Kolmogorov Backward Equations with Singular Diffusion Matrices

Hermann Singer

Diskussionsbeitrag Nr. 518

Dezember 2019

Kolmogorov Backward Equations with Singular Diffusion Matrices

Hermann Singer

December 12, 2019

Abstract The Feynman-Kac formula is derived from the Kolmogorov formula by using a state extension, which permits the absorption of the potential term into the drift part of the backward equation. The resulting homogenous equation with singular diffusion matrix can be solved by using importance sampling techniques based on a drift correction. The method is applied to the Cameron-Martin formula, the quantum mechanical linear oscillator and the pricing of arithmetic Asian options.

Keywords Feynman-Kac formula; Integrated variables; Monte Carlo method; Importance sampling; Drift correction; Stochastic differential equations

FernUniversität in Hagen,
Fakultät für Wirtschaftswissenschaft,
Lehrstuhl für angewandte Statistik
und Methoden der empirischen Sozialforschung,
D-58084 Hagen, Germany
E-mail: Hermann.Singer@FernUni-Hagen.de

1 Introduction

Solutions of deterministic partial differential equations can be written as functional (path) integrals, which may in some cases be interpreted as probabilistic expressions [1, 31], [29, chs. 4, 5.7]. This was first demonstrated by Feynman in 1942 for the case of the Schrödinger equation [5, 12, 13]. Introducing imaginary time, a rigorous probabilistic theory was developed by Kac [25, 26, 27]. Later, the option pricing theory of Black, Scholes and Merton [3, 37, 38] lead to a partial differential equation, which may be solved by classical methods such as separation of variables, or by the discounted expectation value of a terminal condition [8]. This is especially intuitive to economists, who accomplish pricing by considering the mean of future cash flows. One advantage of probabilistic solutions of deterministic equations is the possibility of approximations by statistical averages including standard error estimates [2, 7, 15, 21, 48]. Furthermore, stochastic methods work well in higher dimensions, when spatial discretization methods suffer from excessive computational demand.

The main device used in this paper is the reduction of the backward equation with potential term to a homogenous equation by using a state extension [47, exercise 2.26, p. 421]. Then, the well known Kolmogorov formula implies the Feynman-Kac formula. Consequently, importance sampling (IS) techniques [30, 35, 41, 50, 55] for the former turn over to the latter equation. Moreover, models with integrated variables occurring, for example, in the case of memory effects [23] or in Asian options [16], are included as a special case. A drawback is the occurrence of a singular diffusion matrix, which requires special attention.

In the field of rare events simulation, adaptive importance sampling methods have been used [9, 20, 28, 52, 54, 57]. In this context, a duality between an optimal control problem and importance sampling can be derived. A solution is given in closed form in terms of a path integral. The measure transformation involved in Girsanov's theorem introduces an additional drift term in the stochastic differential equation (SDE) which is used to evaluate the Feynman-Kac formula. This may be seen as a controlled version of the original SDE. The optimal choice of the control function in explicit form was given by Milstein [41, ch. 12] and Kloeden and Platen [30, ch. 16.2], but it requires knowledge of the solution.

In an application context, approximations of the solution must be computed [40], either iteratively or by using parametrized solutions of simpler problems. An adaptive iterative approach uses the Kullback-Leibler entropy [28, 58].

In the present paper, analytically known solutions of the inhomogenous backward equation, e.g. the Black-Scholes formula or Gaussian solutions for the Schrödinger equation with zero potential are used in order to implement suboptimal importance sampling. Nevertheless, substantial variance reduction can be achieved.

The article is organized as follows: after the definition of backward equations (Section 2) the main result is derived in Section 3. Then, integrated

variables are introduced in Section 4 and importance sampling is applied to the extended problem (Section 5). Finally in Section 6, we discuss applications to the multivariate Cameron-Martin formula, the Schrödinger equation with quadratic potential and to arithmetic Asian options.

2 Backward equations

The inhomogenous backward equation for the scalar concentration (density) field $C(\boldsymbol{x}, t)$ is given by

$$C_t(\boldsymbol{x}, t) + L(\boldsymbol{x}, t)C(\boldsymbol{x}, t) + v(\boldsymbol{x}, t)C(\boldsymbol{x}, t) = 0, \tag{1}$$

$t \leq T$, with terminal condition $C(\boldsymbol{x}, T) = h(\boldsymbol{x})$. In Eq. (1), the partial differential operator

$$\begin{aligned}
L(\boldsymbol{x}, t) &= f_\alpha(\boldsymbol{x}, t)\partial_\alpha + \tfrac{1}{2}\Omega_{\alpha\beta}(\boldsymbol{x}, t)\partial_\alpha\partial_\beta \\
&= \boldsymbol{f}^\mathsf{T}(\boldsymbol{x}, t)\nabla + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\Omega}(\boldsymbol{x}, t)\nabla\nabla^\mathsf{T}]
\end{aligned} \tag{2}$$

is the backward operator w.r.t. the 'spatial' variables $\boldsymbol{x} = [x_1, ..., x_p]^\mathsf{T} \in \mathbb{R}^{p \times 1}$, where $^\mathsf{T}$ denotes the transpose of a vector or matrix, $t \in \mathbb{R}$ is time, $\partial_\alpha := \partial/\partial x_\alpha$, $\partial_y := \partial/\partial y$ and $C_t := \partial C(\boldsymbol{x}, t)/\partial t$, $C_{\boldsymbol{x}} := \partial C(\boldsymbol{x}, t)/\partial \boldsymbol{x} = \nabla C(\boldsymbol{x}, t)$, with $\nabla := [\partial_1, ..., \partial_p]^\mathsf{T}$, the gradient (nabla) operator. The drift vector is $\boldsymbol{f} = f_\alpha, \alpha = 1, ..., p$ and the symmetric diffusion matrix is $\boldsymbol{\Omega} = \Omega_{\alpha\beta}; \alpha, \beta = 1, ..., p$. We assume a sum convention $f_\alpha g_\alpha = \sum_\alpha f_\alpha g_\alpha = \boldsymbol{f}^\mathsf{T}\boldsymbol{g}$, and $\mathrm{tr}[\boldsymbol{A}] = A_{\alpha\alpha}$ is the trace of a square matrix.

The backward equation (1) is a generalized diffusion (transport) equation including a drift term $\boldsymbol{f}$ and a source (potential) term $v\,C$ [1, 29, 31, 50]. In a more narrow sense, only the equation without potential term is a Kolmogorov backward equation, but this name is also used for the general case [30, chs. 4.8, 17.1]. Likewise, it is called a Cauchy problem for the backward heat equation with potential. Moreover, a term $g(\boldsymbol{x}, t)$ can be included [29, chs. 4.2, 5.7]. In this paper, the potential term is called an inhomogeneity, although it contains the solution C. Equations including stochastic terms are discussed by Holden et al [22].

More generally, one can also consider so-called quasi-linear equations, where the potential term is a nonlinear function of C and $C_{\boldsymbol{x}}$ [10, 39, 45, 46]. The solution can be represented by a pair of forward-backward stochastic differential equations (BSDE). For a linear function, the Feynman-Kac formula (see Sec. 3) is obtained as a special case. Adaptive importance sampling schemes based on BSDE equations are considered by Gobet and Turkedjiev [17, 18], Gobet et al [19].

3 Kolmogorov- and Feynman-Kac formula

A well known stochastic representation of the solution $C(\boldsymbol{x}, t)$ of Eq. (1) is given by the conditional expectation value (Feynman-Kac formula)

$$C(\boldsymbol{x}, t) = E\left[e^{\int_t^T v(\boldsymbol{X}(\tau), \tau) d\tau} h(\boldsymbol{X}(T)) \,\Big|\, \boldsymbol{X}(t) = \boldsymbol{x}\right], \tag{3}$$

where $\boldsymbol{X}(\tau) \in \mathbb{R}^p, t \leq \tau \leq T$ is a vector stochastic process fulfilling the Itô stochastic differential equation

$$d\boldsymbol{X}(\tau) = \boldsymbol{f}(\boldsymbol{X}, \tau) d\tau + \boldsymbol{G}(\boldsymbol{X}, \tau) d\boldsymbol{W}(\tau) \tag{4}$$

with initial condition $\boldsymbol{X}(t) = \boldsymbol{x}$, volatility matrix $\boldsymbol{G} : p \times r$, $\boldsymbol{G}\boldsymbol{G}^\mathsf{T} = \boldsymbol{\Omega}$, and $d\boldsymbol{W}(\tau)$ denotes the increments of an r-dimensional Wiener process, i.e. $E[W_\alpha(t)] = 0, E[W_\alpha(t) \, W_\beta(t')] = \delta_{\alpha\beta} \min(t, t'); \; \alpha, \beta = 1, ..., r$. For a proof, see Karatzas and Shreve [29, ch. 5.7] and Appendix A. The (formal) derivative $\zeta_\alpha(t) = dW_\alpha(t)/dt$ is a white noise process with covariance function $E[\zeta_\alpha(t)\zeta_\beta(t')] = \delta(t - t')\delta_{\alpha\beta}$. Here, $\delta(t - t')$ is Dirac's delta function [32] and $\delta_{\alpha\beta}$ denotes the Kronecker delta symbol.

This is the starting point for Monte Carlo (MC) approximations [2, 16], in which the expectation value in Eq. (3) is replaced by a mean value over independent or dependent replications.

In order to reduce the variance of the MC estimator, importance sampling using a modified drift term can be utilized [35, 36, 40, 41, 50, 51]. This new drift term in Eq. (4) is chosen such that, (important) trajectories $\boldsymbol{X}(\tau)$ are simulated yielding nonzero values of the terminal condition $h(\boldsymbol{X}(T))$ (see Section 5).

In the work of Melchior and Öttinger [35, 36], no potential term was considered. Therefore in the present context, it is desirable to derive a homogenous equation without potential term. To achieve this, the integrated potential term in Eq. (3) can be generated by the system extension

$$dY(\tau) = v(\boldsymbol{X}(\tau), \tau) d\tau \tag{5}$$

with initial condition $Y(t) = 0$ [41, ch. 12-13], [30, ch. 17.1], [42, ch. 7].

For the extended system of equations (4, 5) we can use the Kolmogorov formula [31]

$$C(\boldsymbol{x}, y = 0, t) = E\left[e^{Y(T)} h(\boldsymbol{X}(T)) \,\Big|\, \boldsymbol{X}(t) = \boldsymbol{x}, Y(t) = y = 0\right]. \tag{6}$$

This representation is the solution of the homogenous backward equation

$$C_t(\boldsymbol{x}, y, t) + L(\boldsymbol{x}, y, t)C(\boldsymbol{x}, y, t) = 0, \; t \leq T, \tag{7}$$

with

$$L(\boldsymbol{x}, y, t) = f_\alpha(\boldsymbol{x}, t)\partial_\alpha + v(\boldsymbol{x}, t)\partial_y + \tfrac{1}{2}\Omega_{\alpha\beta}(\boldsymbol{x}, t)\partial_\alpha\partial_\beta \tag{8}$$

and terminal condition $C(\boldsymbol{x}, y, T) = \exp(y)h(\boldsymbol{x})$. Due to the variable extension (the additional variable y is called *integrated* because of Eq. (5)), the extended diffusion matrix $\tilde{\boldsymbol{\Omega}} = \mathrm{diag}(\boldsymbol{\Omega}, 0)$ in Eq. (8) is singular and the drift of the system Eq. (4, 5) is of the form $\tilde{\boldsymbol{f}} = [\boldsymbol{f}^\mathsf{T}, v]^\mathsf{T}$. One has $\tilde{\nabla}^\mathsf{T} := [\nabla^\mathsf{T}, \partial_y] = [\partial_1, ..., \partial_p, \partial_y]$; $\tilde{\boldsymbol{f}}^\mathsf{T}\tilde{\nabla} = \boldsymbol{f}^\mathsf{T}\nabla + v\,\partial_y$; $\mathrm{tr}[\tilde{\boldsymbol{\Omega}}\tilde{\nabla}\tilde{\nabla}^\mathsf{T}] = \mathrm{tr}[\boldsymbol{\Omega}\nabla\nabla^\mathsf{T}]$.

With the ansatz $C(\boldsymbol{x}, y, t) = \exp(y)C(\boldsymbol{x}, t)$ one recovers the inhomogenous equation (1), since

$$v(\boldsymbol{x}, t)\partial_y C(\boldsymbol{x}, y, t) = v(\boldsymbol{x}, t)C(\boldsymbol{x}, y, t),$$

and the term $\exp(y)$ can be dropped. Thus one can identify

$$C(\boldsymbol{x}, y = 0,\, t) = C(\boldsymbol{x}, t)$$
$$C(\boldsymbol{x}, y = 0, T) = C(\boldsymbol{x}, T) = h(\boldsymbol{x}) \text{ (terminal condition)}.$$

In summary, one has derived the Feynman-Kac formula Eq. (3) for the inhomogenous backward equation (1) by using the Kolmogorov formula for the extended system Eq. (4, 5) and the corresponding homogenous backward equation (7).

The main advantage of the extended system is the possibility to use variance reduction techniques for the Kolmogorov formula, even in the presence of an inhomogeneity (potential term) in Eq. (1).

A slight drawback is the higher dimensionality $p + 1$ of the extended diffusion process $[\boldsymbol{X}^\mathsf{T}(\tau), Y(\tau)]^\mathsf{T}$ and the singular diffusion matrix $\tilde{\boldsymbol{\Omega}}$ (see also Section 4). Equations with zero diffusion matrix correspond to deterministic motions which can be described by the Liouville equation in statistical mechanics [53, p. 411]. Here one has both stochastic motion, Eq. (4), and deterministic motion, Eq. (5), without error term.

