



**HAL**  
open science

# Jacobi Stochastic Volatility factor for the Libor Market Model

Pierre-Edouard Arrouy, Sophian Mehalla, Bernard Lapeyre, Alexandre Boumezoued

► **To cite this version:**

Pierre-Edouard Arrouy, Sophian Mehalla, Bernard Lapeyre, Alexandre Boumezoued. Jacobi Stochastic Volatility factor for the Libor Market Model. 2020. hal-02468583v1

**HAL Id: hal-02468583**

**<https://hal.science/hal-02468583v1>**

Preprint submitted on 13 Feb 2020 (v1), last revised 7 Dec 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Jacobi Stochastic Volatility factor for the Libor Market Model

Pierre-Edouard Arrouy<sup>2</sup>, Alexandre Boumezoued<sup>2</sup>, Bernard Lapeyre<sup>1</sup> and Sophian Mehalla<sup>1,2</sup>  
<sup>1</sup>CERMICS and <sup>2</sup>Milliman R&D

February 13, 2020

## Abstract

We propose a new method to efficiently price swap rates derivatives under the LIBOR Market Model with Stochastic Volatility and Displaced Diffusion (DDSVLMM). This method uses polynomial processes combined with Gram-Charlier expansion techniques.

The standard pricing method for this model relies on dynamics freezing to recover an Heston-type model for which analytical formulas are available. This approach is time consuming and efficient approximations based on Gram-Charlier expansions have been recently proposed.

In this article, we first discuss the fact that for a class of stochastic volatility model, including the Heston one, the classical sufficient condition ensuring the convergence of the Gram-Charlier series can not be satisfied. Then, we propose an approximating model based on Jacobi process for which we can prove the stability of the Gram-Charlier expansion. For this approximation, we have been able to prove a strong convergence toward the original model; moreover, we give an estimate of the convergence rate. We also prove a new result on the convergence of the Gram-Charlier series when the volatility factor is not bounded from below. We finally illustrate our convergence results with numerical examples.

**Keywords:** Stochastic Volatility; Displaced Diffusion; Jacobi dynamics; Gram-Charlier expansions; Polynomial processes; LIBOR Market Model.

## 1 Introduction

Since the work of [30], [8] and [24], so-called market models have spread among practitioners. [25] extended the original LIBOR Market Model to both stochastic volatility and displaced diffusion, whereas [35] proposed a widely used version of the stochastic volatility component which is a Cox-Ingersoll-Ross (CIR) process; on this basis they provided several analytical results such as integral-based formulas for caplets and swaptions prices. Combining that two modelling frameworks yield the so-called Displaced Diffusion with Stochastic Volatility LIBOR Market Model (referred to as DDSVLMM), very popular within insurance market. Our main motivation is indeed to propose a modelling framework allowing for efficient Risk-Neutral valuation of the

balance sheet of insurers. Several other versions of the model proposed in [35] have been developed in the literature, whose differences mainly lie in the way of modelling the stochastic volatility component and the scope of instruments to be addressed. The model introduced in [32] is also very popular: the stochastic volatility factor is again a CIR process but uncorrelated to other drivers and the developed pricing methodology is built on successive suited approximations of the primary dynamics. An overview of existing modelling choices can be found in [10] and references therein.

Due to the complexity of the resulting swap rate dynamics in the DDSVLMM or its variants, as in particular the forward rates are involved in the drift of the stochastic volatility process, one resorts to the so-called freezing technique, see [1]. As a result, the swap rate dynamics can be assimilated to an Heston-type model (with time-varying coefficients), whose characteristic function can be analytically derived and then used for pricing, see [22] or [35]. The famous Fast Fourier Transform described in [13] used to pricing over a grid of strikes is especially suited for pricing when using such models.

Beyond numerical instability issues of the Heston characteristic function, as pointed out and solved by [7], recent interest has been dedicated to the significant computational time cost of these methods and the use of more efficient techniques for pricing. In particular, approximations based on Gram-Charlier and Edgeworth expansion techniques have been proposed by [15], who developed an adjustment of a reference Gaussian distribution (Bachelier model) for skewness and kurtosis, and as a by-product derived a smile formula linking the volatility to the moneyness with interpretable parameters. Pricing formulas under the DDSVLMM involving moments up to order four are provided in analytical form by [15] taking advantage of the explicit knowledge of the characteristic function in Heston-type models.

However, ensuring the convergence of the Gram-Charlier series in the case of such affine dynamics is not solved. In this paper we actually prove that the classical sufficient condition presented in [21] used to secure the convergence of the expansion technique is not satisfied for a class of unbounded stochastic volatility models including the classical Heston model.

In order to provide a modelling framework suitable for Gram-Charlier expansion series, we rely on the theory of polynomial processes. In particular, we use the so-called Jacobi process for the volatility as in [3], whose main feature is to allow for naturally bounding the volatility process while preserving the tractability of moments using matrix exponentials. The Jacobi process was initially introduced to study gene frequencies (see for instance [27]). Furthermore, applications to finance have been studied more recently: [17] worked with an interest-rate model based on a Jacobi dynamics whereas [29] studied a stochastic correlation adjustment modelled by a Jacobi process. This has also been studied in [18] and [12] with further applications in finance in [20].

Combining the works of [3] and [?], we propose a new approximation of the DDSVLMM. We will refer to this suggested model as Jacobi version of the DDSVLMM. Our purpose is to be able to price swap rates derivatives with Gram-Charlier expansions, as in [?], in a framework in which the convergence of the series holds using a Jacobi process, as in [3]. Indeed, [3] justified the use of Gram-Charlier expansion in an equity type context when bounding the stochastic volatility factor. Gram-Charlier expansion is a technique allowing to approximate the density of the modelled swap rate using as reference a Gaussian density. As such, the approximation

is made by a Bachelier-type model, to which adjustments are added in the form of higher moments of the swap rate distribution. It is particularly interesting due to the tractability and the interpretability of the Bachelier model. However, alternatives do exist to develop converging expansions approximations for Heston-type models. As pointed out in [21], a bilateral Gamma density can be used as reference density in the Heston model, at the price of computational difficulties of the orthonormal basis of polynomials, see for instance [4].

Our proposed model can be viewed as an approximation of the standard DDSVLMM. As in [3], the model converges weakly to the original model when volatility bounds vanish. Moreover, we have been able to prove this convergence in the  $\mathbb{L}^1$  space and obtain a convergence rate. Numerical investigations seem to indicate that the obtained convergence speed can be still improved and allow to conjecture the value of the optimal convergence rate. Our conjecture is supported by the derivation of a better speed of convergence for the weak error on the volatility process. While [3] provided a bound on the error made when truncating the Gram-Charlier series, in this work we use the obtained convergence rate to assess the pricing error made when comparing the use of Jacobi and standard CIR processes. We also have been able to prove a conjecture stated in [3] on the convergence of the Gram-Charlier series when the volatility factor is not bounded from below. Finally, note that we work here with time-dependent coefficients (mainly, piecewise constant), which extends the previous mentioned works.

This paper is structured as follows. In the end of this introductory section, we bring in the density approximation technique we are interested in, then we present a preliminary motivation result and finally we introduce the notion of (time-dependent) polynomial processes exemplified by the particular Jacobi process. In Section 2, after recalling the standard DDSVLMM framework, we specify our proposed approximating dynamics, especially by introducing the Jacobi model for the stochastic volatility factor. Our main results about convergence of the Gram-Charlier series,  $\mathbb{L}^1$  convergence of the approximating model along with the derivation of strong and weak convergence results are stated in this section. We provide numerical results and illustrations in Section 3. The paper ends with some concluding remarks.

**Notations** We consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  equipped with a filtration  $(\mathcal{F}_t)_{t \geq 0}$  satisfying usual conditions. In a financial context,  $\mathbb{P}$  can be viewed as the historical probability measure whereas the filtration will represent market information (quoted prices, observed smile, etc.). The latter is assumed to be generated by a multivariate Brownian motion. For two (local) martingales  $(X_t)_{t \geq 0}$  and  $(Y_t)_{t \geq 0}$ ,  $\langle X, Y \rangle_t$  will denote their quadratic variation at time  $t$ .  $Z \stackrel{d}{=} Z'$  means that distributions of  $Z$  and  $Z'$  are in fact the same whereas  $Z \stackrel{a.s.}{=} Z'$  stands for almost sure equality.  $\sigma(Z)$  is the sigma-algebra generated by the random variable  $Z$ .  $\mathbb{E}[\cdot]$  is the expectation associated to  $\mathbb{P}$ ,  $\mathbb{E}_x[X_T]$  is the conditional expectation of  $X_T$  given the starting point  $X_0 = x$ .

We will denote with bold font  $\mathbf{u}$  the vectors; the canonical scalar product between two vectors will be denoted  $\mathbf{u} \cdot \mathbf{v}$ . Unless otherwise stated,  $\|\mathbf{u}\|$  will represent the ( $\mathcal{L}^2$ -) norm induced by the scalar product i.e.  $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$ . If  $x \in \mathbb{R}$ , we will denote by  $\mathbf{x}$  the vector of which all coordinates are equal to  $x$ :  $\mathbf{x} = (x, x, \dots, x)$ .

### 1.1 Density approximation

In this section, we provide more insights on the Gram-Charlier expansion. It is a technique used to approximate an unknown density based on a reference - also named *auxiliary* - Gaussian distribution denoted by

$$g_r(x) = e^{-(x-\mu_r)^2/(2\sigma_r^2)}/\sqrt{2\pi\sigma_r^2}, \quad (1)$$

of mean  $\mu_r$  and variance  $\sigma_r^2$ . Note that generally, Gram-Charlier expansion (of type A) only refers to the particular case where  $\mu_r = 0$  and  $\sigma_r = 1$ .

Let  $Z$  be the random variable of unknown probability density  $f$  that is aimed at being approximated and define the  $\mathcal{L}_r^2$  Hilbert space as

$$\mathcal{L}_r^2 = \left\{ h : \mathbb{R} \rightarrow \mathbb{R} \text{ measurable such that } \|h\|_r^2 := \int_{\mathbb{R}} h(u)^2 g_r(u) du < \infty \right\}.$$

Let us introduce the family of polynomials (lying in  $\mathcal{L}_r^2$ ) defined by

$$H_n(x) = \frac{1}{\sqrt{n!}} \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k n!}{2^k k! (n-2k)!} \left( \frac{x - \mu_r}{\sigma_r} \right)^{n-2k}.$$

$(H_n)_{n \in \mathbb{N}}$  forms an Hilbertian basis of  $\mathcal{L}_r^2$  and in case where  $\mu_r = 0$  and  $\sigma_r^2 = 1$ , the  $H_n$  are generally known as the (normalized) Hermite polynomials (see for instance [19]). Thanks to this basis, we can define for any  $N \in \mathbb{N}$  pseudo-density functions by

$$f^{(N)}(x) = g_r(x) \sum_{n=0}^N c_n H_n(x).$$

These are called *pseudo*-densities since they can take negative values while always integrating to one. As long as the likelihood ratio  $\bar{f} = f/g_r$  belongs to  $\mathcal{L}_r^2$ ,  $f$  can be approximated in this space in the sense that

$$\frac{f^{(N)}}{g_r} \xrightarrow[N \rightarrow \infty]{\mathcal{L}_r^2} \bar{f}.$$

We emphasize that the condition  $f/g_r \in \mathcal{L}_r^2$  is equivalent to

$$\int_{\mathbb{R}} f(u)^2 e^{\frac{(u-\mu_r)^2}{2\sigma_r^2}} du < \infty, \quad (2)$$

and is a *sufficient* condition to ensure the convergence of  $(f^{(N)}/g_r)_{N \in \mathbb{N}}$  in  $\mathcal{L}_r^2$ . In this space, the standard inner product is defined by  $\langle h_1, h_2 \rangle_{\mathcal{L}_r^2} = \int_{\mathbb{R}} h_1(u) h_2(u) g_r(u) du$  for  $h_1, h_2 \in \mathcal{L}_r^2$ . Since the  $(H_n)_{n \in \mathbb{N}}$  are orthogonal in  $\mathcal{L}_r^2$  with respect to this inner product, the coefficients  $c_n$  are uniquely determined as  $c_n = \langle f/g_r, H_n \rangle_{\mathcal{L}_r^2} = \int_{\mathbb{R}} H_n(u) f(u) du = \mathbb{E}[H_n(Z)]$ ,  $n \in \mathbb{N}$ . Thus, the coefficients  $c_n$  are linear combination of *moments* of the unknown density  $f$ : to perform a density approximation following this technique, the moments of  $Z$  have to be computed.

Numerically, the Gram-Charlier expansion is performed up to a given order  $N \in \mathbb{N}$ . The assessment of the error due to the truncation of the series can be an issue. When the condition (2) is not satisfied, we can not conclude on the convergence of the approximating density; in addition, there is no estimation of the error induced by the numerical truncation of the approximating series. More details about approximation density techniques based on orthonormal polynomials expansion can be found in [21] and references given there.

The next subsection proves that for general stochastic volatility models (and in particular in the Heston model), (2) is not satisfied.

## 1.2 Preliminary result

The use of Gram-Charlier type techniques is appealing in financial applications, since pricing can be performed efficiently while only requiring the computation of moments up to some order. This has mostly been explored for modelling equity and associated derivatives, see e.g. [14], [31], [33], [11] and [23] among others, as well as for the DDSVLMM modelling framework for interest rates, see [15]. The aim of the present section is to analyze the core sufficient condition for the convergence of the Gram-Charlier series, and in particular to prove that this condition is not satisfied for general stochastic volatility model with unbounded volatility factor.

Let fix an horizon  $T$  and consider the following model:

$$\begin{cases} dS_t &= u(V_t)\lambda(t)dB_t \\ dV_t &= b(t, V_t)dt + h(t, V_t)dW_t \\ (S_0, V_0) &\in \mathbb{R} \times \mathbb{R}_+ \end{cases} \quad (3)$$

where the functions  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $b : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$  and  $h : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$  are assumed to be regular enough to ensure existence and uniqueness of a solution to (3).  $(S_t)_{t \leq T}$  stands for the financial driver (share price, interest-rate, ...) of interest - possibly at log-scale,  $(V_t)_{t \leq T}$  is a second source of risk, and  $(B_t)_{t \leq T}$  and  $(W_t)_{t \leq T}$  are two Brownian motions under the appropriate probability measure. Correlation between Brownian motions is driven by a coefficient  $\rho \in ]-1, 1[$  such that  $d\langle B, W \rangle_t = \rho dt$ . In the case where  $u(v) = \sqrt{v}$ ,  $(V_t)_{t \leq T}$  is interpreted as the instantaneous volatility of  $(S_t)_{t \leq T}$ . Moreover, when  $(V_t)_{t \leq T}$  is a Cox-Ingersoll-Ross (CIR) process, model (3) is an Heston type model. This is a particular case of interest since the standard approximation of the DDSVLMM is an Heston type model as well. Let  $g_T$  denote the unknown density of the random variable  $S_T$ , and consider the reference Gaussian distribution as introduced in Equation (1). The following result shows that the sufficient condition to the convergence of a Gram-Charlier expansion does not hold under some unboundedness assumption.

