Inference of the Structural Credit Risk Model using MLE

Yuxi Li, Li Cheng and Dale Schuurmans

***Abstract*— Credit risk analysis is not only an important research topic in finance, but also of interest in everyday life. Unfortunately, the non-linear nature of the widely accepted Black-Scholes option price model, which sits at the very heart of the structural credit risk model, causes great difficulty when inferring the latent asset value sequence from observed data. The main contribution of this paper is to address this problem by pursuing maximum likelihood state estimation (MLE) instead of the usual particle filtering approach. Experiments demonstrate the competitiveness of the proposed MLE approach: it achieves a much lower inference error and a much lower running time than particle filtering methods. This work has merit for the general problem of inferring latent values for probabilistic time-series.**

I. Introduction

Credit risk is not only an important research topic in finance, it is also relevant to everyday life. Take the sub-prime mortgage crisis as an example where a failure in credit risk assessment has played a critical role in precipitating a world wide financial crisis that has profoundly affected the global economy, particularly that of the US, since it emerged in the summer of 2007. It is of paramount importance to have a good understanding of the credit risk associated with various assets in financial markets.

The structural credit risk model founded by Merton [8] provides a tool to understand the risk entailed by corporate equities. Before giving an introduction of the structural credit risk model, we need to introduce the concepts of financial derivatives, options and option pricing. A financial derivative is an instrument whose price is determined by another asset. Financial derivatives are critical for risk management since they enable the handling of risk exposure to stock prices, interest rates or exchange rates. Options are a popular type of financial derivatives. An option specifies the maturity time T and the strike price K. A European option can only be exercised at time T, but the owner may choose not to exercise it. The payoff to the owner of a European call option is $\max(S_T - K, 0)$, where S_T is the price of the underlying asset at time T. Black and Scholes [2] developed the option pricing formula. Merton [8] treated the equity of a corporate as an option on its asset, and founded the structural approach to credit risk modeling. Recently, Duan and Fulop [4] have refined the structural credit risk model by adding a noise term to the observed equity price to account for the fact that the observed equity prices may have been contaminated by trading noises, such that the proposed model in [4] allows the observed equity price to diverge from the equilibrium values due to microstructure noises [7].

Yuxi Li is with the Department of Computing Science, the University of Alberta, Edmonton, Alberta, Canada, T6G 2E8 (email: yuxi@cs.ualberta.ca).

Li Cheng is with the Toyota Technological Institute at Chicago, 1427 East 60th Street, Chicago, Illinois 60637, USA (email:licheng@tti-c.org).

Dale Schuurmans is with the Department of Computing Science, the University of Alberta, Edmonton, Alberta, Canada, T6G 2E8 (email: dale@cs.ualberta.ca).

We are interested in inferring the latent asset values from the observed equity prices. Usually, the latent asset states are inferred using particle filtering, which will be introduced in Section II-B. In this paper, we propose to use an offline maximum likelihood estimate (MLE) approach to infer the latent asset value in the structural credit risk model. The offline strategy we introduce is intended to provide better results than an online counterpart such as particle filtering, because each iteration of an offline algorithm guarantees an improvement, while online updates only promise improvement on average over a number of iterations. We compare the performance of our proposed approach with various particle filtering methods, including sampling importance resampling (SIR) [5], auxiliary particle filters (APF) [11] and regularised particle filters (RPF) [9]. Our experimental results validate the claim that the MLE approach achieves a much lower inference error and running time than current particle filtering approaches.

The remainder of this paper is organized as follows. First we introduce the dynamic state space model and discuss two main approaches to inferring the latent state sequence, namely, MLE and particle filtering. After introducing the structural credit risk model, we demonstrate how MLE and particle filtering can, respectively, be used to infer the latent asset value from observed price data. Next we conduct a performance study that compares and contrasts the two approaches before concluding the paper.

II. Dynamic state space model

In [8], the equity of a corporation is treated as an option on its assets, which leads to the structural approach to credit risk modeling. The widely-accepted Black-Scholes option pricing formula [2] relates the latent asset value to the observed equity value through a non-linear equation. Together these lead to the well-known non-linear statistical model used to address the structural credit risk analysis task. We first introduce a generic dynamical system model in this section and then specialize it to the structural credit risk model in Section III-A.

