

# Bayesian Inference of Discretely Sampled Markov Processes with Closed-Form Likelihood Expansions

Osnat Stramer<sup>1</sup>   Matthew Bognar<sup>1</sup>   Paul Schneider<sup>2</sup>

<sup>1</sup>Department of Statistics and Actuarial Science, University of Iowa

<sup>2</sup>Finance Group, Warwick Business School, Coventry, UK

March, 2009

# Diffusions

- Diffusions are used to model continuous time processes and are therefore commonly used in financial models
- A diffusion process is described as a solution to the stochastic differential equation (SDE)

$$dY_t = \mu(Y_t, \theta) dt + \sigma(Y_t, \theta) dW_t \quad 0 \leq t \leq T$$

where

- $Y_t$  takes values in  $\mathbb{R}^d$
- $\mu$  is the drift function of dimension  $d$
- $\sigma$  is the volatility function of dimension  $d \times d$
- $\nu = \sigma\sigma^T$  is a covariance function
- $W_t$  is a  $k$ -dimensional Brownian motion,  $k \geq d$ .

## Diffusions cont.

- We assume the process  $Y$  is observed (possibly with noise) only at discrete time points  $t_i = i\Delta$  ( $i = 0, \dots, n$ ) yielding observations  $\mathbf{x} = (x_0, \dots, x_n)$ .
- By the Markov property, if all components of  $Y$  at time  $t_i$  ( $i = 0, \dots, n$ ) are observed *without* noise, the likelihood function is

$$L(\mathbf{x}|x_0, \theta) = \prod_{i=1}^n p_Y(\Delta, x_i|x_{i-1}, \theta),$$

where the transition density  $p_Y(\Delta, x|x_0, \theta)$  is the conditional density of  $Y_{t+\Delta} = x$  given  $Y_t = x_0$ .

- In all except a few typical cases, the transition densities are not analytically available

# Closed Form Approximation-univariate diffusions

- The analytical, closed-form (CF) approximation of the unknown transition density was introduced by Aït-Sahalia (2002) for **one dimensional** diffusions.
  - Avoids a computationally intensive data augmentation scheme
  - For univariate diffusions, the closed form approximation is a non-Gaussian approximation to the transition density by means of a truncated Hermite series expansion
  - Under certain regularity conditions, the sequence converges to the true transition density as more correction terms are added for any fixed transition interval  $\Delta$  which is smaller than a threshold  $\overline{\Delta}$ .
  - This threshold  $\overline{\Delta}$  depends on both the drift  $\mu$  and covariance  $\nu$ .
  - The effectiveness of the CF approximation is well documented by Jensen and Poulsen (2002), Hurn, Jeisman and Lindsay (2007), and Stramer and Yan (2007).
  - The CF approximation should be used with caution for very volatile models or sparse data-sets (see Stramer and Yan (2007))

# Bayesian Approach Using Data Augmentation

- The main approach in the literature for Bayesian estimation in discretely observed diffusion models are classical missing-data techniques
  - Introduce latent auxiliary data to complete the missing diffusion between each adjacent pair of data points
- These algorithms can break down due to high dependence between the volatility coefficient and the missing data
  - Re-parametrization methods can help break down the high dependence: see Roberts and Stramer (2001), Kalogeropoulos et al. (2007), Chib et al. (2006), Golightly and Wilkinson (2008)

# Bayesian Approach Using the CF Approximation

- DiPietro (2001) used the CF approximation for the Bayesian analysis of one-dimensional diffusions observed without noise
- The CF approximation is a local approximation around the MLE
- It cannot always be used for values of  $\theta$  that are far from the MLE.
- As observed in DiPietro (2001) and by our simulation study, while the second order of the univariate CF approximation is always positive, it can explode to infinity for values of  $\theta$  that are far from the MLE.
- The MCMC sampler may get stuck in the tails of the posterior, typically when  $\theta$  is far from the MLE.
- In fact, the approximated likelihood does not integrate to 1, its normalizer is an intractable function of the parameters  $\theta$  and the state variables.

