A dimension reduction Shannon-wavelet based method for option pricing *

Duy-Minh Dang† Luis Ortiz-Gracia ‡

April 14, 2017

Abstract

We present a robust and highly efficient dimension reduction Shannon-wavelet method for computing European option prices and hedging parameters under a general jump-diffusion model with square-root stochastic variance and multi-factor Gaussian interest rates. Within a dimension reduction framework, the option price can be expressed as a two-dimensional integral that involves only (i) the value of the variance at the terminal time, and (ii) the time-integrated variance process conditional on this value. A Shannon wavelet inverse Fourier technique is developed to approximate the conditional density of the time-integrated variance process. Furthermore, thanks to the excellent approximation properties of Shannon wavelets, the overall pricing procedure is reduced to the evaluation of just a single integral that involves only the density of the terminal variance value. This single integral can be accurately evaluated, since the density of the variance at the terminal time is known in closed-form. We develop sharp approximation error bounds for the option price and hedging parameters. Numerical experiments confirm the robustness and impressive efficiency of the method.

Keywords: Shannon wavelets, dimension reduction, jump diffusions.

AMS Classification 62P05, 60E10, 65T60

1 Introduction

Jump-diffusion models with stochastic variance are very popular in option pricing, due to their ability to capture, in both short and long maturities, the two important empirical phenomena, namely (i) the leptokurtic features of the asset return distribution, and (ii) the observed volatility smile/skew. See, Alizadeh et al. (2002); Andersen et al. (2002); Bakshi et al. (1997); Bates (1996), among many others. In addition, from a risk management point of view, jump-diffusion models are useful as they permit us to explore the effects of severe market crashes on the underlying asset price. Recently, extensions to these models to include one-factor stochastic interest rates, such as

*This research was supported in part by a University of Queensland Early Career Researcher (ECR) Grant (grant number 1006301-01-298-21-609775), and by the Agència de Gestió i d'Ajuts Universitaris i de Recerca (grant number 2014SGR-1307).

†School of Mathematics and Physics, The University of Queensland, St Lucia, Brisbane 4072, Australia, email: duyminh.dang@uq.edu.au

‡Universitat de Barcelona School of Economics, Faculty of Economics and Business, University of Barcelona, John M. Keynes 1-11, 08034 Barcelona, Spain
email: luis.ortiz-gracia@ub.edu

the Hull-White (Hull and White, 1993) and the square-root CIR (Cox et al., 1985a) dynamics, have become more and more common in the finance literature.[1] This is due to increasing popularity of long-dated products, as well as risk-management purposes. For example, see Ahlip and Rutkowski (2013); Cozma and Reisinger (2016); Grzelak and Oosterlee (2011, 2012b); Haastrecht and Pelsser (2011); Haentjens and in 't Hout (2012).

Whilst the use of one-factor interest rate dynamics has been popular in option pricing, a major limitation of these models is their inability to accurately capture de-correlations, i.e. non-perfect correlations, between rates for different maturities. In other words, under a one-factor interest rate model, a shock to the interest rate curve at any given time instant is transmitted equally through all maturities. This property of one-factor interest rate models is not only unrealistic, since interest rates are known to exhibit some de-correlation, but also undesirable from a risk-management standpoint (Brigo and Mercurio, 2006; Jamshidian and Zhu, 1997; Rebonato, 1998). It is suggested in some of the standard text books, such as Brigo and Mercurio (2006), that, in order to sufficiently capture de-correlations in the rates, multi-factor interest rate dynamics should be used. A number of empirical studies of the whole yield curve using principal component analysis also supports the use of multi-factor interest rate dynamics. As examples, in the analysis in Jamshidian and Zhu (1997), where JPY, USD and DEM data are considered, one principal component explains from 68% to 76% of the total variation, whereas three principal components can explain from 93% to 94%. In the analysis in Rebonato (1998) which uses the UK data, one component explains 92% of the total variance, whereas two components can explain 99.1% of the total variance.

While from the modelling and risk-management perspectives, jump-diffusion models with stochastic variance and multi-factor interest rates provide realistic dynamics for the underlying, from the computational viewpoint, these models pose a number of significant challenges. These challenges are high-dimensionality, and jumps in the underlying asset price, as well as the model's non-affinity, due to non-trivial correlations between the underlying asset price and its variance (Ahlip and Rutkowski, 2013; Grzelak and Oosterlee, 2012a,b).[2] The first hurdle in using such a general model is calibration, which typically requires a *very* efficient pricing method for European options. Broadly speaking, existing computational methods in finance can be classified into three major approaches, namely Monte Carlo (MC), partial differential equation (PDE), and numerical integration, linked together via the famous Feynman-Kac theorem. It appears that both the MC and PDE computational approaches are neither feasible nor sufficiently fast for calibration of the type of the afore-mentioned general model.

State-of-the-art numerical integration based methods, such as the COS method of Fang and Oosterlee (2008) or the Shannon-wavelet method of Ortiz-Gracia and Oosterlee (2016), if applicable, are very fast, with the Shannon-wavelet method being significantly more robust. These methods are originally developed upon the availability of a closed-form expression for the characteristic function of the underlying process. For a number of processes, this characteristic function is available, due to the well-known Lévy-Khinchine theorem for Lévy processes or by other means, such as solving an associated PDE (Duffie et al., 2000; Heston, 1993). As a characteristic function is the Fourier transform of the associated density, knowing a closed-form expression for the characteristic function of the underlying process allows us to recover, via an inversion process,

[1] A stochastic factor is usually understood as a source of randomness which is typically modelled by a Brownian motion.

[2] Having a non-trivial correlation between the underlying asset price and its variance is important for capturing the skewness in the underlying asset price.

the coefficients of the projection of the density function onto the respective set of basis functions. These coefficients can then be used in the pricing integral that involves the density of the underlying process. However, such a closed-form expression for the characteristic function of the underlying process is difficult, perhaps impossible, to obtain for many interesting and realistic models. This also holds for the afore-mentioned type of general models, due to its non-affinity, and hence the approach to find the characteristic function of the underlying via solving an associated PDE of (Duffie et al., 2000) is not applicable.

In this paper, we extend the applicabilities of these state-of-the-art numerical integration methods to a general jump-diffusion model having square-root stochastic variance and multi-factor Gaussian interest rates. We focus on the Shannon wavelet method of Ortiz-Gracia and Oosterlee (2016), due to its established robustness. We show that, within a dimension reduction framework, this Shannon wavelet method can be adapted for effective use with this type of models. Due to the very impressive efficiency of the proposed Shannon wavelet method under this general model, we solely devote this paper to European-style options. Its application to tackle early-exercise and barrier features under this model will be covered in a follow-up paper.

To avoid difficulties in obtaining a closed-form expression for the characteristic function of the underlying process under the considered general model, the proposed Shannon wavelet method is developed within the dimension reduction framework put forward in Dang et al. (2015b, 2017). This framework involves (i) applying the conditional MC technique to the variance factor, and (ii) removing completely the noise in the interest rate factors via exact integrations. Under this framework, the option price and hedging parameters can be expressed as a two-dimensional integral that involves only (i) the value of the variance at the terminal time, and (ii) the time-integrated variance process conditional on this value. There are several novel computational aspects and significant efficiency benefits that are central to the evaluation of this two-dimensional integral via Shannon-wavelets.

- The recovery of the density of the conditional time-integrated variance process from its known conditional characteristic function is performed by means of the highly efficient Shannon wavelet inverse Fourier technique, referred to as SWIFT, developed in Ortiz-Gracia and Oosterlee (2016). This approach of approximating the density of the conditional time-integrated variance process is much more computationally efficient than existing methods. For example, in the technique proposed in Broadie and Kaya (2006), the cumulative distribution function is first recovered, and a root-finding method is then applied to generate samples of the density, and hence resulting in a great computational effort.

- Once the conditional density is recovered by the SWIFT technique, the initial two-dimensional integral can be further reduced to a one-dimensional integral that involves only the known density of the terminal value of the variance. This is due to the fact that, as stated in Stenger (2011), the integral of the product of a certain function and a Shannon basis can be accurately approximated just by the function evaluated at a certain point, provided that the modulus of its Fourier transform decays rapidly, which is the case considered in our work.

- A major computational advantage of the proposed method is that, regardless of the number of stochastic factors in the models, the method only relies on the inversion of the known characteristic function of the conditional time-integrated variance process. This is obviously an advantage over numerical integration methods that require a known closed-form expres-

sion for the characteristic function of the underlying process, such as Fast Fourier Transform (FFT) based methods in Pillay and O'Hara (2011); Zhang and Wang (2013). Furthermore, with Shannon wavelets, we can develop sharp approximation error bounds for the option price. It is not clear how this can be achieved by other techniques.

The numerical experiments confirm the robustness and impressive efficiency of the proposed pricing technique, while the computational complexity remains independent of the number of stochastic factors in the model. In about 0.05 seconds on a personal computer, the method can compute the price of a European option under a 6-factor jump-diffusion model within 0.01% relative error of a benchmark solution obtained via a multi-level MC method.

The remainder of the paper is organized as follows. We start by introducing a general pricing model and reviewing the dimension reduction framework in Sections 2 and 3, respectively. In Section 4, we discuss in detail the dimension reduction SWIFT, herein after referred to as drSWIFT. Section 5 develops error bounds for the option price. In Section 6, we present several numerical results to illustrate the robustness, error bounds, and efficiency of the drSWIFT method. Section 7 concludes the paper and outlines possible future work.

2 A general jump-diffusion model

We consider an (international) economy consisting of $c+1$ markets (currencies), $c \in \{0,1\}$, indexed by $i \in \{d,f\}$, where "d" stands for the domestic market (Dang et al., 2015b). We consider a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$, with sample space Ω, sigma-algebra $\mathcal{F}$, filtration $\{\mathcal{F}_t\}_{t\geq 0}$, and "$d$" risk-neutral measure $\mathbb{Q}$ defined on $\mathcal{F}$. We denote by $\mathbb{E}$ the expectation taken under $\mathbb{Q}$ measure. Let the underlying asset $S(t)$, its instantaneous variance $\nu(t)$, and the two short rates $r_d(t)$ and $r_f(t)$ be governed by the following SDEs under the measure $\mathbb{Q}$:

$$\frac{\mathrm{d}S(t)}{S(t^-)} = (r_d(t) - c\, r_f(t) - \lambda\delta)\ \mathrm{d}t + \sqrt{\nu(t)}\,\mathrm{d}W_s(t) + \mathrm{d}J(t)\ , \tag{2.1a}$$

$$r_d(t) = \sum_{i=1}^{n} X_i(t) + \gamma_d(t),$$
$$\text{with } \mathrm{d}X_i(t) = -\kappa_{d_i}(t)\, X_i(t)\,\mathrm{d}t + \sigma_{d_i}(t)\,\mathrm{d}W_{d_i}(t)\ , \quad X_i(0) = 0, \tag{2.1b}$$

$$r_f(t) = \sum_{i=1}^{l} Y_i(t) + \gamma_f(t),$$
$$\text{with } \mathrm{d}Y_i(t) = -\kappa_{f_i}(t)\, Y_i(t)\,\mathrm{d}t + \sigma_{f_i}(t)\,\mathrm{d}W_{f_i}(t) - \rho_{s,f_i}\sigma_{f_i}(t)\sqrt{\nu(t)}\,\mathrm{d}t\ , \quad Y_i(0) = 0, \tag{2.1c}$$

$$\mathrm{d}\nu(t) = \kappa_\nu\,(\bar{\nu} - \nu(t))\ \mathrm{d}t + \sigma_\nu\sqrt{\nu(t)}\,\mathrm{d}W_\nu(t)\ . \tag{2.1d}$$

We work under the following assumptions for model (2.1).

- Processes $W_s(t)$ and $W_\nu(t)$ are correlated Brownian motions (BMs) with a constant correlation coefficient $\rho_{(\cdot)(\cdot)} \in [-1,1]$. So are processes $W_{d_i}(t)$, $i = 1,\ldots,n$, and $W_{f_i}(t)$, $i = 1,\ldots,l$, with a constant correlation between each BM pair.

- Processes $W_s(t)$ and $W_\nu(t)$ are independent of processes $W_{d_i}(t)$, $i = 1,\ldots,n$, as well as of processes $W_{f_i}(t)$, $i = 1,\ldots,l$.

- The process $J(t) = \sum_{j=1}^{\pi(t)} (x_j - 1)$ is a compound Poisson process. Specifically, $\pi(t)$ is a Poisson process with a constant finite jump intensity $\lambda > 0$, and x_j, $j = 1,2,\ldots$, are

independent and identically distributed (i.i.d.) positive random variables representing the jump amplitude, and having the density $\chi(\cdot)$.

Several popular cases for $\chi(\cdot)$ are (i) the log-normal distribution given in Merton (1976), and (ii) the log-double-exponential distribution given in Kou (2002). When a jump occurs at time t, we have $S(t) = xS(t^-)$, where t^- is the instant of time just before the time t. In (2.1a), $\delta = \mathbb{E}[x-1]$ represents the expected percentage change in the underlying asset price.

- The Poisson process $\pi(t)$, and the sequence of random variables $\{x_j\}_{j=1}^{\infty}$ are mutually independent, as well as independent of the BMs $W_s(t)$, $W_{d_i}(t)$, $i = 1, \ldots, n$, $W_{f_i}(t)$, $i = 1, \ldots, l$, and $W_\nu(t)$.

- The functions $\kappa_{d_i}(t)$, $\sigma_{d_i}(t)$, $i = 1, \ldots, n$, $n \geq 1$, $\kappa_{f_i}(t)$, and $\sigma_{f_i}(t)$, $i = 1, \ldots, l$, $l \geq 1$, are strictly positive deterministic functions of t, with $\kappa_{d_i}(t)$, and $\kappa_{f_i}(t)$ being the positive mean-reversion rates. The functions $\gamma_d(t)$ and $\gamma_f(t)$ are also deterministic, and they, respectively, capture the "d" and "f" current term structures. They are defined as

$$\gamma_i(t) = r_i(0)\, e^{-\kappa_{i_1} t} + \kappa_{i_1} \int_0^t e^{-\kappa_{i_1}(t-s)}\, \theta_i(s)\, \mathrm{d}s \;, \quad i \in \{d, f\} \;, \tag{2.2}$$

  where θ_i are deterministic, and represent the interest rates' mean levels. In addition, κ_ν, σ_ν and $\bar{\nu}$ are also positive constants.