The transition density $p(\boldsymbol{x}_J, y_J, T | \boldsymbol{x}, y, t)$, appearing in the conditional expectation Eq. (6), explicitly

$$C(\boldsymbol{x}, y, t) = \int \exp(y_J)h(\boldsymbol{x}_J)p(\boldsymbol{x}_J, y_J, T | \boldsymbol{x}, y, t)d\boldsymbol{x}_J dy_J,$$

is known analytically only for linear systems and some special cases such as the Feller square root process [8, 11]. Therefore, one can insert the Chapman-Kolmogorov equation and use the product representation on the time slices $\tau_j = t + j\delta\tau, j = 0, ...J; \delta\tau = (T - t)/J$

$$p(\boldsymbol{\xi}_J, T | \boldsymbol{\xi}_0, t) = \prod_{j=0}^{J-1} \int p(\boldsymbol{\xi}_{j+1}, \tau_{j+1} | \boldsymbol{\xi}_j, \tau_j)d\boldsymbol{\xi}_{J-1}...d\boldsymbol{\xi}_1, \qquad (9)$$

$\boldsymbol{\xi}_0 = \boldsymbol{\xi} = [\boldsymbol{x}^\mathsf{T}, y]^\mathsf{T}$, $\boldsymbol{\xi}_j = [\boldsymbol{x}_j^\mathsf{T}, y_j]^\mathsf{T}$, $\tau_J = T, \tau_0 = t$. For a small time step $\delta\tau$, one can approximate the transition densities p by the Euler densities (cf. Eq. (4, 5))

$$q(\boldsymbol{\xi}_{j+1}, \tau_{j+1} | \boldsymbol{\xi}_j, \tau_j) = \phi(\boldsymbol{\xi}_{j+1}; \boldsymbol{\xi}_j + \tilde{\boldsymbol{f}}(\boldsymbol{\xi}_j, \tau_j)\delta\tau, \tilde{\boldsymbol{\Omega}}(\boldsymbol{\xi}_j, \tau_j)\delta\tau), \qquad (10)$$

where $\phi(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = |2\pi\boldsymbol{\Sigma}|^{-1/2}\exp[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})]$ is the multivariate Gauss density and $|\cdot|$ is the determinant. In this formula, it is assumed that the covariance matrix $\boldsymbol{\Sigma}$ is nonsingular. The extended drift function is of the form $\tilde{\boldsymbol{f}} = [\boldsymbol{f}^{\mathsf{T}}, v]^{\mathsf{T}}$, but the diffusion matrix $\tilde{\boldsymbol{\Omega}} = \operatorname{diag}(\boldsymbol{\Omega}, 0)$ is singular. In this situation, one may

- either use the singular normal distribution [34, ch. 2.5.4]
- or use a small parameter $\epsilon > 0$ ($\tilde{\boldsymbol{\Omega}} = \operatorname{diag}(\boldsymbol{\Omega}, \epsilon)$) and finally consider the limit $\epsilon \to 0$.

The latter method leads to the desired result Eq. (3), which can be seen as follows. The explicit form of Eq. (10) is

$$q(\boldsymbol{\xi}_{j+1}, \tau_{j+1}|\boldsymbol{\xi}_j, \tau_j) = \phi(\boldsymbol{x}_{j+1}; \boldsymbol{x}_j + \boldsymbol{f}(\boldsymbol{x}_j, \tau_j)\delta\tau, \boldsymbol{\Omega}(\boldsymbol{x}_j, \tau_j)\delta\tau)$$
$$\phi(y_{j+1}; y_j + v(\boldsymbol{x}_j, \tau_j)\delta\tau, \epsilon^2\delta\tau), \qquad (11)$$

where the second line can be considered as a delta sequence $\delta_\epsilon(y_{j+1} - y_j - v(\boldsymbol{x}_j, \tau_j)\delta\tau)$ with the property $\lim_{\epsilon\to 0}\int \delta_\epsilon(x)f(x)dx = f(0)$, where $f(x)$ is a test function. Thus the generalized function $\delta(x)$ is represented by a function sequence $\delta_\epsilon(x)$ [32]. Now the integration over the y-variables in Eq. (9) can be carried out, resulting in the term $y_J = \sum_j v(\boldsymbol{x}_j, \tau_j)\delta\tau$ for $\epsilon \to 0$, which is the Riemann sum of the integral $\int_t^T v(\boldsymbol{x}(\tau), \tau)d\tau$.

In contrast, the singular normal distribution uses the pseudo inverse $\tilde{\boldsymbol{\Omega}}^- = \operatorname{diag}(\boldsymbol{\Omega}^-, 0)$ and the second line in Eq. (11) is 1. Thus the result in Eq. (3) cannot be obtained by using this approach. Therefore in the sequel, we always assume a small parameter ϵ and consider the limit $\epsilon \to 0$ in the final results.

4 Integrated variables

More generally, instead of the scalar process Eq. (5), one may use a q-dimensional vector process

$$d\boldsymbol{Y}(\tau) = \boldsymbol{v}(\boldsymbol{X}(\tau), \tau)d\tau \qquad (12)$$

with initial condition $\boldsymbol{Y}(t) = \boldsymbol{y}$ and consider the Kolmogorov representation

$$C(\boldsymbol{x}, \boldsymbol{y}, t) = E\left[H(\boldsymbol{X}(T), \boldsymbol{Y}(T)) \mid \boldsymbol{X}(t) = \boldsymbol{x}, \boldsymbol{Y}(t) = \boldsymbol{y}\right]$$
$$(13)$$

with terminal condition $C(\boldsymbol{x}, \boldsymbol{y}, T) = H(\boldsymbol{x}, \boldsymbol{y})$. For example, in case of scalar $X(\tau)$ one can write

$$dY_1(\tau) = v(X(\tau), \tau)d\tau$$
$$dY_2(\tau) = X(\tau)d\tau.$$

The integrated variable approach thus permits equations with average values, e.g. memory effects, or arithmetic Asian options with terminal condition $h(y_2)$ (see Section 6.3). In this case, one sets $H(x, y_1, y_2) = \exp(y_1)h(y_2)$ and obtains

$$C(x, y_1 = 0, y_2, t) = E\left[e^{Y_1(T)}h(Y_2(T)) \mid X(t) = x, Y_1(t) = 0, Y_2(t) = y_2\right].$$
(14)

The starting value $Y_2(t) = y_2 = \int_0^t X(\tau)d\tau$ is the cumulated value of the so called *underlying* $X(\tau)$ from 0 to t. At terminal time T, the value of the option is $C(x, 0, y_2, T) = h(y_2) = h(\int_0^T X(\tau)d\tau)$ and thus depends on the time average $\bar{X}_T = T^{-1}\int_0^T X(\tau)d\tau$.

This shows that the system extension Eq. (12) has the double effect, that

1. the original inhomogenous backward equation can be transformed to a homogenous one
2. cumulated variables can be used as well.

The advantage is the possibility of importance sampling for the Kolmogorov formula Eq. (13), as discussed in the next section. A drawback is the occurence of a singular diffusion matrix, but this happens at any rate in the presence of integrated variables. As mentioned above, the problem can be solved by introducing small parameters $\epsilon_\alpha, \alpha = 1, ..., q$, in the Gaussian transition densities Eq. (11), i.e. $\epsilon^2 \to \mathrm{diag}(\epsilon_1^2, ..., \epsilon_q^2)$, and by taking the limit $\epsilon_\alpha \to 0$ in the final results.

5 Importance sampling

The expectation value Eq. (13) can be written explicitly in the form

$$C(\boldsymbol{\xi}, t) = \int H(\boldsymbol{\xi}_J)p(\boldsymbol{\xi}_J, T|\boldsymbol{\xi}, t)d\boldsymbol{\xi}_J$$
(15)

$$= \int H(\boldsymbol{\xi}_J)\frac{p(\boldsymbol{\xi}_J, T|\boldsymbol{\xi}, t)}{p_2(\boldsymbol{\xi}_J, T|\boldsymbol{\xi}, t)}p_2(\boldsymbol{\xi}_J, T|\boldsymbol{\xi}, t)d\boldsymbol{\xi}_J$$
(16)

where $\boldsymbol{\xi} := [\boldsymbol{x}^\mathsf{T}, \boldsymbol{y}^\mathsf{T}]^\mathsf{T}$ and p_2 is an importance density, yielding a smaller variance of the MC estimator. Intuitively, p_2 should attain small values, when $H(\boldsymbol{\xi}_J) \geq 0$ is small, since these values contribute little to the functional. The choice $p_2 \propto Hp$ fulfils this condition and is even optimal [30, ch. 16.3]. However, the normalization $C = \int Hp\, d\boldsymbol{\xi}$ is the desired unknown quantity. Again, the transition densities can be expressed by the Chapman-Kolmogorov formula (9) and subsequent Euler approximation.

Assuming that p_2 is known, a variance reduced estimator of C (see Eq. (16)) can be written in the form

$$\hat{C}(\boldsymbol{\xi}, t; \delta\tau) = N^{-1}\sum_n H(\boldsymbol{\xi}_{nJ})\left(\frac{p}{p_2}\right)(\boldsymbol{\Xi}_n, \boldsymbol{\tau}|\boldsymbol{\xi}, t)$$
(17)

where $\boldsymbol{\Xi}_n = [\boldsymbol{\xi}_{nJ}, ..., \boldsymbol{\xi}_{n1}] \sim p_2$ is a replication of the path $\boldsymbol{\xi}_n(\tau_j); j = 1, ..., J$, starting at $\boldsymbol{\xi}_n(\tau_0) = \boldsymbol{\xi}$, $\tau_0 = t$, and $p = \prod_{j=0}^{J-1} p(\boldsymbol{\xi}_{j+1}, \tau_{j+1}|\boldsymbol{\xi}_j, \tau_j)$, $\boldsymbol{\tau} = [\tau_J, ..., \tau_1]$. The optimal importance density is explicitly given by

$$p_2 = H(\boldsymbol{\xi}_J)p(\boldsymbol{\Xi}, \boldsymbol{\tau}|\boldsymbol{\xi}, t)/C(\boldsymbol{\xi}, t) = p_{\mathrm{opt}}, \tag{18}$$

$\boldsymbol{\Xi} = [\boldsymbol{\xi}_J, ..., \boldsymbol{\xi}_1]$, with corresponding Euler approximation $q_2 = q_{\mathrm{opt}} = Hq/C$. In order to sample paths from p_2, one can find a new drift function $\tilde{\boldsymbol{f}}_2 = \tilde{\boldsymbol{f}}_{\mathrm{opt}}$, so that the density $p(\boldsymbol{\Xi}, \boldsymbol{\tau}|\boldsymbol{\xi}, t; \tilde{\boldsymbol{f}}_2, \tilde{\boldsymbol{\Omega}})$ coincides with p_2. Then, one can simulate independent trajectories $\boldsymbol{\xi}_n \sim p_2$ and compute the variance reduced estimator Eq. (17). Since $\tilde{\boldsymbol{\Omega}}$ remains invariant, the measures (with densities) p and p_2 are absolutely continuous w.r.t. each other and p/p_2 (or q/q_2 in the case of approximate Euler densities) exists in a continuum limit ($\delta\tau \to 0$).

It can be shown [36, 50], that the density Eq. (18) fulfils a backward equation $\partial_t p_2 + L_2(\boldsymbol{\xi}, t)p_2 = 0$ with modified drift $\tilde{\boldsymbol{f}}_2$. The new drift can be written in the form (in the limit $\epsilon \to 0$, $t \le \tau \le T$)

$$\begin{aligned}
\tilde{\boldsymbol{f}}_2(\boldsymbol{\xi}, \tau) &= \tilde{\boldsymbol{f}} + \tilde{\boldsymbol{\Omega}} \, \nabla_{\boldsymbol{\xi}} \log C \\
&= \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{v} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Omega} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \nabla_{\boldsymbol{x}} \log C \\ \nabla_{\boldsymbol{y}} \log C \end{bmatrix} \\
&:= \begin{bmatrix} \boldsymbol{f}_2(\boldsymbol{x}, \boldsymbol{y}, \tau) \\ \boldsymbol{v}(\boldsymbol{x}, \tau) \end{bmatrix}
\end{aligned}$$

[30, ch. 16.2, p. 514] and Appendix B. Thus, the original drift ist distorted (tilted) by an additional term, which produces trajectories with high values of the terminal condition $H(\boldsymbol{x}_J, \boldsymbol{y}_J)$. Clearly, in order to practically compute the importance drift, an approximation of the solution C must be known, e.g. a Gaussian solution for the Schrödinger equation with zero potential (example 6.2) or the Black-Scholes formula (example 6.3). The components of $\tilde{\boldsymbol{f}}_2$ corresponding to the integrated variables are not distorted, since they only depend on the variables $\boldsymbol{x}$, without error term; see Eq. (12). Because of the special form of $\tilde{\boldsymbol{f}}_2$

$$\left(\frac{q}{q_2}\right)(\boldsymbol{\Xi}, \boldsymbol{\tau}|\boldsymbol{\xi}_0, \tau_0) = \exp\left\{\sum_{j=0}^{J-1}(\boldsymbol{f}_j - \boldsymbol{f}_{2j})^{\mathsf{T}}\boldsymbol{\Omega}_j^{-1}[\delta\boldsymbol{x}_j - \tfrac{1}{2}(\boldsymbol{f}_j + \boldsymbol{f}_{2j})\delta\tau]\right\} \tag{19}$$