The generic dynamical system model consists of the state process $\{\mathbf{x}_t, t \in \mathbf{N}\}$ and observation process $\{\mathbf{y}_t, t \in \mathbf{N}\}$

given by:

$$\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}, \mathbf{v}_{t-1}), \quad (1)$$

$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{x}_t, \mathbf{n}_t), \quad (2)$$

respectively, where $\mathbf{f}_t : \Re^{n_x} \times \Re^{n_v} \rightarrow \Re^{n_x}$ and $\mathbf{h}_t : \Re^{n_x} \times \Re^{n_n} \rightarrow \Re^{n_y}$ are possibly non-linear functions, and $\{\mathbf{v}_t, t \in \mathbf{N}\}$ and $\{\mathbf{n}_t, t \in \mathbf{N}\}$ are i.i.d. state and observation noise sequences, with n_x, n_v, n_y and n_n being respectively the dimensions of state, state noise, observation and observation noise. The noise distributions may be non-Gaussian.

The state transitions $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ follow a Markov process; while the observation model $p(\mathbf{y}_t|\mathbf{x}_t)$ implies that the observations are assumed to be conditionally independent given the states. The prior distribution at $t = 0$ is denoted by $p(\mathbf{x}_0)$. In what follows, we describe in detail our MLE approach, and to keep our paper self-contained, we also provide an account of variants of particle filtering algorithms that will be used later in the experimental section. These variants include sampling importance resampling, the auxiliary particle filter and the regularized particle filter.

A. MLE

The penalized negative log likelihood function and its partial derivatives for the dynamic state space model (1, 2) are defined as

$$\begin{aligned} NL(\mathbf{x}) &= -\log\{\prod_{t=1}^{t=T} p(\mathbf{y}_t|\mathbf{x}_t) \prod_{t=2}^{t=T} p(\mathbf{x}_t|\mathbf{x}_{t-1})\} \\ &= -\sum_{t=1}^{t=T} \log\, p(\mathbf{y}_t|\mathbf{x}_t) - \sum_{t=2}^{t=T} \log\, p(\mathbf{x}_t|\mathbf{x}_{t-1}) \quad (3) \end{aligned}$$

$$PNL(\mathbf{x}) = \frac{1}{2}\lambda||\mathbf{x}||^2 + NL(\mathbf{x}) \quad (4)$$

$$\frac{\partial PNL(\mathbf{x})}{\partial \mathbf{x}} = \lambda \mathbf{x} + \frac{\partial NL(\mathbf{x})}{\partial \mathbf{x}} \quad (5)$$

where T is the number of observations.

To infer the latent values $\mathbf{x}$, instead of using the popular particle filtering approach, we propose to directly maximize the penalized negative log likelihood function $PNL(\mathbf{x})$. This MLE strategy possesses in fact quite a few theoretical merits. First it provides asymptotically unbiased estimate of the optimal time-series model in hindsight. In other words, the bias goes to zero as the number of examples increases to infinity. Moreover, it asymptotically achieves the Cramer-Rao lower bound, which guarantees the variance of the estimator reaches the lowest estimation error possible in the mean square sense. Numerically our approach computes a maximum likelihood latent state sequence by using a quasi-Newton method. We use a quasi-Newton approach because it only requires the approximate rather than the exact Hessian matrix, which would be prohibitively expensive in practice. In particular, we employ a limited memory BFGS (L-BFGS) [10] solver, which requires one to evaluate only the PNL function values and gradients, and is both memory and run-time efficient in our scenario.

B. Particle filters

Particle filters allow one to estimate latent variables in non-linear, non-Gaussian dynamical systems [3], [1], like the dynamic state space model (1, 2) introduced above. A filter usually estimates the posterior distribution $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ and the filtering density $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, where $\mathbf{x}_{1:t} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_t\}$ and $\mathbf{y}_{1:t} = \{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_t\}$. Next we introduce the essential components of particle filtering, sequential importance sampling and resampling, and also discuss auxiliary particle filters and regularized particle filters.

