On the choice of the numerical scheme for the simulation of the Heston model

1 Options under the discretized the Heston model

In this section, we consider testing options under the discretized Heston model [6, 4, 8, 2] whose dynamics are given by

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW_t^S = \mu S_t dt + \rho\sqrt{v_t} S_t dW_t^v + \sqrt{1-\rho^2}\sqrt{v_t} S_t dW_t \\ dv_t &= \kappa(\theta - v_t)dt + \xi\sqrt{v_t} dW_t^v, \end{aligned} \tag{1.1}$$

where

- S_t is the price of the asset, v_t the instantaneous variance, given as a CIR process.
- W_t^S, W_t^v are correlated Wiener processes with correlation ρ.
- μ is the rate of return of the asset.
- θ is the mean variance.
- κ is the rate at which v_t reverts to θ.
- ξ is the volatility of the volatility, and determines the variance of v_t.

1.1 Schemes to simulate the Heston dynamics

Given that the SDE for the asset path is now dependent upon the solution of the second volatility SDE in (1.1), it is necessary to simulate the volatility process first and then utilize this "volatility path" in order to simulate the asset path. In the case of the original GBM SDE it is possible to use Itô's Lemma to directly solve for S_t. However, we are unable to utilize that procedure here and must use a numerical approximation in order to obtain both paths. In the following, we try to give an overview of the methods that we can use in our context. These methods mainly differs by the way it simulates the volatility process to ensure its positiveness.

1.1.1 Fixed Euler scheme

In this context, one can use Forward Euler scheme to simulate the Heston model, and to avoid problems with negative values of the volatility process v_t, many fixes were introduced in the literature (see [9]). We introduce f_1, f_2, and f_3, which with different choices implies different schemes, and apply Forward Euler scheme to discretize 1.1 results in

$$\begin{aligned}
\widehat{S}_{t+\Delta t} &= \widehat{S}_t + \mu \widehat{S}_t \Delta t + \sqrt{\widehat{V}_t \Delta t}\widehat{S}_t Z_s \\
\widehat{V}_{t+\Delta t} &= f_1(\widehat{V}_t) + \kappa(\theta - f_2(\widehat{V}_t))\Delta t + \xi\sqrt{f_3(\widehat{V}_t)\Delta t} Z_V \\
\widehat{V}_{t+\Delta t} &= f_3(\widehat{V}_{t+\Delta t}),
\end{aligned} \tag{1.2}$$

where Z_s and Z_V are two correlated standard normal variables with correlation ρ, such that $Z_s = \rho Z_V + \sqrt{1-\rho^2} Z_t$.

Scheme	f_1	f_2	f_3
full truncation scheme	$\widehat{V}_t$	$\widehat{V}_t^+$	$\widehat{V}_t^+$
Partial truncation scheme	$\widehat{V}_t$	$\widehat{V}_t$	$\widehat{V}_t^+$
The reflection scheme	$\left\lvert\widehat{V}_t\right\rvert$	$\left\lvert\widehat{V}_t\right\rvert$	$\left\lvert\widehat{V}_t\right\rvert$

where $\widehat{V}_t^+ = \max(0, \widehat{V}_t)$.

[9] suggest that the Full Truncation method is the optimal option in terms of weak error convergence and robustness with respect to the parameters of the model. Therefore, we use this scheme for this approach.

1.2 Euler schemes with moment matching

We consider two moment matching Euler schemes that were suggested by Andersen and Brotherton-Ratcliffe [3] (we call it ABR scheme) and by Anderson in [2] (we choose the QE method that was reported to have the optimal results).

1.2.1 The ABR method

The ABR method is suggested in [3], where the variance v is assumed to be locally lognormal, where the parameters are determined such that the first two moments of the discretization coincide with the theoretical moments, that is

$$\begin{aligned}
\widehat{V}(t+\Delta t) &= \left(e^{-\kappa \Delta t}\widehat{V}(t) + \left(1 - e^{-\kappa\Delta t}\right)\theta\right) e^{-\frac{1}{2}\Gamma(t)^2\Delta t + \Gamma(t)\Delta W_v(t)} \\
\Gamma(t)^2 &= \Delta t^{-1} \log\left(1 + \frac{\frac{1}{2}\xi^2\kappa^{-1}\widehat{V}(t)(1-e^{-2\kappa\Delta t})}{\left(e^{-\kappa\Delta t}\widehat{V}(t) + (1-e^{-\kappa\Delta t})\theta\right)^2}\right)
\end{aligned} \tag{1.3}$$

As reported in [9], the scheme, being very easy to implement, is much more effective than many of the Euler fixes, however, it was reported that it has a bad weak error behavior that is not robust with parameters values, which is something that we need to check in our setting!

