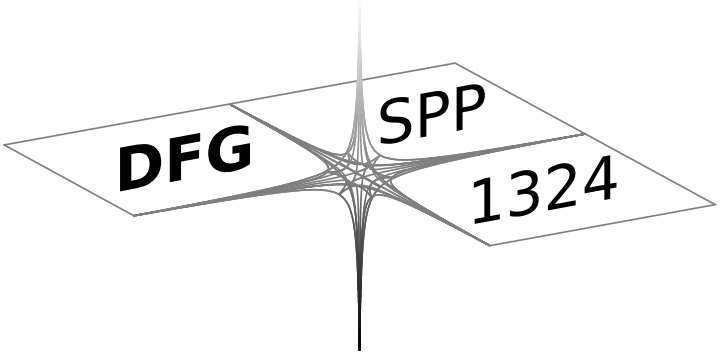
DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts

M. Altmayer, A. Neuenkirch

Preprint 144



Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

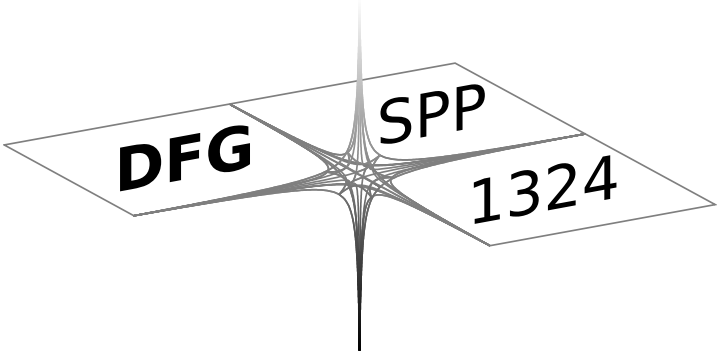
DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts

M. Altmayer, A. Neuenkirch

Preprint 144



The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts

Martin Altmayer*, Andreas Neuenkirch†

Abstract. In this manuscript, we establish an integration by parts formula for the quadrature of discontinuous payoffs in a multidimensional Heston model. For its derivation we use Malliavin calculus techniques and work under mild integrability conditions on the payoff and under the assumption of a strictly positive volatility. Since the integration by parts procedure smoothes the original functional, our formula in combination with a payoff-splitting allows to construct efficient multilevel Monte Carlo estimators, which is illustrated by several numerical examples.

Key words. Generalized Heston model, Malliavin calculus, Multilevel Monte Carlo, Drift-implicit Euler scheme

AMS subject classifications. 60H07; 60H35; 65C05; 91G60

1. Introduction. The Heston Model is a popular stochastic volatility model in mathematical finance. It was introduced in [18] and further developed since then, see e.g. the recent work [16] on the Heston model with stochastic interest rates or [15] on time-dependent coefficients. In its classical form the Heston model uses a Cox-Ingersoll-Ross process (CIR) as volatility and is given by the SDEs

$$\begin{aligned} \mathrm{d}S_t &= bS_t\mathrm{d}t + \sqrt{v_t}S_t(\rho\mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2), && t \in [0,T] \\ \mathrm{d}v_t &= \kappa(\lambda - v_t)\mathrm{d}t + \theta\sqrt{v_t}\mathrm{d}W_t^1, && t \in [0,T] \end{aligned}$$

with $S_0, v_0, \kappa, \lambda, \theta > 0$, $b \in \mathbb{R}$, $\rho \in [-1,1]$ and independent Brownian motions W^1, W^2. In this article we will also consider a generalized Heston model (introduced in [5]) that uses a mean-reverting constant elasticity of variance process (CEV) as volatility, i.e.

$$\mathrm{d}v_t = \kappa(\lambda - v_t)\mathrm{d}t + \theta v_t^\gamma \mathrm{d}W_t^1$$

with $\gamma \in (1/2, 1)$.

An efficient method to compute expectations of (smooth) functionals of SDEs is the multilevel Monte Carlo method, see [17] and [13]. Combining approximations using different step-sizes in a way that reduces the overall variance this method usually is significantly more efficient than standard Monte Carlo. However, the method requires a good L^2-convergence rate for the approximations which is often not easy to achieve for discontinuous functionals, see [6, 14].

In this article we are interested to compute

$$Ef(S_T)$$

*Institut für Mathematik, Universität Mannheim, A5,6, D-68131 Mannheim, Germany, altmayer@uni-mannheim.de

†Institut für Mathematik, Universität Mannheim, A5,6, D-68131 Mannheim, Germany, neuenkirch@kiwi.math.uni-mannheim.de

for general, e.g. discontinuous, payoff functionals $f : [0, \infty) \to \mathbb{R}$ by means of multilevel Monte Carlo methods. For this, we will prove that under certain assumptions such an expected payoff can be written as

$$Ef(S_T) = E\left(\frac{F(S_T)}{S_T} \cdot \Pi\right) \tag{1.1}$$

where $F : [0, \infty) \to \mathbb{R}$ is an antiderivative of f and

$$\Pi = 1 + \frac{1}{\sqrt{1-\rho^2}T} \cdot \int_0^T \frac{1}{\sqrt{v_r}} \mathrm{d}W_r^2 \tag{1.2}$$

Thus even for discontinuous f the expectation on right hand side of (1.1) contains a continuous functional of the price as well as a weight term independent of the functional. The proof of this representation formula relies on a Malliavin integration by parts technique. Moreover, we will show that this approach can be easily extended to multidimensional Heston models.

Based on the representation formula (1.1), which smoothes the original payoff functional, we will then construct a multilevel estimator for $p = Ef(S_T)$ that uses a log-Euler method for the Heston price, an Euler method for the Malliavin weight, and a drift-implicit Lamperti-Euler method for the volatility (see e.g. [2, 9]). The latter method preserves the strict positivity for the volatility, which is crucial for the discretization of the Malliavin weight (1.2). Combined with a payoff-splitting to reduce the variance of the Malliavin weight, this estimator outperforms the direct multilevel Monte Carlo estimator for $Ef(S_T)$ in our numerical experiments. We strongly suppose that this a general feature of the presented Malliavin integration by parts multilevel estimator and a complete error analysis will be carried out in a forthcoming project to justify this conjecture.

The remainder of this manuscript is structured as follows: In the following three sections we derive the required Malliavin regularity of the CIR- and CEV-processes and also of the Heston price. Note that we could treat the CIR- and CEV-process simultaneously (using the notation $\gamma \in [1/2, 1)$), but in our opinion this would be less accessible to the reader. The proofs for the CEV-process follow the same ideas as for the CIR-process, so in the analysis of the CEV-process (which is a bit more technical) we only outline the differences. Formula (1.1) is estalished in Section 5, while Section 6 studies the multidimensional case. Finally, in Section 7 we present the used multilevel Monte Carlo method, while Section 8 contains our numerical results. The appendix lists some required results from Malliavin calculus and other stochastic analysis tools.

2. The CIR-process. The first volatility process considered in this paper is the well known Cox-Ingersoll-Ross (CIR) process defined by the SDE

$$\mathrm{d}v_t = \kappa(\lambda - v_t)\mathrm{d}t + \theta\sqrt{v_t}\mathrm{d}W_t, \qquad v_0 > 0 \tag{2.1}$$

for $\kappa, \lambda, \theta > 0$. A theorem of Yamada and Watanabe (see [22, Prop. 5.2.13]) implies that this SDE possesses a unique strong solution. We will always work under the so called Feller condition for the CIR-process:

Assumption 2.1.

$$2\kappa\lambda > \theta^2$$

It turns out that this assumption is necessary for our mathematical analysis and often met in stock markets, see e.g. [10, 27, 1]. This condition in particular guarantees that

$$P(v_t > 0 \ \ \forall t \in [0,T]) = 1$$

for all $T > 0$, see e.g. [21, Lemma 2.2], and the existence of inverse moments up to a certain order:

Lemma 2.2. *Assume that $2\kappa\lambda > p\theta^2$ for some $p \in \mathbb{R}$. Then*

$$\sup_{t\in[0,T]} E(v_t^{-p}) < \infty$$

Proof. This follows from [19, Thm. 3.1]. See also [9]. ■

Moreover, all moments of the maximum of the CIR-process exist:

Lemma 2.3. *We have $E \sup_{t\in[0,T]} v_t^p < \infty$ for all $p \geq 1$.*

Proof. See Lemma 3.2 in [9]. ■

For a major part of this article we will consider the square root of the volatility, which we denote $\sigma_t := \sqrt{v_t}$, instead of the volatility itself. The Itō formula shows that σ_t follows the SDE

$$\mathrm{d}\sigma_t = \left(\left(\frac{\kappa\lambda}{2} - \frac{\theta^2}{8}\right)\frac{1}{\sigma_t} - \frac{\kappa}{2}\sigma_t\right)\mathrm{d}t + \frac{\theta}{2}\mathrm{d}W_t \tag{2.2}$$

The constant diffusion term is no surprise because the square root is up to the factor $2/\theta$ the Lamperti transformation of the CIR-process. In the following we set $C_\sigma := \kappa\lambda/2 - \theta^2/8$, which is positive by (2.1).

2.1. Approximation of the square root CIR-process. To prove that σ_t and v_t are Malliavin differentiable, we will follow an article of Alòs and Ewald [4] and approximate σ_t by a process whose SDE has Lipschitz coefficients. This allows us to use standard results from Malliavin calculus. Because we later want to show the differentiability of $1/\sigma_t$ and S_t we cannot use the results of [4] directly and have to choose a different function f_ε than Alòs and Ewald.

For each $\varepsilon > 0$ let f_ε be a function that satisfies:

(i) $f_\varepsilon \in C^\infty(\mathbb{R}; \mathbb{R}_{>0})$ with bounded derivative,
(ii) $f_\varepsilon(x) = \frac{1}{x}$ for all $x \geq \varepsilon$,
(iii) f_ε is monotonically decreasing,
(iv) $f_\varepsilon(x) \leq \frac{1}{x}$ for all $x \in \mathbb{R}_{>0}$,
(v) $|f_\varepsilon'(x)| \leq \frac{1}{x^2} = \left|\left(\frac{1}{x}\right)'\right|$ for all $x \in \mathbb{R}_{>0}$.

Now define a process σ_t^ε by the SDE

$$\mathrm{d}\sigma_t^\varepsilon = \left(C_\sigma f_\varepsilon(\sigma_t^\varepsilon) - \frac{\kappa}{2}\sigma_t^\varepsilon\right)\mathrm{d}t + \frac{\theta}{2}\mathrm{d}W_t, \qquad \sigma_0^\varepsilon = \sqrt{v_0} \tag{2.3}$$

Because the coefficients satisfy the usual Lipschitz and linear growth conditions this SDE has a unique strong solution.

Proposition 2.4. *For almost all $\omega \in \Omega$ we have $\sigma_t^\varepsilon(\omega) = \sigma_t(\omega)$ for all $t \in [0,T]$ and all $0 < \varepsilon < \varepsilon_0(\omega) := \inf_{t\in[0,T]} \sigma_t(\omega)$. In particular*

$$\sup_{t\in[0,T]} |\sigma_t^\varepsilon - \sigma_t| \to 0$$

holds almost surely for $\varepsilon \to 0$.

Proof. Comparing the SDEs in integral form yields for all $\varepsilon > 0$ and almost all $\omega \in \Omega$ that

$$|\sigma_t^\varepsilon(\omega) - \sigma_t(\omega)| \le \int_0^t C_\sigma \left| f_\varepsilon(\sigma_\tau^\varepsilon(\omega)) - \frac{1}{\sigma_\tau(\omega)} \right| \mathrm{d}\tau + \int_0^t \frac{\kappa}{2} |\sigma_\tau^\varepsilon(\omega) - \sigma_\tau(\omega)| \, \mathrm{d}\tau, \qquad t \in [0,T]$$

Fix an ω such that the path $\sigma_\cdot(\omega)$ is positive. Then for $0 < \varepsilon < \varepsilon_0(\omega)$ the functions f_ε and $g : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $g(x) = 1/x$ coincide on the path $\sigma_\cdot(\omega)$, i.e. $f_\varepsilon(\sigma_t(\omega)) = g(\sigma_t(\omega))$ for all $t \in [0,T]$. Thus we have

$$|\sigma_t^\varepsilon(\omega) - \sigma_t(\omega)| \le \int_0^t \left(C_\sigma L_\varepsilon + \frac{\kappa}{2} \right) |\sigma_\tau^\varepsilon(\omega) - \sigma_\tau(\omega)| \, \mathrm{d}\tau, \qquad t \in [0,T],$$

where L_ε is a Lipschitz constant for f_ε. By Gronwall's lemma it follows that $\sigma_t(\omega) = \sigma_t^\varepsilon(\omega)$ for all $t \in [0,T]$. ■

Note that the last proposition does not imply that σ^ε has almost surely positive paths. In fact, when σ falls below ε, the approximation σ^ε might fall below 0, too.

Lemma 2.5. *Let u_t be an Ornstein-Uhlenbeck process given by the SDE*

$$\mathrm{d}u_t = -\frac{\kappa}{2} u_t \mathrm{d}t + \frac{\theta}{2} \mathrm{d}W_t, \qquad u_0 = \sigma_0 \tag{2.4}$$

Then $P(u_t \le \sigma_t^\varepsilon \le \sigma_t \ \ \forall t \in [0,T]) = 1$.

Proof. (This follows the proof of [4, Prop. 2.1], but with f_ε instead of Λ_ε.)

(1) We first show $u_t(\omega) \le \sigma_t^\varepsilon(\omega)$ uniformly in t and for almost all ω. The diffusion coefficients of (2.3) and (2.4) are the same while the drift coefficient of u_t is smaller than the one of σ_t^ε. By the Yamada-Watanabe comparison lemma (see Theorem 9.1 in the Appendix) we have $P(u_t \le \sigma_t^\varepsilon \ \ \forall\, t \in [0,T]) = 1$.

(2) The comparison lemma cannot be used to prove the second inequality since the drift term in the SDE for σ_t is not continuous on $\mathbb{R}$. But because $\sigma_t > 0$ it suffices to show the inequality $v_t^\varepsilon := (\sigma_t^\varepsilon)^2 \le \sigma_t^2 = v_t$. The SDE for v_t^ε is

$$\mathrm{d}v_t^\varepsilon = \left(2C_\sigma \sqrt{v_t^\varepsilon} f_\varepsilon(\sqrt{v_t^\varepsilon}) - \kappa v_t^\varepsilon + \frac{\theta^2}{4} \right) \mathrm{d}t + \theta \sqrt{v_t^\varepsilon} \mathrm{d}W_t, \qquad v_0^\varepsilon = v_0 \tag{2.5}$$

Now the drift and diffusion terms are continuous on $\mathbb{R}$ (set $\sqrt{x} = 0$ for $x \le 0$ to define the drift and diffusion coefficients on $\mathbb{R}_{\le 0}$) and we can apply the comparison lemma after we have verified its remaining conditions. Condition (ii) of the comparison lemma can be fulfilled

choosing $h(x) = \theta\sqrt{x}$ *because of* $\sqrt{|x-y|} \geq |\sqrt{x} - \sqrt{y}|$ *for all* $x, y \geq 0$. *Since we assumed* $f_\varepsilon(x) \leq 1/x$ *for all* $x > 0$ *and thus* $f_\varepsilon(\sqrt{x})\sqrt{x} \leq 1$, *condition (iv) follows from*

$$\kappa\lambda \geq \frac{\theta^2}{4} \Rightarrow \left(\kappa\lambda - \frac{\theta^2}{4}\right)\left(\sqrt{x} \cdot f_\varepsilon(\sqrt{x})\right) \leq \kappa\lambda - \frac{\theta^2}{4}$$
$$\Rightarrow C_\sigma\left(\sqrt{x} \cdot f_\varepsilon(\sqrt{x})\right) - \kappa x + \frac{\theta^2}{4} \leq \kappa(\lambda - x)$$

For the final condition (v) it suffices if only one of the drift coefficients is Lipschitz-continuous, which $\mathbb{R} \ni x \mapsto \kappa(\lambda - x) \in \mathbb{R}$ *clearly is.* ■

Proposition 2.6. *For each* $p \geq 1$ *we have* $\sup_{t\in[0,T]} |\sigma_t^\varepsilon - \sigma_t| \to 0$ *in* L^p.