The constant c takes on the value of either zero or one, and essentially serves as an on/off switch of the "f" economy. That is, by setting $c = 0$, the model (2.1) reduces to an option pricing model in a single market. It can be used for stock options, in which case, $S(t)$ denotes the underlying stock price. When $c = 1$, the model (2.1) becomes a foreign exchange (FX) model, with indexes "d" and "f" respectively denoting the domestic and foreign markets (currencies). In this case, $S(t)$ denotes the spot FX rate, which is defined as the number of units of "d" currency per one unit of "f" currency.

We emphasize the generality of the model. A number of widely used pricing models are a special case of (2.1). For example, for stock options, (2.1) covers the Heston model due to Heston (1993), its jump-extension, or the Bates model (Bates, 1996), as well as the popular (3D) Heston-Hull-White (HHW) equity model used in Grzelak and Oosterlee (2012b); Haentjens and in 't Hout (2012). For FX options, the widely used four-factor model with stochastic volatility and one-factor Gaussian interest rates is also a special case of (2.1) (see, for example, Grzelak and Oosterlee (2011, 2012a); Haastrecht et al. (2009); Haastrecht and Pelsser (2011)). Furthermore, this model is highly suitable for long-dated products, such as Power-Reverse Dual-Currency (PRDC) swaps (Sippel and Ohkoshi, 2002), a very popular cross-currency exotic contract, because the prices of these complex FX products are very sensitive to the volatility skews or smiles (Dang et al., 2014, 2015a; Piterbarg, 2006).

3 Review of the dimension reduction framework

We denote by $b = n + 2 + c\,l$, where $c \in \{0, 1\}$, the total number of stochastic factors in the model. To decompose the (correlated) BM processes into a linear combination of independent BM processes, we apply the standard decomposition procedure involving matrix $\mathbf{A} \equiv [a_{ij}] \in \mathbb{R}^{b \times b}$ obtained by a Cholesky factorization. Here, $\mathbf{A}$ is an upper triangular matrix with $a_{b,b} = 1$. The normalization condition on the correlation matrix requires $\sum_{j=1}^{b} a_{i,j}^2 = 1$ for each row. Under the

afore-mentioned independency assumptions between $S(t)$, as well as $\nu(t)$, and $r_d(t)$ and $r_f(t)$, we have that $a_{1,j} = a_{j,b} = 0$, $j = 2, \ldots, b-1$.

We denote by

$$V(S(t), t, \cdot) \equiv V(S(t), t, r_d(t), r_f(t), \nu(t))$$

the price at time t of a plain-vanilla European option under the model (2.1) with payoff $\Phi(S(T))$ We further assume that the payoff $\Phi(x)$ is a continuous function of its argument having at most polynomial (sub-exponential) growth. This condition is satisfied in the case of call and put options, where $\Phi(S(T)) = \max(S(T) - K, 0)$ and $\Phi(S(T)) = \max(K - S(T), 0)$, respectively. Here, K is the strike of the option.

In the following, we briefly review the main steps of the dimension reduction approach for the jump-diffusion model (2.1). The reader is referred to Dang et al. (2015b, 2017) for detailed discussions of the approach and relevant proofs.

- Step 1: Using standard arbitrage theory (Delbaen and Schachermayer, 1994) and the "tower property" of the conditional expectation, the option price under our general model can be expressed as a two-level nested expectation, with the inner expectation being conditioned on all Brownian motions, except the one associated with the underlying asset.

- Step 2: Under certain regularity conditions, which are satisfied in the present case, by the Feynman-Kac theorem for jump-diffusion processes (Cont and Tankov, 2004), the inner expectation in Step 1 can be shown to be equal to the unique solution to an associated (conditional) Partial Integro-Differential Equation (PIDE) (Dang et al., 2017)[Lemma 3.1].

- Step 3: To solve the conditional PIDE, we first transform it into the Fourier space to obtain an ordinary differential equation in terms of a transformed option price. This ordinary differential equation can then be easily solved in closed-form from maturity $t = T$ to time $t = 0$ to obtain the transformed solution of the conditional PIDE at time $t = 0$. This closed-form solution contains the term $\exp(\lambda T \Gamma(\xi))$, which arises from the jump component, where $\Gamma(\xi)$ is the characteristic function of $\ln(x)$, i.e. the log of the jump amplitude x. This leaves only an outer expectation over the Brownian motion associated with the variance to be approximated by numerical methods.

  Another crucial step in our approach is to remove the variances associated with *all* the interest rate factors when evaluating the (outer) expectation. This step is achieved by applying iterated conditioning on the Brownian motion associated with the variance factor, and solving in closed-form for the expectations of expressions of the interest rates conditioned on this Brownian motion

- Step 4: The final step in the dimension reduction framework is to inverse the result Step 3 to obtain the option price. This step can be achieved by the convolution theorem in combination with expanding the term $\exp(\lambda T \Gamma(\xi))$ in a Taylor series.

In the case that the log of the jump amplitude $\ln(x) \sim \text{Normal}(\tilde{\mu}, \tilde{\sigma}^2)$ (Merton, 1976), the European call option value is given by (Dang et al., 2017)[Corollary 3.2]

$$V(S(0), 0, \cdot) = \mathbb{E}\left[\sum_{j=0}^{\infty} \frac{(\lambda T)^j}{j!} \left\{ \exp\left(j\tilde{\mu} + \frac{j\tilde{\sigma}^2}{2}\right) S(0) \mathrm{e}^{(G+F+H)} \mathcal{N}(d_{1,j}) - K \mathrm{e}^{H} \mathcal{N}(d_{2,j}) \right\}\right], \quad (3.1)$$

where

$$d_{1,j} = \frac{\ln\left(\frac{S(0)}{K}\right) + j\tilde{\mu} + F}{\sqrt{2\left(G + \frac{j\tilde{\sigma}^2}{2}\right)}} + \sqrt{2\left(G + \frac{j\tilde{\sigma}^2}{2}\right)}, \quad d_{2,j} = d_{1,j} - \sqrt{2\left(G + \frac{j\tilde{\sigma}^2}{2}\right)}\,. \tag{3.2}$$

Here, the coefficients G, F, and H, are given by

$$G = \frac{a_{11}^2}{2}\int_0^T \nu(t)\,\mathrm{d}t + \frac{1}{2}\sum_{k=2}^{b-1}\int_0^T \left(\sum_{j=1}^{n} a_{(j+1),k}\,\beta_{d_j}(t) - c\sum_{j=1}^{l} a_{(j+n+1),k}\,\beta_{f_j}(t)\right)^2 \mathrm{d}t, \tag{3.3a}$$

$$\begin{aligned} F = &-\frac{1}{2}\int_0^T \nu(t)\,\mathrm{d}t + \int_0^T \left(\gamma_d(t) - c\gamma_f(t)\right)\,\mathrm{d}t + a_{1,b}\int_0^T \sqrt{\nu(t)}\,\mathrm{d}W_\nu(t) - \lambda\delta T \\ &- \sum_{k=2}^{b-1}\int_0^T \left(\sum_{j=1}^{n} a_{(j+1),k}\,\beta_{d_j}(t)\left(\sum_{j=1}^{n} a_{(j+1),k}\,\beta_{d_j}(t) - c\sum_{j=1}^{l} a_{(j+n+1),k}\,\beta_{f_j}(t)\right)\right)\mathrm{d}t \end{aligned} \tag{3.3b}$$

$$H = -\int_0^T \gamma_d(t)\,\mathrm{d}t + \frac{1}{2}\sum_{k=2}^{b-1}\int_0^T \left(\sum_{j=1}^{n} a_{(j+1),k}\beta_{d_j}(t)\right)^2 \mathrm{d}t - \lambda T, \tag{3.3c}$$

In (3.3a)-(3.3c), $\beta_{d_i}(t)$, $i = 1, \ldots, n$, and $\beta_{f_i}(t)$, $i = 1, \ldots, l$, are defined as

$$\beta_{d_i}(t) = \sigma_{d_i}(t)\int_t^T \mathrm{e}^{-\int_t^{t'}\kappa_{d_i}(t'')\,\mathrm{d}t''}\,\mathrm{d}t', \quad \beta_{f_i}(t) = \sigma_{f_i}(t)\int_t^T \mathrm{e}^{-\int_t^{t'}\kappa_{f_i}(t'')\,\mathrm{d}t''}\,\mathrm{d}t'\,. \tag{3.4}$$

We emphasize that quantity H is deterministic, while G and F are stochastic, but depend on the variance factor. The variance coming from the r_d's BMs and the r_f's BMs, if any, is completely removed from the computation. The Delta of the option is (Dang et al., 2017)[Corollary 4.2]

$$\left.\frac{\partial V}{\partial S}\right|_{(S(0),0,\cdot)} = \mathbb{E}\left[\sum_{j=0}^{\infty}\frac{(\lambda T)^j}{j!}\left\{\exp\left(j\tilde{\mu} + \frac{j\tilde{\sigma}^2}{2} + G + F + H\right)\mathcal{N}\left(d_{1,j}\right)\right\}\right], \tag{3.5}$$

where $d_{1,j}$ and $d_{2,j}$ are defined in (3.2). A formula of the Gamma of the option can be found in (Dang et al., 2017)[Corollary 4.2]. See Dang et al. (2017)[Corrolary 3.1] for the results of the double-exponential distribution (Kou, 2002). The formulas for the pure-diffusion model can be easily obtained by setting the jump intensity $\lambda = 0$ and using only $j = 0$ in (3.1)-(3.5) (also see (Dang et al., 2015b)). In our subsequently analysis, we focus on the price of the option under the normal jump case (Merton, 1976). The analysis for the option hedging parameters are the same, and hence omitted.

4 Dimension reduction Shannon wavelet method

Examination of (3.3) shows that G depends only on the time-integrated variance process $\int_0^T \nu(t)\mathrm{d}t$, while F depends on both $\int_0^T \nu(t)\mathrm{d}t$ and $\int_0^T \sqrt{\nu(t)}\,\mathrm{d}W_\nu(t)$. We note from (2.1d) that

$$\int_0^T \sqrt{\nu(t)}\,\mathrm{d}W_\nu(t) = \frac{\nu(T) - \nu(0) - \kappa_\nu\bar{\nu}T + \kappa_\nu\int_0^T \nu(t)\mathrm{d}t}{\sigma_\nu}.$$

As a result, F can be expressed in terms of $\int_0^T \nu(t)\mathrm{d}t$ and the terminal value $\nu(T)$ of the variance. It follows from (3.1) that the option value can be written as

$$V(S(0),0,\cdot) = \mathbb{E}\left[g\left(\int_0^T \nu(t)\mathrm{d}t, \nu(T)\right)\right] = \mathbb{E}\left[\mathbb{E}\left[g\left(\int_0^T \nu(t)\mathrm{d}t, \nu(T)\right)\middle|\nu(T)\right]\right], \tag{4.1}$$

for a function $g(\cdot,\cdot)$ that may take different forms, depending on the model under investigation. Here, the second equality, which comes from the "tower property" of the conditional expectation, allows us to take advantage of the known characteristic function of the time-integrated variance process conditional on the terminal value of the variance $\nu(T)$.

Let $f(\cdot|y) \equiv f(\cdot;\nu(T)=y)$ denote the density of the time-integrated variance process conditional on the terminal value of the variance $\nu(T)$ being y, where $y \in [0, y_0]$ for a $y_0 > 0$. This process can be roughly approximated by a central discretization

$$\int_0^T \nu(t)\mathrm{d}t \approx \frac{T}{2}(\nu(0)+\nu(T)) \ .$$

Taking this into account, without loss of generality, we assume that the conditional density function $f(\cdot|y)$ is supported on the interval $[0,T]$. It is worth remarking that the SWIFT method employed to recover the density is capable to compute the mass underneath the curve as a byproduct, and therefore, this interval can be adaptively modified, if necessary. From (4.1), the option price can be represented by the following double integral

$$V(S(0),0,\cdot) = \int_0^{y_0} \left[\int_0^T g(x,y) f(x|y)\,\mathrm{d}x\right] w(y)\,\mathrm{d}y \ . \tag{4.2}$$

Here, $w(\cdot)$ is the density of the terminal value of the variance $\nu(T)$, which is known in closed-form (Cox et al., 1985b)

$$w(y) := \zeta e^{-\zeta(\nu(0)e^{-\kappa_\nu T}+y)} \cdot \left(\frac{y}{\nu(0)e^{-\kappa_\nu T}}\right)^{\frac{q}{2}} \cdot I_q\left(2\zeta e^{-\frac{1}{2}\kappa_\nu T}\sqrt{\nu(0)y}\right) \ , \tag{4.3}$$

where $q := \frac{2\kappa_\nu\bar{\nu}}{\sigma_\nu^2} - 1$, $\zeta := \frac{2\kappa_\nu}{(1-e^{-\kappa_\nu T})\sigma_\nu^2}$ and $I_q(x)$ is the modified Bessel function of the first kind with order q.

