Universität Konstanz

Forward Simulation of Financial Problems via BSDEs

Christian Bender
Robert Denk

Konstanzer Schriften in Mathematik und Informatik
Nr. 207, Mai 2005

ISSN 1430–3558

Konstanzer Online-Publikations-System (KOPS)
URL: http://www.ub.uni-konstanz.de/kops/volltexte/2007/2227/
URN: http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-22277

© Fachbereich Mathematik und Statistik
© Fachbereich Informatik und Informationswissenschaft
Universität Konstanz
Fach D 188, 78457 Konstanz, Germany
Email: preprints@informatik.uni–konstanz.de
WWW: http://www.informatik.uni–konstanz.de/Schriften/

Forward Simulation of Financial Problems via BSDEs[1]

Christian Bender and Robert Denk
WIAS Berlin and University of Konstanz

Abstract

We introduce a forward scheme to simulate backward SDEs and demonstrate the strength of the new algorithm by solving some financial problems numerically.

1 Introduction

The study of nonlinear backward stochastic differential equations (BSDEs) was initiated by Pardoux and Peng (1990). Mainly motivated by financial problems (see e.g. the survey article by El Karoui et al. (1997)) the theory of BSDEs was developed at high speed during the 1990s. Comparably slow progress has been made on the numerics of BSDEs.

Up to now basically two types of schemes have been considered. Based on the theoretical 4-step-scheme from Ma et al. (1994), numerical algorithms for BSDEs have been developed by Douglas et al. (1996) and more recently by Milstein and Tretyakov (2004). The main focus of these algorithms is the numerical solution of a parabolic PDE which is related to the BSDE.

A second type of algorithms works backwards through time and tries to tackle the stochastic problem directly. Since these algorithms can be motivated from a naive backward analogue of the Euler scheme (see Bouchard and Touzi (2004)), we shall refer to them as backward Euler schemes. Bally (1997) and Chevance (1997) were the first to study this type of algorithm with a (hardly implementable) random time partition respectively under strong regularity assumptions. The work of Ma et al. (2002) is in the same spirit, replacing, however, the Brownian motion by a binary random walk in the approximative equation. Only recently, a new notion of L^2-regularity on the control part of the solution was introduced in Zhang (2004), which allowed to prove convergence of backward Euler schemes with deterministic partitions under rather weak regularity assumptions, see Zhang (2004), Bouchard and Touzi (2004), and Gobet et al. (2004) for slightly different algorithms.

The main drawback of the backward Euler scheme is that it leads to nestings of conditional expectations backwards in time. Therefore, the computational costs for approximating the conditional expectations explode when the mesh

[1]C. Bender is supported by the DFG Research Center MATHEON 'Mathematics for key technologies' in Berlin. R. Denk is partially supported by the AFF grant 28/04 of the University of Konstanz. C. Bender thanks Shanjian Tang for the kind invitation to present this paper at the '4th Colloquium on BSDEs and Their Applications'.

of the partition tends to zero (see the discussion of the backward Euler scheme in section 2).

In Section 3 we introduce a new forward scheme for BSDEs which avoids nestings of conditional expectations backwards in time. Theoretical results as well as simulations of some financial problems (Sections 4–5) show that the computational costs of the new scheme for approximating the conditional expectation grow moderately when the mesh size tends to zero.

Throughout the paper we shall consider the following type of BSDE.

$$\begin{aligned} dX_t &= b(t, X_t)dt + \sigma(t, X_t)dW_t \\ dY_t &= f(t, X_t, Y_t, Z_t)dt + Z_t dW_t \\ X_0 &= x \\ Y_T &= \xi \end{aligned} \tag{1}$$

Here $W_t = (W_{1,t}, \dots, W_{D,t})^*$ is a D-dimensional Brownian motion on $[0, T]$ and $Z_t = (Z_{1,t}, \dots, Z_{D,t})$. The process X is $\mathbb{R}^M$-valued and the process Y is $\mathbb{R}$-valued. We shall always assume:

Standing Assumption 1.1. *There is a constant K such that*

$$\begin{aligned} & |b(t,x) - b(t',x')| + |\sigma(t,x) - \sigma(t',x')| + |f(t,x,y,z) - f(t',x',y',z')| \\ \leq \; & K(\sqrt{|t-t'|} + |x-x'| + |y-y'| + |z-z'|) \end{aligned}$$

for all (t,x,y,z), $(t',x',y',z') \in [0,T] \times \mathbb{R}^M \times \mathbb{R} \times \mathbb{R}^D$,

$$\xi = \Phi(X)$$

where Φ is a functional on the space of RCLL-functions on $[0,T]$ satisfying the L^∞-Lipschitz condition,

$$|\Phi(\mathbf{x}) - \Phi(\mathbf{x}')| \leq K \sup_{0 \leq t \leq T} |\mathbf{x}(t) - \mathbf{x}'(t)|$$

for all RCLL-functions $\mathbf{x}$, $\mathbf{x}'$. Moreover,

$$\sup_{0 \leq t \leq T} (|b(t,0)| + |\sigma(t,0)| + |f(t,0,0,0)|) + |\Phi(\mathbf{0})| \leq K$$

where $\mathbf{0}$ denotes the constant function taking value 0 on $[0,T]$.

Note that that the matrix σ is neither assumed to be quadratic nor $\sigma\sigma^*$ to be invertible.

Remark 1.1. We say that a constant depends on the data if it is depending on K, T, x_0 and the dimensions M and D only. Throughout the paper, C denotes a generic constant depending on the data which may vary from line to line.

2 The Backward Euler Scheme Revisited

In this section we review some recent progress on the backward Euler scheme due to Bouchard and Touzi (2004) and Zhang (2004).

