|\mathbf{y}|<1\}}\omega'\mathbf{y}\right)\nu(\mathrm{d}\mathbf{y})\right\}\mathcal{F}[v](t,\omega) \\
=& \Psi(\omega)\mathcal{F}[v](t,\omega) \qquad (3.32)
\end{aligned}$$

Another result that is important and used widely in the FTS method is the well known fact that

$$\mathcal{F}[\partial^n_{\mathbf{x}}g](\omega) = i\omega\mathcal{F}[\partial^{n-1}_{\mathbf{x}}g](\omega) = \cdots = (i\omega)^n\mathcal{F}[g](\omega) \qquad (3.33)$$

With knowledge of the above, applying the Fourier Transform to the PIDE (3.30) results in

$$\begin{cases}\partial_t\mathcal{F}[v](t,\omega) + \Psi(\omega)\mathcal{F}[v](t,\omega) = 0 \\ \mathcal{F}[v](T,\omega) = \mathcal{F}[\varphi](\omega)\end{cases} \qquad (3.34)$$

a system of ODEs that are easily solved analytically. From this system, given a value of $\mathcal{F}[v](t,\omega)$ at time $t_2 \leq T$, we can obtain the value at $t_1 < t_2$ via

$$v(t_1,\mathbf{x}) = \mathcal{F}^{-1}\left[\mathcal{F}[v](t_2,\omega)\cdot e^{\Psi(\omega)(t_2-t_1)}\right](\mathbf{x}) \tag{3.35}$$

which is the valuation equation in continuous form. To implement this on a computer, we use the Fast Fourier Transform (FFT) in place of the continuous Fourier Transform operator $\mathcal{F}$ in (3.35). The Fourier Time Stepping method is efficient, as inherited from the use of FFTs in the valuation equation. Furthermore, it is easily seen from the valuation equations that stepping in space is not required; the time stepping in the frequency space contains embedded information about the spacial domain via the Nyquist critical frequency.

3.2.1. European Options. Evident in (3.35), European options can be solved in an extremely efficient manner with a single time step. For such a case, one starts with the payoff at the maturity and does a single timestep backward via (3.35). This is not limited to vanilla European options, but also options of European exercise with more exotic payoffs such as digitals.

3.2.2. Greeks. Hedge parameters can be computed in the (FST) method framework similarly to the strategy for the COS method. Namely, the valuation equation is partially differentiated analytically and the resulting equation is solved as if it were an original valuation equation. In the FST framework, this method of obtaining the Greeks is referred to as greekFST. In many cases (most predominantly Delta and Gamma), the Greeks are obtained simply via a scaling of the option values.

3.2.3. Extensions. The FST method is not limited to European options: as mentioned previously, FST method is also applicable to Bermudan and American options. In addition to this, the FST method has been extended in two different manners - namely to more exotic payoffs and complex models of the underlying. Payoff-wise, the FST method is also applicable to exotic options such as Barrier options, Shout options, and Swing options. Model-wise extensions include extending the FST method to be applicable in mean-reverting and regime-switching frameworks.

REFERENCES

[1] L. Abbas-Turki and B. Lapeyre, *American options by Malliavin calculus and nonparametric variance and bias reduction methods*, SIAM Journal on Financial Mathematics, 3 (2012), pp. 479–510.

[2] Y. Achdou and J. Guermond, *Convergence analysis of a finite element projection/Lagrange–Galerkin method for the incompressible Navier–Stokes equations*, SIAM Journal on Numerical Analysis, 37 (2000), pp. 799–826.

[3] Y. Achdou and O. Pironneau, *Computational methods for option pricing*, Society for Industrial and Applied Mathematics (SIAM), (2005).

[4] Y. Achdou and N. Tchou, *Variational analysis for Black and Scholes equation with stochastic volatility*, M2AN Math. Model. Numer. Anal, 36 (2002), pp. 373–395.

[5] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, (1978).

[6] L. Andersen and J. Andreasen, *Jump-diffusion models: Volatility smile fitting and numerical methods for pricing*, Rev. Derivatives Research, 4 (2000), pp. 231–262.

[7] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, (2004).

[8] F. Avner, *Variational Principles and Free-Boundary Problems*, Dover Publications, (2010).

[9] B. Bakshi and D. Madan, *Spanning and derivative-security valuation*, Journal of Financial Economics, 55 (2000), pp. 205–238.

