

Volatility Estimation Using High, Low, and Close Data – a Maximum Likelihood Approach

Malik Magdon-Ismail
CS Department
Lally 207, RPI
110 8th Street
Troy, NY 12180
magdon@cs.rpi.edu

Amir F. Atiya
Learning Systems Group
EE Department
Caltech, MC 136-93
Pasadena, CA 91125
amir@work.caltech.edu

The necessity of an accurate volatility estimate in order to price derivatives, and the time varying nature of volatility, make it imperative to obtain reliable volatility estimates using only the most recent data. More specifically, it is crucial to make use of all the available information. Many volatility estimates are based on the close prices of the instrument alone, despite the fact that high and low data are also available. It is to be expected that throwing away such high-low information will lead to a suboptimal volatility estimate when compared to an estimate that takes this extra information into account. We present a maximum likelihood approach to using the high, low and close information for obtaining volatility estimates. We present experiments to demonstrate that our estimate obtains consistently better performance than existing estimates on simulated data. In addition, we present simulations on real data, demonstrating that our method results in a more stable volatility estimate.

1.1 Introduction

The volatility of a financial instrument is a crucial parameter for a number of reasons. It enters as a parameter in pricing formulas for derivatives instruments, and plays key roles in asset allocation, portfolio optimization strategies and risk management. As a result, considerable attention has been devoted to the accurate estimation of volatility. Because it is recognized that volatility is time varying, it is imperative to use only the most recent price observations to construct an estimate of the volatility. To obtain a good estimator, one should thus attempt to make the utmost use of the small number of observations available. In this paper, we propose an estimator that uses the high and low price information in addition to the closing price used by conventional estimators. In practice this would be of great interest because most historical data is quoted with both the high and low in addition to the close.

Volatility estimates using high and low prices have been considered to some extent in the literature. All previous studies have considered securities characterized by geometric Brownian motion. The logarithm of such a process then follows a Brownian motion. Parkinson [Parkinson 1980] shows that expectation of the high minus the low squared is proportional to σ^2 , and thus constructs an estimate based on the high minus the low. Garman and Klass [Garman and Klass 1980] define a quadratic function of the high, low and close, and derive the parameters of such a function that result in the estimate being unbiased, (their estimate is unbiased only in the case of zero drift). Rogers et al [Rogers and Satchell 1991, Rogers et al 1994, Rogers 1998] propose another formula, and show that it is an unbiased estimate even for non-zero drift. The problem with these approaches is that they are not necessarily optimal estimates. In addition, they consider only one period (one day for example). By taking the average of the estimates over the days considered in the data set, unbiasedness of the estimates will prevail, but optimality will generally not be valid.

In this paper we propose a new estimate for both the volatility and the drift using maximum likelihood. We derive an expression for the joint density of the maximum and the minimum of a Brownian motion, and construct a likelihood function that we maximize by a two-dimensional search.

One of the advantages of the maximum likelihood approach lies in the the fact that it produces estimates that are asymptotically efficient. Further, if one assumes independence among the time periods, as is customarily the case, then multiple time periods can be incorporated by using the product likelihood function. In addition, within this probabilistic framework, it is straightforward to employ a fully Bayesian,

decision theoretic approach, whereby, one enforces certain priors that one might have on the drift and the volatility (for example, in a risk averse world, the drift should be higher than the risk free rate).

This paper is organized as follows. First we develop the maximum likelihood formulation of the problem. We then present extensive simulations to compare our method to existing methods (the close estimator, Parkinson's estimator, Rogers' et al estimator, and the Garman–Klass estimator). Finally we demonstrate our method on real data. We compare our method for estimating the volatility using to the method of Parkinson [Parkinson 1980] and to the estimate based on using the close price alone. Our simulations indicate that, the RMS prediction error of our estimator is about 2.6–2.7 times less than that of the estimator using the closing price alone. In practical terms, using the closing price alone would require about 35–40 days of data to obtain a comparable accuracy to our method on 5 days. Other methods that use the high-low information also obtain reductions in the RMS prediction error when compared to the close, but not by as much as our method. For comparison, Parkinson's method obtains a reduction by a factor of about 2.2. On real data, we demonstrate that our method obtains more stable (and hence more realistic) time varying volatility estimates.