$\boldsymbol{f}_j := \boldsymbol{f}(\boldsymbol{x}_j, \tau_j)$, $\boldsymbol{f}_{2j} := \boldsymbol{f}_2(\boldsymbol{x}_j, \boldsymbol{y}_j, \tau_j)$, etc. [33, vol. I, p. 297], only depends on the original drift components $\boldsymbol{f}(\boldsymbol{x}, \tau)$ and the drift correction in

$$\begin{aligned}
\boldsymbol{f}_2(\boldsymbol{x}, \boldsymbol{y}, \tau) &= \boldsymbol{f}(\boldsymbol{x}, \tau) + \boldsymbol{\Omega}(\boldsymbol{x}, \tau) \, \nabla_{\boldsymbol{x}} \log C(\boldsymbol{x}, \boldsymbol{y}, \tau) \\
&= \boldsymbol{f}(\boldsymbol{x}, \tau) + \delta\boldsymbol{f}(\boldsymbol{x}, \tau). \tag{20}
\end{aligned}$$

The parts of the density function depending only on $\boldsymbol{y}_j$ and $\boldsymbol{v}(\boldsymbol{x}_j, \tau_j)$ cancel (see Eq. (11), 3rd line), as well as the Jacobian terms $|2\pi\boldsymbol{\Omega}_j|^{-1/2}$. Now, in analogy to the discussion following Eq. (11), the integration in Eq. (16) over the y-variables can be carried out and the delta functions $\delta_\epsilon(\boldsymbol{y}_{j+1} - \boldsymbol{y}_j - \boldsymbol{v}(\boldsymbol{x}_j, \tau_j)\delta\tau)$

generate the integrated variables $\boldsymbol{y}_j = \sum_{l=0}^{j} \boldsymbol{v}(\boldsymbol{x}_l, \tau_l)\delta\tau$ in the limit $\epsilon_\alpha \to 0, \alpha = 1, ..., q$ (see Sections 3, 4). An alternative derivation using the Girsanov theorem is given in Appendix B. Thus the variance reduced MC estimator can be expressed in this limit as

$$\hat{C}(\boldsymbol{x}, \boldsymbol{y}, t; \delta\tau) = N^{-1} \sum_n H(\boldsymbol{x}_{nJ}, \boldsymbol{y}_{nJ}) \left(\frac{q}{q_2}\right) (\boldsymbol{X}_n, \boldsymbol{Y}_n, \boldsymbol{\tau} \mid \boldsymbol{x}, \boldsymbol{y}, t), \quad (21)$$

where $\boldsymbol{X}_n = [\boldsymbol{x}_{nJ}, ..., \boldsymbol{x}_{n1}]$, $\boldsymbol{Y}_n = [\boldsymbol{y}_{nJ}, ..., \boldsymbol{y}_{n1}]$, $\boldsymbol{\tau} = [\tau_J, ..., \tau_1]$. The Euler-Maruyama approximation for the stochastic equations (4, 12) with modified drift $\tilde{\boldsymbol{f}}_2$ reads

$$\boldsymbol{x}_{n,j+1} = \boldsymbol{x}_{nj} + \tilde{\boldsymbol{f}}_2(\boldsymbol{x}_{nj}, \boldsymbol{y}_{nj}, \tau_j)\delta\tau + \boldsymbol{G}(\boldsymbol{x}_{nj}, \boldsymbol{y}_{nj}, \tau_j)\delta\boldsymbol{W}_{nj}$$
$$\boldsymbol{y}_{n,j+1} = \boldsymbol{v}(\boldsymbol{x}_{nj}, \tau_j)\delta\tau,$$

$\boldsymbol{x}_{n0} = \boldsymbol{x}, \boldsymbol{y}_{n0} = \boldsymbol{y}$ and $\delta\boldsymbol{W}_{nj} \sim N(\boldsymbol{0}, \boldsymbol{I}_r\delta\tau)$, independent and identically distributed, and $\boldsymbol{I}_r$ is an r-dimensional unit matrix. The solution Eq. (15) is implicitly contained in the drift correction $\boldsymbol{\Omega}(\boldsymbol{x}, \tau)\nabla_{\boldsymbol{x}} \log C(\boldsymbol{x}, \boldsymbol{y}, \tau), t \leq \tau \leq T$. For example, the solution Eq. (6) can be estimated as

$$\hat{C}(\boldsymbol{x}, t; \delta\tau) = N^{-1} \sum_n e^{y_{nJ}} h(\boldsymbol{x}_{nJ}) \left(\frac{q_1}{q_2}\right) (\boldsymbol{X}_n, \boldsymbol{y}_n, \tau \mid \boldsymbol{x}, y = 0, t), \quad (22)$$

$\boldsymbol{y}_n = [y_{nJ}, ..., y_{n1}]$, and the drift correction is

$$\boldsymbol{\Omega}(\boldsymbol{x}, \tau)\nabla_{\boldsymbol{x}} \log(e^y C(\boldsymbol{x}, \tau)) = \boldsymbol{\Omega}(\boldsymbol{x}, \tau)\nabla_{\boldsymbol{x}} \log C(\boldsymbol{x}, \tau), \quad (23)$$

without the integrated potential term y.

In contrast, in the case of Asian options (Eq. 14), one has $C(x, y_1, y_2, \tau) = e^{y_1}C(x, y_2, \tau)$, so only the discount factor e^{y_1} cancels, but not the cumulated variable y_2.

6 Examples

6.1 The Cameron-Martin formula

The functional

$$E\left[e^{-\frac{\gamma^2}{2} \int_0^\rho W^2(\tau)d\tau} \mid W(0) = x\right] = \frac{e^{-\frac{1}{2}x^2\gamma \tanh(\gamma\rho)}}{\sqrt{\cosh(\gamma\rho)}} \quad (24)$$

of the Wiener process was computed analytically by Cameron and Martin [6], Gelfand and Yaglom [14]. See also Borodin and Salminen [4, p. 168, 1.9.3]. According to Eq. (3), it is the solution of the backward equation with the time invariant functions $f = 0, \Omega = 1$, quadratic potential $v(x) = -\frac{1}{2}\gamma^2 x^2$ and terminal condition $C(x, t) = h(x) = 1$. Explicitly, we have

$$C_t(x, t) + \left[\frac{1}{2}\nabla_x^2 - \frac{1}{2}\gamma^2 x^2\right] C(x, t) = 0; \ t \leq T. \quad (25)$$

Now, since the system is autonomous, one can rewrite Eq. (3) as

$$C(x, T) = E\left[e^{\int_0^{T-t} v(X(\tau))d\tau} h(X(T-t)) \,\Big|\, X(0) = x\right].$$

(26)

Setting $\rho = T - t$, one obtains the left hand side of Eq. (24). Since $f = 0, \Omega = 1$, one has $X = W$.

Liptser and Shiryayev [33, vol. 1, ch. 7.7.] derived a multivariate extension of Eq. (24), which serves as a comparison of Monte Carlo and numerical integration, based on a semi-discretization of the backward equation. For a p-dimensional Wiener process starting at $\boldsymbol{x}$, one obtains

$$E\left[e^{-\frac{1}{2}\int_0^\rho \boldsymbol{W}^{\mathsf{T}}(\tau)\boldsymbol{\Gamma}(\tau)\boldsymbol{W}(\tau)d\tau} \,|\, \boldsymbol{W}(0) = \boldsymbol{x}\right] = e^{\frac{1}{2}\int_0^\rho \mathrm{tr}[\boldsymbol{R}(\tau)]d\tau} e^{\frac{1}{2}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{R}(0)\boldsymbol{x}},$$

(27)

where $\boldsymbol{\Gamma} : p \times p$ is a symmetric positive semidefinite weight matrix and $\boldsymbol{R}(\tau)$ is the solution of the matrix Riccati equation ($0 \leq \tau \leq \rho$)

$$\dot{\boldsymbol{R}}(\tau) = \boldsymbol{\Gamma}(\tau) - \boldsymbol{R}(\tau)^2; \ \boldsymbol{R}(\rho) = \boldsymbol{0}.$$

(28)

In the case $p = 2$ one has the system

$$\dot{R}_{11} = \gamma_1^2 - R_{11}^2 - R_{12}^2 \tag{29}$$
$$\dot{R}_{12} = \gamma_{12} - (R_{11} + R_{22})R_{12} \tag{30}$$
$$\dot{R}_{22} = \gamma_2^2 - R_{22}^2 - R_{12}^2. \tag{31}$$

For diagonal $\boldsymbol{\Gamma}$, the equation for R_{12} can be solved by

$$R_{12}(\tau) = e^{\int_\tau^\rho (R_{11} + R_{22})ds} \ R_{12}(\rho) = 0.$$

Therefore for constant $\boldsymbol{\Gamma}$, one obtains the solution (analogous for R_{22})

$$R_{11}(\tau) = \gamma_1 \tanh(\gamma_1(\tau - \rho))$$
$$R_{11}(0) = -\gamma_1 \tanh(\gamma_1 \rho)$$

and $\frac{1}{2}\int_0^\rho R_{11}(\tau)d\tau = -\frac{1}{2}\log\cosh\gamma_1\rho$. Note that $\boldsymbol{R}(0)$ is a function of ρ, denoted as $\boldsymbol{R}(0, \rho)$. This implies the result

$$E\left[e^{-\frac{1}{2}\int_0^\rho \boldsymbol{W}^{\mathsf{T}}(\tau)\mathrm{diag}(\gamma_1^2, \gamma_2^2)\boldsymbol{W}(\tau)d\tau} \,|\, \boldsymbol{W}(0) = [x_1, x_2]^{\mathsf{T}}\right] =$$
$$e^{-\frac{1}{2}\log\cosh\gamma_1\rho - \frac{1}{2}\log\cosh\gamma_2\rho} e^{-\frac{1}{2}x_1^2\gamma_1\tanh(\gamma_1\rho) - \frac{1}{2}x_2^2\gamma_2\tanh(\gamma_2\rho)}$$

(32)

which is, as expected, the product of the univariate formula Eq. (24). With off-diagonal $\gamma_{12} \neq 0$, the solution of Eq. (30) is not identically to zero and the system Eqs. (29–31) must be solved. One can use Runge-Kutta integration and compute the importance drift $\nabla \log C(\boldsymbol{x}, u) = \boldsymbol{R}(0, u)\boldsymbol{x}$ (see Eq. (27) and below).

Fig. 1 shows a Monte Carlo approximation (Eq. 22) of the 2-dimensional solution Eq. (32) as a function of $x_1 = x_2 = x$ with sample size $N = 10$. The standard errors of the simulated solution were estimated as $s/\sqrt{N}$,
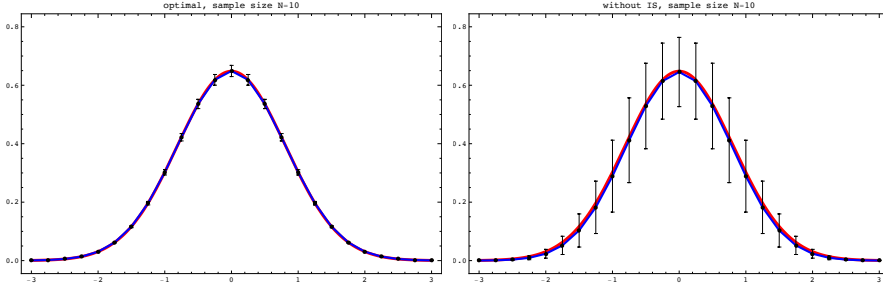
Fig. 1 Two-dimensional Cameron-Martin formula (Monte Carlo simulation). Solution and estimated standard errors as a function of $x_1 = x_2 = x = -3, ..., +3, \delta x = 0.25$, $\rho = 1$, $\gamma_1^2 = \gamma_1^2 = 1, \gamma_{12} = 0$, $\delta\tau = 0.01$. Sample size $N = 10$. Left: Optimal importance sampling. Right: MC solution without importance sampling. Red lines: analytical values computed from Eq. (32).
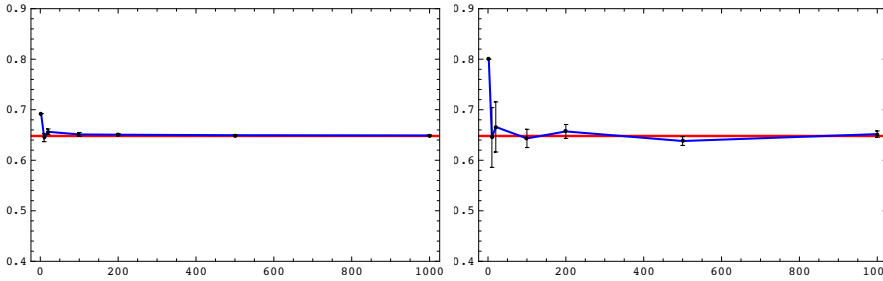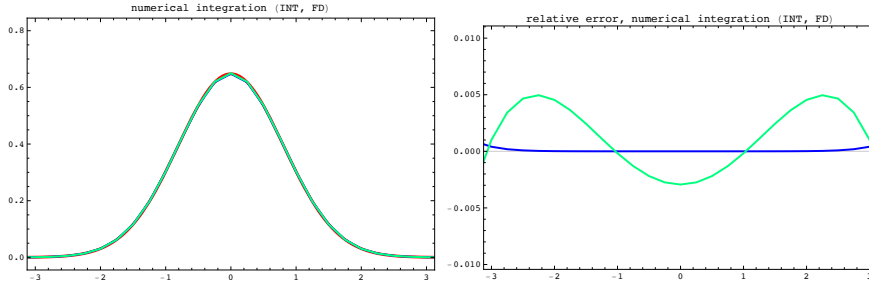

Fig. 2 Two-dimensional Cameron-Martin formula (Monte Carlo simulation). Left: convergence for fixed $\boldsymbol{x} = [0, 0]^{\mathsf{T}}$ over sample size $N = 2, ...1000$ and estimated standard errors (optimal importance sampling). Right: convergence without importance sampling.

where s is the standard deviation of the N random numbers in Eq. (22), i.e. $e^{y_{nT}} h(x_{nT})(q_1/q_2)(\boldsymbol{X}_n, \boldsymbol{y}_n, \rho \mid x, y = 0, t = 0)$.