The basic idea of the backward Euler scheme is to replace the BSDE by sort of a semi-discretized BSDE. Let a partition $\pi = \{t_0, t_1, \ldots, t_N\}$ of $[0, T]$ be given and consider the BSDE

$$\begin{aligned} dY_t^{(\infty,\pi)} &= f^{(\pi)}(t)dt + \tilde{Z}_t^{(\infty,\pi)} dW_t \\ Y_T^{(\infty,\pi)} &= \xi^{(\pi)} \end{aligned} \quad (2)$$

where

$$f^{(\pi)}(t) = \sum_{i=0}^{N-1} f(t_i, X_{t_i}^{(\pi)}, Y_{t_i}^{(\infty,\pi)}, \tilde{Z}_{t_i}^{(\infty,\pi)}) \mathbf{1}_{(t_i, t_{i+1}]}(t)$$

and $X^{(\pi)}$, $\xi^{(\pi)}$ are some square integrable approximations of X, ξ, respectively. Note, the Lebesgue integral is discretized only, not the Itô integral.

It is easily seen that the BSDE (2) admits a unique solution $(Y_t^{(\infty,\pi)}, \tilde{Z}_t^{(\infty,\pi)})$ (satisfying the usual integrability conditions), if $Y_t^{(\infty,\pi)}$ is additionally supposed to be right-continuous, provided the mesh size $|\pi|$ of the partition is sufficiently fine. Indeed, at the time points of the partition one has

$$\begin{aligned} Y_{t_N}^{(\infty,\pi)} &= \xi^{(\pi)} \\ \tilde{Z}_{d,t_i}^{(\infty,\pi)} &= E\left[\frac{W_{t_{i+1},d} - W_{t_i,d}}{t_{i+1} - t_i} Y_{t_{i+1}}^{(\infty,\pi)} \middle| \mathcal{F}_{t_i} \right] \\ Y_{t_i}^{(\infty,\pi)} &= E[Y_{t_{i+1}}^{(\infty,\pi)} | \mathcal{F}_{t_i}] - f(t_i, X_{t_i}^{(\pi)}, Y_{t_i}^{(\infty,\pi)}, \tilde{Z}_{t_i}^{(\infty,\pi)})(t_{i+1} - t_i) \end{aligned} \quad (3)$$

$Y_t^{(\infty,\pi)}$ is obtained by right-continuous constant interpolation, while the interpolation of $\tilde{Z}_t^{(\infty,\pi)}$ is implicitly constructed via the martingale representation theorem. We denote by $Z_t^{(\infty,\pi)}$ the piecewise constant RCLL process, which coincides with $\tilde{Z}_t^{(\infty,\pi)}$ at the points of the partition and has jumps at these points only.

The following theorem is a slight extension of Theorem 3.1 in Bouchard and Touzi (2004) concerning the assumptions on the data, which in particular allows for path-dependent terminal data. It may be obtained by combining ideas of Bouchard and Touzi (2004) and Zhang (2004) (see the appendix of Bender and Denk (2005) for details):

Theorem 2.1. *Suppose assumption 1.1 holds, and the discretization $X^{(\pi)}$ of X satisfies*

$$\sup_{0 \le t \le T} E\left[|X_t - X_t^{(\pi)}|^2 \right] \le C|\pi| \quad (4)$$

for some constant C depending on the data. Then there is a constant C depending on the data such that

$$\begin{aligned} &\sup_{0 \le t \le T} E\left[\left| Y_t - Y_t^{(\infty,\pi)} \right|^2 \right] + E\int_0^T |Z_t - Z_t^{(\infty,\pi)}|^2 dt \\ &\le C\left(|\pi| + E[|\xi - \xi^{(\pi)}|^2] \right) \end{aligned}$$

provided $|\pi|$ is sufficiently small.

Remark 2.1. (i) Note that the condition on the discretization $X^{(\pi)}$ of X is, for instance, satisfied by the Euler scheme.
(ii) In the above theorem $Z_t^{(\infty,\pi)}$ may be replaced by $\tilde{Z}_t^{(\infty,\pi)}$.
(iii) In Zhang (2004) convergence of a slightly different approximation scheme is proved.

To make the above approximation scheme implementable one has to approximate the conditional expectations in (3). As $Y_{t_i}^{(\infty,\pi)}$ is calculated from the conditional expectation of $Y_{t_{i+1}}^{(\infty,\pi)}$ the approximation error of the conditional expectation is expected to propagate backwards in time. We denote by $\widehat{E}^\pi[\cdot|\mathcal{F}_t]$ a generic estimator for the conditional expectation. Let

$$\begin{aligned}
\widehat{Y}_{t_N}^{(\infty,\pi)} &= \xi^{(\pi)} \\
\widehat{Z}_{d,t_i}^{(\infty,\pi)} &= \widehat{E}^\pi\left[\left.\frac{W_{t_{i+1},d}-W_{t_i,d}}{t_{i+1}-t_i}\widehat{Y}_{t_{i+1}}^{(\infty,\pi)}\right|\mathcal{F}_{t_i}\right] \\
\widehat{Y}_{t_i}^{(\infty,\pi)} &= \widehat{E}^\pi[\widehat{Y}_{t_{i+1}}^{(\infty,\pi)}|\mathcal{F}_{t_i}] - f(t_i, X_{t_i}^{(\pi)}, \widehat{Y}_{t_i}^{(\infty,\pi)}, \widehat{Z}_{t_i}^{(\infty,\pi)})(t_{i+1}-t_i) \qquad (5)
\end{aligned}$$

Bouchard and Touzi (2004), Theorem 4.1, prove the following theorem under slightly stronger assumptions than Assumption 1.1:

Theorem 2.2. *There is a constant C depending on the data such that*

$$\begin{aligned}
&\max_{0\le i\le N} E[|\widehat{Y}_{t_i}^{(\infty,\pi)} - Y_{t_i}^{(\infty,\pi)}|^2] \\
\le\; &\frac{C}{|\pi|}\max_{0\le j\le N} E\Bigg(|\widehat{E}^\pi[\widehat{Y}_{t_{i+1}}^{(\infty,\pi)}|\mathcal{F}_{t_i}] - E[\widehat{Y}_{t_{i+1}}^{(\infty,\pi)}|\mathcal{F}_{t_i}]|^2 \\
&+\left|\widehat{E}^\pi\left[\left.\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\widehat{Y}_{t_{i+1}}^{(\infty,\pi)}\right|\mathcal{F}_{t_i}\right] - E\left[\left.\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\widehat{Y}_{t_{i+1}}^{(\infty,\pi)}\right|\mathcal{F}_{t_i}\right]\right|^2\Bigg)
\end{aligned}$$

This generic theorem highlights one of the main drawbacks of the backward Euler scheme: Given the same accuracy of the conditional expectation estimator the error due to the approximation of the conditional expectation explodes when the mesh of the partition tends to zero. Put differently, due to the numerical approximation of the conditional expectation one has to simulate the more paths the finer the partition. This increases the computational costs. This effect is particularly unfavorable when the constant in Theorem 2.1 is large (e.g. due to a large Lipschitz constant or time horizon) and, thus, a fine mesh is needed for $Y_t^{(\infty,\pi)}$ to be a good approximation of Y_t. We note that the described effect has also been observed in the numerical examples by Gobet et al. (2004).

3 A Discretization of the Picard Iteration

We will now introduce a new approximation scheme which avoids nestings of conditional expectation backwards in time. Instead it mimics the Picard iteration for BSDEs.

Recall that the backward part (Y, Z) of the system (1) can be obtained as the limit of a Picard type iteration $(Y^{(n)}, Z^{(n)})$, see e.g. Yong and Zhou (2000), theorem 7.3.4. Here $(Y^{(0)}, Z^{(0)}) \equiv (0,0)$, and $(Y^{(n)}, Z^{(n)})$ is the solution of the simple BSDE

$$\begin{aligned} dY_t^{(n)} &= f(t, X_t, Y_t^{(n-1)}, Z_t^{(n-1)})dt + Z_t^{(n)}dW_t \\ Y_T^{(n)} &= \xi \end{aligned}$$

with X as above.

The solution is given by

$$Y_t^{(n)} = E\left[\xi - \int_t^T f(s, X_s, Y_s^{(n-1)}, Z_s^{(n-1)})ds \middle| \mathcal{F}_t\right]$$

and $Z^{(n)}$ is obtained via the martingale representation theorem. As is emphasized in Yong and Zhou (2000), Ch. 7, the above Picard iteration is still implicit due to the use of the martingale representation theorem.

We will now introduce a time discretization of the above Picard iteration, which is explicit but for the occurrence of conditional expectations.

Suppose a partition $\pi = \{t_0, t_1, \ldots, t_N\}$ of $[0,T]$ is given and a corresponding discretization $X^{(\pi)}$ of X as well as some approximation $\xi^{(\pi)}$ of ξ. Let $(Y^{(0,\pi)}, Z^{(0,\pi)}) \equiv (0,0)$. Then define iteratively, with $\Delta_i = t_{i+1} - t_i$ and $\Delta W_{d,i} = W_{d,t_{i+1}} - W_{d,t_i}$,

$$\begin{aligned} Y_{t_i}^{(n,\pi)} &= E\left[\xi^{(\pi)} - \sum_{j=i}^{N-1} f(t_j, X_{t_j}^{(\pi)}, Y_{t_j}^{(n-1,\pi)}, Z_{t_j}^{(n-1,\pi)})\Delta_j \middle| \mathcal{F}_{t_i}\right] \\ Z_{d,t_i}^{(n,\pi)} &= E\left[\frac{\Delta W_{d,i}}{\Delta_i}\left(\xi^{(\pi)} - \sum_{j=i+1}^{N-1} f(t_j, X_{t_j}^{(\pi)}, Y_{t_j}^{(n-1,\pi)}, Z_{t_j}^{(n-1,\pi)})\Delta_j\right) \middle| \mathcal{F}_{t_i}\right] \end{aligned}$$

The processes $Y^{(n,\pi)}$ and $Z^{(n,\pi)}$ are extended to RCLL processes by constant interpolation.

We can now state one of the main results of the present paper:

Theorem 3.1. *Suppose Assumption 1.1 holds, and suppose that for some constant C depending on the data we have*

$$\begin{aligned} \sup_{0\le t\le T} E\left[|X_t - X_t^{(\pi)}|^2\right] &\le C|\pi|, \\ \sup_{|\pi|\le 1} E\left[|\xi^{(\pi)}|^2\right] &\le C. \end{aligned}$$

Then there is a constant C depending on the data such that

$$\begin{aligned} &\sup_{0\le t\le T} E\left[\left|Y_t - Y_t^{(n,\pi)}\right|^2\right] + E\int_0^T |Z_t - Z_t^{(n,\pi)}|^2 dt \\ \le\ & C\left(|\pi| + E[|\xi - \xi^{(\pi)}|^2] + \left(\frac{1}{2} + C|\pi|\right)^n\right) \end{aligned}$$

provided $|\pi|$ is sufficiently small.

Remark 3.1. The condition on $\xi^{(\pi)}$ is satisfied whenever for $|\pi| \leq 1$

$$E[|\xi - \xi^{(\pi)}|^2] \leq C|\pi|^\alpha$$

with some constant C depending on the data and some $\alpha > 0$. Indeed,

$$E[|\xi^{(\pi)}|^2] \leq 2E[|\xi|^2] + 2E[|\xi - \xi^{(\pi)}|^2],$$

and, thanks to the L^∞-Lipschitz condition and a classical estimate for SDEs,

$$\begin{aligned} E[|\xi|^2] &\leq 2K^2 E[\sup_{0\leq t\leq T} |X_t|^2] + 2|\Phi(\mathbf{0})|^2 \\ &\leq C\left(x^2 + \int_0^T |b(t,0)|^2 + |\sigma(t,0)|^2 dt\right) + 2K^2 \leq C \end{aligned}$$

The proof of Theorem 3.1 basically relies on the following a priori estimates in a weighted L^2-norm, which are proved in Bender and Denk (2005).