1.2 Maximum Likelihood Approach

Our approach will be to obtain the conditional density for the high (h) and the low (l) over a time period T , given that the process starts at x and undergoes brownian motion with drift μ and volatility σ . Thus, we assume that

$$dx_t = \mu dt + \sigma dW \quad (1.1)$$

where dW represents increments due to the standard Wiener process. It is commonly assumed that the above dynamics describes the motion of the log stock price of many widely held securities. We would like to derive the conditional density

$$p(h, l|x, \mu, \sigma, T) \quad (1.2)$$

which is the probability density for obtaining a high of h and a low of l over the interval T given x , μ , σ . Having derived the conditional distribution above, the likelihood formulation is as follows. One observes a set of prices, which consists of the open x_0 of the first day and the triples $\{h_i, l_i, c_i\}_{i=1}^N$ where i indexes the consecutive days for which one has data, and h , l , c represent high, low, close

respectively. We assume that the close of any given day is the open of the next day, hence we can define the series of opens by $o_1 = x_0$ and $o_i = c_{i-1}$, for $i > 1$. If one assumes independence from day to day, then the likelihood for the set of N days becomes the product of the likelihoods of each day. The log likelihood then becomes a sum and is given by the formula

$$L(\mu, \sigma) = \sum_{i=1}^N \log p(l_i, h_i | o_i, \mu, \sigma, T) \quad (1.3)$$

which represents the function that we wish to maximize with respect to μ and σ . The values of μ and σ that maximize (1.3) represent our estimates $\hat{\mu}$ and $\hat{\sigma}$.

What remains is to obtain $p(h, l | x, \mu, \sigma, T)$. In order to do this, we introduce the distribution function that governs the first passage through two barriers given by h_1 and h_2 , with $h_1 \leq x \leq h_2$. We denote this distribution function by:

$$F(h_1, h_2 | x, \mu, \sigma, T) \quad (1.4)$$

which is the probability that the process remains bounded by $h_1 \leq x_t \leq h_2$ for all $t \in [0, T]$ given the parameters of the diffusion. From this distribution function, one obtains the density $p(h, l | x, \mu, \sigma, T)$ by taking derivatives as follows.

$$p(h, l | x, \mu, \sigma, T) = - \frac{\partial^2}{\partial h_1 \partial h_2} F(h_1, h_2 | x, \mu, \sigma, T) \Big|_{h_1=l, h_2=h} \quad (1.5)$$

where the right hand side represents the second order partial derivative with respect to the two barrier levels evaluated at the low and high. Dominé [Dominé 1996] has computed a series expansion for exactly this first passage time distribution F , hence what remains is to compute the necessary derivatives. The formulas are tedious and their explicit form is given in the appendix, see equation (A.24) in section A.1, which is a series representation for the density $p(h, l | x, \mu, \sigma, T)$, and can be computed to any desired accuracy by taking sufficiently many terms. As a practical point, it is found that as the volatility decreases, more terms in the series should be computed to maintain the accuracy of the estimate.

1.3 Simulation Results

In our first simulation, we will assume that μ is known and does not need to be estimated. It is frequently the case that this assumption is made by equating the drift to the risk free rate (this can be done providing that $\mu \gg \sigma^2$). In our

simulations, we compare different estimators by looking at their estimates based on observed data over a window ranging from 5 days to 50 days. We obtain the RMS prediction error ($\sqrt{E[(\hat{\sigma} - \sigma)^2]}$) using 2000 realizations of each window size. For our simulations, we set $T = 1$, $\mu = 0.02$, $\sigma = 0.5$ and $x_0 = 0$. We assume that the drift μ is known and only the volatility σ needs to be estimated. Shown in first four columns of Table 1.1 is a comparison of four methods. The first method uses the close prices only and the estimate is given by

$$\hat{\sigma}_{close} = \sqrt{\frac{1}{NT} \sum_{i=1}^N (c_i - o_i - \mu T)^2} \quad (1.6)$$

The second method (Parkinson's estimator) uses only the high and low values [Parkinson 1980], and the estimate is given by

$$\hat{\sigma}_{Park} = \sqrt{\frac{1}{4NT \ln 2} \sum_{i=1}^N (h_i - l_i)^2} \quad (1.7)$$