The formula was evaluated for the coordinates $x_1 = x_2 = x = -3, ..., +3$, $\delta x = 0.25$ with time step $\delta\tau = 0.01$. The weight matrix $\boldsymbol{\Gamma} = \mathrm{diag}(1, 1)$ was set diagonal to have an analytical reference solution. Clearly, the simulated solution with optimal importance sampling (top) is very accurate. The standard deviations are about a factor 10^{-1} smaller as compared to the naive MC estimator (bottom), and the relative error w.r.t. the exact solution is smaller than 2.5%. In Fig. 2, the convergence over sample size $N = 2, ..., 1000$ is displayed (error bars: means $\pm$ standard errors).

It is interesting to compare the Monte Carlo results with approximations based on numerical integration (for details, see [51]). The functional integral Eq. (3) was first approximated on the time slices $\tau_j = s + j\delta\tau, j = 0, ...J; \delta\tau = (T - t)/J$ by a $(J - 1)$-dimensional integral. Afterwards, in a second step, the integrals were replaced by Riemann sums on a 'spatial' grid $\boldsymbol{x}_a, ..., \boldsymbol{x}_b$ with spacing $\delta\boldsymbol{x}$. The transition densities $p(\boldsymbol{x}_{j+1}, \tau_{j+1} \mid \boldsymbol{x}_j, \tau_j)$ were evaluated on these sample points, resulting in transition matrices. Thus the functional integral is approximated by a $(J - 1)$-fold matrix product.
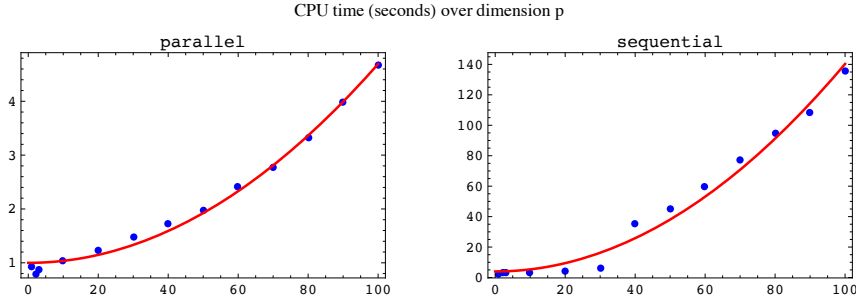
Fig. 3 Two-dimensional Cameron-Martin formula (numerical integration). Left: Solution as a function of x. Same parameters as in Fig. 1, spatial discretization $\boldsymbol{x}_a = [-5, -5]^\mathsf{T}, ..., \boldsymbol{x}_b = [5, 5]^\mathsf{T}$, $\delta\boldsymbol{x} = [0.25, 0.25]^\mathsf{T}$, $\delta\tau = 0.01$. Exact solution and methods INT (blue), FD (green) (for details, see text). Right: relative error w.r.t. the analytical solution as a function of x.
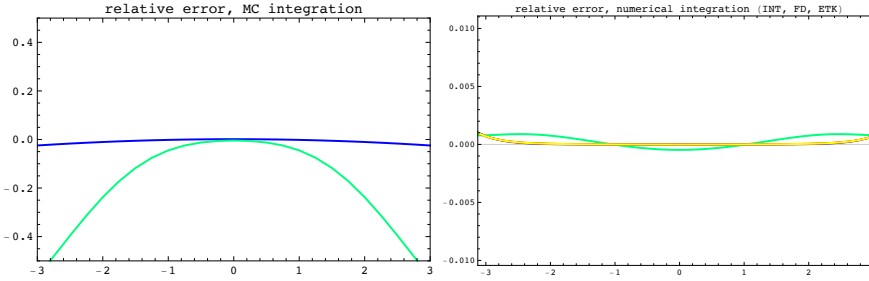

Fig. 4 Two-dimensional Cameron-Martin formula (numerical integration). Left: relative error of MC solution (blue line: with importance sampling, green: without IS). Right: relative error with smaller spatial grid steps $\delta\boldsymbol{x} = [0.1, 0.1]^\mathsf{T}$. Methods INT, FD and ETK (yellow).

| | Monte Carlo | | | | | semi-discretization | | |
	MC/IS		MC/without IS			INT	FD	ETK
N	10	100	10	100	1000			
p					relative error			
1	0.009	0.009	0.728	0.068	0.075	0.0002	0.0025	0.0015
2	0.025	0.017	0.566	0.253	0.066	0.0004	0.005	0.0029
3	0.029	0.027	0.832	0.504	0.148	0.0002	0.027	0.0605
p					CPU time			
1	0.34	3.52	0.17	1.33	14.86	0.02	0.01	0.01
2	0.39	3.80	0.17	1.41	14.80	0.91	0.05	6.00
3	0.41	3.80	0.16	1.48	15.96	40.91	1.23	146.66

Table 1 Cameron-Martin formula: Maximum absolute relative error in the interval $x = -3, ..., 3$ and CPU time of Monte Carlo method versus semi-discretization methods for $p = 1, ..., 3$. Sample size $N = 10, 100, 1000$, $\delta\tau = 0.01$. Spatial discretization $\boldsymbol{x}_a = [-5, ..., -5]^\mathsf{T}, ..., \boldsymbol{x}_b = [5, ..., 5]^\mathsf{T}$, $\delta\boldsymbol{x} = [0.25, ..., 0.25]^\mathsf{T}$. For $p = 3$, $\delta\boldsymbol{x} = [0.5, 0.5, 0.5]^\mathsf{T}$ (see text).

CPU time (seconds) over dimension p

Fig. 5 Cameron-Martin formula: MC method with importance sampling, sample size $N = 10$. CPU time (in seconds) as function of dimension $p = 1, ..., 100$ and quadratic fit. Left: parallel, right: sequential computation.
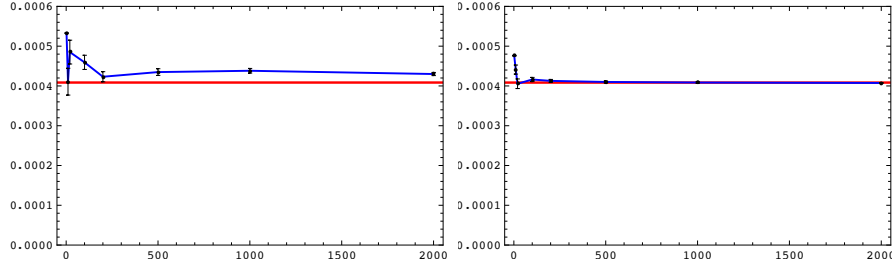
Fig. 6 Cameron-Martin formula, $\boldsymbol{x} = [0, ..., 0]^{\mathsf{T}}$, Nondiagonal weight matrix $\boldsymbol{\Gamma} : 10 \times 10$. Exact importance sampling for several time steps as a function of sample size $N = 2, ..., 2000$. For larger time steps, the estimates are biased. Left: $\delta t = 0.01$. Right: $\delta t = 0.001$. For comparison, the solution of the Riccati equation using Runge-Kutta integration is displayed (red line).
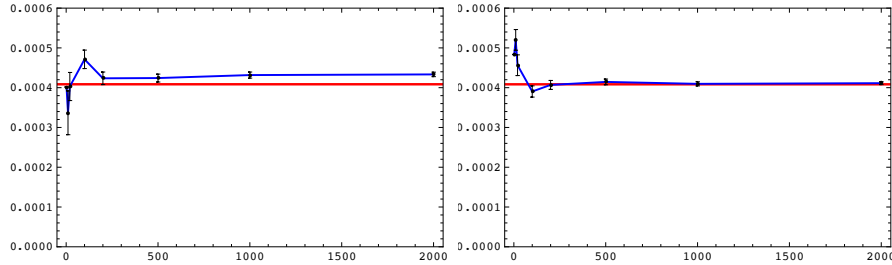
Fig. 7 Cameron-Martin formula, $\boldsymbol{x} = [0, ..., 0]^{\mathsf{T}}$, Nondiagonal weight matrix $\boldsymbol{\Gamma} : 10 \times 10$. Approximative importance sampling (diagonal $\boldsymbol{\Gamma}_{\mathrm{imp}}$) as a function of sample size $N = 2, ..., 2000$. Left: $\delta t = 0.01$. Right: $\delta t = 0.001$.

The transition matrices were computed by 3 methods:

1. Euler transition kernel (ETK) based on the short time Gaussian approximation $\phi(\boldsymbol{x}_{j+1}; \boldsymbol{x}_j + \boldsymbol{f}_j \delta\tau, \boldsymbol{\Omega}_j \delta\tau)$ (cf. Eq. (10)).
2. Matrix exponential $\exp(\boldsymbol{F}\delta\tau)$ of the discretized Fokker-Planck operator $\boldsymbol{F} = \boldsymbol{L}^{\mathsf{T}}$. Here we used either finite differences (FD) or an integral representation (INT) $F(\boldsymbol{x})h(\boldsymbol{x}) = \int F(\boldsymbol{x}, \boldsymbol{x}')h(\boldsymbol{x}')d\boldsymbol{x}'$, where the integral kernel

is $F(\boldsymbol{x}, \boldsymbol{x}') = F(\boldsymbol{x})\delta(\boldsymbol{x} - \boldsymbol{x}')$, and $F = L^*$ is the adjoint of the backward operator Eq. (2). $F(\boldsymbol{x}, \boldsymbol{x}')$ was evaluated on the spatial grid, resulting again in a matrix [51, 56].

Clearly, the size of the matrices scales with the power of model dimension p and gets excessively large. In contrast, the computational demand of the MC method should be of order Np^2, since the simulated data and involved densities depend on the $p \times p$ diffusion matrix $\boldsymbol{\Omega}$ (cf. Eq. 21).

In Fig. (3), top picture, the exact solution Eq. (32) for the 2-dimensional model and the methods INT and FD are displayed. The relative error w.r.t. the exact solution (smaller than 0.04% (INT), 0.5% (FD)) is shown in the bottom picture, and for the MC estimates with sample size $N = 10$ (smaller than 2.5%) in the top picture of Fig. (4). Due to the large grid steps $\delta\boldsymbol{x} = [0.25, 0.25]$ in the range $\boldsymbol{x}_a = [-5, -5],..., \boldsymbol{x}_b = [5, 5]$ and $\delta\tau = 0.01$, the method ETK yields very large errors and is not displayed. In fact a smaller $\delta\boldsymbol{x}$ of order $\sqrt{\delta\tau} = 0.1$ should be used. However, this yields transition matrices of order 10201×10201, which is computationally demanding. Now, the method ETK yields reasonable errors and the result is displayed in Fig. (4), bottom.

A comparison of the dimensions $p = 1, ..., 3$ is shown in table 1. The Cameron-Martin formula was computed for the 25 values $x_1 = ... = x_p = x = -3, ..., 3; \delta x = 0.25$ and $\delta\tau = 0.01$. The interval for the numerical integration was chosen larger ($\boldsymbol{x}_a = [-5, ..., -5], ..., \boldsymbol{x}_b = [5, ..., 5], \delta\boldsymbol{x} = [0.25, ..., 0.25]$), to minimize errors in the region of interest. In the case $p = 3$, one would obtain a $41^3 \times 41^3 = 68921 \times 68921$ transition matrix, therefore the grid steps were doubled to $\delta\boldsymbol{x} = [0.5, 0.5, 0.5]$ (9261×9261 matrix). Then, the maximal absolute relative error of the 25 function values w.r.t. the analytical solution ($\max_{j=1,...,25} |\hat{C}(\boldsymbol{x}_j)/C(\boldsymbol{x}_j) - 1|$) and the CPU time was determined and tabulated. The times relate to a Intel Core i9 processor with 2,9 GHz and Mathematica 11.3. Sparse arrays and parallel computing were utilized. The MC method with importance sampling performs well in all dimensions $p = 1, ..., 3$, already for small sample size $N = 10$. In contrast, without importance sampling, a much bigger sample size $N > 1000$ is required. The semi-discretization methods INT and FD yield very small relative error, except for the case FD, $p = 3$ (similar to MC). In higher dimensions, the performance of these methods declines quickly. As noted above already for $p = 3$, the grid steps must be doubled to permit numerical computations. The small CPU time for the finite difference method FD is due to the tridiagonal form of the backward matrix $\boldsymbol{L} = \boldsymbol{F}^{\mathsf{T}}$.

For the ETK method, the time discretization should be of order $\delta\tau = \delta x^2$, i.e. $0.25^2 = 0.0625, p = 1, 2$ and $0.5^2 = 0.25$ for $p = 3$. Therefore, the last column in the table was computed with these values.

In summary, the numerical integration methods work well for low dimensions $p = 1, 2$, but the size of the transition matrices scales with the power of p and CPU time and storage demand gets excessively large, despite using sparse arrays. In contrast, the MC methods also work for high dimensions, especially when using parallel computing (see Fig. 5).