Lemma 3.2. *Suppose* Γ *and* γ *are positive real numbers* $\tilde{y}^{(\iota)}$, $\tilde{z}^{(\iota)}$, $\iota = 1, 2$ *are adapted processes and*

$$\begin{aligned} \tilde{Y}_{t_i}^{(\iota)} &= E\left[\xi^{(\pi)} - \sum_{j=i}^{N-1} f(t_j, X_{t_j}^{(\pi)}, \tilde{y}_{t_j}^{(\iota)}, \tilde{z}_{t_j}^{(\iota)})\Delta_j \middle| \mathcal{F}_{t_i}\right], \\ \tilde{Z}_{d,t_i}^{(\iota)} &= E\left[\frac{\Delta W_{d,i}}{\Delta_i}\left(\xi^{(\pi)} - \sum_{j=i+1}^{N-1} f(t_j, X_{t_j}^{(\pi)}, \tilde{y}_{t_j}^{(\iota)}, \tilde{z}_{t_j}^{(\iota)})\Delta_j\right) \middle| \mathcal{F}_{t_i}\right]. \end{aligned}$$

Moreover, assume that f *is Lipschitz in* (y,z) *uniformly in* (t,x) *with constant* K *and* $\xi^{(\pi)}$ *is square integrable. Then*

$$\begin{aligned} &\max_{0\leq i\leq N} \lambda_i E\left[|\tilde{Y}_{t_i}^{(1)} - \tilde{Y}_{t_i}^{(2)}|^2\right] + \sum_{i=0}^{N-1} \lambda_i E\left[|\tilde{Z}_{t_i}^{(1)} - \tilde{Z}_{t_i}^{(2)}|^2\right]\Delta_i \\ \leq\ & K^2(T+1)\left(\left(|\pi| + \Gamma^{-1}\right)(\gamma DT + 1) + \frac{D}{\gamma}\right) \\ &\times\left(\max_{0\leq i\leq N} \lambda_i E\left[|\tilde{y}_{t_i}^{(1)} - \tilde{y}_{t_i}^{(2)}|^2\right] + \sum_{i=0}^{N-1} \lambda_i E\left[|\tilde{z}_{t_i}^{(1)} - \tilde{z}_{t_i}^{(2)}|^2\right]\Delta_i\right). \end{aligned}$$

where $\lambda_0 = 1$ *and* $\lambda_i = (1 + \Gamma\Delta_{i-1})\lambda_{i-1}$.

Indeed, it is rather straightforward to derive from these a priori estimates that for sufficiently large Γ and sufficiently fine partition π,

$$\begin{aligned} &\max_{0\leq i\leq N} E\left[\left|Y_{t_i}^{(\infty,\pi)} - Y_{t_i}^{(n,\pi)}\right|^2\right] + \sum_{i=0}^{N-1} E\left[\left|Z_{t_i}^{(\infty,\pi)} - Z_{t_i}^{(n,\pi)}\right|^2\right]\Delta_i \\ \leq\ & e^{\Gamma T}\left(\max_{0\leq i\leq N} E\left[|Y_{t_i}^{(1,\pi)}|^2\right] + \sum_{i=0}^{N-1} E\left[|Z_{t_i}^{(1,\pi)}|^2\right]\Delta_i\right)\left(1 - \sqrt{\frac{\Gamma|\pi|}{4} + \frac{1}{2}}\right)^{-2} \\ &\times\left(\frac{\Gamma|\pi|}{4} + \frac{1}{2}\right)^n \end{aligned}$$

Thus, Theorem 3.1 follows from Theorem 2.1 since

$$\left(\max_{0\le i\le N} E\left[|Y_{t_i}^{(1,\pi)}|^2\right] + \sum_{i=0}^{N-1} E\left[|Z_{t_i}^{(1,\pi)}|^2\right]\Delta_i \right)$$

can be shown to be bounded by a constant which depends on the data only.

As for the backward Euler scheme we shall next investigate the error due to approximation of the conditional expectation for the discrete Picard iteration. Since the discrete Picard iteration has no nestings of conditional expectations backwards in time we expect no explicit dependence of the error from the mesh of the partition. However, there are nestings of conditional expectations in the iterations n. It turns out, that the influence of earlier errors in the Picard iteration becomes smaller and smaller and that the weights of these errors form a convergent series series. To state this result, $\widehat{E}^{\pi}[\cdot|\mathcal{F}_t]$ again denotes a generic estimator of the conditional expectation. Moreover we denote

$$\begin{aligned}
\widehat{b}_i^{(n,\pi)} &= \xi^{(\pi)} - \sum_{j=i}^{N-1} f(t_j, X_{t_j}^{(\pi)}, \widehat{Y}_{t_j}^{(n-1,\pi)}, \widehat{Z}_{t_j}^{(n-1,\pi)})\Delta_j \\
\widehat{Y}_{t_i}^{(n,\pi)} &= \widehat{E}[\widehat{b}_i^{(n,\pi)}|\mathcal{F}_{t_i}] \\
\widehat{Z}_{d,t_i}^{(n,\pi)} &= \widehat{E}\left[\frac{\Delta W_{d,i}}{\Delta i}\widehat{b}_{i+1}^{(n,\pi)} \middle| \mathcal{F}_{t_i} \right]
\end{aligned}$$

initialized at $(\widehat{Y}^{(0,\pi)}, \widehat{Z}^{(0,\pi)}) = (0,0)$.