The third method (the Rogers-Satchell estimator) uses the high low and close prices [Rogers and Satchell 1991], and is given by

$$\sigma_{RS} = \sqrt{\frac{1}{NT} \sum_{i=1}^N (h_i - o_i)(h_i - c_i) + (l_i - o_i)(l_i - c_i)} \quad (1.8)$$

The fourth method is our maximum likelihood based method,

$$\hat{\sigma}_{ML} = \underset{\nu}{\operatorname{argmax}} L(\mu, \nu) \quad (1.9)$$

We show the results of these simulations in the first four columns of Table 1.1. The results are also summarised in Figure 1.1. From these results it is clear that our method produces a superior estimate.

One might note that the Parkinson and Rogers-Satchell estimates do not rely on knowledge of μ . Thus, one might argue that one should not assume that μ is known in comparing these estimators with our Maximum likelihood estimator. One can repeat the previous simulation without assuming that μ is known. In this case, both μ and σ need to be estimated – this is only relevant for the close estimator and our maximum likelihood estimator. The results of this simulation are shown in Table 1.1 in the “ μ unknown” columns. These results are also summarised in Figure 1.2.

Days used	RMS Prediction Error							
	μ known				μ unknown		$\mu = 0$	
	Close	Park	R-S	ML	Close	ML	G-K	ML
5	0.1597	0.0713	0.0642	0.0621	0.1752	0.0639	0.0591	0.0640
10	0.1090	0.0489	0.0448	0.0426	0.1152	0.0434	0.0399	0.0417
15	0.0900	0.0410	0.0375	0.0353	0.0930	0.0354	0.0337	0.0346
20	0.0781	0.0360	0.0317	0.0303	0.0808	0.0307	0.0292	0.0304
25	0.0702	0.0317	0.0289	0.0273	0.0709	0.0270	0.0260	0.0271
30	0.0645	0.0292	0.0270	0.0246	0.0654	0.0248	0.0233	0.0245
35	0.0605	0.0272	0.0245	0.0230	0.0615	0.0229	0.0224	0.0232
40	0.0556	0.0252	0.0227	0.0215	0.0559	0.0215	0.0205	0.0215
45	0.0526	0.0238	0.0215	0.0200	0.0534	0.0202	0.0192	0.0196
50	0.0499	0.0222	0.0204	0.0192	0.0505	0.0191	0.0186	0.0191

Table 1.1
Comparison of different volatility estimation methods.

Our method not only remains superior to all the other methods, but the fact that μ needs to be estimated as well has not significantly worsened the performance.

As one further comparison, we consider the special case of $\mu = 0$. In this case, Garman–Klass [Garman and Klass 1980] have constructed the optimal (in the least squares sense) quadratic estimator as

$$\sigma_{GK} = \sqrt{\frac{1}{NT} \sum_{i=1}^N 0.511(\tilde{h}_i - \tilde{l}_i)^2 - 0.019(\tilde{c}_i(\tilde{h}_i + \tilde{l}_i) - 2\tilde{l}_i\tilde{h}_i) - 0.383\tilde{c}_i^2} \quad (1.10)$$

The tilde over the symbols indicates that one “normalizes” the quantities by subtracting the open prices. For example $\tilde{h}_i = h_i - o_i$. A comparison between this estimator and our maximum likelihood estimator for the case $\mu = 0$ is given in the last two columns of Table 1.1. In this special case of zero drift, our estimator does not beat the optimal quadratic estimator for small window sizes, but it approaches the optimal estimator as the window size increases. These results are summarised in Figure 1.3, along with the performance of the other estimators for this special case of $\mu = 0$. From Figure 1.3, it is seen that the Garman–Klass estimator is slightly better than the maximum likelihood estimator, but, however, one can also note that the maximum likelihood estimator is asymptotically approaching the

