SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric
Analysis of Time Series

CREATES Research Paper 2010-32

Simple simulation of diffusion bridges with application to likelihood inference for diffusions

Mogens Bladt and Michael Sørensen

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Simple simulation of diffusion bridges with application to likelihood inference for diffusions

Mogens Bladt
Universidad Nacional Autónoma de México
IIMAS, A.P. 20-726
01000 Mexico, D.F.
Mexico
bladt@sigma.iimas.unam.mx

Michael Sørensen*
University of Copenhagen and CREATES
Dept. of Mathematical Sciences
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark
michael@math.ku.dk

August 5, 2010

Abstract

With a view to likelihood inference for discretely observed diffusion type models, we propose a simple method of simulating approximations to diffusion bridges. The method is applicable to all one-dimensional diffusion processes and has the advantage that simple simulation methods like the Euler scheme can be applied to bridge simulation. Another advantage over other bridge simulation methods is that the proposed method works well when the diffusion bridge is defined in a long interval because the computational complexity of the method is linear in the length of the interval. In a simulation study we investigate the accuracy and efficiency of the new method and compare it to exact simulation methods. In the study the method provides a very good approximation to the distribution of a diffusion bridge for bridges that are likely to occur in applications to likelihood inference. To illustrate the usefulness of the new method, we present an EM-algorithm for a discretely observed diffusion process. We demonstrate how this estimation method simplifies for exponential families of diffusions and very briefly consider Bayesian inference.

Key words: Bayesian inference; diffusion bridge; discretely sampled diffusions; EM-algorithm; Euler scheme; likelihood inference; time-reversion.

JEL codes: C22, C15.

*Michael Sørensen acknowledges support from Center for Research in Econometric Analysis of Time Series, CREATES, www.creates.au.dk, funded by the Danish National Research Foundation and from the Danish Center for Accounting and Finance funded by the Danish Social Science Research Council.

1 Introduction

In this paper we propose a general and very simple method for simulation of an approximation to diffusion bridges. Our main motivation for doing so is that simulation of diffusion bridges plays an important role in some very useful approaches to likelihood inference (including Bayesian inference) for discretely sampled diffusion processes and other diffusion-type processes like stochastic volatility models. Our approach is based on the simple idea of letting one diffusion process move forward from time zero out of one given point until it meets another diffusion process that independently moves backwards from time one out of another given point. Conditional on the event that the two diffusions intersect, we show that the process constructed in this way is an approximation to a realization of a diffusion bridge between the two points. The diffusions can be simulated by means of simple procedures like the Euler scheme or the Milstein scheme, see Kloeden & Platen (1999). The new method is therefore very easy to implement for likelihood inference for discretely sampled diffusion processes.

It was previously though impossible to simulate diffusion bridges by means of simple procedures like the Euler scheme, because a rejection sampler that tries to hit the prescribed end-point for the bridge (or a small neighbourhood around it) will have an excessively high rejection probability. The rejection sampler presented in this paper has a quite acceptable rejection probability because what must be hit is a sample path rather than a point. MCMC-algorithms for simulation of diffusion bridges were proposed by Roberts & Stramer (2001). Later Beskos, Papaspiliopoulos & Roberts (2006) developed algorithms for exact simulation of diffusion bridges when the drift and diffusion coefficients satisfy certain boundedness conditions. Under strong boundedness conditions the algorithm is relatively simple, whereas it is more involved under weaker condition. Recently Beskos, Papaspiliopoulos & Roberts (2007) has proposed an algorithm that works for a broad class of diffusion processes, but is more complex. A main advantage of the method proposed in this paper is that the same simple algorithm can be used for essentially all one-dimensional diffusions, and that it is very easy to understand and to implement. Another advantage is that a simulation study indicates that the computational complexity of our method is linear in the distance between the two end-points of the diffusion bridge, whereas the computer time seems to grow exponentially with this distance for the exact algorithms.

The result in the present paper, that the proposed method simulates an approximation to a diffusion bridge, is a correction of the previous paper Bladt & Sørensen (2007), where it was claimed that the method simulated an exact diffusion bridge (apart from the discretization error).

Methods for likelihood inference for discretely sampled diffusion processes have been proposed by Ozaki (1985), Pedersen (1995), Poulsen (1999), Durham & Gallant (2002), Aït-Sahalia (2002), Beskos et al. (2006), Aït-Sahalia (2008) and others, and Bayesian inference methods have been developed by Elerian, Chib & Shephard (2001), Eraker (2001), and Roberts & Stramer (2001). The methods by Roberts & Stramer (2001) and Beskos et al. (2006) use bridge simulation in an essential way. To illustrate how our method can be used, we modify an EM-algorithm in Beskos et al. (2006) by using our simple simulation method.

The paper is organized as follows. In Section 2 we present the new bridge simulation method and show in what sense it approximates a diffusion bridge. In Section 3 a stochastic differential equation for a diffusion bridge is presented. In Section 4 the new bridge simulation

method is compared to exact simulation methods in two examples, the Ornstein–Uhlenbeck process and the hyperbolic diffusion. The study indicates that our method provides a very accurate approximation to the distribution of a diffusion bridge except for bridges that are very unlikely to occur when using the method for maximum likelihood estimation. The EM-algorithm is briefly presented in Section 5. It is demonstrated how the algorithm simplifies for an exponential family of diffusions (i.e. when drift is linear in the parameters). In this case Bayesian inference is considered too. Section 6 concludes.

2 Diffusion bridge simulation

Let $X = \{X_t\}_{t\geq 0}$ be a one-dimensional diffusion given by the stochastic differential equation

$$dX_t = \alpha(X_t)dt + \sigma(X_t)dW_t, \tag{2.1}$$

where W is a Wiener process, and where the coefficients α and σ are sufficiently regular to ensure that the equation has a unique weak solution that is a strong Markov process. Let a and b be given points in the state–space of X. We present a method for simulating an approximation to a sample path of X such that $X_0 = a$ and $X_1 = b$. A solution of (2.1) in the interval $[t_1, t_2]$ such that $X_{t_1} = a$ and $X_{t_2} = b$ will in the following be called a (t_1, a, t_2, b)-bridge. When $t_1 = 0$ and $t_2 = 1$, we sometimes simply call it an (a, b)-bridge. We will denote the transition density of X by $p_t(x, y)$. Specifically, the conditional density of X_{s+t} given $X_s = x$ is $y \mapsto p_t(x, y)$. The state space of X is denoted by (ℓ, r) where $\infty \leq \ell < r \leq \infty$.

Let W^1 and W^2 be two independent standard Wiener processes, and define X^1 and X^2 as the solutions to

$$dX_t^i = \alpha(X_t^i)dt + \sigma(X_t^i)dW_t^i, \quad i = 1, 2, \quad X_0^1 = a \quad \text{and } X_0^2 = b.$$

The main idea of the paper is to realize an approximation to a $(0, a, 1, b)$-bridge by simulating the process X^1 from a forward in time and X^2 from b backward in time starting at time one. If the samples paths of the two processes intersect, they can be combined into a realization of a process that approximates a $(0, a, 1, b)$-bridge.

Thus to simulate an approximate diffusion bridge in the interval $[0, \Delta]$, we can use any of the several methods available to simulate the diffusions X^1 and X^2, see e.g. Kloeden & Platen (1999). Let $Y_{\delta i}^1$, $i = 0, 1, \ldots, N$ and $Y_{\delta i}^2$, $i = 0, 1, \ldots, N$ be (independent) simulations of X^1 and X^2 in $[0, \Delta]$ with step size $\delta = \Delta/N$. Then a simulation of an approximation to a $(0, a, \Delta, b)$-bridge is obtained by the following rejection sampling scheme. Keep simulating Y^1 and Y^2 until the sample paths cross, i.e. until there is an i such that either $Y_{\delta i}^1 \geq Y_{\delta(N-i)}^2$ and $Y_{\delta(i+1)}^1 \leq Y_{\delta(N-(i+1))}^2$ or $Y_{\delta i}^1 \leq Y_{\delta(N-i)}^2$ and $Y_{\delta(i+1)}^1 \geq Y_{\delta(N-(i+1))}^2$. Once a trajectory crossing has been obtained, define

$$B_{\delta i} = \begin{cases} Y_{\delta i}^1 & \text{for } \; i = 0, 1, \ldots, \nu - 1 \\ \\ Y_{\delta(N-i)}^2 & \text{for } \; i = \nu, \ldots N, \end{cases} \tag{2.2}$$

where $\nu = \min\{i \in \{1, \ldots N\} | Y_{\delta i}^1 \leq Y_{\delta(N-i)}^2\}$ if $Y_0^1 \geq Y_\Delta^2$, and $\nu = \min\{i \in \{1, \ldots N\} | Y_{\delta i}^1 \geq Y_{\delta(N-i)}^2\}$ if $Y_0^1 \leq Y_\Delta^2$. Then B approximates a $(0, a, \Delta, b)$-bridge. The rejection probability (the probability of no trajectory crossing) depends on the drift and diffusion coefficients, on

the values of a and b, and on the length of the interval Δ. Simulation studies indicate that the number of rejections is small when a and b are not very far apart, see Section 4. This is the typical situation when the simulation algorithm is used to make likelihood inference for discretely observed diffusion processes as explained in Section 5.

The distribution of the process that is simulated by the algorithm above and the sense in which it is an approximation of a diffusion bridge is seen from the following theorem.

