Inference for stochastic volatility model using time change transformations

Kalogeropoulos, Konstantinos and Roberts, Gareth O. and Dellaportas, Petros

University of Cambridge, Engineering Department - Signal Processing Lab, University of Warwick, Statistics Department, Athens University of Economics and Business, Statistics Department

2007

Inference for stochastic volatility models using time change transformations

Konstantinos Kalogeropoulos
University of Cambridge, Department of Engineering - Signal Processing Lab

Gareth O. Roberts
University of Warwick, Department of Statistics

Petros Dellaportas
Athens University of Economics and Business, Department of Statistics

October 14, 2007

Abstract

We address the problem of parameter estimation for diffusion driven stochastic volatility models through Markov chain Monte Carlo (MCMC). To avoid degeneracy issues we introduce an innovative reparametrisation defined through transformations that operate on the time scale of the diffusion. A novel MCMC scheme which overcomes the inherent difficulties of time change transformations is also presented. The algorithm is fast to implement and applies to models with stochastic volatility. The methodology is tested through simulation based experiments and illustrated on data consisting of US treasury bill rates.

Keywords: Imputation, Markov chain Monte Carlo, Stochastic volatility

1 Introduction

Diffusion processes provide natural models for continuous time phenomena. They are used extensively in diverse areas such as finance, biology and physics. A diffusion process is defined through a stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t, \theta)dt + \sigma(t, X_t, \theta)dW_t,\ 0 \leq t \leq T, \tag{1}$$

where W is standard Brownian motion. The drift $\mu(.)$ and volatility $\sigma(.)$ reflect the instantaneous mean and standard deviation respectively. In this paper we assume the existence of a unique weak solution to (1), which translates into some regularity conditions (locally Lipschitz with a linear growth bound) on $\mu(.)$ and $\sigma(.)$; see chapter 5 of Rogers and Williams (1994) for more details.

The task of inference for diffusion processes is particularly challenging and has received remarkable attention in the recent literature; see Sørensen (2004) for an extensive review. The main difficulty is inherent in the nature of diffusions which are infinite dimensional objects. However, only a finite number of points may be observed and the marginal likelihood of these observations is generally unavailable in closed form. This has stimulated the development of various non-likelihood approaches which use indirect inference (Gouriéroux et al., 1993), estimating functions (Bibby and Sorensen, 1995), or the efficient method of moments (Gallant and Tauchen, 1996); see also Gallant and Long (1997).

Most likelihood based methods approach the likelihood function through the transition density of (1). Denote the observations by Y_k, $k = 0, \ldots, n$, and with t_k their corresponding times. If the dimension of Y_k equals that of X (for each k) we can use the Markov property to write the likelihood, given the initial point Y_0, as:

$$\mathcal{L}(Y, \theta | Y_0) = \prod_{k=1}^{n} p_k(Y_k | Y_{k-1}; \theta, \Delta), \quad \Delta = t_k - t_{k-1} \tag{2}$$

The transition densities $p_k(.)$ are not available in closed form but several approximations are available. They may be analytical, see Aït-Sahalia (2002), Aït-Sahalia (2005), or simulation based, see Pedersen (1995), Durham and Gallant (2002). They usually approximate the likelihood in a way so that the discretisation error can become arbitrarily small, although the methodology developed in Beskos et al. (2006) succeeds exact inference in the sense that it allows only for Monte Carlo error. A potential downside of these methods may be their dependence on the Markov property. In many interesting multidimensional diffusion models the observation regime is different and some of their components are not observed at all.

A famous such example is provided by stochastic volatility models, used extensively to model financial time series such as equity prices (Heston, 1993; Hull and White, 1987; Stein and Stein, 1991), or interest rates (Andersen and Lund, 1998; Durham, 2002; Gallant and Tauchen, 1998). A stochastic volatility model is usually represented by a 2-dimensional diffusion

$$\begin{pmatrix} dX_t \\ d\alpha_t \end{pmatrix} = \begin{pmatrix} \mu_x(X_t, \alpha_t, \theta) \\ \mu_\alpha(\alpha_t, \theta) \end{pmatrix} dt + \begin{pmatrix} \sigma_x(\alpha_t, \theta) & 0 \\ 0 & \sigma_\alpha(\alpha_t, \theta) \end{pmatrix} \begin{pmatrix} dB_t \\ dW_t \end{pmatrix}, \qquad (3)$$

where X denotes the observed equity (stock) log-price or the short term interest rate with volatility $\sigma_x(.)$, which is a function of a latent diffusion α. For the diffusion in (3), the Markov property may no longer hold; the distribution of a future stock price depends (besides the current price) on the current volatility which in turn depends on the entire price history. Stochastic volatility models are used

An alternative approach to the problem adopts Bayesian inference utilizing Markov chain Monte Carlo (MCMC) methods. Adhering to the Bayesian framework, a prior $p(\theta)$ is first assigned on the parameter vector θ. Then, given the observations Y, the posterior $p(\theta|Y)$ can be explored through data augmentation (Tanner and Wong, 1987), treating the unobserved paths of X (paths between observations) as missing data. The resulting algorithm alternates between updating θ and X. Initial MCMC schemes following this programme were introduced by Jones (1999); see also Jones (2003), Eraker (2001) and Elerian et al. (2001). However, as noted in the simulation based experiment of Elerian et al. (2001) and established theoretically by Roberts and Stramer (2001), any such algorithm's convergence properties will degenerate as the number of imputed points increases. The problem may be overcome with the reparametrisation of Roberts and Stramer (2001), and this scheme may be applied in all one-dimensional and some multi-dimensional contexts. However this framework does not cover general multidimensional diffusion models. Chib et al. (2005) and Kalogeropoulos (2007) offer appropriate reparametrisations but only for a class of stochastic volatility models. Alternative reparametrisations were introduced in Golightly and Wilkinson (2007); see also Golightly and Wilkinson (2006) for a sequential approach.

In this paper we introduce a novel reparametrisation that, unlike previous MCMC approaches, operates on the time scale of the observed diffusion rather than its path. This facilitates the construction of irreducible and efficient MCMC schemes, designed appropriately to accommodate the time change of the diffusion path. Our approach is general enough to cover almost every stochastic volatility model used in practice. The paper is organized as follows: Section 2 elaborates on the need for a transformation of the diffusion to avoid problematic MCMC algorithms. In Section 3 we introduce time change transformations whereas Section 4 provides the details for the corresponding non-trivial MCMC implementation. The proposed methodology of the paper is tested and illustrated through numerical experiments

in section 5, and on US treasury bill rates in section 6. Finally, section 7 concludes and provides some relevant discussion.

2 The necessity of reparametrisation

A Bayesian data augmentation scheme bypasses a problematic sampling from the posterior $\pi(\theta|Y)$ by introducing a latent variable $\mathcal{X}$ that simplifies the likelihood $\mathcal{L}(Y;\mathcal{X},\theta)$. It usually involves the following two steps:

1. `Simulate` $\mathcal{X}$ `conditional on` Y `and` θ`.`
2. `Simulate` θ `from the augmented conditional posterior which is proportional to` $\mathcal{L}(Y;\mathcal{X},\theta)\pi(\theta)$`.`

It is not hard to adapt our problem to this setting. Y represents the observations of the price process X. The latent variables $\mathcal{X}$ introduced to simplify the likelihood evaluations are discrete skeletons of diffusion paths between observations or entirely unobserved diffusions. In other words, $\mathcal{X}$ is a fine partition of multidimensional diffusion with drift $\mu_X(t,X_t,\theta)$ and diffusion matrix

$$\Sigma_X(t,X_t,\theta) \;=\; \sigma(t,X_t,\theta) \;\times\; \sigma(t,X_t,\theta)',$$

and the augmented dataset is $\mathcal{X}_{i\delta},\ i=0,\ldots,T/\delta$, where δ specifies the amount of augmentation. The likelihood can be approximated via the Euler scheme

$$\mathcal{L}^E(Y;\mathcal{X},\theta)=\prod_{i=1}^{T/\delta} p(\mathcal{X}_{i\delta}|\mathcal{X}_{(i-1)\delta}),\ \mathcal{X}_{i\delta}|\mathcal{X}_{(i-1)\delta}\sim\mathcal{N}\left(\mathcal{X}_{(i-1)\delta}+\delta\mu_{\mathcal{X}}(.),\delta\Sigma_{\mathcal{X}}(.)\right),$$

which is known to converge to the true likelihood $\mathcal{L}(Y;\mathcal{X},\theta)$ for small δ (Pedersen, 1995).

Another property of diffusion processes relates $\Sigma_{\mathcal{X}}(.)$ to the quadratic variation process. Specifically we know that

$$\lim_{\delta\to 0}\sum_{i=1}^{T/\delta}\left(\mathcal{X}_{i\delta}-\mathcal{X}_{(i-1)\delta}\right)\left(\mathcal{X}_{i\delta}-\mathcal{X}_{(i-1)\delta}\right)^T=\int_0^T\Sigma_{\mathcal{X}}(s,\mathcal{X}_s,\theta)ds\ \ a.s.$$

The solution of the equation above determines the diffusion matrix parameters. Hence, there exists perfect correlation between these parameters and $\mathcal{X}$ as $\delta\to 0$. This has disastrous implications for the mixing and convergence of the MCMC chain as it translates into reducibility for $\delta\to 0$. This issue was first noted by Roberts and Stramer (2001) for scalar diffusions and also confirmed by the simulation experiment of Elerian et al. (2001).

Nevertheless, it is not an MCMC specific problem. It turns out that the convergence of its deterministic analogue, EM algorithm, is problematic when the amount of information in the augmented data $\mathcal{X}$ strongly exceeds that of the observations. In our case $\mathcal{X}$ contains an infinite amount of information for $\delta \to 0$.