Particles refer to a set of N weighted samples $\{\mathbf{x}^i_{0:t}; i = 1, 2, ..., N\}$ drawn from a proposal distribution $q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$. Particle filters can approximate the posterior density $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ using these samples. With such particles, intractable integration problems, such as computing expectations, may be reduced to summations. Its convergence is justified by the strong law of large numbers.

Sequential Importance Sampling. It is desirable to compute a sequence estimate of the posterior distribution without modifying the previously simulated states. To this end, we can use the following proposal distribution and the weights:

$$q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = q(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}) q(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t}) \quad (6)$$

$$w^i_t = w^i_{t-1} \frac{p(\mathbf{y}_t|\mathbf{x}^i_t) p(\mathbf{x}^i_t|\mathbf{x}^i_{t-1})}{q(\mathbf{x}^i_t|\mathbf{x}^i_{0:t-1}, \mathbf{y}_{1:t})} \quad (7)$$

Resampling. A common problem for sequential importance sampling is particle degeneracy, where only one particle will have significant weight, after a few iterations. Resampling is an approach to addressing the degeneracy problem.

The effective sample size N_{eff} is defined as a measure of the degeneracy [6]:

$$N_{\mathit{eff}} = \frac{N}{1 + var(w^{*i}_t)} \quad (8)$$

where $w^{*i}_t = p(\mathbf{x}^i_t|\mathbf{y}_{1:t}) / q(\mathbf{x}^i_t|\mathbf{x}^i_{t-1}, \mathbf{y}_{1:t})$, which usually cannot be evaluated exactly. An approximation $\widehat{N_{\mathit{eff}}}$ of N_{eff} can be obtained by:

$$\widehat{N_{\mathit{eff}}} = \frac{N}{\sum_{i=1}^{N} (w^i_t)^2} \quad (9)$$

where w^i_t is the normalized weight obtained using (7).

Resampling works by eliminating particles with low weights and duplicate particles with high weights. It resamples from the following discrete approximation of $p(\mathbf{x}_t|\mathbf{y}_{1:t})$:

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w^i_t \delta(\mathbf{x}_t - \mathbf{x}^i_t) \quad (10)$$

where $\delta(\cdot)$ is the Dirac delta measure. Resampling maps the weighted measure $\{\tilde{\mathbf{x}}^i_t,\ \tilde{w}^i_t\}$ to an unweighted measure $\{\tilde{\mathbf{x}}^i_t, N^{-1}\}$. Several resampling algorithms are available, e.g., the residual, stratified and systematic resampling schemes [1].

Sampling importance resampling. Sampling importance resampling in [5] uses the prior as the proposal distribution.

As a result, we only need the likelihood for computing weights:

$$q(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t}) = p(\mathbf{x}_t|\mathbf{x}_{t-1}), \quad (11)$$

$$w_t^i = w_{t-1}^i p(\mathbf{y}_t|\mathbf{x}_t^i). \quad (12)$$

This proposal is widely used for its simplicity. It conducts resampling at every time step, so that it may lead to a quick loss of particle diversity. Moreover, it does not consider the most recent observation.

Auxiliary particle filters. The auxiliary particle filter [11] takes advantage of the most recent observation. It introduces an important density $q(\mathbf{x}_t, i|\mathbf{y}_{0:t})$ to sample the pair $(\mathbf{x}_t^j, i^j)$, where i^j refers to the index of the particle at time step $t-1$ [1]. One can derive,

$$p(\mathbf{x}_t, i|\mathbf{y}_{1:t}) \propto p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1}^i)w_{t-1}^i; \quad (13)$$

and define,

$$q(\mathbf{x}_t, i|\mathbf{y}_{1:t}) \propto p(\mathbf{y}_t|\mu_t^i)p(\mathbf{x}_t|\mathbf{x}_{t-1}^i)w_{t-1}^i, \quad (14)$$

where μ_t^i is some characterization of $\mathbf{x}_t$, given $\mathbf{x}_{t-1}^i$. For example, it can be a sample $\mu_t^i = p(\mathbf{x}_t|\mathbf{x}_{t-1})$. Therefore, we have,

$$w_t^j = w_{t-1}^{i^j} \frac{p(\mathbf{y}_t|\mathbf{x}_t^j)p(\mathbf{x}_t^j|\mathbf{x}_{t-1}^{i^j})}{q(\mathbf{x}_t^j, i^j|\mathbf{y}_{1:t})} = \frac{p(\mathbf{y}_t|\mathbf{x}_t^j)}{p(\mathbf{y}_t|\mu_t^{i^j})} \quad (15)$$