Theorem 2.1 *Let* $\tau = \inf\{0 \leq t \leq 1 | X_t^1 = X_{1-t}^2\}$ $(\inf \emptyset = +\infty)$ *and define*

$$Z_t = \begin{cases} X_t^1 & \text{if } \ 0 \leq t \leq \tau \\ \\ X_{1-t}^2 & \text{if } \ \tau < t \leq 1. \end{cases}$$

Then the distribution of $\{Z_t\}_{0 \leq t \leq 1}$ *conditional on the event* $\{\tau \leq 1\}$ *equals the distributions of a* $(0, a, 1, b)$*-bridge conditional on the event that the bridge is hit by an independent diffusion with stochastic differential equation (2.1) and initial distribution with density* $p_1(b, \cdot)$*.*

The quality of the approximation obviously depends on the probability π that a $(0, a, 1, b)$-bridge is hit by an independent diffusion with initial distribution $p_1(b, \cdot)$. When π is close to one, the simulated process is essentially a $(0, a, 1, b)$-bridge. It is important to realize that the probability π is not equal to the acceptance probability $P(\tau \leq 1)$. It is quite possible that $P(\tau \leq 1)$ is small while π is close to one. This happens, for instance, for a diffusion with mean reversion to a level μ when $a \ll \mu \ll b$. In Section 4 we shall investigate when π can be expected to be close to one, and when a good approximation to a diffusion bridge is obtained. Simulations indicate that also when π is not close to one (but also not close to zero), the distribution of the simulated bridge is often indistinguishable from the distribution of an exact diffusion bridge.

Before proving Theorem 2.1, we prove a lemma on the distribution of a time-reversed diffusion. To do so we need the density of the speed measure

$$s(x) = \exp\left(-2 \int_{x^\#}^x \frac{\alpha(y)}{\sigma^2(y)} dy\right), \quad x \in (\ell, r), \tag{2.3}$$

with $x^\#$ denoting an arbitrary point in the state space of (ℓ, r).

Lemma 2.2 *The distribution of the time-reversed process* $\{\bar{X}_t\}$ *given by* $\bar{X}_t = X_{1-t}^2$ *has the same distribution as the conditional distribution of the process* $\{X_t\}$ *with* $X_0 \sim p_1(b, \cdot)$ *given that* $X_1 = b$*. The transition density of* $\{\bar{X}_t\}$ *as well as of* $\{X_t\}$ *conditional on* $X_1 = b$ *is*

$$q(x, s, y, t) = \frac{p_{t-s}(x, y) p_{1-t}(y, b)}{p_{1-s}(x, b)} = \frac{p_{t-s}(y, x) p_{1-t}(b, y)}{p_{1-s}(b, x)}, \quad s < t. \tag{2.4}$$

Proof: The joint density of $(\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_{t_n}) = (X_1, X_{1-t_1}, \ldots, X_{1-t_n})$, where $t_1 < \cdots < t_n < 1$, is

$$f(y_0, y_1, \ldots, y_n) = p_{1-t_n}(b, y_n) p_{t_n - t_{n-1}}(y_n, y_{n-1}) \cdots p_{t_2 - t_1}(y_2, y_1) p_{t_1}(y_1, y_0).$$

The conditional joint density of $(X_0, X_{t_1}, \ldots, X_{t_n})$, where $X_0 \sim p_1(b, \cdot)$, conditional on $X_1 = b$ is

$$g(y_0, y_1, \ldots, y_n) = p_1(b, y_0) p_{t_1}(y_0, y_1) p_{t_2 - t_1}(y_1, y_2) \cdots p_{t_n - t_{n-1}}(y_{n-1}, y_n) p_{1-t_n}(y_n, b) / p_1(y_0, b).$$

That $f(y_0, y_1, \ldots, y_n) = g(y_0, y_1, \ldots, y_n)$ follows from the fact that $p_t(x,y)s(x) = p_t(y,x)s(y)$ for any one-dimensional diffusion, where s is the density of the speed measure given by (2.3), see Ito & McKean (1965), p. 149. That the two expressions for the transition density (2.4) coincide follows from the same result. The first expression for q is the well-known expression for the transition density of a diffusion bridge ending in b at time 1, see Fitzsimmons, Pitman & Yor (1992), p. 111. It can be easily established by direct calculation. The second expression for q can similarly be obtained as the transition density of $\bar{X}$ by direct calculation:

$$p_{\bar{X}_s,\bar{X}_t}(x,y)/p_{\bar{X}_s}(x) = p_{X_{1-t},X_{1-s}}(y,x)/p_{X_{1-s}}(x) = p_{1-t}(b,y)p_{t-s}(y,x)/p_{1-s}(b,x).$$

□

Remark: Note that the results of Lemma 2.2 hold for a multivariate diffusion too, provided that there exists a function v such that $p_t(x,y)v(x) = p_t(y,x)v(y)$. Diffusions with this property are called v-symmetric, see the discussion in Kent (1978).

Proof of Theorem 2.1: Let W^3 be a standard Wiener processes independent of W^1, and let X^3 be the solution of

$$dX_t^3 = \alpha(X_t^3)dt + \sigma(X_t^3)dW_t^3,$$

where the distribution of X_0^3 has the density $p_1(b,\cdot)$. Finally, let τ be the first time the diffusion X^3 hits the sample path of X^1. Define a process by

$$Y_t = \begin{cases} X_t^1 & \text{if } \ 0 \le t \le \tau \\ \\ X_t^3 & \text{if } \ \tau < t \le 1 \end{cases}$$

on $\{\tau \le 1\}$, and $Y = X^1$ on $\{\tau = \infty\}$. By the strong Markov property Y has the same distribution as X^1. Now condition on $X_1^3 = b$. Since

$$P(Y \in \cdot \,|\, X_1^3 = b, \tau \le 1) = P(Y \in \cdot \,|\, Y_1 = b, \tau \le 1),$$

the theorem follows because by Lemma 2.2 the distribution of $\{X_{1-t}^2\}_{0\le t\le 1}$ equals that of $\{X_t^3\}_{0\le t\le 1}$ conditional on $X_1^3 = b$. The event $\{Y_1 = b, \tau \le 1\}$ is the event that Y is a diffusion bridge from a to b and that the diffusion bridge is hit by X^3.

□

By symmetry we see that the distribution of the process $\tilde{Z}$ defined by

$$\tilde{Z}_t = \begin{cases} X_t^1 & \text{if } \ 0 \le t \le 1-\tilde{\tau} \\ \\ X_{1-t}^2 & \text{if } \ 1-\tilde{\tau} < t \le 1, \end{cases}$$

where $\tilde{\tau} = \inf\{0 \le t \le 1 | X_{1-t}^1 = X_t^2\}$, is that of a $(0,a,1,b)$-bridge conditional on the event that the bridge is hit by $\{X_{1-t}^3\}$, where X^3 is an independent diffusion with stochastic differential equation (2.1) and initial distribution with density $p_1(a,\cdot)$. Here we use X^1 until the last time it crosses the trajectory of $\{X_{1-t}^2\}$, which happens at time $1-\tilde{\tau}$. Obviously, an approximate diffusion bridge can also be simulated by using $\tilde{Z}$.

3 Stochastic differential equations for diffusion bridges

In this section we give stochastic differential equations for diffusion bridges. These results cannot in general be used to simulate diffusion bridges, because the drift depends on the transition density of the original diffusion, which is usually not known explicitly. With a suitable approximation to the transition density, the stochastic differential equations can possibly be used to generate proposals for a Metropolis-Hastings sampler.

We consider the stochastic differential equation

$$dX_t = \alpha(X_t)dt + \sigma(X_t)dW_t, \tag{3.1}$$

where the solution X is either one-dimensional or a v-symmetric multivariate diffusion. The reader is reminded that a diffusion is called v-symmetric if there exists a function v such that $p_t(x,y)v(x) = p_t(y,x)v(y)$. For a discussion of v-symmetric multivariate diffusions, see Kent (1978). If X is p-dimensional, then α is a p-dimensional vector, σ is a $p \times p$-matrix, and $v(x) = \sigma(x)\sigma(x)^T$. In this section we assume that the drift and diffusion coefficient satisfy the usual global Lipschitz conditions. The domain of X is denoted by D.

Theorem 3.1 *Assume that*

$$\int_t^1 \int_D \left| \sum_{j=1}^{p} \nabla_{x_j} \left(v_{ij}(x) p_s(b,x) \right) \right| dxds < \infty, \quad i = 1, \ldots, p. \tag{3.2}$$

Then the $(0,a,1,b)$-bridge corresponding to (3.1) is a diffusion with diffusion coefficient $\sigma(x)$ and drift

$$\bar{b}_i(t,x) = -b_i(x) + \frac{\sum_{j=1}^{p} \nabla_{x_j} \left(v(x)_{ij} p_{1-t}(b,x) \right)}{p_{1-t}(b,x)}, \quad i = 1, \ldots, p.$$

For $p = 1$,

$$\bar{b}(t,x) = -b(x) + v'(x) + v(x)\partial_x \log p_{1-t}(b,x). \tag{3.3}$$

Proof: By Lemma 2.2, which also holds for v-symmetric multivariate diffusions, the transition density of the $(0,a,1,b)$-bridge is equal to the transition density of the time-reversed process $\bar{X}_t = X_{1-t}$, which by Theorem 2.3 in Millet, Nualert & Sanz (1989) is a diffusion with drift and diffusion coefficients as given in the theorem.