1.2.2 The QE method

On the other hand, and using the same idea of moment matching, Anderson in[2], suggested more discretisations for the Heston model similar to (1.3) and taking the shape of the Heston density function into account. These schemes have negligible bias, at the cost of a more complex implementation than the one introduced by [3]. As reported by [2], the QE method is the optimal scheme to use and it follows the following algorithm.

1. Select some arbitrary level $\Psi_c \in [1, 2]$.
2. Given $\widehat{V}(t)$, compute m and s^2 given by

$$m = \mathrm{E}\left[V(t+\Delta t) \mid V(t) = \widehat{V}(t)\right] = \theta + (\widehat{v}(t) - \theta)e^{-\kappa\Delta t}$$
$$s^2 = \mathrm{Var}\left[V(t+\Delta t) \mid V(t) = \widehat{V}(t)\right] = \frac{\widehat{V}(t)\xi^2 e^{-\kappa\Delta t}}{\kappa}(1-e^{-\kappa\Delta t}) + \frac{\theta\xi^2}{2\kappa}(1-e^{-\kappa\Delta t})^2.$$

3. Compute $\Psi = s^2/m^2$.
4. Draw a uniform random number U_V.
5. If $\Psi \leq \Psi_c$

   i) Compute a and b given by

$$b^2 = 2\Psi^{-1} - 1 + \sqrt{2\Psi^{-1}}\sqrt{2\Psi^{-1} - 1}$$
$$a = \frac{m}{1+b^2}$$

   ii) Compute $Z_V = \Phi^{-1}(U_V)$, where Φ^{-1} is the inverse cumulative Gaussian distribution function.

   iii) Set $\widehat{V}(t+\Delta t) = a(b + Z_V^2)^2$.

6. Otherwise, if $\Psi > \Psi_c$

   i) Compute p and β according to

$$p = \frac{\Psi - 1}{\Psi - 1}$$
$$\beta = \frac{1-p}{m}$$

   ii) Set $\widehat{V}(t+\Delta t) = h^{-1}(U_V; p, \beta)$, where h^{-1} is given by

$$h^{-1}(u; p, \beta) = \begin{cases} 0, & 0 \leq u \leq p \\ \beta^{-1}\log\left(\frac{1-p}{1-u}\right), & p < u \leq 1 \end{cases}$$

Compared to ABR scheme, the QE scheme may have a better weak error behavior with the disadvantage of being more costly. Furthermore, the QE algorithm uses two different distributions to model the volatility depending on the initial value of the volatility!

1.3 Kahl-Jackel Scheme

[8] suggested discretizing the volatility process using an implicit Milstein scheme, coupled with their IJK discretization for the stock process. Specifically, they propose the scheme

$$\log(\widehat{X}(t+\Delta t)) = \log(\widehat{X}(t)) - \frac{\Delta t}{4}\left(\widehat{V}(t+\Delta t)+\widehat{V}(t)\right) + \rho\sqrt{\widehat{V}(t)}Z_V\sqrt{\Delta t}$$
$$+\frac{1}{2}\left(\sqrt{\widehat{V}(t+\Delta t)}+\sqrt{\widehat{V}(t)}\right)\left(Z_x\sqrt{\Delta t}-\rho Z_V\sqrt{\Delta t}\right)+\frac{\xi}{2}\rho\Delta t(Z_V^2-1)$$

$$\widehat{V}(t+\Delta t) = \frac{\widehat{V}(t)+\theta\kappa\Delta t+\xi\sqrt{\widehat{V}(t)}Z_V\sqrt{\Delta t}+\frac{\xi^2}{4}\Delta t(Z_V^2-1)}{1+\kappa\Delta t} \tag{1.4}$$

It is easy to check that this discretization scheme will result in positive paths for the V process if $4\kappa\theta > \xi^2$ (which is rarely satisfied in practice). Unfortunately [8] does not provide a solution for this problem, but it seems reasonable to use a truncation scheme similar to those presented in Section 1.1.1. Therefore, we do not test this scheme in our numerical experiments since we think that it will have similar issues that we observed for the full truncation scheme presented in Section 1.1.1. Furthermore, it was shown that the convergence of the weak rate of the IJK-IMM scheme somewhat erratic, especially for cases $4\kappa\theta \leq \xi^2$.