Proof. From the previous lemma we obtain

$$\sup_{t\in[0,T]} |\sigma_t^\varepsilon - \sigma_t| \leq \sup_{t\in[0,T]} |\sigma_t - u_t| \leq \sup_{t\in[0,T]} |\sigma_t| + \sup_{t\in[0,T]} |u_t|$$

We have $\sup_{t\in[0,T]} \sigma_t \in L^p$ *by Lemma 2.3 and it is a standard application of Doob's martingale inequality that an Ornstein-Uhlenbeck process* u_t *satisfies* $\sup_{t\in[0,T]} |u_t| \in L^p$. *Thus the claim follows from Proposition 2.4 using dominated convergence.* ■

2.2. Malliavin differentiability of the CIR-process. Because the volatility depends on only one Brownian motion, we will use one-dimensional Malliavin calculus in this section. The underlying Hilbert space is thus given by $H = L^2([0,T])$.

Proposition 2.7. *We have* $\sigma_t^\varepsilon \in \mathbb{D}^{1,\infty}$ *and*

$$D_r\sigma_t^\varepsilon = \frac{\theta}{2} \cdot \exp\left(\int_r^t \left(-\frac{\kappa}{2} + C_\sigma f_\varepsilon'(\sigma_s^\varepsilon)\right) \mathrm{d}s\right) \cdot 1_{[0,t]}(r) \leq \frac{\theta}{2} \tag{2.6}$$

Proof. The claim $\sigma_t^\varepsilon \in \mathbb{D}^{1,\infty}$ *and the form of the derivative follow from Theorem 9.4. Note that we can only apply this result because the coefficients of the SDE for* σ^ε *are differentiable with bounded derivatives on the whole of* $\mathbb{R}$. *The* $\theta/2$*-bound is clear because* $f_\varepsilon'(x) \leq 0$ *(this is not true for the corresponding function* Λ_ε *in [4]).* ■

Proposition 2.8. *We have* $\sigma_t \in \mathbb{D}^{1,\infty}$ *and the derivative is given by*

$$D_r\sigma_t = \frac{\theta}{2} \cdot \exp\left(\int_r^t \left(-\frac{\kappa}{2} - \frac{C_\sigma}{\sigma_s^2}\right) \mathrm{d}s\right) \cdot 1_{[0,t]}(r) \leq \frac{\theta}{2} \tag{2.7}$$

Additionally the following uniform convergence holds true almost surely and in L^p *for all* $p \geq 1$*:*

$$\sup_{r,t\in[0,T]} |D_r\sigma_t^\varepsilon - D_r\sigma_t| \to 0$$

as $\varepsilon \to 0$.

Proof. Set

$$v_{r,t} = \frac{\theta}{2} \cdot \exp\left(\int_r^t \left(-\frac{\kappa}{2} - \frac{C_\sigma}{\sigma_s^2}\right) \mathrm{d}s\right) \cdot 1_{[0,t]}(r)$$

From Proposition 2.4 we know that for almost all $\omega \in \Omega$ we have $\varepsilon_0(\omega) := \inf_{s\in[0,T]} \sigma_s(\omega) > 0$ and $\sigma_s^\varepsilon(\omega) = \sigma_s(\omega)$ for all $0 < \varepsilon < \varepsilon_0(\omega)$ and $s \in [0,T]$. Hence comparing the integrals in (2.6) and (2.7) and using the definition of f_ε we have

$$D_r\sigma_t^\varepsilon(\omega) = v_{r,t}(\omega)$$

for $r,t \in [0,T]$ and all $0 < \varepsilon < \varepsilon_0(\omega)$. Since ε does not depend on $r,t \in [0,T]$ this shows

$$\sup_{r,t\in[0,T]} |D_r\sigma_t^\varepsilon - v_{r,t}| \to 0 \qquad a.s.$$

as $\varepsilon \to 0$ and the $L^p(\Omega)$-convergence follows by dominated convergence. Hence $D_\cdot\sigma_t^\varepsilon \to v_{\cdot,t}$ in $L^p(\Omega;H)$ and this implies $D_\cdot\sigma_t = v_{\cdot,t}$ due to the closedness of the Malliavin derivative. ■

Theorem 2.9. *The Heston volatility v_t is in $\mathbb{D}^{1,\infty}$ and has the derivative*

$$D_r v_t = 2\sigma_t \cdot D_r(\sigma_t) = \theta \cdot \exp\left(\int_r^t \left(-\frac{\kappa}{2} - \frac{C_\sigma}{v_t}\right) \mathrm{d}s\right) \cdot \sqrt{v_t} \cdot 1_{[0,t]}(r) \tag{2.8}$$

Proof. This follows from the Malliavin chain rule Theorem 9.2. ■

Our quadrature rule will contain the inverse of the square root volatility. As a preparation we prove now its differentiability.

Proposition 2.10. *The inverse square root volatility σ_t^{-1} is in $\mathbb{D}^{1,2}$ with derivative $D_r\sigma_t^{-1} = -\sigma_t^{-2} \cdot D_r\sigma_t$.*

Proof. Fix $t \in [0,T]$. By the chain rule Theorem 9.2 we have $f_\varepsilon(\sigma_t) \in \mathbb{D}^{1,2}$ with derivative $f_\varepsilon'(\sigma_t) \cdot D_r\sigma_t$. This converges almost surely to $-1/\sigma_t^2 \cdot D_r\sigma_t$ and its absolute value is bounded by $\frac{\theta}{2} \cdot 1/\sigma_t^2$ (see (2.7)) which is in $L^p(\Omega)$ due to Lemma 2.2. This implies $1/\sigma_t \in \mathbb{D}^{1,2}$ and the given form of the derivative. ■

3. The CEV-process. In this section we will prove the same results as in the previous one for a new volatility process which differs from the CIR-process only in the exponent of the diffusion term. For $\gamma \in (1/2,1)$ and $\kappa,\lambda,\theta > 0$ define $v_t^{(\gamma)}$ by the SDE

$$\mathrm{d}v_t^{(\gamma)} = \kappa(\lambda - v_t^{(\gamma)})\mathrm{d}t + \theta\big(v_t^{(\gamma)}\big)^\gamma \mathrm{d}W_t \tag{3.1}$$

Again [22, Prop. 5.2.13] guarantees the existence of a strong solution. In the literature this process is denoted as mean-reverting constant elasticity of variance process (CEV), see e.g. [21, 24]. In the following, we will usually omit the index γ on each process.

In contrast to the CIR-process we do not need further restrictions on the parameters to ensure that v_t remains positive, i.e. we have

$$P\big(v_t > 0 \ \ \forall t \in [0,T]\big) = 1$$

for all $T > 0$, see [5, Prop. 2.1]. Also its moments of any order exist:

Lemma 3.1. *Let $p \geq 0$. Then*

$$E \sup_{t\in[0,T]} v_t^p < \infty$$

and

$$\sup_{t\in[0,T]} E(v_t^{-p}) < \infty$$

Proof. See Lemma 2.1 in [7]. ■

As for the CIR-process we will use a scaled Lamperti's transformation to replace the SDE by one with a constant diffusion coefficient. In this case the transformation takes the form $\sigma_t := v_t^{1-\gamma}$ and leads to the SDE

$$\mathrm{d}\sigma_t = (1-\gamma)\left(\kappa\lambda\sigma_t^{-\frac{\gamma}{1-\gamma}} - \kappa\sigma_t - \frac{\gamma\theta^2}{2}\sigma_t^{-1}\right)\mathrm{d}t + \theta(1-\gamma)\mathrm{d}W_t \tag{3.2}$$

Now set

$$f(x) := \kappa\lambda x^{-\frac{\gamma}{1-\gamma}} - \frac{\gamma\theta^2}{2}x^{-1}$$

so that

$$\mathrm{d}\sigma_t = (1-\gamma)\left(f(\sigma_t) - \kappa\sigma_t\right)\mathrm{d}t + \theta(1-\gamma)\mathrm{d}W_t$$

Note that

$$\sup_{x\searrow 0} f(x) = \infty \quad \text{and} \quad \sup_{x\to\infty} f(x) = 0,$$

so there exists a $C_f \geq 0$ with

$$f(x) \geq -C_f x \qquad \forall x \in \mathbb{R}_{>0}$$

Finally, since

$$f'(x) = -\frac{\kappa\lambda\gamma}{1-\gamma}x^{-\frac{1}{1-\gamma}} + \frac{\gamma\theta^2}{2}x^{-2}$$

we have

$$\sup_{x\in\mathbb{R}_{>0}} f'(x) < \infty$$

3.1. Approximation of the CEV-volatility. Again we will modify (3.2), replace f by a C^∞-function and approximate σ by the solution of this modified SDE. For each $\varepsilon > 0$ choose a function f_ε with

(i) $f_\varepsilon \in C^\infty(\mathbb{R};\mathbb{R})$ with bounded derivative,
(ii) $f_\varepsilon = f$ on $[\varepsilon,\infty)$,
(iii) f_ε is monotonically decreasing on $(-\infty,0]$,
(iv) $f_\varepsilon(x) \leq f(x)$ for $x \in \mathbb{R}_{>0}$,
(v) $f_\varepsilon'(x) \leq f'(x)$ for all $x \in \mathbb{R}_{>0}$,
(vi) $f_\varepsilon(x) \geq -C_f x$ for all $x \in \mathbb{R}_{>0}$.

Now define σ_t^ε by the modified SDE

$$\mathrm{d}\sigma_t^\varepsilon = (1-\gamma)(f_\varepsilon(\sigma_t^\varepsilon) - \kappa\sigma_t^\varepsilon)\mathrm{d}t + \theta(1-\gamma)\mathrm{d}W_t, \qquad \sigma_0^\varepsilon = \sigma_0 \tag{3.3}$$

Proposition 3.2. *For almost all $\omega \in \Omega$ we have $\sigma_t^\varepsilon(\omega) = \sigma_t(\omega)$ for all $t \in [0,T]$ and $0 < \varepsilon < \varepsilon_0(\omega) := \inf_{t\in[0,T]} \sigma_t(\omega)$. In particular we have*

$$\sup_{t\in[0,T]} |\sigma_t^\varepsilon - \sigma_t| \to 0$$

almost surely for $\varepsilon \to 0$.

Proof. This can be shown completely analogously to Proposition 2.4. ■

Lemma 3.3. *Let u_t be the solution of the SDE*

$$\mathrm{d}u_t = -(1-\gamma)(C_f + \kappa)u_t \mathrm{d}t + \theta(1-\gamma)\mathrm{d}W_t, \qquad u_0 = \sigma_0 \tag{3.4}$$

Then $P(u_t \leq \sigma_t^\varepsilon \leq \sigma_t \ \ \forall t \in [0,T]) = 1$ for all $\varepsilon > 0$.

Proof. (1) We first show $u_t \leq \sigma_t^\varepsilon$. The diffusion coefficients of (3.3) and (3.4) are the same while the drift coefficient of u_t is smaller than that of σ_t^ε. So Theorem 9.1 gives $P(u_t \leq \sigma_t^\varepsilon \ \ \forall\, t \in [0,T]) = 1$.

(2) The comparison lemma cannot be used to prove $\sigma_t^\varepsilon \leq \sigma_t$ directly because the drift term in the SDE for σ_t is not continuous on $\mathbb{R}$. But because $\sigma_t > 0$ it suffices to show the corresponding inequality after the back transformation $v_t^\varepsilon := (\sigma_t^\varepsilon)^{1/(1-\gamma)} \leq \sigma_t^{1/(1-\gamma)} = v_t$. The SDE for v_t^ε is

$$\mathrm{d}v_t^\varepsilon = \left[(v_t^\varepsilon)^\gamma f_\varepsilon((v_t^\varepsilon)^{1-\gamma}) - \kappa v_t^\varepsilon + \frac{\gamma\theta^2}{2}(v_t^\varepsilon)^{2\gamma-1}\right]\mathrm{d}t + \theta(v_t^\varepsilon)^\gamma \mathrm{d}W_t \qquad v_0^\varepsilon = v_0 \tag{3.5}$$

Now the drift and diffusion terms are continuous on $\mathbb{R}$, if we set $x^\alpha = 0$ for $\alpha \in (0,1)$ and $x < 0$, and it remains to verify the conditions (ii)-(v) of the comparison lemma. Condition (ii) can be fulfilled choosing $h(x) = \theta \cdot x^\gamma$ since this function is concave and satisfies $\int_0^\varepsilon x^{-2\gamma} dx = \infty$. For condition (iv) note that the drift coefficient of v_t^ε is

$$x \mapsto x^\gamma f_\varepsilon(x^{1-\gamma}) - \kappa x + \frac{\theta^2\gamma}{2}x^{2\gamma-1}$$

for $x > 0$, while for $x < 0$ it is

$$x \mapsto -\kappa x$$

Due to assumption (iv) for f_ε we have

$$x^\gamma f_\varepsilon(x^{1-\gamma}) + \frac{\theta^2\gamma}{2}x^{2\gamma-1} \leq x^\gamma f(x^{1-\gamma}) + \frac{\theta^2\gamma}{2}x^{2\gamma-1} = \kappa\lambda$$

for $x > 0$, while for $x < 0$ we clearly have $-\kappa x \leq \kappa(\lambda - x)$. For the final condition (v) it suffices that the drift coefficient of v_t is Lipschitz-continuous. ■

The following Proposition can be shown again identically as its analogon for the square root CIR-process.

Proposition 3.4. *For each $p \geq 1$ we have $\sup_{t\in[0,T]} |\sigma_t^\varepsilon - \sigma_t| \to 0$ in L^p as $\varepsilon \to 0$.*

3.2. Differentiability of the CEV-volatility. As in Proposition 2.7 we use Theorem 9.4 to obtain $\sigma_t^\varepsilon \in \mathbb{D}^{1,\infty}$ and to compute the derivative as

$$D_r\sigma_t^\varepsilon = \theta(1-\gamma)\cdot\exp\left(\int_r^t (1-\gamma)\left(f_\varepsilon'(\sigma_s^\varepsilon)-\kappa\right)\mathrm{d}s\right)\cdot 1_{[0,t]}(r)$$

Note that this derivative is not necessarily bounded by $\theta(1-\gamma)$ as it was with the CIR-process. However, it is bounded by $\theta(1-\gamma)\cdot\exp\left(T(1-\gamma)\cdot\sup_{x\in\mathbb{R}_{>0}} f'(x)\right)$ because $f_\varepsilon'(x)\le f'(x)$ for $x>0$ and $f_\varepsilon'(x)\le 0$ for $x\le 0$. Using this bound and the previous approximation results one can prove the following proposition exactly as Proposition 2.8.

Proposition 3.5. *We have $\sigma_t\in\mathbb{D}^{1,\infty}$ and the derivative is given by*

$$D_r\sigma_t = \theta(1-\gamma)\cdot\exp\left(\int_r^t (1-\gamma)(f'(\sigma_s)-\kappa)\mathrm{d}s\right)\cdot 1_{[0,t]}(r) \tag{3.6}$$

Additionally the following uniform convergence holds true almost surely and in L^p for all $p\ge 1$:

$$\sup_{r,t\in[0,T]} |D_r\sigma_t^\varepsilon - D_r\sigma_t| \to 0$$

as $\varepsilon\to 0$.