□

Conditions ensuring the local integrability condition (3.2) are discussed in Millet, Nualert & Sanz (1989), where also a similar result under the local Lipschitz condition are given. The condition (3.2) is implied if the two coefficients are twice continuously differentiable and if there exists $\epsilon > 0$ such that $v(x) \geq \epsilon I$, where I is the $p \times p$ identity matrix. Alternative conditions can be found in Haussmann & Pardoux (1986).

In cases, where the transition density is not differentiable with respect to x, the partial derivative in the formula for the drift are in the distributional sense.

Example 3.2 For a standard Brownian motion, (3.3) gives the well-known result that the drift of the Brownian bridge is $-(x-b)/(1-t)$.

□

Example 3.3 For an Ornstein–Uhlenbeck process with drift $-\theta x$ and diffusion coefficient σ, the drift of the $(0, a, 1, b)$-bridge is, by (3.3),

$$\bar{b}(t,x) = \theta\left(x - 2(x - be^{-\theta(1-t)})/(1 - e^{-2\theta(1-t)})\right),$$

which close to the boundary 1 behaves like

$$\bar{b}(t,x) = -(x-b)/(1-t) + \theta(x-b) + O(1-t).$$

□

4 Simulation study

In this section we simulate two examples of diffusion bridges for which exact algorithms are available, the Ornstein–Uhlenbeck process and the hyperbolic diffusion. We compare the distribution of exact bridge simulations to the distribution of the process obtained by the approximate bridge simulation proposed in the present paper. It is found that the approximation is very accurate in most cases, including bridges that are likely to occur in applications to likelihood inference. We also compare CPU execution times.

4.1 The Ornstein–Uhlenbeck bridge

First we consider the case of an Ornstein–Uhlenbeck bridge, which is a solution to the stochastic differential equations

$$dX_t = -\theta X_t dt + \sigma dW_t$$

conditionally on $X_0 = a$ and $X_1 = b$ for some $a, b \in \mathbb{R}$. From the well-known Gaussian transition densities of the Ornstein–Uhlenbeck process we can calculate the transition densities of the Ornstein–Uhlenbeck bridge by (2.4). Thus we could in principle simulate the Ornstein–Uhlenbeck bridge by sampling transitions from these densities. The following alternative method is, however, numerically more stable.

Lemma 4.1 *Generate $X_{t_0}, X_{t_1}, \ldots X_{t_n}, X_{t_{n+1}}$, where $0 = t_0 < t_1 < \cdots < t_n < t_{n+1}$, by $X_0 = x_0$ and*

$$X_{t_i} = e^{-\theta(t_i - t_{i-1})} X_{t_{i-1}} + W_i, \quad i = 1, \ldots, n+1$$

where the W_is are independent and

$$W_i \sim N\left(0, \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta(t_i - t_{i-1})}\right)\right).$$

Define

$$Z_{t_i} = X_{t_i} + (x - X_{t_{n+1}})\frac{e^{\theta t_i} - e^{-\theta t_i}}{e^{\theta t_{n+1}} - e^{-\theta t_{n+1}}}, \quad i = 0, \ldots, n+1.$$

Then $(Z_{t_0}, Z_{t_1}, \ldots, Z_{t_n}, Z_{t_{n+1}})$ is distributed like an Ornstein–Uhlenbeck bridge with $Z_{t_0} = x_0$ and $Z_{t_{n+1}=x}$.

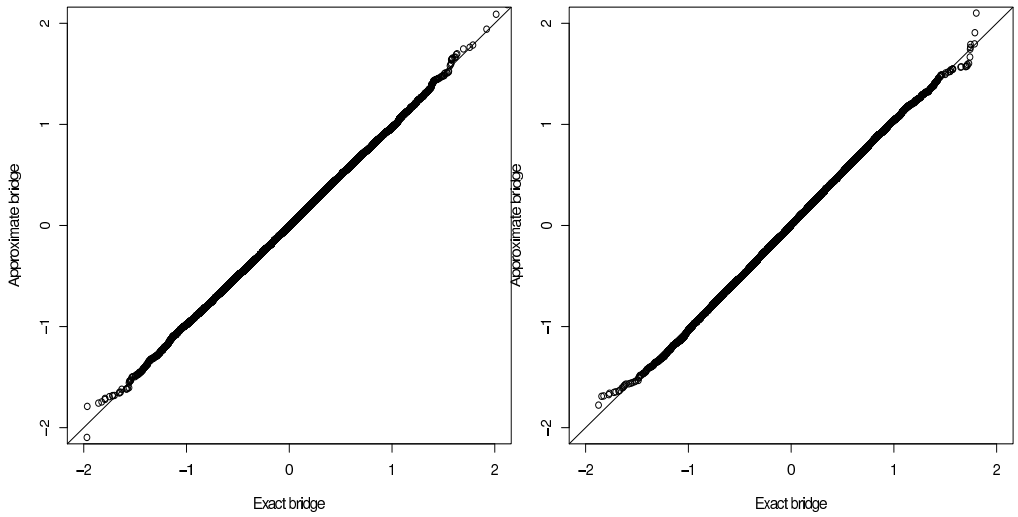

Figure 4.1: Q-Q plots that compare the empirical distribution at time 0.5 based on 25,000 simulated $(0, 0)$ diffusion bridges obtained by our method to that based on 25,000 exactly simulated diffusion bridges. The left plot is for the Ornstein–Uhlenbeck bridge and the right plot is for the hyperbolic diffusion bridge. Exact simulations are obtained by the method in Lemma 4.1 for the Ornstein–Uhlenbeck bridge and by the exact algorithm of Beskos, Papaspiliopoulos & Roberts (2006) for the hyperbolic diffusion bridge.

Proof: The result follows straightforwardly because $Z = (Z_{t_1}, \ldots, Z_{t_n})^T$ is a linear transformation of a normal distribution. For completeness, we give the details. Let Σ denote the covariance matrix of $X = (X_{t_1}, \ldots, X_{t_n})^T$, and define $c^T = (c_1, \ldots, c_n)$, where $c_i = \mathrm{Cov}(X_{t_i}, X_{t_{n+1}}) = \sigma^2 e^{-\theta t_{n+1}}(e^{\theta t_i} - e^{-\theta t_i})/(2\theta)$, $i = 1, \ldots, n$. Since $\mathrm{Var}(X_{t_{n+1}}) = \sigma^2(1 - e^{-2\theta t_{n+1}})/2\theta$, it follows that

$$Y = X - \frac{2\theta X_{t_{n+1}}}{\sigma^2(1 - e^{-2\theta t_{n+1}})} c \sim N_n \left(\xi - \frac{2\theta x_0 e^{-\theta t_{n+1}}}{\sigma^2(1 - e^{-2\theta t_{n+1}})} c,\ \Sigma - \frac{2\theta}{\sigma^2(1 - e^{-2\theta t_{n+1}})} c c^T \right),$$

where $\xi^T = (x_0 e^{-\theta t_1}, \ldots, x_0 e^{-\theta t_n})$. Thus

$$Z = Y + \frac{2\theta x}{\sigma^2(1 - e^{-2\theta t_{n+1}})} \sim N_n \left(\xi + \frac{2\theta(x - x_0 e^{-\theta t_{n+1}})}{\sigma^2(1 - e^{-2\theta t_{n+1}})} c,\ \Sigma - \frac{2\theta}{\sigma^2(1 - e^{-2\theta t_{n+1}})} c c^T \right),$$

which is the conditional distribution of X given $X_{t_{n+1}} = x$.

□

In all examples considered in the following we simulated 25,000 realizations of diffusion bridges over the time interval $[0, 1]$. The Euler scheme was used with discretization level $N = 100$ ($\delta = 0.01$). The methods were implemented in Fortran 90 on a Dell Precision M65 workstation (laptop).

For the Ornstein–Uhlenbeck bridge we chose the parameter values $\theta = 0.5$ and $\sigma = 1.0$. This particular choice of the parameter values is not essential to the conclusions of the simulations study. First we considered a bridge that started at 0 and ended at 0. We compare our method presented in Section 2 and based on Theorem 2.1 to the exact algorithm of Lemma 4.1. To the left in Figure 4.1 we have plotted the quantiles of the empirical