# Cox-Ingersoll-Ross Model

- A simple model is the CIR model

$$dY_t = \beta(\alpha - Y_t) dt + \sigma\sqrt{Y_t} dW_t$$

where

- $\alpha$  is the mean reverting level
  - $\beta$  is the speed of the process
  - $\sigma$  is the volatility parameter
- The true transition density

$$p_Y(\Delta, x|x_0, \theta)$$

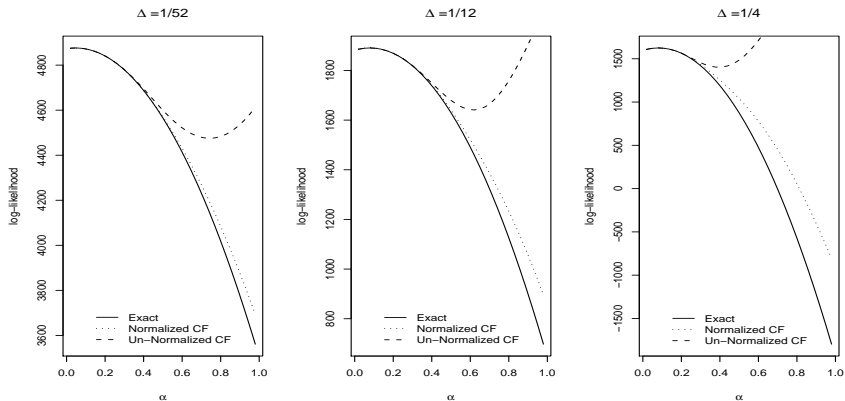
(the conditional density of  $Y_{t+\Delta} = x$  given  $Y_t = x_0$ ) is a non-central chi-squared distribution

## CIR Model - Likelihood and its approximation

- We produce a dataset for  $\Delta = 1/52$  (weekly) with  $n = 1,000$ ,  $\Delta = 1/12$  (monthly) with  $n = 500$ , and  $\Delta = 1/4$  (quarterly) with  $n = 500$ . Datasets were generated from the true CIR transition density with  $\alpha = 0.07, \beta = 0.15, \sigma = 0.07$ , and a Gaussian noise  $N(0, \tau^2)$ ,  $\tau = 0.005$ .
- The commonly analyzed FedFunds dataset yields parameter estimates close to  $\alpha = 0.07, \beta = 0.15$ , and  $\sigma = 0.07$ , so the simulated dataset mimics real-world data.
- In Figure 1 we compare the true likelihood with the estimated normalized and un-normalized likelihood, as a function of  $\alpha$ . The likelihood functions are evaluated at  $(\beta, \sigma, \tau) = (0.15, 0.07, 0.005)$



# CIR Model - Likelihood and its approximation



**Figure:** Comparison of the true likelihood with estimated normalized and un-normalized likelihood, as a function of  $\alpha$ . The likelihood functions are evaluated at  $(\beta, \sigma, \tau) = (0.15, 0.07, 0.005)$

# Dealing with the Normalizer-univariate diffusions

- DiPietro (2001) approximated the unknown normalizer explicitly for the Vasicek model and numerically for the CIR model. The normalized CF likelihood is a density and can be used to derive the posterior.
- Simulation study in DiPietro finds that, for his examples, the **normalized** CF is better behaved than the un-normalized CF and outperforms the high frequency augmentation technique introduced in Elerian et al. (2001).

# Closed Form Approximation-multivariate diffusions

- Ait-Sahalia (2008) provides an extension of the CF approximation for multidimensional diffusions. This includes a broad class of models used in the literature such as stochastic volatility models.
- The CF approximation for the log-transition density,  $\log p_X(\Delta, x|x_0, \theta)$ , is a Taylor expansion around  $\Delta = 0$  and  $x = x_0$ . Away from  $x_0$ , the error is a polynomial in  $(x - x_0)$ . We therefore assume that the CF approximation for  $p_X(\Delta, x|x_0, \theta)$  is zero outside some compact set around  $x_0$ .