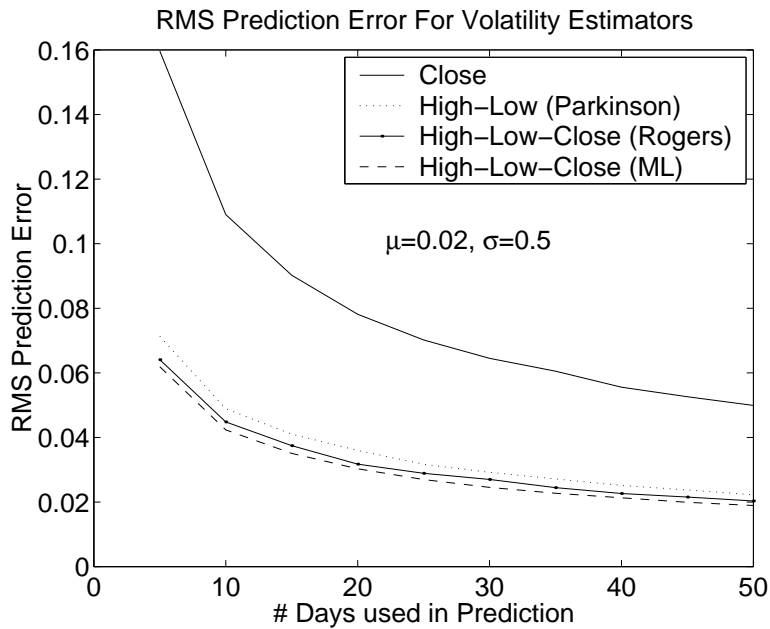


Figure 1.1 Comparison of volatility prediction methods using the RMS prediction error when the drift parameter is known.

Garman–Klass estimator as one might expect due to the asymptotic efficiency of maximum likelihood estimators.

We have also applied our method to real data. Since no ground truth is available in this case, we compare the methods based on stability measures for the volatility estimate. We estimate the volatility from a 10 day moving window using the close estimator, Parkinsons high–low estimator and our maximum likelihood high–low–close estimator. The volatility should not vary drastically from day to day. At most, it should be slowly varying, especially since the model assumes a constant volatility. We have experimented with several stocks and in all cases, the maximum likelihood estimate is clearly the most stable. Representative volatility time series predictions based on a 10 day prediction window are shown for IBM in Figure 1.4. In general, it is desirable to use such a small prediction window as the volatility can change over time. The figures on the left show the volatility predictions. The figures on the right

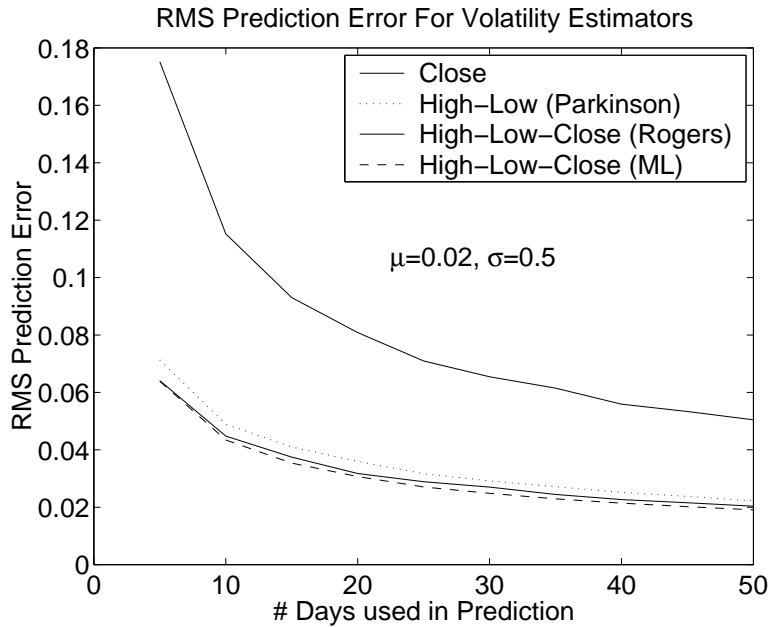


Figure 1.2

Comparison of volatility prediction methods using the RMS prediction error when the drift has to be estimated as well.

show the variability in the volatility predictions over a moving window of size 20 days. The figures on the left clearly show that the maximum likelihood estimate is qualitatively more stable, due in part to the fact that it is a more accurate estimate of the volatility. This qualitative difference is quantified in the figures on the right.