To evaluate the integral (4.2), the conditional density $f(\cdot|y)$, $y \in [0,y_0]$, first needs to be approximated, since it is not known in closed-form. Then, a quadrature rule can be applied to approximate the price or the hedging parameters of the option. In our approach, we recover the conditional density $f(\cdot|y)$ from its Fourier transform, i.e. the characteristic function of the time-integrated variance conditional on the terminal value, denoted by $\Psi(\xi|\nu(T))$, for which a closed-form is (Broadie and Kaya, 2006)

$$\begin{aligned}
\Psi(\xi|y) &= \mathbb{E}\left[\exp\left(-i\xi\int_0^T \nu(t)\mathrm{d}t\right)\middle|\nu(T)=y,\nu(0)\right] \\
&= \frac{I_q\left(\sqrt{\nu(T)\nu(0)}\frac{4\gamma(\xi)e^{-\frac{1}{2}\gamma(\xi)T}}{\sigma_\nu^2(1-e^{-\gamma(\xi)T})}\right)}{I_q\left(\sqrt{\nu(T)\nu(0)}\frac{4\kappa_\nu e^{-\frac{1}{2}\kappa_\nu T}}{\sigma_\nu^2(1-e^{-\kappa_\nu T})}\right)} \times \frac{\gamma(\xi)e^{-\frac{1}{2}(\gamma(\xi)-\kappa_\nu)T}(1-e^{-\kappa_\nu T})}{\kappa_\nu(1-e^{-\gamma(\xi)T})} \\
&\times \exp\left(\frac{\nu(0)+\nu(T)}{\sigma_\nu^2}\left[\frac{\kappa_\nu(1+e^{-\kappa_\nu T})}{1-e^{-\kappa_\nu T}} - \frac{\gamma(\xi)(1+e^{-\gamma(\xi)T})}{1-e^{-\gamma(\xi)T}}\right]\right) \ .
\end{aligned} \tag{4.4}$$

Here, $\gamma(\xi) := \sqrt{\kappa_\nu^2 - 2i\sigma_\nu^2\xi}$. This step can be very efficiently achieved by means of the SWIFT technique (Ortiz-Gracia and Oosterlee, 2016). We then show that the double integral (4.2) can be further simplified to a single integral, thanks to certain local approximation properties of the Shannon wavelets. Furthermore, we can also develop sharp approximation error bound for the option price. In the following subsection, we first present a brief review on Shannon wavelets and the SWIFT method, and then discuss the dimension reduction SWIFT (drSWIFT) method in detail.

4.1 Shannon wavelets and SWIFT

4.1.1 Multi-resolution analysis and Shannon wavelets

Consider the space of square-integrable functions, denoted by $L^2(\mathbb{R})$, where

$$L^2(\mathbb{R}) = \left\{ f : \int_{-\infty}^{+\infty} |f(x)|^2 \, \mathrm{d}x < \infty \right\} .$$

A general structure for wavelets in $L^2(\mathbb{R})$ is called a *multi-resolution analysis*. We start with a family of closed nested subspaces in $L^2(\mathbb{R})$

$$\ldots \subset \mathcal{V}_{-2} \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \ldots ,$$

where

$$\bigcap_{m\in\mathbb{Z}} \mathcal{V}_m = \{0\} , \qquad \overline{\bigcup_{m\in\mathbb{Z}} \mathcal{V}_m} = L^2(\mathbb{R}) ,$$

and

$$f(x) \in \mathcal{V}_m \iff f(2x) \in \mathcal{V}_{m+1} .$$

If these conditions are met, then there exists a function $\varphi \in \mathcal{V}_0$ that generates an orthonormal basis, denoted by $\{\varphi_{m,k}\}_{k\in\mathbb{Z}}$, for each $\mathcal{V}_m$ subspace, where

$$\varphi_{m,k}(x) = 2^{m/2}\varphi(2^m x - k) .$$

The function $\varphi(\cdot)$ is usually referred to as the *scaling function* or *father wavelet*.

For any $f \in L^2(\mathbb{R})$, a projection map of $L^2(\mathbb{R})$ onto $\mathcal{V}_m$, denoted by $\mathcal{P}_m : L^2(\mathbb{R}) \to \mathcal{V}_m$, is defined by means of

$$\mathcal{P}_m f(x) = \sum_{k\in\mathbb{Z}} c_{m,k}\varphi_{m,k}(x) . \tag{4.5}$$

Here,

$$c_{m,k} = \langle f, \varphi_{m,k} \rangle , \tag{4.6}$$

where $< f, g >= \int_{\mathbb{R}} f(x)\overline{g(x)}\,\mathrm{d}x$ denotes the inner product in $L^2(\mathbb{R})$, with $\overline{g(\cdot)}$ being the complex conjugation of $g(\cdot)$, and $\mathcal{P}_m f$ converges to f in $L^2(\mathbb{R})$, i.e. $\|f - \mathcal{P}_m f\|_2 \to 0$, when $m \to +\infty$.

Considering higher m values (i.e. when more terms are used), the accuracy of the truncated series representation of the function f improves. As opposed to Fourier series, a key fact regarding the use of wavelets is that wavelets can be moved (by means of the k value), stretched or compressed (by means of the m value) to accurately represent the local properties of a function.

In this paper, we employ Shannon wavelets (Cattani, 2008). Shannon wavelets represent the real part of the so-called harmonic wavelets. They have a slow decay in the time domain but a very sharp compact support in the frequency, i.e. Fourier, domain. A set of Shannon scaling functions $\varphi_{m,k}(\cdot)$ in the subspace $\mathcal{V}_m$ is defined as

$$\varphi_{m,k}(x) = 2^{m/2}\frac{\sin(\pi(2^m x - k))}{\pi(2^m x - k)} = 2^{m/2}\varphi(2^m x - k) , \quad k \in \mathbb{Z} , \tag{4.7}$$

where

$$\varphi(x) = \mathrm{sinc}(x) = \begin{cases} \dfrac{\sin(\pi x)}{\pi x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0, \end{cases} \tag{4.8}$$

is the basic (Shannon) scaling function. We note that the Fourier transform of $\varphi_{m,k}(x)$ can easily be obtained.

4.1.2 SWIFT method

In this subsection, we present the SWIFT method which can be used to effectively invert the function $\Psi(\xi|y)$, given in (4.4), to obtain an approximation to the conditional density function $f(\cdot|y)$ to be used in (4.2).

We assume that $f(\cdot|y) \in L^2(\mathbb{R})$, and we consider its expansion in terms of the Shannon scaling functions at the level of resolution m. Our aim is to recover the coefficients of this approximation from the Fourier transform of the function $f(\cdot|y)$ which, as mentioned before, is known in closed-form (4.4). Following the wavelets theory

$$f(x|y) \approx \mathcal{P}_m f(x|y) = \sum_{k\in\mathbb{Z}} c_{m,k}(y)\varphi_{m,k}(x) \ , \tag{4.9}$$

In our context, the infinite series in (4.9) can be well-approximated by a finite summation without loss of density mass, since the function f is supported on the finite interval $[0, T]$. More specifically, we have the following approximation

$$\mathcal{P}_m f(x|y) \approx f_m(x|y) := \sum_{k=0}^{\lceil 2^m T\rceil} c_{m,k}(y)\varphi_{m,k}(x) \ , \tag{4.10}$$

where $\lceil x\rceil$ denotes the smallest integer greater than or equal to x.

The next step is the computation of the coefficients in (4.10). Recalling (4.6) and (4.7), we have that

$$c_{m,k}(y) = \langle f(\cdot|y), \varphi_{m,k}\rangle = \int_{\mathbb{R}} f(x|y)\overline{\varphi}_{m,k}(x)\,\mathrm{d}x = 2^{m/2}\int_{\mathbb{R}} f(x|y)\varphi(2^m x - k)\mathrm{d}x \ . \tag{4.11}$$

Using the classical Vieta's formula (Gearhart and Shultz, 1990), the cardinal sinus can be expressed as the following infinite product

$$\varphi(t) = \mathrm{sinc}(t) = \prod_{j=1}^{+\infty} \cos\left(\frac{\pi t}{2^j}\right) \ . \tag{4.12}$$

If we truncate the infinite product (4.12) to a finite product with a total of J terms, then, thanks to the cosine product-to-sum identity, we have

$$\prod_{j=1}^{J} \cos\left(\frac{\pi t}{2^j}\right) = \frac{1}{2^{J-1}}\sum_{j=1}^{2^{J-1}} \cos\left(\frac{2j-1}{2^J}\pi t\right) \ . \tag{4.13}$$

The parameter J plays an important role in the efficiency of the method, and hereinafter is referred to as the *truncation parameter*. By (4.12) and (4.13) the $\varphi(t) = \mathrm{sinc}(\cdot)$ function can thus be approximated as

$$\varphi(t) = \mathrm{sinc}(t) \approx \mathrm{sinc}^*(t) := \frac{1}{2^{J-1}}\sum_{j=1}^{2^{J-1}} \cos\left(\frac{2j-1}{2^J}\pi t\right) \ . \tag{4.14}$$

Replacing the function $\varphi(\cdot)$ in (4.11) by the approximation (4.14) gives the following approximation for coefficients $c_{m,k}(y)$:

$$c_{m,k}(y) \approx c^*_{m,k}(y) := \frac{2^{m/2}}{2^{J-1}}\sum_{j=1}^{2^{J-1}} \int_{\mathbb{R}} f(x|y)\cos\left(\frac{2j-1}{2^J}\pi(2^m x - k)\right)\mathrm{d}x \ . \tag{4.15}$$

Next, by taking into account that any function f with Fourier transform $\hat{f}$ satisfies $\Re\left(\hat{f}(\xi)\right) = \int_{\mathbb{R}} f(x)\cos(\xi x)\,\mathrm{d}x$, where $\Re(z)$ denotes the real part of z, and observing that

$$\hat{f}(\xi)e^{ik\pi\frac{2j-1}{2^J}} = \int_{\mathbb{R}} e^{-i\left(\xi x - \frac{k\pi(2j-1)}{2^J}\right)} f(x)\mathrm{d}x \ ,$$

we can simplify (4.15) to

$$c_{m,k}(y) \approx c^*_{m,k}(y) = \frac{2^{m/2}}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \Re\left[\Psi\left(\frac{(2j-1)\pi 2^m}{2^J}\middle| y\right) e^{\frac{ik\pi(2j-1)}{2^J}}\right] \ . \tag{4.16}$$

Putting everything together gives the following approximation to $f(x|y)$:

$$f(x|y) \approx f^*_m(x|y) := \sum_{k=0}^{\lceil 2^m T\rceil} c^*_{m,k}(y)\varphi_{m,k}(x) \ , \tag{4.17}$$

where $\varphi_{m,k}(x)$ and $c^*_{m,k}(y)$ are defined in (4.7) and (4.16), respectively.

4.2 Option pricing with drSWIFT

For a fixed level of resolution m and a fixed truncation parameter J used in (4.13), replacing the conditional density function $f(\cdot|y)$ in (4.2) by the finite approximation $f^*_m(\cdot|y)$ in (4.17), gives us the approximation $V_1(S(0),0,\cdot)$ to the option price $V(S(0),0,\cdot)$

$$V(S(0),0,\cdot) \approx V_1(S(0),0,\cdot) := \int_0^{y_0} \left[\sum_{k=0}^{\lceil 2^m T\rceil} c^*_{m,k}(y) \int_0^T g(x,y)\varphi_{m,k}(x)\,\mathrm{d}x\right] w(y)\,\mathrm{d}y \ , \tag{4.18}$$

with coefficients $c^*_{m,k}$ defined in (4.16).

It turns out that, thanks to certain local approximation properties of wavelets, the expression (4.18) can be further simplified to a single integral by using a highly accurate approximation for the inner integral terms. To this end, we recall the following theorem in Stenger (2011)

Theorem 4.1 (Theorem 1.3.2 of Stenger (2011))**.** *Let f be defined on $\mathbb{R}$, and let its Fourier transform, denoted by $\hat{f}$, be such that, for some positive constant d*

$$|\hat{f}(\xi)| = \mathcal{O}\left(e^{-d|\xi|}\right) \ , \quad \xi \to \pm\infty \ . \tag{4.19}$$

Then, as $a \to 0$,

$$\frac{1}{a}\int_{\mathbb{R}} f(y)\mathcal{S}(k,a)(y)\,\mathrm{d}y - f(ka) = \mathcal{O}\left(e^{-\frac{\pi d}{a}}\right) \ ,$$

where $\mathcal{S}(k,a)(y) := sinc\left(\frac{y}{a} - k\right)$.

To apply this theorem to function $g\left(\int_0^T \nu(t)\mathrm{d}t, \nu(T)\right)$, we need to check whether its Fourier transform satisfies the condition (4.19). It turns out that the Fourier transform of $g(\cdot,\cdot)$ is the term

$$\hat{\Phi}(\xi)\exp\left(-G\,\xi^2 + iF\xi + H + \lambda T\Gamma(\xi)\right),$$

where $\hat{\Phi}(\xi)$ is the Fourier transform of the payoff. First, we notice that coefficient G in the quadratic term in the exponent of this term is strictly positive (see (3.3)). Furthermore, G, F, and H are a also bounded, due to the boundedness of the variance $\nu(t)$ (Andersen and Piterbarg, 2007). It follows that the Fourier transform of $g(\cdot,\cdot)$ satisfies the hypothesis of Theorem 4.1.

Hence, we can apply Theorem 4.1 with $a = \frac{1}{2^m}$ to the inner integral terms in expression (4.18). This gives

$$\int_0^T g(x,y)\varphi_{m,k}(x)\mathrm{d}x \approx \frac{1}{2^{m/2}} g\left(\frac{k}{2^m}, y\right) \ . \tag{4.20}$$

Thus, we arrive at the approximation $V_2(S(0),0,\cdot)$ to $V_1(S(0),0,\cdot)$

$$V_1(S(0),0,\cdot) \approx V_2(S(0),0,\cdot) := \frac{1}{2^{m/2}} \int_0^{y_0} \left[\sum_{k=0}^{\lceil 2^m T \rceil} c^*_{m,k}(y)\, g\left(\frac{k}{2^m}, y\right) \right] w(y)\, \mathrm{d}y \ , \tag{4.21}$$

where $c^*_{m,k}(y)$ are defined in (4.16). Finally, the integral in (4.21) can be approximated by means of the composite trapezoidal rule.