## Closed Form Approximation-multivariate diffusions cont.

- Clearly, the normalization constant cannot be easily approximated, and even if it could, it would require tremendous computational effort.
- It is therefore not feasible to extend the results in DiPietro (2001) to the CF approximation for most multivariate diffusions.

## Other Work

- There are numerous instances in which the likelihood of interest contains an intractable normalizer that is a function of the parameters (Gibbs processes, for example)
  - Numerous *approximate* inferential approaches have been proposed in the literature
  - See, for example, Berthelsen and Møller (2003), Heikkinen and Penttinen (1999), and Bognar (2008)
- The first method that avoids such approximations was first proposed by Møller et al. (2006)
  - Introduce a cleverly chosen auxiliary variable into the Metropolis-Hastings (M-H) algorithm so that the normalizing constants cancel in the M-H ratio
  - A simpler and more efficient version, which inspired our work, is proposed in Murray et al. (2006)

# Closed Form Likelihood & Posterior

- The CF (approximated) likelihood is

$$\begin{aligned} L_{CF}^N(\mathbf{x}|x_0, \theta) &= \prod_{i=1}^n p_{CF}^N(\Delta, x_i|x_{i-1}, \theta) \\ &= \prod_{i=1}^n \frac{g_{CF}(\Delta, x_i|x_{i-1}, \theta)}{Z(x_{i-1}, \theta)} \end{aligned}$$

- $g_{CF}(\Delta, x|x_0, \theta)$  is the CF approximation of  $p_X(\Delta, x|x_0, \theta)$
- $p_{CF}^N(\Delta, x|x_0, \theta) = g_{CF}(\Delta, x|x_0, \theta)/Z(x_0, \theta)$  is the *normalized* CF transition density where  $Z(x_0, \theta) = \int g_{CF}(\Delta, x|x_0, \theta) dx$  is analytically intractable
- Goal: use MCMC techniques to sample from the posterior distribution

$$\pi_{CF}^N(\theta|\mathbf{x}) \propto L_{CF}^N(\mathbf{x}|x_0, \theta)\pi(\theta)$$

where  $\pi(\theta)$  is the prior distribution on  $\theta$

# Standard Metropolis-Hastings Update

- Can not use a standard M-H algorithm for updating  $\theta$  since the acceptance ratio which involves the intractable normalizing constants
- If  $\theta$  is the current value and  $\theta^*$  is the proposed value (generated from some proposal density  $q(\theta^*|\theta)$ ), the M-H algorithm has acceptance probability  $\min[1, \mathcal{R}_{MH}]$  where

$$\begin{aligned}\mathcal{R}_{MH} = & \frac{\prod_{i=1}^n g_{CF}(\Delta, x_i | x_{i-1}, \theta^*) / Z(x_{i-1}, \theta^*)}{\prod_{i=1}^n g_{CF}(\Delta, x_i | x_{i-1}, \theta) / Z(x_{i-1}, \theta)} \\ & \times \frac{\pi(\theta^*)}{\pi(\theta)} \frac{q(\theta|\theta^*)}{q(\theta^*|\theta)}\end{aligned}$$

- The intractable normalizers  $Z(\cdot, \cdot)$  do *not* cancel

# The Exchange Algorithm: Murray Update

- Murray et al. (2006) suggested a clever auxiliary variable algorithm to simulate from the posterior density
- Assumes the likelihood takes the form  $L(\mathbf{x}|\theta) = \prod_{i=1}^n g(x_i|\theta)/Z(\theta)$  where  $Z(\theta)$  is an intractable normalizer
- Update procedure
  - Generate  $\theta^*$  from some proposal density  $q(\theta^*|\theta)$
  - Generate a sample  $\mathbf{w}$  from  $L(\mathbf{w}|\theta^*) = \prod_{i=1}^n g(w_i|\theta^*)/Z(\theta^*)$
  - Accept  $\theta^*$  with probability  $\min[1, \mathcal{R}_{Mur}]$  where