**Theorem 1.** *We assume that the map  $t \rightarrow \lambda(t)$  is bounded:*

$$0 < \lambda_{min}^2 \leq \lambda(t)^2 \leq \lambda_{max}^2 < \infty.$$

*We also assume that the cumulated variance  $\int_0^T u(V_s)ds$  is unbounded and is positive:*

$$\begin{aligned} \text{for any constant } M > 0, \mathbb{P}\left(\int_0^T u(V_s)ds \geq M\right) &> 0 \\ \text{and } \mathbb{P}\left(\int_0^T u(V_s)ds = 0\right) &= 0. \end{aligned} \quad (4)$$

*Then the sufficient condition (2) to the  $\mathcal{L}_r^2$ -convergence of the approximating density series of  $g_T$  is not satisfied, that is:*

$$\int_{\mathbb{R}} g_T(u)^2 e^{\frac{(u-\mu_T)^2}{2\sigma_r^2}} du = \infty.$$

*Proof.* We first derive an analytical expression for the density  $g_T$  of  $S_T$ . Consider  $(W_t^\perp)_{t \geq 0}$  a Brownian motion independent of  $(W_t)_{t \geq 0}$ . Since

$$S_T \stackrel{a.s.}{=} S_0 + \int_0^T u(V_s)\lambda(s)dB_s \stackrel{d}{=} S_0 + \rho \int_0^T u(V_s)\lambda(s)dW_s + \sqrt{1-\rho^2} \int_0^T u(V_s)\lambda(s)dW_s^\perp,$$

one can note that, conditionally to  $\mathcal{F}_T^W = \sigma(W_t, t \leq T)$ ,  $S_T$  is normally distributed with mean  $\tilde{S}_T := S_0 + \rho \int_0^T u(V_s) \lambda(s) dW_s$  and variance  $C_T^2 := (1 - \rho^2) \int_0^T u(V_s)^2 \lambda(s)^2 ds$ . By application of Jensen's inequality:

$$C_T^2 \geq \frac{(1 - \rho^2)}{T} \lambda_{\min}^2 \left( \int_0^T u(V_s) ds \right)^2 \quad \text{a.s.}, \quad (5)$$

and thus with Assumption (4), we deduce  $\mathbb{P}(C_T^2 = 0) = 0$ . Then, for any measurable function  $f$ , we have thanks to Fubini theorem

$$\begin{aligned} \mathbb{E}[f(S_T)] &= \mathbb{E}[\mathbb{E}[f(S_T) | \mathcal{F}_T^W]] = \mathbb{E} \left[ \int_{\mathbb{R}} f(x) \frac{1}{\sqrt{2\pi C_T^2}} e^{-\frac{(x - \tilde{S}_T)^2}{2C_T^2}} dx \right] \\ &= \int_{\mathbb{R}} f(x) \mathbb{E} \left[ \frac{1}{\sqrt{2\pi C_T^2}} e^{-\frac{(x - \tilde{S}_T)^2}{2C_T^2}} \right] dx \end{aligned}$$

which allows to identify the density of  $S_T$  as  $g_T(x) = \mathbb{E}[G_T(x)]$  where  $G_T(x) := (2\pi C_T^2)^{-1/2} e^{-\frac{(x - \tilde{S}_T)^2}{2C_T^2}}$ . Now, Equation (5) along with Assumption (4) shows that  $\mathbb{P}(C_T^2 > 2\sigma_r^2) > 0$  for all  $\sigma_r > 0$  and thus with positive probability  $Y_T := -\left(\frac{1}{2C_T^2} - \frac{1}{4\sigma_r^2}\right) > 0$ . Now, we have

$$\begin{aligned} &\int_{\mathbb{R}} \frac{g_T^2(x)}{g_r(x)} dx = \sqrt{2\pi\sigma_r^2} \int_{\mathbb{R}} \mathbb{E}[G_T(x)]^2 e^{(x - \mu_r)^2 / (2\sigma_r^2)} dx \\ &\geq \sqrt{2\pi\sigma_r^2} \int_{-a}^a \mathbb{E} \left[ \mathbb{1}_{C_T^2 > 2\sigma_r^2} \times G_T(x) e^{(x - \mu_r)^2 / (4\sigma_r^2)} \right]^2 dx \\ &\geq \frac{\sqrt{2\pi\sigma_r^2}}{2a} \left( \int_{-a}^a \mathbb{E} \left[ \mathbb{1}_{C_T^2 > 2\sigma_r^2} \times G_T(x) e^{(x - \mu_r)^2 / (4\sigma_r^2)} \right] dx \right)^2 \quad (\text{Jensen's inequality}) \\ &= \frac{\sigma_r}{2a} \mathbb{E} \left[ \frac{\mathbb{1}_{C_T^2 > 2\sigma_r^2}}{C_T} \int_{-a}^a e^{-\frac{(x - \tilde{S}_T)^2}{2C_T^2} + \frac{(x - \mu_r)^2}{4\sigma_r^2}} dx \right]^2 \quad (\text{Fubini theorem}) \\ &= \frac{\sigma_r}{2a} \mathbb{E} \left[ \frac{\mathbb{1}_{Y_T > 0}}{C_T} \exp \left( \frac{\mu_r^2}{4\sigma^2} - \frac{\tilde{S}_T^2}{2C_T^2} - \frac{1}{4Y_T} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)^2 \right) \times \right. \\ &\quad \left. \int_{-a}^a e^{\left( \sqrt{Y_T} x + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right) \right)^2} dx \right]^2 \\ &= \frac{\sigma_r}{2} \mathbb{E} \left[ \frac{\mathbb{1}_{Y_T > 0}}{\sqrt{Y_T} C_T} \exp \left( \frac{\mu_r^2}{4\sigma^2} - \frac{\tilde{S}_T^2}{2C_T^2} - \frac{1}{4Y_T} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)^2 \right) \times \right. \\ &\quad \left. \frac{1}{a} \int_{-\sqrt{Y_T} a + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)}^{\sqrt{Y_T} a + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)} e^{y^2} dy \right]^2. \end{aligned}$$

By Fatou's lemma

$$\begin{aligned} &\lim_{a \rightarrow \infty} \mathbb{E} \left[ \frac{\mathbb{1}_{Y_T > 0}}{\sqrt{Y_T} C_T} \exp \left( \frac{\mu_r^2}{4\sigma^2} - \frac{\tilde{S}_T^2}{2C_T^2} - \frac{1}{4Y_T} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)^2 \right) \frac{1}{a} \int_{-\sqrt{Y_T} a + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)}^{\sqrt{Y_T} a + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)} e^{y^2} dy \right] \\ &\geq \mathbb{E} \left[ \frac{\mathbb{1}_{Y_T > 0}}{\sqrt{Y_T} C_T} \exp \left( \frac{\mu_r^2}{4\sigma^2} - \frac{\tilde{S}_T^2}{2C_T^2} - \frac{1}{4Y_T} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)^2 \right) \times \right. \\ &\quad \left. \lim_{a \rightarrow \infty} \frac{1}{a} \int_{-\sqrt{Y_T} a + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)}^{\sqrt{Y_T} a + \frac{1}{2\sqrt{Y_T}} \left( \frac{\tilde{S}_T}{C_T} - \frac{\mu_r}{2\sigma^2} \right)} e^{y^2} dy \right]. \end{aligned}$$

Finally, the lower bound can be proved to be infinite since for any  $\lambda > 0$  and  $\mu \in \mathbb{R}$ ,

$$\frac{1}{a} \int_{-\lambda a + \mu}^{\lambda a + \mu} e^{y^2} dy \geq \frac{1}{a} \int_{\frac{\lambda a + \mu}{2}}^{\lambda a + \mu} e^{y^2} dy \geq \frac{\lambda a + \mu}{2a} e^{\frac{(\lambda a + \mu)^2}{4}},$$

so that  $\lim_{a \rightarrow \infty} \frac{1}{a} \int_{-\lambda a + \mu}^{\lambda a + \mu} e^{y^2} dy = \infty$ . This concludes the proof.  $\square$

Thus, to be able to perform a Gram-Charlier expansion and assess the truncation error for a stochastic volatility model, we should work with a bounded volatility factor.

In the next subsection, we introduce the class of polynomial processes which is a generalization of the class of affine processes; as such, it still benefits from analytical tractability of their marginal moments based on matrix exponentials, while loosing in general that of the moment generating function.

### 1.3 Polynomial processes

In the original papers [12] and [20], a theory of polynomial processes is developed for time-homogeneous dynamics: [12] focus on Markov process possibly with jumps whereas [20] deal with continuous dynamics but non-necessarily Markov. For such models, the expectation of any polynomial function of the process at a given date reduces to a polynomial in the initial condition with equal or lower degree.

The theory of polynomial processes has been extended by [16] to time-dependent parameters under appropriate smoothness assumption. In this subsection, we present continuous Markov processes whose diffusion parameters are piecewise constant with respect to time. We derive matrix exponential representations of the moments for polynomial processes with such time-dependency as in [?].

#### 1.3.1 Moments computation for polynomial processes

Let us denote by  $\mathcal{P}_k(\mathbb{R}^m)$  the set of polynomial functions of degree at most  $k \in \mathbb{N}$  on  $\mathbb{R}^m$ :  $\mathcal{P}_k(\mathbb{R}^m) = \left\{ \mathbb{R}^m \ni \mathbf{x} \mapsto \sum_{\alpha} c_{\alpha} x_1^{\alpha_1} \dots x_m^{\alpha_m} : c_{\alpha} \in \mathbb{R}, \max \{ \alpha_1 + \dots + \alpha_m : c_{\alpha} \neq 0 \} \leq k \right\}$ . We consider an  $E$ -valued time-inhomogeneous continuous stochastic process,  $(\mathbf{X}_t)_{0 \leq t \leq T}$ , specified by a Stochastic Differential Equation and whose infinitesimal generator at time  $t$  is denoted by  $\mathcal{A}_t$ ; its domain is denoted by  $\text{Dom}(\mathcal{A}_t)$ . We fix a time horizon  $T$ . We assume that for all  $t \leq T$ ,  $\mathcal{P}_k(\mathbb{R}^m) \subset \text{Dom}(\mathcal{A}_t)$ . For technical reasons, we assume that  $E \subset \mathbb{R}^m$  has a non-empty interior (see Section 2 in [20] for a detailed discussion). Following the definition of [20], a diffusion is said to be polynomial if  $\mathcal{A}_t$  maps  $\mathcal{P}_k(\mathbb{R}^m)$  to itself for all  $k \in \mathbb{N}$  and  $t \leq T$ .

In our work, we deal with piecewise constant infinitesimal generators. Let us split the whole time interval  $[0, T]$ : let  $0 \leq t_1 \leq \dots \leq t_J \leq T$  be instants such that for any  $t \in [t_j, t_{j+1}[$ ,  $\mathcal{A}_t \equiv \mathcal{A}_j$ ,  $1 \leq j \leq J - 1$ . As a linear operator, the action of each  $\mathcal{A}_j$  on  $\mathcal{P}_k(\mathbb{R}^m)$  can be represented in a unique way through a  $M \times M$  matrix,  $A_j^{(k)}$ , where  $M = \dim(\mathcal{P}_k(\mathbb{R}^m))$ . Let us consider a polynomial function  $p \in \mathcal{P}_k(\mathbb{R}^m)$ . In a given basis  $\mathbf{b}(\mathbf{x}) := (b_1(\mathbf{x}), b_2(\mathbf{x}), \dots, b_M(\mathbf{x}))$  of  $\mathcal{P}_k(\mathbb{R}^m)$ ,  $p \in \mathcal{P}_k(\mathbb{R}^m)$  can be uniquely represented thanks to a vector  $\mathbf{p} \in \mathbb{R}^M$  such that  $p(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \cdot \mathbf{p}$ . Thus,  $\mathcal{A}_j(p)(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \cdot A_j^{(k)} \mathbf{p}$ . We can now adapt one of the main property of polynomial process to the time-dependent case.



**Proposition 2.** Let  $k \in \{1, \dots, q\}$  and  $p \in \mathcal{P}_k(\mathbb{R}^m)$ . Assume that  $\mathbb{E}[\|\mathbf{X}_0\|^{2k}] < \infty$ . For  $t_1 < t_2 < \dots < t_J \leq t \leq T$ ,

$$\mathbb{E}[p(\mathbf{X}_t)|\mathcal{F}_0] = \mathbf{b}(\mathbf{X}_0) \cdot \left( \prod_{j=1}^J e^{(t_j - t_{j-1})A_{j-1}^{(k)}} \right) e^{(t-t_J)A_J^{(k)}} \mathbf{p}. \quad (6)$$

*Proof.* As a direct application of theorem 3.1 in [20], we have at any time  $t_j \leq t \leq t_{j+1}$

$$\mathbb{E}[p(\mathbf{X}_t)|\mathcal{F}_{t_j}] = \mathbf{b}(\mathbf{X}_{t_j}) \cdot e^{(t-t_j)A_j^{(k)}} \mathbf{p}.$$

Now, assuming in addition  $\mathbb{E}[\|\mathbf{X}_{t_{j-1}}\|^{2k}] < \infty$ , two successive conditionings at times  $t_{j-1}$  and  $t_j$  give for  $t \in [t_j, t_{j+1}[$ ,

$$\begin{aligned} & \mathbb{E}[p(\mathbf{X}_t)|\mathcal{F}_0] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{E}[p(\mathbf{X}_t)|\mathbf{X}_{t_j}]|\mathbf{X}_{t_{j-1}}]|\mathbf{X}_0] \\ &= \mathbb{E}[\mathbb{E}[\mathbf{b}(\mathbf{X}_{t_j})|\mathbf{X}_{t_{j-1}}]|\mathbf{X}_0] \cdot e^{(t-t_j)A_j^{(k)}} \mathbf{p} \\ &= \mathbb{E}\left[\left(\mathbb{E}[b_1(\mathbf{X}_{t_j})|\mathbf{X}_{t_{j-1}}], \dots, \mathbb{E}[b_M(\mathbf{X}_{t_j})|\mathbf{X}_{t_{j-1}}]\right)|\mathbf{X}_0\right] \cdot e^{(t-t_j)A_j^{(k)}} \mathbf{p} \\ &= \mathbb{E}\left[\left(\mathbf{b}(\mathbf{X}_{t_{j-1}}) \cdot e^{(t_j-t_{j-1})A_{j-1}^{(k)}} \mathbf{b}_1, \dots, \mathbf{b}(\mathbf{X}_{t_{j-1}}) \cdot e^{(t_j-t_{j-1})A_{j-1}^{(k)}} \mathbf{b}_M\right)|\mathbf{X}_0\right] e^{(t-t_j)A_j^{(k)}} \mathbf{p} \\ &= \mathbb{E}[\mathbf{b}(\mathbf{X}_{t_{j-1}})|\mathbf{X}_0] \cdot e^{(t_j-t_{j-1})A_{j-1}^{(k)}} e^{(t-t_j)A_j^{(k)}} \mathbf{p}. \end{aligned}$$

For the fifth equality, coordinates vector of the basis vectors  $(b_1(\mathbf{x}), \dots, b_M(\mathbf{x}))$  have been introduced as  $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_M)$ . Note that  $\mathbf{b}_i$  is only composed of zeros except at  $i$ -th coordinate that is a 1. A recursive reasoning gives the result.  $\square$

Note that the order in (6) matters as in general matrix exponentials may not commute. This way, taking a monomial function  $p(\mathbf{x}) = x_i^l$  for some  $i \in \llbracket 1, M \rrbracket$  and  $l \in \mathbb{N}$  and applying the formula (6) allows to compute marginal moments of the process  $X$ .

**Remark 1.** An alternative characterisation of the polynomial process can be done using the coefficients of the diffusion. Namely, the following stochastic differential equation, driven by a Brownian motion  $(W_t)_{t \leq T}$ ,

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t$$

would define a polynomial diffusion, according to the terminology of [20], if  $b \in \mathcal{P}_1(\mathbb{R}^m)$  and  $\sigma^2 \in \mathcal{P}_2(\mathbb{R}^m)$ . The case  $\sigma^2 \in \mathcal{P}_1(\mathbb{R}^m)$  defines an affine process.

We can now introduce the so-called Jacobi process and its dynamics which can be viewed as an alternative to the CIR process. It is a stochastic process which is naturally bounded over time under some technical conditions on the parameters. Moreover, it is a one-dimensional stochastic process belonging to the class of the polynomial processes. It will be useful for us in the following of this work.

### 1.3.2 Jacobi process

Let us consider the function

$$Q(v) = \frac{(v_{max} - v)(v - v_{min})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \quad (7)$$

where  $0 \leq v_{min} < v_{max} \leq \infty$ . We can observe that  $Q(v) \leq v$  for any  $v \in \mathbb{R}$  and that  $Q(v) \geq 0$  for  $v \in [v_{min}, v_{max}]$ .

Let us consider now the following dynamics

$$\begin{cases} dV_t = \kappa(\theta - V_t) dt + \epsilon\sqrt{Q(V_t)}dB_t, \\ V_0 \in [v_{min}, v_{max}], \end{cases} \quad (8)$$

where  $\kappa > 0$ ,  $\theta \in [v_{min}, v_{max}]$ ,  $\epsilon > 0$  and  $(B_t)_{t \geq 0}$  is a one-dimensional Brownian motion. Strong existence and uniqueness of the process  $V$  is studied in theorem 2.1 of [3]. Moreover, Feller' condition allows to ensure the process remains bounded through time. Namely,  $\mathbb{P}(\forall t \geq 0, V_t \in [v_{min}, v_{max}]) = 1$  if and only if  $V_0 \in ]v_{min}, v_{max}[$  and Feller condition

$$\frac{\epsilon^2(v_{max} - v_{min})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \leq 2\kappa \min(v_{max} - \theta, \theta - v_{min}) \quad (9)$$

is satisfied (see [26], Section 5.5.C, for a detailed treatment of the topic). The infinitesimal generator,  $\mathcal{G}_J$ , of (8) applied to a twice differentiable function  $f \in \mathcal{C}^2(\mathbb{R})$  writes as

$$\mathcal{G}_J f(v) = \kappa(\theta - v) \frac{\partial f}{\partial v} + \frac{\epsilon^2}{2} Q(v) \frac{\partial^2 f}{\partial v^2}.$$

Straightforward computations show that  $\mathcal{G}_J(\mathcal{P}_k(\mathbb{R})) \subset \mathcal{P}_k(\mathbb{R})$ : if  $p(v) = \sum_{i=0}^k c_i v^i$ ,

$$\begin{aligned} \mathcal{G}_J p(v) &= - \sum_{i=0}^k \left( i\kappa + \frac{i(i-1)\epsilon^2}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \right) c_i v^i + \sum_{i=0}^k \left( i\kappa\theta + \frac{i(i-1)\epsilon^2(v_{min} + v_{max})}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \right) c_i v^{i-1} \\ &\quad - \frac{\epsilon^2 v_{max} v_{min}}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \sum_{i=0}^k i(i-1) c_i v^{i-2} \in \mathcal{P}_k(\mathbb{R}). \end{aligned}$$

As a result, the Jacobi process is polynomial.