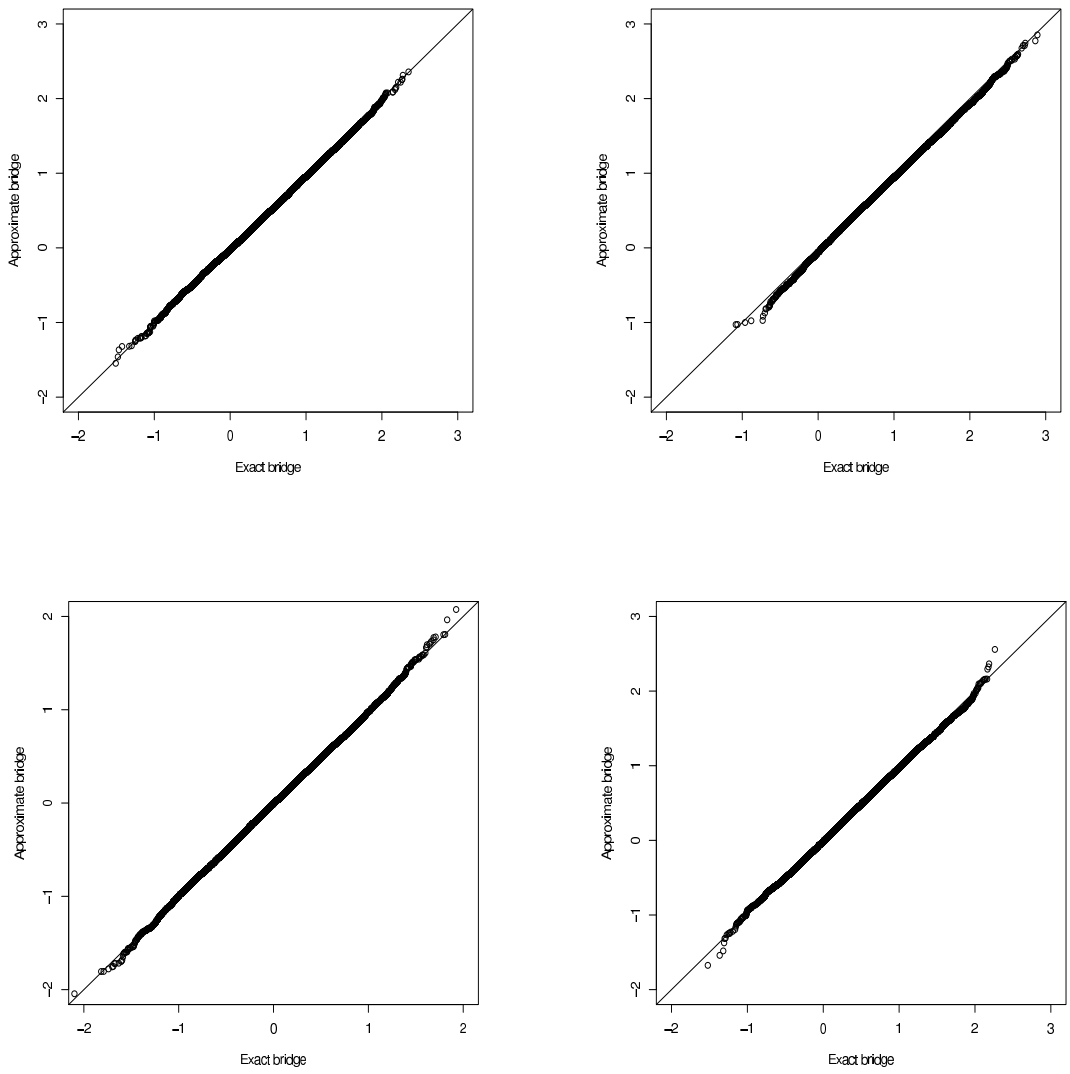

Figure 4.2: Q-Q plots that compare the empirical distributions at time 0.5 based on 25,000 simulated $(0, 1)$, $(0, 2)$, $(-1, 1)$ and $(-1, 2)$ Ornstein–Uhlenbeck bridges obtained by our method to that based on 25,000 exactly simulated Ornstein–Uhlenbeck bridges. Exact simulations are obtained by the method in Lemma 4.1.

distribution at the time point 0.5 obtained by our method against the quantiles of the empirical distribution obtained by the exact algorithm. The two distributions appear to be equal. Similar comparisons of quantiles at time 0.5 for our method to quantiles of an exact bridge are presented in Figure 4.2 for $(0,1)$, $(0,2)$, $(-1,1)$ and $(-1,2)$ Ornstein–Uhlenbeck bridges. In all four cases the two distributions seem to be essentially equal, except for a very small negative bias for the $(0,2)$-bridge. Similar results were found for several other comparisons of distributions with similar values of the start and end points, a and b.

The CPU execution time (in seconds) to simulate 10,000 Ornstein–Uhlenbeck bridges using our method for the various starting points, a, and end points, b, are given in Table 4.1 together with estimated rejection probabilities. The table also gives the probabilities that an Ornstein–Uhlenbeck process moves from a to b or farther in the time interval $[0,1]$. We see that for moves that are likely to appear in data sets, the CPU times and rejection probabilities are small, and the CPU times are only slightly larger than the execution time for the exact algorithm which is about 0.5 CPU seconds. For more unlikely moves the rejection probability is quite large, but also in these cases the execution time is not a problem in applications. The last column of Table 4.1 gives the (estimated) probability of the event that an exact (a,b)-bridge is not hit by an independent diffusion with initial distribution $p_1(b,\cdot)$. These probabilities were found by simulating exact Ornstein–Uhlenbeck bridges and independent Ornstein–Uhlenbeck processes with initial distribution $p_1(b,\cdot)$. If this probability were zero, our method would simulate an exact diffusion bridge. The probabilities are small, but not negligible. It is remarkable that our method gives an accurate approximation to a diffusion bridge in spite of this. The reason must be that the diffusion bridges are not hit by the independent diffusion in a systematic way for the a and b values considered here.

$a \mapsto b$	CPU (sec.)	rejection prob.	probability of move	$1-\pi$
$0 \mapsto 0$	0.5	0.17		0.28
$0 \mapsto 1$	0.7	0.41	0.1	0.21
$0 \mapsto 2$	1.7	0.77	0.006	0.08
$-1 \mapsto 1$	1.9	0.80	0.02	0.16
$-1 \mapsto 2$	11.9	0.97	0.0005	0.06

Table 4.1: The CPU execution time (in seconds) used to simulate 10,000 Ornstein–Uhlenbeck bridges using our method for various starting points, a, and end points, b. Also estimated rejection probabilities and the probabilities of a move from a to b or farther. The last column gives the probability that an exact (a,b)-bridge is not hit by an independent diffusion with initial distribution $p_1(b,\cdot)$.

In order to test out method in an extreme situation we simulated 25,000 Ornstein–Uhlenbeck bridges that started from -2 and ended in 2. The probability that an Ornstein–Uhlenbeck process with parameters $\theta = 0.5$ and $\sigma = 1.0$ moves from -2 to 2 or farther in the time interval $[0,1]$ equals the probability that a standard normal distribution is larger than 4.04, which equals 0.00003, so this a indeed a very extreme test of the method. Not surprisingly that rejection rate was very high, but as appears from Figure 4.3 the distribution at time 0.5 fits the distribution obtained by exact simulation very well.

The only situation we have been able to find where the distribution obtained by our simulation method differs appreciably from the distribution of an exact bridge is when the

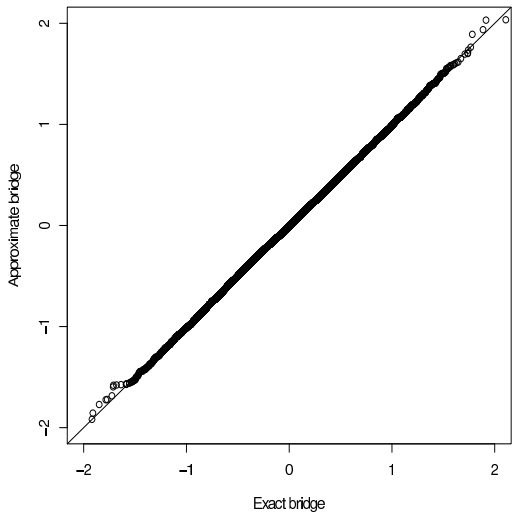

Figure 4.3: Q-Q plot that compares the empirical distribution at time 0.5 based on 25,000 simulated $(-2, 2)$ Ornstein–Uhlenbeck bridges obtained by our method to that based on 25,000 exactly simulated Ornstein–Uhlenbeck bridges.

start and end points, a and b, have the same sign and are both far from the equilibrium point zero. This is to be expected because we simulate an exact bridge conditional on the event that it is hit by an independent diffusion with initial distribution $p_1(b, \cdot)$. When b is far from zero, most of the probability mass of $p_1(b, \cdot)$ is located considerably closer to zero than b (because of the drift towards zero). Moreover the independent diffusion will tend to move towards zero. On the other hand, the (a, b)-bridge will tend to stay relatively close to a and b. Only trajectories of the bridge that move sufficiently towards zero has a reasonable chance of being hit by the independent diffusion. This creates a bias towards zero. The comparison of quantiles at time 0.5 for $(-1, -1)$, $(-2, -2)$, $(-3, -3)$ and $(2, 3)$ bridges are presented in Figure 4.4. For the $(-1, -1)$-bridge the bias is very small, and for the $(-2, -2)$-bridge there is a relatively small positive bias (as expected from the consideration above). For the $(-3, -3)$-bridge and the $(2, 3)$ bridge there is a more considerable bias (positive and negative, respectively, as expected). Note that the points are, to a good approximation, parallel to the identity line, which indicates that essentially the only problem is that the position of the distribution has been shifted a little.

Figure 4.5 illustrates Theorem 2.1. The plot to the left compares approximate simulation of the $(-3, -2)$-bridge to the exact $(-3, -2)$-bridge in a way similar to the previous plots, whereas the plot to the right is a Q-Q plot that compares the empirical distributions at time 0.5 based on 25,000 simulated $(-3, -2)$ Ornstein–Uhlenbeck bridges obtained by our method to that based on 25,000 exactly simulated Ornstein–Uhlenbeck bridges, where the exactly simulated bridges were removed from the sample if the bridge was not hit by an independent diffusion with initial distribution $p_1(b, \cdot)$. As expected from Theorem 2.1 the two distributions in the plot to the right appear to be equal.