Finally, we used a nondiagonal matrix $\boldsymbol{\Gamma} : 10 \times 10$ with main diagonal $\gamma_{ii} = i$ and $\gamma_{ij} = 1/2, i \neq j$ and computed the solution with exact and approximate importance sampling. In the former case, the Riccati equation (28) was solved using an Euler scheme yielding the log derivative $\nabla \log C(\boldsymbol{x}, u) = \boldsymbol{R}(0, u)\boldsymbol{x}$. Approximate importance sampling was implemented using the product of univariate explicit formulas as in Eq. (32), equivalent to a diagonal $\boldsymbol{\Gamma}_{\text{imp}}$ in Eq. (20). The results are displayed in Fig. (6) for fixed $\boldsymbol{x} = [0, ..., 0]^{\mathsf{T}}$.

The MC estimates with standard errors were plotted over sample size $N = 2, ..., 2000$ with time steps $\delta t = 0.01$ (top) and $\delta t = 0.001$ (second from the top), both with exact importance sampling. Clearly, the bias of the estimates gets smaller with decreasing time step δt. The red line is the result computed with a 4th order Runge-Kutta scheme for Eq. (28). In Fig. (7), approximate importance sampling is displayed. The standard errors are larger, but also a substantial variance reduction is achieved.

6.2 Schrödinger equation

With respect to the time difference $\rho = T - t$, one obtains from Eq. (25) in forward direction $\rho \geq 0$

$$u_\rho(x, \rho) = [\tfrac{1}{2}\Delta - \tfrac{1}{2}\gamma^2 x^2]u(x, \rho), \tag{33}$$

where $u(x, \rho) = C(x, T - \rho)$. This can be considered as Euclidean version of the Schrödinger equation for the harmonic oscillator with quadratic potential $v(x) = \tfrac{1}{2}\gamma^2 x^2$, $m = \hbar = 1$,

$$i\psi_\rho(x, \rho) = [-\tfrac{1}{2}\Delta + v(x)]\psi(x, \rho), \tag{34}$$

$\rho \geq 0$, with the replacement $i\rho \to \rho$ (imaginary time) [44, 49, p. 7]. Using the initial condition $u(x, 0) = \delta(x - z)$, the solution of equation (33) reads (cf. Eq. (26))

$$u(x, \rho) = E\left[e^{-\frac{\gamma^2}{2}\int_0^\rho W^2(\tau)d\tau} \, \delta(W(\rho) - z) \, \Big| \, W(0) = x\right] =$$
$$\sqrt{\frac{\gamma}{2\pi \sinh(\gamma\rho)}} \exp\left(\frac{\gamma[2xz - (x^2 + z^2)\cosh(\gamma\rho)]}{2\sinh(\gamma\rho)}\right) \tag{35}$$

[4, 13, p. 168]. A simulated solution of Eq. (33) using optimal importance sampling (Eq. (22)) is shown in Fig. 8, top row, as a function of initial localization $z = -1, ..., 2$. As importance drift, the representation Eq. (23) was inserted, using the analytical solution Eq. (35). In contrast, without importance sampling, one obtains much larger simulation errors (Fig. 8, middle). Now, since in practice the exact solution is unknown, one must use a suitable approximation to compute a suboptimal drift correction. Setting $\gamma = 0$, the oscillator equation coincides with the diffusion equation. The solution is $u(x, \rho) = \phi(x; z, \rho)$, $u(x, 0) = \delta(x - z)$. This function is used in the drift correction. In Fig. 8, bottom, the result is displayed. It is very similar to the optimal
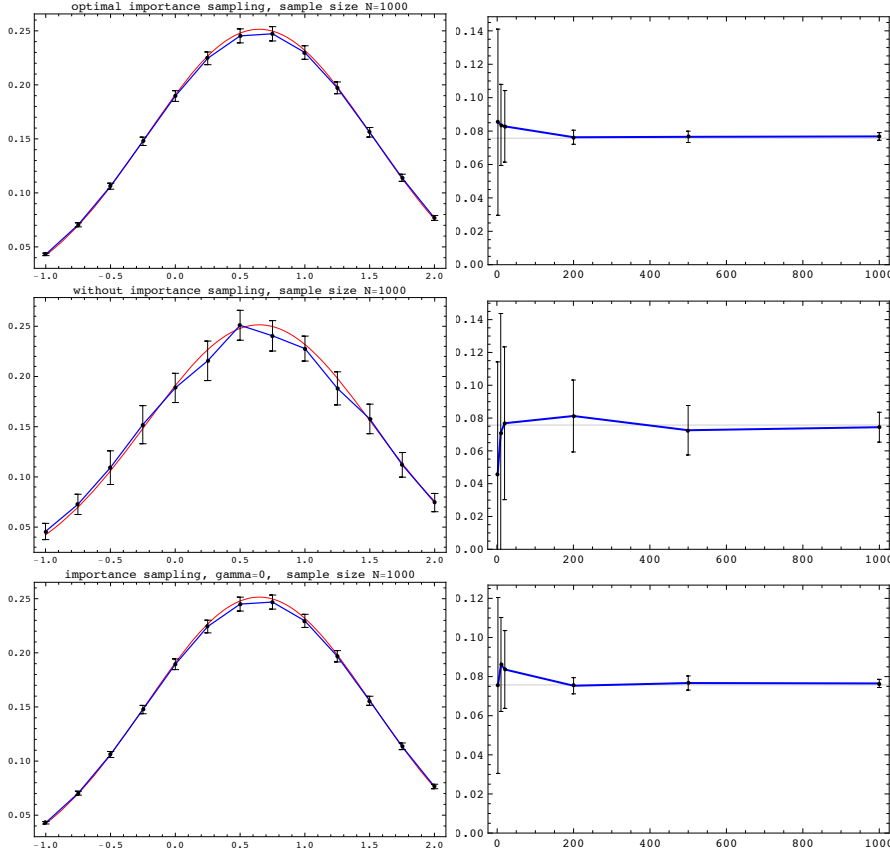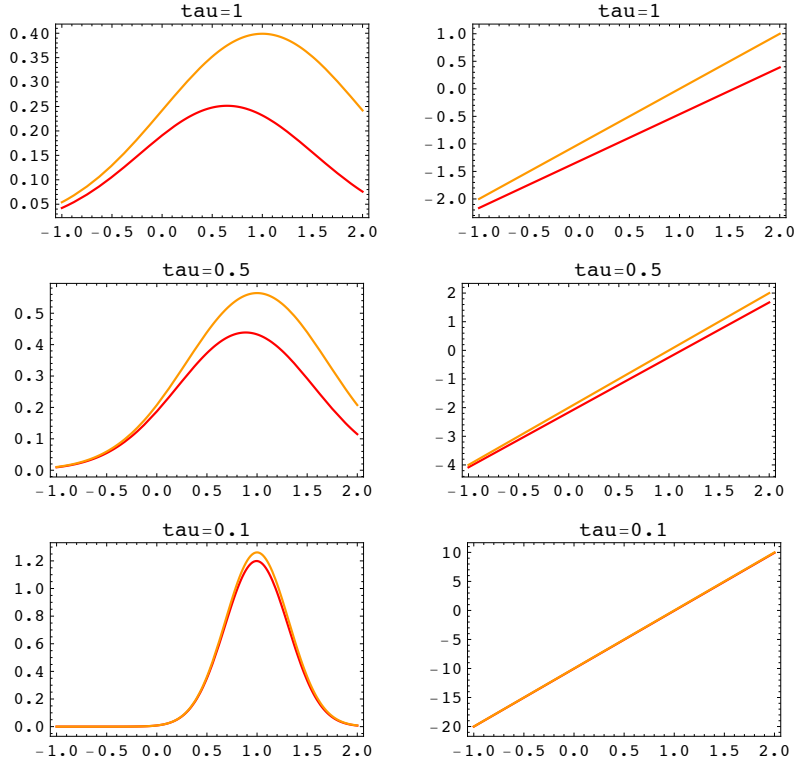
Fig. 8 Schrödinger equation (linear oscillator). Left: Simulated Feynman-Kac formula as a function of $z = -1, ..., 2$, $x = 1$, $\rho = 1$, $\gamma = 1$, $\delta\tau = 0.01$. Sample size $N = 1000$, means and standard deviations in $M = 10$ replications. Top: optimal importance sampling. Middle: without importance sampling. Bottom: approximate importance sampling with $\gamma = 0$ (see text). Right: means and standard errors as a function of sample size $N = 2, ..., 1000$, $x = 1, z = 2$.

approach (top). A plot over sample size displays some minor differences (Fig. 8, right column). Without importance sampling, very large sample sizes are required to obtain small standard deviations (middle, right). In Fig. 8, means and standard deviations (error bars) were computed over $M = 10$ replications of Eq. (22).

In Fig. 9, left column, the exact solution of the oscillator equation and the corresponding diffusion equation ($v(x) = 0; \gamma = 0$) are displayed for several time points. The logarithmic derivatives (right) for the importance drift are linear in both cases. For small times, the derivatives and thus the importance sampling Monte Carlo algorithms are very similar. Furthermore, Fig. 10 illustrates the tilted dynamics of the optimal trajectories, which are determined by the drift correction.

Fig. 9 Schrödinger equation. Left column: Comparison of exact (red) and approximate (orange) solution of the oscillator equation for three time points $\tau = 1, 0.5, 0.1$. Right column: logarithmic derivative $\nabla_x \log u(x, \tau)$ for the importance drift.
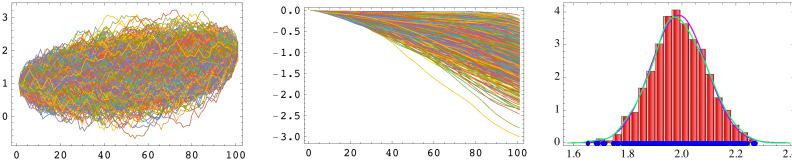
Fig. 10 Schrödinger equation. Simulation of solution $u(x, \rho)$, $\rho = 1, x = 1, z = 2, \delta\tau = 0.01$. Left: trajectories $W_2(\tau)$ with importance drift f_2 (as function of $\tau/\delta\tau = 0, ..., 100$). Middle: trajectories $-\gamma^2/2 \int_0^\tau W_2^2(\tau') d\tau'$, Right: histogram of $W_2(\rho)$, mean $= 1.98804$, std $= 0.10231$.

As explained in Sec. 6.1, one can solve the backward equation on a spatial grid (here from $x_a = -5, ..., x_b = 5, \delta x = 0.1$). The results are displayed in Fig. 11. In the top picture, the exact solution (red) is displayed together with the three methods as described above. The relative error (smaller than 2%) w.r.t. the analytical solution is displayed in the middle picture. The results of INT and FTK are very similar. Similar small errors are obtained for the MC method ($N = 1000$) with exact and approximate (zero potential) importance
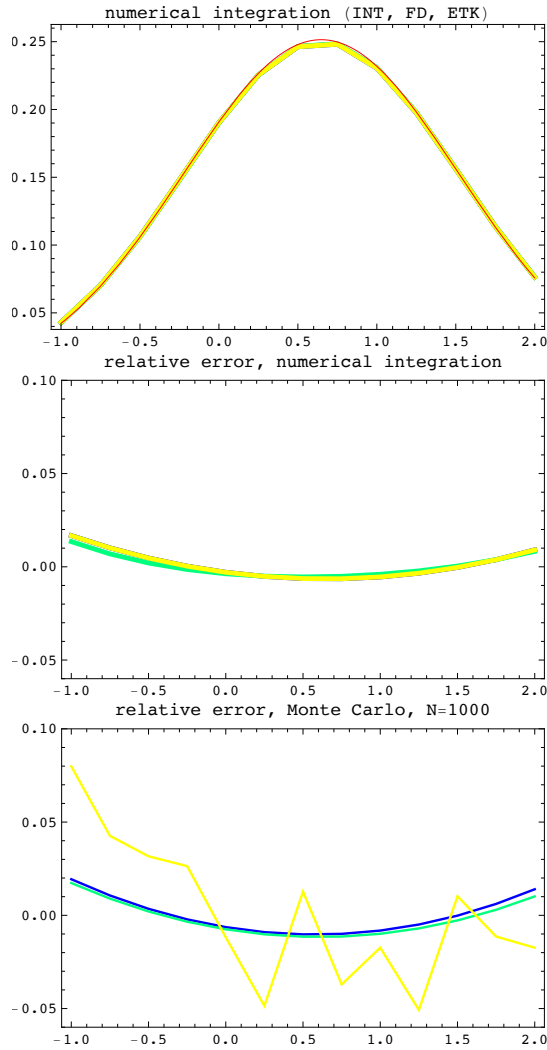
Fig. 11 Schrödinger equation. Comparison of MC with numerical integration (see text). Top: Analytical solution (red), finite differences (FD, green), Euler transition kernel (yellow) and integral operator (INT, blue) as a function of $z = -1, ..., 2$. Middle: relative error w.r.t. the analytical solution. The results of INT and FTK are very similar. Bottom: relative error of Monte Carlo approximations (optimal IS = blue, approximate IS (Gaussian) = green, $N = 1000$) w.r.t. the analytical solution. Without importance sampling (yellow), the error is very large.

sampling (bottom). The error without IS is much larger and not a smooth function of the variable z.