Regularized particle filter. Resampling is a popular approach to reduce the degeneracy problem. However, the diversity among particles might be lost, which can finally lead to the outcome of "particle collapse" where all particles occupy a single point in the state space. Regularized particle filters [9] were proposed as a potential solution.

The RPF differs from SIR in the resampling step. RPF resamples from a continuous approximation of the posterior density $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, given in (16) below, while SIR resamples from the discrete approximation (10):

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w_t^i K_h(\mathbf{x}_t - \mathbf{x}_t^i) \quad (16)$$

where

$$K_h(\mathbf{x}) = \frac{1}{h^{n_x}} K(\frac{\mathbf{x}}{h}) \quad (17)$$

is the rescaled Kernel density, h is the Kernel bandwidth, n_x is the dimension of the state $\mathbf{x}$, and w_t^i are normalized weights. The Kernel density is a symmetric probability density function such that,

$$\int K(\mathbf{x})d\mathbf{x} = 1, \int \mathbf{x}K(\mathbf{x})d\mathbf{x} = 0, \int \|\mathbf{x}\|^2 K(\mathbf{x})d\mathbf{x} < \infty$$

The Kernel and bandwidth are chosen to minimize the mean integrated square error between the true posterior density and the regularized approximation (16). When all the samples having the same weights, the optimal choice of kernel is the Epanechnikov kernel:

$$K_{opt} = \begin{cases} \frac{n_x+2}{2c_{n_x}}(1 - \|\mathbf{x}\|^2), & \text{if } \|\mathbf{x}\| < 1 \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

where c_{n_x} is the volume of the unit hypersphere in $\Re^{n_x}$. The optimal bandwidth h for a Gaussian density with unit covariance matrix is:

$$h_{opt} = [8c_{n_x}^{-1}(n_x + 4)(2\sqrt{\pi})^{n_x} N]^{1/(n_x+4)} \quad (19)$$

The optimal solution (18) and (19) applies to the special case where particles have equal weights and the underlying density is Gaussian. However, these results can achieve sub-optimal performance for general cases. As our experiments will show, RPF with (18) and (19) achieves rather good results.

III. Structural Credit Risk Model

In the following, we present the structural credit risk model founded by Merton [8], based on the work on option pricing [2], and its recent extension by Duan and Fulop [4] that accounts for trading noises. We also briefly discussion credit risk applications.

A. Structural credit risk model

Merton founded the structural approach to credit risk modeling by treating the equity of a corporate as an option on its assets [8]. Suppose the value of the corporate at time t, V_t, follows a geometric Brownian Motion, with drift and volatility parameters μ and σ:

$$dV_t = \mu V_t dt + \sigma V_t dW_t \quad (20)$$

The corporate finances its assets with two claims: equity and a zero-coupon bond maturing at time T with a principle payment of F. At time T, equity holders repay the bond if the asset is greater than the principle, $V_T \geq F$. Otherwise, equity holders declare bankruptcy and the bond holders own the corporate. Therefore, the equity holders receive at time T, $\hat{S}_t = \max(V_t - F, 0)$. This shows that the equity can be regarded as a call option on the asset. Thus we can use Black-Scholes formula [2] to determine its price:

$$\hat{S}_t = V_t\Phi(d_t) - Fe^{-r(T-t)}\Phi(d_t - \sigma\sqrt{T-t}) \quad (21)$$

$$d_t = \frac{\ln(V_t/F) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \quad (22)$$

It is usually more accurate to simulate $\ln V_t$ in practice. Using Itô's lemma, the process followed by $\ln V_t$ is:

$$d\ln \mathrm{V}_t = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t. \quad (23)$$

We can obtain its discrete-time form:

$$\ln V_t = \ln V_{t-1} + (\mu - \frac{\sigma^2}{2})h + \sigma\sqrt{h}\varepsilon_t \quad (24)$$

where h is the time interval and ε_t are i.i.d. standard normal variables.