1.4 Conclusions

We have presented a formula for obtaining the joint distribution of the high and low given the open and the parameters of the Wiener process. Using this formula, one can construct a likelihood for the observed data given the parameters and hence obtain a maximum likelihood estimate. One could also employ a fully Bayesian framework to obtain a Bayes optimal estimator under some risk measure, if one has some prior information on the possible values of the volatility. We have shown that

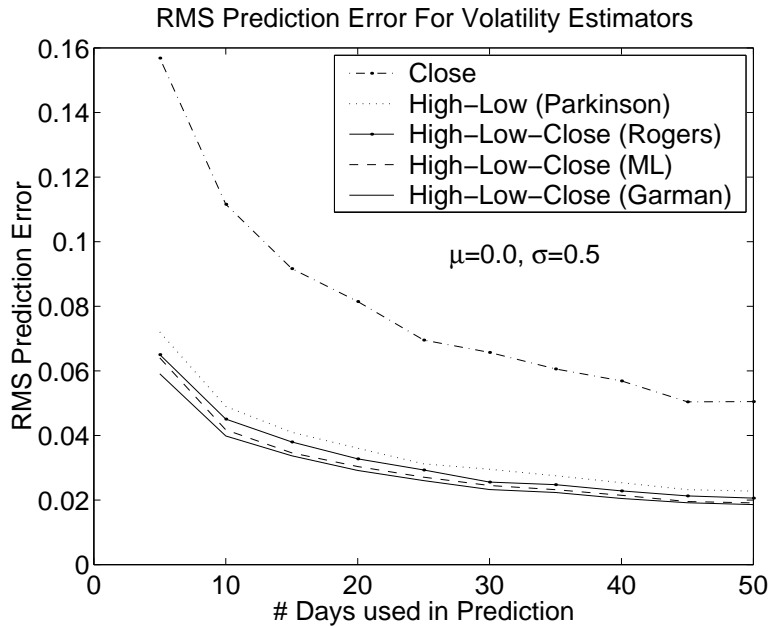


Figure 1.3 Comparison of volatility prediction methods using the RMS prediction error for the special case of zero drift.

in simulations, our estimator obtains a significant improvement over the conventional close estimator as well as other estimators based on the high, low and close values. Further we have demonstrated that our method produces a much more stable estimator on real data, thus enabling one to make reliable volatility estimates using the few most recent data points. It is expected that a more accurate estimate of the time varying volatility will lead to more efficient pricing of volatility based derivative instruments such as options.

A.1 Joint Density of the High and Low

In this appendix, we compute a series expansion for the joint density of the high and the low given the open, the drift and volatility parameters. This is the expression that is needed for the computation of the likelihood as given in (1.3). We start