The problem may be resolved if we apply a transformation so that the algorithm based on the transformed diffusion is no longer reducible as $\delta \to 0$. Roberts and Stramer (2001) provide appropriate diffusion transformations for scalar diffusions. In a multivariate context this requires a transformation to a diffusion with unit volatility matrix; see for instance Kalogeropoulos et al. (2007). Aït-Sahalia (2005) terms such diffusions as reducible and proves the non-reducibility of stochastic volatility models that obey (3). The transformations introduced in this paper follow a slightly different route and target the time scale of the diffusion. One of the appealing features of such a reparametrisation is the generalisation to stochastic volatility models.

3 Time change transformations

For ease of illustration we first provide the time change transformation and the relevant likelihood function for scalar diffusion models with constant volatility. Nevertheless, one of the main advantages of this technique is the applicability to general stochastic volatility models.

3.1 Scalar diffusions

Consider a diffusion X defined through the following SDE:

$$dX_t = \mu(t, X_t, \theta)dt + \sigma dW_t^X, \;\; 0 < t < 1 \;\; \sigma > 0. \tag{4}$$

Without loss of generality, we assume a pair of observations $X_0 = y_0$ and $X_1 = y_1$. For more data, note that the same operations are possible for every pair of successive observations that are linked together through the Markov property. We introduce the latent 'missing' path of X for $0 \leq t \leq 1$, denoted by X^{mis}, so that $X = (y_0, X^{mis}, y_1)$. In the spirit of Roberts and Stramer (2001), the goal is to write the likelihood for θ, σ with respect to a parameter-free dominating measure. Using Girsanov's theorem we can get the Radon-Nikodym derivative between the law of the diffusion X, denoted by $\mathbb{P}^X$, and that of the

driftless diffusion $M = \sigma dW_t^X$ which represents Wiener measure and is denoted by $\mathbb{W}^X$. We can write

$$\frac{d\mathbb{P}(X)}{d\mathbb{W}^X} = G(t, X, \theta, \sigma) = \exp\left\{\int_0^T \frac{\mu(s, X_s, \theta)}{\sigma^2} dX_s - \frac{1}{2}\int_0^T \frac{\mu(s, X_s, \theta)^2}{\sigma^2} ds\right\}.$$

By factorizing $\mathbb{W}^X = \mathbb{W}_y^X \times Leb(y_1) \times f(y_1; \sigma^2)$, where $y_1 \sim \mathcal{N}(y_0, \sigma^2)$ and $Leb(.)$ denotes Lebesgue measure, we obtain

$$\frac{d\mathbb{P}(X^{mis}, y_0, y_1)}{d\left\{\mathbb{W}_y^X \times Leb(y)\right\}} = G(t, X, \theta, \sigma) f(y_1; \sigma),$$

where clearly the dominating measure depends on σ, since it reflects a Brownian bridge with volatility σ.

Now consider the time change transformation which first introduces a new time scale $\eta(t, sigma))$

$$\eta(t, \sigma) = \int_0^t \sigma^2 ds = \sigma^2 t, \tag{5}$$

and then defines the new transformed diffusion U as

$$U_t = \begin{cases} X_{\eta^{-1}(t,\sigma)}, & 0 \le t \le \sigma^2, \\ M_{\eta^{-1}(t,\sigma)}, & t > \sigma^2. \end{cases}$$

The definition for $t > \sigma^2$ is needed to ensure that tU is well defined for different values of $\sigma^2 > 0$ which is essential in the context of a MCMC algorithm. Using standard time change properties, see for example Oksendal (2000), the SDE for U is

$$dU_t = \begin{cases} \frac{1}{\sigma^2}\mu(t, U_t, \theta)dt + dW_t^U & 0 \le t \le \sigma^2, \\ dW_t^U, & t > \sigma^2, \end{cases}$$

where W^U is another Brownian motion at the time scale of U. By using Girsanov's theorem again, the law of U, denoted by $\mathbb{P}$, is given through its Radon-Nikodym derivative with respect to the law $\mathbb{W}^U$ of the Brownian motion W^U at the $U-$time scale:

$$\begin{aligned} \frac{d\mathbb{P}}{d\mathbb{W}^U} = G(t, U, \theta, \sigma) &= \exp\left\{\int_0^{+\infty} \frac{\mu(s, U_s, \theta)}{\sigma^2} dU_s - \frac{1}{2}\int_0^{+\infty} \frac{\mu(s, U_s, \theta)^2}{\sigma^4} ds\right\} \\ &= \exp\left\{\int_0^{\sigma^2} \frac{\mu(s, U_s, \theta)}{\sigma^2} dU_s - \frac{1}{2}\int_0^{\sigma^2} \frac{\mu(s, U_s, \theta)^2}{\sigma^4} ds\right\}. \end{aligned} \tag{6}$$

If we condition the Wiener measure on y at the new time scale, the likelihood can be written with respect to a Brownian bridge measure $\mathbb{W}_y^U$ as

$$d\mathbb{P}(U, y_0, y_1) = G(t, U, \theta, \sigma) f(y_1; \sigma) d\left\{\mathbb{W}_y^U \times Leb(y)\right\}.$$

However, this Brownian bridge is conditioned on the event $U_{\sigma^2} = y_1$ and therefore contains the parameter σ. For this reason we introduce a second transformation which applies to both the diffusion's time scale and its path. Define

$$\begin{aligned} U_t^0 &= (\sigma^2 - t) Z_{t/\{\sigma^2(\sigma^2 - t)\}}, \;\; 0 \le t < \sigma^2, \\ U_t^0 &= U_t - (1 - \frac{t}{\sigma^2}) y_0 - \frac{t}{\sigma^2} y_1, \;\; 0 \le t < \sigma^2. \end{aligned} \tag{7}$$

Note that this transformation is 1-1. Its inverse is given by

$$Z_t = \frac{1 + \sigma^2 t}{\sigma^2} U^0_{\sigma^4 t/(1+\sigma^2 t)}, \; 0 \le t < +\infty.$$

Applying Ito's formula and using time change properties we can also obtain the SDE of Z based on another driving Brownian motion W^Z operating at the $Z-$time:

$$dZ_t = \frac{\mu\left(t, \frac{\sigma^2}{1+t\sigma^2}\nu(Z_t, \sigma), \theta\right) + \nu(Z_t)\sigma^2}{1 + t\sigma^2} \, dt + dW_t^Z, \; 0 \le t < +\infty, \tag{8}$$

where $\nu(Z_t, \sigma) = U_t$. This operation essentially transforms to a diffusion that runs from 0 to $+\infty$ preserving the unit volatility. We can re-attempt to write the likelihood using Girsanov theorem and condition the dominating measure on y_1 to obtain $\mathbb{W}_y^Z$,

$$\frac{d\mathbb{P}(Z, y_0, y_1)}{d\left\{\mathbb{W}_y^Z \times Leb(y)\right\}} = G(t, Z, \theta, \sigma) f(y_1; \sigma). \tag{9}$$

Despite the fact that $G(Z, \theta, \sigma)$ contains integrals defined in $(0, +\infty)$, it is always finite being an $1-1$ transformation of the Radon-Nikodym derivative between $\mathbb{P}(U)$ and $\mathbb{W}^U$ given by (6). Using the following lemma, we prove that $\mathbb{W}_y^Z$ is the law of the standard Brownian motion and hence the likelihood is written with respect to a dominating measure that does not depend on any parameters.

Lemma 3.1 *Let W be a standard Brownian motion in $[0, +\infty)$. Consider the process defined for $0 \le t \le T$*

$$B_t = (T - t) W_{t/\{T(T-t)\}} + (1 - \frac{t}{T}) y_0 + \frac{t}{T} y_1, \; 0 \le t < T$$

Then B is a Brownian bridge from y_0 at time 0 to y_1 at time T.

Proof: See (Rogers and Williams, 1994, IV.40.1) for the case $y_0 = 0$, $T = 1$. The extension for general y_0 and T is trivial.

Corollary 3.1 *The measure $\mathbb{W}_y^Z$ is standard Wiener measure.*

Proof: Note that $\mathbb{W}_y^U$ reflects a Brownian bridge from y_0 at time 0 to y_1 at time T and we obtained $\mathbb{W}_y^Z$ by using the transformation of Lemma 3.1. Since this transformation is $1-1$, U is a Brownian bridge (under the dominating measure) if and only if Z is standard Brownian motion.

Note that $\mathbb{W}_y^Z$ is the probability law of the driftless version of the conditional diffusion Z, whereas the SDE in (8) corresponds to the unconditional version of Z itself. The conditional SDE of Z is generally not available but this does not create a problem. For the path updates we may use the fact that

$$\frac{d\mathbb{P}}{d\mathbb{W}_y^Z}(Z|y_0, y_1) = G(t, Z, \theta, \sigma)\frac{f(y_1;\sigma)}{f^P(y_1;\sigma)} \propto G(t, Z, \theta, \sigma), \tag{10}$$

where $\mathbb{P}_y$ is the law of the conditional version of Z and $f^P(.)$ is the density of y_1 under $\mathbb{P}$. Both $\mathbb{P}_y$ and $f^P(.)$ are generally unknown but $G(.)$ and $f(.)$, which appear in (9) and (10), are available.