$$\begin{aligned}\mathcal{R}_{Mur} &= \frac{\prod_{i=1}^n g(x_i|\theta^*)/Z(\theta^*)}{\prod_{i=1}^n g(x_i|\theta)/Z(\theta)} \\ &\quad \times \frac{\pi(\theta^*)}{\pi(\theta)} \frac{q(\theta|\theta^*)}{q(\theta^*|\theta)} \frac{\prod_{i=1}^n g(w_i|\theta)/Z(\theta)}{\prod_{i=1}^n g(w_i|\theta^*)/Z(\theta^*)}\end{aligned}$$

- The intractable  $Z$ 's cancel



## Murray Update Continued

- The Murray algorithm is not applicable in our situation
- Murray procedure:
  1. Generate  $\theta^*$  from some proposal density  $q(\theta^*|\theta)$
  2. Generate a sample from

$$L_{CF}^N(\mathbf{w}|w_0, \theta^*) = \prod_{i=1}^n \frac{g_{CF}(\Delta, w_i|w_{i-1}, \theta^*)}{Z(w_{i-1}, \theta^*)}$$

3. Accept  $\theta^*$  with probability  $\min[1, \mathcal{R}_{Mur}]$  where

$$\begin{aligned}\mathcal{R}_{Mur} &= \frac{\prod_{i=1}^n g_{CF}(\Delta, x_i|x_{i-1}, \theta^*)/Z(x_{i-1}, \theta^*)}{\prod_{i=1}^n g_{CF}(\Delta, x_i|x_{i-1}, \theta)/Z(x_{i-1}, \theta)} \\ &\quad \times \frac{\pi(\theta^*)}{\pi(\theta)} \frac{q(\theta|\theta^*)}{q(\theta^*|\theta)} \\ &\quad \times \frac{\prod_{i=1}^n g_{CF}(\Delta, w_i|w_{i-1}, \theta)/Z(w_{i-1}, \theta)}{\prod_{i=1}^n g_{CF}(\Delta, w_i|w_{i-1}, \theta^*)/Z(w_{i-1}, \theta^*)}\end{aligned}$$

- The intractable normalizers  $Z(\cdot, \cdot)$  do *not* cancel

## Our Modified Exchange Algorithm

0. Choose a starting value  $\theta^{(t)}$  where  $t = 0$ .
1. Propose a new value  $\theta^*$  from some proposal density  $q(\theta^*|\theta^{(t)})$ ; the proposal density may update one randomly chosen component of  $\theta^{(t)}$  at a time, or may attempt to update multiple components of  $\theta^{(t)}$  simultaneously
2. Generate  $\mathbf{w} = (w_1, \dots, w_n)$  from

$$p(\mathbf{w}|w_0, \mathbf{x}, \theta^*) = \prod_{i=1}^n \frac{g_{CF}(\Delta, w_i|x_{i-1}, \theta^*)}{Z(x_{i-1}, \theta^*)},$$