Table 2 displays the maximum relative absolute error w.r.t. the exact solution and CPU time for several MC sample sizes $N = 10, 1000, 2000$ ($M = 10$ replications) and grid steps $\delta x = 0.2, 0.1, 0.05$. For exact or approximate importance sampling, similar error estimates are obtained ($< 2\%$ for $N =$

	Monte Carlo			semi-discretization			
N	opt. IS	appr. IS	no IS	δx	INT	FD	ETK
relative error							
10	0.1238	0.1395	0.8081	0.2	0.0289	0.0196	3.5775
1000	0.0194	0.0173	0.0799	0.1	0.0167	0.0135	0.0167
2000	0.0164	0.0142	0.045	0.05	0.0167	0.0159	0.0167
CPU time							
10	2.25	2.33	0.95	0.2	0.09	0.05	0.01
1000	296.53	299.37	115.46	0.1	0.25	0.06	0.28
2000	587.31	575.13	244.98	0.05	1.29	0.10	0.92

Table 2 Schrödinger equation. Maximum absolute relative error in interval $z = -1, ..., 2$ and CPU time (seconds) of Monte Carlo method versus semi-discretization methods. Sample size $N = 10, 1000, 2000$, $M = 10$ replications. Spatial discretization $x_a = -5, ..., x_b = 5$, $\delta x = 0.2, 0.1, 0.05$, $\delta \tau = 0.01$ (see text).

$1000, 2000$). Without importance sampling, larger sample sizes are required. For the semi-discretization methods, a spatial discretization of $\delta x \leq 0.1$ is required to obtain errors $< 2\%$. In this one-dimensional example, the numerical integration method is much faster than the MC method.

Numerical considerations

In the simulations, the initial condition $u(0, x) = \delta(x - z)$ was represented by the delta sequence $\delta_\epsilon(x) = \phi(x; 0, \epsilon^2)$ [32] with the smoothing parameter $\epsilon^2 = \delta \tau$. This can be motivated from the quadratic variation $dW^2(\tau) = d\tau$ of the Wiener process. Furthermore, the drift correction $\delta f = \nu \, \Omega \nabla_x \log u(x, \tau)$ was multiplied by a scaling factor $\nu \approx 0.8$ in order to obtain a minimal standard deviation. The optimal value may be determined by a minimization algorithm. The scaling factor can be motivated by the observation, that the function $h = \delta(W(\rho) - z)$ is very narrow (of order $\sqrt{\delta \tau} = 0.1$) and too large drifts lead to approximation errors. Alternatively, a scaled smoothing parameter can be used.

6.3 European and Asian options

As already mentioned in Sec. 4, the terminal value of a so called arithmetic Asian option depends on the time average of the underlying $X(t)$ (stock price etc.) in the interval $t_0 \leq t \leq T$. This time average can be generated by the inclusion of an integrated variable $Y(t) = \int_{t_0}^t X(u) du$.

In contrast, the terminal condition of a European option is given only by the last value $X(t) = x$. For example, a call option has terminal condition $C(x, T) = h(x) = \max(x - K, 0)$ where K is the strike price. In this simpler case, the option fulfils the Black-Scholes partial differential equation (PDE)

$$C_t(x, t) + L(x, t)C(x, t) - r \, C(x, t) = 0, \; t \leq T, \tag{36}$$
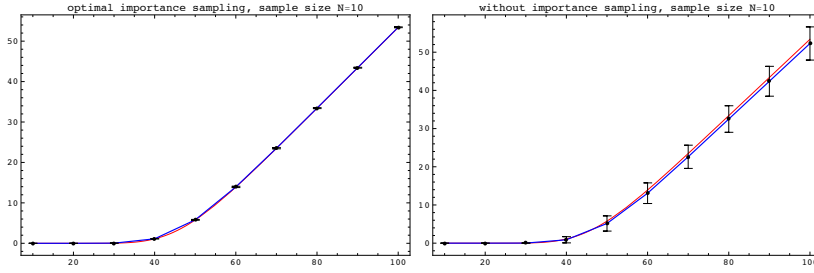
Fig. 12 Black-Scholes formula and simulated call option. Left: with optimal importance sampling. Right: without importance sampling. Sample size $N = 10$, strike price $K = 50$, volatility $\sigma = 0.2$, interest rate $r = 0.07$, $T = 1, t = 0$.

with terminal condition $C(x, T) = h(x)$ and backward operator

$$L(x, t) = rx\nabla_x + \tfrac{1}{2}\sigma^2 x^2 \nabla_x^2 \tag{37}$$

[see, e.g. 3, 29, 50]. Solving the PDE (36) leads to the Black-Scholes formula. Originally [3], separation of variables was used, but the stochastic representation

$$C(x, t) = E\left[e^{-r(T-t)}h(X(T)) \,\Big|\, X(t) = x\right] \tag{38}$$

gives the same results. The stochastic process $X(\tau)$ fulfils the Itô equation $(t \le \tau \le T; X(t) = x)$

$$dX(\tau) = rX(\tau)d\tau + \sigma X(\tau)dW(\tau), \tag{39}$$

a geometric Brownian motion. The integral in Eq. (38) can be calculated analytically by noting that $X(t)$ is log-normally distributed. The results are displayed in Fig. (12) for a call option with sample size $N = 10$, strike price $K = 50$, volatility $\sigma = 0.2$, interest rate $r = 0.07$, time to maturity $T - t = 1$ and terminal condition $h(x) = \max(x - K, 0)$. Clearly, importance sampling yields a very strong variance reduction. Other simulation methods, option types and dynamics of the underlying are discussed in [15, 16, 43, 48, 50, 59].

In the case of arithmetic Asian options, according to Section 4, one has the stochastic representation

$$C(x, y, t) = E\left[e^{-r(T-t)}h(Y(t)) \,\Big|\, X(t) = x, Y(t) = y\right] \tag{40}$$

where $Y(t) = \int_{t_0}^{t} X(u)du$ is an integrated variable. The processes X and Y fulfil the system of Itô equations (one could also join an equation for the state-dependent interest rate $v(x, \tau) = -r(x, \tau)$; see sect. 4)

$$dX(\tau) = rX(\tau)d\tau + \sigma X(\tau)dW(\tau)$$
$$dY(\tau) = X(\tau)d\tau,$$

with initial conditions $X(t) = x, Y(t) = y$, $t \le \tau \le T$. Since the integrated variable does not have an error term, a singular transition density
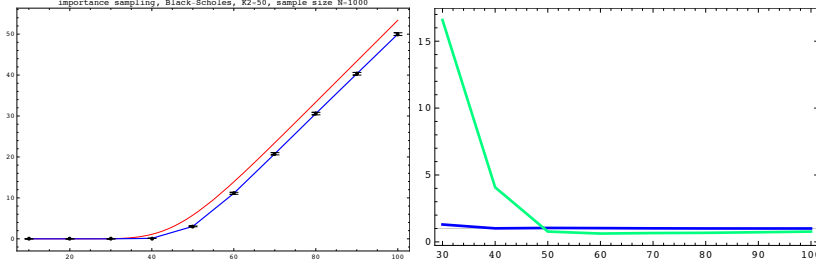
Fig. 13 Asian call option. Left: Simulated values (Black-Scholes importance sampling, $K_2 = K$). Black-Scholes value (red) for comparison. Same parameters as in Fig. 12, sample size $N = 1000$ and $t = t_0 = 0, T = 1$. Right: Ratios of means (blue) and standard deviations (green), each ratio computed without/with importance sampling. Strong variance reduction for options out of the money ($x < 50$).

$p(x_{j+1}, y_{j+1}|x_j, y_j)$ with covariance matrix $\mathrm{diag}(\Omega, 0)$ occurs (see Section 5). Clearly, Eq. (40) is the Feynman-Kac formula for the PDE (Ingersoll [24, ch. 17.10], Zvan et al [60])

$$C_s + rxC_x + xC_y + \tfrac{1}{2}\sigma^2 x^2 C_{xx} - rC = 0; \ t \leq T.$$

$C \equiv C(x, y, t)$ with terminal condition $C(x, y, T) = h(y)$. For a call option, one sets

$$h(y) = \max\left(\frac{y}{T - t_0} - K, 0\right).$$

In this case, the solution is not known, but one can use the Black-Scholes formula to compute a suboptimal drift correction $\Omega(x, \tau)\nabla_x C_{\mathrm{ref}}(x, \tau)$. In this reference solution $C_{\mathrm{ref}}(x, \tau)$, appropriate parameters may be used, which give a better fit to the solution to be approximated. For example, the strike price K can be enlarged to generate smaller values of C_{ref}. Fig. 13 (left) shows that the Asian option value is below the European one (used as reference solution) for x-values above K. This may be understood from the observation that

$$\mu(t) = E[X(t)] = x\exp(rt)$$
$$E[Y(t)/t] = \frac{1}{t}\int_0^t \mu(\tau)d\tau = \frac{x}{rt}(\exp(rt) - 1) < x\exp(rt) = \mu(t).$$

Therefore, the expected time average is always smaller, at least for the geometric Brownian motion. From this it is plausible that also the expectation values of the terminal condition fulfil $E[h(Y(t)/t)] < E[h(X(t))]$, for $h(x)$ is linear above the strike price K, leading to a higher option value for a European option 'in the money' ($x > K$). In general, this is not true, since a lower expectation value can be compensated by a higher variance leading to a higher expectation of h.

Fig. 13 (right) displays the ratio of mean values and standard deviations, computed with and without importance sampling (see Eq. 21). The ratio of

standard deviations (green) shows a strong variance reduction for options out of the money ($x < K = 50$). In this case, many trajectories are lifted above the strike price K, whereas without importance sampling, they are supressed by the terminal condition h. Actually, for Asian options only the time average of $X(u)$ must be above K.

7 Conclusion

The inhomogenous backward equation with potential term and integrated variables was transformed, using a state extension, to a singular form, where the inhomogeneity was absorbed in the drift term. Then, the Feynman-Kac formula can be represented as a Kolmogorov formula and the importance sampling method with distorted drift can be used, although the Feynman-Kac formula contains terms depending on the complete history from present time t to final time T. This embraces, e.g., state-dependent potentials (interest rates) and models with integrated variables, e.g. arithmetic Asian options, where the terminal value depends on the cumulated history. The resulting new probability measure p_2 expresses the optimal importance density in terms of a stochastic process with distorted drift, which only concerns the nonintegrated part. The integrated potential does not appear in the drift correction. Other integrated variables still appear in the correction term.

In one or two dimensions, methods using a spatial discretization are fast and accurate, but in higher dimensions, only the Monte Carlo approach yields feasible computations, since the computational demand only scales quadratically with dimension. Although the drift correction requires the knowledge of the solution, suboptimal importance sampling using approximate solutions leads to a substantial variance reduction.

In conclusion, the variance reduced Monte Carlo estimator for the Feynman-Kac representation can be computed by using a Kolmogorov formula with state extension, even in high dimensions.

Appendix A: The Feynman-Kac formula

The Feynman-Kac formula Eq. (3) can be proved by using the Itô lemma for the 'discounted' function $C^* = C(\boldsymbol{X}(\tau), \tau)\, \beta(t, \tau)$, where $C(\boldsymbol{x}, t)$ is a solution of Eq. 1 and $\beta(t, \tau) = \exp[\int_t^\tau v(\boldsymbol{X}(u), u)du]$ is the discount factor [see 29, ch. 5.7]. One obtains ($C_{,\alpha} := \partial_\alpha C$)

$$
\begin{aligned}
d[C(\boldsymbol{X}(\tau), \tau)\beta(t, \tau)] &= dC\,\beta + C\,d\beta \\
&= (C_{,\alpha}dX_\alpha(\tau) + \tfrac{1}{2}C_{,\alpha\beta}\Omega_{\alpha\beta}d\tau + C_\tau d\tau)\beta + C\beta v d\tau \\
&= \beta(LC + C_\tau + vC)d\tau + \beta C_{,\alpha}G_{\alpha\beta}dW_\beta(\tau) \qquad (41)
\end{aligned}
$$

The first term in Eq. (41) is null due to the backward equation. Integrating from t to T one obtains

$$C(\boldsymbol{X}(T), T)\beta(t, T) - C(\boldsymbol{X}(t), t)\beta(t, t) = \int_t^T \beta(t, \tau)C_{,\alpha}G_{\alpha\beta}dW_\beta(\tau).$$

Taking the conditional expectation $E[\,\cdot\,|\boldsymbol{X}(t) = \boldsymbol{x}]$ the result Eq. (3) is proved.