3.2 Stochastic volatility models

Consider the general class of stochastic volatility models with SDE given by (3). Without loss of generality, we may assume a pair of observations ($X_0 = y_0$, $X_1 = y_1$) due to the Markov property of the 2-dimensional diffusion (X, α). The likelihood can then be divided into two parts: The first contains the marginal likelihood of the diffusion α and the remaining part corresponds to the diffusion X conditioned on the path of α

$$\mathbb{P}_\theta(X, \alpha) = \mathbb{P}_\theta(\alpha)\mathbb{P}_\theta(X|\alpha).$$

Denote the marginal likelihood for α by $\mathcal{L}_\alpha(\alpha, \theta)$. To overcome reducibility issues arising from the paths of α one may use the reparametrisations of Chib et al. (2005) or Kalogeropoulos (2007). The relevant transformations of the latter are

$$\beta_t = h(\alpha_t, \theta),\ \frac{\partial h(\alpha_t, \theta)}{\partial \alpha_t} = \{\sigma_\alpha(\alpha_t, \theta)\}^{-1},$$

$$\gamma_t = \beta_t - \beta_0, \;\; \beta_t = \eta(\gamma_t),$$

and the marginal likelihood for the transformed latent diffusion γ becomes

$$\mathcal{L}_\gamma(\gamma, \theta) = \frac{d\mathbb{P}}{d\mathbb{P}}(\gamma) = G\left\{\eta(\gamma), \theta\right\}. \tag{11}$$

By letting $\alpha_t = g_t^\gamma = h^{-1}(\eta(\gamma_t), \theta)$, the SDE of X conditional on γ becomes:

$$dX_t = \mu_x(X_t, g_t^\gamma, \theta)dt + \sigma_x(g_t^\gamma, \theta)dB_t, \;\; 0 \le t \le 1.$$

Given the paths of the diffusion α, the volatility function $\sigma_x(g_t^\gamma, \theta)$ may be viewed as a deterministic function of time. The situation is similar to that of the previous section. We can introduce a new time scale

$$\eta(t, \gamma, \theta) = \int_0^t \sigma_x^2(g_t^\gamma, \theta)ds,$$
$$T = \eta(t_k, \gamma, \theta),$$

and define U with the new time scale as before (M is a Brownian motion on the $U-$time scale)

$$U_t = \begin{cases} X_{\eta^{-1}(t)}, & 0 \le t \le T, \\ M_{\eta^{-1}(t)}, & t > T. \end{cases} \tag{12}$$

The SDE for U now becomes

$$dU_t = \left\{ \frac{\mu_x\left(U_t, \gamma_{\eta^{-1}(t,\gamma,\theta)}, \theta\right)}{\sigma_x^2(\gamma_{\eta^{-1}(t,\gamma,\theta)}, \theta)} \right\} dt + dW_t^U, \;\; 0 \le t \le T.$$

We obtain the Radon Nikodym derivative between the distribution of U with respect to that of the Brownian motion W^U,

$$\frac{d\mathbb{P}}{d\mathbb{W}^U} \;=\; G(U, \gamma, \theta),$$

and introduce $\mathbb{W}_y^U$ as before. The density of y_1 under $\mathbb{W}^U$, denoted by $f(y, \gamma, \theta)$, is just

$$f(y_1; \gamma, \theta) \equiv N(y_0, T).$$

The dominating measure $\mathbb{W}_y^U$ reflects a Brownian motion conditioned to equal y at a parameter depended time $T = \eta(t_{k+1}, \gamma, \theta)$. To remove this dependency we introduce a second time change

$$U_t^0 = (T-t)Z_{t/\{T(T-t)\}},\ 0 \le t < T, \tag{13}$$
$$U_t^0 = U_t - (1-\frac{t}{T})y_0 - \frac{t}{T}y_1,\ 0 \le t < T.$$

Therefore, the SDE for Z is now given by

$$dZ_t = \frac{T}{1+tT}\left\{\frac{\mu\left(\frac{T}{1+tT}\nu(Z_t), \gamma_{k(t,\gamma,\theta)}, \theta\right)}{\sigma_x^2(\gamma_{k(t,\gamma,\theta)}, \theta)} + \nu(Z_t)\right\}dt + dW_t^Z,\ \ 0 \le t < \infty,$$

where $k(t,\gamma,\theta)$ denotes the initial time scale of X and $\nu(Z_t) = U_t$.
Conditional on γ, the likelihood can be written in a similar manner as in (9):

$$\frac{d\mathbb{P}}{d\left\{\mathbb{W}_y^Z \times Leb(y)\right\}}(Z|y_0, y_1, \gamma) = G(Z,\gamma,\theta)f(y_1;\gamma,\theta) \tag{14}$$

It is not hard to see that $\mathbb{W}_y^Z$ reflects a standard Wiener measure and therefore the dominating measure is independent of parameters. To obtain the full likelihood we need to multiply the two parts given by (11) and (14).

3.3 Incorporating leverage effect

In the previous section we made the assumption that the increments of X and γ are independent, in other words we assumed no leverage effect. This assumption can be relaxed in the following way: In the presence of a leverage effect ρ, the SDE of X conditional on γ can be written as

$$dX_t = \mu_x(X_t, g_t^\gamma, \theta)dt + \rho\sigma_x(g_t^\gamma, \theta)dW_t + \sqrt{1-\rho^2}\sigma_x(g_t^\gamma, \theta)dB_t,\ \ 0 \le t \le t_k,$$

where W is the driving Brownian motion of γ). Note that given γ, W can be regarded as a function of γ and its parameters θ. Therefore, the term $\rho\sigma_x(g_t^\gamma, \theta)dW_t$ can be viewed as a deterministic function of time, and it can be treated as part of the drift of X_t. However, this operation introduces additional problems as the assumptions ensuring a weakly unique solution to the SDE of X are violated. To avoid this issue we introduce the infinitesimal transformation

$$X_t = \mathcal{H}(H_t, \rho, \gamma, \theta) = H_t + \int_0^t \rho\sigma_x(g_s^\gamma, \theta)dW_s,$$

which leads us to the following SDE for H:

$$dH_t = \mu_x \left\{ \mathcal{H}(X_t, \rho, \gamma, \theta), g_t^\gamma, \theta \right\} dt + \sqrt{1-\rho^2}\sigma_x(g_t^\gamma, \theta) dB_t, \;\; 0 \leq t \leq t_k.$$

We can now proceed as before, defining U and Z based on the SDE of H in a similar manner as in (12) and (13) respectively.

3.4 State dependent volatility

Consider the family of state dependent stochastic volatility models where conditional on γ, the SDE of X may be written as:

$$dX_t = \mu_x(X_t, g_t^\gamma, \theta)dt + \sigma_1(g_t^\gamma, \theta)\sigma_2(X_t, \theta)dB_t, \;\; 0 \leq t \leq t_k.$$

This class contains among others, the models of Andersen and Lund (1998), Gallant and Tauchen (1998), Durham (2002), Eraker (2001). In order to apply the time change transformations of section 3.2, we should first transform X to $\dot{X}_t$, through $\dot{X}_t = h(X_t, \theta)$, so that it takes the form of (3). Such a transformation, which may be viewed as the first transformation in Roberts and Stramer (2001), should satisfy the following differential equation

$$\frac{\partial h(X_t, \theta)}{\partial X_t} = \frac{1}{\sigma_2(X_t, \theta)}.$$

The time change transformations for U and Z may then be defined on the basis of $\ddot{X}$ that will now have volatility $\sigma_1(g_t^\gamma, \theta)$. The transformation $h(.)$ also applies to the observations $(\dot{y}_0, \; \dot{y}_1)$ which are now functions of parameters. This would translated in a parameter dependent likelihood dominating measure, had it not been for the second step in (12) which in this case acts like the second transformation in Roberts and Stramer (2001). Note that the parameters of $\sigma_2(X_t, \theta)$ enter the reparametrised likelihood in two ways: first through the $f(y; \gamma, \theta)$ which now should include the relevant Jacobian term, and second through the drift of Z which is centered at 0 based on the transformed observations.

3.5 Multivariate stochastic volatility models

We may use the techniques of section 3.3 to define time change transformations for multidimensional diffusions. Consider a $d-$dimensional version of the SDE in (4) where σ now is a 2×2 matrix ($[\sigma]_{ij} = \sigma_{ij}$). As noted in Kalogeropoulos et al. (2007), the mapping between σ and the volatility matrix $\sigma\sigma^T$ should be 1-1 in order to ensure identifiability of the σ

parameters. A way to achieve this, is by allowing σ to be the lower triangular matrix that produces the Cholesky decomposition of $\sigma\sigma^T$. For $d = 2$, the SDE of such a diffusion is given by

$$dX_t^{\{1\}} = \mu(X_t^{\{1\}}, X_t^{\{2\}}, \theta)dt + \sigma_{11}dB_t,$$
$$dX_t^{\{2\}} = \mu(X_t^{\{1\}}, X_t^{\{2\}}, \theta)dt + \sigma_{21}dB_t + \sigma_{22}dW_t.$$

The time change transformations for $X^{\{1\}}$ will be exactly as in section 3.1. For $X^{\{2\}}$ note that given $X^{\{1\}}$ the term $\sigma_{21}dB_t$ is now a deterministic function of time and may be treated as part of the drift. Thus, we may proceed following the route of the section 3.3.

Similar transformations can be applied for diffusions that have, or may be transformed to have, volatility functions independent of their paths. For example we may assume two correlated price processes with correlation ρ_x:

$$[\sigma]_{11} = \sigma_x^{\{1\}}(g_t^\gamma, \theta),$$
$$[\sigma]_{21} = \rho_x\sigma_x^{\{2\}}(g_t^\gamma, \theta),$$
$$[\sigma]_{22} = \sqrt{1-\rho_x^2}\sigma_x^{\{2\}}(g_t^\gamma, \theta).$$

We may proceed in a similar manner for multivariate stochastic volatility models of general dimension d.