Table 4.2 gives estimated rejection probabilities and the probability that an exact (a, b)-bridge is not hit by an independent diffusion with initial distribution $p_1(b, \cdot)$. As expected the probabilities of not being hit is quite substantial in most cases. The last column gives the probability that a stationary Ornstein–Uhlenbeck process is a distance $|a|$ or more from

zero. We see that the process will only spend very little time in the parts of the state space, where our method is biased. This information can be supplemented by the fact that the probability that an Ornstein–Uhlenbeck process goes from 2 to 3 or farther in the time interval $[0, 1]$ is 0.01, so it is highly unlikely to observe an Ornstein–Uhlenbeck process at 2 at time zero and at 3 at time one. Therefore such a bridge will extremely rarely be needed in simulation–based likelihood inference for an Ornstein–Uhlenbeck process.

$a \mapsto b$	rejection prob.	$1 - \pi$	$2P(X_t > \lvert a \rvert)$
$-1 \mapsto -1$	0.15	0.37	0.32
$-2 \mapsto -2$	0.09	0.55	0.05
$-3 \mapsto -3$	0.05	0.74	0.003
$2 \mapsto 3$	0.24	0.41	0.05
$-3 \mapsto 2$	0.24	0.74	0.003

Table 4.2: Estimated rejection probabilities for the Ornstein–Uhlenbeck bridges using our method for various starting points, a, and end points, b. The second column gives the probability that an exact (a, b)-bridge is not hit by an independent diffusion with initial distribution $p_1(b, \cdot)$. The last column gives the probability of finding an Ornstein–Uhlenbeck a distance $|a|$ or more from zero.

The time point 0.5 was chosen in the simulation study because it is expected that this is the time point where it is most difficult to to get a good approximation of the distribution of a diffusion bridge. In order to check that our method works better close to the end-points of the time interval, we have compared the distribution at time 0.1 obtained by our method to that of an exact Ornstein–Uhlenbeck bridge for a $(-1, -1)$-bridge and a $(-3, -3)$-bridge. The results are given in Figure 4.6, which shows that our method works better at time 0.1 than at time 0.5. For the $(-1, -1)$-bridge the bias has essentially disappeared at time 0.1, and for the $(-3, -3)$-bridge the bias is considerably smaller at time 0.1 than at time 0.5. Several similar comparisons confirm that our method works better close to the end-points than at time 0.5.

4.2 The hyperbolic bridge

Next we consider the hyperbolic diffusion which is the solution to

$$dX_t = -\frac{\theta X_t}{\sqrt{1 + X_t^2}} dt + \sigma dW_t,$$

with $\theta > 0$ and $\sigma > 0$. The hyperbolic diffusion was introduced by Barndorff-Nielsen (1978). It is ergodic with the standardized symmetric hyperbolic distribution as invariant distribution, see e.g. Bibby & Sørensen (2003). In this case the transition density is not explicitly known, but we can compare our method to the exact algorithm by Beskos, Papaspiliopoulos & Roberts (2006). This method is referred to as EA1 in Beskos et al. (2006). It is is applicable to diffusion processes on the form

$$dX_t = \alpha(X_t)dt + dW_t, \tag{4.1}$$

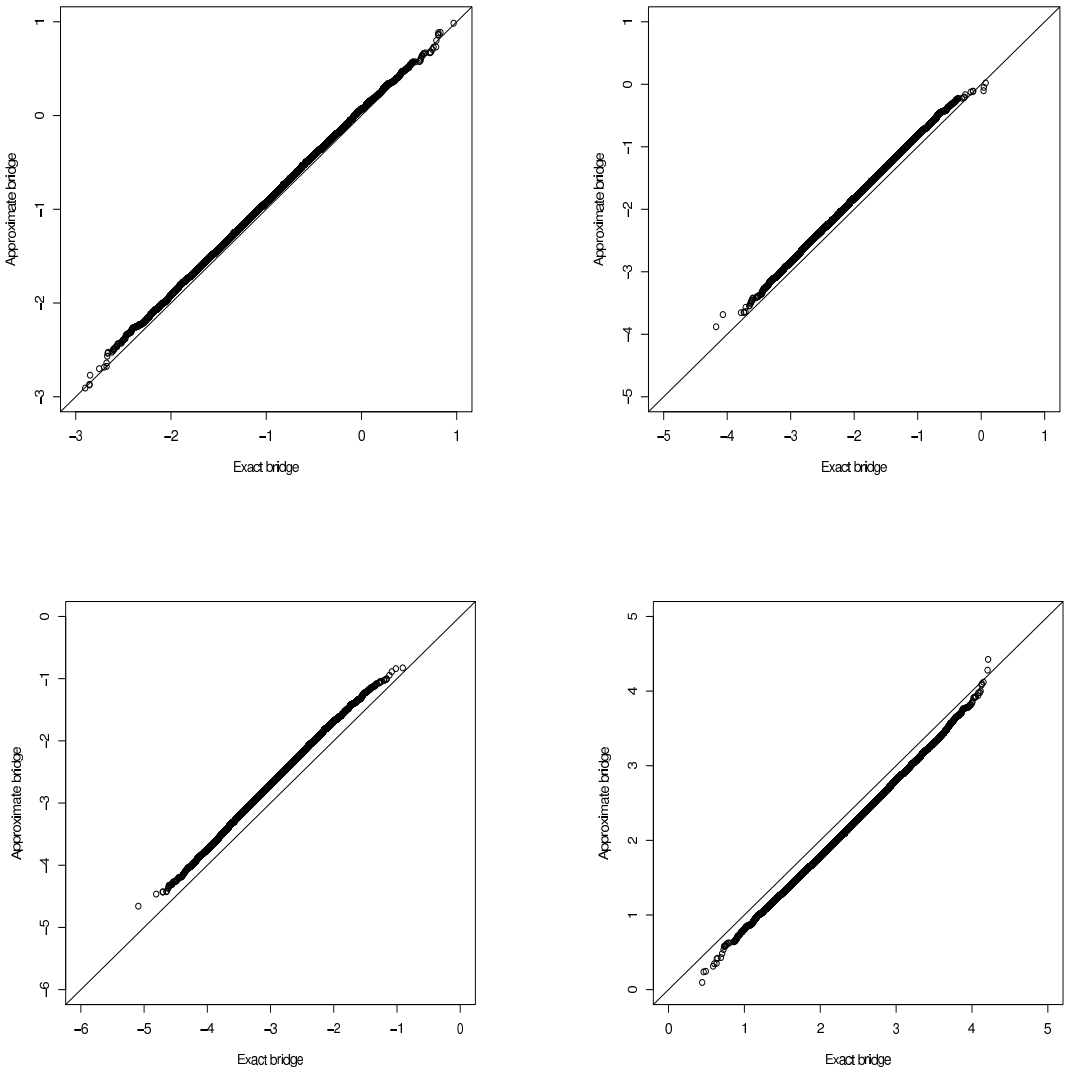

Figure 4.4: Q-Q plots that compare the empirical distributions at time 0.5 based on 25,000 simulated $(-1,-1)$, $(-2,-2)$, $(-3,-3)$ and $(2,3)$ Ornstein–Uhlenbeck bridges obtained by our method to that based on 25,000 exactly simulated Ornstein–Uhlenbeck bridges. Exact simulations are obtained by the method in Lemma 4.1.

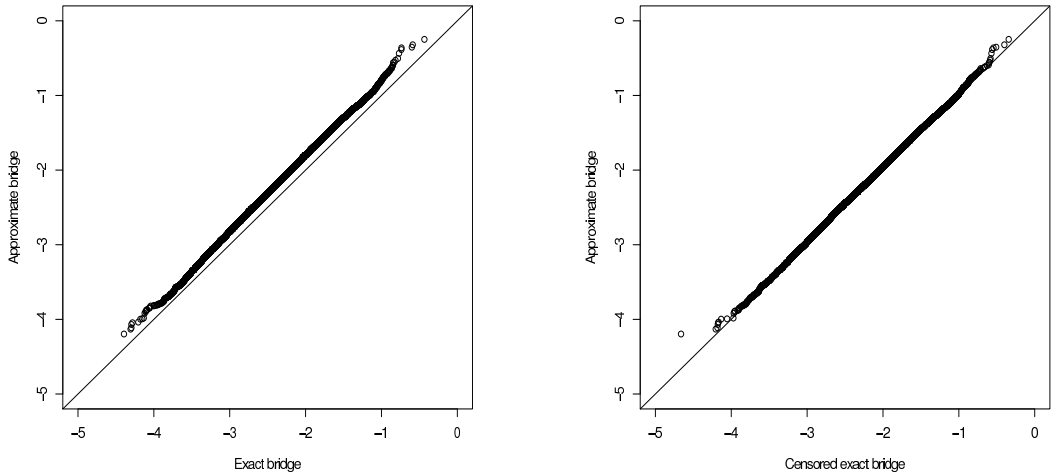

Figure 4.5: Q-Q plots that compare the empirical distributions at time 0.5 based on 25,000 simulated $(-3,-2)$ Ornstein–Uhlenbeck bridges obtained by our method to that based on 25,000 exactly simulated Ornstein–Uhlenbeck bridges (left plot), and to that where the exactly simulated Ornstein–Uhlenbeck bridges were removed from the sample if the bridge was not hit by an independent diffusion with initial distribution $p_1(b,\cdot)$ (right plot). Exact simulations are obtained by the method in Lemma 4.1.