Appendix B: Milstein-Girsanov transformation

The drift correction used in Section 5, namely $\delta\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}, \tau) = \boldsymbol{\Omega}(\boldsymbol{x}, \tau)\nabla_{\boldsymbol{x}}\log C$ $(\boldsymbol{x}, \boldsymbol{y}, \tau)$, can also be derived by the measure transformation approach [41, ch. 12], [30, ch. 16.2]. The aim is to compute the functional

$$C(\boldsymbol{\xi}, t) = E[H(\boldsymbol{\xi}(T)) \mid \boldsymbol{\xi}(t) = \boldsymbol{\xi}],$$

where $d\boldsymbol{\xi} = \tilde{\boldsymbol{f}}dt + \tilde{\boldsymbol{G}}d\boldsymbol{W}$. By the Girsanov theorem, the 'drifted' Wiener process

$$\boldsymbol{W}_d(t) = \boldsymbol{W}(t) - \int_0^t \boldsymbol{d}(\boldsymbol{\xi}_d, \tau)d\tau$$

(cf. Eq. 42) is a Wiener process w.r.t. a transformed probability measure P_d with Radon-Nikodym density $dP_d/dP = \Theta(t)/\Theta(0)$. The correction process $\Theta(t)$ fulfils the stochastic equation $d\Theta(\tau) = \Theta(\tau)\langle\boldsymbol{d}(\boldsymbol{\xi}_d, \tau), d\boldsymbol{W}(\tau)\rangle$ with the explicit solution

$$\Theta(T)/\Theta(t) = \exp\left[\int_t^T \langle\boldsymbol{d}, d\boldsymbol{W}(\tau)\rangle - \tfrac{1}{2}\langle\boldsymbol{d}, \boldsymbol{d}\rangle d\tau\right],$$

$\boldsymbol{d} = \boldsymbol{d}(\boldsymbol{\xi}_d, \tau)$, where $\langle\boldsymbol{f}, \boldsymbol{g}\rangle = \sum_\alpha f_\alpha g_\alpha$ denotes the scalar product. Using the drifted Wiener process $\boldsymbol{W}_d$,

$$d\boldsymbol{\xi}_d = \tilde{\boldsymbol{f}}d\tau + \tilde{\boldsymbol{G}}d\boldsymbol{W}_d = \tilde{\boldsymbol{f}}d\tau - (\tilde{\boldsymbol{G}}\boldsymbol{d})d\tau + \tilde{\boldsymbol{G}}d\boldsymbol{W} \qquad (42)$$

is a transformed Itô process, which has the same properties as $\boldsymbol{\xi}(t)$ (w.r.t. the probability P_d). Therefore, the functional can be written as expectation w.r.t. the transformed process including a correction factor

$$E[H(\boldsymbol{\xi}(T)) \mid \boldsymbol{\xi}(t) = \boldsymbol{\xi}] = E[H(\boldsymbol{\xi}_d(T))\Theta(T)/\Theta(t) \mid \boldsymbol{\xi}_d(t) = \boldsymbol{\xi}].$$

It can be shown (Milstein [41], ch. 12, Kloeden and Platen [30], ch. 16.2.) that the optimal choice of the drift $\boldsymbol{d}$ is given by

$$\boldsymbol{d}(\boldsymbol{\xi}, \tau) = -\tilde{\boldsymbol{G}}^\mathsf{T}\nabla_{\boldsymbol{\xi}}\log C(\boldsymbol{\xi}, \tau).$$

Now, since $\tilde{\boldsymbol{G}}^\mathsf{T} = [\boldsymbol{G}^\mathsf{T}, \boldsymbol{0}]$ and $\boldsymbol{\xi} = [\boldsymbol{x}^\mathsf{T}, \boldsymbol{y}^\mathsf{T}]^\mathsf{T}$, we have $\boldsymbol{d}(\boldsymbol{\xi}, \tau) = -\boldsymbol{G}^\mathsf{T}\nabla_{\boldsymbol{x}}\log C$ $(\boldsymbol{x}, \boldsymbol{y}, \tau) := -\boldsymbol{G}^\mathsf{T}\boldsymbol{s}$. Therefore, the drift correction appearing in the transformed Itô equation (42) is

$$-\tilde{\boldsymbol{G}}\boldsymbol{d} = \begin{bmatrix} \boldsymbol{G}\boldsymbol{G}^\mathsf{T}\nabla_{\boldsymbol{x}}\log C(\boldsymbol{x}, \boldsymbol{y}, \tau) \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \delta\boldsymbol{f} \\ \boldsymbol{0} \end{bmatrix},$$

$\boldsymbol{\Omega} = \boldsymbol{G}\boldsymbol{G}^{\mathsf{T}}$. Now, $\langle \boldsymbol{d}, \boldsymbol{d} \rangle = \langle \boldsymbol{G}^{\mathsf{T}}\boldsymbol{s}, \boldsymbol{G}^{\mathsf{T}}\boldsymbol{s} \rangle = \langle \boldsymbol{\Omega}\boldsymbol{s}, \boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}\boldsymbol{s} \rangle = \langle \delta\boldsymbol{f}, \boldsymbol{\Omega}^{-1}\delta\boldsymbol{f} \rangle$ and $\langle \boldsymbol{d}, d\boldsymbol{W} \rangle = \langle \boldsymbol{G}^{\mathsf{T}}\boldsymbol{s}, d\boldsymbol{W} \rangle = \langle \boldsymbol{s}, d\boldsymbol{x} - \boldsymbol{f}d\tau - \delta\boldsymbol{f}d\tau \rangle = -\langle \delta\boldsymbol{f}, \boldsymbol{\Omega}^{-1}(d\boldsymbol{x} - \boldsymbol{f}d\tau - \delta\boldsymbol{f}d\tau) \rangle$, assuming a nonsingular diffusion matrix. Thus, the correction factor can be written as

$$\Theta(T)/\Theta(t) = \exp \int_t^T -\langle \delta\boldsymbol{f}, \boldsymbol{\Omega}^{-1}(d\boldsymbol{x}_d - \boldsymbol{f}d\tau - \tfrac{1}{2}\delta\boldsymbol{f}d\tau) \rangle,$$

which coincides with Eq. (19), after setting $\boldsymbol{f}_2 = \boldsymbol{f} + \delta\boldsymbol{f}$ and discretization of the integrals.

References

1. Arató M (1994) On the solution of Kac-type partial differential equations. Journal of Applied Probability 31(A):311–324
2. Asmussen S, Glynn PW (2007) Stochastic simulation: algorithms and analysis, vol 57. Springer Science & Business Media
3. Black F, Scholes M (1973) The pricing of options and corporate liabilities. Journal of Political Economy 81:637–654
4. Borodin A, Salminen P (2002) Handbook of Brownian Motion – Facts and Formulae, 2nd edn. Birkhäuser-Verlag, Basel
5. Brown LM (2005) Feynman's Thesis. A New Approach to Quantum Theory. World Scientific
6. Cameron RH, Martin WT (1945) Transformations of Wiener Integrals Under a General Class of Linear Transformations. Transactions of the American Mathematical Society 58(2):184–219
7. Cockayne J, Oates C, Sullivan T, Girolami M (2017) Bayesian probabilistic numerical methods. arXiv preprint arXiv:170203673
8. Cox J, Ross S (1976) The valuation of options for alternative stochastic processes. Journal of Financial Economics 3:145–166
9. Dupuis P, Spiliopoulos K, Wang H (2012) Importance sampling for multiscale diffusions. Multiscale Modeling & Simulation 10(1):1–27
10. El Karoui N, Peng S, Quenez MC (1997) Backward stochastic differential equations in finance. Mathematical Finance 7(1):1–71
11. Feller W (1951) Two singular diffusion problems. Annals of Mathematics 54:173–182
12. Feynman R (1948) Space-time approach to nonrelativistic quantum mechanics. Review of Modern Physics 20:367–387
13. Feynman R, Hibbs A (1965) Quantum Mechanics and Path Integrals. McGraw-Hill, New York
14. Gelfand I, Yaglom A (1960) Integration in Functional Spaces and its Application in Quantum Physics. Journal of Mathematical Physics 1(1):48–69
15. Giles MB (2008) Multilevel Monte Carlo path simulation. Operations Research 56(3):607–617
16. Glasserman P (2003) Monte Carlo Methods in Financial Engineering. Springer, New York

17. Gobet E, Turkedjiev P (2016) Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. Mathematics of Computation 85(299):1359–1391
18. Gobet E, Turkedjiev P (2017) Adaptive importance sampling in least-squares Monte Carlo algorithms for backward stochastic differential equations. Stochastic Processes and their Applications 127(4):1171–1203
19. Gobet E, López-Salas JG, Turkedjiev P, Vázquez C (2016) Stratified regression Monte-Carlo scheme for semilinear PDEs and BSDEs with large scale parallelization on GPUs. SIAM Journal on Scientific Computing 38(6):C652–C677
20. Hartmann C, Schütte C, Zhang W (2016) Model reduction algorithms for optimal control and importance sampling of diffusions. Nonlinearity 29(8):2298
21. Hennig P, Osborne MA, Girolami M (2015) Probabilistic numerics and uncertainty in computations. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 471(2179)
22. Holden H, Øksendal B, Ubøe J, Zhang T (2009) Stochastic Partial Differential Equations: A Modeling, White Noise Functional Approach, 2nd edn. Springer, Berlin
23. Huang K, Szlufarska I (2014) Green-Kubo relation for friction at liquid-solid interfaces. Physical Review E 89(3):032,119
24. Ingersoll J (1987) Theory of Financial Decision Making. Rowman & Littlefield, Savage, Maryland
25. Kac M (1949) On distributions of certain Wiener functionals. Transactions of the American Mathematical Society 65:1–13
26. Kac M (1966) Wiener and integration in function spaces. Bulletin of the American Mathematical Society 72:52–68
27. Kac M (1980) Integration in function spaces and some of its applications. Accademia Nazionale dei Lincei, Scuola Normale Superiore
28. Kappen HJ, Ruiz HC (2016) Adaptive importance sampling for control and inference. Journal of Statistical Physics 162(5):1244–1266
29. Karatzas I, Shreve S (1991) Brownian Motion and Stochastic Calculus, 2nd edn. Springer, New York, Berlin, Heidelberg
30. Kloeden P, Platen E (1999) Numerical Solution of Stochastic Differential Equations. Springer, Berlin, corrected third printing
31. Kolmogoroff A (1931) Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Mathematische Annalen 104:415–458
32. Lighthill M (1958) Introduction to Fourier Analysis and Generalised Functions. Cambridge University Press, Cambridge, U.K.
33. Liptser R, Shiryayev A (2001) Statistics of Random Processes, Volumes I and II, 2nd edn. Springer, New York, Heidelberg, Berlin (1974 in Russian)
34. Mardia K, Kent J, Bibby J (1979) Multivariate Analysis. Academic Press, London
35. Melchior M, Öttinger HC (1995) Variance reduced simulations of stochastic differential equations. The Journal of Chemical Physics 103(21):9506–9509

36. Melchior M, Öttinger HC (1996) Variance reduced simulations of polymer dynamics. The Journal of Chemical Physics 105(8):3316–3331
37. Merton R (1973) Theory of Rational Option Pricing. Bell Journal of Economics and Mangement Science 4:141–183, reprinted in Merton, 1990
38. Merton R (1990) Continuous-Time Finance. Blackwell, Cambridge MA, Oxford UK
39. Milstein G, Tretyakov MV (2006) Numerical algorithms for forward-backward stochastic differential equations. SIAM Journal on Scientific Computing 28(2):561–582
40. Milstein G, Tretyakov MV (2009) Practical variance reduction via regression for simulating diffusions. SIAM Journal on Numerical Analysis 47(2):887–910
41. Milstein GN (1995) Numerical integration of stochastic differential equations, vol 313. Springer Science & Business Media (1988 in Russian)
42. Milstein GN, Tretyakov MV (2004) Stochastic Numerics for Mathematical Physics. Springer, Berlin, Heidelberg, New York
43. Müller A (2016) Improved Variance Reduced Monte-Carlo Simulation of in-the-Money Options. Journal of Mathematical Finance 6(3):361–367
44. Nelson E (1983) Functional Integration and Quantum Physics; Quantum Physics, A Functional Integral Point of View (book review). Bull Amer Math Soc 8, 2:378–380
45. Pardoux E, Peng S (1990) Adapted solution of a backward stochastic differential equation. Systems & Control Letters 14(1):55–61
46. Peng S (1991) Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. Stochastics and Stochastics Reports 37(1-2):61–74
47. Revuz D, Yor M (1998) Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin, Heidelberg
48. Rhee CH, Glynn PW (2015) Unbiased Estimation with Square Root Convergence for SDE Models. Operations Research 63(5):1026–1043
49. Simon B (2005) Functional Integration and Quantum Physics, 2nd edn. American Mathematical Society
50. Singer H (2014) Importance Sampling for Kolmogorov Backward Equations. Advances in Statistical Analysis 98:345–369
51. Singer H (2018) Langevin and Kalman Importance Sampling for Nonlinear Continuous-Discrete State Space Models. In: van Montfort K, Oud J, Voelkle M (eds) Continuous Time Modeling in the Behavioral and Related Sciences, Springer
52. Thijssen S, Kappen H (2015) Path integral control and state-dependent feedback. Physical Review E 91(3):032,104
53. Van Kampen N (1992) Stochastic Processes in Physics and Chemistry, 3rd edn. North Holland, Amsterdam, New York, Oxford
54. Vanden-Eijnden E, Weare J (2012) Rare event simulation of small noise diffusions. Communications on Pure and Applied Mathematics 65(12):1770–1803

55. Wagner W (1988) Monte Carlo Evaluation of Functionals of Solutions of Stochastic Differential Equations. Variance Reduction and Numerical Examples. Stochastic Analysis and Applications 6:447–468
56. Wei GW, Zhang DS, Kouric DJ, Hoffman DK (1997) Distributed approximating functional approach to the Fokker-Planck equation: Time propagation. The Journal of Chemical Physics 107 (8):3239–3246
57. Zhang W, Hartmann C, Weber M, Schütte C (2013) Importance sampling in path space for diffusion processes. Multiscale Modeling and Simulation
58. Zhang W, Wang H, Hartmann C, Weber M, Schütte C (2014) Applications of the cross-entropy method to importance sampling and optimal control of diffusions. SIAM Journal on Scientific Computing 36(6):A2654–A2672
59. Zhao Q, Liu G, Gu G (2013) Variance Reduction Techniques of Importance Sampling Monte Carlo Methods for Pricing Options. Journal of Mathematical Finance 3(04):431
60. Zvan R, Forsyth PA, Vetzal KR (1996) Robust numerical methods for PDE models of Asian options. Tech. rep., University of Waterloo

Die Diskussionspapiere ab Nr. 403 (2007) bis heute, können Sie im Internet unter
http://www.fernuni-hagen.de/wirtschaftswissenschaft/forschung/beitraege.shtml einsehen
und zum Teil downloaden.

Ältere Diskussionspapiere selber erhalten Sie nur in den Bibliotheken.