3. Accept  $\theta^*$  (i.e. set  $\theta^{(t+1)} = \theta^*$ ) with probability  $\min[1, \mathcal{R}_\theta]$  where

$$\begin{aligned} \mathcal{R}_\theta &= \frac{\prod_{i=1}^n g_{CF}(\Delta, x_i|x_{i-1}, \theta^*)/Z(x_{i-1}, \theta^*)}{\prod_{i=1}^n g_{CF}(\Delta, x_i|x_{i-1}, \theta^{(t)})/Z(x_{i-1}, \theta^{(t)})} \frac{\pi(\theta^*)}{\pi(\theta^{(t)})} \frac{q(\theta^{(t)}|\theta^*)}{q(\theta^*|\theta^{(t)})} \\ &\quad \times \frac{\prod_{i=1}^n g_{CF}(\Delta, w_i|x_{i-1}, \theta^*)/Z(x_{i-1}, \theta^{(t)})}{\prod_{i=1}^n g_{CF}(\Delta, w_i|x_{i-1}, \theta^{(t)})/Z(x_{i-1}, \theta^*)} \\ &= \frac{\prod_{i=1}^n g_{CF}(\Delta, x_i|x_{i-1}, \theta^*)}{\prod_{i=1}^n g_{CF}(\Delta, x_i|x_{i-1}, \theta^{(t)})} \frac{\pi(\theta^*)}{\pi(\theta^{(t)})} \frac{q(\theta^{(t)}|\theta^*)}{q(\theta^*|\theta^{(t)})} \\ &\quad \times \frac{\prod_{i=1}^n g_{CF}(\Delta, w_i|x_{i-1}, \theta^*)}{\prod_{i=1}^n g_{CF}(\Delta, w_i|x_{i-1}, \theta^{(t)})} \end{aligned}$$

# CIR Model - Data Observed With Error

- 

$$dY_t = \beta(\alpha - Y_t)dt + \sigma\sqrt{Y_t}dW_t$$

- Datasets were generated from the true CIR transition density with  $\alpha = 0.07$ ,  $\beta = 0.15$ ,  $\sigma = 0.07$ ,  $\Delta = 1/52$  (weekly),  $\Delta = 1/12$  (monthly),  $\Delta = 1/4$  (quarterly), and a Gaussian noise  $N(0, \tau^2)$ ,  $\tau = 0.005$ .
- The prior was specified similar to DiPietro (2001):

$$\pi(\theta) = \pi(\alpha, \beta, \sigma, \tau) = I_{(0,1)}(\alpha) \cdot I_{(0,\infty)}(\beta) \cdot \sigma^{-1} I_{(0,\infty)}(\sigma) \cdot I_{(0,1)}(\tau)$$

# CIR-Mixing of Different Algorithms

- To determine the likelihood of a chain getting stuck, we produce 10 datasets each for  $\Delta = 1/52(n=2000)$ ,  $\Delta = 1/12(n = 500)$ ,  $\Delta = 1/4(n = 500)$ , respectively.
- Each dataset was analyzed by 10 different chains of length 10,000 (not including a 1,000 iteration burn-in for each chain). Each chain started near the MLE using a  $MV\ t(r)$ ,  $r = 2, 4, 40$ , proposal density for a  $\alpha/\beta$ -move.
- The analysis was run using the
  - exact transition densities
  - normalized CF
  - un-normalized CF

## CIR-The Likelihood of a Chain Getting Stuck

- The proportion of chains that became stuck (out of 100 chains total) for the **un-normalized CF** are summarized in the following table.
- The exact and normalized CF analyses did not become stuck for any combination of proposal density and  $\Delta$ .

	$\Delta = 1/52$	$\Delta = 1/12$	$\Delta = 1/4$
$MV\ t(2)$	0.53	0.96	0.99
$MV\ t(4)$	0	0.04	0.04
$MV\ t(40)$	0	0.01	0

**Table:** Proportion of chains that became stuck (un-normalized CF)

## CIR: Monthly Data Recorded with Error

- A dataset was generated from true CIR transition density with  $\Delta = 1/12, n = 500, \alpha = 0.07, \beta = 0.15, \sigma = 0.07, \tau = 0.005$
- The commonly analyzed FedFunds dataset yields parameter estimates close to  $\alpha = 0.07, \beta = 0.15$ , and  $\sigma = 0.07$ , so the simulated dataset mimics real-world data.
- The joint  $\alpha/\beta$ -move used an independence proposal  $(\alpha^*, \beta^*) \leftarrow MV\ t(4)$  where the mean vector and covariance matrix were chosen via a simple regression.
- The random-walk  $\sigma$ -move generated candidate values according to  $\sigma^* \leftarrow Unif(\sigma^{(t)} - 0.01, \sigma^{(t)} + 0.01)$ ,
- A  $\tau$ -move used  $\tau^* \leftarrow Unif(\tau^{(t)} - 0.001, \tau^{(t)} + 0.001)$ .