4 MCMC implementation

The construction of an appropriate data augmentation algorithm involves several issues. The time change transformations introduce three interesting features to the MCMC algorithm which we address separately: the presence of three time scales; the need to update diffusion paths that run from 0 to $+\infty$; and the fact that time scales depend on parameters. For ease of illustration we will assume the simple case of a univariate diffusion with constant volatility and a pair observations ($X_0 = y_0$ and $X_1 = y_1$). Extensions and generalisations of the algorithm for stochastic volatility models are noted where appropriate.

4.1 Three time scales

We introduce m intermediate points of X at equidistant times between 0 and 1, to give $X = \{X_{i/(m+1)},\ i = 0, 1, \ldots, m+1\}$. In addition, we make the assumption that m is

large enough for accurate likelihood approximations and any error induced by the time discretisation is negligible for the purposes of our analysis.

Given a value of the time scale parameter σ, we can get the $U-$time points by applying (5) to each one of the existing points X, so that

$$U_{\sigma^2 i/(m+1)} = X_{i/(m+1)},\ i = 0, 1, \ldots, m+1.$$

Note that it is only the times that change, the values of the diffusion remain intact. In a stochastic volatility model we would use the quantities

$$\int_{\frac{i}{m+1}}^{\frac{i+1}{m+1}} \sigma_x^2(.)ds$$

for each pair of consecutive imputed points.

The points of Z are multiplied by a time factor which corrects the deviations from unit volatility. The $Z-$time points may be obtained by

$$t_i^Z = \frac{\sigma^2 i/(m+1)}{\sigma^2(\sigma^2 - \sigma^2 i/(m+1))},\ i = 0, 1, \ldots, m.$$

Clearly this does not apply to the last point which occurs at time $+\infty$. Therefore, the paths of X, or U, are more convenient and may be used for likelihood evaluations exploiting the fact that the relevant transformations are 1-1. However, the component of the relevant Gibbs sampling scheme is the diffusion Z.

Figure 1 shows a graphical representation of X, U and Z plotted against their corresponding time scales for $\sigma = \sqrt{2}$ and $m = 7$. Although X and U have the same values, their volatilities are $\sqrt{2}$ and 1 respectively. The ending point of Z does not appear on the graph as it occurs at time $+\infty$.

[Figure 1 about here.]

4.2 Updating the paths of Z

The paths of Z may be updated using an independence sampler with the reference measure as a proposal. Here $\mathbb{W}^Z$ reflects a Brownian motion at the $Z-$time which is fixed given the current values of the time-scale parameter(s). An appropriate algorithm is given by the following steps.

- `Step 1: Propose a Brownian motion on the` $Z-$`time, say` Z^*`. The value at the endpoint (time` $+\infty$`) is not needed.`

- `Step 2: Transform back to` U^*`, using (7).`
- `Step 3: Accept with probability:` $\min\left\{1, \frac{G(U^*,\theta,\sigma)}{G(U^*,\theta,\sigma)}\right\}$.

4.3 Updating time scale parameters

The updates of parameters that define the time scale, such as σ, are of particular interest. In almost all cases, their conditional posterior density is not available in closed form, and Metropolis steps are inevitable. However, the proposed values of these parameters will imply different $Z-$ time scales. In other words, for each potential proposed value for σ there exists a different set of $Z-$ points needed for accurate approximations of the likelihood the Metropolis accept-reject probabilities. In theory, this would pose no issues had we been able to store an infinitely thin partition of Z, but of course this is not possible.

We use retrospective sampling ideas; see Papaspiliopoulos and Roberts (2005) and Beskos and Roberts (2005) for applications in different contexts. Under the assumption of a sufficiently fine partition of Z, all the non-recorded intermediate points contribute nothing to the likelihood and they are irrelevant in that respect; the set of recorded points is sufficient for likelihood approximation purposes. Alternatively, we may argue that their distribution is given by the likelihood reference measure which reflects a Brownian motion. Thus, they can be drawn after the proposal of the candidate value of the time scale parameter. To ensure compatibility with the recorded partition of Z, it suffices to condition on their neighboring points. This is easily done using standard Brownian bridge properties: Suppose that we want to simulate the value of Z at time t_b which fall between the recorded values at times t_a and t_c, so that $t_a \leq t_b \leq t_c$. Denote by Z_{t_a} and Z_{t_c} the corresponding Z values. Under the assumption that Z is distributed according to $\mathbb{W}_y^Z$ between t_a and t_c we have that

$$Z_{t_b} |\, Z_{t_a}, Z_{t_c} \sim N\left\{\frac{(t_b - t_a)Z_{t_c} + (t_c - t_b)Z_{t_a}}{t_c - t_a},\ \frac{(t_b - t_a)(t_c - t_b)}{t_c - t_a}\right\}. \tag{15}$$

The situation is pictured in Figure 2, where the black bullets represent the stored points and the triangles the new points required for a proposed value of σ. The latter should be drawn retrospectively given the former via (15).

[Figure 2 about here.]

A suitable algorithm for the $\sigma-$updates may be summarized through the following steps:

- `Step 1: Propose a candidate value for` σ`, say` σ^*.

- Step 2: Repeat for each pair of successive points:
  - Use (5) and (7) to get the new times associated with it.
  - Draw the values of Z at the new times using (15).
  - Transform back to U^*, using (7).

  Form the entire path U^* by appropriately joining its bits.
- Step 3: Accept with probability: $\min\left\{1, \frac{G(U^*,\theta,\sigma^*)f(y;\sigma^*)}{G(U^*,\theta,\sigma)f(y;\sigma)}\right\}$.

Note that in a stochastic volatility model the paths of the transformed diffusion γ_t are associated with the time scale of the Z. Therefore a similar algorithm may be used for their updates.

5 Simulations

As discussed in section 2, appropriate reparametrisations are necessary to avoid issues regarding the mixing and convergence of the MCMC algorithm. In fact, the chain becomes reducible as the level of augmentation increases. This is also verified by the numerical examples performed in Kalogeropoulos (2007) even in very simple stochastic volatility models. In this section we perform a simulation based experiment to check the immunity of MCMC schemes to increasing levels of augmentation, as well as the ability of our estimation procedure to retrieve the correct values of the diffusion parameters despite the fact that the series is partially observed at only a finite number of points. We simulate data from the following stochastic volatility model

$$\begin{aligned} dX_t &= \kappa_x(\mu_x - X_t)dt + \rho\exp(\alpha_t/2)dW_t + \sqrt{1-\rho^2}\exp(\alpha_t/2)dB_t, \\ d\alpha_t &= \kappa_\alpha(\mu_\alpha - \alpha_t)dt + \sigma dW_t, \end{aligned}$$

where B and W are independent Brownian motions, and ρ reflects the correlation between the increments of X and α, also term as leverage effect. A high frequency Euler approximating scheme with a step of 0.001 was used for the simulation of the diffusion paths. Specifically, $500,001$ points were drawn and one value of X for every 1000 was recorded, thus forming a

dataset of 501 observations of X at $0 \leq t \leq 500$. The parameter values were set to $\rho = -0.5$, $\sigma = 0.4$, $\kappa_x = 0.2$, $\mu_x = 0.1$, $\kappa_\alpha = 0.3$ and $\mu_\alpha = -0.2$

The transformations required to construct an irreducible data augmentation scheme are listed below. First we transform α to γ through

$$\gamma_t = \frac{\alpha_t - \alpha_0}{\sigma},\ 0 \leq t \leq 500,$$
$$\alpha_t = \nu(\gamma_t, \sigma, \alpha_0) = \alpha_0 + \sigma\gamma_t.$$

Given γ, and for each pair of consecutive observation times t_{k-1} and t_k $(k = 1, 2, \ldots, 500)$ on X, we transform as follows: First, we remove the term introduced from the leverage effect

$$H_t = X_t - \int_{t_{k-1}}^{t} \rho \exp\left\{\nu(\gamma_s, \sigma, \alpha_0)/2\right\} dW_s,\ t_{k-1} \leq t \leq t_k,$$

and consequently we set

$$\eta(t) = \int_{t_{k-1}}^{t} (1-\rho)^2 \exp\left\{\nu(\gamma_s, \sigma, \alpha_0\right\} ds.$$

Then, U and Z may be defined again from 12 and 13 respectively, but based on H rather on X. The elements of the MCMC scheme are Z,γ, α_0 and the parameters $(\kappa_x, \mu_x, \kappa_\alpha, \mu_\alpha, \rho, \sigma)$.

We proceed by assigning flat priors to all the parameters, restricting κ_x, κ_α, σ to be positive and ρ to be in $(-1, 1)$. The number of imputed points was set to 30 and 50, the length of the overlapping blocks needed for the updates of γ was 2, and the relevant acceptance rate 75% whereas the acceptance rate for X was 95%. Figure 3 shows autocorrelation plots for the 2-dimensional diffusion's $(X, \alpha)'$ volatility parameters ρ and σ. There is no sign of any increase in the autocorrelation to raise suspicions against the irreducibility of the chain. Figure 4 shows density plots for all parameters and both values of m. These plots indicate a sufficiently fine discretisation and a good agreement with true values of the parameters. The latter is also confirmed by Table 1.

[Figure 3 about here.]

[Figure 4 about here.]

[Table 1 about here.]

6 Application: US treasury bill rates

To illustrate the time change methodology we fit a stochastic volatility model to US treasury bill rates. The dataset consists of 1809 weekly observations (Wednesday) of the 3−month US Treasury bill rate from the 5th of January 1962 up to the 30th of August 1996. The data are plotted in Figure 5.

[Figure 5 about here.]