It is worth remarking that (4.21) is in terms of the variance process. As pointed out by Fang and Oosterlee (2011), the Feller condition for the variance process, which is equivalent to $q = \frac{2\kappa_\nu \bar{\nu}}{\sigma_\nu^2} - 1 \geq 0$, is difficult to satisfy in a practical situation. Specifically, one often finds $2\kappa_\nu\bar{\nu} < \sigma_\nu^2$ from market data, this is $q < 0$, in which case the left tail of the variance density $w(y)$, defined in (4.3), grows extremely fast in value, and this may affect the accuracy of the composite trapezoidal rule applied to (4.21). Based on these insights, we perform the change of variables $v = \ln(y)$ in (4.21), and transform the problem from the (terminal) variance domain to the (terminal) log-variance domain

$$V_2(S(0),0,\cdot) = \frac{1}{2^{m/2}} \int_{-\infty}^{\ln(y_0)} \left[\sum_{k=0}^{\lceil 2^m T \rceil} c^*_{m,k}(e^v)\, g\left(\frac{k}{2^m}, e^v\right) \right] \bar{w}(v)\, \mathrm{d}v \ , \tag{4.22}$$

where

$$\bar{w}(v) = e^v \tilde{w}(v), \text{ with } \quad \tilde{w}(v) := \zeta e^{-\zeta(\nu(0)e^{-\kappa_\nu T} + e^v)} \cdot \left(\frac{e^v}{\nu(0)e^{-\kappa_\nu T}}\right)^{\frac{q}{2}} \cdot I_q\left(2\zeta e^{-\frac{1}{2}\kappa_\nu T}\sqrt{\nu(0)e^v}\right) \ . \tag{4.23}$$

The first step to follow for a practical implementation of the option pricing formula (4.22) is to determine an appropriate truncated integration domain for the log-variance density $\bar{w}(v)$. In what follows, we briefly describe an iterative procedure to determine this truncated integration domain, denoted by $[a_v, b_v]$, according to a pre-defined tolerance ϵ_{tol}. We denote by $[a_v^{(j)}, b_v^{(j)}]$, $j = 0, 1, \ldots$, the interval at the j-th iteration. Given an initial guess $[a_v^{(0)}, b_v^{(0)}]$, we iteratively modify the interval until the condition $\bar{w}(v) < \epsilon_{\text{tol}}$ for $v \in \mathcal{D}$ is met, where $\mathcal{D} = (-\infty, a_v^{(j)}) \cup (b_v^{(j)}, \ln(y_0))$, for some j, after which the truncated integration domain is taken to be $[a_v^{(j)}, b_v^{(j)}]$.

We start by estimating a proper initial guess $[a_v^{(0)}, b_v^{(0)}]$. As pointed out in Cox et al. (1985b), the expected value and the variance of $\nu(T)$ can be calculated as

$$\begin{aligned} \mathbb{E}[\nu(T)] &= \nu(0)e^{-\kappa_\nu T} + \bar{\nu}\left(1 - e^{-\kappa_\nu T}\right) \ , \\ \mathbb{V}[\nu(T)] &= \nu(0)\frac{\sigma_\nu^2}{\kappa_\nu} e^{-\kappa_\nu T} - e^{-2\kappa_\nu T} + \bar{\nu}\frac{\sigma_\nu^2}{2\kappa_\nu}\left(1 - e^{-\kappa_\nu T}\right)^2 \ . \end{aligned} \tag{4.24}$$

If we consider a first-order Taylor expansion of $\ln(\nu(T))$ then the expected value and the variance of the log-variance process at terminal time T can be approximated as follows

$$\mathbb{E}[\ln(\nu(T))] \approx \ln\left(\mathbb{E}[\nu(T)]\right), \quad \mathbb{V}[\ln(\nu(T))] \approx \frac{\mathbb{V}[\nu(T)]}{\mathbb{E}[\nu(T)]^2} \ . \tag{4.25}$$

Taking into account that the left tail of the density decays slower than the right tail, we therefore consider the following initial interval

$$[a_v^{(0)}, b_v^{(0)}] := \left[\ln\left(\mathbb{E}[\nu(T)]\right) - 7\frac{\mathbb{V}[\nu(T)]}{\mathbb{E}[\nu(T)]^2}, \ln\left(\mathbb{E}[\nu(T)]\right) + 3\frac{\mathbb{V}[\nu(T)]}{\mathbb{E}[\nu(T)]^2}\right] . \tag{4.26}$$

Now, given $[a_v^{(0)}, b_v^{(0)}]$, we propose two methods for finding the final interval $[a_v^{(j)}, b_v^{(j)}]$. The first one involves the Newton iteration, for which we need the derivative of $\tilde{w}(v)$

$$\tilde{w}'(v) := \zeta e^{-u-\zeta e^v+v}\left(\frac{\zeta e^v}{u}\right)^{\frac{q}{2}} \cdot \left[(-\zeta e^v + q + 1)\cdot I_q\left(2\sqrt{\zeta e^v u}\right) + \zeta\sqrt{\nu(0)e^{v-\kappa_\nu T}}\cdot I_{q+1}\left(2\sqrt{\zeta e^v u}\right)\right] , \tag{4.27}$$

where $u := \zeta\nu(0)e^{-\kappa_\nu T}$. We suggest to use this method when the Feller condition for the variance process is not satisfied. This method is considered in one of the examples for a six-factor model in Section 6. As showed later, numerical results show that only a few iterations are needed to achieve convergence, even for a heavy left-tail distribution. In the second method, we just update the interval $[a_v^{(j)}, b_v^{(j)}]$ by subtracting and adding the approximated value for the variance in (4.25) to $a_v^{(j)}$ and $b_v^{(j)}$, respectively. We suggest to use this method when the Feller condition for the variance process is satisfied.

Once the truncated integration domain $[a_v, b_v]$ has been identified via the above steps, then $V_2(S(0), 0, \cdot)$ in (4.22) can be approximated as follows

$$V_2(S(0),0,\cdot) \approx V_3(S(0),0,\cdot) := \frac{1}{2^{m/2}}\int_{a_v}^{b_v}\left[\sum_{k=0}^{\lceil 2^m T\rceil} c_{m,k}^*(e^v)\, g\left(\frac{k}{2^m}, e^v\right)\right]\bar{w}(v)\,\mathrm{d}v . \tag{4.28}$$

Finally, we consider a partition of the interval $[a_v, b_v]$ into N_I subintervals, and by the composite trapezoidal rule, we obtain the approximation $V_4(S(0), 0, \cdot)$ to $V_3(S(0), 0, \cdot)$

$$V_3(S(0),0,\cdot) \approx V_4(S(0),0,\cdot) := \frac{h}{2}\sum_{\ell=0}^{N_I-1}\left(\mathcal{S}_m(v_\ell) + \mathcal{S}_m(v_{\ell+1})\right) , \tag{4.29}$$

where

$$\mathcal{S}_m(v) = \frac{1}{2^{m/2}}\left[\sum_{k=0}^{\lceil 2^m T\rceil} c_{m,k}^*(e^v) g\left(\frac{k}{2^m}, e^v\right)\right]\bar{w}(v) , \tag{4.30}$$

and $h = \frac{b_v - a_v}{N_I}$ and $v_\ell = a_v + \ell h$, $\ell = 0, \ldots, N_I$.

5 Error analysis

In practice, the parameters to the interest rate dynamics are such that it is possible to compute in closed-form deterministic integrals in (3.3), namely $\int_0^T \gamma_d(t)\,\mathrm{d}t$; $\int_0^T \gamma_f(t)\,\mathrm{d}t$; $\int_0^T \beta_{d_i}(t)\beta_{f_j}(t)\mathrm{d}t$, where $i = 1, \ldots, n$, $j = 1, \ldots, l$; $\int_0^T \beta_{d_i}(t)\beta_{d_i'}(t)\,\mathrm{d}t$, where $i, i' = 1, \ldots, n$; and $\int_0^T \beta_{f_i}(t)\beta_{f_i'}(t)\,\mathrm{d}t$, where $i, i' = 1, \ldots, l$. For the case of a diffusion model, i.e. the jump intensity $\lambda = 0$ and $j = 0$ in (3.1), the function g in (4.30) is known in closed-form. For the case of a jump-diffusion model, g is known analytically, as the infinite series (3.1). However, we can achieve any level of accuracy for this quickly converging series, taking into account the boundedness of the numerator of each term. Furthermore, $\Psi(\cdot|y)$ is known in closed-form. As a result, we can assume that there are no numerical errors in evaluating g in (4.30), and hence the total numerical error of the drSWIFT

method comes from the computation of the integrals in (4.1). In this section, we perform an error analysis on the drSWIFT method, and discuss the computational complexity of the method. We also explain how to determine the value of the level of resolution m and the truncation parameter J in (4.13) to achieve a pre-determined error bound.

There are four sources of error in the evaluation process of the drSWIFT method:

(i) in (4.18) when f is approximated by f_m^* defined in (4.17);

(ii) in (4.21) when the approximation (4.20) is used in place of the inner integral (from zero and maturity time T) in (4.18);

(iii) in (4.28) when truncating the infinite interval $(-\infty, \ln(y_0))$ into the finite interval $[a_v, b_v]$; and

(iv) in (4.29), due to the use of the composite trapezoidal rule as an approximation to the outer integral (from a_v to b_v) in (4.28).

We denote by E the total numerical error of the drSWIFT method in evaluation the outer expectation. This error can be bounded as follows

$$E := |V(S(0),0,\cdot) - V_4(S(0),0,\cdot)| \leq E_{m,1}^* + E_{m,2}^* + E_{m,3}^* + E_{m,h}^* \ , \tag{5.1}$$

where $E_{m,1}^*$, $E_{m,2}^*$, $E_{m,3}^*$, and $E_{m,h}^*$ respectively are the errors in (i)-(iv). Here,

$$E_{m,1}^* := |V(S(0),0,\cdot) - V_1(S(0),0,\cdot)| = \left| \int_0^{y_0} \left[\int_0^T g(x,y) \left(f(x|y) - f_m^*(x|y) \right) \mathrm{d}x \right] w(y) \mathrm{d}y \right| \ , \tag{5.2}$$

where, by (4.17),

$$f_m^*(x|y) := \sum_{k=0}^{\lceil 2^m T \rceil} c_{m,k}^*(y) \varphi_{m,k}(x) \ , \tag{5.3}$$

$$\begin{aligned} E_{m,2}^* :&= |V_1(S(0),0,\cdot) - V_2(S(0),0,\cdot)| \\ &= \left| \int_0^{y_0} \left[\sum_{k=0}^{\lceil 2^m T \rceil} c_{m,k}^*(y) \left(\int_0^T g(x,y) \varphi_{m,k}(x) \mathrm{d}x - \frac{1}{2^{m/2}} g\left(\frac{k}{2^m}, y \right) \right) \right] w(y) \mathrm{d}y \right| \ , \end{aligned} \tag{5.4}$$

and

$$E_{m,3}^* := |V_2(S(0),0,\cdot) - V_3(S(0),0,\cdot)| = \left| \frac{1}{2^{m/2}} \int_{\mathcal{D}} \left[\sum_{k=0}^{\lceil 2^m T \rceil} c_{m,k}^*(e^v)\, g\left(\frac{k}{2^m}, e^v \right) \right] \bar{w}(v)\, \mathrm{d}v \right| \ , \tag{5.5}$$

where

$$\mathcal{D} = (-\infty, a_v) \bigcup (b_v, \log(y_0)),$$

as well as

$$E_{m,h}^* := |V_3(S(0),0,\cdot) - V_4(S(0),0,\cdot)| \ . \tag{5.6}$$

We observe that all $E_{m,1}^*$, $E_{m,2}^*$, $E_{m,3}^*$ and $E_{m,h}^*$ depend on the level of resolution m. In addition, $E_{m,h}^*$ also depends on the number of subintervals N_I via $h = (b_v - a_v)/N_I$.

5.1 Bound for error term $E^*_{m,1}$

We define the projection error, denoted by $\epsilon_p(x,y)$, as

$$\epsilon_p(x,y) = |f(x|y) - \mathcal{P}_m f(x|y)| = \left| f(x|y) - \sum_{k\in\mathbb{Z}} c_{m,k}(y)\varphi_{m,k}(x) \right| , \tag{5.7}$$

where, as defined earlier, $c_{m,k}(y) = \int_{-\infty}^{+\infty} f(x|y)\varphi_{m,k}(x)\mathrm{d}x$. We also define the truncation error, denoted by $\epsilon_t(x,y)$, as

$$\epsilon_t(x,y) = |\mathcal{P}_m f(x|y) - f_m(x|y)| = \left| \sum_{k\notin\{0,\dots,\lceil 2^m T\rceil\}} c_{m,k}(y)\varphi_{m,k}(x) \right| .$$

We denote by $\epsilon_c(x,y)$ the error arising from using approximated coefficients $c^*_{m,k}(y)$ instead of the exact ones $c_{m,k}(y)$. We have

$$\epsilon_c(x,y) = |f_m(x|y) - f^*_m(x|y)| = \left| \sum_{k=0}^{\lceil 2^m T\rceil} (c_{m,k}(y) - c^*_{m,k}(y))\varphi_{m,k}(x) \right| .$$

Then, we have

$$|f(x|y) - f^*_m(x|y)| \le \epsilon_p(x,y) + \epsilon_t(x,y) + \epsilon_c(x,y) .$$

First, we consider the projection error $\epsilon_p(\cdot,\cdot)$. The projection $\mathcal{P}_m f$ can be written as (Maree et al., 2017)

$$\mathcal{P}_m f(x|y) = \frac{1}{2\pi}\int_{-2^m\pi}^{2^m\pi} \Psi(\xi|y)e^{i\xi x}\mathrm{d}\xi . \tag{5.8}$$

By definition of the inverse Fourier transform of f, we have

$$f(x|y) = \frac{1}{2\pi}\int_{\mathbb{R}} \Psi(\xi|y)e^{i\xi x}\mathrm{d}\xi . \tag{5.9}$$

Let

$$K(v,y) = \frac{1}{2\pi}\int_{|\xi|>v} |\Psi(\xi|y)|\mathrm{d}\xi , \tag{5.10}$$

then

$$\epsilon_p(x,y) \le K(2^m\pi, y) . \tag{5.11}$$

Next, we consider the truncation error $\epsilon_t(\cdot,\cdot)$. We observe that

$$c_{m,k}(y) = \int_{-\infty}^{+\infty} f(x|y)\varphi_{m,k}(x)\mathrm{d}x = \int_0^T f(x|y)\varphi_{m,k}(x)\mathrm{d}x ,$$

since the density function f is supported on the interval $[0,T]$. Therefore, the truncation error ϵ_t can be neglected when $k \notin \{0,\dots,\lceil 2^m T\rceil\}$.