# Monthly Data, With Error: Comparison of Algorithms

- The aforementioned dataset was analyzed via the three algorithms under study:
  - the exact (non-central chi-square) CIR transition density (and likelihood) in a standard M-H sampler,
  - the un-normalized CF likelihood in a standard M-H sampler, and
  - the normalized CF likelihood.
- For each algorithm, the successful (i.e. non-stuck) chains were combined,
  - posterior mean and quantiles were computed
  - acceptance rate was determined for each move-type  $(\alpha/\beta, \sigma, \tau)$
  - A Kolmogorov-Smirnov test was run to determine if a significant difference existed between the output from the exact CIR chain and the CF chains.
  - the effective sample size (ESS) was also computed; i.e. the number of independent samples that would carry the same amount of information as the available correlated samples.

## CIR: Monthly Data, With Error

		Exact	Norm CF	UnNorm CF
$\alpha x$	Mean	0.09960	0.10597	0.10070
	0.99	0.73601	0.82556	0.80446
	0.95	0.34242	0.39288	0.33734
	0.75	0.08560	0.08578	0.08373
	0.50	0.06049	0.06053	0.06021
	0.25	0.04930	0.04946	0.04915
	0.05	0.03719	0.03751	0.03698
	0.01	0.02963	0.02923	0.02935
	AR	0.53664	0.30654	0.54208
	K-S		0.39273	0.92282
	ESS	1514.6	778.2	881.7

**Table:** CIR: Marginal posterior summary statistics for the mean reverting level parameter  $\alpha$ . Dataset generated with  $\alpha = 0.07$



## CIR: Monthly Data, With Error

		Exact	Norm CF	UnNorm CF
$\beta x$	Mean	0.15951	0.16128	0.16229
	0.99	0.41951	0.43084	0.42153
	0.95	0.33983	0.34742	0.34189
	0.75	0.22742	0.22985	0.23043
	0.50	0.15113	0.15172	0.15378
	0.25	0.07908	0.08197	0.08212
	0.05	0.01375	0.01157	0.01389
	0.01	0.00662	0.00487	0.00550
	AR	0.53664	0.30654	0.54208
	K-S		0.22021	0.96395
	ESS	4750.5	2427.8	3985.4

**Table:** CIR: Marginal posterior summary statistics for the speed parameter  $\beta$ .  
Dataset generated with  $\beta = 0.15$

## CIR: Monthly Data, With Error

		Exact	Norm CF	UnNorm CF
$\sigma x$	Mean	0.07599	0.07618	0.07605
	0.99	0.08905	0.09129	0.08877
	0.95	0.08494	0.08584	0.08474
	0.75	0.07949	0.07981	0.07945
	0.50	0.07585	0.07596	0.07593
	0.25	0.07241	0.07225	0.07242
	0.05	0.06737	0.06716	0.06796
	0.01	0.06457	0.06395	0.06481
	AR	0.37528	0.27328	0.37500
	K-S		0.71124	0.46532
	ESS	840.5	383.3	890.3

**Table:** CIR: Marginal posterior summary statistics for the mean reverting level parameter  $\sigma$ . Dataset generated with  $\sigma = 0.07$

## CIR: Monthly Data, With Error

$\tau   \mathbf{x}$	Mean	0.00493	0.00491	0.00491
	0.99	0.00558	0.00560	0.00557
	0.95	0.00537	0.00538	0.00539
	0.75	0.00511	0.00510	0.00510
	0.50	0.00493	0.00490	0.00491
	0.25	0.00474	0.00471	0.00471
	0.05	0.00447	0.00445	0.00446
	0.01	0.00428	0.00424	0.00430
	AR	0.25080	0.24780	0.24876
	K-S		0.71124	0.54414
	ESS	1021.8	862.2	1053.1

**Table:** CIR: Marginal posterior summary statistics for the variance of the noise parameter  $\tau$ . Dataset generated with  $\tau = 0.005$

# Timing

- All of the computer code was written in *R* and *C++*, with *C++* handling many of the computational routines (*C++* was called from within the main *R* program).
- The time to execute 11,000 iterations (including a 1,000 iteration burn-in) on an Intel Core 2 Duo 2.0 GHz processor was
  - 114 seconds for the exact analysis (using the *C*-based non-central  $\chi^2$  function in *R*),
  - 141 seconds for the normalized CF analysis,
  - 69 seconds for the un-normalized CF analysis.