Theorem 3.3. *Under Assumption 1.1 there is a constant C depending on the data such that for any sufficiently fine partition π,*

$$\begin{aligned}
&\max_{0\le i\le N} E[|\widehat{Y}_{t_i}^{(n,\pi)} - Y_{t_i}^{(n,\pi)}|^2] + \sum_{i=0}^{N-1} E[|\widehat{Z}_{t_i}^{(n,\pi)} - Z_{t_i}^{(n,\pi)}|^2]\Delta_i \\
\le\ & C \max_{1\le\nu\le n}\left(\max_{0\le i\le N} E\left[|\widehat{E}^{\pi}[\widehat{b}_i^{(\nu,\pi)}|\mathcal{F}_{t_i}] - E[\widehat{b}_i^{(\nu,\pi)}|\mathcal{F}_{t_i}]|^2\right]\right. \\
&\left. + E\sum_{i=0}^{N-1}\left|\widehat{E}^{\pi}\left[\frac{\Delta W_i}{\Delta_i}\widehat{b}_{i+1}^{(\nu,\pi)}\middle|\mathcal{F}_{t_i}\right] - E\left[\frac{\Delta W_i}{\Delta_i}\widehat{b}_{i+1}^{(\nu,\pi)}\middle|\mathcal{F}_{t_i}\right]\right|^2\Delta_i\right)
\end{aligned}$$

Sketch of the proof. By Young's inequality and the a priori estimates (Lemma 3.2) we obtain the following estimates in the weighted norm (when γ and Γ are chosen appropriately),

$$\begin{aligned}
&\max_{0\le i\le N}\lambda_i E[|\widehat{Y}^{(n,\pi)}_{t_i}-Y^{(n,\pi)}_{t_i}|^2]+\sum_{i=0}^{N-1}\lambda_i E[|\widehat{Z}^{(n,\pi)}_{t_i}-Z^{(n,\pi)}_{t_i}|^2]\Delta_i\\
\le\;& 2\Bigg(\max_{0\le i\le N}\lambda_i E\left[|\widehat{E}^{\pi}[\widehat{b}^{(n,\pi)}_i|\mathcal{F}_{t_i}]-E[\widehat{b}^{(n,\pi)}_i|\mathcal{F}_{t_i}]|^2\right]\\
&+E\sum_{i=0}^{N-1}\lambda_i\left|\widehat{E}^{\pi}\left[\frac{\Delta W_i}{\Delta_i}\widehat{b}^{(n,\pi)}_{i+1}\middle|\mathcal{F}_{t_i}\right]-E\left[\frac{\Delta W_i}{\Delta_i}\widehat{b}^{(n,\pi)}_{i+1}\middle|\mathcal{F}_{t_i}\right]\right|^2\Delta_i\Bigg)\\
&+\left(\frac{1}{4}+\Gamma|\pi|\right)\Bigg(\max_{0\le i\le N}\lambda_i E[|\widehat{Y}^{(n-1,\pi)}_{t_i}-Y^{(n-1,\pi)}_{t_i}|^2]\\
&+\sum_{i=0}^{N-1}\lambda_i E[|\widehat{Z}^{(n-1,\pi)}_{t_i}-Z^{(n-1,\pi)}_{t_i}|^2]\Delta_i\Bigg)
\end{aligned}$$

Now for $|\pi|$ sufficiently small (e.g. less or equal $(4\Gamma)^{-1}$) the above estimate can be iterated to obtain the theorem. □

Theorem 3.3 clarifies the main advantage of the discretized Picard iteration compared to the backward Euler scheme: Given the accuracy of the estimator of the conditional expectation, the total error of approximating the conditional expectations does neither explode when the mesh of the partition goes to zero nor when the number of iterations tends to infinity.

4 A Numerical Forward Scheme

Before we can present some simulations we have to specify the approximation of the conditional expectation. We shall utilize the so-called least-squares Monte-Carlo regression method, which was introduced in Longstaff and Schwartz (2001) in the context of American options and is also applied to the backward Euler scheme in Gobet et al. (2004). For notational covenience we shall assume $D = 1$.

Discretization of X: We discretize X by the Euler scheme

$$\begin{aligned}
X^{(\pi)}_0 &= x\\
X^{(\pi)}_{t_i} &= X^{(\pi)}_{t_{i-1}}+b(t_{i-1},X^{(\pi)}_{t_{i-1}})\Delta_{i-1}+\sigma(t_{i-1},X^{(\pi)}_{t_{i-1}})\Delta W_{i-1}
\end{aligned}$$

and extend $X^{(\pi)}$ to an RCLL process by piecewise constant interpolation. When X is known to be strictly positive, it can be more convenient to apply the Euler scheme to $\ln(X)$ instead of X, see Gobet et al. (2004). Note that $(X^{(\pi)}_{t_i},\mathcal{F}_{t_i})$ forms a Markov chain.

Discretization of the terminal condition: The approximation $\xi^{(\pi)}$ of the terminal condition is supposed to be of the form

$$\xi^{(\pi)}=\Phi^{(\pi)}(\Xi^{(\pi)}_{t_N})$$

where $(\Xi^{(\pi)}_{t_i},\mathcal{F}_{t_i})$ is an M'-dimensional Markov chain with $X^{(\pi)}_{t_i}$ as its first M components and $\Phi^{(\pi)}$ is a deterministic function. We refer to Zhang (2004), Corollary 4.4, for some convergence results of such $\xi^{(\pi)}$ to terminal conditions which satisfy the standing assumption.