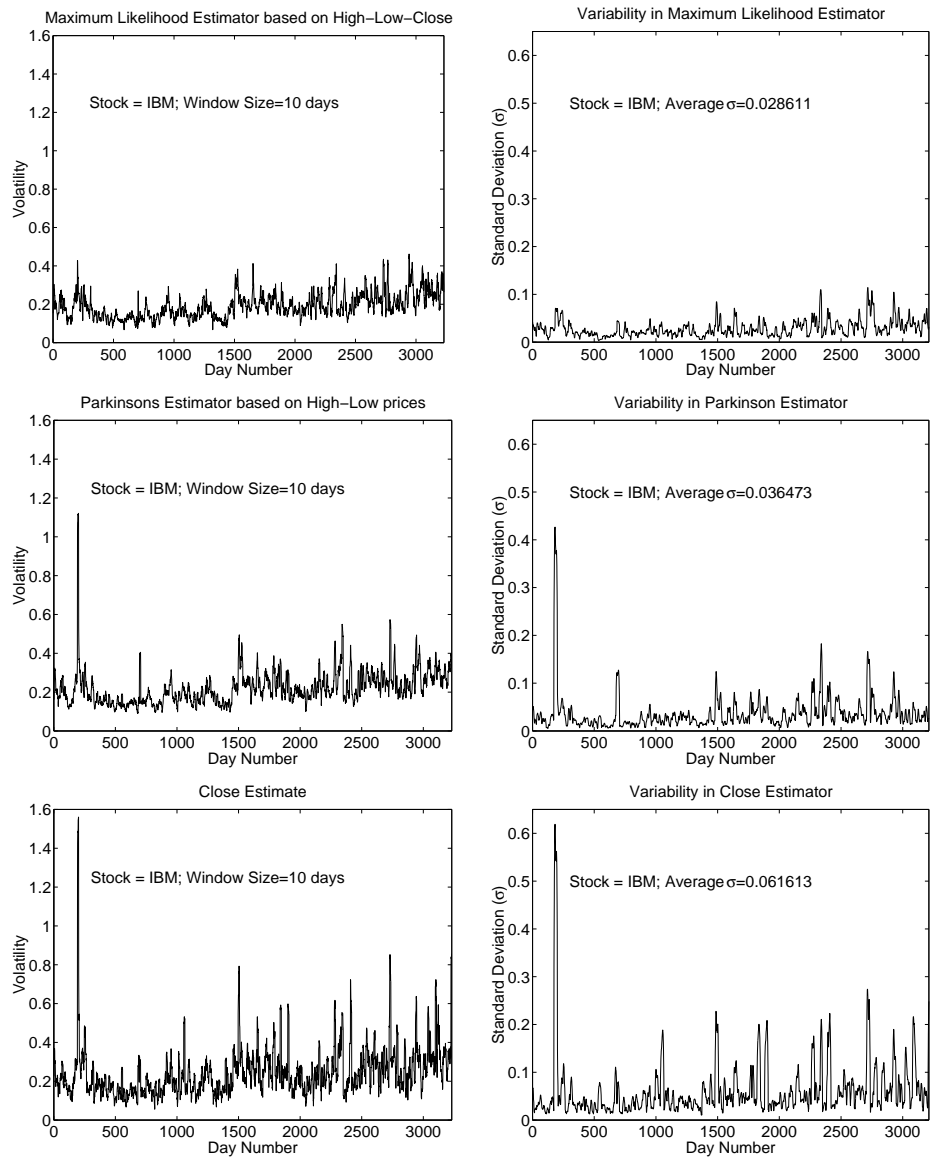


Figure 1.4
Volatility estimates on IBM's return series for various predictors.

with the distribution function (1.4). A series expansion is derived in [Dominé 1996] which we reproduce here for convenience.

$$F(h_1, h_2 | x, \mu, \sigma, T) = \sum_{k=1}^{\infty} 2\sigma^4 k\pi \tilde{C}(k, x, \mu, \sigma, h_1, h_2) f(k, \mu, \sigma, T, (h_1 - h_2)^2) \quad (\text{A.11})$$

where the functions \tilde{C} and f are given by

$$\begin{aligned} \tilde{C}(k, x, \mu, \sigma, h_1, h_2) &= \left[\exp\left(\frac{\mu}{\sigma^2}(h_2 - x)\right)(-1)^{k+1} + \exp\left(\frac{\mu}{\sigma^2}(h_1 - x)\right) \right] \\ &\quad \times \sin\left(k\pi \frac{x - h_1}{h_2 - h_1}\right) \end{aligned} \quad (\text{A.12})$$

and

$$f(k, \mu, \sigma, T, u) = \frac{\exp\left(-\frac{g(k, \mu, \sigma, u)T}{2\sigma^2 u}\right)}{g(k, \mu, \sigma, u)} \quad (\text{A.13})$$

where

$$g(k, \mu, \sigma, u) = \mu^2 u + \sigma^4 k^2 \pi^2 \quad (\text{A.14})$$

In order to make the notation more concise, we will suppress the k, x, μ, σ, T dependence when referring to the above functions (keeping only the h_1, h_2 dependence), for example we will write $f(u)$ instead of $f(k, \mu, \sigma, u)$. The derivatives of $\tilde{C}(h_1, h_2)$ and $f(u)$ will be needed. We will use the usual subscript notation to denote the partial derivatives with respect to the arguments, for example, $\tilde{C}_{i,j}(h_1, h_2)$ is the i^{th} partial derivative with respect to the first argument and the j^{th} partial derivative with respect to the second argument. Using this notation, the joint density (1.5) is given by

$$p(h, l | x, \mu, \sigma, T) = -F_{1,1}(h_1, h_2)|_{h_1=l, h_2=h} \quad (\text{A.15})$$