**Remark 2.** We propose to give details on the computation of marginal expectation of  $V$ . In the basis  $\{1, v\}$  of  $\mathcal{P}_1(\mathbb{R})$ , the action of  $\mathcal{G}_J$  can be represented by the  $2 \times 2$  matrix

$$G_J = \begin{pmatrix} 0 & \kappa\theta \\ 0 & -\kappa \end{pmatrix}.$$

Due to its upper triangular form, the exponential of  $G_J$  can be analytically computed. First observe that

$$G_J = \begin{pmatrix} 1 & -\theta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -\kappa \end{pmatrix} \begin{pmatrix} 1 & \theta \\ 0 & 1 \end{pmatrix}$$

and that the left sided matrix is the inverse of the right sided one. Then, for a given  $t \leq T$

$$\exp(tG_J) = \begin{pmatrix} 1 & \theta(1 - e^{-\kappa t}) \\ 0 & e^{-\kappa t} \end{pmatrix},$$

and thus

$$\mathbb{E}_{v_0}[V_t] = (1, v_0) \cdot \exp(tG_J) (0, 1)^T = v_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}).$$

Note that the expectation does not depend on  $v_{min}$  and  $v_{max}$  and is equal to the expectation of a CIR process. Recall that a CIR process is defined by taking  $Q(v) = v$  in (8) and  $V_0 \in \mathbb{R}_+$ .

## 2 Swaption pricing with Gram-Charlier expansion

### 2.1 Reference model

We introduce in this subsection the standard dynamics generally used for pricing under the Displaced Diffusion LIBOR Market Model with Stochastic Volatility. [25] proposed an extension of the standard LMM with respect to two features: a stochastic volatility factor has been added to reproduce the observed implied smile of volatility while a displacement coefficient (also called *shift*) allows to generate negative interest-rates, which has become necessary in view of late market conditions. [35] proposed a tractable and widely used (among insurance market) version of the model with stochastic volatility (but no displacement factor) modelled through a Feller process (Cox-Ingersoll-Ross dynamics) often named by practitioners DDSVLM.

Let  $P(t, T)$  be the time- $t$  price of a Zero-Coupon bond maturing at time  $T > t$  with par value 1. Let us consider a finite tenor structure  $0 \leq T_1 \leq T_2 \leq \dots \leq T_n$ . We denote by  $\Delta T_j = T_{j+1} - T_j$ . The swap rate seen at time  $t \leq T_m$  that prevails over the period  $[T_m, T_n]$ , can be expressed as

$$S_t^{m,n} = \frac{P(t, T_m) - P(t, T_n)}{\sum_{j=m}^{n-1} \Delta T_j P(t, T_{j+1})} \quad (10)$$

according to an arbitrage-free reasoning (see Section 1.5 in [10]). We will denote by  $B^S(t) := \sum_{j=m}^{n-1} \Delta T_j P(t, T_{j+1})$  the *annuity* of the swap rate. Under the probability measure  $\mathbb{P}^S$  (the *forward swap measure*, named after [24]) associated to the *numéraire*  $B^S(t)$ , the swap rate is a martingale. The standard modelling consists in giving to the swap rate (10) the following dynamics:

$$\begin{aligned} dS_t^{m,n} &= \sqrt{V_t} \left( \rho(t) \|\boldsymbol{\lambda}^{m,n}(t)\| dW_t + \sqrt{1 - \rho(t)^2} \boldsymbol{\lambda}^{m,n}(t) \cdot d\mathbf{W}_t^{S,*} \right), \\ dV_t &= \kappa (\theta - \xi^0(t) V_t) dt + \epsilon \sqrt{V_t} dW_t, \end{aligned} \quad (11)$$

where  $(\mathbf{W}_t^{S,*})_{t \geq 0}$  is a  $D$ -dimensional Brownian motion under  $\mathbb{P}^S$ , whose components are all independent from the Brownian motion  $(W_t)_{t \geq 0}$ . The coefficients  $\kappa$ ,  $\theta$ , and  $\epsilon$  are non-negative parameters. They are assumed to satisfy Feller condition  $2\kappa\theta \geq \epsilon^2$  (it coincides with (9) when  $v_{max} = \infty$  and  $v_{min} = 0$ ) that ensures the process  $V$  to remain non-negative almost surely, as long as  $V_0 > 0$ . The time-dependent coefficients are all bounded. In particular,  $\xi^0$  is positive and bounded:  $0 < \xi_{min}^0 \leq \xi^0(t) \leq \xi_{max}^0$ . Namely, in standard set-up all time dependent quantities are assumed to be piecewise constant on time intervals  $[T_i, T_{i+1}[$ . The function  $\rho$  accounts for the correlation between the swap rate and its instantaneous volatility;  $D$ -dimensional vector function  $\boldsymbol{\lambda}^{m,n}$  distorts the volatility structure over time;  $\xi^0$  is a deterministic adjustment in the drift term of the instantaneous volatility due to the correlation structure between swap rate and instantaneous volatility (see for instance [35]). To understand in detail the dynamics (11) and especially how the time-dependent functions are defined, we refer the reader to [35]. We also refer the interested reader to [9] for the treatment of the Heston model with time-dependent parameters.

Observe that (11) defines an Heston-type process. As an affine dynamics, (11) offers the ability of explicitly knowing the moment generating function of  $S^{m,n}$  through Riccati equations. This has been developed and solved in [?] or in [35] for the swap rate modelling. Exploiting

the explicit knowledge of the moment generating function allows to derive closed-form formulas for prices of swap rate derivatives (especially, for swaptions). In our work, we aim at pricing such derivatives using Gram-Charlier expansions. As discussed in Theorem 1, the convergence of such expansion can not be ensured *a priori* for dynamics (11). This is why, an approximation of this dynamics is suggested in the following subsection.

## 2.2 Jacobi dynamics approximation

We consider the following approximation of the model (11) referred to as Jacobi version of the DDSVLMM:

$$\begin{aligned} dS_t^{m,n} &= \sqrt{Q(V_t)}\rho(t)\|\boldsymbol{\lambda}^{m,n}(t)\|dW_t + \sqrt{V_t - \rho(t)^2Q(V_t)}\boldsymbol{\lambda}^{m,n}(t) \cdot d\mathbf{W}_t^{S,*} \\ dV_t &= \kappa(\theta - \xi^0(t)V_t)dt + \epsilon\sqrt{Q(V_t)}dW_t. \end{aligned} \quad (12)$$

where  $\boldsymbol{\lambda}^{m,n}$ ,  $\xi^0$  and  $\rho$  are the same as in (11),  $Q$  is defined in (7). We recall that all components of  $\mathbf{W}^{S,*}$  are independent and are also all independent from  $W$ . We also recall that we impose  $v_{min} < v_{max}$ . The volatility factor  $(V_t)_{t \geq 0}$  follows a Jacobi dynamics introduced in Section 1.3 with additional time dependency in the drift. For this dynamics, the Feller condition writes:

$$\frac{\epsilon^2(v_{max} - v_{min})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \leq 2\kappa \min(\xi_{min}^0 v_{max} - \theta, \theta - \xi_{max}^0 v_{min}). \quad (13)$$

It ensures the process  $V$  in (12) to remain bounded at any date  $t$ :  $\mathbb{P}(\forall t \geq 0 : V_t \in [v_{min}, v_{max}]) = 1$ . In this setting, the coefficient  $\rho(t)$  defined in (30) is interpreted in dynamics (12) as a scaling factor of the instantaneous correlation between the swap rate and its volatility since the following holds:

$$\frac{d\langle V, S^{m,n} \rangle_t}{\sqrt{d\langle V, V \rangle_t} \sqrt{d\langle S^{m,n}, S^{m,n} \rangle_t}} = \rho(t) \sqrt{\frac{Q(V_t)}{V_t}}.$$

We observed that  $0 \leq Q(v) \leq v$  for  $v \in [v_{min}, v_{max}]$  and thus the instantaneous correlation is smaller than  $\rho(t)$  at each time. The time-dependent infinitesimal generator of the diffusion (12) applied to a function  $[v_{min}, v_{max}] \times \mathbb{R} \ni (v, s) \mapsto f(v, s)$  is given by

$$\begin{aligned} \mathcal{A}_t f(v, s) &= \kappa(\theta - \xi^0(t)v) \frac{\partial f}{\partial v}(v, s) + \frac{\epsilon^2}{2} Q(v) \frac{\partial^2 f}{\partial v^2}(v, s) + \frac{v}{2} \|\boldsymbol{\lambda}^{m,n}(t)\|^2 \frac{\partial^2 f}{\partial s^2}(v, s) \\ &\quad + \epsilon Q(v) \rho(t) \|\boldsymbol{\lambda}^{m,n}(t)\| \frac{\partial^2 f}{\partial s \partial v}(v, s), \quad f \in \text{Dom}(\mathcal{A}_t), \quad t \leq T. \end{aligned} \quad (14)$$

Note that for all  $k \in \mathbb{N}$ , at all date  $t \in [0, T]$ ,  $\mathcal{A}_t(\mathcal{P}_k(\mathbb{R}^2)) \subset \mathcal{P}_k(\mathbb{R}^2)$ . Then, the dynamics (12) is a polynomial diffusion in the terminology of [20]; in addition it is a Markov process and thus (12) is a also 2-dimensional polynomial model in the terminology of [12]. Marginal moments of  $S^{m,n}$  solution of (12) can thus be computed by matrix exponentials following method presented in Section 1.3.

**Remark 3.** *The dynamics (12) is an extension of affine models as observed in [3] in the special case of the Heston model. Reference dynamics (11) belongs to the class of affine model. When  $v_{min} = 0$  and  $v_{max} = \infty$ , (12) coincides with the case when  $Q(v) = v$  which is the standard DDSVLMM (11). Thus (12) can be viewed as an approximation of the reference model (11). Namely, the proposed process (12) converges toward the standard one (11) in the path space*

of processes as  $v_{min} \rightarrow 0$  and  $v_{max} \rightarrow \infty$ . This result is proved in [3], Theorem 2.3. In Subsection 2.4, we will be able to further prove a  $\mathbb{L}^1$  convergence in Theorem 6 coming along with a convergence rate.

### 2.3 Gram-Charlier expansion

The Gram-Charlier expansion is justified now. We assume that the two following conditions hold:

$$\text{Assumption (A): } \begin{cases} 4\kappa\theta & > \epsilon^2, \\ 2\kappa(v_{max} - \theta) & \geq \epsilon^2, \end{cases} \quad (15)$$

and

$$\text{Assumption (B): } \sup_{t \in [0, T]} |\rho(t)| < 1. \quad (16)$$

Note that when  $v_{min} = 0$ , Feller condition (13) implies that  $4\kappa\theta > \epsilon^2$  and thus assumption (A) is stronger than (13).

#### 2.3.1 Convergence result

Each coordinate of the  $D$ -dimensional function  $t \in [0, T] \mapsto \boldsymbol{\lambda}^{m,n}(t)$  is piecewise constant and so is its norm  $t \in [0, T] \mapsto \|\boldsymbol{\lambda}^{m,n}(t)\|$ . Its bounds are denoted by  $\lambda_{min}$  and  $\lambda_{max}$  so that  $0 < \lambda_{min} \leq \|\boldsymbol{\lambda}^{m,n}(t)\| \leq \lambda_{max} < \infty$ . We define the cumulated volatility process

$$\Xi_t := \int_0^t \|\boldsymbol{\lambda}^{m,n}(s)\|^2 (V_s - \rho(s)^2 Q(V_s)) ds. \quad (17)$$

Let us denote by  $f_T$  the density of  $S_T^{m,n}$ . Recall that it depends on  $v_{min}$  and  $v_{max}$ . Following the notations in Section 1.1,  $f_T$  will be approximated with a reference Gaussian density  $g_r$ .

**Theorem 3.** *We suppose assumptions (A) and (B) hold,  $v_{min} \geq 0$  and  $v_{max} < \infty$ . Consider now a centered Gaussian density  $g_r$  of variance  $\sigma_r^2$  satisfying*

$$\sigma_r^2 > \frac{T v_{max}}{2} \lambda_{max}^2. \quad (18)$$

*Then, a Gram-Charlier expansion can be performed on the density  $f_T$  using the reference density  $g_r$ . In particular, the sufficient condition (2) to the  $\mathcal{L}_r^2$ -convergence of the family of approximating densities is satisfied; that is*

$$\int_{\mathbb{R}} \frac{f_T(u)^2}{g_r(u)} du < \infty.$$

The proof of the theorem is based on the following lemma.

**Lemma 4.** *For  $\eta < 1/(2T v_{max} \lambda_{max}^2)$ ,  $f_T$  is such that*

$$\int_{\mathbb{R}} e^{\eta u^2} f_T(u) du < \infty.$$

*In addition, if*

$$\mathbb{E}^S \left[ \Xi_T^{-1/2} \right] < \infty \quad (19)$$

*holds, then  $u \mapsto f_T(u)$  and  $u \mapsto e^{\eta u^2} f_T(u)$  are uniformly bounded over  $\mathbb{R}$ .*

**Remark 4.** In [3], the results claimed in Theorem 3 and Lemma 4 are proved (in a time-independent framework) under the following assumptions:

- $\rho^2 < 1$  which is equivalent to present Assumption (B);
- Feller condition (13) holds (with  $\xi^0(t) \equiv 1$ ), that is  $\frac{\epsilon^2(v_{max}-v_{min})}{(\sqrt{v_{max}}-\sqrt{v_{min}})^2} \leq 2\kappa \min(v_{max} - \theta, \theta - v_{min})$ , with  $0 < v_{min} < v_{max} < \infty$ .

They conjecture the results are still valid for  $v_{min} = 0$ . In our framework, Feller condition is again required to bound the volatility process, and Assumption (A) is used to prove the mentioned conjecture.