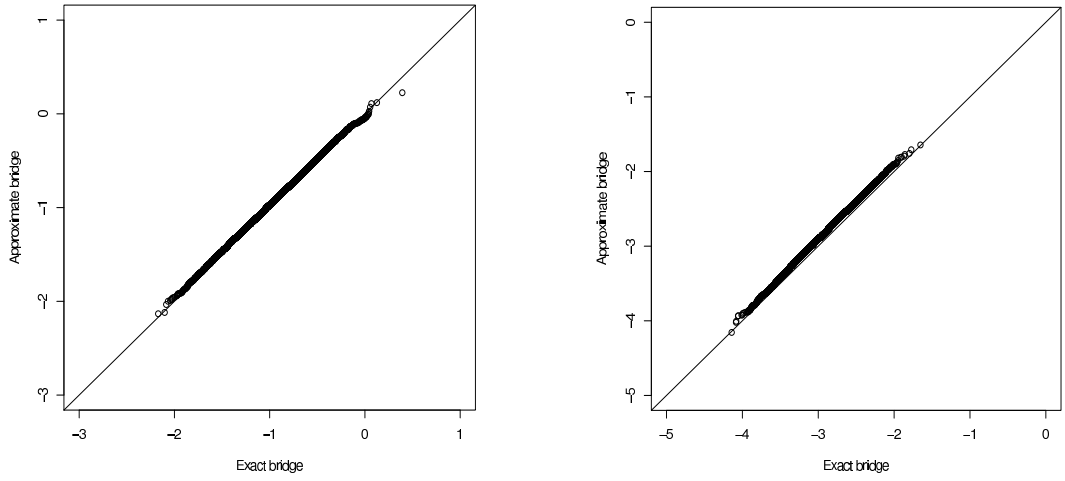

Figure 4.6: Q-Q plots that compare the empirical distributions at time 0.1 based on 25,000 simulated $(-1,-1)$ and $(-3,-3)$ Ornstein–Uhlenbeck bridges obtained by our method to that based on 25,000 exactly simulated Ornstein–Uhlenbeck bridges. Exact simulations are obtained by the method in Lemma 4.1.

provided that α is continuously differentiable, and the function $\alpha(x)^2 + \alpha'(x)$ is bounded from above and below for all x. The hyperbolic diffusion process obviously satisfies this boundedness condition. The algorithm by Beskos, Papaspiliopoulos & Roberts (2006) is very quick for short intervals as it essentially only requires one simulation of a Brownian bridge if not rejected. Rejection in the EA1 algorithm is not very costly computationally in our example since it is only a few points that are thrown away per rejection. Thus we compare our algorithm to a very efficient method.

Again we simulated 25,000 bridges using the Euler scheme with a 100 points subdivision of $[0, 1]$. The parameter values were $\theta = \sigma = 1$. We start with a bridge from 0 to 0 and compare our method to the exact algorithm of Beskos, Papaspiliopoulos & Roberts (2006). To the right in Figure 4.1 we have plotted the quantiles of the empirical distribution at the time point 0.5 obtained by our method against the quantiles of the empirical distribution obtained by the exact algorithm. Also for this example the two distributions appear to be equal. Table 4.3 shows CPU execution times to simulate 10,000 hyperbolic diffusion bridges by our method for various starting points, a, and end points, b. Also estimated rejection probabilities are given. The pattern is similar to that for the Ornstein–Uhlenbeck process. For moves that are likely to appear in data sets, the CPU times and rejection probabilities are small, and for unlikely moves the execution time is not a problem in applications, even though the rejection probability is quite large. The execution time for the EA1 algorithm was 0.3 CPU seconds, which, as expected, is faster than our method. Note that there is no reason to consider diffusions for which the EA1 algorithm does not work in order to compare our method to the more complicated simulation methods EA2 and EA3 in Beskos, Papaspiliopoulos & Roberts (2006) and Beskos, Papaspiliopoulos & Roberts (2007). The EA2 and EA3 algorithms are clearly more time consuming than EA1, while execution times for our methods can be expected to be approximately as for the two examples considered here.

$a \mapsto b$	CPU (sec.)	rejection prob.
$0 \mapsto 0$	0.6	0.14
$0 \mapsto 1$	0.8	0.36
$0 \mapsto 2$	2.1	0.77
$-1 \mapsto 1$	2.0	0.76
$-1 \mapsto 2$	12.6	0.96

Table 4.3: The CPU execution time (in seconds) used to simulate 10,000 hyperbolic diffusion bridges with $\theta = \sigma = 1$ by our method for various starting points, a, and end points, b. Also estimated rejection probabilities are given.

Beskos, Papaspiliopoulos & Roberts (2006) noted that the computing time of their exact algorithm is large for diffusion bridges over long time intervals. It is therefore of interest to compare computer time and rejection probabilities for our algorithm to the algorithm in Beskos, Papaspiliopoulos & Roberts (2006). To do so, we simulated 10,000 trajectories of the $(0, 0, \Delta, 0)$-bridge for the hyperbolic diffusion with $\theta = \sigma^2 = 4$. This was done for values of the interval length Δ ranging from 0.5 to 5. The CPU execution time (in seconds) used to simulate the 10,000 trajectories are given in Table 4.4. We see that for this particular diffusion the two methods use the same CPU time for an interval length of two. For smaller

interval lengths the exact algorithm is somewhat faster, whereas the method proposed in Section 2 is much faster for long intervals. The simulations indicate that the computational complexity of the proposed method is linear in the interval length Δ, whereas it appears to grow at least exponentially with Δ for the exact algorithm; see Figure 4.7. The main reason is that for long intervals the number of rejections becomes very large for the algorithm in Beskos, Papaspiliopoulos & Roberts (2006), while the algorithm proposed in Section 2 has a small rejection probability for long intervals. It is not surprising that the trajectories of the two ergodic diffusions will intersect with a large probability for a long interval, which implies a small rejection probability.

	Present paper		Beskos et al. (2006)	
Δ	CPU time	# rejections	CPU time	# rejections
0.5	0.52	819	0.28	14497
1.0	0.99	307	0.59	53087
1.5	1.45	102	1.05	163599
2.0	1.93	44	1.92	457226
2.5	2.40	17	4.00	1242922
3.0	2.88	6	10.01	3491838
3.5	3.36	2	26.86	9357310
4.0	3.83	0	75.79	25232418
4.5	4.31	0	222.09	69299642
5.0	4.79	0	641.70	187069771

Table 4.4: The CPU execution time (in seconds) used to simulate 10,000 hyperbolic $(0, 0, \Delta, 0)$-bridges with $\theta = \sigma^2 = 4$ for the method proposed in Section 2 and for the method in Beskos, Papaspiliopoulos & Roberts (2006) for different interval lengths Δ. Also the number of rejections while simulating the 10,000 trajectories is given.

5 Maximum likelihood estimation

To give an example of an application of our diffusion bridge simulation method, we present a way of finding the maximum likelihood estimator for discretely observed diffusion processes that is a modification of a method by Beskos et al. (2006). The advantage of our method is that it works for all one-dimensional diffusion processes. We also briefly discuss aspects of Bayesian inference.

Consider the diffusion process

$$dX_t = b_\alpha(X_t)dt + \sigma_\beta(X_t)dW_t, \tag{5.1}$$

where α and β are unknown parameters to be estimated, and W is the standard Wiener process. We assume that $\sigma_\beta(x) > 0$ for all x in the state interval. Suppose that the only data available from a realization of the diffusion process are observations at times $t_1 < t_2 < ... < t_n$, $x_i = X_{t_i}$, $i = 1, ..., n$.

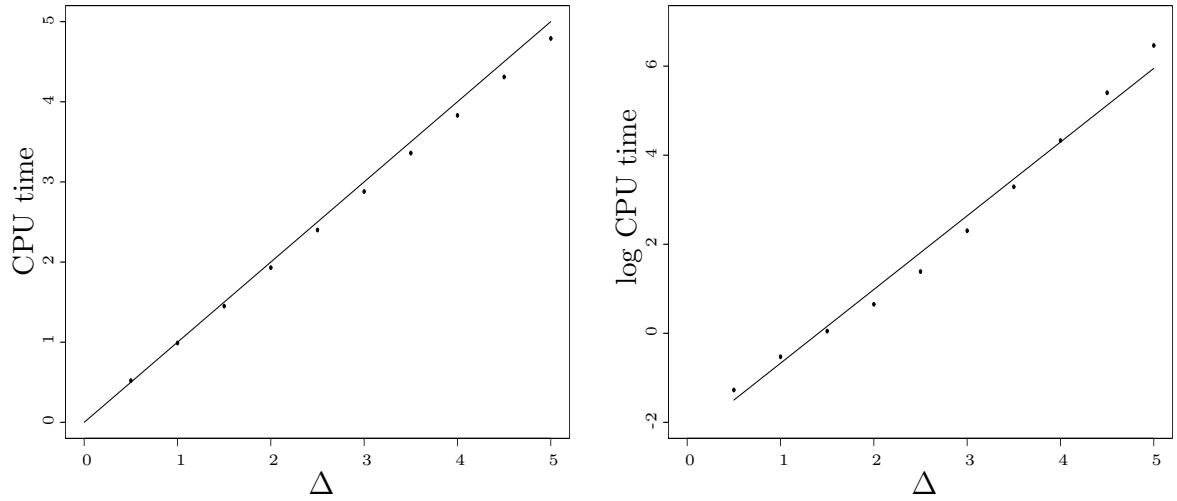

Figure 4.7: The CPU execution time (in seconds) used to simulate 10,000 hyperbolic $(0, 0, \Delta, 0)$-bridges with $\theta = \sigma^2 = 4$. In the left plot the CPU time is plotted against Δ for the method proposed in Section 2, while in the right plot the logarithm of the CPU time is plotted against Δ for the method in Beskos, Papaspiliopoulos & Roberts (2006).