Finally, we consider $\epsilon_c(\cdot,\cdot)$. The coefficients $c_{m,k}(y)$ are to be calculated by means of Vieta's formula and the numerical error can be estimated as

$$\epsilon_c(x,y) \le \sum_{k=0}^{\lceil 2^m T\rceil} |c_{m,k}(y) - c^*_{m,k}(y)||\varphi_{m,k}(x)| \le 2^{m/2}\sum_{k=0}^{\lceil 2^m T\rceil} |c_{m,k}(y) - c^*_{m,k}(y)| . \tag{5.12}$$

Since $f(\cdot|\cdot)$ is supported on the interval $[0,T]$, it follows that

$$\begin{aligned}|c_{m,k}(y) - c^*_{m,k}(y)| &= 2^{m/2}\left|\int_0^T f(x|y)\left(\mathrm{sinc}(2^m x - k) - \mathrm{sinc}^*(2^m x - k)\right)\mathrm{d}x\right| \\ &\leq 2^{m/2}\int_0^T f(x|y)\left|\mathrm{sinc}(2^m x - k) - \mathrm{sinc}^*(2^m x - k)\right|\mathrm{d}x\ .\end{aligned} \tag{5.13}$$

Applying the Cauchy-Schwarz inequality to the right-hand-side of the inequality in (5.13) gives

$$|c_{m,k}(y) - c^*_{m,k}(y)| \leq 2^{m/2}\|f(\cdot,y)\|_2\left(\int_0^T \left(\mathrm{sinc}(2^m x - k) - \mathrm{sinc}^*(2^m x - k)\right)^2\mathrm{d}x\right)^{\frac{1}{2}}\ . \tag{5.14}$$

To further bound (5.14), we make use of the following lemma in Ortiz-Gracia and Oosterlee (2016) which gives us an estimate of the error when approximating the sinus cardinal function.

Lemma 5.1 (Lemma 2 of Ortiz-Gracia and Oosterlee (2016))**.** *Define the absolute error* $\mathcal{E}_V(t) := sinc(t) - sinc^*(t)$*. Then,*

$$|\mathcal{E}_V(t)| \leq \frac{(\pi c)^2}{2^{2(J+1)} - (\pi c)^2},$$

for $t \in [-c,c]$*, where* $c \in \mathbb{R}, c > 0$ *and* $J \geq \log_2(\pi c)$*.*

We observe that, since $0 \leq x \leq T$, it follows $-\lceil 2^m T\rceil \leq 2^m x - k \leq \lceil 2^m T\rceil$. Thus, by Lemma 5.1, we have the following bound for (5.14)

$$|c_{m,k}(y) - c^*_{m,k}(y)| \leq 2^{m/2}\|f(\cdot,y)\|_2\sqrt{T}\frac{(\lceil 2^m T\rceil\pi)^2}{2^{2(J+1)} - (\lceil 2^m T\rceil\pi)^2}\ , \quad \text{where} \quad J \geq \log_2\left(\lceil 2^m T\rceil\pi\right)\ . \tag{5.15}$$

Putting everything together, we have

$$\epsilon_c(x,y) \leq L(J,y) := 2^m\left(\lceil 2^m T\rceil + 1\right)\|f(\cdot,y)\|_2\sqrt{T}\frac{(\lceil 2^m T\rceil\pi)^2}{2^{2(J+1)} - (\lceil 2^m T\rceil\pi)^2}\ . \tag{5.16}$$

Thus,

$$E^*_{m,1} \leq \max_{(x,y)\in[0,T]\times[0,y_0]}|g(x,y)|T\left(\max_{y\in[0,y_0]}K(2^m\pi,y) + \max_{(x,y)\in[0,T]\times[0,y_0]}\epsilon_t(x,y) + \max_{y\in[0,y_0]}L(J,y)\right)\ , \tag{5.17}$$

where $K(2^m\pi,\cdot)$ and $L(J,\cdot)$ are defined in (5.10) and (5.16), respectively.

5.2 Bound for error term $E^*_{m,2}$

From (5.4), we have

$$E^*_{m,2} \leq \int_0^{y_0}\left[\sum_{k=0}^{\lceil 2^m T\rceil}\left|c^*_{m,k}(y)\right|\left|\int_0^T g(x,y)\varphi_{m,k}(x)\mathrm{d}x - \frac{1}{2^{m/2}}g\left(\frac{k}{2^m},y\right)\right|\right]w(y)\mathrm{d}y\ . \tag{5.18}$$

From (4.15), we have

$$\left|c^*_{m,k}(y)\right| \leq \frac{2^{m/2}}{2^{J-1}}\sum_{j=1}^{2^{J-1}}\int_{\mathbb{R}} f(x|y)\mathrm{d}x = 2^{m/2}\ , \tag{5.19}$$

and from Theorem 4.1,

$$\left|\int_0^T g(x,y)\varphi_{m,k}(x)\mathrm{d}x - \frac{1}{2^{m/2}}g\left(\frac{k}{2^m},y\right)\right| \leq \frac{1}{2^{m/2}}M(y)e^{-\pi d(y)2^m}\ , \tag{5.20}$$

where $d(y)$ and $M(y)$ are positive bounded constants depending on $y \in [0, y_0]$. As a result, we have the following bound for $E^*_{m,2}$

$$E^*_{m,2} \leq U(m) := y_0 \left(\lceil 2^m T \rceil + 1\right) \max_{y\in[0,y_0]} M(y) e^{-\pi d(y) 2^m} \ . \tag{5.21}$$

5.3 Bound for error term $E^*_{m,3}$

From (5.5), we have

$$E^*_{m,3} \leq \frac{1}{2^{m/2}} \int_{\mathcal{D}} \left[\sum_{k=0}^{\lceil 2^m T \rceil} \left| c^*_{m,k}(e^v) \right| \left| g\left(\frac{k}{2^m}, e^v\right) \right| \right] \bar{w}(v)\, \mathrm{d}v \ . \tag{5.22}$$

From (5.19) we have that $|c^*_{m,k}(e^v)| \leq 2^{m/2}$ and from Section 4.2 we know that $\bar{w}(v) < \epsilon_{\text{tol}}, v \in \mathcal{D}$. Thus,

$$E^*_{m,3} \leq \epsilon_{\text{tol}} \sum_{k=0}^{\lceil 2^m T \rceil} \int_{\mathcal{D}} \left| g\left(\frac{k}{2^m}, e^v\right) \right| \mathrm{d}v \ . \tag{5.23}$$

If we assume that the integrals in (5.23) are convergent and define $\bar{Y}(k,m) := \int_{\mathcal{D}} |g\left(\frac{k}{2^m}, e^v\right)| \, \mathrm{d}v$ and

$$Y(m) := \max_{k\in\{0,\ldots,\lceil 2^m T\rceil\}} \bar{Y}(k,m) \ , \tag{5.24}$$

then

$$E^*_{m,3} \leq \epsilon_{\text{tol}} (\lceil 2^m T \rceil + 1) Y(m) \ . \tag{5.25}$$

5.4 Bound for the error term $E^*_{m,h}$ and the total error E

The error of the composite trapezoidal rule in (4.29) is

$$E^*_{m,h} = \frac{(b_v - a_v)^3}{12 N_I^2} |\mathcal{S}''_m(\xi)|, \quad \xi \in (a_v, b_v) \ .$$

If $|\mathcal{S}''_m(\cdot)|$ is bounded over (a_v, b_v) by a positive constant $C(m)$, then the total error term E is bounded by

$$\begin{aligned} E \leq \|g\|_\infty T \left(\max_{y\in[0,y_0]} K(2^m\pi, y) + \|\epsilon_t\|_\infty + \max_{y\in[0,y_0]} L(J,y) \right) + U(m) + \epsilon_{\text{tol}}(\lceil 2^m T\rceil + 1) Y(m) \\ + \frac{(b_v - a_v)^3}{12 N_I^2} C(m) \ , \end{aligned} \tag{5.26}$$

where $\|g\|_\infty := \max_{(x,y)\in[0,T]\times[0,1]} |g(x,y)|$ denotes the infinite norm of g, and $K(2^m\pi, \cdot)$, $L(J, \cdot)$, $U(m)$ and $Y(m)$ are defined in (5.10), (5.16), (5.21), and (5.24), respectively.

5.5 Choice of m and J for Fourier inversion

It is observed from (4.29) and (4.30) that for each discretization point v_ℓ, a Fourier inversion needs to be performed to compute the coefficients $c^*_{m,k}(e^{v_\ell})$, $k = 0, \ldots, \lceil 2^m T \rceil$, by the formula (4.16). From (4.16), we note that the two parameters, namely the level of resolution m and the truncation parameter J, need to be determined before this inversion. In this section, we discuss how to select m and J. Once these values have been chosen, the discretization error introduced by the composite trapezoidal rule can be controlled by varying N_I.

From (5.15), we know that once an appropriate value for m has been selected, we can pick J such that $J \geq \log_2\left(\lceil 2^m T \rceil \pi\right)$, so we first discuss how to select an appropriate value for m. We proceed by finding m such that the projection error ϵ_p, defined in (5.7), is below a pre-determined tolerance `tol`. We denote by $\epsilon_p^{(m)}$ an approximation to ϵ_p, given the level of resolution m. From the bound (5.11), together with (5.10), we approximate $\epsilon_p^{(m)}$ by

$$\epsilon_p^{(m)} := \frac{1}{2\pi} \max_{v_\ell} \left(|\Psi(-2^m \pi | e^{v_\ell})| + |\Psi(2^m \pi | e^{v_\ell})| \right) \ . \tag{5.27}$$

We can find the level of resolution by iteratively computing the first m such that $\epsilon_p^{(m)} \leq \texttt{tol}$.

While we can choose a different m for each discretization point $v_\ell \in [a_v, b_v]$, the above procedure selects a common m that first satisfies (5.27) for all v_ℓ. This leads us to a more conservative estimation of the error at the cost of extra computational complexity, since the higher the level of resolution m, the more coefficients are used for the approximation at a particular discretization point v_ℓ. Nonetheless, timing results indicate that, even under this choice, the drSWIFT method is already extremely efficient.

Once the parameter m has been selected by the above-described procedure, we consider $J = \lceil \log_2\left(\lceil 2^m T \rceil \pi\right) \rceil$. However, inspection of (4.16) show that, in evaluating $c_{m,k}^*$, $k = 0, \ldots, \lceil 2^m T \rceil$, a different J can be selected for each k. For simplicity and efficiency, we prefer the above fixed value $J = \lceil \log_2\left(\lceil 2^m T \rceil \pi\right) \rceil$ for all k. If we use this value of J in (5.16), it appears that the resulting bound for $\epsilon_c(\cdot, \cdot)$ may not be very sharp. Nonetheless, we observe that, in practice, this selection of J gives us a good balance between accuracy and computational complexity. More specifically, the most computationally involved part in (4.16) is the evaluation of $\Psi(\cdot, \cdot)$ at the grid points v_ℓ, $\ell = 1, \ldots, N_I$. Those values need to be computed only once for each value of v_ℓ, and then be used by an FFT algorithm to compute the set of coefficients $c_{m,k}^*(e^{v_\ell})$, for all $k = 0, \ldots, \lceil 2^m T \rceil$. More specifically, assuming $\Psi\left(\frac{(2j+1)\pi 2^m}{2^J} \middle| e^{v_\ell} \right) = 0$, from 2^{J-1} to $2^J - 1$, we have (Ortiz-Gracia and Oosterlee, 2016)

$$\begin{aligned} c_{m,k}^*(e^{v_\ell}) &= \frac{2^{m/2}}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \Re\left[\Psi\left(\frac{(2j-1)\pi 2^m}{2^J} \middle| e^{v_\ell} \right) e^{\frac{ik\pi(2j-1)}{2^J}} \right] \\ &= \frac{2^{m/2}}{2^{J-1}} \Re\left[e^{\frac{ik\pi}{2^J}} \sum_{j=1}^{2^{J-1}} \Psi\left(\frac{(2j+1)\pi 2^m}{2^J} \middle| e^{v_\ell} \right) e^{\frac{2ijk\pi}{2^J}} \right] \ , \end{aligned} \tag{5.28}$$

and hence the FFT algorithm can be applied to compute $c_{m,k}^*(e^{v_\ell})$. An algorithm to approximate $V(S(0), 0, \cdot)$ using the drSWIFT method is given in Algorithm 5.1. We study the overall computational complexity of the algorithm in the next subsection.

5.6 Computational complexity

Examination of (4.29) reveals that a total of $(N_I + 1)$ terms $\mathcal{S}_m(v_\ell)$, $\ell = 0, \ldots, N_I$, need to be evaluated for the computation of $V_4(S(0), 0, \cdot)$. Further examination of (4.30) reveals the following complexity for evaluating each of these $\mathcal{S}_m(v_\ell)$ terms.

- For a given v_ℓ, all the coefficients $c_{m,k}^*(e^{v_\ell})$, $k = 0, \ldots, \lceil 2^m T \rceil$, need to be computed via (5.28) using the FFT algorithm. So the complexity of this step is $\mathcal{O}(N_J \log(N_J))$, where $N_J = 2^J - 1$ is the number of terms required to compute each coefficient $c_{m,k}^*(e^{v_\ell})$.