# Summary of Results-CIR Models

- The un-normalized CF analysis yields correct posterior inferences when the MCMC chain mixes properly (i.e. when the chain does not become stuck).
- However, normalizing dramatically decreases the tendency of the MCMC chain becoming stuck
- Therefore, normalizing increases the applicability of the CF approximation in MCMC samplers.

# The Heston's Model

- $Y_t$  is the log-price for the stock  $S_t$  and  $V_t$  is the volatility process.

$$dY_t = (\mu - \frac{1}{2}V_t)dt + \rho\sqrt{V_t}dW_t + \sqrt{1 - \rho^2}\sqrt{V_t}dB_t \quad (1)$$

$$dV_t = \beta(\alpha - V_t)dt + \sigma\sqrt{V_t}dW_t, \quad (2)$$

where  $B$  and  $W$  are independent standard B.M's, and the correlation between  $dY_t$  and  $dV_t$  is  $\rho$ .

- To keep the simulation study simple we make the assumption of zero risk premia such that  $W_t = W_t^{\mathbb{Q}}$  and  $B_t = B_t^{\mathbb{Q}}$  and no adjustments in the drift are necessary.
- *Instantaneous* stochastic variance is latent, even though we use the VIX implied volatility index published by the CBOE.
- To account for the stochastic nature and mean reversion of index variance, we use the fact that for short-maturity at-the-money options the Black-Scholes formula is approximately linear in volatility. (Aït-Sahalia and Kimmel (2007))

# The Heston's Model- Simulation Study

- To exploit available market information we will consider the joint time series of log stock price and implied variance with and without observation error:

$$[Y_t, IV_t] \quad \text{no observation error} \quad (3)$$

$$[Y_t, IV_t + \varepsilon_t] \quad \text{observation error,} \quad (4)$$

where  $\varepsilon_t \sim N(0, \sigma_\varepsilon)$ ,  $t = t_0, \dots, t_N$ .

- We produce ten datasets each for  $\Delta = 1/52, 1/12, 1/4$ , using the parametrization  $\beta = 3, \alpha = 0.1, \mu = 0.05, \rho = -0.8, \sigma = 0.25$  and  $\sigma_\varepsilon = 0.001$ . (Ait-Sahalia and Kimmel (2007))
- We collect 11,000 draws from the posterior density of the parameters for all datasets, where we discard the first 1,000 to account for the burn-in period.

## Summary- Without Error

- An investigation of the chains for the 10 datasets observed **without** error reveals excellent performance of both normalized and un-normalized CF for the weekly data.
- For the monthly data, the sampler using un-normalized CF explodes once.
- All chains using un-normalized CF diverge for quarterly data, while the chains using normalized CF all converge. CF estimates for quarterly data exhibit a tendency to underestimate the volatility function  $\sigma$  and the speed-of-mean-reversion coefficient  $\beta$ .



## Summary- With Error

- Estimations of the systems observed **with** error show that all chains using un-normalized CF converge for weekly data
- The sampler explodes for all monthly and quarterly spaced datasets.
- All chains using the true density and the CF together with our normalization algorithm converge.

# The Heston's model- Noisy Monthly Data

- We investigate posterior distributions of parameters for noisy monthly data using 101,000 draws, where the first 1,000 are discarded accounting for the burn in.
- The tables show only results for the normalized CF chain and the chain using exact density.
- The quantiles in Tables reveal that there is strong agreement between the posterior distributions of the parameters.
- There is higher autocorrelation in the MCMC chains than in the univariate experiment however. As K-S statistics are very sensitive to autocorrelation they are not reported.