We will need the partial derivatives $f_1(u), f_2(u), \tilde{C}_{0,1}(h_1, h_2), \tilde{C}_{1,0}(h_1, h_2)$ and $\tilde{C}_{1,1}(h_1, h_2)$. Tedious but straightforward computations yield the following expressions.

$$f_1(u) = f(u) \left[\frac{T}{2\sigma^2 u} \left(\frac{g(u)}{u} - \mu^2 \right) - \frac{\mu^2}{g(u)} \right] \quad (\text{A.16})$$

$$f_2(u) = \frac{f_1(u)^2}{f(u)} + f(u) \left[\frac{T}{\sigma^2 u^2} \left(\mu^2 - \frac{g(u)}{u} \right) + \frac{\mu^4}{g(u)^2} \right] \quad (\text{A.17})$$

In order to write the derivatives of \tilde{C} more compactly, introduce the function $A(h_1, h_2) = k\pi(x - h_1)/(h_2 - h_1)$. The derivatives of A are then given by

$$A_{0,1}(h_1, h_2) = -\frac{A(h_1, h_2)}{(h_2 - h_1)} \quad (\text{A.18})$$

$$A_{1,0}(h_1, h_2) = A(h_1, h_2) \left(\frac{1}{h_2 - h_1} - \frac{1}{x - h_1} \right) \quad (\text{A.19})$$

$$A_{1,1}(h_1, h_2) = \frac{A(h_1, h_2)}{h_2 - h_1} \left(\frac{1}{x - h_1} - \frac{2}{h_2 - h_1} \right) \quad (\text{A.20})$$

One then finds for the derivatives of \tilde{C}

$$\tilde{C}_{0,1}(h_1, h_2) = A_{0,1} \cot(A) \tilde{C} + (-1)^{k+1} \frac{\mu}{\sigma^2} \sin(A) \exp\left(\frac{\mu}{\sigma^2}(h_2 - x)\right) \quad (\text{A.21})$$

$$\tilde{C}_{1,0}(h_1, h_2) = A_{1,0} \cot(A) \tilde{C} + \frac{\mu}{\sigma^2} \sin(A) \exp\left(\frac{\mu}{\sigma^2}(h_1 - x)\right) \quad (\text{A.22})$$

$$\begin{aligned} \tilde{C}_{1,1}(h_1, h_2) &= A_{1,1} \cot(A) \tilde{C} + \frac{A_{1,0} A_{0,1}}{\sin^2(A)} \tilde{C} + A_{1,0} \cot(A) \tilde{C}_{0,1} \\ &\quad + \frac{\mu}{\sigma^2} A_{0,1} \cos(A) \exp\left(\frac{\mu}{\sigma^2}(h_1 - x)\right) \end{aligned} \quad (\text{A.23})$$

where we have suppressed the (h_1, h_2) dependence of the A and \tilde{C} functions on the right hand side. Finally the function $-F_{1,1}(h_1, h_2)$ can be obtained from the following expression

$$\begin{aligned} F_{1,1} &= 2\sigma^4 \pi \sum_{k=1}^{\infty} k \left[\tilde{C}_{1,1} f(u) + 2(h_2 - h_1) f_1(u) (\tilde{C}_{1,0} - \tilde{C}_{0,1}) \right. \\ &\quad \left. - 4\tilde{C} f_2(u) (h_2 - h_1)^2 - 2\tilde{C} f_1(u) \right] \end{aligned} \quad (\text{A.24})$$

References

- Dominé, M. 1996. First passage time distribution of a wiener process with drift concerning two elastic barriers. *Journal of Applied Probability*, 33:164–175.
- Garman, M. and M. Klass. 1980. On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1):67–78.
- Parkinson, M. 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65.
- Rogers, L. 1998. Volatility estimation with price quanta. *Mathematical Finance*, 8(3):277–290.

Rogers, L. and S. Satchell. 1991. Estimating variance from high, low and closing prices. *Annals of Applied Probability*, 1(4):504–512.

Rogers, L., S. Satchell, and Y. Yoon. 1994. Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Applied Financial Economics*, 4:241–247.