The chain rule Theorem 9.2 shows:

Theorem 3.6. *The volatility process v_t is in $\mathbb{D}^{1,\infty}$ and has the derivative*

$$\begin{aligned} D_r v_t &= \frac{1}{1-\gamma}(\sigma_t)^{\frac{\gamma}{1-\gamma}}\cdot D_r\sigma_t \\ &= \theta v_t^\gamma\cdot\exp\left(\int_r^t\left(-\frac{\kappa\lambda\gamma}{v_s}+\frac{\theta^2}{2}\frac{\gamma(1-\gamma)}{v_s^{2(1-\gamma)}}-\kappa(1-\gamma)\right)\mathrm{d}s\right)\cdot 1_{[0,t]}(r)\end{aligned} \tag{3.7}$$

Later on we need:

Proposition 3.7. *The inverse square root volatility $1/\sqrt{v_t}$ is in $\mathbb{D}^{1,2}$ with derivative*

$$D_r\left(\frac{1}{\sqrt{v_t}}\right) = -\frac{1}{2(1-\gamma)}v_t^{-\frac{3}{2}+\gamma}\cdot D_r\sigma_t$$

Proof. By the chain rule Theorem 9.2 we have $\sqrt{f_\varepsilon(v_t)}\in\mathbb{D}^{1,2}$ with

$$D_r(\sqrt{f_\varepsilon(v_t)}) = \frac{1}{2}\cdot\frac{f_\varepsilon'(v_t)}{\sqrt{f_\varepsilon(v_t)}}\cdot D_r v_t = \frac{1}{2(1-\gamma)}\cdot\frac{f_\varepsilon'(v_t)}{\sqrt{f_\varepsilon(v_t)}}\cdot\sigma_t^{\frac{\gamma}{1-\gamma}}\cdot D_r\sigma_t$$

where f_ε is the approximation of $\mathbb{R}_{>0}\ni x\mapsto 1/x\in\mathbb{R}_{>0}$ from Section 2. By dominated convergence $\sqrt{f_\varepsilon(v_t)}\to 1/\sqrt{v_t}$ in $L^p(\Omega)$ and

$$D_r(\sqrt{f_\varepsilon(v_t)})\to -\frac{1}{2(1-\gamma)}v_t^{-3/2}\cdot\sigma_t^{\frac{\gamma}{1-\gamma}}\cdot D_r\sigma_t$$

in $L^p(\Omega;H)$. Because the Malliavin derivative is a closed operator, the last term must be the derivative $D_r(1/\sqrt{v_t})$. ■

4. The Price Process. This section deals with the price process S_t given by

$$\mathrm{d}S_t = bS_t\mathrm{d}t + \sqrt{v_t}S_t\mathrm{d}B_t \tag{4.1}$$

Here v_t is either the CIR- or the CEV-process and $B_t = \rho \cdot W_t^1 + \sqrt{1-\rho^2} \cdot W_t^2$ is a Brownian motion composed of two indepedent Brownian motions, with W^1 being the Brownian motion driving v_t. To calculate the derivative of the price we will concentrate on the log-price $X_t := \ln S_t$. Itō's formula shows that X_t can be given explicitly in terms of the volatility:

$$\mathrm{d}X_t = \left(b - \frac{1}{2}v_t\right)\mathrm{d}t + \sqrt{v_t}\mathrm{d}B_t \tag{4.2}$$

We will again use an approximation by an SDE with Lipschitz coefficients. For $\varepsilon > 0$ let ψ_ε be a function such that

(i) $\psi_\varepsilon : \mathbb{R} \to \mathbb{R}$ is bounded and continuously differentiable,
(ii) $|\psi_\varepsilon(x)| \leq |x|$ for all $x \in \mathbb{R}$,
(iii) $\psi_\varepsilon(x) = x$ on $[0, 1/\varepsilon]$,
(iv) $|\psi_\varepsilon'(x)| \leq 1$ for all $x \in \mathbb{R}$.

Denote by σ^ε the approximation to the transformed volatility $\sigma_t = v_t^{1-\gamma}$ of Section 2 (with $\gamma = 1/2$) or Section 3. To obtain the derivative we define X_t^ε by

$$X_t^\varepsilon = X_0 + \int_0^t \left(b - \frac{1}{2}\psi_\varepsilon^{\frac{1}{1-\gamma}}(\sigma_s^\varepsilon)\right)\mathrm{d}s + \int_0^t \psi^{\frac{1}{2(1-\gamma)}}(\sigma_s^\varepsilon)\mathrm{d}B_s$$

Lemma 4.1. *For each $p, q \geq 1$ we have*

$$E \sup_{t\in[0,T]} |\sigma_t^q - \psi_\varepsilon^q(\sigma_t^\varepsilon)|^p \to 0$$

as well as

$$E \sup_{t\in[0,T]} \left|\sigma_t^q - \psi_\varepsilon^q(\sigma_t^\varepsilon) \cdot \psi_\varepsilon'(\sigma_t^\varepsilon)\right|^p \to 0$$

Proof. By Propositions 2.4 and 3.2 for almost all ω there exists an $\varepsilon_0(\omega) > 0$ such that $\varepsilon < \sigma_t(\omega) = \sigma_t^\varepsilon(\omega) < 1/\varepsilon$ for all $t \in [0,T]$ and $0 < \varepsilon < \varepsilon_0(\omega)$. For such ω and ε we have

$$\sup_{t\in[0,T]} |\sigma_t^q(\omega) - \psi_\varepsilon^q(\sigma_t^\varepsilon(\omega))| = \sup_{t\in[0,T]} \left|\sigma_t^q(\omega) - \psi_\varepsilon^q(\sigma_t^\varepsilon(\omega)) \cdot \psi_\varepsilon'(\sigma_t^\varepsilon(\omega))\right| = 0$$

Also, both suprema are bounded by $2\sup_{t\in[0,T]} |\sigma_t|^q$. Since this is L^p-integrable, the assertion follows by dominated convergence. ■

Proposition 4.2. *We have $\sup_{t\in[0,T]} |X_t^\varepsilon - X_t| \to 0$ in L^p for all $p \geq 1$.*

Proof. From the SDEs we obtain the following inequality for some constant $c > 0$:

$$E \sup_{t\in[0,T]} |X_t^\varepsilon - X_t|^p \leq c \cdot E \sup_{t\in[0,T]} \left|\int_0^t \left(\sigma_s^{\frac{1}{1-\gamma}} - \psi_\varepsilon^{\frac{1}{1-\gamma}}(\sigma_s^\varepsilon)\right)\mathrm{d}s\right|^p + c \cdot E \sup_{t\in[0,T]} \left|\int_0^t \left(\sigma_s^{\frac{1}{2(1-\gamma)}} - \psi_\varepsilon^{\frac{1}{2(1-\gamma)}}(\sigma_s^\varepsilon)\right)\mathrm{d}B_s\right|^p$$

The first term converges to zero by the previous lemma. For the second iterm we use the Burkholder-Davis-Gundy inequality as well as $|x-y|^2 \leq |x^2-y^2|$ for $x, y \geq 0$: There is a constant c such that

$$E \sup_{t\in[0,T]} \left| \int_0^t \left(\sigma_s^{\frac{1}{2(1-\gamma)}} - \psi_\varepsilon^{\frac{1}{2(1-\gamma)}}(\sigma_s^\varepsilon) \right) \mathrm{d}B_s \right|^p \leq c \cdot E \left(\int_0^T \left| \sigma_s^{\frac{1}{1-\gamma}} - \psi_\varepsilon^{\frac{1}{1-\gamma}}(\sigma_s^\varepsilon) \right| \mathrm{d}s \right)^{p/2}$$

which converges to zero again by the previous lemma. ■

Theorem 4.3. *The log-price X_t is in $\mathbb{D}^{1,p}$ for all $p \geq 1$ and its derivative is given by*

$$\begin{aligned} D_r^1 X_t &= \rho\sqrt{v_r} - \frac{1}{2(1-\gamma)} \cdot \int_0^t \sigma_s^{\frac{\gamma}{1-\gamma}} D_r\sigma_s \mathrm{d}s + \frac{1}{2(1-\gamma)} \cdot \int_0^t \sigma_s^{\frac{1}{2(1-\gamma)}-1} D_r\sigma_s \mathrm{d}B_s \\ D_r^2 X_t &= \sqrt{1-\rho^2}\sqrt{v_r} \end{aligned}$$

for $r \leq t$ and $D_r^1 X_t = D_r^2 X_t = 0$ else. Moreover for each $r \in [0,T]$ and $i \in \{1,2\}$ the uniform convergence

$$\sup_{t\in[0,T]} |D_r^i X_t - D_r^i X_t^\varepsilon| \to 0$$

holds in $L^p(\Omega)$.

Proof. Because the system of SDEs for X_t^ε and σ_t^ε has globally Lipschitz coefficients, we can use Theorem 9.4 to calculate the derivatives: For $r \leq t$ we have

$$\begin{aligned} D_r^1 X_t^\varepsilon &= \rho\psi_\varepsilon^{\frac{1}{2(1-\gamma)}}(\sigma_r^\varepsilon) - \frac{1}{2(1-\gamma)} \cdot \int_r^t \psi_\varepsilon^{\frac{\gamma}{1-\gamma}}(\sigma_s^\varepsilon) t\psi_\varepsilon'(\sigma_s^\varepsilon) D_r^1\sigma_s^\varepsilon \mathrm{d}s \\ &\quad + \frac{1}{2(1-\gamma)} \cdot \int_r^t \psi_\varepsilon^{\frac{1}{2(1-\gamma)}-1}(\sigma_s^\varepsilon)\psi_\varepsilon'(\sigma_s^\varepsilon) D_r^1\sigma_s^\varepsilon \mathrm{d}B_s \\ D_r^2 X_t^\varepsilon &= \sqrt{1-\rho^2}\psi_\varepsilon^{\frac{1}{2(1-\gamma)}}(\sigma_r^\varepsilon) \end{aligned}$$

The second derivative and the first term of the first one converge to $\rho\sqrt{v_r}$ and $\sqrt{1-\rho^2}\sqrt{v_r}$, respectively, using Lemma 4.1. Now consider the second term of the first derivative:

$$\begin{aligned} &\sup_{t\in[0,T]} \left| \int_r^t \left(\psi_\varepsilon^{\frac{\gamma}{1-\gamma}}(\sigma_s^\varepsilon)\psi_\varepsilon'(\sigma_s^\varepsilon) D_r^1\sigma_s^\varepsilon - \sigma_s^{\frac{\gamma}{1-\gamma}} D_r^1\sigma_s \right) \mathrm{d}s \right| \\ &\qquad \leq T \cdot \sup_{s\in[0,T]} \left| \psi_\varepsilon^{\frac{\gamma}{1-\gamma}}(\sigma_s^\varepsilon)\psi_\varepsilon'(\sigma_s^\varepsilon) - \sigma_s^{\frac{\gamma}{1-\gamma}} \right| \cdot \sup_{s\in[0,T]} |D_r^1\sigma_s^\varepsilon| \\ &\qquad\quad + T \cdot \sup_{s\in[0,T]} \left| \sigma_s^{\frac{\gamma}{1-\gamma}} \right| \cdot \sup_{s\in[0,T]} \left| D_r^1\sigma_s^\varepsilon - D_r^1\sigma_s \right| \end{aligned}$$

This converges to 0 in $L^p(\Omega)$ using Lemma 4.1, Lemma 2.2/3.1 and Proposition 2.8/3.5. Convergence of the third term of the first derivative can be shown analogously, using additionally the Burkholder-Davis-Gundy inequality. ■

It is well known that in the generalized Heston model moment explosions may appear, see e.g. [5]. For the existence of the moments for all $T > 0$, one has the following result (see [12]):

Theorem 4.4. *Let $u \in [1, \infty)$ and define*

$$T^*(u) = \inf\{t \geq 0 : ES_t^u = \infty\}$$

(1) For the Heston model with the CIR-process ($\gamma = \frac{1}{2}$) as volatility process we have

$$T^*(u) = \infty \Leftrightarrow \rho \leq -\sqrt{\frac{u-1}{u}} + \frac{\kappa}{\theta u}$$

(2) For the Heston model with the CEV-process ($\gamma \in (\frac{1}{2}, 1)$) as volatility process we have

$$T^*(u) = \begin{cases} \infty & \text{for} \quad \rho < -\sqrt{\frac{u-1}{u}} \\ 0 & \text{for} \quad \rho > -\sqrt{\frac{u-1}{u}} \end{cases}$$

Based on the following result one can check the Malliavin smoothness of S_t:

Theorem 4.5. *Assume $S_t \in L^{p+\varepsilon}(\Omega)$ for some $p \geq 1$ and $\varepsilon > 0$. Then the Heston price S_t is in $\mathbb{D}^{1,p}$ and its derivative is $DS_t = S_t \cdot DX_t$.*

Proof. *Because $X_t \in \mathbb{D}^{1,\infty}$ we have $S_t \cdot DX_t \in L^p(\Omega; H)$. Now the claim follows directly from the chain rule Theorem 9.2.* ■

5. Quadrature. Now we can turn to the derivation of our quadrature formula. We will make the following assumption on the payoff function f:

Assumption 5.1. *Assume that $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a measurable function bounded by a polynomial p and that $F\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ is an antiderivative of f given by $F(x) = \int_0^x f(z)\mathrm{d}z$. Furthermore assume that there is an $\varepsilon > 0$ such that $S_T \in L^{\max\{2,\, \deg p\}+\varepsilon}$.*

The following Lemma is a final technical preparation.

Lemma 5.2. *For all $p \geq 0$ we have*

$$E\left(\int_0^T \frac{1}{\sqrt{v_s}}\,\mathrm{d}s\right)^p < \infty$$

Proof. *If v is the CIR-process this is [9, Lemma 3.1]. In the CEV-case the claim follows from Lemma 3.1.* ■

Theorem 5.3. *If Assumption 5.1 holds true and if in the case $\gamma = 1/2$ Assumption 2.1 is satisfied, then we have*

$$E(f(S_T)) = E\left(\frac{F(S_T)}{S_T} \cdot \left(1 + \frac{1}{\sqrt{1-\rho^2}T} \cdot \int_0^T \frac{1}{\sqrt{v_s}}\,\mathrm{d}W_r^2\right)\right)$$

Proof. *(1) First assume that f is bounded and that there exists an $\varepsilon > 0$ such that $f(x) = 0$ for all $x < 3\varepsilon$. Moreover, let $\psi = \psi_\varepsilon$ be a function such that*

(i) $\psi\colon \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ is continuously differentiable,

(ii) $\psi(x) = 1$ for $x > 2\varepsilon$,

(ii) $\psi(x) = 0$ *for* $x < \varepsilon$.