*Proof of the lemma.* We have from Equation (12):

$$\begin{aligned} S_T^{m,n} &\stackrel{a.s.}{=} S_0^{m,n} + \frac{1}{\epsilon} \int_0^T \rho(u) \|\boldsymbol{\lambda}^{m,n}(u)\| dV_u - \frac{\kappa}{\epsilon} \int_0^T \rho(u) \|\boldsymbol{\lambda}^{m,n}(u)\| (\theta - \xi^0(u)V_u) du \\ &\quad + \int_0^T \sqrt{V_u - \rho(u)^2 Q(V_u)} \boldsymbol{\lambda}^{m,n}(u) \cdot d\mathbf{W}_u^{S,*} \\ &=: \tilde{S}_T^{m,n} + \int_0^T \sqrt{V_u - \rho(u)^2 Q(V_u)} \boldsymbol{\lambda}^{m,n}(u) \cdot d\mathbf{W}_u^{S,*}. \end{aligned}$$

Since  $\mathbf{W}^{S,*}$  and  $V$  are independent,  $S_T^{m,n}$  is a Gaussian variable of  $\tilde{S}_T$  and variance  $\Xi_T$  conditionally to  $\sigma(V_t, t \leq T)$ . Since  $\lambda_{min} > 0$  and  $\sup_{t \in [0, T]} |\rho(t)|^2 < 1$ , note that  $\Xi_T > 0$  almost surely. The conditional density of  $S_T^{m,n}$  writes

$$F_T(u) = \frac{1}{\sqrt{2\pi\Xi_T}} \exp\left(-\frac{(u - \tilde{S}_T)^2}{2\Xi_T}\right).$$

Fubini's theorem allows to identify  $f_T(u) = \mathbb{E}^S[F_T(u)]$ . Hence,  $f_T$  is uniformly bounded as long as  $\mathbb{E}^S\left[\frac{1}{\sqrt{\Xi_T}}\right] < \infty$ . The maps  $[0, T] \ni t \mapsto \|\boldsymbol{\lambda}^{m,n}(t)\|$  and  $[0, T] \ni t \mapsto \rho(t)$  are piecewise constant, so their derivatives are almost everywhere zero. They also have a finite number of discontinuities over  $[0, T]$  occurring at times  $(T_i)_{1 \leq i \leq n(T)}$  where  $n(T) = \max\{i : T_i \leq T\}$ . Integrating with respect to the volatility process  $V$  over whole interval  $[0, T]$  gives

$$\tilde{S}_T \stackrel{a.s.}{=} S_0^{m,n} + \frac{1}{\epsilon} \sum_{i \leq n(T)} \rho(T_i) \|\boldsymbol{\lambda}^{m,n}(T_i)\| (V_{T_{i+1}} - V_{T_i}) - \frac{\kappa}{\epsilon} \int_0^T \rho(u) \|\boldsymbol{\lambda}^{m,n}(u)\| (\theta - \xi^0(u)V_u) du.$$

As a consequence, the following stands almost surely

$$|\tilde{S}_T| \leq S_0^{m,n} + \frac{2v_{max}}{\epsilon} \lambda_{max} + \frac{T\kappa}{\epsilon} \lambda_{max} (\theta + \sup_{0 \leq u \leq T} |\xi^0(u)| v_{max}) =: \Lambda < \infty.$$

Moreover, as

$$|\Xi_T| \leq T v_{max} \lambda_{max}^2,$$

we set

$$1 - 2\eta\Xi_T \geq 1 - 2\eta T v_{max} \lambda_{max}^2 =: \nu,$$

where  $0 < \nu$  since we assumed  $\eta < 1/(2T v_{max} \lambda_{max}^2)$ . Furthermore,

$$e^{\eta u^2} F_T(u) = \frac{1}{\sqrt{2\pi\Xi_T}} \exp\left(-\frac{1 - 2\eta\Xi_T}{2\Xi_T} \left(u - \frac{\tilde{S}_T}{1 - 2\eta\Xi_T}\right)^2 + \frac{\eta\tilde{S}_T^2}{1 - 2\eta\Xi_T}\right). \quad (20)$$

Integration of (20) gives

$$\int_{\mathbb{R}} e^{\eta u^2} F_T(u) du = \frac{1}{\sqrt{1-2\eta\Xi_T}} \exp\left(\frac{\eta\tilde{S}_T^2}{1-2\eta\Xi_T}\right) \leq \frac{e^{\eta\Lambda^2/\nu}}{\sqrt{\nu}} < \infty \text{ almost surely.}$$

Taking expectation in (20) leads to

$$e^{\eta u^2} f_T(u) = \mathbb{E}^S \left[ e^{\eta u^2} F_T(u) \right] \leq \mathbb{E}^S \left[ \frac{1}{\sqrt{\Xi_T}} \right] \frac{e^{\eta\Lambda^2/\nu}}{\sqrt{2\pi}}.$$

Thus,  $e^{\eta u^2} f_T(u)$  is uniformly bounded as soon as (19) holds.  $\square$

We can now prove Theorem 3.

*Proof of the theorem.* We proceed in two steps. First, we show that (19) holds under assumptions (A) and (B), for any  $v_{min} \geq 0$  and  $v_{max} < \infty$ ; secondly, we explain why (19) implies (2) when (18) is satisfied.

The treatment of the case  $v_{min} > 0$  is done in [3]. Assume that  $v_{min} = 0$ . Almost surely  $\lambda_{min}^2 \int_0^T (V_s - \rho(s)^2 Q(V_s)) ds \leq \Xi_T \leq \lambda_{max}^2 \int_0^T (V_s - \rho(s)^2 Q(V_s)) ds$ . Since  $\mathbb{R}_+^* \ni x \mapsto \frac{1}{\sqrt{x}}$  is convex, Jensen inequality implies

$$\frac{1}{\sqrt{\frac{1}{T} \int_0^T \|\lambda^{m,n}(s)\|^2 (V_s - \rho(s)^2 Q(V_s)) ds}} \leq \frac{1}{T \lambda_{min}} \int_0^T \frac{ds}{\sqrt{V_s - \rho(s)^2 Q(V_s)}} \text{ a.s..}$$

Taking expectation and applying Fubini theorem lead to

$$\mathbb{E}^S \left[ \Xi_T^{-1/2} \right] \leq \frac{1}{\lambda_{min} T^{3/2}} \int_0^T \mathbb{E}^S \left[ \frac{1}{\sqrt{V_s - \rho(s)^2 Q(V_s)}} \right] ds$$

with possibly infinite values. Since  $Q(v) \leq v$ , for any  $t \geq 0$

$$\sqrt{V_s(1 - \rho(t)^2)} \leq \sqrt{V_s - \rho(t)^2 Q(V_s)} \text{ a.s.}$$

and thus

$$\mathbb{E}^S \left[ \Xi_T^{-1/2} \right] \leq (1 - \sup_{t \in [0, T]} \rho(t)^2)^{-1/2} T^{-3/2} \lambda_{min}^{-1} \int_0^T \mathbb{E}^S \left[ \frac{1}{\sqrt{V_s}} \right] ds. \quad (21)$$

In order to exhibit a control of the right hand side, we consider the stopped process  $X_t^{\tau_n} := \sqrt{V_{t \wedge \tau_n}}$  where  $\tau_n, n \in \mathbb{N}^*$ , is the first time when the volatility process goes below the threshold  $1/n$ :  $\tau_n = \inf\{t \geq 0 : V_t \leq \frac{1}{n}\}$ . Ito's lemma applied to the stopped process gives

$$\begin{aligned} X_T^{\tau_n} &\stackrel{\text{a.s.}}{=} \sqrt{V_0} + \left( \frac{\kappa\theta}{2} - \frac{\epsilon^2}{8} \right) \int_0^{T \wedge \tau_n} \frac{du}{\sqrt{V_u}} + \int_0^{T \wedge \tau_n} \left( \frac{\epsilon^2}{8v_{max}} - \frac{\kappa\xi^0(s)}{2} \right) \sqrt{V_u} du \\ &\quad + \frac{\epsilon}{2} \int_0^{T \wedge \tau_n} \sqrt{1 - \frac{V_u}{v_{max}}} dW_u \end{aligned} \quad (22)$$

where we used that  $Q(v) = v - \frac{v^2}{v_{max}}$ . Ito integral is a martingale since  $1 - \frac{v}{v_{max}} \leq 1$  for  $v \in [0, v_{max}]$  and thus vanishes when taking the expectation in (22):

$$\begin{aligned} \mathbb{E}^S \left[ \sqrt{V_{T \wedge \tau_n}} - \sqrt{V_0} \right] &= \mathbb{E}^S \left[ \int_0^{T \wedge \tau_n} \left( \frac{\epsilon^2}{8v_{max}} - \frac{\kappa\xi^0(s)}{2} \right) \sqrt{V_s} ds \right] \\ &= \left( \frac{\kappa\theta}{2} - \frac{\epsilon^2}{8} \right) \mathbb{E}^S \left[ \int_0^{T \wedge \tau_n} \frac{ds}{\sqrt{V_s}} \right]. \end{aligned}$$

First recall that  $V_0 \in [0, v_{max}]$  (almost surely) and thus  $\mathbb{E}[\sqrt{V_0}] < \infty$ . Now, we aim at taking the limit as  $n$  goes to  $\infty$  in the previous equality to show that the right-hand side is finite. Since  $\tau_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \tau_0 \stackrel{\text{a.s.}}{=} \infty$  where  $\tau_0$  is the first time the volatility process hits 0,  $T \wedge \tau_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} T$ .  $\sqrt{V_t} \leq \sqrt{v_{max}}$  for any  $t \geq 0$ , and thus  $\mathbb{E}^S[\sqrt{V_{T \wedge \tau_n}}] \leq \sqrt{v_{max}}$  for all  $n \in \mathbb{N}$ . Lebesgue's dominated convergence theorem then provides that  $\mathbb{E}^S[\sqrt{V_T}] = \lim_n \mathbb{E}^S[\sqrt{V_{T \wedge \tau_n}}] \leq \sqrt{v_{max}}$ . Then, we recall that  $\xi^0$  is bounded on  $[0, T]$  which gives

$$\left| \mathbb{E}^S \left[ \int_0^{T \wedge \tau_n} \left( \frac{\epsilon^2}{8v_{max}} - \frac{\kappa \xi^0(s)}{2} \right) \sqrt{V_s} ds \right] \right| \leq \left( \frac{\epsilon^2}{8v_{max}} + \frac{\kappa |\xi_{max}^0|}{2} \right) \sqrt{v_{max}} T =: K_1 < \infty.$$

Using monotone convergence theorem, we have that

$$\int_0^{T \wedge \tau_n} \sqrt{V_u} du \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \int_0^T \sqrt{V_u} du \quad \text{and} \quad \int_0^{T \wedge \tau_n} \xi^0(u) \sqrt{V_u} du \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \int_0^T \xi^0(u) \sqrt{V_u} du.$$

This combined with dominated convergence theorem gives

$$\begin{aligned} \mathbb{E}^S \left[ \int_0^T \left( \frac{\epsilon^2}{8v_{max}} - \frac{\kappa \xi^0(s)}{2} \right) \sqrt{V_s} ds \right] &= \lim_{n \rightarrow \infty} \mathbb{E}^S \left[ \int_0^{T \wedge \tau_n} \left( \frac{\epsilon^2}{8v_{max}} - \frac{\kappa \xi^0(s)}{2} \right) \sqrt{V_s} ds \right] \\ &\leq K_1 < \infty. \end{aligned}$$

Observe that assumption **(A)** implies  $\left(\frac{\kappa\theta}{2} - \frac{\epsilon^2}{8}\right) > 0$ . Together with (23) gives the existence of a constant  $K < \infty$  such that

$$\lim_{n \rightarrow \infty} \mathbb{E}^S \left[ \int_0^{T \wedge \tau_n} \frac{du}{\sqrt{V_u}} \right] \leq K.$$

Fatou's lemma provides

$$\mathbb{E}^S \left[ \int_0^T \frac{du}{\sqrt{V_u}} \right] \leq \lim_{n \rightarrow \infty} \mathbb{E}^S \left[ \int_0^{T \wedge \tau_n} \frac{du}{\sqrt{V_u}} \right] < \infty,$$

which allows to deduce the claimed result

$$\mathbb{E}^S \left[ \Xi_T^{-1/2} \right] < \infty.$$

Now, following Lemma 4, as long as (19) holds,  $x \in \mathbb{R} \mapsto e^{\eta x^2} f_T(x)$  is uniformly bounded and integrable for  $\eta < 1/(2Tv_{max}\lambda_{max}^2)$ :  $e^{\eta x^2} f_T(x) \leq C < \infty$  and  $\int_{\mathbb{R}} e^{\eta x^2} f_T(x) dx < \infty$ . We apply this result with  $\eta = 1/(4\sigma_r^2) < 1/(2Tv_{max}\lambda_{max}^2) \iff \sigma_r^2 > v_{max}\lambda_{max}^2 T/2$ :

$$\begin{aligned} \int_{\mathbb{R}} \frac{f_T(x)^2}{g_r(x)} dx &= \sqrt{2\pi\sigma_r^2} \int_{\mathbb{R}} \left( e^{\frac{x^2}{4\sigma_r^2}} f_T(x) \right)^2 dx \\ &\leq C \sqrt{2\pi\sigma_r^2} \int_{\mathbb{R}} e^{\frac{x^2}{4\sigma_r^2}} f_T(x) dx < \infty. \end{aligned}$$

Hence the result.  $\square$

**Remark 5.** When using a non-centered Gaussian density  $g_r$  with mean  $\mu_r \in \mathbb{R}$ , a slight modification of the proof of Theorem 3 shows that Lemma 4 implies the convergence of the Gram-Charlier expansion for  $\sigma_r^2 > v_{max}\lambda_{max}^2 T$ .



### 2.3.2 Application to pricing of swap rate derivatives

In this section, we use some notations introduced in Section 1.1. The convergence of approximating densities built with Gram-Charlier method is ensured in  $\mathcal{L}_r^2$ . An application of the Cauchy-Schwarz inequality shows that the convergence of approximating prices can be deduced for square-integrable payoffs. Let us consider a (discounted) payoff  $\varphi \in \mathcal{L}_r^2$  and  $f_T$  the density function of  $S_T^{m,n}$  that is modelled with dynamics (12). The spot price of a European derivative  $P_T(\varphi)$  expiring at time  $T > 0$  associated to the considered payoff can be computed thanks to the likelihood ratio  $\bar{f}_T = f_T/g_r$ :

$$P_T(\varphi) = \int_{\mathbb{R}} \varphi(x) f_T(x) dx = \langle \varphi, \bar{f}_T \rangle_{\mathcal{L}_r^2} = \sum_{p \geq 0} h_p \varphi_p, \quad (23)$$

where the coefficients  $(\varphi_p)_{p \in \mathbb{N}}$  are given by

$$\varphi_p = \langle \varphi, H_p \rangle_{\mathcal{L}_r^2},$$

while the *Hermite moments*  $(h_p)_{p \in \mathbb{N}}$  are defined by

$$h_p = \langle \bar{f}_T, H_p \rangle_{\mathcal{L}_r^2} = \int_{\mathbb{R}} H_p(x) f_T(x) dx = \mathbb{E}^S [H_p(S_T^{m,n})].$$

The coefficients  $h_p$  are linear combinations of the moments of  $f_T$ . Since the polynomial functions  $H_p$  are analytically known (see Section 1.1), the only matter here is to be able to compute the moments of  $f_T$ . This can be achieved using the polynomial property of (12) discussed in Subsection 2.2. The approximating price

$$P_T^N(\varphi) = \sum_{p=0}^N h_p \varphi_p$$

is so that  $P_T^N(\varphi) \xrightarrow[N \rightarrow \infty]{} P_T(\varphi)$  once the Gram-Charlier convergence is ensured. Numerically, we have to truncate (23) at a given order  $N \in \mathbb{N}$ . However, the estimation of the speed of convergence of the series as  $N$  increases is not clear ([3] provided a numerical estimation of the truncation error).

**Remark 6.** For operational efficiency, a recursive relation can be derived as in [3] to compute the coefficients  $\varphi_p$ .

Pricing under model (12) using expansion techniques introduces two types of errors when compared to reference model (11): first one is due to the truncation of the Gram-Charlier series as discussed above, and second one is due to the boundedness of the volatility factor. Estimation of this second kind of error is the purpose of next section.

## 2.4 Rates of convergence and pricing error

We prove in this section the results about strong and weak convergence in the particular case where  $v_{min} = 0$ . We consider a sequence of upper bounds parameters  $(v_{max}^{(p)})_{p \in \mathbb{N}}$  such that  $v_{max}^{(p)} \xrightarrow[p \rightarrow \infty]{} +\infty$ . Let us denote by  $(S_t^{m,n,J(p)}, V_t^{J(p)})_{t \leq T}$  the solution of (12) associated to the bound parameters  $v_{max}^{(p)}$ .  $(S_t^{m,n,ref}, V_t^C)_{t \leq T}$  represents the solution of the reference dynamics (11). In particular,  $(V_t^C)_{t \geq 0}$  is a CIR process. For all  $p \in \mathbb{N}$ , we will consider an initial condition  $\mathbb{R}_+^* \ni v_0 = V_0^C = V_0^{J(p)} \leq \inf_{p \in \mathbb{N}} v_{max}^{(p)}$  and  $S_0^{m,n,J(p)} = S_0^{m,n,ref} = s_0^{m,n} \in \mathbb{R}$ .

### 2.4.1 Study of the moments

The following result shows that the marginal moments of the Jacobi process can be bounded independently from the parameters  $v_{max}^{(p)}$ .