1.4 Discretization of Heston model with the volatility process Simulated using the sum of Ornstein-Uhlenbeck (Bessel) processes

Since we are using ASGQ which is very sensitive to the smoothness of the integrand. We found numerically (see section 1.5) that using a non smooth transformation to ensure the positiveness of the volatility process deteriorates the performance of the ASGQ. An alternative way to guarantee the positiveness of the volatility process is to simulate it as the sum of Ornstein-Uhlenbeck (Bessel) processes.

In fact, it is well known that any Ornstein-Uhlenbeck (OU) process is normally distributed. Thus, the sum of n squared OU processes is chi-squared distributed with n degrees of freedom, where $n \in \mathbb{N}$. Let us define $\mathbf{X}$ to be a n-dimensional vector valued OU process with

$$\mathrm{d}X_t^i = \alpha X_t^i \mathrm{d}t + \beta \mathrm{d}W_t^i, \tag{1.5}$$

where $\mathbf{W}$ is a n-dimensional vector of independent Brownian motions.

We define also the process Y_t as

$$Y_t = \sum_{i=1}^{n}\left(X_t^i\right)^2. \tag{1.6}$$

Then, using the fact that

$$\begin{aligned}\mathrm{d}\left(X_t^i\right)^2 &= 2X_t^i\mathrm{d}X_t^i + 2\mathrm{d}\langle X^i\rangle_t \\ &= \left(2\alpha\left(X_t^i\right)^2+\beta^2\right)\mathrm{d}t + 2\beta X_t^i\mathrm{d}W_t^i,\end{aligned} \tag{1.7}$$

we can write

$$
\begin{aligned}
\mathrm{d}Y_t &= \mathrm{d}\left(\sum_{i=1}^{n}\left(X_t^i\right)^2\right) \\
&= \sum_{i=1}^{n}\mathrm{d}\left(X_t^i\right)^2 \\
&= \left(2\alpha Y_t + n\beta^2\right)\mathrm{d}t + 2\beta\sum_{i=1}^{n}X_t^i \mathrm{d}W_t^i, \qquad (1.8)
\end{aligned}
$$

where the second step follows from the independence of the Brownian motions. Furthermore, we note that the process

$$
Z_t = \int_0^t \sum_{i=1}^{n} X_u^i \mathrm{d}W_u^i \qquad (1.9)
$$

is a martingale with quadratic variation

$$
\begin{aligned}
\langle Z\rangle_t &= \int_0^t \sum_{i=1}^{n}\left(X_u^i\right)^2 \mathrm{d}u \\
&= \int_0^t Y_u \mathrm{d}u. \qquad (1.10)
\end{aligned}
$$

Consequently, by Lévy's characterization theorem, we can show that the process

$$
\widetilde{W}_t = \int_0^t \frac{1}{\sqrt{Y_u}} \sum_{i=1}^{n} X_u^i \mathrm{d}W_u^i \qquad (1.11)
$$

is a Brownian motion.

Finally, we have

$$
\begin{aligned}
\mathrm{d}Y_t &= \left(2\alpha Y_t + n\beta^2\right)\mathrm{d}t + 2\beta\sqrt{Y_t}\mathrm{d}\tilde{W}_t \\
&= \kappa\left(\theta - Y_t\right)\mathrm{d}t + \xi\sqrt{Y_t}\mathrm{d}W_t, \qquad (1.12)
\end{aligned}
$$

where $\kappa = -2\alpha$, $\theta = -n\beta^2/2\alpha$ and $\xi = 2\beta$.

Equations (1.5), (1.8), and (1.12) show in order to simulate the process Y_t given by (1.12), we can simulate the OU process $\mathbf{X}$, with dynamics given by (1.5) such that its parameters (α, β) are expressed in terms of those of the process Y_t, that is

$$
\alpha = -\frac{\kappa}{2}, \quad \beta = \frac{\xi}{2}, \quad n = \frac{-2\theta\alpha}{\beta^2} = \frac{4\theta\kappa}{\xi^2}.
$$

Therefore, we conclude that we can simulate the volatility of the Heston model using a sum of OU processes.