Under our assumptions Theorem 9.2 together with Theorem 9.3 prove that $F(S_T)$ *belongs to* $\mathbb{D}^{1,2}$ *with derivative* $DF(S_T) = f(S_T) \cdot DS_T$. *Thus*

$$\begin{aligned} E(f(S_T)) &= \frac{1}{T} \cdot E\left(\int_0^T f(S_T) \cdot D_r^2 S_T \cdot \frac{1}{D_r^2 S_T} \mathrm{d}r \right) \\ &= \frac{1}{T} \cdot E\left(\int_0^T f(S_T) D_r^2 S_T \cdot \frac{1}{D_r^2 S_T} \psi(S_T) \mathrm{d}r \right) \\ &= \frac{1}{T} \cdot E\left(\int_0^T D_r^2 (F(S_T)) \cdot \frac{1}{D_r^2 S_T} \psi(S_T) \mathrm{d}r \right) \end{aligned}$$

Now we can use the integration by parts rule of Malliavin calculus and Theorems 4.3 and 4.5 to obtain

$$\begin{aligned} E(f(S_T)) &= \frac{1}{T} \cdot E\left(F(S_T) \cdot \delta\left(\frac{\psi(S_T)}{D_r^2 S_T} \right) \right) \\ &= \frac{1}{T} \cdot E\left(F(S_T) \cdot \frac{1}{\sqrt{1-\rho^2}} \cdot \delta\left(\frac{\psi(S_T)}{S_T} \cdot \frac{1}{\sqrt{v_r}} \right) \right) \end{aligned}$$

where the Skorohod integral integrates over r *and with respect to* W^2. *In order to calculate this integral, we use Proposition 1.3.3 from [26], whose prerequisites we will check now (see Lemma 9.5 in the appendix of this manuscript):*

(i) Due to our assumptions we have $S_T \in \mathbb{D}^{1,2}$. *Since* $\mathbb{R}_{\geq 0} \ni x \mapsto \frac{\psi(x)}{x} \in \mathbb{R}$ *is bounded with bounded first derivative, the chain rule implies that* $\frac{\psi(S_T)}{S_T} \in \mathbb{D}^{1,2}$.

(ii) Proposition 2.10 or 3.7 and the discussion of $\mathbb{L}^{1,2}$ *in [26, p. 42] imply that* $1/\sqrt{v_\cdot} \in \operatorname{dom} \delta$.

(iii) $\psi(S_T)/(S_T\sqrt{v_\cdot}) \in L^2(\Omega; H)$: *Here we have*

$$E \int_0^T \left(\frac{\psi(S_T)}{S_T \sqrt{v_r}} \right)^2 \mathrm{d}r \leq \varepsilon^{-2} \cdot E\left(\int_0^T \frac{1}{v_r} \, \mathrm{d}r \right) < \infty$$

by Lemma 5.2.

(iv) $\psi(S_T)/S_T \cdot \delta(1/\sqrt{v}) \in L^2(\Omega)$ *follows similarly using the Itō-isometry because the Skorohod integral equals* $\int_0^T 1/\sqrt{v_r} \, \mathrm{d}W_r^2$.

Since moreover

$$\begin{aligned} \left\langle D_\cdot^2 \left(\frac{\psi(S_T)}{S_T} \right), \frac{1}{\sqrt{v_\cdot}} \right\rangle_H &= \left\langle \left(\frac{\psi'(S_T)}{S_T} - \frac{\psi(S_T)}{S_T^2} \right) D_\cdot^2 S_T, \frac{1}{\sqrt{v_\cdot}} \right\rangle_H \\ &= \sqrt{1-\rho^2} \left\langle \psi'(S_T)\sqrt{v_\cdot}, \frac{1}{\sqrt{v_\cdot}} \right\rangle - \sqrt{1-\rho^2} \left\langle \frac{\psi(S_T)}{S_T} \sqrt{v_\cdot}, \frac{1}{\sqrt{v_\cdot}} \right\rangle \\ &= \sqrt{1-\rho^2} T \cdot \left(\psi'(S_T) - \frac{\psi(S_T)}{S_T} \right), \end{aligned}$$

which is square-integrable by our assumptions, we have

$$\delta\left(\frac{\psi(S_T)}{S_T}\cdot\frac{1}{\sqrt{v_\cdot}}\right)=\frac{\psi(S_T)}{S_T}\cdot\int_0^T\frac{1}{\sqrt{v_r}}\,\mathrm{d}W_r^2-\sqrt{1-\rho^2}T\cdot\left(\psi'(S_T)-\frac{\psi(S_T)}{S_T}\right)$$

Hence we have shown that

$$\begin{aligned}E(f(S_T))=&E\left(F(S_T)\cdot\frac{1}{\sqrt{1-\rho^2}T}\cdot\frac{\psi(S_T)}{S_T}\cdot\int_0^T\frac{1}{\sqrt{v_r}}\,\mathrm{d}W_r^2\right)\\&-E\left(F(S_T)\cdot\left(\psi'(S_T)-\frac{\psi(S_T)}{S_T}\right)\right)\end{aligned}$$

But since $F(x)=0$ for $x\leq 3\varepsilon$ and $\psi(x)=1$ for $x>2\varepsilon$ we have

$$F(S_T)\frac{\psi(S_T)}{S_T}=\frac{F(S_T)}{S_T}=F(S_T)\cdot\left(\frac{\psi(S_T)}{S_T}-\psi'(S_T)\right)$$

and the assertion follows.

(2) Now let f be bounded. Then there exists functions f_n such that $|f_n(x)|\leq|f(x)|$ for $x\in\mathbb{R}_{\geq 0}$, $f_n(x)=0$ for $x<1/n$ and $f_n\to f$ Lebesque-almost everywhere. From (1) we have

$$E(f_n(S_T))=E\left(\frac{F_n(S_T)}{S_T}\cdot\Pi\right)$$

where $F_n(x):=\int_0^x f_n(z)\mathrm{d}z$ and the weight

$$\Pi:=1+(\sqrt{1-\rho^2}T)^{-1}\int_0^T\sqrt{v_s}^{-1}\mathrm{d}W_s^2$$

is in $L^q(\Omega)$ for all $q\geq 0$ by the Burkholder-Davis-Gundy inequality and Lemma 5.2. Since $f_n\to f$ almost everywhere and $|f_n|\leq|f|$, we have $F_n\to F$ almost everywhere due to dominated convergence. Moreover, $\sup_{x\in\mathbb{R}_{\geq 0}}|F_n(x)/x|$ is uniformly bounded in $n\in\mathbb{N}$.

Using that S_T is absolute continuous with respect to the Lebesgue measure by Theorem 9.3 we have by dominated convergence

$$E(f(S_T))=\lim_{n\to\infty}E(f_n(S_T))=\lim_{n\to\infty}E\left(\frac{F_n(S_T)}{S_T}\cdot\Pi\right)=E\left(\frac{F(S_T)}{S_T}\cdot\Pi\right)$$

(3) Now consider the general case. Choose bounded functions f_n such that $|f_n|\leq|f|$ and $f_n\to f$ almost everywhere. Define $F_n(x):=\int_0^x f_n(z)\mathrm{d}z$ and $P(x):=\int_0^x p(z)\mathrm{d}z$. Then $F_n\to F$ almost everywhere due to dominated convergence. By our assumptions we have $|f_n(S_T)|\leq|f(S_T)|\leq|p(S_T)|\in L^1(\Omega)$ and, because $P(x)/x$ is a polynomial of degree $\deg p$,

$$\left|\frac{F_n(S_T)}{S_T}\right|\leq\left|\frac{F(S_T)}{S_T}\right|\leq\left|\frac{P(S_T)}{S_T}\right|\in L^{1+\varepsilon/\deg p}(\Omega)$$

Since the weight Π is in $L^q(\Omega)$ for all $q\geq 0$ we obtain

$$\frac{F(S_T)}{S_T}\cdot\Pi\in L^1(\Omega)$$

Again the assertion follows by dominated convergence. ■

6. Multidimensional Heston Models. In this section we will extend our quadrature formula of Theorem 5.3 to functionals in the multidimensional Heston model, i.e. to (S^i, v^i), $i = 1, \dots, d$, given by the SDEs

$$\begin{pmatrix} \mathrm{d}S_t^i \\ \mathrm{d}v_t^i \end{pmatrix} = \begin{pmatrix} b_i & S_t^i \\ 0 & \kappa_i(\lambda_i - v_t^i) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sqrt{v_t^i} S_t^i & 0 \\ 0 & \theta_i (v_t^i)^{\gamma_i} \end{pmatrix} \begin{pmatrix} \mathrm{d}B_t^i \\ \mathrm{d}W_t^i \end{pmatrix} \tag{6.1}$$

with $b_i \in \mathbb{R}$, $\kappa_i, \lambda_i, \theta_i > 0$ and $\gamma_i \in [1/2, 1)$ for $i = 1, \dots, d$. Here B^i, W^i, $i = 1, \dots, d$, are (possibly) correlated Brownian motions.

It remains to specify the covariance matrix of (B, W), i.e.

$$\Sigma^{(B,W)} = \begin{pmatrix} (EB_1^i B_1^j)_{i,j=1,\dots,d} & (EB_1^i W_1^j)_{i,j=1,\dots,d} \\ (EW_1^i B_1^j)_{i,j=1,\dots,d} & (EW_1^i W_1^j)_{i,j=1,\dots,d} \end{pmatrix}$$

Assumption 6.1.

(a) Assume $\frac{2\kappa_i\lambda_i}{\theta_i^2} > 1$ *or* $\gamma_i > 1/2$ *for all* $i = 1, \dots, d$.

(b) Assume that $\Sigma^{(B,W)}$ *is positive definite.*

Under assumption (b), we can find an upper $2d \times 2d$-triangular matrix R with positive values on the diagonal such that

$$Z := R^{-1} \begin{pmatrix} B \\ W \end{pmatrix}$$

is a standard $2d$-dimensional Brownian motion, i.e. the components of Z are independent Brownian motions. For this define the linear transformation

$$g \colon \mathbb{R}^{2d,2d} \to \mathbb{R}^{2d,2d}, \qquad g(A)_{ij} := A_{2d+1-i,\, 2d+1-j}, \qquad i, j = 1, \dots, 2d$$

Then $g(\Sigma) := g(\Sigma^{(B,W)})$ is the covariance matrix of $(W_n, \dots, W_1, B_n, \dots, B_1)$, and let $LL^T = g(\Sigma)$ be its Cholesky decomposition. Then $R = g(L)$ is an upper triangular matrix and fulfills $RR^T = \Sigma^{(B,W)}$.

Lemma 6.2. *Let Z be defined as above and $S_t^i \in L^{2+\varepsilon}$, $i = 1, \dots, d$. The Brownian motion Z^1 is independent of $(S_t^j)_{t\in[0,T]}$ if and only if $j \neq 1$. Using the partial Malliavin derivative D^{Z^1} defined in the appendix we have in particular*

$$D_r^{Z^1} X_t^j = \begin{cases} R_{11} \cdot \sqrt{v_r^1} \cdot 1_{[0,t]}(r) & j = 1 \\ 0 & j \neq 1 \end{cases} \tag{6.2}$$

and

$$D_r^{Z^1} S_t^j = \begin{cases} S_t^1 \cdot R_{11} \cdot \sqrt{v_r^1} \cdot 1_{[0,t]}(r) & j = 1 \\ 0 & j \neq 1 \end{cases} \tag{6.3}$$

Proof. By construction we have that Z^1 and $B^j, W^i, j = 2, \dots, d, i = 1, \dots, d$, are independent, which implies the first claim.

These independence relations allow us to easily calculate Malliavin derivatives with respect to Z^1 (see the appendix for this type of derivative).

$$D_r^{Z^1} B_t^1 = \sum_{k=1}^{2d} R_{1,k} \cdot D_r^{Z^1} Z_t^k = R_{11} \cdot 1_{[0,t]}(r)$$

Similarly we obtain

$$D^{Z^1}(B_t^j) = 0 \quad \text{for } j \neq 1$$
$$D^{Z^1}(W_t^i) = 0 \quad \text{for } i = 1, \dots, d$$

Now we can easiliy calculate the derivatives of the price processes. ■

An application of Theorem 4.5 and Theorem 9.3 shows that for fixed $i = 1, \dots, d$ the Heston price S_t^i is absolute continuous with respect to the Lebesgue measure if $S_t^i \in L^{2+\varepsilon}(\Omega)$. For our multidimensional quadrature formula we also need the existence of a joint density.

Theorem 6.3. *Let $t \in [0, T[$ and $S_t^i \in L^{2+\varepsilon}(\Omega)$, $i = 1, \dots, d$. Then the law of the random vector S_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.*

Proof. It suffices to prove that the law of X_t is absolutely continuous. This assertion follows from [26, Lemma 2.1.1], if we can show that

$$|E(\partial_i \varphi(X_t))| \leq c \cdot \frac{1}{t} \cdot \|\varphi\|_\infty$$

for all $\varphi \in C_b^\infty(\mathbb{R}^d; \mathbb{R})$ and each $i = 1, \dots, d$ for some constant $c > 0$.

Case 1: $i = 1$. Applying the standard chain rule (φ is differentiable) and writing D^1 for D^{Z^1} and δ^1 for δ^{Z^1}, we have

$$\begin{aligned} |E(\partial_1 \varphi(X_t))| &= \frac{1}{t} \cdot \left| E\left(\int_0^t \partial_1 \varphi(X_t) \cdot D_r^1 X_t^1 \cdot \frac{1}{D_r^1 X_t^1} \mathrm{d}r \right) \right| \\ &= \frac{1}{t} \cdot \left| E\left(\int_0^t D_r^1 \varphi(X_t) \cdot \frac{1}{D_r^1 X_t^1} \mathrm{d}r \right) \right| \end{aligned}$$

By (6.2) we have $(D_\cdot^1 X_t^1)^{-1} = (R_{11} \sigma_\cdot^1)^{-1}$. As seen in the proof of Theorem 5.3, this is in the domain of δ^1 and thus we can apply the integration by parts rule.

$$\begin{aligned} &= \frac{1}{t} \cdot \left| E\left(\varphi(X_t) \cdot \delta^i\left(\frac{1}{D^1 X_t^1} \right) \right) \right| \\ &\leq \frac{1}{t} \cdot \|\varphi\|_\infty \cdot E\left| \delta^1\left(\frac{1}{D^1 X_t^1} \right) \right| \end{aligned}$$

Case 2: $i \neq 1$. This can be shown analogously to the first case by using a reordering of the Brownian motions so that B^i is the first one.

■

Now we will extend the quadrature formula to the multidimensional setting. For notational convenience we restrict ourselves to the smoothing of the first component.