**Lemma 5.** *For any  $k \geq 0$ ,  $\sup_{p \in \mathbb{N}} \sup_{0 \leq t \leq T} \mathbb{E}^S \left[ \left( V_t^{J(p)} \right)^k \right] \leq C_k$ .*

*Proof.* Fix a  $p \in \mathbb{N}$ . We first observe that for any time  $t \geq 0$ ,  $\mathbb{E}^S[V_t^{J(p)}] = \mathbb{E}^S[V_t^C]$  does not depend on  $v_{max}^{(p)}$  since the deterministic drifts in dynamics of  $(V_t^{J(p)})_{t \leq T}$  and  $(V_t^C)_{t \leq T}$  are the same (the computation has been done in Remark 2). Take  $k = 2$ . Applying Ito's formula to the process  $((V_t^{J(p)})^2)_{t \leq T}$  and taking the expectation lead to

$$\begin{aligned} & \mathbb{E}^S[(V_t^{J(p)})^2] \\ &= (v_0)^2 + \int_0^t \left\{ 2\kappa\theta \mathbb{E}^S[V_s^C] - 2\kappa\xi(s) \mathbb{E}^S[(V_s^{J(p)})^2] + \epsilon^2 \left( \mathbb{E}^S[V_s^C] - \frac{1}{v_{max}^{(p)}} \mathbb{E}^S[(V_s^{J(p)})^2] \right) \right\} ds \\ &\leq (v_0)^2 + \int_0^t \left\{ (2\kappa\theta + \epsilon^2) \mathbb{E}^S[V_s^C] - 2\kappa\xi(s) \mathbb{E}^S[(V_s^{J(p)})^2] \right\} ds. \end{aligned}$$

Gronwall's lemma gives

$$\mathbb{E}^S[(V_t^{J(p)})^2] \leq \left\{ (v_0)^2 + \int_0^t (2\kappa\theta + \epsilon^2) \mathbb{E}^S[V_s^C] ds \right\} \exp \left( -2\kappa \int_0^t \xi(s) ds \right)$$

and the claim is true for the constant  $C_2 = \left( (v_0)^2 + \int_0^T (2\kappa\theta + \epsilon^2) \mathbb{E}^S[V_s^C] ds \right) \exp \left( 2\kappa \|\xi\|_\infty T \right)$  that does not depend on  $v_{max}^{(p)}$  nor on  $t \leq T$ . Now assume that

$$\forall p \in \mathbb{N}, \forall t \leq T, \mathbb{E}^S[(V_t^{J(p)})^{k-1}] \leq C_{k-1}$$

holds for a given  $k \geq 2$ . Applying Ito's lemma again to  $((V_t^{J(p)})^k)_{t \geq 0}$  and taking expectation lead to, for  $t \leq T$ ,

$$\begin{aligned} \mathbb{E}^S[(V_t^{J(p)})^k] &= (v_0)^k + \int_0^t \left\{ \left( 2\kappa\theta + k(k-1) \frac{\epsilon^2}{2} \right) \mathbb{E}^S[(V_s^{J(p)})^{k-1}] \right. \\ &\quad \left. - \left( 2\kappa\xi(s) + \frac{\epsilon^2 k(k-1)}{2v_{max}^{(p)}} \right) \mathbb{E}^S[(V_s^{J(p)})^k] \right\} ds \\ &\leq (v_0)^k + \int_0^t \left\{ \left( 2\kappa\theta + k(k-1) \frac{\epsilon^2}{2} \right) \mathbb{E}^S[(V_s^{J(p)})^{k-1}] - 2\kappa\xi(s) \mathbb{E}^S[(V_s^{J(p)})^k] \right\} ds \end{aligned}$$

and Gronwall's lemma again shows that

$$\mathbb{E}^S[(V_t^{J(p)})^k] \leq \left[ (v_0)^k + t \left( 2\kappa\theta + k(k-1) \frac{\epsilon^2}{2} \right) C_{k-1} \right] \exp \left( - \int_0^t 2\kappa\xi(s) ds \right).$$

Hence the result holds for  $k$  with

$$C_k = \left( (v_0)^k + T \left( 2\kappa\theta + k(k-1) \frac{\epsilon^2}{2} \right) C_{k-1} \right) \exp \left( 2\kappa \|\xi\|_\infty T \right).$$

□

This result is useful to get a rate of convergence in the  $\mathbb{L}^1$  space and in a weak sense.

### 2.4.2 Strong convergence

We recall that  $(V_t^C)_{t \leq T}$  denotes the solution of the second stochastic differential equation (11). Note that the same Brownian motion  $(W_t)_{t \leq T}$  is used to build Cox-Ingersoll-Ross process in Equation (11) and Jacobi process in Equation (12).

**Theorem 6.** *There exists finite constants  $C, K \in \mathbb{R}$  such that for any  $n \in \mathbb{N}$*

$$\sup_{0 \leq t \leq T} \mathbb{E}^S \left[ |V_t^{J(n)} - V_t^C| \right] \leq C / \log \left( \frac{v_{max}^{(n)}}{v_0} \right),$$

and

$$\mathbb{E}^S \left[ \sup_{0 \leq t \leq T} |V_t^{J(n)} - V_t^C| \right] \leq K / \sqrt{\log \left( \frac{v_{max}^{(n)}}{v_0} \right)}.$$

*Proof.* In this proof, we draw from A. Alfonsi in [5], Section 4.3, who studied discretization scheme for CIR process and proved strong convergence.

Let us consider the family of Yamada functions,  $\psi_{\eta, m}$ , parametrized by two positive numbers  $\eta$  and  $m$ . Observe that for any  $m \geq 1$ ,  $\int_{\eta e^{-\epsilon^2 m}}^{\eta} \frac{1}{\epsilon^2 u} du = m$ . Then, we can find a continuous function  $\rho_{\eta, m}$ , with compact support included in  $]\eta e^{-\epsilon^2 m}, \eta[$ , and such that  $\rho_{\eta, m}(x) \leq \frac{2}{\epsilon^2 x m} \mathbb{1}_{x \in ]\eta e^{-\epsilon^2 m}, \eta[}$  and  $\int_{\eta e^{-\epsilon^2 m}}^{\eta} \rho_{\eta, m}(u) du = 1$ . Then,  $\psi_{\eta, m}$  is defined as

$$\psi_{\eta, m}(x) = \int_0^{|x|} \int_0^y \rho_{\eta, m}(u) du dy,$$

so that these functions satisfy

$$|x| - \eta \leq \psi_{\eta, m}(x) \leq |x|, \quad |\psi'_{\eta, m}(x)| \leq 1, \quad 0 \leq \psi''_{\eta, m}(x) = \rho_{\eta, m}(|x|) \leq \frac{2}{\epsilon^2 |x| m}.$$

$\psi_{\eta, m}$  is a nice smooth (twice differentiable) and even function approximating the absolute value function. Take  $n \in \mathbb{N}$ . Since

$$|V_t^{J(n)} - V_t^C| \leq \eta + \psi_{\eta, m}(V_t^{J(n)} - V_t^C), \quad (24)$$

we apply Ito's formula to the right-hand side to get

$$\begin{aligned} \psi_{\eta, m}(V_t^{J(n)} - V_t^C) &\stackrel{a.s.}{=} -\kappa \int_0^t (V_s^{J(n)} - V_s^C) \xi(s) \psi'_{\eta, m}(V_s^{J(n)} - V_s^C) ds \\ &\quad + \frac{\epsilon^2}{2} \int_0^t \left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right)^2 \psi''_{\eta, m}(V_s^{J(n)} - V_s^C) ds \\ &\quad + \epsilon \int_0^t \left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right) \psi'_{\eta, m}(V_s^{J(n)} - V_s^C) dW_s. \end{aligned} \quad (25)$$

Using that  $\sup_{s \geq 0} \mathbb{E}^S[V_s^C] < \infty$ ,  $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}$  (due to  $\frac{1}{2}$ -Hölder regularity of the square-root) and  $|\psi'_{\eta, m}(x)| \leq 1$  lead to

$$\left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right)^2 \psi'_{\eta, m}(V_s^{J(n)} - V_s^C) \leq \left| Q(V_s^{J(n)}) - V_s^C \right| \leq v_{max}^{(n)} + V_s^C \text{ a.s.}$$

which allows to deduce that the expectation of the Ito integral in (25) is zero. Using again  $\frac{1}{2}$ -Hölder regularity of the square-root, the second term in the right-hand side of (25) can be decomposed in the following way

$$\begin{aligned} & \int_0^t \left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right)^2 \psi''_{\eta,m} \left( V_s^{J(n)} - V_s^C \right) ds \\ & \leq \int_0^t \left( |Q(V_s^{J(n)}) - V_s^{J(n)}| \psi''_{\eta,m} \left( V_s^{J(n)} - V_s^C \right) + |V_s^{J(n)} - V_s^C| \psi''_{\eta,m} \left( V_s^{J(n)} - V_s^C \right) \right) ds. \end{aligned}$$

The first term in this integral is handled using that  $\|\psi''_{\eta,m}\|_\infty \leq \frac{2e^{\epsilon^2 m}}{\epsilon^2 \eta m}$  and

$$|Q(v) - v| \leq \frac{v^2}{v_{max}^{(n)}}. \quad (26)$$

The second term is bounded with  $|x| \psi''_{\eta,m}(x) \leq \frac{2}{\epsilon^2 m}$ . This leads to the almost sure inequality

$$\int_0^t \left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right)^2 \psi''_{\eta,m} \left( V_s^{J(n)} - V_s^C \right) ds \leq \int_0^t \left\{ \frac{(V_s^{J(n)})^2}{v_{max}^{(n)}} \frac{2e^{\epsilon^2 m}}{\epsilon^2 \eta m} + \frac{2}{\epsilon^2 m} \right\} ds.$$

Finally, for the first integral in (25), we use that  $\|\psi'_{\eta,m}\|_\infty \leq 1$  and thus

$$\left| -\kappa \int_0^t (V_s^{J(n)} - V_s^C) \xi(s) \psi'_{\eta,m}(V_s^{J(n)} - V_s^C) ds \right| \leq \kappa \int_0^t |V_s^{J(n)} - V_s^C| \|\xi\|_\infty ds \text{ a.s..}$$

Taking the expectation in (24) using the inequalities that have just been derived and using the fact that the moments of the Jacobi process are uniformly bounded with respect to  $v_{max}^{(n)}$  (cf. Lemma 5), we get that

$$\mathbb{E}^S[|V_t^{J(n)} - V_t^C|] \leq \eta + \kappa \|\xi\|_\infty \int_0^t \mathbb{E}^S[|V_s^{J(n)} - V_s^C|] ds + \frac{C' T e^{\epsilon^2 m}}{v_{max}^{(n)} \eta m} + \frac{T}{m}.$$

Gronwall's lemma again shows that

$$\mathbb{E}^S[|V_t^{J(n)} - V_t^C|] \leq e^{\kappa \|\xi\|_\infty T} \left( \eta + \frac{C' T e^{\epsilon^2 m}}{v_{max}^{(n)} \eta m} + \frac{T}{m} \right).$$

Taking  $\eta = \frac{1}{\log(v_{max}^{(n)}/v_0)}$  and  $m = \frac{\log(v_{max}^{(n)}/v_0)}{2\epsilon^2}$ , we get the existence of a constant  $C \in \mathbb{R}$  satisfying

$$\sup_{0 \leq t \leq T} \mathbb{E}^S[|V_t^{J(n)} - V_t^C|] \leq C / \log\left(\frac{v_{max}^{(n)}}{v_0}\right). \quad (27)$$

The Ito's integral in right-hand side of (25) can be more accurately handled. We have

$$\begin{aligned} & \mathbb{E}^S \left[ \sup_{0 \leq t \leq T} \int_0^t \left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right) \psi'_{\eta,m}(V_s^{J(n)} - V_s^C) dW_s \right] \\ & \leq C_{BDG} \mathbb{E}^S \left[ \left( \int_0^T \left( \sqrt{Q(V_s^{J(n)})} - \sqrt{V_s^C} \right)^2 \|\psi'_{\eta,m}\|_\infty^2 ds \right)^{1/2} \right] \\ & \leq C_{BDG} \sqrt{\int_0^T \left( \mathbb{E}^S[|Q(V_s^{J(n)}) - V_s^{J(n)}|] + \mathbb{E}^S[|V_s^{J(n)} - V_s^C|] \right) ds} \\ & \leq C_{BDG} \sqrt{\left( \frac{C_2}{v_{max}^{(n)}/v_0} + \frac{C}{\log(v_{max}^{(n)}/v_0)} \right) T} \end{aligned} \quad (28)$$

where we used successively: Burkholder-Davis-Gundy inequality,  $\|\psi'_{\eta,m}\|_\infty \leq 1$ , Jensen's one,  $\frac{1}{2}$ -Hölder regularity of the square-root and triangle inequality. This allows to deduce (28) after using inequality (26) combined with Lemma 5 for the first term in the square-root and the previous result (27) for the second term. Combining this with previous inequalities used in this proof leads to the existence of a constant  $K \in \mathbb{R}$  such that

$$\mathbb{E}^S \left[ \sup_{0 \leq t \leq T} |V_t^{J(n)} - V_t^C| \right] \leq K / \sqrt{\log \left( v_{max}^{(n)} / v_0 \right)}.$$

□

### 2.4.3 Weak convergence

We now state and prove the weak convergence result. In this subsection only, we assume that  $\xi^0(t) \equiv \xi^0$  does not depend on time. We first introduce some useful notations. Considering  $E$  a subset of  $\mathbb{R}^d$ , we will denote, for  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$  and a smooth function  $f$  defined over  $E$ , by  $\partial_\alpha f$  the following differentiation operator

$$x = (x_1, \dots, x_d) \in E \mapsto \partial_\alpha f(x) = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} f(x).$$

We then introduce the set of functions with derivatives of polynomial growth:

$$\mathcal{C}_{pol}^\infty(E) = \left\{ f : E \rightarrow \mathbb{R} : f \in \mathcal{C}^\infty, \forall \alpha \in \mathbb{N}^d, \exists C_\alpha > 0, \exists e_\alpha \in \mathbb{N}^*, \forall x \in E, |\partial_\alpha f(x)| \leq C_\alpha (1 + \|x\|^{e_\alpha}) \right\}.$$

We recall the following proposition proved in [6], Proposition 3.3.1.

**Proposition 7.** *Assume that  $f \in \mathcal{C}_{pol}^\infty(\mathbb{R})$ . The function  $u^C(t, x) = \mathbb{E}[f(V_{T-t}^C) | V_0^C = x]$  is well defined over  $[0, T] \times \mathbb{R}_+$ , is  $\mathcal{C}^\infty$ , satisfies the following backward Kolmogorov equation*

$$(P) : \begin{cases} t \in [0, T], x \in \mathbb{R}_+, \partial_t u^C + \kappa(\theta - \xi^0 x) \partial_x u^C + \frac{\epsilon^2}{2} x \partial_x^2 u^C = 0, \\ u^C(T, x) = f(x), \end{cases}$$

and is such that

$$\forall (l, m) \in \mathbb{N}^2, \exists (C_{l,m}, e_{l,m}) \in (\mathbb{R}_+^*)^2, \forall x \in \mathbb{R}_+, \forall t \in [0, T], \\ |\partial_t^l \partial_x^m u^C(t, x)| \leq C_{l,m} (1 + x^{e_{l,m}}).$$

We can now state our weak error estimation.

**Theorem 8.** *Let us consider  $f \in \mathcal{C}_{pol}^\infty(\mathbb{R})$ . Then there exists a constant  $K > 0$  such that for any  $n \in \mathbb{N}$ ,*

$$|\mathbb{E}[f(V_T^C)] - \mathbb{E}[f(V_T^{J(n)})]| \leq \frac{K}{v_{max}^{(n)}}.$$

**Remark 7.** *Numerical investigations suggest that this rate of convergence can be extended to strong convergence. As discussed in the numerical analysis Section 3.3, we conjecture that we can find a constant  $K$  (independent of  $(v_{max}^{(n)})_{n \geq 0}$ ) such that*

$$\mathbb{E}^S \left[ \sup_{0 \leq t \leq T} |V_t^{J(n)} - V_t^C| \right] \leq \frac{K}{v_{max}^{(n)}}.$$

*Proof.* The proof is built as in [34]: the authors proved a weak convergence of order 1 for numerical schemes approximating solutions of stochastic differential equations under proper assumptions. Consider  $u^C$  as defined in Proposition 7; in particular,  $u^C(T, x) = f(x)$ . We have

$$\mathbb{E}[f(V_T^C)] - \mathbb{E}[f(V_T^{J(n)})] = \mathbb{E}[u^C(T, V_T^C)] - \mathbb{E}[u^C(T, V_T^{J(n)})]$$

Thanks to Proposition 7, we have  $u^C(0, v_0) = \mathbb{E}[u^C(T, V_T^C)]$  and thus:

$$\mathbb{E}[f(V_T^C)] - \mathbb{E}[f(V_T^{J(n)})] = - \left( \mathbb{E}[u^C(T, V_T^{J(n)}) - u^C(0, v_0)] \right).$$

Ito's formula gives:

$$\begin{aligned} & u^C(T, V_T^{J(n)}) - u^C(0, v_0) \\ & \stackrel{a.s.}{=} \int_0^T \left\{ \partial_t u^C(s, V_s^{J(n)}) + \kappa(\theta - \xi^0 V_s^{J(n)}) \partial_x u^C(s, V_s^{J(n)}) \right. \\ & \quad \left. + \frac{\epsilon^2 Q(V_s^{J(n)})}{2} \partial_x^2 u^C(s, V_s^{J(n)}) \right\} ds + \epsilon \int_0^T \sqrt{Q(V_s^{J(n)})} \partial_x u^C(s, V_s^{J(n)}) dW_s. \end{aligned}$$