Approximation of the conditional expectations: Choose a set of basis functions $e_1(x), \ldots, e_\kappa(x)$ such that $\eta_k^i = e_k(\Xi_{t_i}^{(\pi)})$ are square integrable. In a first step the conditional expectation at time t_i is replaced by the orthogonal projection on span$\{\eta_1^i, \ldots, \eta_\kappa^i\}$. Then the coefficients of the orthogonal projection are approximated by the least squares estimator. Precisely, suppose that $L > \kappa$ independent samples $(\Delta W_i^\lambda)_{1\leq\lambda\leq L}$ of ΔW_i are given. The corresponding copies of $\Xi^{(\pi)}$, $\xi^{(\pi)}$, and η_k^i are denoted by $\Xi^{(\pi,\lambda)}$, $\xi^{(\pi,\lambda)}$, and $\eta_k^{(i,\lambda)}$. Define

$$\mathcal{A}_i^L = \frac{1}{\sqrt{L}} \left(\eta_k^{(i,\lambda)}\right)_{\lambda=1,\ldots,L,k=1,\ldots,K(i)}$$

and denote its pseudo-inverse by $(\mathcal{A}_i^L)^+$. Then the following recursion approximates the coefficients of the orthogonal projection:

$$\begin{aligned}
\alpha_{i,k}^{(0,\pi,L)} &= \tilde{\alpha}_{i,k}^{(0,\pi,L)} = 0 \\
Y_{t_i}^{(n-1,\pi,\lambda)} &= \sum_{k=1}^{\kappa} \alpha_{i,k}^{(n-1,\pi,L)} \eta_k^{(i,\lambda)} \\
Z_{t_i}^{(n-1,\pi,\lambda)} &= \sum_{k=1}^{\kappa} \tilde{\alpha}_{i,k}^{(n-1,\pi,L)} \eta_k^{(i,\lambda)} \\
\alpha_{i,\cdot}^{(n,\pi,L)} &= \frac{1}{\sqrt{L}} (\mathcal{A}_i^L)^+ \left(\xi^{(\pi,\cdot)} - \sum_{j=i}^{N-1} f(t_j, X_{t_j}^{(\pi,\cdot)}, Y_{t_j}^{(n-1,\pi,\cdot)}, Z_{t_j}^{(n-1,\pi,\cdot)}) \Delta_j \right) \\
\tilde{\alpha}_{i,\cdot}^{(n,\pi,L)} &= \frac{1}{\sqrt{L}} (\mathcal{A}_i^L)^+ \\
&\times \left(\frac{\Delta W_i^{(\cdot)}}{\Delta_i} \left(\xi^{(\pi,\cdot)} - \sum_{j=i+1}^{N-1} f(t_j, X_{t_j}^{(\pi,\cdot)}, Y_{t_j}^{(n-1,\pi,\cdot)}, Z_{t_j}^{(n-1,\pi,\cdot)}) \Delta_j \right) \right)
\end{aligned}$$

The simulation based estimators are now defined by,

$$\begin{aligned}
Y_{t_i}^{(n,\pi,L,*)} &= \sum_{k=1}^{\kappa} \alpha_{i,k}^{(n,\pi,L)} \eta_k^i \\
Z_{t_i}^{(n,\pi,L,*)} &= \sum_{k=1}^{\kappa} \tilde{\alpha}_{i,k}^{(n,\pi,L)} \eta_k^i
\end{aligned}$$

The approximation error of this scheme is analyzed in Bender and Denk (2005). In particular L^2-convergence of a truncated version of this scheme is shown, when $\{e_1(x), \ldots, e_\kappa(x)\}$ are the initial elements of a sequence $(e_k)_{k\in\mathbb{N}}$ such that

$$(e_k(\Xi_{t_i}^{(\pi)}))_{k\in\mathbb{N}}$$

is total in $L^2(\sigma(\Xi_{t_i}^{(\pi)}))$ and are linearly independent for all $0 \leq i \leq N-1$.

5 Simulation of Financial Problems

Throughout this section the process X is one-dimensional representing a stock in the standard Black-Scholes model, i.e.

$$X_t = X_0 \exp\{\sigma W_t + \mu t - 1/2\sigma^2 t\}$$

It is discretized by the log-Euler scheme. In all cases we will apply an equidistant partition of the interval $[0,T]$ with $N+1$ points denoted by π_N.

5.1 Different Interest Rate for Borrowing

In the first example we numerically evaluate a straddle, i.e. the sum of a call and a put option, under different rates for borrowing and investing in the money market account. The rate for borrowing is denoted by R, the one for investing by r. The fair price of a straddle in this model is given by Y_0, where (Y,Z) is the solution of the nonlinear BSDE

$$\begin{aligned} dY_t &= \left[rY_t + \frac{\mu - r}{\sigma} Z_t - (R-r)\left(Y_t - \frac{Z_t}{\sigma}\right)_- \right] dt + Z_t dW_t \\ Y_T &= |X_T - K|, \end{aligned}$$

see Bergman (1995). In the following we fix the parameters $X_0 = 100$, $\sigma = 0.2$, $\mu = 0.05$, $r = 0.01$, $R = 0.06$, and the straddle is supposed to be at the money, i.e. $K = 100$. In the figures below this situation is the 'nonlinear case', which will be compared with the standard 'linear case' where $R = 0.01$, i.e. the same interest rate is applied for borrowing and investing. We stop the Picard iteration, when the distance of two subsequent time-zero-values is less than 0.0001. The total number of calculated iterations is denoted by n_{stop}. We compare two different bases. The first basis consists of monomials and the straddle payoff, the second of characteristic functions. Precisely,

$$\begin{aligned} e_1^{(1)}(x) &= |x-K|, \quad e_k^{(1)}(x) = (x-X_0)^{k-2},\ 2 \le k \le \kappa \\ e_1^{(2)}(x) &= \mathbf{1}_{[0,l)}(x), \quad e_2^{(2)}(x) = \mathbf{1}_{[u,\infty)}(x), \\ e_k^{(2)}(x) &= \mathbf{1}_{[l+(k-3)(u-l)/(\kappa-2)),l+(k-2)(u-l)/(\kappa-2))}(x), \quad 3 \le k \le \kappa \end{aligned}$$