Assumption 6.4. *Assume that $f\colon (\mathbb{R}_{\geq 0})^n \to \mathbb{R}$ is a measurable function, which is almost everywhere continuous and is bounded by a polynomial p. Furthermore let $F\colon (\mathbb{R}_{\geq 0})^n \to \mathbb{R}$ be given by $F(x) = \int_0^{x_1} f(\xi, x_2, \ldots, x_n)\mathrm{d}\xi$ for all $x \in (\mathbb{R}_{\geq 0})^n$. Finally assume that there is an $\varepsilon > 0$ such that $S_T^i \in L^{\max\{2,\, \deg p\}+\varepsilon}$ for all $i = 1, \ldots, d$.*

Theorem 6.5. *If Assumptions 6.1 and 6.4 hold true, we have*

$$E(f(S_T)) = E\left(\frac{F(S_T)}{S_T^1} \cdot \left(1 + \frac{1}{R_{11}T} \cdot \int_0^T \frac{1}{\sqrt{v_s^1}}\, \mathrm{d}Z_r^1\right)\right)$$

Proof. First assume again that f is bounded and that there exists an $\varepsilon > 0$ such that $f(x) = 0$ for all $|x| < 3\varepsilon$. Moreover, let $\psi = \psi_\varepsilon$ be again the smooth localizing function with $\psi(x) = 0$ for $x < \varepsilon$ and $\psi(x) = 0$ for $x > 2\varepsilon$ from the proof of Theorem 5.3. By Lemma 6.2 the first Brownian motion Z^1 is independent of all price processes but the first. To simplify notation we will write D for the partial derivative D^{Z^1} and δ for δ^{Z^1}. The previous theorem allows us to apply the chain rule Proposition 9.6 to $F(S_T)$. Using the integration by parts rule of Malliavin calculus and formula (6.3) we obtain

$$\begin{aligned}
E(f(S_T)) &= \frac{1}{T} \cdot E\left(\int_0^T f(S_T) \cdot D_r S_T^1 \cdot \frac{1}{D_r S_T^1} \mathrm{d}r\right) \\
&= \frac{1}{T} \cdot E\left(\int_0^T D_r(F(S_T)) \cdot \frac{\psi(S_T^1)}{D_r S_T^1} \mathrm{d}r\right) \\
&= \frac{1}{T} \cdot E\left(F(S_T) \cdot \delta\left(\frac{\psi(S_T^1)}{D_r S_T^1}\right)\right) \\
&= \frac{1}{T} \cdot E\left(F(S_T) \cdot \frac{1}{R_{11}} \cdot \delta\left(\frac{\psi(S_T^1)}{S_T^1} \cdot \frac{1}{\sqrt{v_r^1}}\right)\right)
\end{aligned}$$

where the Skorohod integral integrates over r. To compute the Skorohod integral we can use the same proof as in Theorem 5.3 but using D^{Z^1} instead of D^2 and

$$D\left(\frac{\psi(S_T^1)}{S_T^1}\right) = \left(\frac{\psi'(S_T^1)}{S_T^1} - \frac{\psi(S_T^1)}{(S_T^1)^2}\right) \cdot DS_T^1$$

The extension to general f is exactly the same as in Theorem 5.3. ■

In the case of a so called basket option

$$f(S_T) = g\left(\sum_{i=1}^d a_i S_T^i\right)$$

with $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ and $a_i \geq 0$, $i = 1, \ldots, d$, the function $F\colon (\mathbb{R}_{\geq 0})^d \to \mathbb{R}$ is given by

$$F(x) = \frac{1}{a_1}\left(G\left(\sum_{i=1}^d a_i x_i\right) - G\left(\sum_{i=2}^d a_i x_i\right)\right)$$

with an arbitrary antiderivative G of g. For options of the form

$$f(S_T) = g(S_T^1)h(S_T^2, \dots, S_T^d)$$

with $h\colon (\mathbb{R}_{\geq 0})^{d-1} \to \mathbb{R}$ we get

$$F(x) = G(x_1)h(x^2, \dots, x^d)$$

for $G(x) = \int_0^x g(\xi)\mathrm{d}\xi$. Turning to log-coordinates again, i.e. $X_T^i = \log(S_T^i)$, we have the following smoothness result:

Proposition 6.6. *If $f\colon (\mathbb{R}_{\geq 0})^d \to \mathbb{R}$ is bounded by $C \geq 0$, then*

$$\widetilde{G}\colon \mathbb{R} \to \mathbb{R} \qquad \widetilde{G}(x) := \frac{1}{e^{x_1}} \cdot \int_0^{e^{x_1}} f(\xi, x_2, \dots, x_d)\mathrm{d}\xi$$

is bounded by C and globally Lipschitz-continuous.

Proof. Clearly $\widetilde{G}(x) \leq e^{-x_1} C e^{x_1} = C$. For $x, y \in (\mathbb{R}_{\geq 0})^d$ with $x_1 < y_1$ we have

$$\begin{aligned}
|\widetilde{G}(e^{y_1}) - \widetilde{G}(e^{x_1})| &= \left| \frac{1}{e^{y_1}} \cdot \int_{e^{x_1}}^{e^{y_1}} f(\xi, x_2, \dots, x_d)\mathrm{d}\xi + \left(\frac{1}{e^{y_1}} - \frac{1}{e^{x_1}} \right) \cdot \int_0^{e^{x_1}} f(\xi, x_2, \dots, x_d)\mathrm{d}\xi \right| \\
&\leq \left| \frac{C}{e^{y_1}}(e^{y_1} - e^{x_1}) \right| + \left| \frac{C}{e^{x_1} e^{y_1}}(e^{y_1} - e^{x_1}) e^{x_1} \right| \\
&= 2C(1 - e^{x_1 - y_1}) \\
&\leq 2C(1 - (1 + x_1 - y_1)) \\
&= 2C(y_1 - x_1)
\end{aligned}$$

■

7. Multilevel Monte Carlo. Multilevel Monte Carlo is a powerful method for variance reduction introduced by S. Heinrich [17] for parametric integration and by M. Giles [13] for the quadrature of SDEs.

Assume we want to compute

$$p = Ef(S_T),$$

where $f : (\mathbb{R}_{\geq 0})^d \to \mathbb{R}$ and S_T is given by the d-dimensional generalized Heston model (6.1). While in the one-dimensional case PDE methods, see e.g. [20], or FFT methods [8] can be applied, one has to rely on Monte Carlo methods in higherdimensional settings. The standard Monte Carlo approach uses a simulatable approximation $S_T^{(h)}$ of S_T based on a discretization of the driving Brownian motions with stepsize h and an average of N independent copies of $S_T^{(h)}$ to estimate p, i.e.

$$p \approx \frac{1}{N} \sum_{i=1}^{N} f\big(S_T^{(h),(i)}\big)$$

The multilevel Monte Carlo method instead relies on approximations $S_T^{(h_l)}$ to S_T with different stepsizes $h_l = M^{-l}T$, $l = 0, \dots, L$, with $M \in \mathbb{N}, M \geq 2$. Its idea is to use the

telescoping sum

$$Ef\big(S_T^{(h_L)}\big) = Ef\big(S_T^{(h_0)}\big) + \sum_{l=1}^{L} E\left(f\big(S_T^{(h_l)}\big) - f\big(S_T^{(h_{l-1})}\big)\right)$$

to estimate the expectations on the right hand side using independent standard Monte Carlo estimators

$$Ef\big(S_T^{(h_0)}\big) \approx \widehat{P}_0 = \frac{1}{N_0}\sum_{i=1}^{N_0} f\big(S_T^{(h_0),(i)}\big),$$

$$Ef\big(S_T^{(h_l)}\big) \approx \widehat{P}_l = \frac{1}{N_l}\sum_{i=1}^{N_l}\left(f\big(S_T^{(h_l),(i)}\big) - f\big(S_T^{(h_{l-1}),(i)}\big)\right), \qquad l = 1,\dots,L,$$

and to balance the variance and the computational cost of the summands, i.e. the number of random variables, arithmetic operations and function evaluations, in an optimal way. Here the $S_T^{(h_0),(i)}$ are independent copies of $S_T^{(h_0)}$ and the $(S_T^{(h_l),(i)}, S_T^{(h_{l-1}),(i)})$ are independent copies of $(S_T^{(h_l)}, S_T^{(h_{l-1})})$, which use the same sample paths of the Brownian motion. The multilevel Monte Carlo estimator of the quantity p is then given by

$$\widehat{P} = \sum_{l=0}^{L} \widehat{P}_l,$$

where the estimators $\widehat{P}_l$, $l = 0,\dots,L$, of the different levels have to be independent.

Based on estimates for the weak error

$$\big|Ef(S_T^{(h)}) - p\big| \le c_\alpha \cdot h^\alpha$$

and the strong error

$$E\big|f(S_T^{(h)}) - f(S_T)\big|^2 \le c_\beta \cdot h^{2\beta}$$

with $c_\alpha, c_\beta, \alpha, \beta > 0$, the number of levels L and the number of repetitions N_l, $l = 1,\dots,L$, can be chosen such that the mean square error satisfies

$$E|p - \widehat{P}|^2 \le \varepsilon^2$$

with

$$\text{computational cost} \le c_{\alpha,\beta}\begin{cases} \varepsilon^{-2} & \text{if} \quad \beta > 1 \\ (\log(\varepsilon))^2\varepsilon^{-2} & \text{if} \quad \beta = 1 \\ \varepsilon^{-2-\frac{1-\beta}{\alpha}} & \text{if} \quad \beta < 1 \end{cases}$$

for a given accuracy $\varepsilon > 0$. Here the constant $c_{\alpha,\beta}$ depends only on $\alpha, \beta, c_\alpha, c_\beta$. In particular, the multilevel Monte Carlo estimator outperforms the standard estimator for moderate and high accuracies $\varepsilon > 0$.

The above multilevel estimator has two drawbacks if applied to the generalized Heston model with general payoffs. It relies on good strong convergence properties, i.e. $\beta \ge 1$,

and also on the knowledge of $c_\alpha, c_\beta, \alpha$ and β. However for discontinuous functionals f the strong convergence rates of approximations schemes deteriorate: In [6, 14] it is shown that a discontinuity leads to a halving of the strong convergence rate of the approximation scheme. Also, sharp weak convergence rates for the generalized Heston model have not been established so far up to the best of our knowledge.

Our integration by parts result, i.e. Theorem 6.5, circumvents one of theses problem. Here we have shown that

$$E(f(S_T)) = E\left(\frac{F(S_T)}{S_T^1} \cdot \left(1 + \frac{1}{R_{11}T} \cdot \int_0^T \frac{1}{\sqrt{v_s^1}} \mathrm{d}Z_r^1\right)\right) \tag{7.1}$$

and for several types of functionals f, the arising functional

$$(\mathbb{R}^d_{>0}) \ni x \mapsto \frac{F(x)}{x_1} \in \mathbb{R}$$

is locally or globally Lipschitz, see Section 6. Clearly, in the globally Lipschitz case the strong convergence rate of the approximation scheme for S_T is retained when applied to $F(S_T)/S_T^1$.

Concerning the problem of knowing $c_\alpha, c_\beta, \alpha$ and β, M. Giles presented in [13] a heuristic algorithm, which chooses the N_l's adaptively and can be used without knowledge of $c_\alpha, c_\beta, \alpha$ and β, see Algorithm 1.

Input: error tolerance ε, initial samples N_{ini}
Output: estimation for $E(f(S_T))$
1 $L := 0$
2 Generate $N_L = N_{ini}$ samples for level L
3 Estimate the variance V_L from these samples
4 Compute N_l for $l = 0, \ldots, L$ using the formula $N_l := \left\lceil 2 \cdot \varepsilon^{-2} \sqrt{V_l h_l} \left(\sum_{l=0}^{L} \sqrt{V_l / h_l} \right) \right\rceil$
5 If an N_l surpasses the number of samples generated on level l so far, generate additional samples
6 For $l = 0, \ldots, L$ set $\widehat{P}_l$ to the mean of the samples on level l
7 **if** $L \geq 2$ and the convergence condition $\max\left\{M^{-1}|\widehat{P}_{L-1}|, |\widehat{P}_L|\right\} < \frac{1}{\sqrt{2}}(M-1)\varepsilon$ holds
8 **then return** $\sum_{l=0}^{L} \widehat{P}$
9 **else** $L := L + 1$ and **goto** step 2

Algorithm 1: Adaptive Multilevel Monte Carlo

We will use this algorithm with $M = 2$ and $N_{ini} = 500$.

Before we can finally apply this adaptive multilevel estimator together with a suitable approximation scheme for

$$E\left(\frac{F(S_T)}{S_T^1} \cdot \left(1 + \frac{1}{R_{11}T} \cdot \int_0^T \frac{1}{\sqrt{v_s^1}} \, \mathrm{d}Z_r^1\right)\right)$$

we need to take care of the fact that the variance of the weight

$$\Pi = 1 + \frac{1}{R_{11}T} \cdot \int_0^T \frac{1}{\sqrt{v_s^1}} \, \mathrm{d}Z_r^1$$

is typically very high, because it contains the inverse of the typically low volatility. When applied directly, our multilevel estimator would benefit from the smoothing of the functional only for very small error tolerances ε. To solve this problem we split the payoff in a Lipschitz continuous part g and a discontinuous part h with small support and apply the quadrature formula only to the latter part. To be precise, given $f = g + h$ and H with $H(x) = \int_0^{x_1} h(\xi, x_2, \dots, x_d)\mathrm{d}\xi$, where $f, g, h, H : (\mathbb{R}_{\geq 0})^d \to \mathbb{R}$, we apply the multilevel estimator to the functional

$$g(S_T) + \frac{H(S_T)}{S_T^1} \cdot \Pi \tag{7.2}$$

8. Numerical Results. It remains to specify the numerical methods for the approximation of the Heston prices and the weight Π. As pointed out the diffusion coefficients of the volatility processes do not satisfy the usual global Lipschitz assumption, so the standard theory for the approximation of SDEs (see e.g. [23]) cannot be applied. However, recently ([2, 9, 3, 25]) it has been shown that the discretization of the Lamperti-transformed volatility process with a backward Euler-scheme leads to a positivity preserving scheme that also attains the standard convergence rates from the global Lipschitz case.

So we consider the transformed volatility processes $\sigma_t^i = (v_t^i)^{1-\gamma_i}$ and price processes $X_t^i = \log(S_t^i)$, i.e.

$$\begin{pmatrix} \mathrm{d}X_t^i \\ \mathrm{d}\sigma_t^i \end{pmatrix} = \begin{pmatrix} b_i - \frac{1}{2}(\sigma_t^i)^2 \\ (1-\gamma_i)\left(f_i(\sigma_t^i) - \kappa_i \sigma_t^i\right) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sigma_t^i & 0 \\ 0 & \theta_i(1-\gamma_i) \end{pmatrix} \begin{pmatrix} \mathrm{d}B_t^i \\ \mathrm{d}W_t^i \end{pmatrix}$$

with

$$f_i(x) := \kappa_i \lambda_i x^{-\frac{\gamma_i}{1-\gamma_i}} - \frac{\gamma_i \theta_i^2}{2} x^{-1}$$

Recall that the Brownian motions B^i, W^i are given as $\begin{pmatrix} B \\ W \end{pmatrix} = R \cdot Z$ where Z is a $2d$-dimensional Brownian motion and $R \in \mathbb{R}^{2n \times 2n}$ is an invertible upper triangular matrix. Using these notations we finally set

$$P = g(e^{X_T^1}, \dots, e^{X_T^d}) + \frac{H(e^{X_T^1}, \dots, e^{X_T^d})}{e^{X_T^1}} \cdot \Pi \tag{8.1}$$

with

$$\Pi = 1 + \frac{1}{R_{11}T} \cdot \int_0^T \frac{1}{\sigma_s^1} \mathrm{d}Z_s^1 \tag{8.2}$$

As already mentioned, we discretize the transformed volatility processes with the backward Euler scheme with step size $\Delta > 0$, i.e.

$$\begin{aligned} \hat{\sigma}_{k+1}^i = \hat{\sigma}_k^i &+ (1-\gamma_i)\left(f_i(\hat{\sigma}_{k+1}^i) - \kappa_i \hat{\sigma}_{k+1}^i\right)\Delta \\ &+ (1-\gamma_i)\theta_i(W_{(k+1)\Delta}^i - W_{k\Delta}^i), \qquad k = 0, 1 \dots \end{aligned}$$

where $\hat{\sigma}_0^i = \sigma_0^i$. Under Assumption 6.1(a) this scheme is positivity-preserving, i.e.

$$P\left(\widehat{\sigma}_k^i > 0,\ k = 0, 1, \ldots\right) = 1 \tag{8.3}$$

Moreover, in the case $\gamma = 1/2$ the above implicit equation has the explicit solution

$$\hat{\sigma}_{k+1}^i = \frac{\hat{\sigma}_k^i + \frac{\theta_i}{2}(W_{(k+1)\Delta}^i - W_{k\Delta}^i)}{2 + \kappa_i \Delta} + \sqrt{\frac{\left(\hat{\sigma}_k^i + \frac{\theta_i}{2}(W_{(k+1)\Delta}^i - W_{k\Delta}^i)\right)^2}{(2 + \kappa_i\Delta)^2} + \frac{(\kappa_i\lambda_i - \frac{\theta_i^2}{4})\Delta}{2 + \kappa_i\Delta}},$$

while in the general case it can be solved using standard zero-finding methods. We discretize then the log-price X_T and the weight Π using the Euler scheme, which yields

$$\hat{X}_N^i = \log(S_0^i) + bT - \frac{1}{2}\sum_{k=0}^{N-1}(\hat{\sigma}_k^i)^2\Delta + \sum_{k=0}^{N-1}\hat{\sigma}_k^i\left(B_{(k+1)\Delta}^i - B_{k\Delta}^i\right)$$

and

$$\hat{\Pi} = 1 + \frac{1}{R_{11}T}\sum_{k=0}^{N-1}\frac{1}{\hat{\sigma}_{k\Delta}^i}(Z_{(k+1)\Delta}^1 - Z_{k\Delta}^1)$$

where $N \in \mathbb{N}$ with $N\Delta = T$. Note that our approximation of Π is well defined due to (8.3) and that the results of [9, 3] imply the strong error estimate

$$\left(E\left|X_T^i - \hat{X}_N^i\right|^p\right)^{1/p} \leq K_{p,i} \cdot \sqrt{\Delta} \qquad \text{for} \qquad \begin{cases} 1 \leq p < \frac{2\kappa_i\lambda_i}{\theta_i^2} & \text{if} \quad \gamma_i = \frac{1}{2} \\ 1 \leq p < \infty & \text{else} \end{cases}$$

with $K_{p,i} > 0$. Furthermore set $\hat{S}_T^i := \exp(\hat{X}_T^i)$, $\hat{Y}_T = \sum_{i=1}^d \hat{S}_T^i$ and $Y_T = \sum_{i=1}^d S_T^i$.