Using Proposition 7 and that the moments of the Jacobi process are all finite (see Lemma 5), we deduce that the expectation of Ito's integral is zero. With Fubini's theorem, we get

$$\mathbb{E}[u^C(T, V_T^{J(n)}) - u^C(0, v_0)] = \int_0^T \mathbb{E}[g^C(s, V_s^{J(n)})] ds$$

where  $g^C(t, x) = \partial_t u^C(t, x) + \kappa(\theta - \xi^0 x) \partial_x u^C(t, x) + \frac{\epsilon^2 Q(x)}{2} \partial_x^2 u^C(t, x)$ . Since  $u^C$  is solution of (P), we obtain

$$\begin{aligned} g^C(t, V_t^{J(n)}) &= \frac{\epsilon^2(Q(V_t^{J(n)}) - V_t^{J(n)})}{2} \partial_x^2 u^C(t, V_t^{J(n)}) \\ &= -\frac{\epsilon^2}{2v_{max}^{(n)}} (V_t^{J(n)})^2 \partial_x^2 u^C(t, V_t^{J(n)}), \end{aligned}$$

where we recall that we set  $v_{min} = 0$  in the bounding function  $Q$  defined in (7). Hence finally

$$\mathbb{E}[f(V_T^C)] - \mathbb{E}[f(V_T^{J(n)})] = \frac{\epsilon^2}{2v_{max}^{(n)}} \int_0^T \mathbb{E}[(V_s^{J(n)})^2 \partial_x^2 u^C(s, V_s^{J(n)})] ds.$$

Using that the derivatives of  $u^C$  are all of polynomial growth for a proper terminal condition  $f$  and that the moments of the Jacobi process can be bounded independently from  $v_{max}^{(n)}$  (see Lemma 5), we obtain the claimed result.  $\square$

#### 2.4.4 Pricing error

We study in this section the error made on swaption prices when using Jacobi dynamics for the volatility process compared to the standard model. The following proposition outlines the result for exotic options whose payoff depends on a path of the swap rate process. We focus on payoff  $\varphi$  that is Lipschitz in the following sense: given two continuous processes  $(x(t))_{t \in [0, T]}$  and  $(y(t))_{t \in [0, T]}$ , there exists a constant  $C_{Lip} > 0$  such that:

$$\left| \varphi \left( (x(t))_{t \in [0, T]} \right) - \varphi \left( (y(t))_{t \in [0, T]} \right) \right| \leq C_{Lip} \sup_{0 \leq t \leq T} |x(t) - y(t)|.$$

The pricing error is defined by:

$$\epsilon_{model} := \left| \mathbb{E}^S \left[ \varphi \left( \left( S_t^{m, n, ref} \right)_{t \leq T} \right) \right] - \mathbb{E}^S \left[ \varphi \left( \left( S_t^{m, n, J} \right)_{t \leq T} \right) \right] \right|.$$

**Proposition 9.** *There exists constants  $K_1$  and  $K_2 \in \mathbb{R}$  such that*

$$\epsilon_{model} \leq \sqrt{\frac{K_1}{\log(v_{max}/v_0)} + \frac{K_2}{v_{max}/v_0}}.$$

*Proof.* Similarly to the preceding proof, we will use successively triangle inequality, the Burkholder-Davis-Gundy inequality,  $\frac{1}{2}$ -Hölder regularity of the square-root function ( $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}$ , for  $x, y \geq 0$ ), the fact that for all  $t \leq T$ ,  $\rho(t)^2 Q(v) \leq \rho(t)^2 v \leq v$  since  $\rho(t)^2 \leq 1$  and that the square-root function is concave to get the following:

$$\begin{aligned} \epsilon_{model} &\leq C_{\text{Lip}} \mathbb{E}^S \left[ \sup_{0 \leq t \leq T} |S_t^{m,n,ref} - S_t^{m,n,J}| \right] \\ &= C_{\text{Lip}} \mathbb{E}^S \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \rho(s) \left( \sqrt{V_s^C} - \sqrt{Q(V_s^J)} \right) \|\boldsymbol{\lambda}^{m,n}(s)\| dW_s \right. \right. \\ &\quad \left. \left. + \int_0^t \left( \sqrt{V_s^C - \rho(s)^2 V_s^C} - \sqrt{V_s^J - \rho(s)^2 Q(V_s^J)} \right) \boldsymbol{\lambda}^{m,n}(s) \cdot dW_s^{S,*} \right| \right] \\ &\leq C_{\text{Lip}} C_{BDG} \left\{ \mathbb{E}^S \left[ \sqrt{\int_0^T \rho(s)^2 \left( \sqrt{V_s^C} - \sqrt{Q(V_s^J)} \right)^2 \|\boldsymbol{\lambda}^{m,n}(s)\|^2 ds} \right] \right. \\ &\quad \left. + \mathbb{E}^S \left[ \sqrt{\int_0^T \left( \sqrt{V_s^C - \rho(s)^2 V_s^C} - \sqrt{V_s^J - \rho(s)^2 Q(V_s^J)} \right)^2 \|\boldsymbol{\lambda}^{m,n}(s)\|^2 ds} \right] \right\} \\ &\leq C_{\text{Lip}} C_{BDG} \lambda_{max} \left( \mathbb{E}^S \left[ \sqrt{\int_0^T \rho(s)^2 |V_s^C - Q(V_s^J)| ds} \right] \right. \\ &\quad \left. + \mathbb{E}^S \left[ \sqrt{\int_0^T |V_s^C - V_s^J + \rho(s)^2 (Q(V_s^J) - V_s^C)| ds} \right] \right) \\ &\leq C \left( \sqrt{\int_0^T \rho(s)^2 \mathbb{E}^S [|V_s^C - Q(V_s^J)|] ds} \right. \\ &\quad \left. + \sqrt{\int_0^T \mathbb{E}^S [|V_s^C - V_s^J + \rho(s)^2 (Q(V_s^J) - V_s^C)|] ds} \right) \\ &=: C(I_1 + I_2) \end{aligned}$$

where  $C$  is a non-negative number. Observing that

$$\begin{aligned} I_1 &= \left( \int_0^T \rho(s)^2 \mathbb{E}^S [|V_s^C - V_s^J + V_s^J - Q(V_s^J)|] ds \right)^{1/2} \\ &\leq \left( \int_0^T (\mathbb{E}^S [|V_s^C - V_s^J|] + \mathbb{E}^S [|V_s^J - Q(V_s^J)|]) ds \right)^{1/2} \end{aligned}$$

and that

$$\begin{aligned} I_2 &= \left( \int_0^T \mathbb{E}^S [|V_s^C - V_s^J + \rho(s)^2 (Q(V_s^J) - V_s^C)|] ds \right)^{1/2} \\ &\leq \left( \int_0^T ((1 + \rho(s)^2) \mathbb{E}^S [|V_s^C - V_s^J|] + \rho(s)^2 \mathbb{E}^S [|Q(V_s^J) - V_s^C|]) ds \right)^{1/2} \\ &\leq \left( \int_0^T (2\mathbb{E}^S [|V_s^C - V_s^J|] + \mathbb{E}^S [|Q(V_s^J) - V_s^C|]) ds \right)^{1/2}, \end{aligned}$$

we are led to study the quantities  $\mathbb{E}^S [|V_s^C - V_s^J|]$  and  $\mathbb{E}^S [|Q(V_s^J) - V_s^J|]$ . Recall that  $v - Q(v) = \frac{v^2}{v_{max}}$  for  $v_{min} = 0$ . Regarding the results of Lemma 5 and Theorem 6, we can find some constants  $c_1, c_2$  such that

$$I_1 \leq \sqrt{T} \sqrt{\frac{c_1}{\log(v_{max}/v_0)} + \frac{c_2}{v_{max}/v_0}} \text{ and } I_2 \leq \sqrt{T} \sqrt{\frac{2c_1}{\log(v_{max}/v_0)} + \frac{c_2}{v_{max}/v_0}}$$

and thus for some constants  $K_1$  and  $K_2$ ,

$$\epsilon_{model} \leq C' \sqrt{\frac{K_1}{\log(v_{max}/v_0)} + \frac{K_2}{v_{max}/v_0}}.$$

where  $C' \geq 0$ . □

### 3 Numerical analysis

#### 3.1 Matrix exponential computation

To perform Gram-Charlier type expansion at an arbitrary order  $k \in \mathbb{N}$ , we need to represent the action of the infinitesimal generator through a matrix before taking its exponential. Consider the standard basis of  $\mathcal{P}_k(\mathbb{R}^2)$  composed of monomial functions:  $((v, s) \mapsto v^p s^q)_{p+q \leq k}$ . Given an enumeration  $\Lambda : E_k \rightarrow \mathbb{N}$  of the set of exponents  $E_k = \{(p, q) \in \mathbb{N}^2 : p + q \leq k\}$ , the non-zero elements of the matrix  $A^{(k)}$  representing the action of  $\mathcal{A}_t$  on  $\mathcal{P}_k(\mathbb{R}^2)$  are given by:

$$\begin{aligned} A_{\Lambda(p,q), \Lambda(p,q)}^{(k)} &= - \left( \frac{p(p-1)\epsilon^2}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2} + p\kappa\xi^0(t) \right), \\ A_{\Lambda(p-1,q), \Lambda(p,q)}^{(k)} &= \kappa p\theta + \frac{p(p-1)\epsilon^2(v_{min} + v_{max})}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \\ A_{\Lambda(p-2,q), \Lambda(p,q)}^{(k)} &= - \frac{p(p-1)\epsilon^2 v_{min} v_{max}}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \\ A_{\Lambda(p+1, q-2), \Lambda(p,q)}^{(k)} &= \frac{q(q-1) \|\boldsymbol{\lambda}^{m,n}(t)\|^2}{2}, \\ A_{\Lambda(p+1, q-1), \Lambda(p,q)}^{(k)} &= - \frac{pq\epsilon\rho(t) \|\boldsymbol{\lambda}^{m,n}(t)\|}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \\ A_{\Lambda(p-1, q-1), \Lambda(p,q)}^{(k)} &= - \frac{pq\epsilon v_{min} v_{max} \rho(t) \|\boldsymbol{\lambda}^{m,n}(t)\|}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \\ A_{\Lambda(p, q-1), \Lambda(p,q)}^{(k)} &= \frac{pq\epsilon\rho(t) \|\boldsymbol{\lambda}^{m,n}(t)\| (v_{min} + v_{max})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2}. \end{aligned}$$

#### 3.2 Specification of the Jacobi version of the DDSVLM

In our modelling framework, the vectors and multi-dimensional Brownian motions are of length 2. We consider a piecewise constant parametrization of the time dependency. This set-up corresponds to a *frozen* approximation of the Libor Market Model as motivated in [8], [10] or [35]. The coefficients appearing in dynamics (11) and (12) are defined as:

$$\xi^0(t) = 1 + \frac{\epsilon}{\kappa} \sum_{j=m}^{n-1} \alpha_j(0) \xi_j^0(t), \quad \xi_j^0(t) = \sum_{k=1}^j \frac{\Delta T_k (F_k(0) + \delta)}{1 + \Delta T_k F_k(0)} \rho_k(t) \|\boldsymbol{\gamma}_k(t)\|, \quad (29)$$



and

$$\boldsymbol{\lambda}^{m,n}(t) = \sum_{j=m}^{n-1} \omega_j(0) \boldsymbol{\gamma}_j(t), \quad \rho(t) = \frac{1}{\|\boldsymbol{\lambda}^{m,n}(t)\|} \sum_{j=m}^{n-1} \omega_j(0) \|\boldsymbol{\gamma}_j(t)\| \rho_j(t). \quad (30)$$

The quantities  $(F_j(0))_{j=m,\dots,n}$  are forward rates quoted on markets and the coefficients  $\omega_j$  defined for  $m \leq j \leq n-1$ , by

$$\omega_j(0) := \frac{\Delta T_j P(0, T_{j+1})}{B^S(0)} \left( 1 + \frac{\Delta T_j}{1 + \Delta T_j F_j(0)} \sum_{l=m}^{j-1} (F_l(0) - S_0^{m,n}) \right) (F_j(0) + \delta),$$

$\delta \in \mathbb{R}$  is a parameter often named *shift*. The volatility vectors are specified as  $\boldsymbol{\gamma}_j(T_i) = g(T_j - T_i) \boldsymbol{\beta}_{j-i+1}$  over the interval  $[T_i, T_{i+1}[$ . The  $\boldsymbol{\beta}_k$  are 2-dimensional vectors with unitary Euclidian norm, while  $g(u) = (bu + a)e^{-cu} + d$ , where  $a, b, c$  and  $d$  are non-negative constants. Finally, the coefficients  $\rho_j$  are parametrized thanks to a coefficient  $\rho \in [-1, 1]$  as

$$\rho_j(t) = \frac{\rho}{\sqrt{2}} \frac{\boldsymbol{\gamma}_j^{(1)}(t) + \boldsymbol{\gamma}_j^{(2)}(t)}{\|\boldsymbol{\gamma}_j(t)\|}.$$

### Theoretical assumptions for Gram-Charlier convergence

We rewrite the theoretical assumptions made in Section 2. First about assumption **(B)**: we assume that  $\rho^2 < 1$ . In view of the form of the correlation coefficients  $\rho_k$ , we have that

$$\begin{aligned} \rho(t) &= \frac{1}{\|\boldsymbol{\lambda}^{m,n}(t)\|} \sum_{j=m}^{n-1} \omega_j(0) \|\boldsymbol{\gamma}_j(t)\| \rho_j(t) \\ &= \frac{\rho}{\sqrt{2} \|\boldsymbol{\lambda}^{m,n}(t)\|} \sum_{j=m}^{n-1} \omega_j(0) \left( \boldsymbol{\gamma}_j^{(1)}(t) + \boldsymbol{\gamma}_j^{(2)}(t) \right). \end{aligned}$$

Observe that for any time  $t$ ,

$$\begin{aligned} \frac{1}{\rho(t)^2} &= \frac{2 \|\boldsymbol{\lambda}^{m,n}(t)\|^2}{\rho^2 \left( \sum_{j=m}^{n-1} \omega_j(0) \left( \boldsymbol{\gamma}_j^{(1)}(t) + \boldsymbol{\gamma}_j^{(2)}(t) \right) \right)^2} \\ &= \frac{2}{\rho^2} \frac{\left( \sum_{j=m}^{n-1} \omega_j(0) \boldsymbol{\gamma}_j^{(1)}(t) \right)^2 + \left( \sum_{j=m}^{n-1} \omega_j(0) \boldsymbol{\gamma}_j^{(2)}(t) \right)^2}{\left( \sum_{j=m}^{n-1} \omega_j(0) \boldsymbol{\gamma}_j^{(1)}(t) + \sum_{j=m}^{n-1} \omega_j(0) \boldsymbol{\gamma}_j^{(2)}(t) \right)^2} \geq \frac{1}{\rho^2}, \end{aligned}$$

and thus Assumption **(B)** is satisfied. Assumption **(A)**, that is  $4\kappa\theta > \epsilon^2$  and  $2\kappa(v_{max} - \theta) \geq \epsilon^2$ , is straightforward to check.

Besides assumptions **(A)** and **(B)**, the following has to hold in order to ensure the convergence of the Gram-Charlier expansion (see Theorem 3 and assumption in Equation (18)):

$$\sigma_r^2 > \frac{T v_{max}}{2} \lambda_{max}^2 \Leftrightarrow \frac{2\sigma_r^2}{T v_{max}} > \lambda_{max}^2$$

where we recall that  $T$  is the maturity of the priced derivative and  $\lambda_{max}^2 = \max\{t \geq 0 : \|\boldsymbol{\lambda}^{m,n}(t)\| = \sum_{j=m}^{n-1} \omega_j(0) \boldsymbol{\gamma}_j(t)\}$ . This constraint can be numerically checked.

### 3.3 Convergence illustration

#### Convergence of the Jacobi process toward the CIR process

The  $\mathbb{L}^1$  convergence of the Jacobi process toward the CIR has been discussed in Section 2.4. It is illustrated here. Firstly, we describe the discretization schemes used to simulate volatility processes. We chose to adapt the scheme presented in [28]. More elaborated schemes can be found in [6] for the CIR process. The discretized Jacobi process we used writes

$$V_{t_{i+1}} = V_{t_i} + \kappa \left( \theta - \xi^0(t_i) \min((V_{t_i})_+, v_{max}) \right) (t_{i+1} - t_i) + \epsilon \sqrt{Q(\min(V_{t_i})_+, v_{max})} (W_{t_{i+1}} - W_{t_i}), \quad (31)$$

with  $x_+ = \max(x, 0)$ . CIR process is simulated by putting  $v_{max} = \infty$  (and thus  $Q(v) = v$ ) in this scheme.