Remark 1.1. The previous derivation can can be generalized to cases where n is not an integer by considering a time-change of a squared Bessel process (see Chapter 6 in [7] for details).

1.5 On the choice of the simulation scheme of the Heston model

In this section, we try to determine the optimal scheme for simulating the Heston model defined in 1.1. In the literature [2, 9, 1], more focus was on designing schemes that i) ensures the positiveness of the volatility process and ii) has a good weak error behavior. In our setting, an optimal scheme is defined through two properties: i) the behavior of mixed rates which is an important feature for an optimal performance of ASGQ, and ii) the behavior of the weak error in order to apply Richardson extrapolation when it is needed.

1.5.1 Comparison in terms of mixed differences rates

In this section, we try to compare the four approaches, discussed in Section 1.1, which are : i) the full truncation scheme discussed in Section 1.1.1, ii) the ABR scheme discussed in Section 1.2.1, ii) the QE scheme discussed in Section 1.2.2, and iv) the smooth scheme bases on sum of square of OU processes discussed in Section 1.4. The comparison is done in terms of mixed differences convergence (defined in (1.13)) rates which, as pointed out in [5], are a key determinant for the optimal performance of ASGQ.

$$\Delta E_{\boldsymbol{\beta}} = \left| Q_N^{\mathcal{I}\cup\{\boldsymbol{\beta}\}} - Q_N^{\mathcal{I}} \right|. \tag{1.13}$$

For comparison, we use basically four sets of examples given in table 1.1. The first two sets of parameters we used we test two different cases where the first case corresponds to $n = 1$ and the second one corresponds to $n = 2$. On the other hand, for the two remaining sets we used similar parameters tested in [9], where the set 3 stems from Broadie and Kaya [4], and is the hardest of all examples. The example of set 4 parameters stems from Andersen [2], where it is used to represent the market for long-dated FX options. We also point out that only the second set of parameters satisfies the Feller condition, that is $4\kappa\theta > \xi^2$.

Parameters	Reference solution
Set 1: $S_0 = K = 100$, $v_0 = 0.04$, $\mu = 0$, $\rho = -0.9$, $\kappa = 1$, $\xi = 0.1$, $\theta = \frac{\xi^2}{4\kappa} = 0.0025$, $(n = 1)$.	6.332542
Set 2: $S_0 = K = 100$, $v_0 = 0.04$, $\mu = 0$, $\rho = -0.9$, $\kappa = 1$, $\xi = 0.1$, $\theta = \frac{\xi^2}{4\kappa} = 0.005$, $(n = 2)$.	6.445535
Set 3: $S_0 = K = 100$, $v_0 = 0.09$, $\mu = 0$, $\rho = -0.3$, $\kappa = 2.7778$, $\xi = 1$, $\theta = \frac{\xi^2}{4\kappa} = 0.09$, $(n = 1)$.	10.86117
Set 4: $S_0 = K = 100$, $v_0 = 0.04$, $\mu = 0$, $\rho = -0.9$, $\kappa = 0.5$, $\xi = 1$, $\theta = 0.04$.	4.403384

Table 1.1: Reference solution using Premia with cf_call_heston method as explained in [6], for different parameter constellations. By n we refer to the number of OU processes for simulating the volatility process in approach given by Section 1.4

In our numerical experiments, we only observed differences of mixed differences rates related to volatility coordinates, which is expected since we use basically schemes that only differ by the way it simulates the volatility process. Figures 1.1, 1.2, 1.3 and 1.4 show a comparison of first differences rates related to volatility coordinates for the different schemes and for the different examples. From this Figures, we have the following conclusions

1. In all examples, the full truncation scheme is the worst scheme to use in terms of first differences rates.

2. The scheme based on the sum of square of OU processes have a good performance for the first two sets. However, surprisingly, we observed a very bad performance for the example of Set 3 parameters!