[10] O.E. Barndorff-Nielsen, *Processes of normal inverse Gaussian type*, Finance and Stochastics, 2 (1997), pp. 41–68.

[11] O.E. Barndorff-Nielsen and S.Z. Levendorskii, *Feller processes of normal inverse Gaussian type*, Quantitative Finance, 1 (2001), pp. 318–331.

[12] O.E. Barndorff-Nielsen and N. Shephard, *Normal modified stable processes*, Working paper, (2001).

[13] R. Bellman, *Functional equations in the theory of Dynamic Programming V. positivity and quasi-linearity*, Proc Natl Acad Sci USA, 41 (1955), pp. 743–746.

[14] ———, *Dynamic Programming*, Princeton University Press, (1957).

[15] J. Bertoin, *Lévy Processes*, Cambridge University Press, (1996).

[16] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy, 81 (1973), pp. pp. 637–654.

[17] D. Blackwell, *Discounted Dynamic Programming*, Ann. Math. Statist., 36 (1965), pp. 226–235.

[18] O. Bokanwski, S. Maroso, and H. Zidani, *Some convergence results for Howard's algorithm*, SIAM Journal on Numerical Analysis, 47 (2009), pp. 3001–3026.

[19] S. Boyarchenko and S. Levendorskii, *Non-Gaussian Merton-Black-Scholes Theory*, Advanced Series on Statistical Science and Applied Probability, Volume 9. World Scientific, (2002).

[20] ———, *New efficient versions of Fourier transform method in applications to option pricing*, SSRN, (2011), pp. 1–62.

[21] ———, *Fast pricing and calculation of sensitivities of out-of-the-money European options under Lévy processes*, Journal of Computational Finance, 15 (2012), pp. 71–133.

[22] P. Carr, H. Geman, D.B. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Computational Finance, 75 (2002), pp. 305–332.

[23] P. Carr and D.B. Madan, *Option valuation using the fast Fourier transform*, Journal of Computational Finance, 3 (1999), pp. 463–520.

[24] C.G. Chang, S.L. Chung, and R.C. Stapleton, *Richardson extrapolation techniques for the pricing of American-style options*, Journal of Futures Markets, 27 (2007), pp. 791–817.

[25] F. Chen, J. Shen, and H. Yu, *A new spectral element method for pricing european options under the Black-Scholes and Merton jump diffusion models*, Journal of Scientific Computing, 52 (2012), pp. 499–518.

[26] A. Cohen, *Numerical analysis of wavelet methods*, Studies in Mathematics and Its Applications, JAI Press, (2003).

[27] A. Cohen and R. Masson, *Wavelet methods for second-order elliptic problems, preconditioning, and adaptivity*, SIAM Journal on Numerical Analysis, 21 (1999), pp. 1006–1023.

[28] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman and Hall/CRC, (2004).

[29] R. Cont and E. Voltchkova, *A finite difference scheme for option pricing in jump diffusion and exponential Lévy models*, SIAM Journal on Numerical Analysis, 43 (2005), pp. 1596–1626.

[30] ———, *Integro-Differential equations for option prices in exponential Lévy models*, Finance and Stochastics, 9 (2005), pp. 299–325.