Previous analyses of these data include Andersen and Lund (1998), Gallant and Tauchen (1998), Durham (2002), Durham and Gallant (2002), Eraker (2001), and Golightly and Wilkinson (2006). Apart from some slight deviations the adopted stochastic volatility models consisted of the following SDE.

$$\begin{aligned} dr_t &= (\theta_0 - \theta_1 r_t)dt + r_t^{\psi} \exp(\alpha_t/2) dB_t, \\ d\alpha_t &= \kappa(\mu - \alpha_t)dt + \sigma dW_t, \end{aligned} \tag{16}$$

with independent Brownian motions B and W. In some cases the following equivalent model was used:

$$\begin{aligned} dr_t &= (\theta_0 - \theta_1 r_t)dt + \sigma_r r_t^{\psi} \exp(\alpha_t/2) dB_t, \\ d\alpha_t &= -\kappa\alpha_t dt + \sigma dW_t. \end{aligned} \tag{17}$$

We proceed with the model in (16), as posterior draws of its parameters exhibit substantially less autocorrelation. In line with Gallant and Tauchen (1998) and Golightly and Wilkinson (2006), we also set $\psi = 1$. Eraker (2001), Durham (2002) and Durham and Gallant (2002) assume general 'elasticity of variance' ψ but their estimates do not indicate a significant deviation from 1. By setting $X_t = log(r_t)$, the volatility of X_t becomes $\exp(\alpha_t/2)$. Therefore the $U-$time for two consecutive observation times t_{k-1} and t_k is defined as

$$\eta(t) = \int_{t_{k-1}}^{t} \exp(\alpha_t) ds,$$

and U and Z are given by (12) and (13) respectively . We also transform α to γ as before:

$$\gamma_t = \frac{\alpha_t - \alpha_0}{\sigma},$$

$$\alpha_t = \nu(\gamma_t, \sigma, \alpha_0) = \alpha_0 + \sigma\gamma_t.$$

We applied MCMC algorithms based on Z and γ to sample from the posterior of the parameters θ_0, θ_1, κ, μ and σ. The time was measured in years setting the distance between successive Wednesdays to 5/252. Non-informative priors were assigned to all the parameters, restricting κ and σ to be positive to ensure identifiability and eliminate the possibility of explosion. The algorithm was run for $50,000$ iterations and for m equal to 10 and 20. To optimize the efficiency of the chain we set the length of the overlapping blocks of γ to 10 which produced an acceptance rate of 51.9%. The corresponding acceptance rate for Z was 98.6% .

The kernel density plots of the posterior parameters and likelihood (Figure 6) indicate that a discretisation from an m of 10 or 20 provide reasonable approximations. The corresponding autocorrelation plots of Figure 7 do not show increasing autocorrelation in m, a feature that would reveal reducibility issues. Finally, summaries of the posterior draws for all the parameters are provided in Table 2. The parameters κ, μ and σ are different from 0 verifying the existence of stochastic volatility. On the other hand, there is no evidence to support the existence of mean reversion on the rate, as θ_0 and θ_1 are not far from 0. The results are in line with those of Durham (2002), Durham and Gallant (2002) and Golightly and Wilkinson (2006).

[Figure 6 about here.]

[Figure 7 about here.]

[Table 2 about here.]

7 Discussion

Data augmentation MCMC schemes constitute a very useful tool for likelihood-based inference on diffusion models. They may not have the appealing properties of complete elimination of the time discretisation error (Beskos et al., 2006), or the closed form approximate likelihood expressions of Aït-Sahalia (2002), but nevertheless they give a satisfactory and very general solution to the problem. However data augmentation schemes require careful construction to avoid the degeneracy issues described at the beginning of this paper.

Here, we introduce an innovative transformation which operates by altering the time axis of the diffusion. To accommodate the special features of time change transformations we also

introduce a novel efficient MCMC scheme which mixes rapidly and is not provibitively computationally expensive. Our method is also easy to implement and introduces no additional approximation error other than that included in methodologies based on a discretisation of the diffusion path. Moreover it is general enough to include general stochastic volatility models.

Further work will consider problems with state-dependent volatility and models which involve jump diffusions, to which the methodology introduced here can be easily applied. Fundamental to our approach here has been the introduction of a non-centered parameterisation to decouple dependence inherent in the model between missing data and volatility parameters. However non-centered constructions are not unique, as illustrated by the choice in the diffusion context between the state rescaling approaches of Golightly and Wilkinson (2007); Roberts and Stramer (2001) and the time-stretching strategy adopted here. Clearly, further work is required to investigate the relative merits of these approaches in different situations.

8 Acknowledgements

This work was supported by the EPSRC grant GR/S61577/01 and by the Marie Curie fellowships. The authors would also like to thank Omiros Papaspiliopoulos for helpful discussions.

References

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: a closed form approximation approach. *Econometrica*, 70:223–262.

Aït-Sahalia, Y. (2005). Closed form likelihood expansions for multivariate diffusions. In preparation.

Andersen, T. G. and Lund, J. (1998). Estimating continuous-time stochastic volatility models of the short term interest rate. *Journal of Econometrics*, 77:343–377.

Beskos, A., Papaspiliopoulos, O., Roberts, G., and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):333–382.

Beskos, A. and Roberts, G. O. (2005). Exact simulation of diffusions. *Ann. Appl. Probab.*, 15(4):2422–2444.

Bibby, B. and Sorensen, M. (1995). Martingale estimating functions for discretely observed diffusion processes. *Bernoulli*, 1:17–39.

Chib, S., Pitt, M. K., and Shephard, N. (2005). Likelihood based inference for diffusion models. Submitted.

Durham, G. B. (2002). Likelihod based specification analysis of continuous time models of the short term interest rate. *Journal of Financial Economics*. Forthcoming.

Durham, G. B. and Gallant, A. R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *J. Bus. Econom. Statist.*, 20(3):297–316. With comments and a reply by the authors.

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

Eraker, B. (2001). Markov chain Monte Carlo analysis of diffusion models with application to finance. *J. Bus. Econom. Statist.*, 19(2):177–191.

Gallant, A. R. and Long, J. R. (1997). Estimating stochastic differential equations efficiently by minimum chi-squared. *Biometrika*, 84(1):125–141.

Gallant, A. R. and Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12(4):657–681.

Gallant, A. R. and Tauchen, G. (1998). Reprojecting partially observed systems with applications to interest rate diffusions. *Journal of American Statistical Association*, 93(441):10–24.

Golightly, A. and Wilkinson, D. (2006). Bayesian sequential inference for nonlinear multivariate diffusions. *Statistics and Computing*, 16:323–338.

Golightly, A. and Wilkinson, D. (2007). Bayesian inference for nonlinear multivariate diffusions observed with error. *Computational Statistics and Data Analysis*. In press.

Gouriéroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics*, 8:S85–S118.

Heston, S. (1993). A closed-form solution for options with stochastic volatility. with applications to bonds and currency options. *Review of Financial Studies*, 6:327–343.

Hull, J. C. and White, A. D. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300.

Jones, C. S. (1999). Bayesian estimation of continuous-time finance models. Unpublished paper, Simon School of Business, University of Rochester.

Jones, C. S. (2003). Nonlinear mean reversion in the short-term interest rate. *The review of financial studies*, 16:793–843.

Kalogeropoulos, K. (2007). Likelihood based inference for a class of multidimensional diffusions with unobserved paths. *Journal of Statistical Planning and Inference*. To appear.

Kalogeropoulos, K., Dellaportas, P., and Roberts, G. (2007). Bayesian inference for correllated diffusions. In submision.

Oksendal, B. (2000). *Stochastic differential equations*. Springer, 5th edition.

Papaspiliopoulos, O. and Roberts, G. (2005). Retrospective MCMC for Dirichlet process hierarchical models. Submitted.

Pedersen, A. R. (1995). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Statist.*, 22(1):55–71.

Roberts, G. and Stramer, O. (2001). On inference for partial observed nonlinear diffusion models using the metropolis-hastings algorithm. *Biometrika*, 88(3):603–621.

Rogers, L. C. G. and Williams, D. (1994). *Diffusions, Markov processes and martingales, 2, Ito calculus*. Wiley, Chicester.

Sørensen, H. (2004). Parametric inference for diffusion processes observed at discrete points in time: a survey. *International Statistical Review*, 72(3):337–354.

Stein, E. M. and Stein, J. C. (1991). Stock proce distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4(4):727–752.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540.

List of Figures

1 Plots of a sample path for X, U and Z against their corresponding times for $\sigma = \sqrt{2}$ and $m = 7$. Z equals 0 at time $+\infty$. 24

2 Updates of time scale parameters: For every proposed value of them, new points are required and should obtained conditional on the stored points. . . 25

3 Autocorrelation plots for the posterior draws of ρ and σ for different numbers of imputed points $m = 30, 50$. Simulation example of Chapter 3. 26

4 Kernel densities of the posterior draws of all the parameters for different numbers of imputed points $m = 30, 50$. Simulation example of Chapter 3. 27

5 Weekly 3−month US Treasury bill rate from the 5th of January 1962 up to the 30th of August 1996. 28

6 Kernel densities of the posterior draws of all the parameters and the log-likelihood for different values of imputed points $m = 10, 20$. Example on Weekly 3−month US Treasury bill rates. 29

7 Autocorrelation plots for the posterior draws of the model parameters for different numbers of imputed points $m = 10, 20$ for the analysis of Weekly 3−month US Treasury bill rates. 30