3. Both schemes based on moment matching (ABR and QE schemes) have the best performance in terms of mixed rates convergence and also they show robust behavior with respect to all examples with almost similar performance

Mixed differences for the case of Set 1 parameters

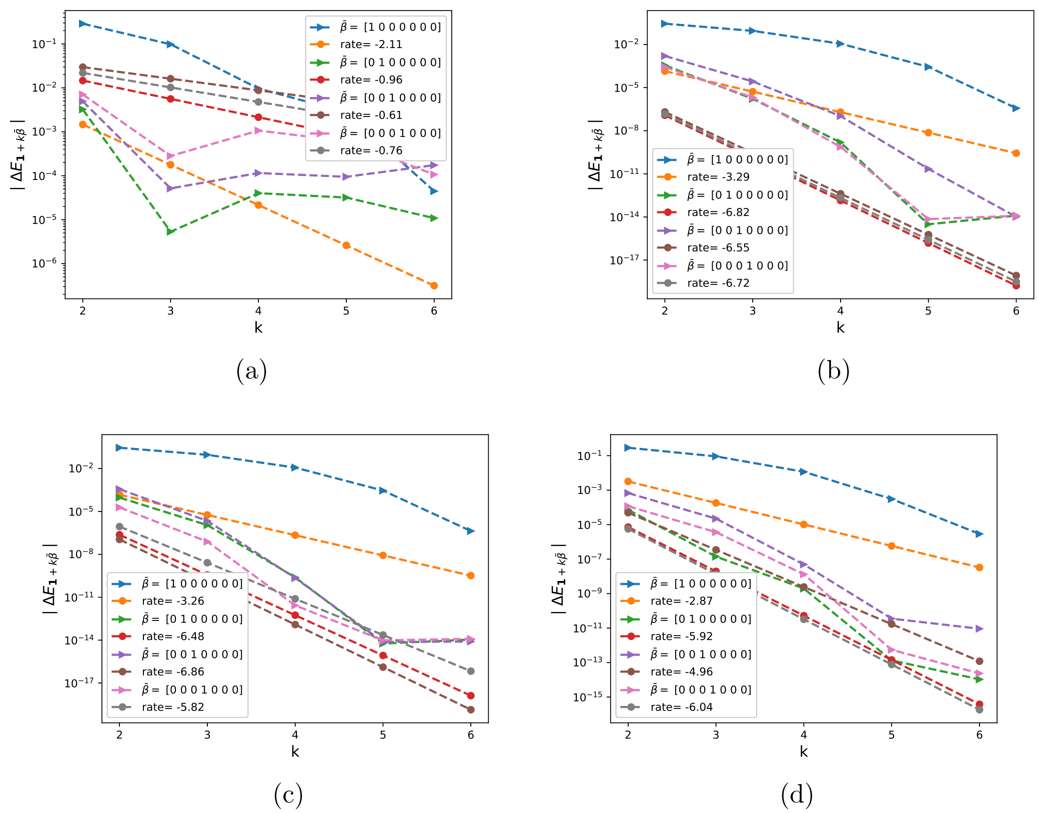
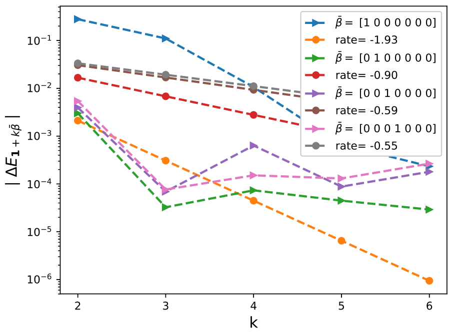
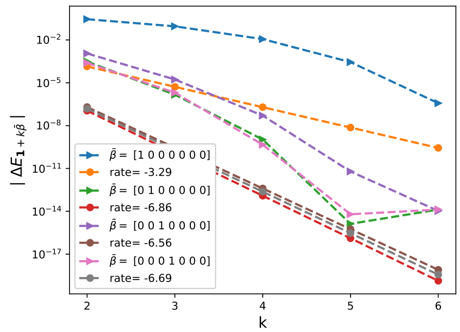


Figure 1.1: The rate of error convergence of first order differences $|\Delta \mathrm{E}_{\boldsymbol{\beta}}|$, defined by (1.13), ($\boldsymbol{\beta} = \mathbf{1} + k\overline{\boldsymbol{\beta}}$) for the example of single call option under Heston model, with parameters given by Set 1 in Table 1.1, using $N = 4$ time steps. In this case, we just show the first 4 dimensions which are used for the volatility noise (mainly dW_v in (1.1)). (a) using full truncation as in Section 1.1.1, (b) using the ABR scheme as in Section 1.2.1, (c) using the QE scheme as in Section 1.2.2, (d) using the smooth transformation as in Section 1.4.