Noises exist in the financial market, as reported in the microstructure literature, for example, see [7]. Thus the relationship between the latent asset value and the observed equity value predicted by (21) is contaminated by noises.

Duan and Fulop [4] suggested a multiplicative noise structure to express the logarithmic equity value as:

$$\ln S_t = \ln \hat{S}_t + \delta \nu_t \tag{25}$$

where $\hat{S}_t$ is given in (21) and ν_t are i.i.d. standard normal variables.

B. Credit risk applications

It is desirable to know the credit spread of a corporate bond over the Treasury rate and the default probability of the corporate. The latent asset values inferred are essential for these credit risk applications. For example, the default probability is the probability that the asset value at time T is less than the face value of the bond F. With V_t, μ and σ, the formula to compute the default probability is:

$$P(V_t, \mu, \sigma) = \Phi(\frac{\ln(F/V_t) - (\mu - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}) \tag{26}$$

IV. Inference of the structural credit risk model

In the following, we describe how to infer the latent asset value with particle filtering or MLE, assuming model parameters are given.

A. Filtering out asset values

We have the state process by (24) and the observation process (25), given μ, σ and δ. The following gives the sketch of the algorithm for particle filtering. It is a modified sampling importance resampling, where resampling may not take place in every step. N is the number of particles. N_T is a predefined resampling threshold.

For time step t:

1) For $i = 1 : N$
   - Draw $\ln V_t^i \sim p(\ln V_t | \ln V_{t-1}^i)$. Calculate $w_t^i = w_{t-1}^i p(\ln S_t | \ln V_t^i)$.
2) Normalize weights: $w_t^i = w_t^i / \sum_{j=1}^N w_t^j$.
3) Resampling
   - Calculate $\widehat{N_{eff}}$ using (9). If $\widehat{N_{eff}} < N_T$, resampling. Set $w_t^i = 1/N$.

B. Inference of asset values with MLE

For the state process (24) and the observation process (25), the negative log likelihood function $NL(\ln V)$ follows:

$$\begin{aligned} & NL(\ln V) \\ = & -\sum_{t=1}^{T} \log\, p(\ln S_t | \ln V_t, \sigma, \delta) - \sum_{t=2}^{T} \log\, p(\ln S_t | \ln V_t, \mu, \sigma) \\ = & \frac{1}{2\delta^2}\sum_{t=1}^{t=T}(\ln\frac{S_t}{\hat{S}_t})^2 + \frac{1}{2h\sigma^2}\sum_{t=2}^{T}(\ln\frac{V_t}{V_{t-1}} - (\mu - \frac{\sigma^2}{2})h)^2 \\ & + \log\, \delta^T \sigma^{T-1} \end{aligned} \tag{27}$$

	μ	σ	δ
Experiment Set 1	0.2	0.3	0.001:0.001:0.020
Experiment Set 2	0.2	0.7	0.001:0.001:0.020
Experiment Set 3	0.2	0.05:0.05:1.00	0.004
Experiment Set 4	0.2	0.05:0.05:1.00	0.016

TABLE I

Parameter settings for μ, σ and δ. $a : \Delta : b$ means from a to b with step size Δ.

We can compute the partial derivative for an optimization procedure:

$$\frac{\partial PNL(\ln V)}{\partial \ln V} = \lambda \ln V + \frac{\partial NL(\ln V)}{\partial \ln V} \tag{28}$$

In particular,

$$\begin{aligned} & \frac{\partial NL(\ln V)}{\partial \ln V_t} \\ = & -\frac{1}{\delta^2}\frac{\ln S_t - \ln \hat{S}_t}{\hat{S}_t}\frac{\partial \hat{S}}{\partial \ln V_t} \\ & + \frac{1}{h\sigma^2}[(\ln V_t - \ln V_{t-1}) - (\ln V_{t+1} - \ln V_t)] \end{aligned} \tag{29}$$

for $t = 2, \cdots, T-1$. We can derive the partial derivatives for $t = 1$ and $t = T$ similarly.