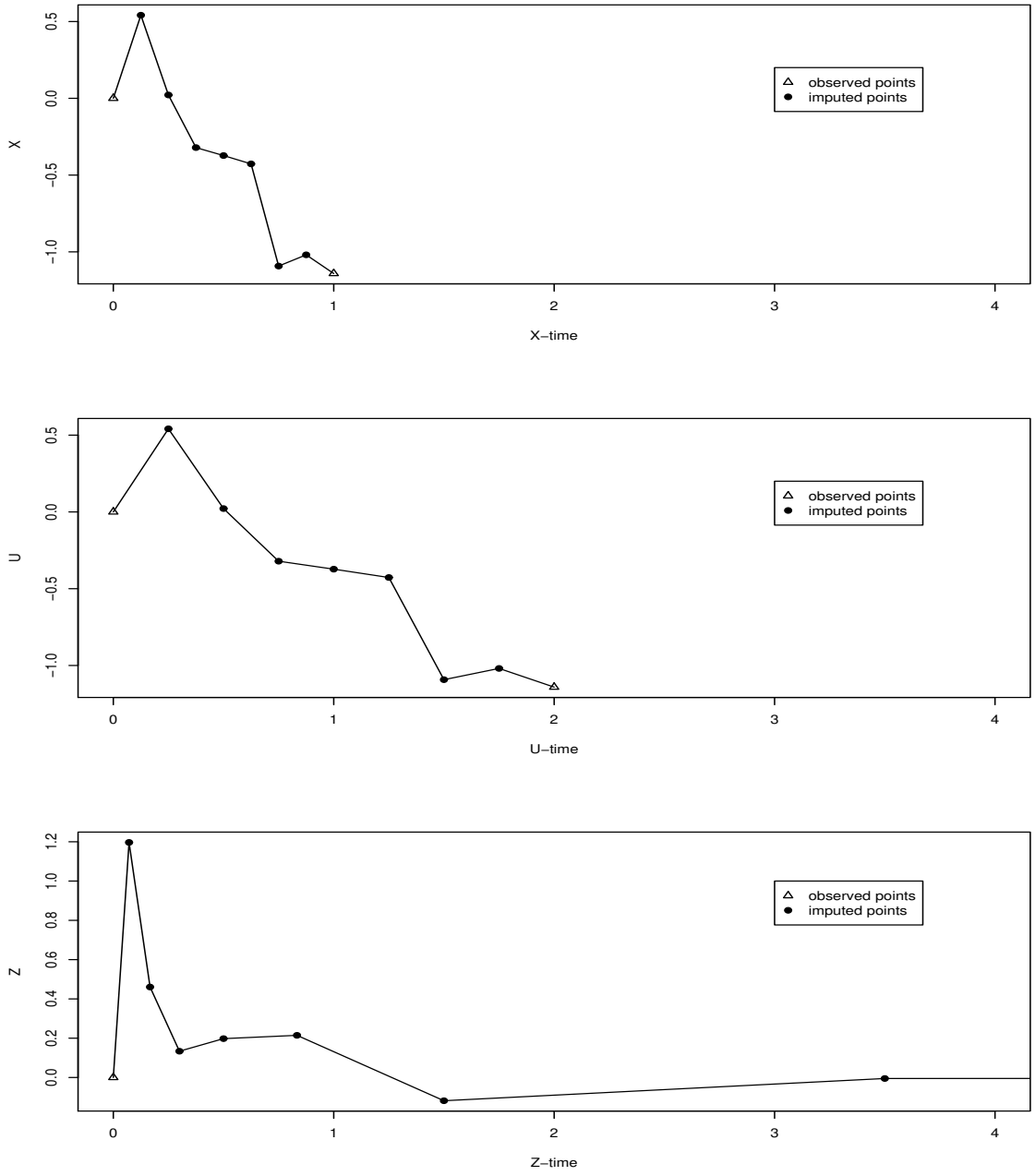


Figure 1: Plots of a sample path for X, U and Z against their corresponding times for $\sigma = \sqrt{2}$ and $m = 7$. Z equals 0 at time $+\infty$.

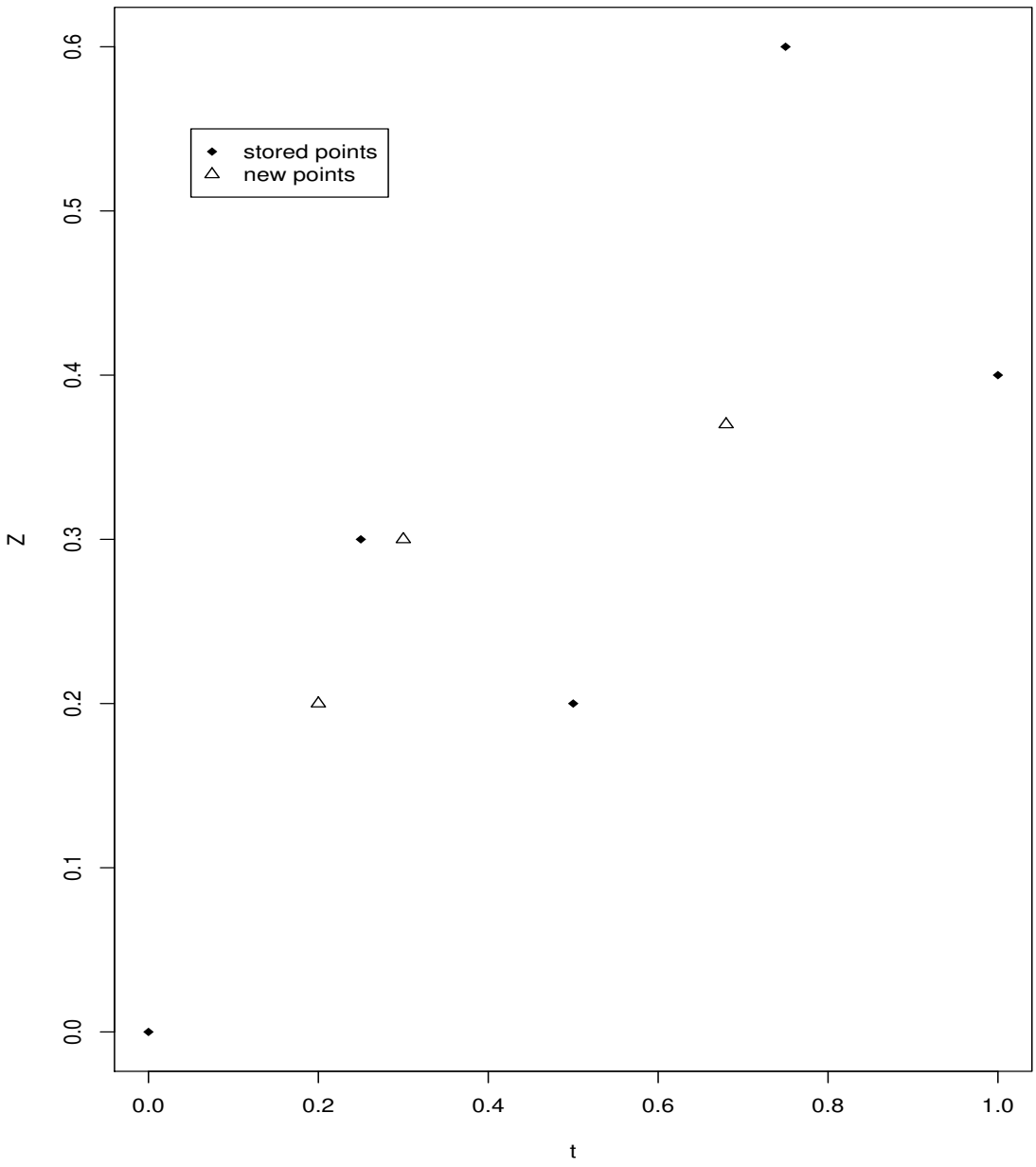
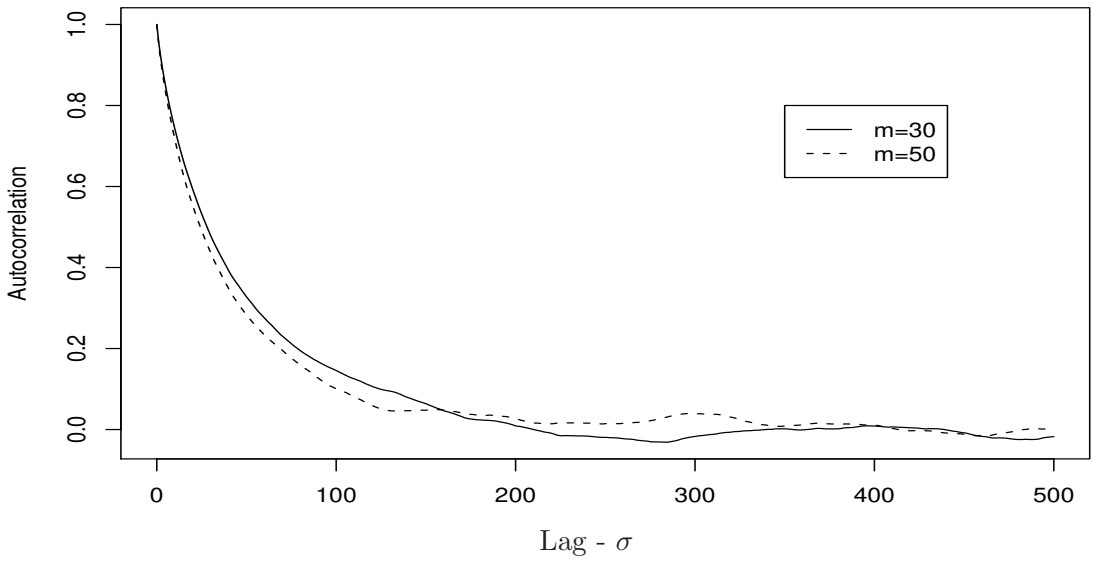


Figure 2: Updates of time scale parameters: For every proposed value of them, new points are required and should obtained conditional on the stored points.