As prototype examples for options with discontinuous payoffs we will consider digital options in the one-dimension (extended) Heston model as well as digital basket options in a three-dimensional model. More precisely, we will consider the functional $1_{[0,d\cdot S_0]}$ and approximate $p = E(1_{[0,d\cdot S_0]}(Y_T))$ in the following models:

(a) One-dimensional Heston model. Here the parameters are taken from [1]:

$$T = 2,\ b = 0,\ \kappa = 5.07,\ \lambda = 0.0457,\ \theta = 0.48,\ \rho = -0.767,\ v_0 = \lambda,\ S_0 = 100$$

Note that this choice of parameters ensures that $S_T \in L^{2+\varepsilon}$ for some $\varepsilon > 0$ due to Theorem 4.4.

(b) One-dimensional extended Heston model. We use $T = 2$, $b = 0$, and again parameters from [1]:

$$\kappa = 4.1031,\ \lambda = 0.0451,\ \theta = 0.8583,\ \gamma = 0.6545,\ \rho = -0.760,\ v_0 = \lambda,\ S_0 = 100$$

The correlation is again sufficiently negative to ensure that $S_T \in L^{2+\varepsilon}$.

(c) Three-dimensional Heston model. Again we set $T = 2$, $b_1 = b_2 = b_3 = 0$. Here the parameters are taken from [10]. However, we modified ρ_1, ρ_2 to avoid negative correlations close to -1.

$$\kappa_1 = 1.0121,\ \lambda_1 = 0.2874,\ \theta_1 = 0.7627,\ \rho_1 = -0.7137,\ v_0^1 = 0.2723,\ S_0^1 = 100,$$
$$\kappa_2 = 0.5217,\ \lambda_2 = 0.2038,\ \theta_2 = 0.4611,\ \rho_2 = -0.8322,\ v_0^2 = 0.2536,\ S_0^2 = 100,$$
$$\kappa_3 = 0.5764,\ \lambda_3 = 0.1211,\ \theta_3 = 0.3736,\ \rho_3 = -0.4835,\ v_0^3 = 0.1539,\ S_0^3 = 100$$

The different volatility processes are independent while the different price processes are correlated according to the following correlation matrix

$$\Sigma^S = \begin{pmatrix} 1 & 0.0246 & 0.0598 \\ 0.0246 & 1 & 0.6465 \\ 0.0598 & 0.6465 & 1 \end{pmatrix}$$

Again all prices belong to $L^{2+\varepsilon}$.

(d) Three-dimensional extended Heston model. Because we did not find parameters for this model in the literature, we took the same as in (c), but with $\gamma_1 = 0.63$, $\gamma_2 = 0.68$, $\gamma_3 = 0.71$.

Figures 1 show a comparison of the following three algorithms at different error tolerances ε in these models.

(i) Standard Monte Carlo (using $\lceil \varepsilon^{-2} \rceil$ iterations and a stepsize of $\lceil T/\varepsilon \rceil$)

(ii) Adaptive Multilevel Monte Carlo using the discontinuous payoff directly, i.e. simulating $1_{[0,d\cdot S_0]}(\hat{Y}_T)$

(iii) Adaptive Multilevel Monte Carlo using the quadrature formula and payoff splitting. For a $\delta \in (0,1)$ we use the splitting

$$g(x) = \begin{cases} 1 & x < (1-\delta)K \\ -\frac{1}{2\delta}(x-K) + 0.5 & x \in [(1-\delta)K, (1+\delta)K] \\ 0 & x > (1+\delta)K \end{cases}$$

and $h = 1_{[0,K]} - g$, $H(x) = \int_0^x h(z)\mathrm{d}z$. As discussed before the algorithm then simulates

$$g(\hat{Y}_T) + \frac{H(\hat{Y}_T)}{\hat{S}_T^1} \cdot \hat{\Pi}$$

For each algorithm a series of error tolerances was chosen and the algorithm was executed 500 times for each tolerance. The figures plot the the mean cost against the relative mean-square error, both on logarithmic scale (base 2). Since we the exact value of p is unknown, the reference value p_{ref} has been computed using algorithm (iii) with a very small error tolerance. For each algorithm the figures contain a least squares line that has been fitted to the data points to estimate the convergence rates.

The (random) cost of the multilevel algorithms was measured by the total number of discretization steps, i.e.

$$\text{cost} := d \cdot \left(N_0 + \sum_{l=1}^{L} N_l \cdot (M^l + M^{l-1}) \right) \tag{8.4}$$

where d is the dimension of the model. Note that we account for the additional cost of computing two approximations for each level $l \geq 1$. The mean cost is then given by the average over 500 repetitions of the algorithm. Analogously, the cost of the standard Monte Carlo is also measured by the total number of discretization steps, i.e. $\lceil \varepsilon^{-2} \rceil \cdot \lceil T/\varepsilon \rceil$. Finally, the empirical relative root mean-square error is given by

$$\text{rmsq error} = \sqrt{\frac{1}{500} \sum_{j=1}^{500} \frac{|\widehat{P}^{(j)} - p_{ref}|^2}{p_{ref}^2}}$$

Figure 8.1 shows that algorithm (iii) is between 2 and 4 times faster than algorithm (ii), i.e. acchieves the same accuracy with between one half and one fourth of the computational cost, and both are much faster than the standard Monte Carlo algorithm (i). However, further numerical tests have shown that it is important to choose a good splitting parameter: For the figures the parameter δ which governs the payoff-splitting was set to values between 0.1 and 0.4. Higher values of δ lead in fact to an algorithm worse than algorithm (ii).

Table 8.1 contains the estimated convergence rates. As already mentioned, a detailed error analysis will be carried out in a forthcoming project.

Algorithm	Model (a)	Model (b)	Model (c)	Model (d)
(i) Monte Carlo	0.333	0.327	0.335	0.334
(ii) Adaptive MLMC	0.379	0.370	0.353	0.355
(iii) Ad. MLMC+Malliavin Quadrature	0.392	0.385	0.381	0.388

Table 8.1: Estimated convergence rates

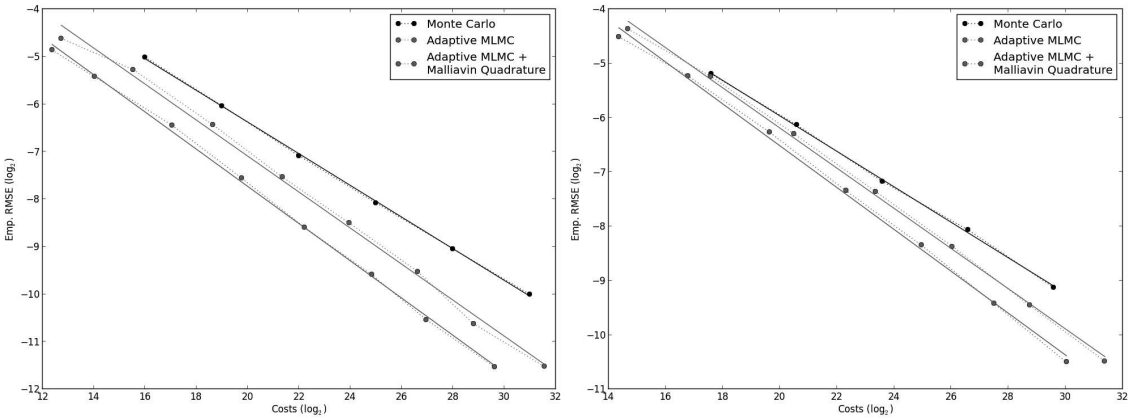
Table 8.2 contains the measured running times for a single run (averaged over 100 runs) using error tolerance $\varepsilon = 2^{-8}$ for algorithms (ii) and (iii) and $\varepsilon = 2^{-7.5}$ for algorithm (i). Using these tolerances all three algorithms given an empirical relative mean-square error of approximately $2^{-7.5}$ and the running times are thus comparable.

Algorithm	Model (a)	Model (b)	Model (c)	Model (d)
(i) Monte Carlo	0.91	5.27	7.84	28.85
(ii) Adaptive MLMC	0.32	1.20	1.47	5.68
(iii) Ad. MLMC+Malliavin Quadrature	0.11	0.62	0.96	2.03

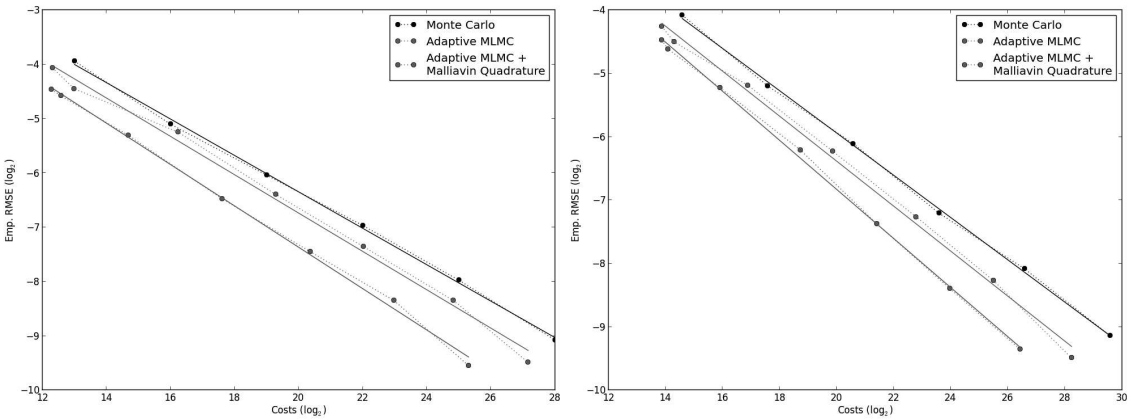
Table 8.2: Measured running times in seconds

Figure 8.1

(a) $d = 1$, $\gamma = 0.5, \delta = 0.3$ (b) $d = 3$, $\gamma = 0.5, \delta = 0.1$



(c) $d = 1$, $\gamma = 0.6545, \delta = 0.2$ (d) $d = 3$, $\gamma = (0.63, 0.68, 0.7), \delta = 0.15$



9. Appendix.

9.1. Comparison Theorem. The following theorem from [22] allows to compare two diffusions with the same diffusion coefficient but different drifts. In Lemma 2.5 and 3.4 we applied it with $h(x) = x^\gamma$ with $\gamma \in [1/2, 1)$.

Theorem 9.1 (Prop. 5.2.18 in [22]). *For $j \in \{1, 2\}$ let X^j be a continuous adapted process such that*

$$X_t^j = X_0^j + \int_0^t b_j(s, X_s^j)\mathrm{d}s + \int_0^t \sigma(s, X_s^j)\mathrm{d}W_s, \qquad t \in \mathbb{R}_{\geq 0}$$

Assume that the following conditions hold:

(i) the coefficients $\sigma(t, x)$ and $b_j(t, x)$ are continuous functions on $\mathbb{R}_{\geq 0} \times \mathbb{R}$,

(ii) $\sigma(t, x)$ satisfies $|\sigma(t, x) - \sigma(t, y)| \leq h(|x - y|)$ for every $t \geq 0$ and all $x, y \in \mathbb{R}$ and a

strictly increasing function $h\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *with* $h(0) = 0$ *and*

$$\int_0^\varepsilon h^{-2}(u)\mathrm{d}u = \infty$$

for all $\varepsilon > 0$,

(iii) $X_0^1 \leq X_0^2$ *almost surely,*

(iv) $b_1(t,x) \leq b_2(t,x)$ *for all* $t \geq 0$, $x \in \mathbb{R}$,

(v) either b_1 *or* b_2 *is Lipschitz continuous, i.e. for* $j = 1$ *or* $j = 2$ *there exists a* $K > 0$ *such that* $|b_j(t,x) - b_j(t,y)| \leq K \cdot |x - y|$ *for all* $x, y \in \mathbb{R}$

Then $P(X_t^1 \leq X_t^2 \ \forall\, t \geq 0) = 1$.

9.2. Malliavin Calculus. The Malliavin calculus extends stochastic analysis by adding a derivative operator, which can be interpreted to measure the influence of the underlying Brownian motion at a specific time on the differentiated random variable.

Let $W = (W^1, \ldots, W^d)$ be a d-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and fix an endtime $T > 0$. Denote by $C^\infty_{\mathrm{pol}}(\mathbb{R}^n; \mathbb{R})$ the functions $f\colon \mathbb{R}^n \to \mathbb{R}$ which are infinitely often differentiable such that f and all of its partial derivatives of any order are bounded by polynomials. We call a random variable X *smooth*, if it can be written as

$$X = f\left(\int_0^T h_1(s)\mathrm{d}W_s, \ldots, \int_0^T h_n(s)\mathrm{d}W_s\right)$$

with $f \in C^\infty_{\mathrm{pol}}(\mathbb{R}^n; \mathbb{R})$ and $h_i \in H := L^2([0,T]; \mathbb{R}^d)$ for $i = 1, \ldots, n$. The set of smooth random variables is denoted by $\mathcal{S}$.

For this class of random variables the Malliavin derivative is defined by the H-valued random variable

$$DX = \sum_{i=1}^n \partial_i f\left(\int_0^T h_1(s)\mathrm{d}W_s, \ldots, \int_0^T h_n(s)\mathrm{d}W_s\right) \cdot h_i$$

In particular $DW_t^i = 1_{[0,t]}(\cdot) \cdot e_i$ where e_i is the i-th unit vector in $\mathbb{R}^d$. It can be shown that this definition is independent of the representation of X in terms of the used function f and the Wiener integrals $\int_0^T h(s)dW_s$. Furthermore, the above operator is closable from $\mathcal{S} \subset L^p(\Omega, \mathcal{F}, P)$ to $L^p(\Omega, \mathcal{F}, P; H)$ for every $p \geq 1$ and the final Malliavin derivative is defined as the closure of the above operator. The class of differentiable random variables $\mathbb{D}^{1,p}$ is then given by

$$\mathbb{D}^{1,p} = \left\{ X \in L^p(\Omega) : \exists (X_n)_{n\in\mathbb{N}} \subset \mathcal{S}, Y \in L^p(\Omega; H) \text{ s.t. } \begin{array}{c} X_n \to X \text{ in } L^p(\Omega) \\ DX_n \to Y \text{ in } L^p(\Omega; H) \end{array} \right\}$$

and the derivative $DX \in L^p(\Omega; H)$ is defined as the limit $Y = \lim_{n\to\infty} DX_n$ for any approximating sequence X_n (the closability of the operator precisely guarantees that this is well-defined). The derivative is a d-dimensional random function and we write $D_r^i X$ for the i-th component evaluated at $r \in [0,T]$. Note that while the set of differentiable random variables depends on the chosen L^p-norm, the actual derivative does not depend on it. In any

case only random variables which are measurable with respect to the underlying Brownian motions can be treated in Malliavin calculus. We denote by $\mathbb{D}^{1,\infty} := \cap_{p\geq 1}\mathbb{D}^{1,p}$ the set of random variables differentiable w.r.t. all $p \geq 1$.