Algorithm 5.1 Algorithm to approximate $V(S(0),0,\cdot)$

1: compute matrix A using a Cholesky decomposition;
2: compute $\beta_{d_i}(t)$, $i=1,\ldots,n$, and $\beta_{f_i}(t)$, $i=1,\ldots,l$, using (3.4);
3: compute the deterministic terms of (3.3), namely
$\int_0^T \gamma_d(t)\,\mathrm{d}t$; $\int_0^T \gamma_f(t)\,\mathrm{d}t$; $\int_0^T \beta_{d_i}(t)\beta_{f_j}(t)\mathrm{d}t$, $i=1,\ldots,n$, $j=1,\ldots,l$;
$\int_0^T \beta_{d_i}(t)\beta_{d_i'}(t)\,\mathrm{d}t$, $i,i'=1,\ldots,n$; $\int_0^T \beta_{f_i}(t)\beta_{f_i'}(t)\,\mathrm{d}t$, $i,i'=1,\ldots,l$;
4: compute the interval $[a_v,b_v]$ as explained in Section 4.2;
5: compute the first m such that $\epsilon_p^{(m)} \le \texttt{tol}$ by iteratively using (5.27);
6: set $J=\lceil \log_2(\lceil 2^m T\rceil \pi)\rceil$;
7: for each v_ℓ compute coefficients $c^*_{m,k}(e^{v_\ell})$, $k=0,\ldots,\lceil 2^m T\rceil$, by FFT using (5.28), where $\Psi(\cdot|\cdot)$ defined in (4.4);
8: compute $V_4(S(0),0,\cdot)$ using (4.29);
9: return $V(S(0),0,\cdot)\approx V_4(S(0),0,\cdot)$;

- Given the computed $c^*_{m,k}(e^{v_\ell})$, $k=0,\ldots,\lceil 2^m T\rceil$, each term $\mathcal{S}_m(v_\ell)$ in (4.30) can be computed with $\mathcal{O}(N_J)$ complexity.

As a result, the total complexity of the drSWIFT method is $\mathcal{O}(N_I N_J \log(N_J))$.

We note that this is an upper bound of the computational complexity, since as explained in Section 5.5 we can select a smaller value of the scale m for each $v_\ell \in [a_v,b_v]$. It is worth underlining that the computational complexity remains the same regardless of the number of factors in the underlying model.

6 Numerical experiments

In this section, we present selected numerical results to illustrate the performance of the drSWIFT method. For verification purposes, we will start with the well-known two-factor Heston (Heston, 1993) and Bates, i.e. jump-extended Heston, Bates (1996)) models, for which a semi closed-form or an analytical solution does exist for a European option. We then consider the jump-extended version of the popular three-factor Heston-Hull-White (HHW) model, and finally, a 6-factor pure- and jump-diffusion FX model, under all of which, an analytical solution does not exist for a European option. In these examples, the correlation between the underlying asset and its instantaneous variance is non-zero, and, where relevant, the interest rate factor(s) and the underlying asset, as well as the instantaneous variance, are pairwise independent.

For all the experiments, in determining the integration interval $[a_v,b_v]$, we consider $\epsilon_{\text{tol}}=10^{-6}$, and follow the procedure explained in Section 4.2, where a Newton search is used when the Feller condition is not satisfied, and the alternative method otherwise. While in the first three models considered, namely Heston, Bates and jump-extended HHW, the Feller condition is satisfied for the variance process, in the 6-factor FX model, we also experiment with a variance process in which the Feller condition is not satisfied to illustrate the benefit of the log-variance transformation discussed in Section 4.2. The programs were coded in MATLAB, and run on a Surface Pro 3 with Intel Dual Core i7-4650U @ 1.70GHz 2.30GHz processor and 8GB RAM.

6.1 Heston model

For the Heston model (Heston, 1993), the function $g(x,y)$ is defined as

$$g(x,y) = S(0)\mathrm{e}^{(G(x)+F(x,y)+H)}\mathcal{N}\left(d_{1,0}(x,y)\right) - K\mathrm{e}^{H}\mathcal{N}\left(d_{2,0}(x,y)\right) \ , \tag{6.1}$$

where

$$\begin{aligned} G(x) &= \frac{a_{1,1}^2}{2}x \ , \quad \text{with} \quad a_{1,1} = \sqrt{1-\rho_{s,\nu}^2} \ , \\ F(x,y) &= -\frac{1}{2}x + r_d(0)T + a_{1,2}\left(\frac{y-(\nu(0)+\kappa_\nu\bar{\nu}T-\kappa_\nu x)}{\sigma_\nu}\right) \quad \text{with} \quad a_{1,2} = \rho_{s,\nu} \ , \\ H &= -r_d(0)T \ . \end{aligned} \tag{6.2}$$

where $d_{1,0}(x,y)$ and $d_{2,0}(x,y)$ are defined in (3.2). Here, we use $(\cdot,\cdot)$ to clearly indicate the dependence of the quantities under discussion on the parameters x and/or y.

In Table 6.1, we present computed prices of a European call option under the Heston dynamics for different maturities T. The payoff in this case is $\Phi(S(T)) = \max(S(T)-K,0)$, with K being the strike. In this test, for each maturity, we also consider different levels of resolution m, namely $m = \{6,7,8\}$ and different number of subintervals N_I for the composite trapezoidal rule, namely $N_I = \{15,25,50\}$. For each value of m, we also report the corresponding error $\epsilon_p^{(m)}$, defined in (5.27). (Note that $\epsilon_p^{(m)}$ is independent of N_I.) Finally, for each parameter combination, we also report the absolute error ("abs. error") between the computed price and the exact price obtained via formulas in Gatheral (2006).

N_I	m	$T=0.2$			$T=1$			$T=5$		
		$\epsilon_p^{(m)}$	abs. error	time (sec.)	$\epsilon_p^{(m)}$	abs. error	time (sec.)	$\epsilon_p^{(m)}$	abs. error	time (sec.)
15	6	2.89e-01	5.26e-01	0.03	2.84e-03	1.66e-03	0.03	6.04e-13	1.00e-05	0.04
	7	2.17e-01	6.54e-02	0.03	4.16e-06	1.03e-06	0.04	5.50e-24	9.53e-06	0.05
	8	7.01e-02	1.53e-04	0.03	1.22e-10	1.24e-06	0.04	6.79e-41	9.52e-06	0.09
25	6		5.25e-01	0.04		1.66e-03	0.04		1.00e-05	0.05
	7		6.53e-02	0.04		3.67e-06	0.04		9.59e-06	0.07
	8		4.62e-05	0.04		3.87e-06	0.05		9.59e-06	0.10
50	6		5.25e-01	0.04		1.66e-03	0.05		8.40e-06	0.05
	7		6.52e-02	0.04		2.32e-06	0.05		7.94e-06	0.07
	8		3.99e-05	0.04		2.52e-06	0.05		7.95e-06	0.15

TABLE 6.1: *European call option under Heston dynamics with parameters:* $S(0)=100$, $K=100$, $r_d(0)=0.15$, $\rho_{s,\nu}=0.4$, $\nu(0)=0.2$, $\kappa_\nu=3$, $\bar{\nu}=0.09$, $\sigma_\nu=0.3$. *The Feller's condition is satisfied for the variance process. Reference values are obtained via Gatheral (2006):* 8.831873326617753 *for* $T=0.2$, 20.967685183036807 *for* $T=1$, *and* 55.881189957646598 *for* $T=5$.

We make the following observations.

- First, for the case $N_I = 15$, we observe that when $T=0.2$, the absolute error decreases when the level of resolution m increases (e.g. 5.26e-01 when $m=6$ versus 1.53e-04 when $m=8$); however, when $T=5$, the absolute error is approximately the same for all three levels of resolution m, (e.g. 1.00e-05 when $m=6$ versus 9.52e-06 when $m=8$), and the approximation is already significantly accurate with the smallest $m=6$.

- Next, across different values of N_I, we observe that, for a given m, an increase in N_I does not seem to improve the accuracy, and this appears to hold true for all maturities T. For example, for $m = 6$, with $T = 0.2$, the absolute errors are 5.26e-01 and 5.25e-01 for $N_I = 15$ and $N_I = 50$, respectively; with $T = 5$, the respective absolute errors are 1.00e-05 and 8.40e-06, which are almost the same.

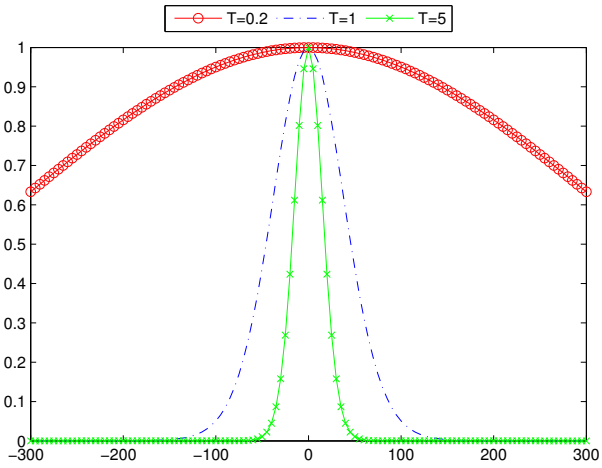

FIGURE 6.1: *Modulus of the characteristic function of the conditional time-integrated variance process* $\Psi(\cdot|\cdot)$ *for different maturities and parameters* $\nu(0) = 0.2, \kappa_\nu = 3, \bar{\nu} = 0.09, \sigma_\nu = 0.3$. *The terminal value of the log-variance in this case is* $v = \ln(\mathbb{E}(\nu(T)))$.

To investigate this further, in Figure 6.1, we plot the modulus of the characteristic function of the conditional time-integrated variance process $\Psi(\cdot|e^v)$, when $v = \ln(\mathbb{E}[\nu(T)])$, for the three maturities considered in this example. From this plot, taking into account the computed $\epsilon_p^{(m)}$ values in Table 6.1, we conclude that the bound of the total error in the method, given in (5.26), is dominated by $K(2^m\pi, y)$, defined in (5.10), which essentially measures the mass in the tails of the modulus of Ψ.

In view of these insights, in the remaining examples, we will consider $N_I = 15$ and the $\texttt{tol} = 10^{-2}$ in estimating the level resolution m, i.e. find the first level of resolution m such that for $\epsilon_p^{(m)} \leq \texttt{tol}$, as discussed in Subsection 5.5. We emphasize that with this choice of m and $N_I = 15$, the price under the Heston model is obtained in less than 0.05 seconds.

6.2 Bates model

Next, we consider the Bates model in Bates (1996), where log of the jump amplitude follows a normal distribution with mean $\tilde{\mu}$ and variance $\tilde{\sigma}^2$. For this model, the function $g(x, y)$ is

$$g(x,y) = \sum_{j=0}^{\infty} \frac{(\lambda T)^j}{j!} \left\{ \exp\left(j\tilde{\mu} + \frac{j\tilde{\sigma}^2}{2}\right) S(0) \mathrm{e}^{(G(x)+F(x,y)+H)} \mathcal{N}\left(d_{1,j}(x,y)\right) - K\mathrm{e}^H \mathcal{N}\left(d_{2,j}(x,y)\right) \right\}, \tag{6.3}$$

where

$$G(x) = \frac{a_{1,1}^2}{2}x \ , \quad \text{with} \quad a_{1,1} = \sqrt{1-\rho_{s,\nu}^2} \ ,$$
$$F(x,y) = -\frac{1}{2}x + r_d(0)T + a_{1,2}a_{1,2}\left(\frac{y-(\nu(0)+\kappa_\nu\bar{\nu}T-\kappa_\nu x)}{\sigma_\nu}\right) - \lambda\delta T \ ,$$
$$\text{with } a_{1,2} = \rho_{s,\nu}, \delta = e^{\tilde{\mu}+\frac{1}{2}\tilde{\sigma}^2} - 1 \ ,$$
$$H = -(r_d(0)+\lambda)T \ .$$

In this test, the parameters for the model are $T = 1$, $S(0) = 80$, $r_d(0) = 0.15$, $\rho_{s,\nu} = -0.5$, $\nu(0) = 0.04$, $\kappa_\nu = 3$, $\bar{\nu} = 0.09$, $\sigma_\nu = 0.3$, $\lambda = 1$, $\tilde{\mu} = -0.08$, $\tilde{\sigma} = 0.3$.

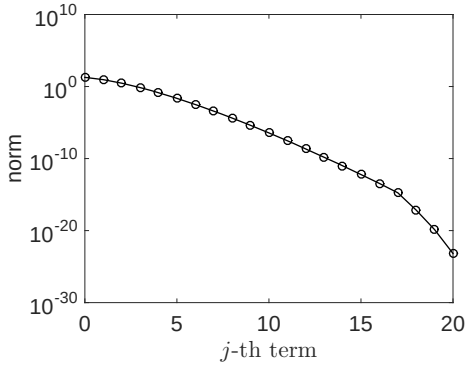

FIGURE 6.2: *Norm of the j-th term, $j = 0, 1, \ldots$, in the infinite series* (6.3).

In the implementation of the infinite series (6.3), we need to determine the number of terms to keep. In Figure 6.2, we plot in log-scale the norm of the j-th term, $j = 0, 1, \ldots$, in the infinite series (6.3). As shown in this figure, the infinite series (6.3) converges very quickly, and we choose to keep the first 20 terms of (6.3) in the implementation, for which the truncation is already much less than 10^{-10}.

In determining the level of resolution m, we find the first m such that $\epsilon_p^{(m)} < \texttt{tol} = 10^{-2}$. For the above set of parameters, the computed level of resolution is $m = 7$. In Table 6.2, we present

K	reference	abs. error	rel. error (%)
66.2563	26.1843	2.79e-03	0.01
70.5529	23.4604	2.83e-03	0.01
75.1281	20.7564	2.69e-03	0.01
80.0000	18.1113	2.50e-03	0.01
85.1878	15.5675	1.94e-03	0.01
90.7121	13.1693	1.37e-03	0.01
96.5945	10.9581	5.13e-04	< 0.01

TABLE 6.2: *European call under the Bates model with parameters: $T = 1$, $S(0) = 80$, $r_d(0) = 0.15$, $\rho_{s,\nu} = -0.5$, $\nu(0) = 0.04$, $\kappa_\nu = 3$, $\bar{\nu} = 0.09$, $\sigma_\nu = 0.3$, $\lambda = 1$, $\tilde{\mu} = -0.08$, $\tilde{\sigma} = 0.3$. The Feller's condition is satisfied for the variance process. Other parameters are $m = 7$ and $N_I = 15$. All the results are obtained in less than 0.05 seconds.*

selected results for a European call option for different strikes. The reference prices ("reference") are those obtained by the exact formula in Bates (1996). The absolute and relative errors, "abs.

error" and "rel. error", respectively, are computed based on these reference prices. As observed from Table 6.2, all the option prices computed by the drSWIFT method are highly accurate. The efficiency of the method is impressive. The method is able to achieve, for the option price, a relative error of about 0.01% in less than 0.05 seconds.