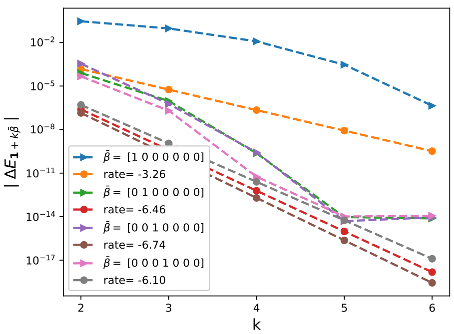
Mixed differences for the case of Set 2 parameters



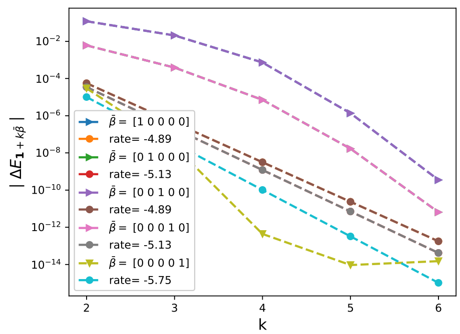
(a)



(b)



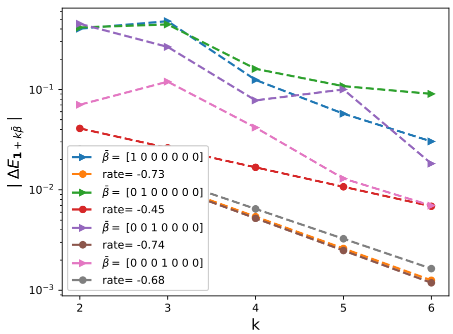
(c)



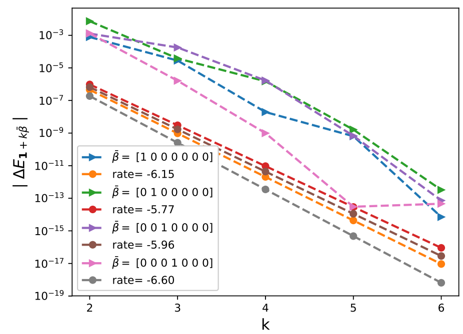
(d)

Figure 1.2: The rate of error convergence of first order differences $|\Delta \mathrm{E}_{\boldsymbol{\beta}}|$, defined by (1.13), ($\boldsymbol{\beta} = \mathbf{1} + k\overline{\boldsymbol{\beta}}$) for the example of single call option under Heston model, with parameters given by Set 2 in Table 1.1, using $N = 4$ time steps. In this case, we just show the first 4 dimensions which are used for the volatility noise (mainly dW_v in (1.1)). (a) using full truncation as in Section 1.1.1, (b) using the ABR scheme as in Section 1.2.1, (c) using the QE scheme as in Section 1.2.2, (d) using the smooth transformation as in Section 1.4, here $N = 2$.

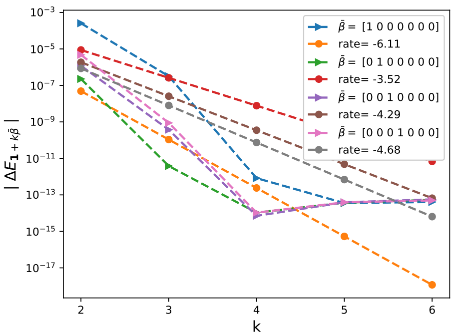
Mixed differences for the case of Set 3 parameters



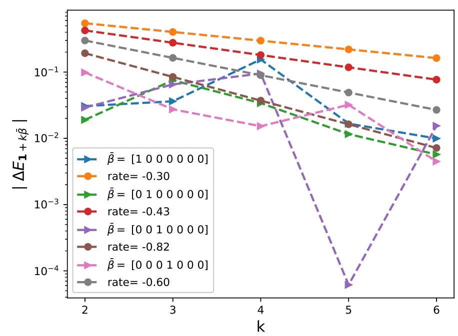
(a)