A main tool for the derivation of the quadrature rule is the chain rule. Because we apply it to the antiderivative of a discontinuous function, it is important that the chain rule does not require the function to be differentiable.

Theorem 9.2. *Let $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ be a locally Lipschitz function. Assume $X = (X_1, \dots, X_m)$ is a random vector with components in $\mathbb{D}^{1,p}$. If*

- *$\varphi(X) \in L^p(\Omega)$,*
- *$\partial_i\varphi(X) \cdot DX_i \in L^p(\Omega; H)$ for each $i = 1, \dots, m$,*
- *and the set of $x \in \mathbb{R}^m$ where φ is not partially continuously differentiable is a P_X-null set,*

then the chain rule holds: $\varphi(X) \in \mathbb{D}^{1,p}$ and its derivative is given by

$$D\varphi(X) = \sum_{i=1}^{n} \partial_i\varphi(X) \cdot DX_i$$

Here we define $\partial_i\varphi := 0$ on the set of points, where $\partial_i\varphi$ does not exist.

Note that by the Theorem of Rademacher [11, §5.8.3, Theorem 6] this set is a Lebesgue-null set and hence the third assumption is always fulfilled, if the law of X is absolute continuous with respect to the Lebesgue measure.

Proof. The derivative is a closed operator, so in order to prove the chain rule for a function φ, it is sufficient to find functions φ_n such that $\varphi_n(X) \to \varphi(X)$ in $L^p(\Omega)$, the chain rule holds for all φ_n and $\partial_i\varphi_n(X) \cdot DX_i \to \partial_i\varphi(X) \cdot DX_i$ in $L^p(\Omega; H)$ for all $i = 1, \dots, m$.

In the first step assume that φ is globally Lipschitz and bounded. For each $n \in \mathbb{N}$ choose a function $\psi_n\colon \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ such that $\psi_n \in C^\infty$, $\operatorname{supp}\psi_n \subset [-1/n, 1/n]^m$ and $\int \psi_n \mathrm{d}x = 1$ (see for example [28, pp. 82-83]). Define $\varphi_n := \varphi * \psi_n$. It is easily shown that $\varphi_n(X) \to \varphi(X)$ in $L^p(\Omega)$ and that $\partial_i\varphi_n$ is bounded by the Lipschitz-constant of φ. By Proposition 1.2.3 in [26] the chain rule holds for functions which are continuously differentiable with bounded partial derivatives and thus for all φ_n. Furthermore

$$|\partial_i\varphi_n(X) - \partial_i\varphi(X)| \leq \int_{\mathbb{R}^m} |\partial_i\varphi(y) - \partial_i\varphi(X)| \cdot \psi_n(X - y)\mathrm{d}y$$

For all $\omega \in \Omega$ such that $\partial_i\varphi$ is continuous at $X(\omega)$ this converges to zero. By dominated convergence we have $\partial_i\varphi_n(X) \cdot DX_i \to \partial_i\varphi(X) \cdot DX_i$ in $L^p(\Omega; H)$ and the chain rule is shown.

In the next step assume that φ is only locally Lipschitz, but bounded. Choose an increasing sequence of natural numbers n_k such that $P(\|X\|_\infty = n_k) = 0$ for all $k \in \mathbb{N}$. Define $\varphi_k(x) := \varphi(-n_k \vee x \wedge n_k)$. Note that by construction the set of $x \in \mathbb{R}^m$ where φ_k is not differentiable, is still a P_X-null set. By the first part of the proof the chain rule holds for φ_k. Because φ is bounded we have $\varphi_k(X) \to \varphi(X)$ in $L^p(\Omega)$ and by dominated convergence $\partial_i\varphi_k(X) \cdot DX_i \to \partial_i\varphi(X) \cdot DX_i$ in $L^p(\Omega; H)$. This proves the chain rule for φ.

To extend the chain rule finally to unbounded functions choose an increasing sequence of natural numbers n_k such that $P(|\varphi(X)| = n_k) = 0$ and use the approximation $\varphi_k(x) := -n \vee \varphi(x) \wedge n$ in a similar way as above. ■

To check the third condition of the previous theorem, the following result is particularly useful (Theorem 2.1.3 in [26]):

Theorem 9.3. *Let $X \in \mathbb{D}^{1,p}$ for some $p \geq 1$ and assume $\|DX\|_H > 0$ almost surely. Then X is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

Finally, we need a theorem to compute derivatives of stochastic processes given by SDEs.

Theorem 9.4. *Let W_t, $t \in [0,T]$, be an m-dimensional Brownian motion and let X_t, $t \in [0,T]$, be given by the d-dimensional SDE*

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

Assume that both $b\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are continuously differentiable with bounded derivative. Then $X^i(t) \in \mathbb{D}^{1,\infty}$ for all $t \in [0,T]$, $i = 1,\dots,d$. The derivative $D_r^j X^i(t)$ satisfies almost everywhere the equation

$$D_r^j X^i(t) = \sigma_{ij}(X(r)) + \sum_{k=1}^{d}\sum_{l=1}^{m}\int_r^t \partial_k\sigma_{il}(s) D_r^j X^k(s)\mathrm{d}W^l(s) + \sum_{k=1}^{d}\int_r^t \partial_k b_i(X(s)) D_r^j X^k(s)\mathrm{d}s$$

for $r \leq t$ and the equation $D_r^j X^i(t) = 0$ for $r > t$.

Proof. This is a simplified version of Theorem 2.2.1 in [26] together with the remark following that theorem. ■

The second important operator in Malliavin calculus is the divergence operator which in the case of an underlying Brownian motion is called *Skorohod integral*. It is denoted by δ and defined as adjoint operator of the derivative on $\mathbb{D}^{1,2}$, i.e. the domain is given by

$$\operatorname{dom}\delta := \left\{u \in L^2(\Omega;H) : X \mapsto \langle DX, u\rangle_{L^2(\Omega;H)} \text{ is continuous on } \mathbb{D}^{1,2}\right\}$$

and on this set the Skorohod integral is defined (by Riesz' representation theorem) as the unique $\delta(u) \in L^2(\Omega)$ such that

$$\langle DX, u\rangle_{L^2(\Omega;H)} = \langle X, \delta(u)\rangle_{L^2(\Omega)} \qquad \forall\, X \in \mathbb{D}^{1,2}$$

This is also known as the integration by parts rule of Malliavin calculus and can be written as

$$E\left(\int_0^T \langle D_r(X), u(r)\rangle_{\mathbb{R}^d}\,\mathrm{d}r\right) = E(X\cdot\delta(u)) \qquad \forall\, X \in \mathbb{D}^{1,2},\ u \in \operatorname{dom}\delta$$

Further analysis shows that for an adapted process u the Skorohod integral $\delta(u)$ coincides with $\int_0^T u(s)\mathrm{d}W_s$ and hence it can be viewed as an extension of the Itō-integral.

To compute Skorohod integrals of non-adapted integrands, the following lemma is useful:

Lemma 9.5 (Prop. 1.3.3 in [26]). *Let $F \in L^2(\Omega)$ and $u \in L^2(\Omega;H)$ such that the following conditions hold:*

(i) $F \in \mathbb{D}^{1,2}$,

(ii) $u \in \operatorname{dom}\delta$,

(iii) $Fu \in L^2(\Omega; H)$,
(iv) $F\delta(u) - \langle DF, u\rangle_H \in L^2(\Omega)$.
Then $Fu \in \operatorname{dom}\delta$ *and* $\delta(Fu) = F\delta(u) - \langle DF, u\rangle_H$.

To derive the quadrature formula for a large class of multidimensional functionals in Theorem 6.5, we will need a derivative operator that differentiates only with respect to one specific Brownian motion. Let $a \in \mathbb{R}^d$ with $\|a\|_2 = 1$. Then $B_t := \langle W_t, a\rangle_{\mathbb{R}^d}$ defines a Brownian motion. For $X \in \mathbb{D}^{1,p}$ we define the derivative with respect to B as the $L^2([0,T])$-valued random variable given by

$$D^B_r X = \langle D_r X, a\rangle_{\mathbb{R}^d}, \qquad r \in [0,T]$$

On random variables which are measurable w.r.t. to B this derivative coincides with the usual derivative in the one-dimensional Malliavin calculus w.r.t. B. If $X \in L^p(\Omega)$ is independent of B, then $D^B = 0$.

The adjoint of D^B is defined for all $u \in L^p(\Omega; L^2([0,T]))$ such that $u \cdot a \in \operatorname{dom}\delta$. For such u it is given by $\delta^B(u) = \delta(u \cdot a)$. The chain rule Theorem 9.2 and Lemma 9.5 are easily extended to D^B and δ^B.

The chain rule can even be extended to functions which are not locally Lipschitz continuous in some directions, as long as the random variables in those directions are independent of B:

Proposition 9.6. *Let B be a Brownian motion on $(\Omega, \mathcal{F}, P)$ and $X = (X_1, \dots, X_m)$ a random vector with components in $\mathbb{D}^{1,p}$ and set $A := \{i = 1, \dots, m : D^B X_i \neq 0\}$. Assume $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ is continuous except on a P_X-zero set and assume that φ is locally Lipschitz continuous in all directions e_i, $i \in A$: For all $i \in A$, $x \in \mathbb{R}^d$ there exists a constant $L_i(x)$ and $\epsilon_i(x) > 0$ such that $|\varphi(x + he_i) - \varphi(x)| \leq L_i \cdot h$ for all $0 \leq h \leq \epsilon_i(x)$. Then $\varphi(X) \in \mathbb{D}^{1,p}$ and its derivative is given by*

$$D^B\varphi(X) = \sum_{i \in A} \partial_i\varphi(X) \cdot D^B X_i$$

Proof. First assume that φ is bounded and globally Lipschitz continuous in all directions e_i, $i \in A$. We will reuse the mollifier functions from the proof of Theorem 9.2. Because the chain rule holds for $\varphi * \psi_n$ we only have to prove $\varphi * \psi_n(X) \to \varphi(X)$ in $L^p(\Omega)$ and $\partial_i\varphi * \psi_n(X) \cdot D^B X_i \to \partial_i\varphi(X) \cdot D^B X_i$ in $L^p(\Omega; L^2([0,T]))$ for $i \in A$. As in the proof of Theorem 9.2 we have

$$|\varphi * \psi_n(X) - \varphi(X)|^p \to 0,$$

because φ is continuous P_X-almost everywhere. Moreover, this expression is bounded by $2^p \sup_{x \in \mathbb{R}^m} |\varphi(x)|^p$, so we have also convergence to 0 in $L^p(\Omega)$. Furthermore

$$\left|(\partial_i(\varphi * \psi_n)(X) - \partial_i\varphi(X)) \cdot D^B X_i\right|^p \leq \int_{\mathbb{R}^d} |\partial_i\varphi(y) - \partial_i\varphi(X)|^p \cdot \psi_n(X - y)^p \cdot |D^B X_i|^p \mathrm{d}y$$

This converges to 0 for all ω such that $\partial_i\varphi$ is continuous at $X(\omega)$. By dominated convergence we also have convergence in $L^p(\Omega; H)$ and the assertion is proven.

The extension to unbounded and locally Lipschitz functions is the same as in Theorem 9.2.

■

REFERENCES

[1] Y. Aït-Sahalia and R. Kimmel, *Maximum likelihood estimation of stochastic volatility models*, J. Financial Economics, 83 (2007), pp. 413–452.

[2] A. Alfonsi, *On the discretization schemes for the CIR (and Bessel squared) processes*, Monte Carlo Methods Appl., 11 (2005), pp. 355–384.

[3] ———, *Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process*, Statist. Probab. Let., 83 (2013), pp. 602–607.

[4] E. Alòs and C.-O. Ewald, *Malliavin differentiability of the Heston volatility and applications to option pricing*, Adv. Appl. Probab., 40 (2008), pp. 144–162.

[5] L.B.G. Andersen and V.V. Piterbarg, *Moment explosions in stochastic volatility models*, Finance Stoch., 11 (2006), pp. 29–50.

[6] R. Avikainen, *On irregular functionals of SDEs and the Euler scheme*, Finance Stoch., 13 (2009), pp. 381–401.

[7] A. Berkaoui, M. Bossy, and A. Diop, *Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence*, ESAIM Probab. Stat., 12 (2007), pp. 1–11.

[8] P. Carr, D.B. Madan, and R.H. Smith, *Option valuation using the fast Fourier transform*, J. Comput. Finance, 2 (1999), pp. 61–73.

[9] S. Dereich, A. Neuenkirch, and L. Szpruch, *An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 468 (2012), pp. 1105–1115.

[10] G. Dimitroff, S. Lorenz, and A. Szimayer, *A parsimonious multi-asset Heston model: Calibration and derivative pricing*, Int. J. Theor. Appl. Finance, 14 (2011), pp. 1299–1333.

[11] L.C. Evans, *Partial Differential Equations*, Amer. Math. Soc., Providence, Rhode Island, 2nd ed., 2010.

[12] P.K. Friz and M. Keller-Ressel, *Moment explosions in stochastic volatility models*, in Encyclopedia of Quantitative Finance, John Wiley & Sons, Ltd, 2010.

[13] M.B. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.

[14] M.B. Giles, D.J. Higham, and X. Mao, *Analysing multilevel Monte Carlo for options with non-globally Lipschitz payoff*, Finance Stoch., 13 (2009), pp. 403–413.

[15] E. Gobet, E. Benhamou, and M. Miri, *Time dependent Heston model*, SIAM J. Financial Math., 1 (2010), pp. 289–325.

[16] L.A. Grzelak and C.W. Oosterlee, *On the Heston model with stochastic interest rates*, SIAM J. Financial Math., 2 (2011), pp. 255–286.

[17] S. Heinrich, *Multilevel Monte Carlo methods*, in Large-scale scientific computing, Margenov, S. (ed.) et al., Lect. Notes Comput. Sci. 2179, 2001, pp. 58–67.

[18] S.L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Financial Studies, 6 (1993), pp. 327–343.

[19] T.R. Hurd and A. Kuznetsov, *Explicit formulas for Laplace transforms of stochastic integrals*, Markov Process. Related Fields, 14 (2008), pp. 277–290.

[20] K.J. in't Hout and S. Foulon, *ADI finite difference schemes for option pricing in the Heston model with correlation*, Int. J. Numer. Anal. Model, 7 (2010), pp. 303–320.

[21] C. Kahl, M. Günther, and T. Rossberg, *Structure preserving stochastic integration schemes in interest rate derivative modeling.*, Appl. Numer. Math., 58 (2008), pp. 284–295.

[22] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, N.Y., 2nd ed., 2010.

[23] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.

[24] R. Lord, R. Koekkoek, and D. Van Dijk, *A comparison of biased simulation schemes for stochastic volatility models*, Quant. Finance, 10 (2010), pp. 177–194.

[25] A. Neuenkirch and L. Szpruch, *First order strong approximations of scalar SDEs with values in a domain*, arXiv:1209.0390, (2012).

[26] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, Berlin, 2nd ed., 2006.

[27] W. Schoutens, E. Simons, and J. Tistaert, *A perfect calibration! Now what?*, Wilmott Magazine, (March 2004).

[28] Dirk Werner, *Funktionalanalysis*, Springer-Verlag Berlin Heidelberg, 6th ed., 2007.

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.