The quantity  $\mathbb{E}^S \left[ \sup_{0 \leq t \leq T} |V_s^{J, v_{max}} - V_s^C| \right]$  is estimated by Monte-Carlo samples and plotted in Figure 1 as a function of  $v_{max}$  with  $10^5$  Monte-Carlo samples. The upper volatility bound  $v_{max}$  takes values in

$$\{0.5, 0.7, 1.0, 1.2, 1.5, 1.8, 2, 3, 4, 5, 10, 10^2, 10^3, 10^4, 10^5\}.$$

All our results are obtained with  $v_{min} = 0$ : this choice is justified by the fact that theoretical results are of particular interest when  $v_{min} = 0$ . The time horizon is fixed to  $T = 5$  years and the volatility processes are discretized with a time step  $\Delta t = 0.05$ . The parameters are  $\kappa = 1$ ,  $\theta = 0.3$ ,  $\epsilon = 0.6$ ,  $V_0^C = V_0^{J, v_{max}} = v_0 = 0.2$ . We set  $\xi_0(t) \equiv 1$  for simplicity. We recall that Theorem 6 gives the following bounds  $\sup_{0 \leq t \leq T} \mathbb{E}^S \left[ |V_t^{J(n)} - V_t^C| \right] \leq C / \log \left( \frac{v_{max}^{(n)}}{v_0} \right)$  and  $\mathbb{E}^S \left[ \sup_{0 \leq t \leq T} |V_t^{J(n)} - V_t^C| \right] \leq K / \sqrt{\log \left( \frac{v_{max}^{(n)}}{v_0} \right)}$ . In Figure 1, the theoretical upper bound  $1 / \sqrt{\log(v_{max}/v_0)}$  is also plotted. It clearly overestimates realized error, especially for small values of  $v_{max}$ .

To complete, the Figure 2 provides the  $\mathbb{L}^1$  convergence of marginal moments as a function of the upper bound parameter. Namely,  $\mathbb{E}^S \left[ |V_T^{J, v_{max}} - V_T^C| \right]$  is plotted for maturities  $T = 1, 10$ . The theoretical upper bound obtained for this error is  $1 / \log(v_{max})$ , and again it clearly overestimates the empirical quantities, notably for small values of  $v_{max}$ .

In Figure 3 the logarithms of previous estimated errors are plotted, that is  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|] \right)$  and  $\log \left( \mathbb{E}^S [|V_T^{J, v_{max}} - V_T^C|] \right)$ ,  $T = 1$  and  $T = 10$  with respect to  $\log \left( \log \left( \frac{v_{max}}{v_0} \right) \right)$ .

**Remark 8.** We can distinguish between two regimes in the curves:  $\log(\log(v_{max}/v_0)) \in [0.0, 1.5]$  and  $\log(\log(v_{max}/v_0)) \geq 1.5$ . A linear regression applied on each part of the curve give satisfactory results, as illustrated in Figure 4 for the quantity  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|] \right)$ . The obtained slopes of each line are respectively  $-1.88$  and  $-7.51$  (with  $R^2$  coefficients of, resp.,  $0.966$  and  $9.969$ ). From these results, it appears that the convergence rate estimated in Theorem 6, although not optimal, can be used to estimate the marginal convergence gain to the original model when increasing  $v_{max}$ .

Finally, we investigate if the optimal convergence rate can be assimilated to a (negative) power of  $v_{max}$ . By analyzing the dependency of the logarithms of the empirical errors toward

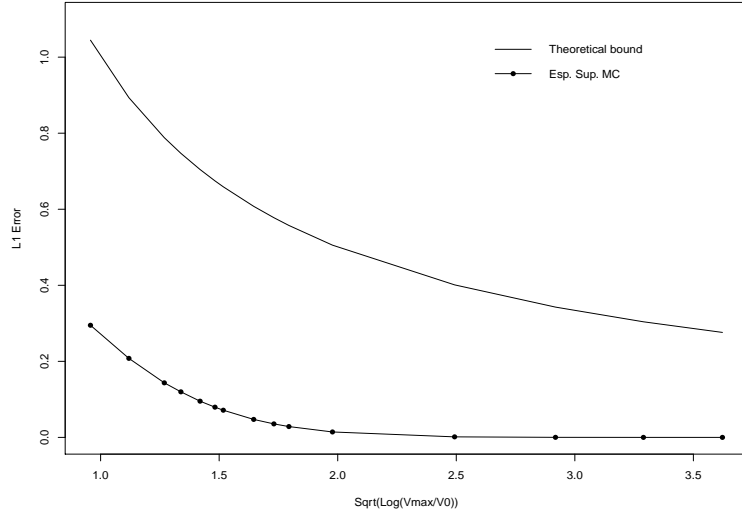


Figure 1:  $\mathbb{E}^S[\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|]$  as a function of  $\sqrt{\log(v_{max}/v_0)}$ . "Theoretical bound":  $1/\sqrt{\log(v_{max}/v_0)}$ ; "Esp. Sup. MC":  $\mathbb{E}^S[\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|]$  estimated by Monte-Carlo simulations.

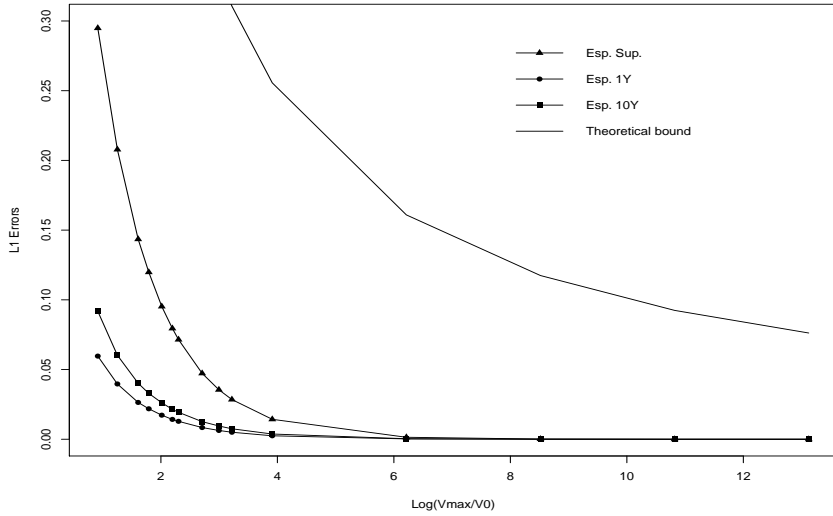


Figure 2:  $\mathbb{E}^S[|V_T^{J, v_{max}} - V_T^C|]$  as a function of  $\log(v_{max}/v_0)$ . "Theoretical bound":  $1/\log(v_{max}/v_0)$ ; Circles:  $\mathbb{E}^S[|V_1^{J, v_{max}} - V_1^C|]$ ; Squares:  $\mathbb{E}^S[|V_{10}^{J, v_{max}} - V_{10}^C|]$ ; Triangles:  $\mathbb{E}^S[\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|]$ .

$\log\left(\frac{v_{max}}{v_0}\right)$  in Figure 5, we observe now an (almost) perfect linear behaviour. Linear regression performed indicate a slope of around  $-1.0$  for all errors considered (respectively, in order of the mentioned errors in legend of Figure 5, slopes are  $-1.0084$ ,  $-1.0096$  and  $-1.0014$ ), with very high accuracy for each ( $R^2 \geq 0.999$ ). These numerical results lead to conjecture that the optimal rate would be such that  $\text{Error} \propto \frac{1}{v_{max}}$ .

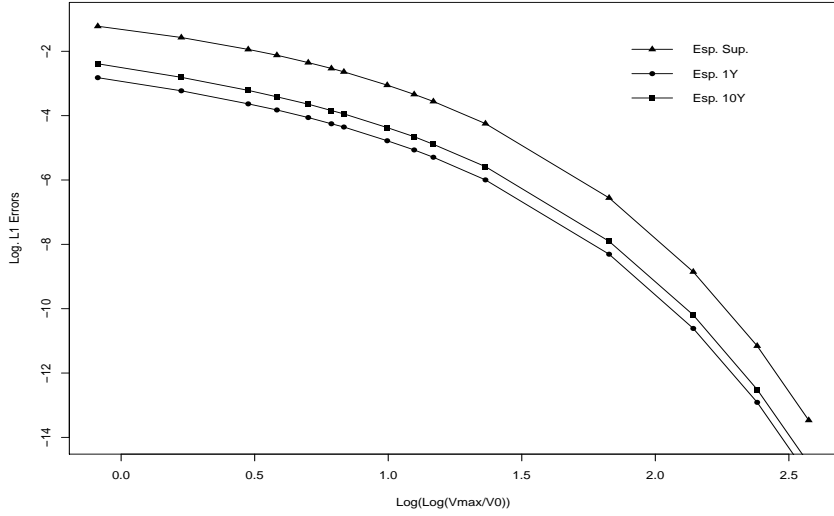


Figure 3:  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|] \right)$  and  $\log \left( \mathbb{E}^S [|V_T^{J, v_{max}} - V_T^C|] \right)$ ,  $T = 1$  and  $T = 10$ , as a function of  $\log(\log(v_{max}/v_0))$ . Circles:  $\log \mathbb{E}^S [|V_1^{J, v_{max}} - V_1^C|]$ ; Squares:  $\log \mathbb{E}^S [|V_{10}^{J, v_{max}} - V_{10}^C|]$ ; Triangles:  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|] \right)$ .

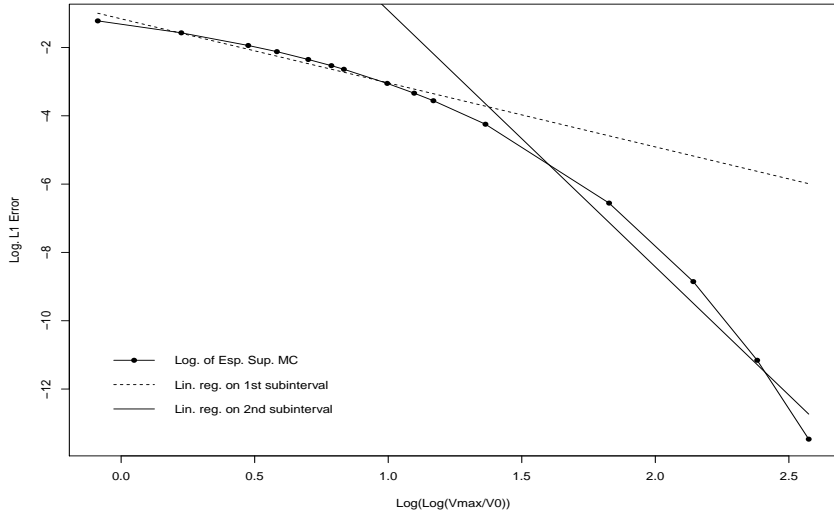


Figure 4:  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J, v_{max}} - V_s^C|] \right)$  as a function of  $\log(v_{max}/v_0)$  and fitted lines on each subinterval.

### Convergence of Gram-Charlier approximating series

We illustrate the approximating pricing of swaptions with the Gram-Charlier method under the model (12). A swaption is a call option on the swap rate: its discounted payoff writes  $\varphi(x) = B^S(0) \times (x - K)_+$  where  $K$  denotes the strike of the swaption (recall that  $B^S$  is annuity of the swap). In the following, parameters of the model (12) change from plot to plot. Reference prices are computed thanks to Monte-Carlo simulations and compared with approximating prices coming from series expansion. The volatility scheme has been introduced

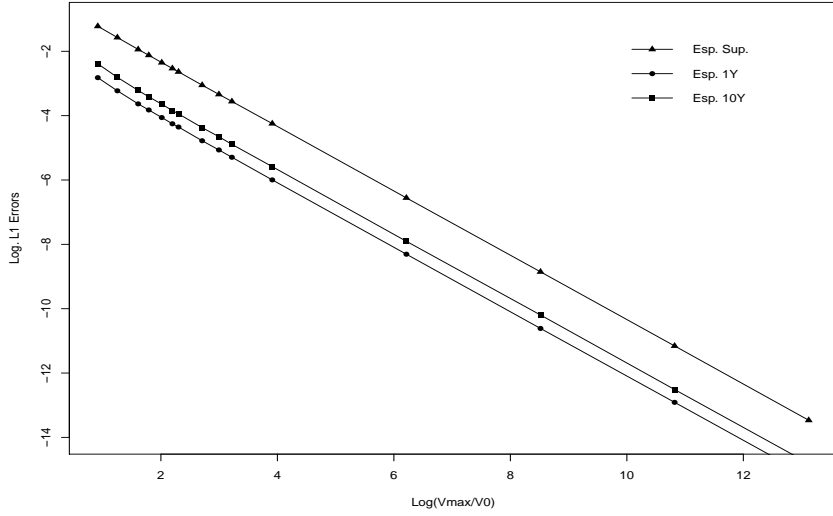


Figure 5:  $\log \left( \mathbb{E}^S [|V_T^{J,v_{max}} - V_T^C|] \right)$ ,  $T = 1$  and  $T = 10$ , and  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J,v_{max}} - V_s^C|] \right)$  as a function of  $\log(v_{max}/v_0)$ . Circles:  $\log \mathbb{E}^S [|V_1^{J,v_{max}} - V_1^C|]$ ; Squares:  $\log \mathbb{E}^S [|V_{10}^{J,v_{max}} - V_{10}^C|]$ ; Triangles:  $\log \left( \mathbb{E}^S [\sup_{0 \leq s \leq 5} |V_s^{J,v_{max}} - V_s^C|] \right)$ .

above in Equation (31); the numerical scheme used for the swap rate is

$$S_{t_{i+1}}^{m,n} = S_{t_i}^{m,n} + \sqrt{Q(\min(V_{t_i})_+, v_{max})\rho(t_i)\|\boldsymbol{\lambda}^{m,n}(t_i)\|} (W_{t_{i+1}} - W_{t_i}) \\ + \sqrt{\min((V_{t_i})_+, v_{max}) - \rho(t_i)^2 Q(\min(V_{t_i})_+, v_{max})\boldsymbol{\lambda}^{m,n}(t_i)} \cdot (W_{t_{i+1}}^{S,*} - W_{t_i}^{S,*}).$$

Except for Figure 8 in which a narrow confidence interval was required, Monte-Carlo prices were computed using  $10^5$  paths and a time-step of 0.001.

The choice of auxiliary density  $g_r$  used to build the space  $\mathcal{L}_r^2$  is now specified. As discussed in [21], [3] or [2], the parameters of the auxiliary density should be chosen so that a maximum number of moments of this auxiliary density match that of the unknown one. This is not always possible due to the number of constraints imposed to the model parameters, and notably (18) ensuring the convergence of the Gram-Charlier approximation. The value of  $\sigma_r^2$  is indicated for each experiment and whether it matches  $\text{Var}^S(S_T^{m,n})$  or not. For the parameter  $\mu_r$ , there is no condition on it and it is always chosen so that  $\mu_r = \mathbb{E}^S[S_T^{m,n}] = S_0^{m,n}$ .

### Example 1

We first chose parameters so that  $\sigma_r^2 = \text{Var}^S(S_T^{m,n}) = 8.726 \cdot 10^{-8}$  but condition (18) is not satisfied. We observe in Figure 6 the divergence of the Gram-Charlier expansion.

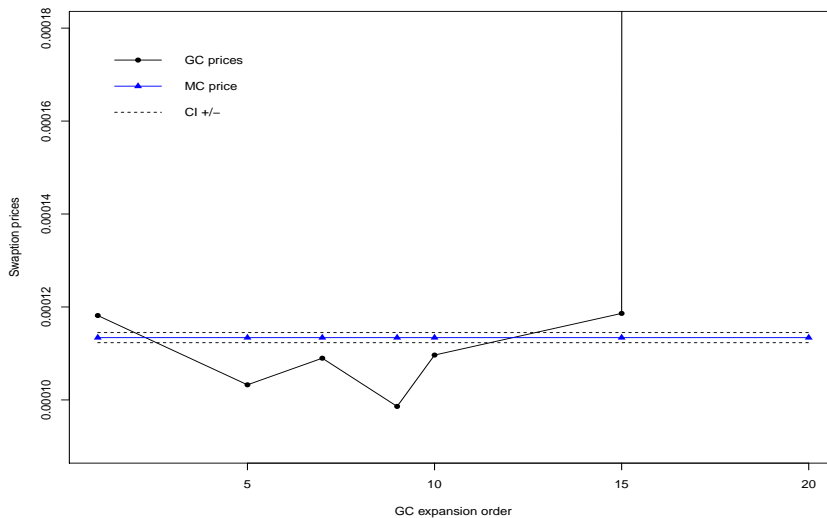


Figure 6: Divergence of Gram-Charlier expansion in case where condition (2) is not satisfied (auxiliary Gaussian of variance  $\sigma_r = \text{Var}^S(S_T^{m,n})$ ).

$$\begin{aligned}
 s_0^{m,n} &= 0, v_0 = 0.025, v_{max} = 0.5, \\
 a = b = c &= 0, d = \kappa = 1, \theta = 0.25, \epsilon = 0.6, \\
 \rho = \delta &= 0, K = 0, \sigma_r = 0.000295403. \\
 \text{Tenor of the swaption: } &T = T_m = 1 \text{ and } T_n - T_m = 1.
 \end{aligned}$$

**Example 2**

The parameter  $\sigma_r$  is now slightly modified: we chose to take  $\sigma_r = \sqrt{\text{Var}^S(S_T^{m,n})} + 10^{-4}$  so that the sufficient condition (18) is now satisfied. We observe in Figure 7 that the Gram-Charlier expansion is stable and now converge. In [3], the authors observed, in their framework, that the approximating implied volatilities are within 10 basis points of their reference implied volatilities from expansion order  $N = 10$ . In [?], the authors proved the numerical efficiency for calibration purposes of approximating swaptions prices using only skewness and kurtosis adjustments (N=4). We see here that the approximating error may be significantly reduced by increasing the expansion order, provided that convergence has been theoretically established here.