6.3 Jump-extended Heston-Hull-White model

In the third example considered in this work, we focus on the jump-extended Heston-Hull-White (HHW) model, where log of the jump amplitude follows a normal distribution with mean $\tilde{\mu}$ and variance $\tilde{\sigma}^2$. While the diffusion version of this model is considered in a number of works, such as Grzelak and Oosterlee (2012b); Haentjens and in 't Hout (2012), pricing methods for European options under this model have not been discussed in the literature. In this case, the model is

$$\begin{aligned}
\frac{\mathrm{d}S(t)}{S(t)} &= r_d(t)\,\mathrm{d}t + \sqrt{\nu(t)}\,\mathrm{d}W_s(t) + \mathrm{d}J(t)\ , \\
r_d(t) &= r_d(0)\,\mathrm{e}^{-\kappa_d t} + \kappa_d \int_0^t \mathrm{e}^{-\kappa_d(t-t')}\,\theta_d(t')\,\mathrm{d}t' + X(t)\ , \\
&\qquad \text{with } \mathrm{d}X(t) = -\kappa_d\,X(t)\,\mathrm{d}t + \sigma_d\,\mathrm{d}W_d(t)\ , \quad X(0) = 0, \\
\mathrm{d}\nu(t) &= \kappa_\nu\left(\bar{\nu} - \nu(t)\right)\,\mathrm{d}t + \sigma_\nu\sqrt{\nu(t)}\,\mathrm{d}W_\nu(t)\ ,
\end{aligned}$$

where κ_d, σ_d, κ_ν, σ_ν and $\bar{\nu}$ are constants. In this example, the $g(x,y)$ is defined as in (6.3), where

$$\begin{aligned}
G(x) &= \frac{a_{1,1}^2}{2}x + \frac{1}{2}a_{2,2}^2\int_0^T (\beta_d(t))^2\,\mathrm{d}t\ , \quad \text{with} \quad a_{1,1} = \sqrt{1-\rho_{s,\nu}^2},\ \text{and } a_{2,2} = 1\ , \\
F(x,y) &= -\frac{1}{2}x + \int_0^T \gamma_d(t)\,\mathrm{d}t + a_{1,3}\left(\frac{y - (\nu(0) + \kappa_\nu\bar{\nu}T - \kappa_\nu x)}{\sigma_\nu}\right) - a_{2,2}^2\int_0^T (\beta_d(t))^2\,\mathrm{d}t - \lambda\delta T\ , \\
&\qquad \text{with } a_{1,3} = \rho_{s,\nu}\ , \delta = e^{\tilde{\mu}+\frac{1}{2}\tilde{\sigma}^2} - 1\ , \\
H &= -\int_0^T \gamma_d(t)\,\mathrm{d}t + \frac{1}{2}a_{2,2}^2\int_0^T (\beta_d(t))^2\,\mathrm{d}t - \lambda T\ ,
\end{aligned}$$

and

$$\begin{aligned}
\int_0^T (\beta_d(t))^2\,\mathrm{d}t &= \left(\frac{\sigma_d}{\kappa_d}\right)^2\left[T + \frac{1-e^{-2\kappa_d T}}{2\kappa_d} - \frac{2\left(1-e^{-2\kappa_d T}\right)}{\kappa_d}\right]\ , \\
\int_0^T \gamma_d(t)\,\mathrm{d}t &= \theta_d T + \frac{r_d(0)-\theta_d}{\kappa_d}\left(1-e^{-\kappa_d T}\right)\ .
\end{aligned}$$

In the implementation of the drSWIFT method, after carrying out the same test as in the Bates example, we choose to keep only the first 20 terms of the series $g(x,y)$.

With this setting, we price a European call option with different maturities. In these tests, similar to previous tests, the level of resolution is the first m such that $\epsilon_p^{(m)} < \texttt{tol} = 10^{-2}$. To compute benchmark solutions, we use the multi-level MC method presented in Dang (2017), where the multi-level MC technique is applied only to the variance factor. To simulate $\nu(t)$, we use the Lamperti-Backward-Euler timestepping method that preserves the positivity of the original dynamics (2.1d), and has a good strong convergence property, recently established in Neuenkirch and Szpruch (2014). In the experiment with multi-level MC, the root-mean-square error is 10^{-3}.

In Table 6.3 we present selected results. The standard deviations in the benchmark option prices all are $\leq \frac{10^{-3}}{\sqrt{2}} \approx 0.000707$, as expected from analysis of multi-level MC methods (Giles, 2008). We note that prices computed by the drSWIFT lie within the 95% confidence intervals

T	multi-level MC		drSWIFT				
(years)	(price, std. dev.)	95% CI	m	price	abs. error	rel. error (%)	time (sec.)
0.5	(14.8127, 0.0007)	[14.8113 14.8140]	7	14.8129	2e-4	< 0.01	< 0.05
1	(21.3952, 0.0007)	[21.3939, 21.3966]	7	21.3948	4e-4	< 0.01	0.05
1.5	(26.7991, 0.0007)	[26.7979, 26.8001]	6	26.7987	4e-4	< 0.01	0.05

TABLE 6.3: *European call prices under jump-extended HHW dynamics with parameters:* $S(0) = 100$, $K = 100$, $\nu(0) = 0.2$, $\kappa_\nu = 3$, $\bar{\nu} = 0.1$, $\sigma_\nu = 0.3$, $r_d(0) = 0.05$, $\kappa_d = 1.5$, $\theta_d = 0.1$, $\sigma_d = 0.1$, $\lambda = 1$, $\tilde{\mu} = -0.08$, $\tilde{\sigma} = 0.3$. *The correlations are* $\rho_{s,\nu} = -0.3$, $\rho_{s,d} = \rho_{d,\nu} = 0$.

obtained by the multi-level MC method. Moreover, they are in excellent agreement with the benchmark prices. Finally, we note again the impressive efficiency of the drSWIFT method, being able to achieve, for the option price, a relative error of about 0.01% in 0.05 seconds.

6.4 A 6-factor foreign exchange model

Finally, we consider the valuation of a European option under a 6-factor FX model. We consider both the pure-diffusion and jump-diffusion versions of the model, for which $g(x, y)$ are respectively defined in (6.1) and (6.3). The functions $G(\cdot)$, $F(\cdot, \cdot)$ and H for the jump-diffusion case are given by

$$
\begin{aligned}
G(x) &= \frac{a_{1,1}^2}{2}x + \frac{1}{2}\sum_{k=2}^{5}\sum_{j=1}^{2}\sum_{j'=1}^{2} a_{j+1,k}a_{j'+1,k}\int_0^T \beta_{d_j}(t)\beta_{d_{j'}}(t)\,\mathrm{d}t \\
&\quad + \frac{1}{2}\sum_{k=2}^{5}\sum_{j=1}^{2}\sum_{j'=1}^{2} a_{j+3,k}a_{j'+3,k}\int_0^T \beta_{f_j}(t)\beta_{f_{j'}}(t)\,\mathrm{d}t + \frac{1}{2}\sum_{k=2}^{5}\sum_{j=1}^{2}\sum_{j'=1}^{2} a_{j+1,k}a_{j'+3,k}\int_0^T \beta_{d_j}(t)\beta_{f_{j'}}(t)\,\mathrm{d}t \\
F(x,y) &= -\frac{1}{2}x + \int_0^T (\gamma_d(t) - \gamma_f(t))\,\mathrm{d}t - \sum_{k=2}^{5}\sum_{j=1}^{2}\sum_{j'=1}^{2} a_{j+1,k}a_{j'+1,k}\int_0^T \beta_{d_j}(t)\beta_{d_{j'}}(t)\,\mathrm{d}t \\
&\quad + \sum_{k=2}^{5}\sum_{j=1}^{2}\sum_{j'=1}^{2} a_{j+1,k}a_{j'+3,k}\int_0^T \beta_{d_j}(t)\beta_{f_{j'}}(t)\,\mathrm{d}t + a_{1,6}\left(\frac{y - (\nu(0) + \kappa_\nu\bar{\nu}T - \kappa_\nu x)}{\sigma_\nu}\right) - \lambda\delta T \\
H &= -\int_0^T \gamma_d(t)\,\mathrm{d}t + \frac{1}{2}\sum_{k=2}^{5}\sum_{j=1}^{2}\sum_{j'=1}^{2} a_{j+1,k}a_{j'+1,k}\int_0^T \beta_{d_j}(t)\beta_{d_{j'}}(t)\,\mathrm{d}t - \lambda T\ .
\end{aligned}
$$

Here,

$$
\int_0^T \gamma_d(t)\,\mathrm{d}t = \theta_d T + \frac{r_d(0) - \theta_d}{\kappa_{d_1}}\cdot\left(1 - e^{-\kappa_{d_1}T}\right)\ ,\quad \int_0^T \gamma_f(t)\,\mathrm{d}t = \theta_f T + \frac{r_f(0) - \theta_f}{\kappa_{f_1}}\cdot\left(1 - e^{-\kappa_{f_1}T}\right)\ ,
$$

$$
\int_0^T \beta_i(t)\beta_j(t)\,\mathrm{d}t = \frac{\sigma_i\sigma_j}{\kappa_i\kappa_j}\cdot\left[T - \frac{1 - e^{-\kappa_i T}}{\kappa_i} - \frac{\left(1 - e^{-\kappa_j T}\right)}{\kappa_j} + \frac{\left(1 - e^{-(\kappa_i+\kappa_j)T}\right)}{\kappa_i + \kappa_j}\right]\ ,
$$

where $i, j \in \{d_1, d_2, f_1, f_2\}$, and κ_{d_i}, κ_{f_i}, σ_{d_i}, σ_{f_i}, $i = 1, 2$, κ_ν, σ_ν and $\bar{\nu}$ are constants. For the pure-diffusion case, the respective functions $G(\cdot)$ and $F(\cdot, \cdot)$ and H can be obtained by setting the jump intensity $\lambda = 0$. For the jump-diffusion case, after carrying out the same test as in the Bates example, we also choose to keep only the first 20 terms of the series $g(x, y)$.

To perform the numerical experiments, we consider two different sets of parameters for the variance.

- Set 1: $\nu(0) = 0.2$, $\kappa_\nu = 2.5$, $\bar{\nu} = 0.6$, $\sigma_\nu = 0.5$ for which Feller's condition is satisfied

- Set 2: $\nu(0) = 0.2$, $\kappa_\nu = 0.1$, $\bar{\nu} = 0.6$, $\sigma_\nu = 0.5$ for which Feller's condition is not satisfied.

For the maturity, we choose $T = 5$ (years).

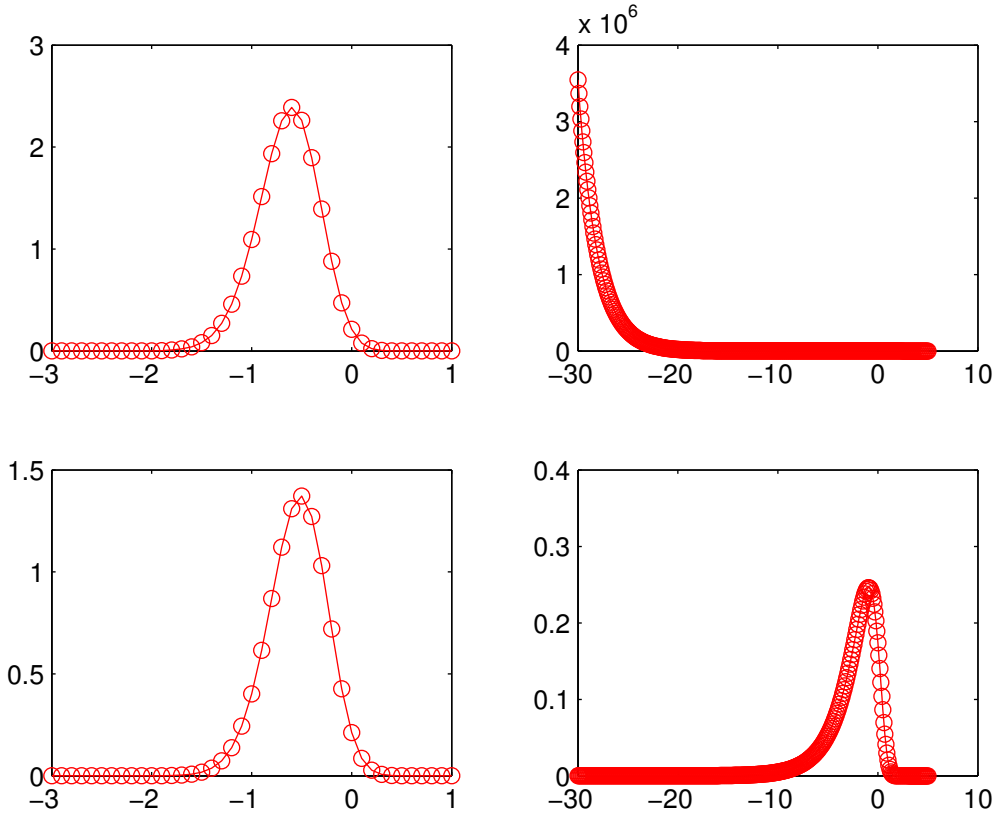

FIGURE 6.3: *Density plots. Top-left: $\tilde{w}(\cdot)$ when the Feller's condition is satisfied (Set 1). Top-right: $\tilde{w}(\cdot)$ when the Feller's condition is unsatisfied (Set 2). Bottom-left: $\bar{w}(\cdot)$ when the Feller's condition is satisfied (Set 1). Bottom-right: $\bar{w}(\cdot)$ when the Feller's condition is unsatisfied (Set 2).*

To illustrate the benefit of changing from the variance to the log-variance, as discussed in Section 4.2, in Figure 6.3, we plot the densities of the terminal variance value $\tilde{w}(\cdot)$ and of terminal log-variance value $\bar{w}(\cdot)$, defined in (4.23), for both set of the variance parameters. As clearly shown in Figure 6.3's top- and bottom-left panels, when the Feller's condition is satisfied, both $\tilde{w}(\cdot)$ and $\bar{w}(\cdot)$ present a similar shape, with tails decaying very quickly. However, when the Feller's condition is not satisfied, we observe a very heavy left tail in $\tilde{w}(\cdot)$, see top-right panel, but not in $\bar{w}(\cdot)$, see the bottom-right panel.