[55] A. Zeiser. Wavelet Approximation in Weighted Sobolev Spaces of Mixed Order with Applications to the Electronic Schrödinger Equation. Preprint 55, DFG-SPP 1324, July 2010.

[56] G. Kutyniok, J. Lemvig, and W.-Q Lim. Compactly Supported Shearlets. Preprint 56, DFG-SPP 1324, July 2010.

[57] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, July 2010.

[58] S. Jokar. Sparse Recovery and Kronecker Products. Preprint 58, DFG-SPP 1324, August 2010.

[59] T. Aboiyar, E. H. Georgoulis, and A. Iske. Adaptive ADER Methods Using Kernel-Based Polyharmonic Spline WENO Reconstruction. Preprint 59, DFG-SPP 1324, August 2010.

[60] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the Convergence of Generalized Polynomial Chaos Expansions. Preprint 60, DFG-SPP 1324, August 2010.

[61] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds of Tensors of Fixed TT-Rank. Preprint 61, DFG-SPP 1324, September 2010.

[62] J. Ballani, L. Grasedyck, and M. Kluge. Black Box Approximation of Tensors in Hierarchical Tucker Format. Preprint 62, DFG-SPP 1324, October 2010.

[63] M. Hansen. On Tensor Products of Quasi-Banach Spaces. Preprint 63, DFG-SPP 1324, October 2010.

[64] S. Dahlke, G. Steidl, and G. Teschke. Shearlet Coorbit Spaces: Compactly Supported Analyzing Shearlets, Traces and Embeddings. Preprint 64, DFG-SPP 1324, October 2010.

[65] W. Hackbusch. Tensorisation of Vectors and their Efficient Convolution. Preprint 65, DFG-SPP 1324, November 2010.

[66] P. A. Cioica, S. Dahlke, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Spatial Besov Regularity for Stochastic Partial Differential Equations on Lipschitz Domains. Preprint 66, DFG-SPP 1324, November 2010.

[67] E. Novak and H. Woźniakowski. On the Power of Function Values for the Approximation Problem in Various Settings. Preprint 67, DFG-SPP 1324, November 2010.

[68] A. Hinrichs, E. Novak, and H. Woźniakowski. The Curse of Dimensionality for Monotone and Convex Functions of Many Variables. Preprint 68, DFG-SPP 1324, November 2010.

[69] G. Kutyniok and W.-Q Lim. Image Separation Using Shearlets. Preprint 69, DFG-SPP 1324, November 2010.

[70] B. Jin and P. Maass. An Analysis of Electrical Impedance Tomography with Applications to Tikhonov Regularization. Preprint 70, DFG-SPP 1324, December 2010.

[71] S. Holtz, T. Rohwedder, and R. Schneider. The Alternating Linear Scheme for Tensor Optimisation in the TT Format. Preprint 71, DFG-SPP 1324, December 2010.

[72] T. Müller-Gronbach and K. Ritter. A Local Refinement Strategy for Constructive Quantization of Scalar SDEs. Preprint 72, DFG-SPP 1324, December 2010.

[73] T. Rohwedder and R. Schneider. An Analysis for the DIIS Acceleration Method used in Quantum Chemistry Calculations. Preprint 73, DFG-SPP 1324, December 2010.

[74] C. Bender and J. Steiner. Least-Squares Monte Carlo for Backward SDEs. Preprint 74, DFG-SPP 1324, December 2010.

[75] C. Bender. Primal and Dual Pricing of Multiple Exercise Options in Continuous Time. Preprint 75, DFG-SPP 1324, December 2010.

[76] H. Harbrecht, M. Peters, and R. Schneider. On the Low-rank Approximation by the Pivoted Cholesky Decomposition. Preprint 76, DFG-SPP 1324, December 2010.

[77] P. A. Cioica, S. Dahlke, N. Döhring, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Adaptive Wavelet Methods for Elliptic Stochastic Partial Differential Equations. Preprint 77, DFG-SPP 1324, January 2011.

[78] G. Plonka, S. Tenorth, and A. Iske. Optimal Representation of Piecewise Hölder Smooth Bivariate Functions by the Easy Path Wavelet Transform. Preprint 78, DFG-SPP 1324, January 2011.

[79] A. Mugler and H.-J. Starkloff. On Elliptic Partial Differential Equations with Random Coefficients. Preprint 79, DFG-SPP 1324, January 2011.

[80] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. A Derandomization of the Euler Scheme for Scalar Stochastic Differential Equations. Preprint 80, DFG-SPP 1324, January 2011.

[81] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. Preprint 81, DFG-SPP 1324, January 2011.

[82] K. Grella and C. Schwab. Sparse Tensor Spherical Harmonics Approximation in Radiative Transfer. Preprint 82, DFG-SPP 1324, January 2011.

[83] D.A. Lorenz, S. Schiffler, and D. Trede. Beyond Convergence Rates: Exact Inversion With Tikhonov Regularization With Sparsity Constraints. Preprint 83, DFG-SPP 1324, January 2011.

[84] S. Dereich, M. Scheutzow, and R. Schottstedt. Constructive quantization: Approximation by empirical measures. Preprint 84, DFG-SPP 1324, January 2011.

[85] S. Dahlke and W. Sickel. On Besov Regularity of Solutions to Nonlinear Elliptic Partial Differential Equations. Preprint 85, DFG-SPP 1324, January 2011.

[86] S. Dahlke, U. Friedrich, P. Maass, T. Raasch, and R.A. Ressel. An adaptive wavelet method for parameter identification problems in parabolic partial differential equations. Preprint 86, DFG-SPP 1324, January 2011.

[87] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and Variational Stabilization for Convection-Diffusion Equations. Preprint 87, DFG-SPP 1324, January 2011.

[88] T. Jahnke. On Reduced Models for the Chemical Master Equation. Preprint 88, DFG-SPP 1324, January 2011.

[89] P. Binev, W. Dahmen, R. DeVore, P. Lamby, D. Savu, and R. Sharpley. Compressed Sensing and Electron Microscopy. Preprint 89, DFG-SPP 1324, March 2011.

[90] P. Binev, F. Blanco-Silva, D. Blom, W. Dahmen, P. Lamby, R. Sharpley, and T. Vogt. High Quality Image Formation by Nonlocal Means Applied to High-Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). Preprint 90, DFG-SPP 1324, March 2011.

[91] R. A. Ressel. A Parameter Identification Problem for a Nonlinear Parabolic Differential Equation. Preprint 91, DFG-SPP 1324, May 2011.

[92] G. Kutyniok. Data Separation by Sparse Representations. Preprint 92, DFG-SPP 1324, May 2011.

[93] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Introduction to Compressed Sensing. Preprint 93, DFG-SPP 1324, May 2011.

[94] H.-C. Kreusler and H. Yserentant. The Mixed Regularity of Electronic Wave Functions in Fractional Order and Weighted Sobolev Spaces. Preprint 94, DFG-SPP 1324, June 2011.

[95] E. Ullmann, H. C. Elman, and O. G. Ernst. Efficient Iterative Solvers for Stochastic Galerkin Discretizations of Log-Transformed Random Diffusion Problems. Preprint 95, DFG-SPP 1324, June 2011.

[96] S. Kunis and I. Melzer. On the Butterfly Sparse Fourier Transform. Preprint 96, DFG-SPP 1324, June 2011.

[97] T. Rohwedder. The Continuous Coupled Cluster Formulation for the Electronic Schrödinger Equation. Preprint 97, DFG-SPP 1324, June 2011.

[98] T. Rohwedder and R. Schneider. Error Estimates for the Coupled Cluster Method. Preprint 98, DFG-SPP 1324, June 2011.

[99] P. A. Cioica and S. Dahlke. Spatial Besov Regularity for Semilinear Stochastic Partial Differential Equations on Bounded Lipschitz Domains. Preprint 99, DFG-SPP 1324, July 2011.

[100] L. Grasedyck and W. Hackbusch. An Introduction to Hierarchical (H-) Rank and TT-Rank of Tensors with Examples. Preprint 100, DFG-SPP 1324, August 2011.

[101] N. Chegini, S. Dahlke, U. Friedrich, and R. Stevenson. Piecewise Tensor Product Wavelet Bases by Extensions and Approximation Rates. Preprint 101, DFG-SPP 1324, September 2011.

[102] S. Dahlke, P. Oswald, and T. Raasch. A Note on Quarkonial Systems and Multilevel Partition of Unity Methods. Preprint 102, DFG-SPP 1324, September 2011.

[103] A. Uschmajew. Local Convergence of the Alternating Least Squares Algorithm For Canonical Tensor Approximation. Preprint 103, DFG-SPP 1324, September 2011.

[104] S. Kvaal. Multiconfigurational time-dependent Hartree method for describing particle loss due to absorbing boundary conditions. Preprint 104, DFG-SPP 1324, September 2011.

[105] M. Guillemard and A. Iske. On Groupoid C*-Algebras, Persistent Homology and Time-Frequency Analysis. Preprint 105, DFG-SPP 1324, September 2011.

[106] A. Hinrichs, E. Novak, and H. Woźniakowski. Discontinuous information in the worst case and randomized settings. Preprint 106, DFG-SPP 1324, September 2011.

[107] M. Espig, W. Hackbusch, A. Litvinenko, H. Matthies, and E. Zander. Efficient Analysis of High Dimensional Data in Tensor Formats. Preprint 107, DFG-SPP 1324, September 2011.

[108] M. Espig, W. Hackbusch, S. Handschuh, and R. Schneider. Optimization Problems in Contracted Tensor Networks. Preprint 108, DFG-SPP 1324, October 2011.

[109] S. Dereich, T. Müller-Gronbach, and K. Ritter. On the Complexity of Computing Quadrature Formulas for SDEs. Preprint 109, DFG-SPP 1324, October 2011.

[110] D. Belomestny. Solving optimal stopping problems by empirical dual optimization and penalization. Preprint 110, DFG-SPP 1324, November 2011.

[111] D. Belomestny and J. Schoenmakers. Multilevel dual approach for pricing American style derivatives. Preprint 111, DFG-SPP 1324, November 2011.

[112] T. Rohwedder and A. Uschmajew. Local convergence of alternating schemes for optimization of convex problems in the TT format. Preprint 112, DFG-SPP 1324, December 2011.

[113] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. Preprint 113, DFG-SPP 1324, December 2011.

[114] Y. Dong, T. Görner, and S. Kunis. An iterative reconstruction scheme for photoacoustic imaging. Preprint 114, DFG-SPP 1324, December 2011.

[115] L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along generated sets. Preprint 115, DFG-SPP 1324, February 2012.

[116] H. Chen and R. Schneider. Numerical analysis of augmented plane waves methods for full-potential electronic structure calculations. Preprint 116, DFG-SPP 1324, February 2012.

[117] J. Ma, G. Plonka, and M.Y. Hussaini. Compressive Video Sampling with Approximate Message Passing Decoding. Preprint 117, DFG-SPP 1324, February 2012.

[118] D. Heinen and G. Plonka. Wavelet shrinkage on paths for scattered data denoising. Preprint 118, DFG-SPP 1324, February 2012.

[119] T. Jahnke and M. Kreim. Error bound for piecewise deterministic processes modeling stochastic reaction systems. Preprint 119, DFG-SPP 1324, March 2012.

[120] C. Bender and J. Steiner. A-posteriori estimates for backward SDEs. Preprint 120, DFG-SPP 1324, April 2012.

[121] M. Espig, W. Hackbusch, A. Litvinenkoy, H.G. Matthiesy, and P. Wähnert. Effcient low-rank approximation of the stochastic Galerkin matrix in tensor formats. Preprint 121, DFG-SPP 1324, May 2012.

[122] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. Preprint 122, DFG-SPP 1324, June 2012.

[123] A. Mugler and H.-J. Starkloff. On the convergence of the stochastic Galerkin method for random elliptic partial differential equations. Preprint 123, DFG-SPP 1324, June 2012.

[124] P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R.L. Schilling. On the convergence analysis of Rothe's method. Preprint 124, DFG-SPP 1324, July 2012.

[125] P. Binev, A. Cohen, W. Dahmen, and R. DeVore. Classification Algorithms using Adaptive Partitioning. Preprint 125, DFG-SPP 1324, July 2012.

[126] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation of hierarchical Tucker and Tensor-Train tensors. Preprint 126, DFG-SPP 1324, July 2012.

[127] M. Kovács, S. Larsson, and K. Urban. On Wavelet-Galerkin methods for semilinear parabolic equations with additive noise. Preprint 127, DFG-SPP 1324, August 2012.

[128] M. Bachmayr, H. Chen, and R. Schneider. Numerical analysis of Gaussian approximations in quantum chemistry. Preprint 128, DFG-SPP 1324, August 2012.

[129] D. Rudolf. Explicit error bounds for Markov chain Monte Carlo. Preprint 129, DFG-SPP 1324, August 2012.

[130] P.A. Cioica, K.-H. Kim, K. Lee, and F. Lindner. On the $L_q(L_p)$-regularity and Besov smoothness of stochastic parabolic equations on bounded Lipschitz domains. Preprint 130, DFG-SPP 1324, December 2012.

[131] M. Hansen. $n-$term Approximation Rates and Besov Regularity for Elliptic PDEs on Polyhedral Domains. Preprint 131, DFG-SPP 1324, December 2012.

[132] R. E. Bank and H. Yserentant. On the H^1 -stability of the L_2 -projection onto finite element spaces. Preprint 132, DFG-SPP 1324, December 2012.

[133] M. Gnewuch, S. Mayer, and K. Ritter. On Weighted Hilbert Spaces and Integration of Functions of Infinitely Many Variables. Preprint 133, DFG-SPP 1324, December 2012.

[134] D. Crisan, J. Diehl, P.K. Friz, and H. Oberhauser. Robust Filtering: Correlated Noise and Multidimensional Observation. Preprint 134, DFG-SPP 1324, January 2013.

[135] Wolfgang Dahmen, Christian Plesken, and Gerrit Welper. Double Greedy Algorithms: Reduced Basis Methods for Transport Dominated Problems. Preprint 135, DFG-SPP 1324, February 2013.

[136] Aicke Hinrichs, Erich Novak, Mario Ullrich, and Henryk Wozniakowski. The Curse of Dimensionality for Numerical Integration of Smooth Functions. Preprint 136, DFG-SPP 1324, February 2013.

[137] Markus Bachmayr, Wolfgang Dahmen, Ronald DeVore, and Lars Grasedyck. Approximation of High-Dimensional Rank One Tensors. Preprint 137, DFG-SPP 1324, March 2013.

[138] Markus Bachmayr and Wolfgang Dahmen. Adaptive Near-Optimal Rank Tensor Approximation for High-Dimensional Operator Equations. Preprint 138, DFG-SPP 1324, April 2013.

[139] Felix Lindner. Singular Behavior of the Solution to the Stochastic Heat Equation on a Polygonal Domain. Preprint 139, DFG-SPP 1324, May 2013.

[140] Stephan Dahlke, Dominik Lellek, Shiu Hong Lui, and Rob Stevenson. Adaptive Wavelet Schwarz Methods for the Navier-Stokes Equation. Preprint 140, DFG-SPP 1324, May 2013.

[141] Jonas Ballani and Lars Grasedyck. Tree Adaptive Approximation in the Hierarchical Tensor Format. Preprint 141, DFG-SPP 1324, June 2013.

[142] Harry Yserentant. A short theory of the Rayleigh-Ritz method. Preprint 142, DFG-SPP 1324, July 2013.

[143] M. Hefter and K. Ritter. On Embeddings of Weighted Tensor Product Hilbert Spaces. Preprint 143, DFG-SPP 1324, August 2013.

[144] M. Altmayer and A. Neuenkirch. Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts. Preprint 144, DFG-SPP 1324, August 2013.