(b)



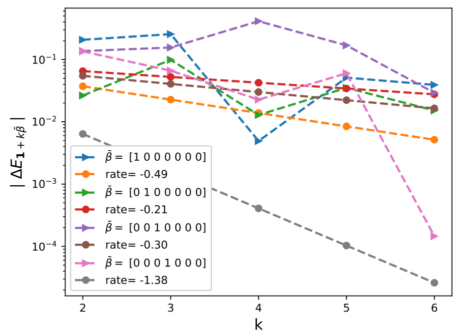
(c)



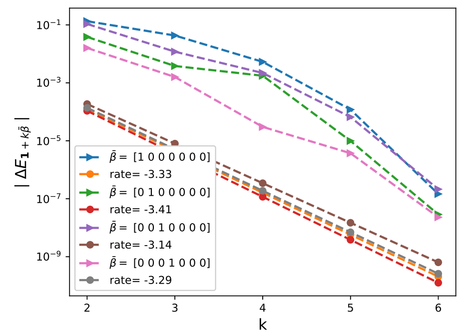
(d)

Figure 1.3: The rate of error convergence of first order differences $|\Delta \mathrm{E}_{\boldsymbol{\beta}}|$, defined by (1.13), ($\boldsymbol{\beta} = \mathbf{1} + k\overline{\boldsymbol{\beta}}$) for the example of single call option under Heston model, with parameters given by Set 3 in Table 1.1, using $N = 4$ time steps. In this case, we just show the first 4 dimensions which are used for the volatility noise (mainly dW_v in (1.1)). (a) using full truncation as in Section 1.1.1, (b) using the ABR scheme as in Section 1.2.1, (c) using the QE scheme as in Section 1.2.2, (d) using the smooth transformation as in Section 1.4.

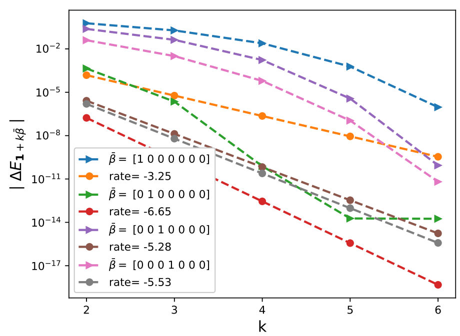
Mixed differences for the case of Set 4 parameters



(a)



(b)



(c)

Figure 1.4: The rate of error convergence of first order differences $|\Delta \mathrm{E}_{\boldsymbol{\beta}}|$, defined by (1.13), ($\boldsymbol{\beta} = \mathbf{1} + k\overline{\boldsymbol{\beta}}$) for the example of single call option under Heston model, with parameters given by Set 4 in Table 1.1, using $N = 4$ time steps. In this case, we just show the first 4 dimensions which are used for the volatility noise (mainly dW_v in (1.1)).(a) using full truncation as in Section 1.1.1, (b) using the ABR scheme as in Section 1.2.1, (c) using the QE scheme as in Section 1.2.2.

1.5.2 Comparison in terms of the weak error behavior (under process)

References Cited

[1] Aurélien Alfonsi. High order discretization schemes for the cir process: application to affine term structure and heston models. *Mathematics of Computation*, 79(269):209–237, 2010.

[2] Leif BG Andersen. Efficient simulation of the heston stochastic volatility model. *Available at SSRN 946405*, 2007.

[3] Leif BG Andersen and Rupert Brotherton-Ratcliffe. Extended libor market models with stochastic volatility. *Journal of Computational Finance*, 9(1), 2005.

[4] Mark Broadie and Özgür Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations research*, 54(2):217–231, 2006.

[5] Abdul-Lateef Haji-Ali, Fabio Nobile, Lorenzo Tamellini, and Raul Tempone. Multi-index stochastic collocation for random PDEs. *Computer Methods in Applied Mechanics and Engineering*, 306:95–122, 2016.

[6] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[7] Monique Jeanblanc, Marc Yor, and Marc Chesney. *Mathematical methods for financial markets*. Springer Science & Business Media, 2009.

[8] Christian Kahl and Peter Jäckel. Fast strong approximation monte carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[9] Roger Lord, Remmert Koekkoek, and Dick Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.