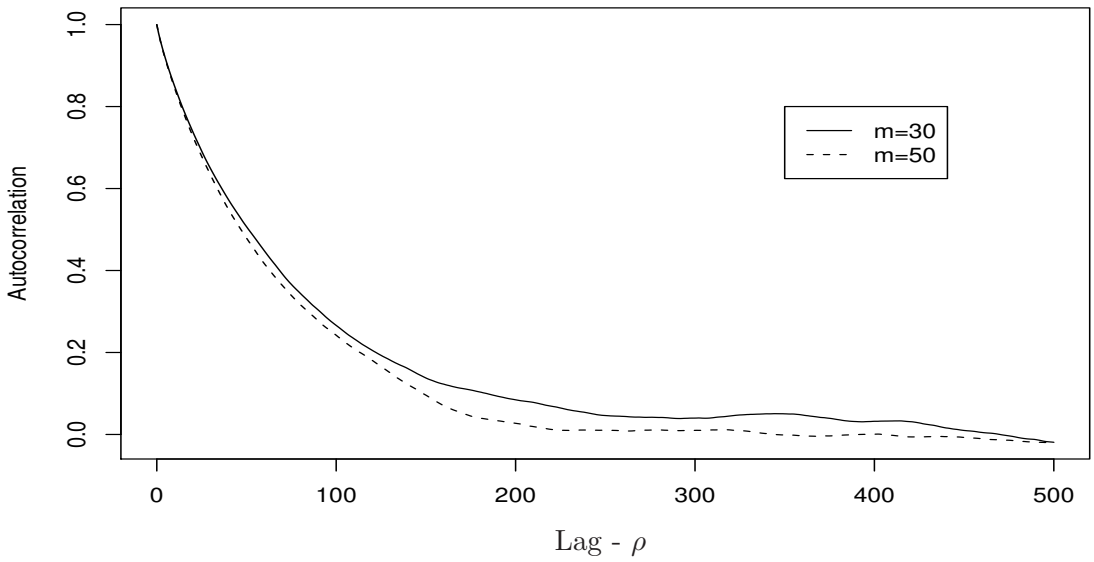


Figure 3: Autocorrelation plots for the posterior draws of ρ and σ for different numbers of imputed points $m = 30, 50$. Simulation example of Chapter 3.

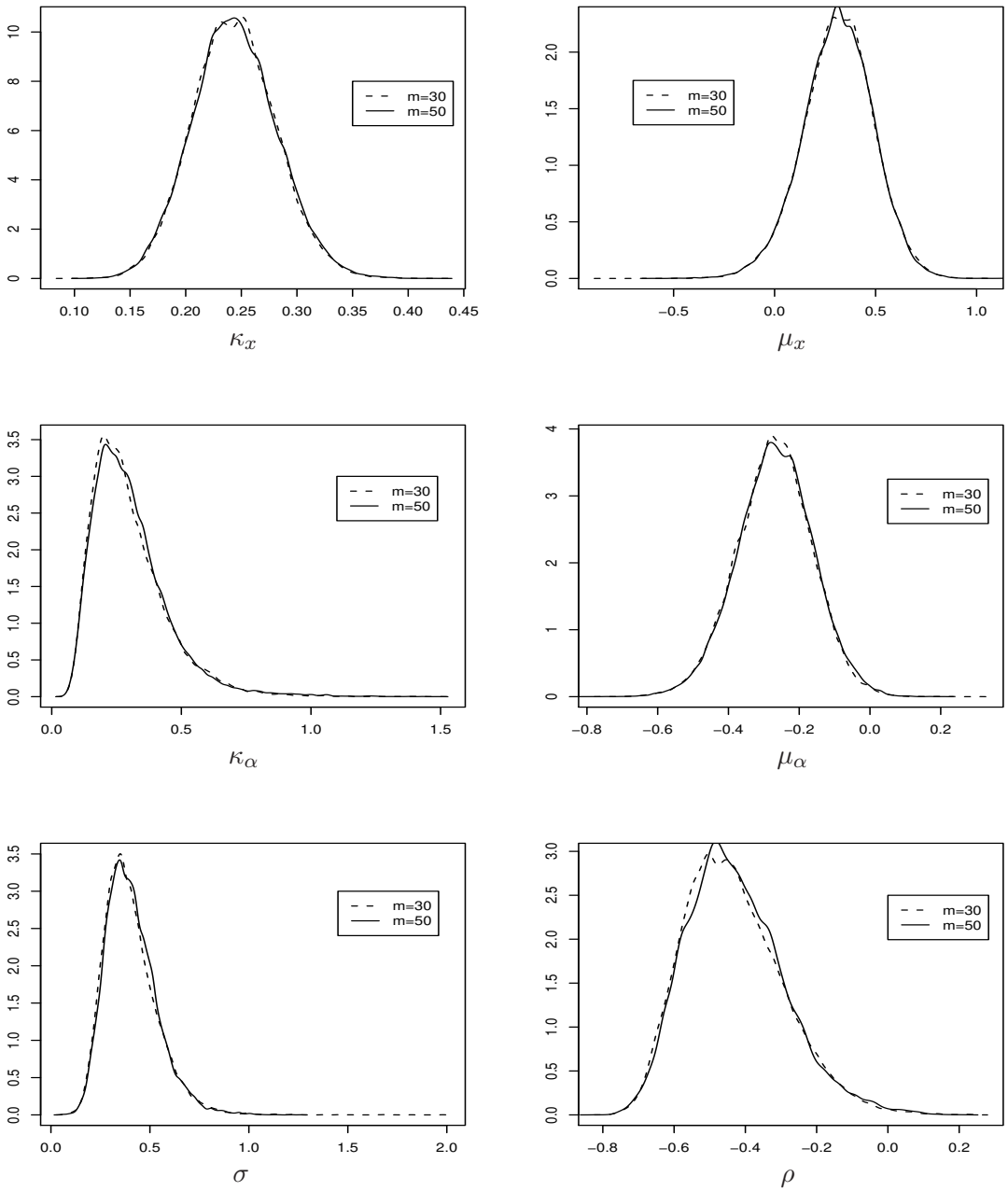


Figure 4: Kernel densities of the posterior draws of all the parameters for different numbers of imputed points $m = 30, 50$. Simulation example of Chapter 3.

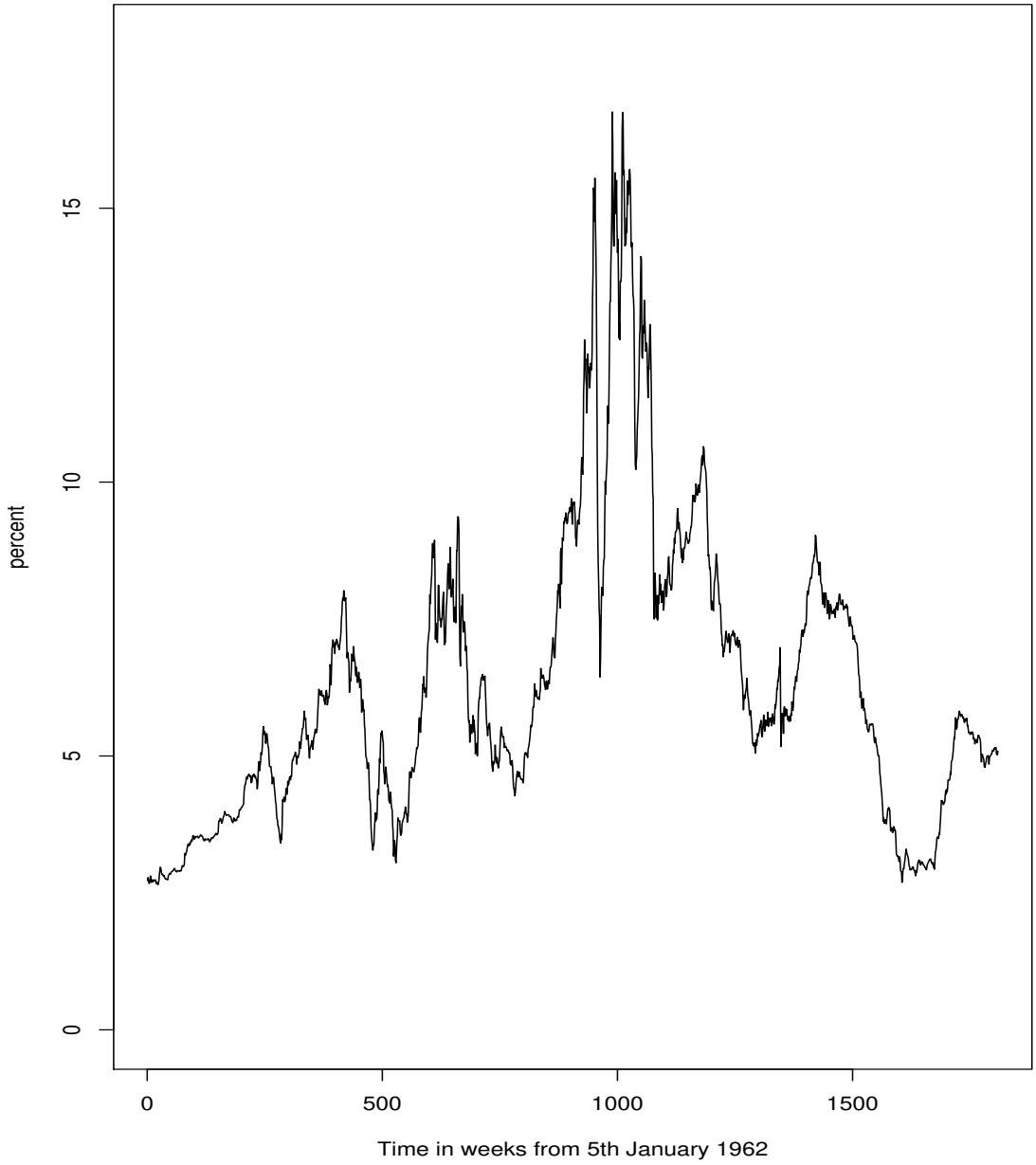


Figure 5: Weekly 3−month US Treasury bill rate from the 5th of January 1962 up to the 30th of August 1996.