To further investigate the decay of the left tails, in Figure 6.4, we plot in log-scale $\tilde{w}(\cdot)$ and $\bar{w}(\cdot)$. It is clearly from this plot that the decay of $\bar{w}(\cdot)$'s left tail is very fast. As such, we clearly benefit from the use of $\bar{w}(\cdot)$ when we apply the composite trapezoidal rule in (4.29) to get the final approximation for the option value.

In Table 6.4 we present selected pricing results of a European put option. In this test, the benchmark solutions are again obtained by the multi-level MC in Dang (2017) as described in the previous experiment. As noted earlier, the standard deviations in the benchmark option prices all are $\leq \frac{10^{-3}}{\sqrt{2}} \approx 0.000707$, as expected. For the drSWIFT, the level of resolution m is chosen with the error $\epsilon_p^{(m)} = 10^{-3}$. We observe from Table 6.4 that all prices computed by the drSWIFT

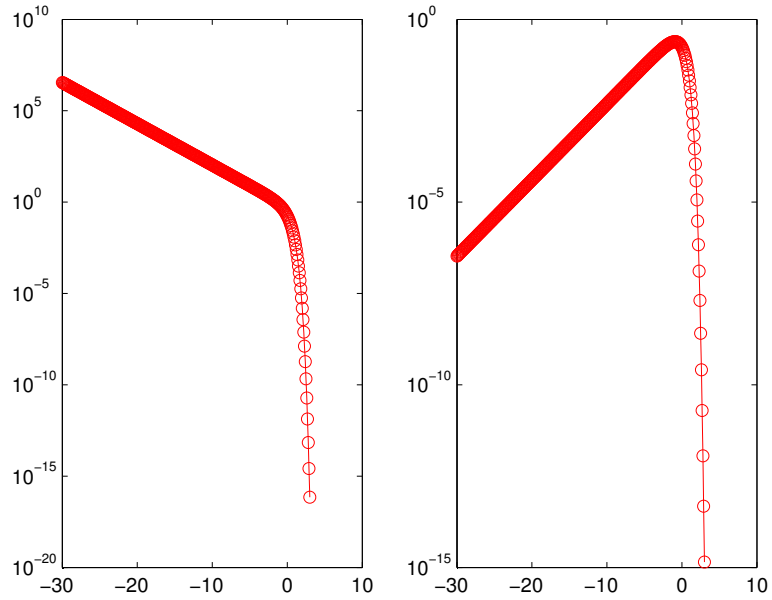

FIGURE 6.4: *Density plots in log-scale when Feller condition is not satisfied (Set 2). Left:* $\tilde{w}(\cdot)$. *Right:*$\bar{w}(\cdot)$.

	variance	multi-level MC		drSWIFT				
	param.	(price, std. dev.)	95% CI	m	price	abs. error	rel. error (%)	time (sec.)
pure	1	(7.1934, 0.0007)	[7.1921, 7.1947]	3	7.1928	5.5e-04	< 0.01	< 0.05
diffusion	2	(5.5730, 0.0007)	[5.5716, 5.5743]	5	5.5724	6.0e-04	0.01	0.05
jump	1	(7.4835, 0.0007)	[7.4821, 7.4847]		7.4839	4.5e-04	< 0.01	0.05
diffusion	2	(6.1230, 0.0007)	[6.1216, 6.1243]		6.1238	8.0e-04	0.01	0.06

TABLE 6.4: *European put prices under the 6-factor FX model dynamics with parameters:* $S(0) = 10$, $K = 10$, $T = 5$, $\kappa_{d_1} = 0.97$, $\kappa_{d_2} = 0.24$, $\sigma_{d_1} = 0.20$, $\sigma_{d_2} = 0.16$, $r_d(0) = 0.02$, $\theta_d = 0.02$, $\kappa_{f_1} = 0.77$, $\kappa_{f_2} = 0.08$, $\sigma_{f_1} = 0.02$, $\sigma_{f_2} = 0.012$, $r_f(0) = 0.05$, $\theta_f = 0.05$, $\lambda = 1$, $\tilde{\mu} = -0.08$, $\tilde{\sigma} = 0.3$. *The correlations between the asset and the interest rate factors, as well as those between the interest rate factor and the variance are zero. The other correlations are* $\rho_{s,\nu} = -0.2$, $\rho_{d_1,d_2} = -0.590$, $\rho_{d_1,f_1} = 0.125$, $\rho_{d_1,f_2} = 0.125$, $\rho_{d_2,f_1} = 0.125$, $\rho_{d_2,f_2} = 0.125$, $\rho_{f_1,f_2} = -0.702$.

lie within the 95% confidence intervals obtained by the multi-level MC method. Moreover, they are in excellent agreement with the benchmark prices, regardless of whether or not the Feller condition is satisfied. Finally, we note the impressive efficiency of the drSWIFT method.

7 Summary and future work

In addition to being useful for risk-management purposes, jump-diffusion models with square-root stochastic variance and multi-factor Gaussian interest rates can provide realistic dynamics for the underlying. Nonetheless, the first hurdle in using such these models is calibration, which typically requires a very efficient pricing method for European options. A direct application of existing state-of-the-art numerical integration technique to these models appears impossible, since a closed-form expression for the characteristic function of the underlying process under these models is not known

In this paper, we show that under the dimension reduction framework put forward in Dang et al. (2015b, 2017), it is possible to extend the applicabilities of existing state-of-the-art numerical integration methods to this broad class of models. We focus on the Shannon wavelet method of Ortiz-Gracia and Oosterlee (2016), due to its established robustness. Within this dimension reduction framework, the option price and hedging parameters can be expressed as a two-dimensional integral that involves only the densities of (i) the value of the variance at the terminal time, and (ii) the time-integrated variance process conditional on this value. We develop a highly efficient Shannon wavelet inverse Fourier technique to recover the density of the conditional time-integrated variance process from its known conditional characteristic function. Furthermore, excellent approximation properties of Shannon wavelets allow to reduce the overall pricing procedure to the evaluation of just a single integral that involves only the density of the terminal variance value. This single integral can be accurately evaluated, since the density of the variance at the terminal time is known in closed-form. We develop sharp approximation error bounds for the option price and hedging parameters.

We present a number of examples to validate the method and to illustrate its robustness. Numerical results validate the methods and its impressive efficiency. In about 0.05 seconds on a personal computer, the method can compute the price of a European option under a 6-factor jump-diffusion model within 0.01% relative error of a benchmark solution obtained via a multi-level Monte Carlo method (Dang, 2017). In addition, the complexity of the method is independent of the number of factors in the model. These advantages of the method make it particularly suitable for calibration of high-dimensional models.

Future work includes extensions of the method to pricing early exercise options. Within the dimension reduction framework, the key challenge in tackling the early exercise feature is the development of efficient computation of the solution of (i) the conditional PIDE and (ii) the conditional continuation value. Preliminary results indicate that the developed Shannon method can be modified to effectively handle this challenge. Another research direction is to extend the method to handle interest rates following multi-factor square-root CIR dynamics (Cox et al., 1985a). It turns out that the developed Shannon methods can also be effectively employed for this purpose.

References

Ahlip, R. and M. Rutkowski (2013). Pricing of foreign exchange options under the Heston stochastic volatility model and CIR interest rates. *Quantitative Finance* 13, 955–966.

Alizadeh, S., M. Brandt, and F. Diebold (2002). Range-based estimation of stochastic volatility models. *Journal of Finance* 57, 1047–1091.

Andersen, L. and V. Piterbarg (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics* 11, 29–50.

Andersen, T. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239–1284.

Bakshi, G., C. Cao, and C. Zhiwu (1997). Empirical performance of alternativee option pricing models. *Journal of Finance* 52, 2003–2049.

Bates, D. (1996). Jumps and stochastic volatility: Exchange rate process implicit in Deutsche Mark options. *Review of Financial Studies* pp. 69–107.

Brigo, D. and F. Mercurio (2006). *Interest Rate Models - Theory and Practice*. Springer, second edition.

Broadie, M. and O. Kaya (2006). Exact simulation of stochastic volatility and other affine jump difusion processes. *Operations Research* 54, 217–231.

Cattani, C. (2008). Shannon wavelets theory. *Mathematical Problems in Engineering* 2008, article ID 164808.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. Chapman and Hall.

Cox, J., J. Ingersoll, and S. Ross (1985a). A theory of the term structure of interest rates. *Econometrica* 53, 385–407.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985b). A theory of the term structure of interest rates. *Econometrica* 53, 385–407.

Cozma, A. and C. Reisinger (2016). A mixed Monte Carlo and PDE variance reduction method for foreign exchange options under the Heston-CIR model. *Journal of Computational Finance* To appear.

Dang, D. M. (2017). A multi-level dimension reduction Monte-Carlo method for jump-diffusion models. *Journal of Computational and Applied Mathematics* To appear.

Dang, D. M., C. Christara, and K. Jackson (2014). GPU pricing of exotic cross-currency interest rate derivatives with a foreign exchange volatility skew model. *Journal of Concurrency and Computation: Practice and Experience* 26, 1609–1625.

Dang, D. M., C. Christara, K. Jackson, and A. Lakhany (2015a). An efficient numerical PDE approach for pricing foreign exchange interest rate hybrid derivatives. *Journal of Computational Finance* 18, 1–55.

Dang, D. M., K. R. Jackson, and M. Mohammadi (2015b). Dimension and variance reduction for Monte-Carlo methods for high-dimensional models in finance. *Applied Mathematical Finance* 22(6), 522–552.

Dang, D. M., K. R. Jackson, and S. Sues (2017). A dimension and variance reduction Monte-Carlo method for option pricing under jump-diffusion models. *Applied Mathematical Finance* To appear.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* 300, 463–520.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.

Fang, F. and C. W. Oosterlee (2008). A novel pricing method for European options based on Fourier-Cosine series expansions. *SIAM Journal on Scientific Computing* 31, 826–848.

691 Fang, F. and C. W. Oosterlee (2011). A Fourier-based valuation method for Bermudan and barrier
692 options under Heston's model. *SIAM Journal on Financial Mathematics* 2, 439–463.

693 Gatheral, J. (2006). *The volatility surface: A practitioner's guide*. Wiley Finance, New York.

694 Gearhart, W. B. and H. S. Shultz (1990). The function $\sin(x)/x$. *The College Mathematics*
695 *Journal* 21, 90–99.

696 Giles, M. B. (2008). Multi-level Monte Carlo path simulation. *Operations Research* 56, 607–617.

697 Grzelak, L. A. and C. W. Oosterlee (2011). The affine Heston model with correlated Gaussian
698 interest rates for pricing hybrid derivatives. *Quantitative Finance* 11, 1647–1663.

699 Grzelak, L. A. and C. W. Oosterlee (2012a). On cross-currency models with stochastic volatility
700 and correlated interest rates. *Applied Mathematical Finance* 19, 1–35.

701 Grzelak, L. A. and C. W. Oosterlee (2012b). On the Heston model with stochastic interest rates.
702 *SIAM Journal of Fianancial Mathematics* 2, 255–286.

703 Haastrecht, A. V., R. Lord, A. Pelsseri, and D. Schrager (2009). Generic pricing of FX, inflation
704 and stock options under stochastic interest rates and stochastic. *Insurance: Math. Econ.* 45,
705 436–448.

706 Haastrecht, A. V. and A. Pelsser (2011). Generic pricing of FX, inflation and stock options under
707 stochastic interest rates and stochastic volatility. *Quantitative Finance* 11, 665–691.

708 Haentjens, T. and K. J. in 't Hout (2012). Alternating direction implicit finite difference schemes
709 for the Heston-Hull-White partial differential equation. *Journal of Computational Finance*
710 16(1), 83–110.

711 Heston, S. (1993). A closed form solution for options with stochastic volatility with applications
712 to bond and currency options. *Review of Financial Studies* 6, 327–343.

713 Hull, J. and A. White (1993). One factor interest rate models and the valuation of interest rate
714 derivative securities. *Journal of Financial and Quantitative Analysis* 28(2), 235–254.

715 Jamshidian, F. and Y. Zhu (1997). Scenario Simulation: Theory and methodology. *Finance and*
716 *Stochastic* 13, 4367.

717 Kou, S. G. (2002). A jump diffusion model for option pricing. *Management Science* 48, 1086–1101.

718 Maree, S. C., L. Ortiz-Gracia, and C. W. Oosterlee (2017). Pricing early-exercise and discrete
719 barrier options by Shannon wavelet expansions. *Numerische Mathematik* To appear.

720 Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of*
721 *Financial Economics* 3, 125–144.

722 Neuenkirch, A. and L. Szpruch (2014). First order strong approximations of scalar SDEs with
723 values in a domain. *Numerische Mathematik* 128, 103–136.

724 Ortiz-Gracia, L. and C. W. Oosterlee (2016). A highly efficient Shannon wavelet inverse Fourier
725 technique for pricing European options. *SIAM Journal on Scientific Computing* 38(1), B118–
726 B143.

727 Pillay, E. and J. O'Hara (2011). FFT based option pricing under a mean reverting process
728 with stochastic volatility and jumps. *Journal of Computational and Applied Mathematics* 235,
729 3378–3384.

730 Piterbarg, V. (2006). Smiling hybrids. *Risk magazine* 19(5), 66–70.

731 Rebonato, R. (1998). *Interest Rate Option Models.* John Wiley & Sons, Inc, second edition.

732 Sippel, J. and S. Ohkoshi (2002). All power to PRDC notes. *Risk magazine* 15(11), 1–3.

733 Stenger, F. (2011). *Handbook of Sinc Numerical Methods.* CRC Press.

734 Zhang, S. and L. Wang (2013). Fast Fourier transform option pricing with stochastic interest rate,
735 stochastic volatility and double jumps. *Applied Mathematics and Computation* 219, 10928–
736 10933.