Discrete time observation of a continuous time process can be viewed as an incomplete observation problem so that the EM-algorithm (Dempster, Laird & Rubin (1977)) is a potential method for finding the maximum likelihood estimator of the parameters. For instance maximum likelihood estimation for discretely observed Markov jump processes was treated in this way by Bladt & Sørensen (2005) and Bladt & Sørensen (2009). Unfortunately, the probability measures corresponding to complete continuous time observation of the diffusion model given by (5.1) are singular because the diffusion coefficient depends on the parameter β. It is therefore not straightforward to implement the EM-algorithm, but an approach in the spirit of Roberts & Stramer (2001) was proposed by Beskos et al. (2006). In the following we summarize a modification of this approach using our diffusion bridge simulation technique. Moreover, we show how a considerable simplification of the algorithm can be achieved when the model (5.1) is an exponential family of diffusions in the sense of Küchler & Sørensen (1997).

The transformation

$$h_\beta(x) = \int_{x^*}^{x} \frac{1}{\sigma_\beta(y)} dy$$

is essential. Here x^* is some arbitrary, but appropriately chosen, point in the state interval. The point is that by Ito's formula, $Y_t = h_\beta(X_t)$ solves

$$dY_t = \mu_{\alpha,\beta}(Y_t)dt + dW_t, \tag{5.2}$$

where

$$\mu_{\alpha,\beta}(y) = \frac{b_\alpha(h_\beta^{-1}(y))}{\sigma_\beta(h_\beta^{-1}(y))} - \frac{1}{2}\sigma_\beta'\left(h_\beta^{-1}(y)\right).$$

In (5.2) the diffusion coefficient does not depend on the parameters, so the probability measures are equivalent and the likelihood function can be found. To do so the function

$$g_{\alpha,\beta}(x) = s_{\alpha,\beta}(x) - \tfrac{1}{2}\log(\sigma_\beta(x)), \tag{5.3}$$

where

$$s_{\alpha,\beta}(x) = \int_{x^*}^{x} \frac{b_\alpha(z)}{\sigma_\beta^2(z)} dz, \tag{5.4}$$

is needed. Note that $\int_{y^*}^{y} \mu_{\alpha,\beta}(z)dz = g_{\alpha,\beta}(h_\beta^{-1}(y)) - g_{\alpha,\beta}(h_\beta^{-1}(y^*))$, and that the functions $g_{\alpha,\beta}$ and $s_{\alpha,\beta}$ are closely related to the density $\varphi_{\alpha,\beta}$ of the stationary distribution of the original diffusion model given by (5.1). Specifically, $s_{\alpha,\beta}(x)$ equals $\frac{1}{2}\log(\sigma_\beta(x)^2\varphi_{\alpha,\beta}(x))$ apart from an additive constant. Thus when the stationary density is known, the only problem is to find h_β and its inverse. This is for instance the case for the Pearson diffusions studied by Forman & Sørensen (2008).

The problem with the transformation h_β is that it is parameter dependent, while we need to keep the original discrete time data fixed when running the EM-algorithm. To get around this problem, define

$$Y_t^*(\beta,\beta_0) = Z_t^{(i,\alpha_0,\beta_0)} + \frac{(t_i - t)(h_\beta(x_{i-1}) - h_{\beta_0}(x_{i-1})) + (t - t_{i-1})(h_\beta(x_i) - h_{\beta_0}(x_i))}{t_i - t_{i-1}},$$

for $t_{i-1} \le t \le t_i$, $i = 2, \ldots, n$. Here $Z_t^{(i,\alpha_0,\beta_0)}$ denotes the $(t_{i-1}, h_{\beta_0}(x_{i-1}), t_i, h_{\beta_0}(x_i))$-bridge for the diffusion (5.2) with parameter values α_0 and β_0, and $Z_t^{(i,\alpha_0,\beta_0)}$, $i = 2, \ldots, n$ are independent. Then the EM-algorithm works as follows. Let α_0, β_0 be initial values of the parameters.

(1) (E–step) Calculate the function

$$\begin{aligned} q(\alpha,\beta) &= g_{\alpha,\beta}(x_n) - g_{\alpha,\beta}(x_1) - \frac{1}{2}\sum_{i=2}^{n}[h_\beta(x_i) - h_\beta(x_{i-1})]^2/(t_i - t_{i-1}) \\ &\quad - \sum_{i=2}^{n}\log(\sigma_\beta(x_i)) - \frac{1}{2}\sum_{i=2}^{n}\mathbb{E}_{Z^{(i,\alpha_0,\beta_0)}}\left(\int_{t_{i-1}}^{t_i}\left[\mu'_{\alpha,\beta}(Y_t^*(\beta,\beta_0)) + \mu_{\alpha,\beta}(Y_t^*(\beta,\beta_0))^2\right]dt\right). \end{aligned}$$

(2) (M–step) $(\alpha_0, \beta_0) = \text{argmax}_{\alpha,\beta} q(\alpha,\beta)$.

(3) GO TO (1).

In the E–step, $\mathbb{E}_{Z^{(i,\alpha_0,\beta_0)}}$ means that the data points are fixed so that only the diffusion bridge is random, and expectation is with respect to the distribution of the diffusion bridge. Thus the expectations in the E–step can be approximated by simulating independent diffusion bridges by our method and averaging. Arguments that $q(\alpha,\beta)$ is the conditional expectation of the relevant continuous time likelihood function can be found in Roberts & Stramer (2001) and Beskos et al. (2006). As pointed out in the latter paper, the conditional expectation can also be calculated as

$$\mathbb{E}_{Z^{(i,\alpha_0,\beta_0)},U}\left(\mu'_{\alpha,\beta}(Y_U^*(\beta,\beta_0)) + \mu_{\alpha,\beta}(Y_U^*(\beta,\beta_0))^2\right),$$

where U is a uniformly distributed random variable on $[t_{i-1}, t_i]$ that is independent of $Z_t^{(i,\alpha_0,\beta_0)}$ and the data.

In the M–step the maximization of $q(\alpha, \beta)$ must in general be done by a suitable maximization algorithm. However when the drift of the original diffusion model (5.1) depends linearly on the vector of parameters α, i.e. when

$$b_\alpha(x) = \alpha_1 a_1(x) + \cdots + \alpha_k a_k(x), \tag{5.5}$$

where $a_1, \ldots, a_k$ are known functions, then the maximization problem is simplified somewhat. When the drift has this form, and when the diffusion parameter β is fixed, the model for continuous time observation of X as well as the transformed process Y is an exponential family of stochastic processes, see Küchler & Sørensen (1997). We can therefore take advantage of well known properties of exponential families of diffusions.

For the EM–algorithm the specification (5.5) implies that the function $q(\alpha, \beta)$ has the form

$$q(\alpha, \beta) = \sum_{i=1}^{k} \alpha_i H_{i,\beta} - \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} \alpha_i \alpha_j B_{i,j,\beta} + G_\beta,$$

where

$$\begin{aligned} H_{i,\beta} &= s_{i,\beta}(x_n) - s_{i,\beta}(x_1) \\ &+ \sum_{j=2}^{n} \mathbb{E}_{Z^{(j,\alpha_0,\beta_0)}} \left(\int_{t_{j-1}}^{t_j} \left[a_i(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))(\log \sigma_\beta)'(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) - \tfrac{1}{2} a_i'(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) \right] dt \right), \end{aligned}$$

with $s_{i,\beta}(x) = \int_{x^*}^{x} a_i(y)/\sigma_\beta^2(y) dy$,

$$B_{i,j,\beta} = \sum_{j=2}^{n} \mathbb{E}_{Z^{(j,\alpha_0,\beta_0)}} \left(\int_{t_{j-1}}^{t_j} \frac{a_i(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) a_j(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))}{\sigma_\beta^2(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))} dt \right),$$

and

$$\begin{aligned} G_\beta &= -\frac{1}{2} \log\left(\sigma_\beta(x_n)/\sigma_\beta(x_1)\right) - \frac{1}{2} \sum_{i=2}^{n} [h_\beta(x_i) - h_\beta(x_{i-1})]^2/(t_i - t_{i-1}) - \sum_{i=2}^{n} \log(\sigma_\beta(x_i)) \\ &+ \frac{1}{4} \sum_{j=2}^{n} \mathbb{E}_{Z^{(j,\alpha_0,\beta_0)}} \left(\int_{t_{j-1}}^{t_j} \left[\sigma_\beta''(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) \sigma_\beta(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) - \tfrac{1}{2} \{\sigma_\beta'(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))\}^2 \right] dt \right). \end{aligned}$$

For a fixed value of β, the function $\alpha \mapsto q(\alpha, \beta)$ is maximal for

$$\hat{\alpha}(\beta) = \boldsymbol{B}_\beta^{-1} \boldsymbol{H}_\beta,$$

where $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_k)^T$, $\boldsymbol{H}_\beta = (H_{1,\beta}, \ldots, H_{k,\beta})^T$ and $\boldsymbol{B}_\beta = \{B_{i,j,\beta}\}$. This is provided that $\boldsymbol{B}_\beta$ is invertible, which it is when the functions a_i, $i = 1, \ldots, k$ are linearly independent. Thus $q(\alpha, \beta)$ attains its maximal value at $(\hat{\alpha}(\hat{\beta}), \hat{\beta})$, where $\hat{\beta}$ maximizes

$$\beta \mapsto q(\hat{\alpha}(\beta), \beta) = \tfrac{1}{2} H_\beta^T B_\beta^{-1} H_\beta + G_\beta.$$