Here, the lower bound l and the upper bound u depend on i and the simulations. They are calculated as the empirical mean of $X_{t_i}^{(\pi_N,\lambda)}$ minus (resp. plus) two times their empirical standard deviation. Figure 1 shows the simulated price of the straddle for a maturity of $T = 2$ years as a function of the number of partition points for both bases. We choose $\kappa = 7$ for the basis $(e_k^{(1)})_k$, respectively $\kappa = 21$ for $(e_k^{(2)})_k$. In both cases we simulate $L = 100000$ paths. The relative standard error in the calculation of $Y_0^{(n_{stop},\pi_N,100000,*)}$ is about 0.28% for the nonlinear case and 0.29% for the linear case for both bases. The relative standard error does not change significantly in the number of partition points N. Thus, the simulation complements the assertion of theorem 3.3.

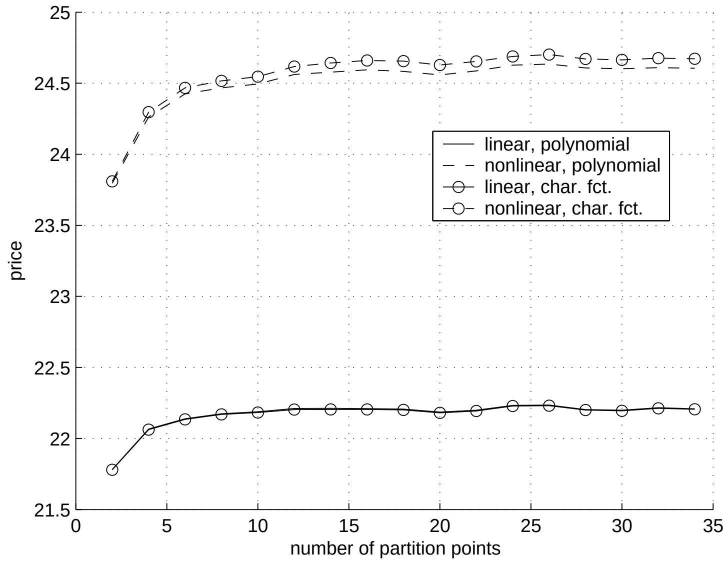

Figure 1: $Y_0^{(n_{stop},\pi_N,100000,*)}$ as a function of N for $T = 2$.

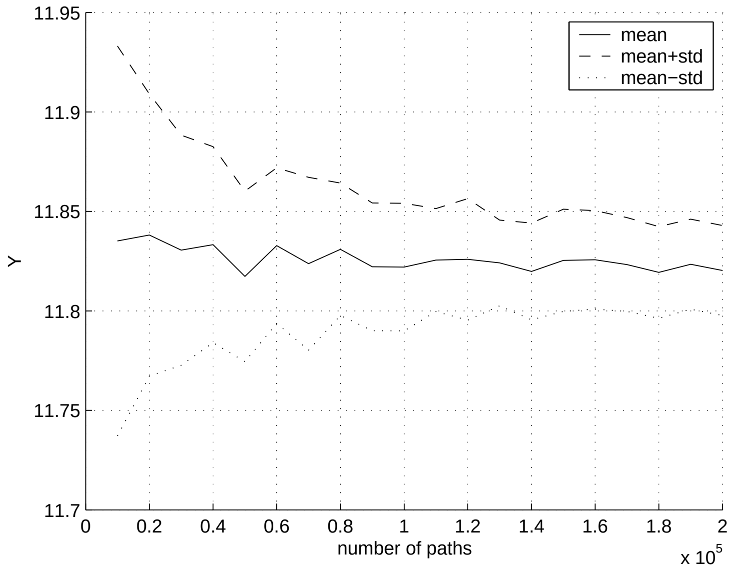

Figure 2: Empirical mean and standard deviation of 100 launches as function of L for $T = 0.5$.

Figure 2 shows the empirical mean and the empirical standard deviation of the simulated price calculated from 100 launches of the algorithm as a function of the number of simulated paths L per launch. Here $N = 20$ and $T = 0.5$. The simulations have been performed with the monomial basis and $\kappa = 5$ for the nonlinear case.

5.2 Constraints on Borrowing

The second example concerns borrowing constraints. Suppose an investor must not borrow an arbitrary amount of money from the money market account but a given fraction of his total wealth only. His goal is to super-replicate a given contingent claim (in our case a call option) with minimal initial wealth. This problem is known as superhedging problem. It is shown in Bender and Kohlmann (2004), extending results of El Karoui et al. (1997), that for quite general constraints the solution of the superhedging problem can be obtained as a limit of a sequence of nonlinear BSDEs. This sequence has an intuitive meaning: The investor is bound to yield an increasing penalization payment when he fails to meet the constraint. In the simple borrowing constraint under consideration the optimal superhedging price can be obtained as the limit of Y_0^ϵ (as ϵ tends to zero), where

$$\begin{aligned} dY_t^\epsilon &= \left[rY_t^\epsilon + \frac{\mu - r}{\sigma} Z_t^\epsilon - \frac{1}{\epsilon} \left(\frac{Z_t^\epsilon}{\sigma} - \rho Y_t^\epsilon \right)_+ \right] dt + Z_t^\epsilon dW_t \\ Y_T^\epsilon &= (X_T - K)_+. \end{aligned}$$

Here $\rho - 1$ is the fraction of his total wealth, which the investor is allowed to borrow. We consider the case $\rho = 10$ with the parameters $\sigma = 0.2$, $\mu = r = 0.05$, and $X_0 = K = 100$. The maturity is $T = 0.5$ years. Note, in this example the superhedging price can be determined analytically by calculating an equivalent dominating, but unconstrained, claim, see Broadie et al. (1998). It is 8.058.