Nr	Jahr	Titel	Autor/en
420	2008	Stockkeeping and controlling under game theoretic aspects	Fandel, Günter Trockel, Jan
421	2008	On Overdissipation of Rents in Contests with Endogenous Intrinsic Motivation	Schlepütz, Volker
422	2008	Maximum Entropy Inference for Mixed Continuous-Discrete Variables	Singer, Hermann
423	2008	Eine Heuristik für das mehrdimensionale Bin Packing Problem	Mack, Daniel Bortfeldt, Andreas
424	2008	Expected A Posteriori Estimation in Financial Applications	Mazzoni, Thomas
425	2008	A Genetic Algorithm for the Two-Dimensional Knapsack Problem with Rectangular Pieces	Bortfeldt, Andreas Winter, Tobias
426	2008	A Tree Search Algorithm for Solving the Container Loading Problem	Fanslau, Tobias Bortfeldt, Andreas
427	2008	Dynamic Effects of Offshoring	Stijepic, Denis Wagner, Helmut
428	2008	Der Einfluss von Kostenabweichungen auf das Nash-Gleichgewicht in einem nicht-kooperativen Disponenten-Controller-Spiel	Fandel, Günter Trockel, Jan
429	2008	Fast Analytic Option Valuation with GARCH	Mazzoni, Thomas
430	2008	Conditional Gauss-Hermite Filtering with Application to Volatility Estimation	Singer, Hermann
431	2008	Web 2.0 auf dem Prüfstand: Zur Bewertung von Internet-Unternehmen	Christian Maaß Gotthard Pietsch
432	2008	Zentralbank-Kommunikation und Finanzstabilität – Eine Bestandsaufnahme	Knütter, Rolf Mohr, Benjamin
433	2008	Globalization and Asset Prices: Which Trade-Offs Do Central Banks Face in Small Open Economies?	Knütter, Rolf Wagner, Helmut
434	2008	International Policy Coordination and Simple Monetary Policy Rules	Berger, Wolfram Wagner, Helmut
435	2009	Matchingprozesse auf beruflichen Teilarbeitsmärkten	Stops, Michael Mazzoni, Thomas
436	2009	Wayfindingprozesse in Parksituationen - eine empirische Analyse	Fließ, Sabine Tetzner, Stefan
437	2009	ENTROPY-DRIVEN PORTFOLIO SELECTION a downside and upside risk framework	Rödder, Wilhelm Gartner, Ivan Ricardo Rudolph, Sandra
438	2009	Consulting Incentives in Contests	Schlepütz, Volker

439	2009	A Genetic Algorithm for a Bi-Objective Winner-Determination Problem in a Transportation-Procurement Auction"	Buer, Tobias Pankratz, Giselher
440	2009	Parallel greedy algorithms for packing unequal spheres into a cuboidal strip or a cuboid	Kubach, Timo Bortfeldt, Andreas Tilli, Thomas Gehring, Hermann
441	2009	SEM modeling with singular moment matrices Part I: ML-Estimation of time series	Singer, Hermann
442	2009	SEM modeling with singular moment matrices Part II: ML-Estimation of sampled stochastic differential equations	Singer, Hermann
443	2009	Konsensuale Effizienzbewertung und -verbesserung – Untersuchungen mittels der Data Envelopment Analysis (DEA)	Rödder, Wilhelm Reucher, Elmar
444	2009	Legal Uncertainty – Is Hamonization of Law the Right Answer? A Short Overview	Wagner, Helmut
445	2009	Fast Continuous-Discrete DAF-Filters	Mazzoni, Thomas
446	2010	Quantitative Evaluierung von Multi-Level Marketingsystemen	Lorenz, Marina Mazzoni, Thomas
447	2010	Quasi-Continuous Maximum Entropy Distribution Approximation with Kernel Density	Mazzoni, Thomas Reucher, Elmar
448	2010	Solving a Bi-Objective Winner Determination Problem in a Transportation Procurement Auction	Buer, Tobias Pankratz, Giselher
449	2010	Are Short Term Stock Asset Returns Predictable? An Extended Empirical Analysis	Mazzoni, Thomas
450	2010	Europäische Gesundheitssysteme im Vergleich – Effizienzmessungen von Akutkrankenhäusern mit DEA –	Reucher, Elmar Sartorius, Frank
451	2010	Patterns in Object-Oriented Analysis	Blaimer, Nicolas Bortfeldt, Andreas Pankratz, Giselher
452	2010	The Kuznets-Kaldor-Puzzle and Neutral Cross-Capital-Intensity Structural Change	Stijepic, Denis Wagner, Helmut
453	2010	Monetary Policy and Boom-Bust Cycles: The Role of Communication	Knütter, Rolf Wagner, Helmut
454	2010	Konsensuale Effizienzbewertung und –verbesserung mittels DEA – Output- vs. Inputorientierung –	Reucher, Elmar Rödder, Wilhelm
455	2010	Consistent Modeling of Risk Averse Behavior with Spectral Risk Measures	Wächter, Hans Peter Mazzoni, Thomas

456	2010	Der virtuelle Peer – Eine Anwendung der DEA zur konsensualen Effizienz-bewertung –	Reucher, Elmar
457	2010	A two-stage packing procedure for a Portuguese trading company	Moura, Ana Bortfeldt, Andreas
458	2010	A tree search algorithm for solving the multi-dimensional strip packing problem with guillotine cutting constraint	Bortfeldt, Andreas Jungmann, Sabine
459	2010	Equity and Efficiency in Regional Public Good Supply with Imperfect Labour Mobility – Horizontal versus Vertical Equalization	Arnold, Volker
460	2010	A hybrid algorithm for the capacitated vehicle routing problem with three-dimensional loading constraints	Bortfeldt, Andreas
461	2010	A tree search procedure for the container relocation problem	Forster, Florian Bortfeldt, Andreas
462	2011	Advanced X-Efficiencies for CCR- and BCC-Modell – Towards Peer-based DEA Controlling	Rödder, Wilhelm Reucher, Elmar
463	2011	The Effects of Central Bank Communication on Financial Stability: A Systematization of the Empirical Evidence	Knütter, Rolf Mohr, Benjamin Wagner, Helmut
464	2011	Lösungskonzepte zur Allokation von Kooperationsvorteilen in der kooperativen Transportdisposition	Strangmeier, Reinhard Fiedler, Matthias
465	2011	Grenzen einer Legitimation staatlicher Maßnahmen gegenüber Kreditinstituten zur Verhinderung von Banken- und Wirtschaftskrisen	Merbecks, Ute
466	2011	Controlling im Stadtmarketing – Eine Analyse des Hagener Schaufensterwettbewerbs 2010	Fließ, Sabine Bauer, Katharina
467	2011	A Structural Approach to Financial Stability: On the Beneficial Role of Regulatory Governance	Mohr, Benjamin Wagner, Helmut
468	2011	Data Envelopment Analysis - Skalenerträge und Kreuzskalenerträge	Wilhelm Rödder Andreas Dellnitz
469	2011	Controlling organisatorischer Entscheidungen: Konzeptionelle Überlegungen	Lindner, Florian Scherm, Ewald
470	2011	Orientierung in Dienstleistungsumgebungen – eine explorative Studie am Beispiel des Flughafen Frankfurt am Main	Fließ, Sabine Colaci, Antje Nesper, Jens

471	2011	Inequality aversion, income skewness and the theory of the welfare state	Weinreich, Daniel
472	2011	A tree search procedure for the container retrieval problem	Forster, Florian Bortfeldt, Andreas
473	2011	A Functional Approach to Pricing Complex Barrier Options	Mazzoni, Thomas
474	2011	Bologna-Prozess und neues Steuerungsmodell – auf Konfrontationskurs mit universitären Identitäten	Jost, Tobias Scherm, Ewald
475	2011	A reduction approach for solving the rectangle packing area minimization problem	Bortfeldt, Andreas
476	2011	Trade and Unemployment with Heterogeneous Firms: How Good Jobs Are Lost	Altenburg, Lutz
477	2012	Structural Change Patterns and Development: China in Comparison	Wagner, Helmut
478	2012	Demografische Risiken – Herausforderungen für das finanzwirtschaftliche Risikomanagement im Rahmen der betrieblichen Altersversorgung	Merbecks, Ute
479	2012	"It's all in the Mix!" – Internalizing Externalities with R&D Subsidies and Environmental Liability	Endres, Alfred Friehe, Tim Rundshagen, Bianca
480	2012	Ökonomische Interpretationen der Skalenvariablen u in der DEA	Dellnitz, Andreas Kleine, Andreas Rödder, Wilhelm
481	2012	Entropiebasierte Analyse von Interaktionen in Sozialen Netzwerken	Rödder, Wilhelm Brenner, Dominic Kulmann, Friedhelm
482	2013	Central Bank Independence and Financial Stability: A Tale of Perfect Harmony?	Berger, Wolfram Kißmer, Friedrich
483	2013	Energy generation with Directed Technical Change	Kollenbach, Gilbert
484	2013	Monetary Policy and Asset Prices: When Cleaning Up Hits the Zero Lower Bound	Berger, Wolfram Kißmer, Friedrich
485	2013	Superknoten in Sozialen Netzwerken – eine entropieoptimale Analyse	Brenner, Dominic, Rödder, Wilhelm, Kulmann, Friedhelm
486	2013	Stimmigkeit von Situation, Organisation und Person: Gestaltungsüberlegungen auf Basis des Informationsverarbeitungsansatzes	Julmi, Christian Lindner, Florian Scherm, Ewald
487	2014	Incentives for Advanced Abatement Technology Under National and International Permit Trading	Endres, Alfred Rundshagen, Bianca
488	2014	Dynamische Effizienzbewertung öffentlicher Dreispartentheater mit der Data Envelopment Analysis	Kleine, Andreas Hoffmann, Steffen

489	2015	Konsensuale Peer-Wahl in der DEA -- Effizienz vs. Skalenertrag	Dellnitz, Andreas Reucher, Elmar
490	2015	Makroprudenzielle Regulierung – eine kurze Einführung und ein Überblick	Velauthapillai, Jeyakrishna
491	2015	SEM modeling with singular moment matrices Part III: GLS estimation	Singer, Hermann
492	2015	Die steuerliche Berücksichtigung von Aufwendungen für ein Studium − Eine Darstellung unter besonderer Berücksichtigung des Hörerstatus	Meyering, Stephan Portheine, Kea
493	2016	Ungewissheit versus Unsicherheit in Sozialen Netzwerken	Rödder, Wilhelm Dellnitz, Andreas Gartner, Ivan
494	2016	Investments in supplier-specific economies of scope with two different services and different supplier characters: two specialists	Fandel, Günter Trockel, Jan
495	2016	An application of the put-call-parity to variance reduced Monte-Carlo option pricing	Müller, Armin
496	2016	A joint application of the put-call-parity and importance sampling to variance reduced option pricing	Müller, Armin
497	2016	Simulated Maximum Likelihood for Continuous-Discrete State Space Models using Langevin Importance Sampling	Singer, Hermann
498	2016	A Theory of Affective Communication	Julmi, Christian
499	2016	Approximations of option price elasticities for importance sampling	Müller, Armin
500	2016	Variance reduced Value at Risk Monte-Carlo simulations	Müller, Armin
501	2016	Maximum Likelihood Estimation of Continuous-Discrete State-Space Models: Langevin Path Sampling vs. Numerical Integration	Singer, Hermann
502	2016	Measuring the domain-specificity of creativity	Julmi, Christian Scherm, Ewald
503	2017	Bipartite Strukturen in Sozialen Netzen – klassische versus MaxEnt-Analysen	Rödder, Wilhelm Dellnitz, Andreas Kulmann, Friedhelm Litzinger, Sebastian Reucher, Elmar
504	2017	Langevin and Kalman Importance Sampling for Nonlinear Continuous-Discrete State Space Models	Singer, Hermann
505	2017	Horizontal versus vertical fiscal Equalization	Anetsberger, Georg Arnold, Volker
506	2017	Formative and Reflective Measurement Models	Singer, Hermann
507	2017	Identifizierung von führenden Köpfen in terroristischen Netzwerken – ein entropiebasiertes Verfahren –	Dellnitz, Andreas Litzinger, Sebastian Rödder, Wilhelm
508	2017	Die Bedeutung der steuerlichen Norm § 5 Abs. 2 EStG für die handelsrechtliche Rechnungslegung	Meyering, Stephan
509	2018	Ein erweitertes Effizienzmaß für DMUs im BCC-Modell – eine ökonomiegerechte DEA-Anpassung –	Rödder, Wilhelm Dellnitz, Andreas Litzinger, Sebastian

510	2018	A concise proof of Gaussian smoothing	Singer, Hermann
511	2018	Empirical evidence on the topological properties of structural paths and some notes on its theoretical explanation	Stijepic, Denis
512	2018	On the predictability of economic structural change by the Poincaré-Bendixson theory	Stijepic, Denis
513	2018	On development paths minimizing the aggregate labor-reallocation costs in the three-sector framework and an application to structural policy	Stijepic, Denis
514	2018	Models of Continuous Dynamics on the 2-Simplex and Applications in Economics	Stijepic, Denis
515	2018	A Note on the Ideological Content of Modern Economic Dynamics Models and Ideology-Reducing Meta-Modeling	Stijepic, Denis
516	2018	Logistik für Versand von Studienmaterialien der FernUniversität in Hagen – Optimierte Bereitstellung bei der Kommissionierung	Brenner, Dominic Gädeke, Andre Kulmann, Friedhelm Kleine, Andreas
517	2019	Zur Quantifizierung von Macht und Machtallianzen – ein struktureller Ansatz in sozialen Netzwerken –	Dellnitz, Andreas Rödder, Wilhelm
518	2019	Kolmogorov Backward Equations with Singular Diffusion Matrices	Singer, Hermann