V. Performance study

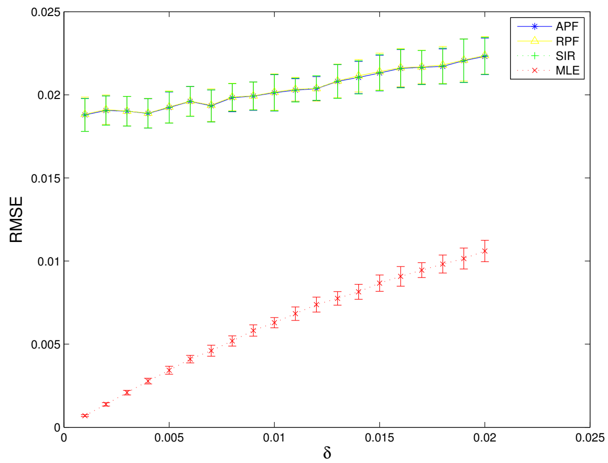
We simulate 252 days' (one trading year) of data. We set μ, δ and σ according to the parameter setting in Table I, as suggested in the simulation study in [4]. With these parameters, we simulate trajectories of $\ln V$ according to (24). With asset values V_t, we simulate equity prices S_t contaminated by noises according to (25). We examine the performance of auxiliary particle filters (APF), regularised particle filters (RPF), sampling importance resampling (SIR) and MLE for the inference of the latent asset value. We use the L-BFGS package to solve the optimization problem in MLE.

We measure the root mean square error (RMSE) and the running time. Since in a simulation study, we know the true latent asset value, we define the RMSE as:

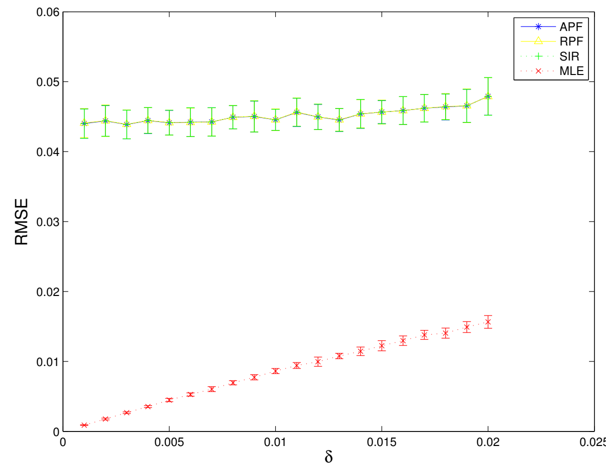
$$\text{RMSE} = \sqrt{\frac{1}{T-1}\sum_{i=2}^{i=T}(\ln V - \widetilde{\ln V})^2} \tag{30}$$

where $\ln V$ is the true latent log asset value and $\widetilde{\ln V}$ is the latent log asset value inferred by APF, RPF, SIR, or MLE.

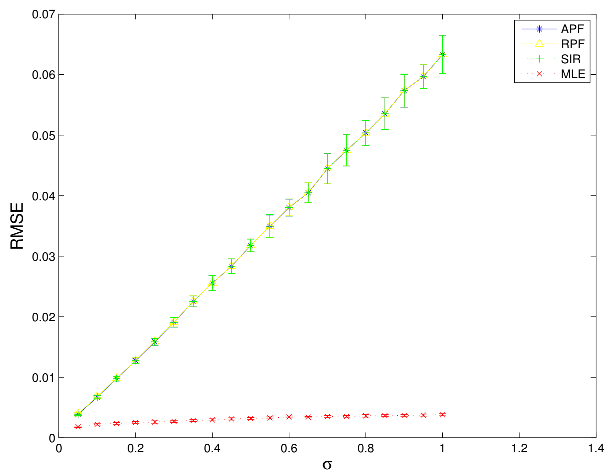
Figures 1 report the results for RMSE. The curves for APF, RPF and SIR overlap; and the curve for MLE is much lower than the curves for particle filtering approaches. Figures on the right show the results for the running time. The curve for MLE is much lower than the curves for particle filtering approaches. These results indicate clearly that the MLE approach has a much lower RMSE and a much lower running time. On the other hand, MLE makes a point-inference; while a particle filtering approach returns a distribution represented