[31] R.W. Cottle, J. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, (1992).
[32] M. Crandall and P.L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.
[33] W. Dahmen, *Wavelet and multiscale methods for operator equations*, Acta Numerica, (1997), pp. 55–28.
[34] I. Daubechies, *Ten lectures on wavelets*, SIAM, (1992).
[35] Y. d'Halluin, P.A. Forsyth, and G. Labahn, *A penalty method for American options with jump diffusion processes*, Numer. Math., 97 (2004), pp. 321–352.
[36] D.A. Di Petro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer, New York, (2012).
[37] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and option pricing for affine jump-diffusions*, Econometrica, 68 (2000), pp. 1343–76.
[38] E. Eberlein and U. Keller, *Hyperbolic distributions in finance*, Bernoulli, 1 (1995), pp. 281–299.
[39] E. Eberlein, K. Ulrich, and K. Prause, *New insights into smile, mispricing, and value at risk: The hyperbolic model*, The Journal of Business, 71 (1998), pp. pp. 371–405.
[40] J. Eckstein and D.P. Bertsekas, *On the Douglas Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, 55 (1992), pp. 293–318.
[41] F. Fang and C. Oosterlee, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM Journal on Scientific Computing, 31 (2009), pp. 826–848.
[42] F. Fang and C. W. Oosterlee, *Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions*, Numerische Mathematik, 114 (2009), pp. 27–62.
[43] L. Feng and V. Linetsky, *Pricing discretely monitored barrier options and defaultable bonds in Lévy process models: a fast Hilbert transform approach*, Mathematical Finance, 18 (2008), pp. 337–384.
[44] L. Feng, V. Linetsky, J.L. Morales, and J. Nocedal, *On the solution of complementarity problems arising in American options pricing*, Optimization Methods and Software, 26 (2011), pp. 813–825.
[45] P.A. Forsyth and G. Labahn, *Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance*, Journal of Computational Finance, 11 (2007), pp. 1–44.
[46] P.A. Forsyth and K. Vetzal, *Quadratic convergence for valuing American options using a penalty method*, SIAM Journal on Scientific Computing, 23 (2002), pp. 2095–2122.
[47] P. Garreau and D. Kopriva, *A spectral element framework for option pricing under general exponential Lévy processes*, Journal of Scientific Computing, 53 (2013), pp. 1–24.
[48] H. Geman, *Pure jump Lévy processes for asset price modelling*, Journal of Banking and Finance, 26 (2002), pp. 1297–1316.
[49] J.S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods*, Springer, New York, (2008).
[50] R.A Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, (1994).
[51] R. Howard, *Dynamc Programming and Markov Processes*, MIT Press, (1960).
[52] S. Howison, C. Reisinger, and J. Witte, *The effect of nonsmooth payoffs on the penalty approximation of American options*, SIAM Journal on Financial Mathematics, 4 (2013), pp. 539–574.
[53] J. Huang and J. Pang, *Option pricing and linear complementarity*, Journal of Computational Finance, 2 (1998), pp. 31–60.
[54] S. Ikonen and J. Toivanen, *Operator splitting methods for American option pricing*, Applied Mathematical Letters, 17 (2004), pp. 809–814.
[55] ———, *Componentwise splitting methods for pricing American options under stochastic volatility*, International Journal of Theoretical and Applied Finance, 10 (2007), pp. 331–361.
[56] K.R. Jackson, S. Jaimungal, and V. Surkov, *Fourier space time-stepping for option pricing with Lévy models*, Journal of Computational Finance, 12 (2008), pp. 1–29.
[57] H.J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, (2001).
[58] Y. Kwon and Y. Lee, *A second-order tridiagonal method for American options under jump-diffusion models*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1860–1872.
[59] A.E. Kyprianou, *Fluctuations of Lévy Processes with Applications*, Springer, (2006).
[60] R. Lee, *Option pricing by transform methods: extensions, unification and error control*, Journal of Computational Finance, 7 (2004), pp. 51–86.
[61] A. Lewis, *A simple option formula for general jump-diffusion and other exponential Lévy processes*, Envision Finanal Systems, (2001).

[62] R. Lord, F. Fang, F. Bervoets, and C. W. Oosterlee, *A fast and accurate FFT-based method for pricing early-exercise options under Lévy process*, SIAM Journal on Scientific Computing, 30 (2008), pp. 1678–1705.

[63] D.B. Madan, P. Carr, and C. Chang, *The variance gamma process and option pricing*, European Finance Review, 2 (1998), pp. 79–105.

[64] R.N. Mantegna and H.E. Stanley, *Stochastic processes with ultraslow convergence to a Gaussian: the truncated Lévy flight*, Physical Review Letters, 73 (1994), pp. 2946–2949.

[65] A. Matache, P.A. Nitsche, and C. Schwab, *Wavelet Galerkin pricing of American options on Lévy driven assets*, Quantitative Finance, 5 (2005), pp. 403–424.

[66] A. Matache, C. Schwab, and T. Wihler, *Fast numerical solution of parabolic integrodifferential equations with applications in finance*, SIAM Journal on Scientific Computing, 27 (2005), pp. 369–393.

[67] A. Matache, T. von Petersdorff, and C. Schwab, *Fast deterministic pricing of options on Lévy driven assets*, ESAIM: Mathematical Modelling and Numerical Analysis, 38 (2004), pp. 37–71.

[68] B.F. Nielsen, O. Skavhaug, and A. Tveito, *Penalty and front-fixing methods for the numerical solution of the American option problems*, Journal of Computational Finance, 5 (2002), pp. 96–97.

[69] E.A. Novikov, *Infinitely divisible distributions in turbulence*, Physics Review E, 50 (1996), pp. 3303–3305.