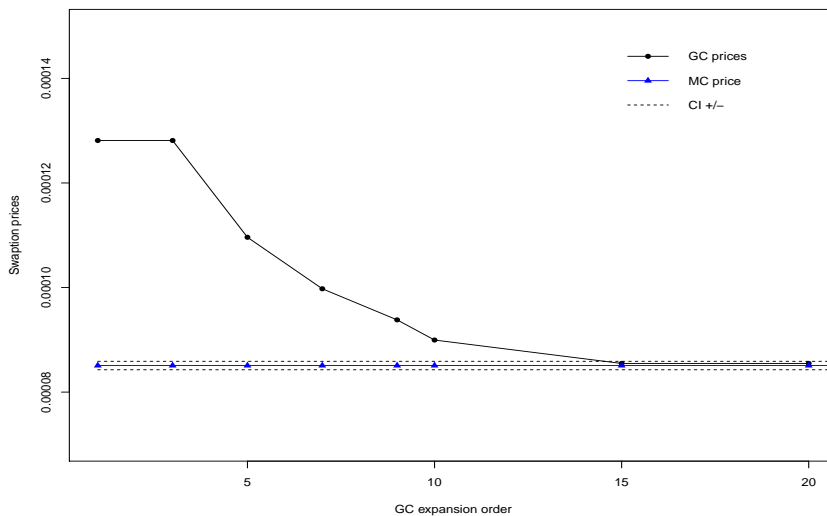


Figure 7: Gram-Charlier convergence with auxiliary Gaussian of variance

$$\sigma_r^2 \neq \text{Var}^S(S_T^{m,n}).$$

$$s_0^{m,n} = 0, v_0 = 0.025, v_{max} = 0.25, a = b = c = 0,$$

$$d = \kappa = 1, \theta = 0.12, \epsilon = 0.4, \rho = \delta = 0.0$$

$$\sigma_r = 3.203183 \cdot 10^{-4}, K = 0.$$

Maturity and tenor of the swaption:  $T = T_m = 1$  and  $T_n - T_m = 1$ .

### Example 3

The parameters are selected such that the first two moments of approximated density are matched *and* condition (18) is satisfied. From the first orders in the Gram-Charlier expansion, the approximated prices are really close to the reference one. Moreover, Gram-Charlier series is very stable: it remains in the confidence interval as the expansion order increases. This is illustrated in Figure 8. In this particular case, the convergence is excellent and the kurtosis/skewness adjustments proposed in [?] are satisfactory. However, this is a very special case, hard to establish. Indeed, condition (18) prompts us to chose a small  $v_{max}$ . But a small  $v_{max}$  leads to a small  $\text{Var}^S(S_T^{m,n})$ . Then, choosing  $\sigma_r^2 = \text{Var}^S(S_T^{m,n})$  while satisfying (18) is unlikely. We are thus lead to boost the volatility in the model, for instance by reducing  $\kappa$  or increasing  $\epsilon$ . However, Feller condition restricts the choice of acceptable parameters. Thus, parameters that satisfy all required constraints while allowing to match moments of the unknown density must be carefully chosen, if possible.

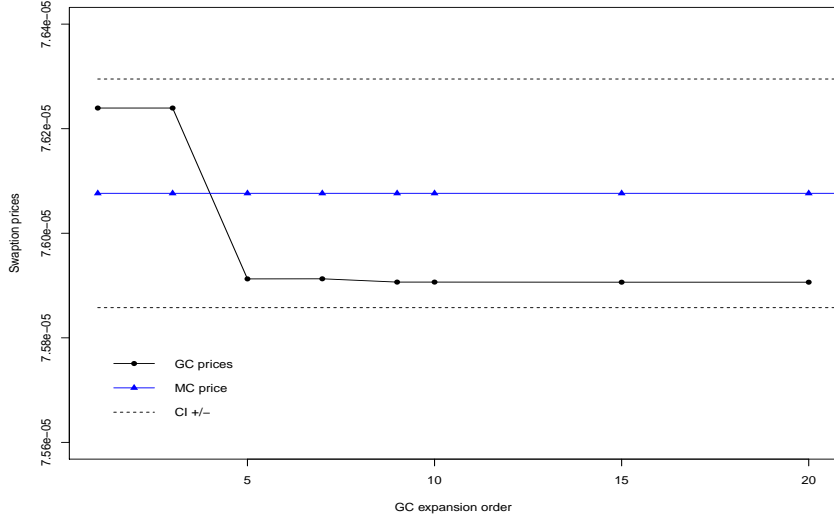


Figure 8: Gram-Charlier convergence with auxiliary Gaussian of variance

$$\sigma_r^2 = \text{Var}^S(S_T^{m,n}).$$

$$s_0^{m,n} = 0, v_0 = 0.025, v_{max} = 0.089, a = b = c = 0, d = 1,$$

$$\kappa = 2, \theta = 0.06, \epsilon = 0.1, \rho = \delta = 0.0,$$

$$\sigma_r = 1.906037 \cdot 10^{-4}, K = 0.$$

Maturity and tenor of the swaption:  $T = T_m = 1$  and  $T_n - T_m = 1$ .

#### Example 4

On the contrary, the more the parameter  $\sigma_r$  is chosen "far" from  $\text{Var}^S(S_T^{m,n})$ , the more the Gram-Charlier series will converge slowly (and sometimes may diverge). This is illustrated in Figure 9. This case is more sophisticated since all the parameters are non-zero and  $T_n - T_m > 1$ . In particular, the correlation structure between swap rate and its volatility plays a notable role. The maturity of the swaption is  $T_m = 2$ . The mean of the auxiliary Gaussian density is no longer zero, and  $\sigma_r = \sqrt{\text{Var}^S(S_T^{m,n}) + 0.002}$  (note that in this case,  $\frac{0.002}{\sqrt{\text{Var}^S(S_T^{m,n})}} \approx 0.63$ ). We observe that the approximating prices do not reach the confidence interval by the expansion order  $N = 50$ .



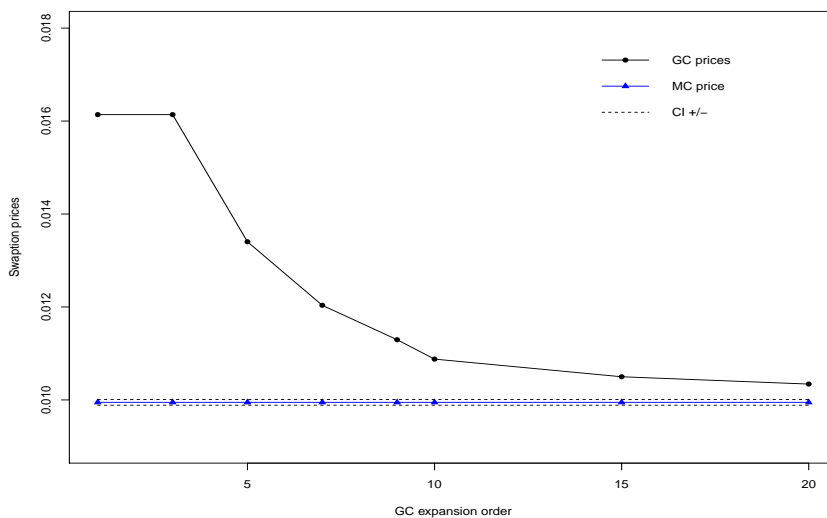


Figure 9: Gram-Charlier convergence with auxiliary Gaussian of variance

$$\sigma_r^2 \neq \text{Var}^S(S_T^{m,n}).$$

$$s_0^{m,n} = 0.008982445, v_0 = 0.025, v_{max} = 0.089, a = 10^{-4}, b = 10^{-1},$$

$$c = 2.5, d = 10^{-1}, \kappa = 1.5, \theta = 0.06, \epsilon = 0.13,$$

$$\rho = 0.4, \text{ and } \delta = 0.1. \sigma_r = 0.005183955. K = S_0^{m,n}.$$

Maturity and tenor of the swaption:  $T = T_m = 2$  and  $T_n - T_m = 8$ .

## Concluding comments

This work is based on and extends two works: [3] motivated the use of Jacobi dynamics as approximation of the famous Cox-Ingersoll-Ross one used to build the DDSVLMM while [?] proved the numerical efficiency of the use of Gram-Charlier expansions for pricing under the DDSVLMM. In this paper, we first discuss the fact that Gram-Charlier technique is not justified in most of stochastic volatility models. Introducing the Jacobi process in the standard DDSVLMM framework after some approximations allows to circumvent this difficulty. A strong convergence of the proposed dynamic towards the standard one along with a convergence rate justify the use of this *approximating* model to price some swap rate derivatives. Numerical investigations suggest that this rate could be theoretically improved, supported by the convergence rate obtained in a weak sense; it is left for further research.

Under the suggested model pricing is done by means of Gram-Charlier expansion. The polynomial property of the proposed dynamics offers (semi) analytical tractability of the marginal moments computations of the modelled underlying. As a result, it fully justifies the use of such density approximation techniques under some parametric conditions. Numerical experiments provide satisfactory results as Gram-Charlier prices do converge to empirical ones rather quickly. By carefully choosing the auxiliary density used to perform the density approximation, the convergence can be hastened: as of the very first order expansion, Gram-Charlier prices are very close to the target. These satisfactory results pave the way to the calibration of the DDSVLMM using expansions techniques as in [?] with ensured convergence here.

## References

- [1] L. Andersen and J. Andreasen. *Volatility skews and extensions of the Libor market model*. Applied Mathematical Finance, Taylor & Francis, Vol.7(1), pp. 1-32, 2000. [2](#)
- [2] D. Ackerer and D. Filipović. *Option pricing with orthogonal polynomial expansions*. Mathematical Finance, Wiley Online Library, 2019. [28](#)
- [3] D. Ackerer, D. Filipović and S. Pulido. *The Jacobi stochastic volatility model*. Finance and Stochastics, Springer, Vol.22(3), pp.667-700, 2018. [2](#), [3](#), [9](#), [11](#), [12](#), [13](#), [14](#), [16](#), [28](#), [29](#), [32](#)
- [4] S. Asmussen, P.-O. Goffard and P. J. Laub. *Orthonormal polynomial expansions and log-normal sum densities*. arXiv preprint arXiv:1601.01763, 2016. [3](#)
- [5] A. Alfonsi. *Credit Risk modelling. Calibration and discretization of financial models*. PhD thesis, Ecole des Ponts ParisTech, 2006. [18](#)
- [6] A. Alfonsi. *Affine Diffusions and Related Processes: Simulation, Theory and Applications*. Springer, Vol.6, 2015. [20](#), [25](#)
- [7] H. Albrecher, P. Mayer, W. Schoutens and J. Tistaert. *The little Heston trap*. Wilmott, No.1, pp.83-92, 2007. [2](#)
- [8] A. Brace, D. G atarek and M. Musiela. *The market model of interest rate dynamics*. Mathematical finance, Wiley Online Library, Vol.7(2), pp. 127-155, 1997. [1](#), [23](#)
- [9] E. Benhamou, E. Gobet, and M. Miri. *Time dependent Heston model*. SIAM Journal on Financial Mathematics, SIAM, Vol.1(1), pp.289–325, 2010. [10](#)
- [10] D. Brigo and F. Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007. [2](#), [10](#), [23](#)
- [11] J.-P. D. Chateau. *Valuing European Put Options under Skewness and Increasing (Excess) Kurtosis*. Journal of Mathematical Finance, Scientific Research Publishing, Vol.4(03), p. 160, 2014. [5](#)
- [12] C. Cuchiero, M. Keller-Ressel and J. Teichmann. *Polynomial processes and their applications to mathematical finance*. Finance and Stochastics, Springer, Vol.16(4), pp. 711-740, 2012. [2](#), [7](#), [11](#)
- [13] P. Carr and D. Madan. *Option valuation using the fast Fourier transform*. Journal of computational finance, Vol.2(4), pp. 61–73, 1999. [2](#)
- [14] C.J. Corrado and T. Su. *Skewness and kurtosis in S&P 500 index returns implied by option prices*. Journal of Financial research, Wiley Online Library, Vol.19(2), pp. 175–192, 1996. [5](#)
- [15] L. Devineau, P.-E. Arrouy, P. Bonnefoy and A. Boumezoued. *Fast calibration of the Libor Market Model with Stochastic Volatility and Displaced Diffusion*. Journal of Industrial and Management Optimization, Vol.13, p.1, 2017. [2](#), [5](#)

- [16] M. F. del Carmen Agoitia Hurtado. *Time-inhomogeneous polynomial processes in electricity spot price models*. PhD thesis, 20187. [7](#)
- [17] F. Delbaen and H. Shirakawa. *An interest rate model with upper and lower bounds*. Asia-Pacific Financial Markets, Springer, Vol.9(3-4), pp. 191-209, 2002. [2](#)
- [18] B. Eriksson and M. Pistorius. *Method of moments approach to pricing double barrier contracts in polynomial jump-diffusion models*. International Journal of Theoretical and Applied Finance, World Scientific, Vol.14(07), pp.1139-1158, 2011. [2](#)
- [19] W. Feller. *An Introduction to Probability Theory and Its Applications: Volume 1.* J. Wiley and sons, Vol.1, 1960. [4](#)
- [20] D. Filipović and M. Larsson. *Polynomial diffusions and applications in finance*. Finance and Stochastics, Springer, Vol.20(4), pp.931-972, 2016. [2](#), [7](#), [8](#), [11](#)
- [21] D. Filipović, E. Mayerhofer and P. Schneider. *Density approximations for multivariate affine jump-diffusion processes*. Journal of Econometrics, Elsevier, Vol.176(2), pp.93-111, 2013. [2](#), [3](#), [4](#), [28](#)
- [22] S. L. Heston. *A closed-form solution for options with stochastic volatility with applications to bond and currency options*. The review of financial studies, Oxford University Press, Vol.6(2), pp.327-343, 1993. [2](#)
- [23] S. L. Heston and A. G. Rossi. *A spanning series approach to options*. Review of Asset Pricing Studies, Oxford University Press, 2016. [5](#)
- [24] F. Jamshidian. *LIBOR and swap market models and measures*. Finance and Stochastics, Springer, Vol.4(1), pp.293-330, 1997. [1](#), [10](#)
- [25] M. Joshi and R. Rebonato. *A stochastic-volatility, displaced-diffusion extension of the LIBOR market model*. Quantitative Finance, Vol.3(6), pp.458-469, 2003. [1](#), [10](#)
- [26] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Springer, 1991. [9](#)
- [27] S. Karlin and H. E. Taylor. *A second course in stochastic processes*. Elsevier, 1981. [2](#)
- [28] R. Lord, R. Koekkoek and D. Van Dijk. *A comparison of biased simulation schemes for stochastic volatility models*. Quantitative Finance, Taylor & Francis, Vol.10(2), pp.177-194, 2010. [25](#)
- [29] J. Ma. *Pricing Foreign Equity Options with Stochastic Correlation and Volatility.* Annals of Economics & Finance, Vol.10(2), 2009. [2](#)
- [30] M. Musiela and M. Rutkowski. *Continuous-time term structure models: Forward measure approach*. Finance and Stochastics, Springer, Vol.1(4), pp.261-291, 1997. [1](#)
- [31] M. Potters, R. Cont and J.-P. Bouchaud. *Financial markets as adaptive systems*. EPL (Europhysics Letters), IOP Publishing, Vol.41(3), p.239, 1998. [5](#)

- [32] V. Piterbarg. *A stochastic volatility forward Libor model with a term structure of volatility smiles*. Available at SSRN 472061, 2003. [2](#)
- [33] E. Schlögl. *Option pricing where the underlying assets follow a Gram/Charlier density of arbitrary order*. Journal of Economic Dynamics and Control, Elsevier, Vol.37(3), pp. 611-632, 2013. [5](#)
- [34] D. Talay and L. Tubaro. *Expansion of the global error for numerical schemes solving stochastic differential equations*. Stochastic analysis and applications, Taylor & Francis, Vol.8(4), pp.483-509, 1990. [21](#)
- [35] L. Wu and F. Zhang. *LIBOR market model with stochastic volatility*. Journal of industrial and management optimization, American Institute of Mathematical Sciences, Vol.2(2), p.199, 2006. [1](#), [2](#), [10](#), [23](#)