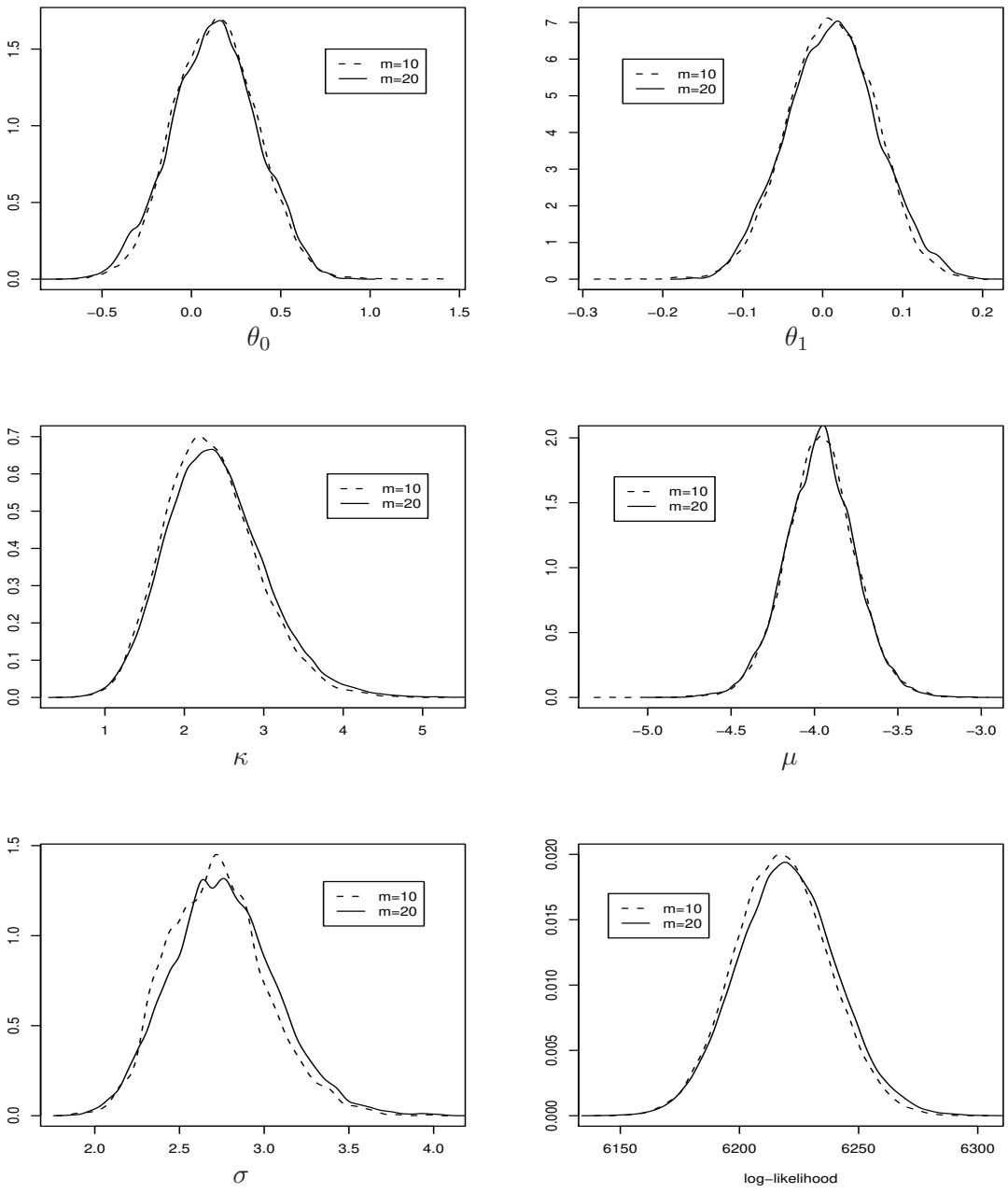


Figure 6: Kernel densities of the posterior draws of all the parameters and the log-likelihood for different values of imputed points $m = 10, 20$. Example on Weekly 3−month US Treasury bill rates.

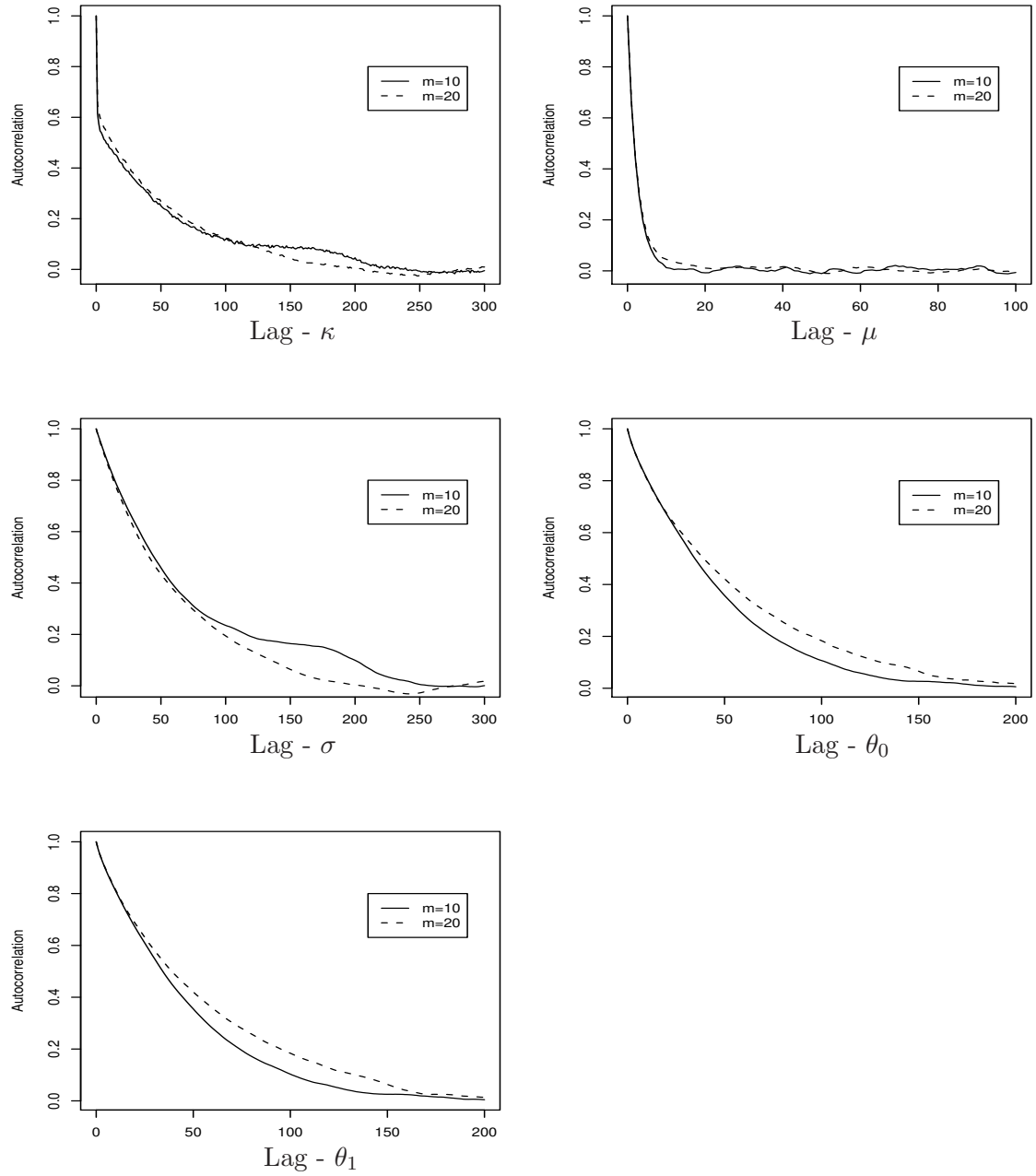


Figure 7: Autocorrelation plots for the posterior draws of the model parameters for different numbers of imputed points $m = 10, 20$ for the analysis of Weekly 3−month US Treasury bill rates.

List of Tables

1 Summaries of the posterior draws for the simulation example of Chapter 3 for $m = 50$. 32
2 Summaries of the posterior draws for the stochastic volatility model of Weekly 3−month US Treasury bill rates. 33

Parameter	True value	Post. mean	Post. SD	Post 2.5%	Post median	Post 97.5%
κ_x	0.2	0.244	0.038	0.173	0.243	0.321
μ_x	0.1	0.313	0.174	-0.046	0.317	0.641
κ_α	0.3	0.304	0.148	0.110	0.277	0.672
μ_α	-0.2	-0.268	0.107	-0.484	-0.267	-0.059
σ	0.4	0.406	0.130	0.202	0.390	0.705
ρ	-0.5	0.477	0.138	-0.657	-0.491	-0.066

Table 1: Summaries of the posterior draws for the simulation example of Chapter 3 for $m = 50$.

Parameter	Post. mean	Post. SD	Post 2.5%	Post median	Post 97.5%
θ_0	0.130	0.238	-0.347	0.132	0.589
θ_1	0.013	0.057	-0.096	0.013	0.125
κ	2.403	0.620	1.319	2.360	3.745
μ	-3.966	0.211	-4.384	-3.964	-3.547
σ	2.764	0.311	2.199	2.750	3.420

Table 2: Summaries of the posterior draws for the stochastic volatility model of Weekly 3−month US Treasury bill rates.