		Exact	Norm CF
$\alpha x$	Mean	0.1007254	0.1033851
	0.99	0.10797464	0.11065052
	0.95	0.10559627	0.10845317
	0.75	0.10304525	0.10568659
	0.50	0.10096617	0.10340112
	0.25	0.09856890	0.10120600
	0.05	0.09523693	0.09812024
	0.01	0.09323680	0.09589357
	AR	0.700717	0.4924349
$\beta x$	Mean	3.182851	3.059048
	0.99	3.785301	3.590590
	0.95	3.641120	3.424803
	0.75	3.372505	3.202920
	0.50	3.171415	3.063390
	0.25	3.007350	2.929160
	0.05	2.730879	2.713008
	0.01	2.562318	2.506056
	AR	0.700717	0.4924349

**Table:** Heston Model: Marginal posterior summary statistics for the mean reverting level  $\alpha$  and speed  $\beta$ .

		Exact	Norm CF
$\mu x$	Mean	-0.002620006	-0.03691722
	0.99	0.10613650	0.07109900
	0.95	0.07533960	0.03918260
	0.75	0.02863925	-0.00655050
	0.50	-0.00410637	-0.03728695
	0.25	-0.03435290	-0.06749430
	0.05	-0.07718900	-0.11102300
	0.01	-0.10896900	-0.14246310
	AR	0.6098661	0.4857749
$\sigma x$	Mean	0.2489901	0.2459580
	0.99	0.269207	0.2657930
	0.95	0.263040	0.2599930
	0.75	0.254661	0.2521470
	0.50	0.248883	0.2468180
	0.25	0.243105	0.2410190
	0.05	0.235256	0.2328630
	0.01	0.230043	0.2244595
	AR	0.1414214	0.1012010

**Table:** Heston Model: Marginal posterior summary statistics for the constant in the stock drift  $\mu$  and volatility of variance parameter  $\sigma$ .

		Exact	Norm CF
$\rho x$	Mean	-0.8057692	-0.8095715
	0.99	-0.769370	-0.770569
	0.95	-0.780608	-0.782672
	0.75	-0.795758	-0.799165
	0.50	-0.805969	-0.809946
	0.25	-0.815921	-0.820241
	0.05	-0.830071	-0.834857
	0.01	-0.841406	-0.845533
	AR	0.5697357	0.4523645
$\tau x$	Mean	0.001358199	0.00235106
	0.99	0.0037247201	0.0045968411
	0.95	0.0030240200	0.0038040055
	0.75	0.0019355600	0.0030373525
	0.50	0.0012030400	0.0024140650
	0.25	0.0006016370	0.0016619775
	0.05	0.0002210365	0.0003748274
	0.01	0.0001769038	0.0001773979
	AR	1	1

**Table:** Heston Model: Marginal posterior summary statistics for the correlation parameter  $\rho$  and noise parameter  $\tau$ .

# Timing

Computation times are very similar to the CIR model, except for estimations using the true density, where 11,000 draws take well over 24 hours due to heavy use of adaptive numeric integration and the use of complex-valued special functions from the density representation of Lamoureux et. al. (2005).

# Conclusions

- The CF transition density is a powerful tool for the analysis of a broad class of jump-diffusion models.
- The intractable normalizer in the CF likelihood is close to 1 when near the MLE, but can markedly differ when far in the tails of the posterior, hindering Bayesian analysis.
- We have provided a Bayesian approach for inference using the CF transition density. Our algorithms are based on the exchange algorithm, proposed in Murray (2006) and thus avoid computation of the normalizing constant.
- Our simulation study shows that our algorithms using the normalizing approach greatly increase the stability and mixing behavior of the MCMC sampler. Furthermore, they are quite efficient, mix well, and are relatively easy to implement.