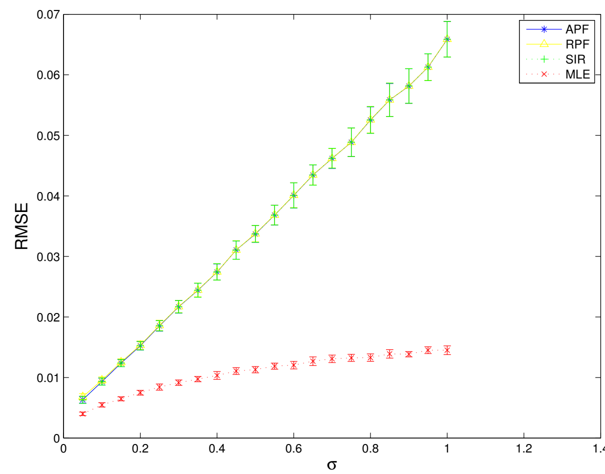
(a) RMSE for Experiment Set 1
$\mu = 0.2$, $\sigma = 0.3$, $\delta = 0.001 : 0.001 : 0.020$



(b) RMSE for Experiment Set 2
$\mu = 0.2$, $\sigma = 0.7$, $\delta = 0.001 : 0.001 : 0.020$



(c) RMSE for Experiment Set 3
$\mu = 0.2$, $\sigma = 0.05 : 0.05 : 1.00$, $\delta = 0.004$



(d) RMSE for Experiment Set 4
$\mu = 0.2$, $\sigma = 0.05 : 0.05 : 1.00$, $\delta = 0.016$

Fig. 1. Experimental results for RMSE.

by the particles. Moreover, the MLE approach conducts the inference in an offline, batch mode, in contrast to the online manner by the particle filtering approaches.

VI. Conclusion

Credit risk analysis is not only an important research topic in finance, but also of interest in everyday life. The non-linearity nature of the structural credit risk model makes it rather difficult to infer the latent asset value sequence. The main contribution of this paper is to address this inference problem by an MLE approach and is solved by the L-BFGS algorithm that is usually done using particle filtering. Experiments demonstrates the superiority of the proposed MLE approach to the particle filter approaches, in particular, auxiliary particle filters, regularised particle filters and sampling importance resampling. This work has the merit for the general problem of inferring latent values for probabilistic time-series.

References

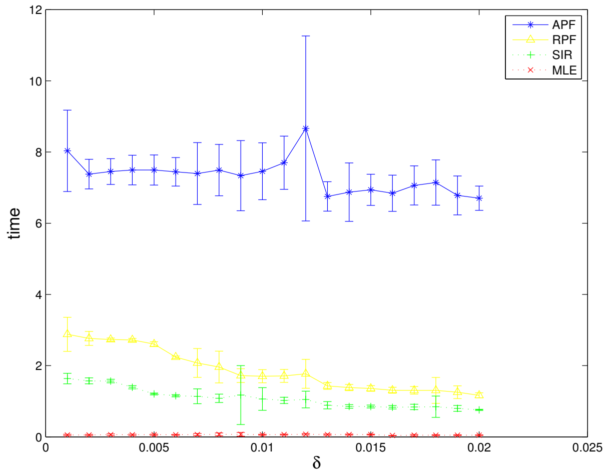
[1] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, Feburary 2002.

[2] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, May-June 1973.

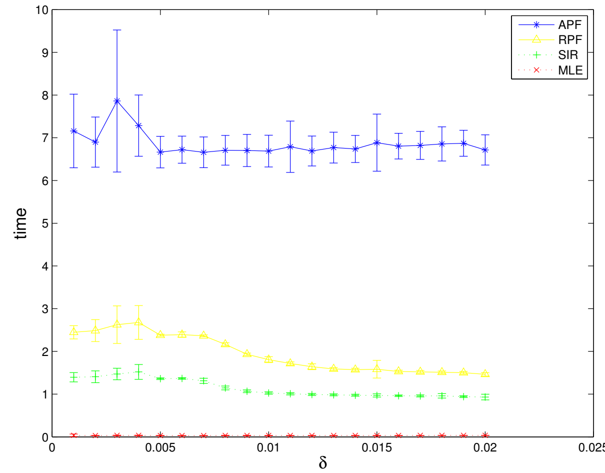
[3] A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo methods in practice*. Springer-Verlag, New York, 2001.