We compute numerical approximations for different values of ϵ. The stop criterion for the Picard iteration is 0.001 and we choose $N = 40$ and the monomial basis with $\kappa = 5$, but the straddle payoff replaced by the call payoff. Figure 3 shows the corresponding approximation of Y_0^ϵ as function of ϵ for different numbers of simulated paths.

Figure 3 indicates that, due to the nonlinearity, the estimator for the conditional expectation has a positive bias. Indeed, the simulated ϵ-approximation tend to merge into a straight line (as function of ϵ^{-1}), when ϵ (depending on the number of paths) is sufficient small. Since the curves for 100000 paths and 200000 paths are almost parallel this effect can not be mended by solely enlarging the number of simulations. Preliminary simulations suggest that a larger number of partition points, an enlarged basis, and the simulation of more paths are needed to obtain accurate approximations of the ϵ-price, the higher the penalization. To achieve this with reasonable computational cost, variance reduction techniques are called for. This issue is left to future researches.

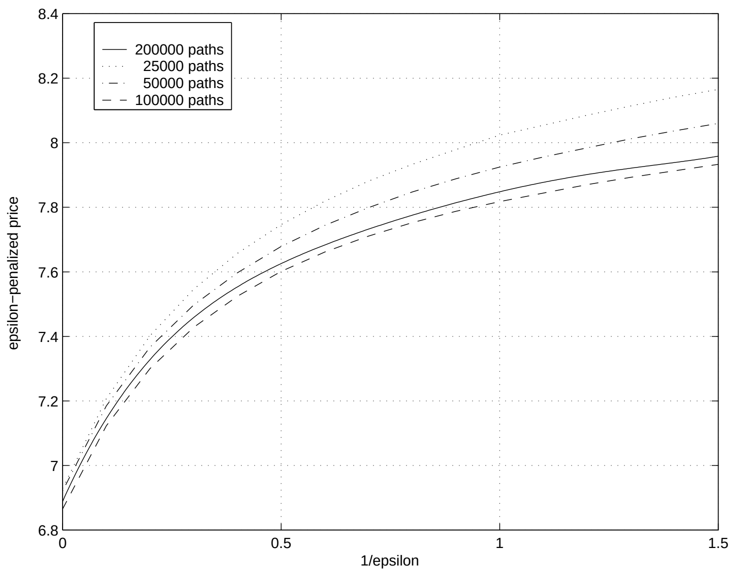

Figure 3: ϵ-approximation of the superhedging price as function of ϵ^{-1}.

References

Bally, V. (1997) Approximation Scheme for Solutions of BSDE. In: El Karoui, N., Mazliak, L. (eds.) *Backward Stochastic Differential Equations*, Addison Wesley Longman, 177-191.

Bender, C., Denk, R. (2005) Forward Simulation of Backward SDEs. Preprint.

Bender, C., Kohlmann, M. (2004) Optimal Superhedging under Nonconvex Constraints – a BSDE Approach. WIAS-Preprint 928.

Bergman, Y. Z. (1995) Option Pricing with Differential Interest Rates. *Rev. of Financial Studies*, **8**, 475-500.

Bouchard, B., Touzi, N. (2004) Discrete-Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations. *Stochastic Process. Appl.*, **111**, 175-206.

Broadie, M., Cvitanić, J., Mete Soner, H. (1998) Optimal Replication of Contingent Claims under Portfolio Constraints. *Rev. of Financial Studies*, **11**, 59-79.

Chevance, D. (1997) Numerical Methods for Backward Stochastic Differential Equations. In: Rogers, L. C. G., Talay, D. (eds.), *Numerical Methods in Finance*, Cambridge: University Press, 232-244.

Douglas, J., Ma, J., Protter, P. (1996) Numerical Methods for Forwad Backward Stochastic Differential Equations. *Ann. Appl. Probab.*, **6**, 940-968.

El Karoui, N., Peng, S., Quenez, M. C. (1997) Backward Stochastic Differential Equations in Finance. *Math. Finance*, **7**, 1-71.

Gobet, E., Lemor, J.-P., Warin, X. (2004) A Regression-Based Monte-Carlo Method to Solve Backward Stochastic Differential Equations. Preprint.

Longstaff, F. A., Schwartz, R. S. (2001) Valuing American Options by Simulation: A Simple Least-Square Approach. *Review of Financial Studies*, **14**, 113-147.

Ma, J., Protter, P. San Martín, J., Soledad, S. (2002) Numerical Method for Backward Stochastic Differential Equations. *Ann. Appl. Probab.*, **12**, 302-316.

Ma, J., Protter, P., Yong, J. (1994) Solving Forward-Backward Stochastic Differential Equations Explicitly – a Four Step Scheme. *Prob. Th. Rel. Fields*, **98**, 339-359.

Milstein, G. N., Tretyakov, M. V. (2004) Numerical Algorithms for Forward-Backward Stochastic Differential Equations Connected with Semilinear Parabolic Equations. Preprint.

Pardoux, E., Peng, S. (1990) Adapted Solutions of a Backward Stochastic Differential Equation. *Syst. Contr. Lett.*, **14**, 55-61.

Yong, J., Zhou, X. Y. (2000) *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Berlin: Springer.

Zhang, J. (2004) A Numerical Scheme for BSDEs. *Ann. Appl. Prob.*, **14**, 459-488.

Christian Bender
Weierstrass Institute
for Applied Analysis
and Stochastics
Mohrenstr. 39
D-10117 Berlin
Germany
bender@wias-berlin.de

Robert Denk
Department of Mathematics
and Statistics
University of Konstanz
D-78457 Konstanz
Germany
robert.denk@uni-konstanz.de