The Gibbs sampler for Bayesian inference for discretely observed diffusion processes proposed by Roberts & Stramer (2001) can also be modified by replacing the MCMC algorithm

for simulating diffusion bridges in that paper by our diffusion bridge simulation method. We will not go into any detail for general diffusions, but will limit ourselves to pointing out that when the drift has the form (5.5), then the (continuous time) posterior distribution of α simplifies. Choose as the prior for α the conjugate prior for an exponential family of diffusions, which is a multivariate normal distribution with expectation $\bar{\alpha}$ and covariance matrix $\boldsymbol{\Sigma}$. Then the posterior of α (given $\beta = \beta_0$ and given simulated diffusion bridges) is a k-dimensional normal distribution with expectation $(\boldsymbol{\Sigma}^{-1} + \tilde{\boldsymbol{B}}_{\beta_0})^{-1}(\boldsymbol{\Sigma}^{-1}\bar{\alpha} + \tilde{\boldsymbol{H}}_{\beta_0})$ and covariance matrix $(\boldsymbol{\Sigma}^{-1} + \tilde{\boldsymbol{B}}_{\beta_0})^{-1}$, where $\tilde{\boldsymbol{H}}_\beta = (\tilde{H}_{1,\beta}, \ldots, \tilde{H}_{k,\beta})^T$, $\boldsymbol{B}_\beta = \{B_{i,j,\beta}\}$,

$$\begin{aligned}\tilde{H}_{i,\beta} &= s_{i,\beta}(x_n) - s_{i,\beta}(x_1) \\ &\quad + \sum_{i=2}^{n} \int_{t_{i-1}}^{t_i} \left[a_i(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))(\log \sigma_\beta)'(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) - \tfrac{1}{2} a_i'(h_\beta^{-1}(Y_t^*(\beta,\beta_0))) \right] dt,\end{aligned}$$

and

$$\tilde{B}_{i,j,\beta} = \sum_{i=2}^{n} \int_{t_{i-1}}^{t_i} \frac{a_i(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))a_j(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))}{\sigma_\beta^2(h_\beta^{-1}(Y_t^*(\beta,\beta_0)))} dt.$$

6 Conclusion

We have presented a straightforward way of simulating an approximation to a diffusion bridge. The advantage of the new method is that it is very easy to understand and to implement, and that the same simple algorithm can be used for essentially all one-dimensional diffusion processes. The method allows the use of simple simulation procedures like the Euler scheme or the Milstein scheme for bridge simulation. The simulation study showed that the one-dimensional distributions obtained by the simple method compare accurately to the results from exact simulations for bridges corresponding to data that are likely in discrete–time samples from diffusion models.

The simulation study also showed that the computing time for the proposed algorithm for small time intervals is of the same order of magnitude as for the exact method and for long intervals is much faster than the exact method. Thus the new approximate diffusion bridge simulation method seems to be suitable for likelihood inference for discretely observed diffusions and can be used to simplify and in some cases speed up methods for likelihood inference and Bayesian inference (the EM-algorithm and the Gibbs sampler) for discretely observed diffusion processes. The method is also potentially useful for inference for more general diffusion type models like stochastic volatility models.

References

Aït-Sahalia, Y. (2002). "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach". *Econometrica*, 70:223–262.

Aït-Sahalia, Y. (2008). "Closed-form likelihood expansions for multivariate diffusions". *Ann. Statist.*, 36:906–937.

Barndorff-Nielsen, O. E. (1978). "Hyperbolic Distributions and Distributions on Hyperbolae". *Scand. J. Statist.*, 5:151–157.

Beskos, A.; Papaspiliopoulos, O. & Roberts, G. O. (2006). "Retrospective exact simulation of diffusion sample paths with applications". *Bernoulli*, 12:1077–1098.

Beskos, A.; Papaspiliopoulos, O. & Roberts, G. O. (2007). "A new factorization of diffusion measure with view towards simulation". To be published.

Beskos, A.; Papaspiliopoulos, O.; Roberts, G. O. & Fearnhead, P. (2006). "Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes". *J. R. Statist.Soc. B*, 68:333–382.

Bibby, B. M. & Sørensen, M. (2003). "Hyperbolic processes in finance". In Rachev, S., editor, *Handbook of Heavy Tailed Distributions in Finance*, pages 211–248. Elsevier Science.

Bladt, M. & Sørensen, M. (2005). "Statistical inference for discretely observed Markov jump processes". *J. R. Statist. Soc. B*, 67:395 – 410.

Bladt, M. & Sørensen, M. (2007). "Simple simulation of diffusion bridges with application to likelihood inference for diffusions". Working Paper 225, Centre for Analytical Finance, Aarhus.

Bladt, M. & Sørensen, M. (2009). "Efficient estimation of transition rates between credit ratings from observations at discrete time points". *Quantitative Finance*, 9:147 – 160.

Dempster, A. P.; Laird, N. M. & Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm (with discussion)". *J. Roy. Statist. Soc. B*, 39:1 – 38.

Durham, G. B. & Gallant, A. R. (2002). "Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes". *J. Business & Econom. Statist.*, 20:297–338.

Elerian, O.; Chib, S. & Shephard, N. (2001). "Likelihood inference for discretely observed non-linear diffusions". *Econometrica*, 69:959–993.

Eraker, B. (2001). "MCMC Analysis of Diffusion Models with Application to Finance". *Journal of Business and Economic Statistics*, 19:177–191.

Fitzsimmons, P.; Pitman, J. & Yor, M. (1992). "Markovian bridges: construction, Palm interpretation, and splicing". In et al., E. C., editor, *Seminar on Stochastic Processes*, pages 101–134. Birkhäuser. Prog. Probab., Vol. 32.

Forman, J. L. & Sørensen, M. (2008). "The Pearson diffusions: A class of statistically tractable diffusion processes". *Scand. J. Statist.*, 35:438–465.

Haussmann, U. G. & Pardoux, E. (1986). "Time reversal of diffusions". *Annals of Probability*, 14:1188–1205.

Ito, K. & McKean, H. P. (1965). *Diffusion Processes and their Sample Paths*. Springer, Berlin.

Kent, J. (1978). "Time-reversible diffusions". *Adv. Appl. Prob.*, 10:819–835.

Kloeden, P. E. & Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*. 3rd revised printing. Springer-Verlag, New York.

Küchler, U. & Sørensen, M. (1997). *Exponential Families of Stochastic Processes*. Springer, New York.

Millet, A.; Nualert, D. & Sanz, M. (1989). "Integration by parts and time reversal for diffusion processes". *Annals of Probability*, 17:208–238.

Ozaki, T. (1985). "Non-linear time series models and dynamical systems". In Hannan, E. J.; Krishnaiah, P. R. & Rao, M. M., editors, *Handbook of Statistics, Vol. 5*, pages 25–83. Elsevier Science Publishers.

Pedersen, A. R. (1995). "A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations". *Scand. J. Statist.*, 22:55–71.

Poulsen, R. (1999). "Approximate maximum likelihood estimation of discretely observed diffusion processes". Working Paper 29, Centre for Analytical Finance, Aarhus.

Roberts, G. O. & Stramer, O. (2001). "On inference for partially observed nonlinear diffusion models using Metropolis-Hastings algorithms". *Biometrika*, 88:603–621.

Research Papers 2010

2010-18: Ole E. Barndorff–Nielsen, Fred Espen Benth and Almut E. D. Veraart: Modelling energy spot prices by Lévy semistationary processes

2010-19: Jeroen V.K. Rombouts and Lars Stentoft: Multivariate Option Pricing with Time Varying Volatility and Correlations

2010-20: Charlotte Christiansen: Intertemporal Risk-Return Trade-off in

2010-21: Marco Aiolfi, Carlos Capistrán and Allan Timmermann : Forecast Combinations

2010-22: Ivan Nourdin, Giovanni Peccati and Mark Podolskij: Quantitative Breuer-Major Theorems

2010-23: Matias D. Cattaneo, Richard K. Crump and Michael Jansson: Bootstrapping Density-Weighted Average Derivatives

2010-24: Søren Johansen and Morten Ørregaard Nielsen: Likelihood inference for a fractionally cointegrated vector autoregressive model

2010-25: Tom Engsted and Bent Nielsen: Testing for rational bubbles in a co-explosive vector autoregression

2010-26: Robinson Kruse: On European monetary integration and the persistence of real effective exchange rates

2010-27: Sanne Hiller and Robinson Kruse: Milestones of European Integration: Which matters most for Export Openness?

2010-28: Robinson Kruse: Forecasting autoregressive time series under changing persistence

2010-29: Nikolaus Hautsch and Mark Podolskij: Pre-Averaging Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence

2010-30: Martin M. Andreasen: Non-linear DSGE Models and The Central Difference Kalman Filter

2010-31: Morten Ørregaard Nielsen and Per Frederiksen: Fully Modified Narrow-Band Least Squares Estimation of Weak Fractional Cointegration

2010-32: Mogens Bladt and Michael Sørensen: Simple simulation of diffusion bridges with application to likelihood inference for diffusions