[4] J.-C. Duan and A. Fulop. Estimating the structural credit risk model when equity prices are contaminated by trading noises. *(to appear in Journal of Econometrics)*, 2008.

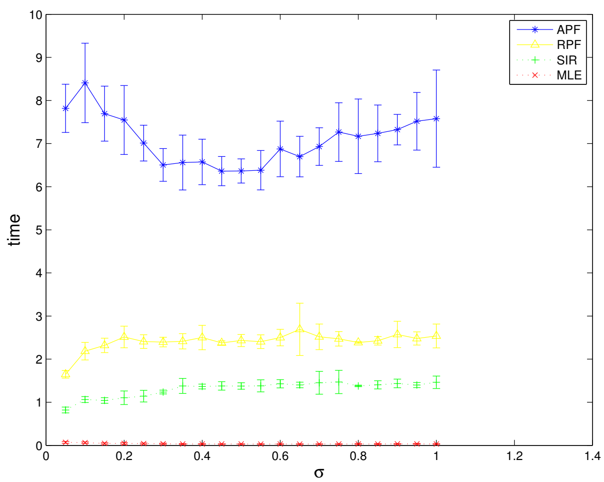
[5] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F: Radar and Signal Processing*, 140(2):107–113, April 1993.



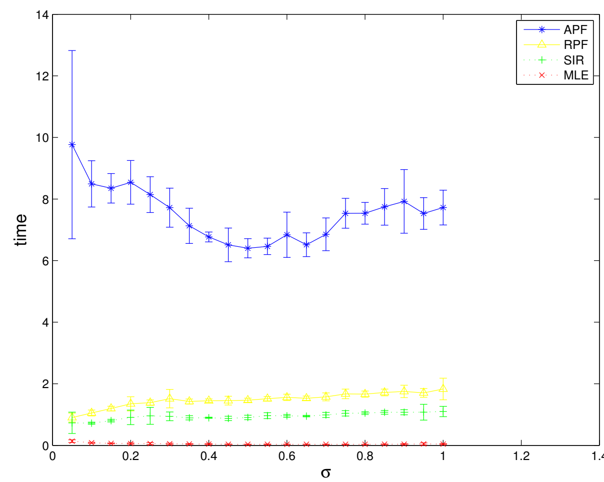
(a) Running time for Experiment Set 1
$\mu = 0.2$, $\sigma = 0.3$, $\delta = 0.001 : 0.001 : 0.020$



(b) Running time for Experiment Set 2
$\mu = 0.2$, $\sigma = 0.7$, $\delta = 0.001 : 0.001 : 0.020$



(c) Running time for Experiment Set 3
$\mu = 0.2$, $\sigma = 0.05 : 0.05 : 1.00$, $\delta = 0.004$



(d) Running time for Experiment Set 4
$\mu = 0.2$, $\sigma = 0.05 : 0.05 : 1.00$, $\delta = 0.016$

Fig. 2. Experimental results for running time.

[6] J. S. Liu and R. Chen. Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443):1032–1044, September 1998.

[7] A. Madhavan, M. Richardson, and M. Roomans. Why do security prices change? a transaction-level analysis of NYSE stocks. *Review of Financial Studies*, 10:1035–1064, 1997.

[8] R. C. Merton. On the pricing of corporate debt: the risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, May 1974.

[9] C. Musso, N. Oudjane, and F. L. Gland. Improving regularised particle filters. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*, New York, 2001. Springer-Verlag.

[10] J. Nocedal. Updating quasi-newton matrices with limited storage. *Mathematics of Computation*, 35(151):773–782, 1980.

[11] M. K. Pitt and N. Shephard. Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association*, 94(446):590–599, June 1999.