[70] C. W. Oosterlee, *On multigrid for linear complementarity problems with application to American-style options*, Electron. Trans. Numer. Anal, 15 (2003), pp. 165–185.

[71] Lions. P.L. and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM Journal on Numerical Analysis, 16 (1979), pp. 964–979.

[72] M.L. Puterman and S.L. Brumelle, *On the convergence of policy iteration in stationary Dynamic Programming*, Math. Oper. Res., (1979), pp. 60–69.

[73] C. Reisinger and J. Witte, *On the use of policy iteration as an easy way of pricing American options*, SIAM Journal on Financial Mathematics, 3 (2012), pp. 459–478.

[74] D. Robinson, L. Feng, J. Nocedal, and J. Pang, *Subspace accelerated matrix splitting algorithms for asymmetric and symmetric linear complementarity problems*, SIAM Journal on Optimization, 23 (2013), pp. 1371–1397.

[75] M. Ruijter and C. Oosterlee, *Two-dimensional Fourier cosine series expansion method for pricing financial options*, SIAM Journal on Scientific Computing, 34 (2012), pp. B642–B671.

[76] M.S. Santos and J. Rust, *Convergence analysis of a finite element projection/Lagrange–Galerkin method for the incompressible Navier–Stokes equations*, SIAM Journal on Numerical Analysis, 37 (2000), pp. 799–826.

[77] K-I. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Camberidge Studies in Advanced Mathemathics. Cambridge University Press, (1999).

[78] R. Schneider, *Multiskalen- und wavelet matrixkompression*, Teubner Series in Applied Mathematics, Stuttgart, (1998).

[79] F. Stenger, *Numerical methods based on sinc and analytic functions*, Springer Series in Computational Mathematics, (1993).

[80] R. Stevenson, *Adaptive solution of operator equations using wavelet frames*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 1074–1100.

[81] ———, *On the compressibility of operators in wavelet coordinates*, SIAM Journal on Numerical Analysis, 35 (2004), pp. 1110–1132.

[82] ———, *Composite wavelet bases with extended stability and cancellation properties*, SIAM Journal on Numerical Analysis, 35 (2007), pp. 1110–1132.

[83] D.W. Stroock, *Introduction to Markov Processes*, Springer, (2005).

[84] C. Van Loan, *Computational frameworks for the fast Fourier transform*, SIAM Frontiers in Applied Mathematics, Stuttgart, (1992).

[85] T. von Petersdorff and C. Schwab, *Wavelet discretizations of parabolic integrodifferential equations*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 159–180.

[86] ———, *Numerical solution of parabolic equations in high dimensions*, ESAIM: Mathematical Modelling and Numerical Analysis, 38 (2004), pp. 93–127.

[87] I. Wang, J.W.L. Wan, and P.A. Forsyth, *Robust numerical valuation of European and American options under the CGMY process*, Journal of Computational Finance, 10 (2007), pp. 31–69.

[88] X Wang, *Enhancing quasi-Monte-Carlo methods by exploiting additive approximation for problems in finance*, SIAM Journal on Scientific Computing, 34 (2012), pp. A283–A308.

[89] X. Wang and I. H. Sloan, *Why are high-dimensional finance problems often of low effective*

dimension, SIAM Journal on Scientific Computing, 27 (2005), pp. 159–183.

[90] ———, *Efficient weighted lattice rules with applications to finance*, SIAM Journal on Scientific Computing, 28 (2006), pp. 728–750.

[91] P. Wilmott, J. Dewynne, and S. Howison, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, (1994).

[92] J. Witte and C. Reisinger, *A penalty method for the numerical solution of Hamilton-Jacobi-Bellman (HJB) equations in finance*, SIAM Journal on Numerical Analysis, 49 (2011), pp. 213–231.

[93] B. Zhang and C. W. Oosterlee, *An efficient pricing algorithm for swing options based on Fourier cosine expansions*, Journal of Computational Finance, 16 (2013), pp. 3–34.

[94] ———, *Efficient pricing of European-style Asian options under exponential Lévy processes based on Fourier Cosine expansions*, SIAM Journal on Finanical Mathematics, 4 (2013), pp. 399–426.

[95] X. Zhang, *Valuation of American options in a jump-diffusion model*, in Numerical Methods in Finance, Cambridge University Press, Cambridge, UK, (1997), pp. 93–114.