High Frequency Trade Prediction with Bivariate Hawkes Process[1]

John Carlsson, Mao-Ching Foo, Hui-Huang Lee, Howard Shek
Stanford University
10 June 2007

Summary

In this project, we used a bivariate Hawkes process to model conditional arrival intensities of buy and sell orders of liquid stocks. We then look into simple trading strategies using MLE parameters of the model. For some of the stocks to which we have fitted the model and applied the strategy, we seem to be able to extract significant positive trading gain[2]*.*

1 Introduction

In the first part of the report, we introduce both the univariate and the bivariate Hawkes processes, with sketch proofs. Simulations of these processes are then carried out to illustrate the self-excitation and cross-excitation features of the model. Next, we introduce the MLE procedure for parameter estimation, first with some background theory and then followed by application - both on simulated process and on actual tick data from TAQ on select stocks. Second part of the report explore a simple trading strategy based on the fitted model.

2 Hawkes Process

2.1 Univariate case

Here we model the intensity λ_t of the counting process by the particular form of Hawkes process that satisfies the following SDE

$$d\lambda_t = \kappa \left(\rho \left(t\right) - \lambda_t\right) dt + \delta dN_t$$

The solution for λ_t can be written (see Appendix A)

$$\lambda_t = \lambda_\infty + \delta \int_0^t e^{-\kappa(t-u)} dN_u \tag{1}$$

where we can think of λ_∞ as the long run "base" intensity, i.e. the intensity if there have been no past arrival.

The linkage between the intensity and the underlying counting process N_t is via the Doob-Meyer decomposition and the two filtrations $\mathcal{H}_t \subset \mathcal{F}_t$, one for the intensity and the other for the jump time

$$\mathcal{H}_t = \sigma \left\{\lambda_s : s \leq t\right\}$$

and

$$\mathcal{F}_t = \sigma \left\{N_s : s \leq t\right\}$$

Then it can be shown [1]

$$E\left[e^{iv(N_s - N_t)} \middle| \mathcal{F}_t \right] = e^{-\Psi(v)(A_s - A_t)}$$

where $\Psi\left(v\right) = 1 - e^{iv}$ and $M_t = N_t - A_t$ is a $\mathcal{F}_t$-adapted martingale. Hence, conditional on the realization of the compensator $A_t = \int_0^t \lambda_u du$ i.e. on $\mathcal{H}_t$, the process is non-homogenous Poisson with

[1] This report forms part of the coursework requirements for MS&E444. Code and data are available upon request by contacting the authors at: johnnyc@stanford.edu, maoching@stanford.edu, huihuang@stanford.edu, shek@stanford.edu. The authors would like to acknowledge our appreciation for guidance from Professor Kay Giesecke and course TA Benjamin Armbruster.

[2] Note we assumed: no transaction cost, no canonical impact of trade order, no short sell limit, transaction at trade price

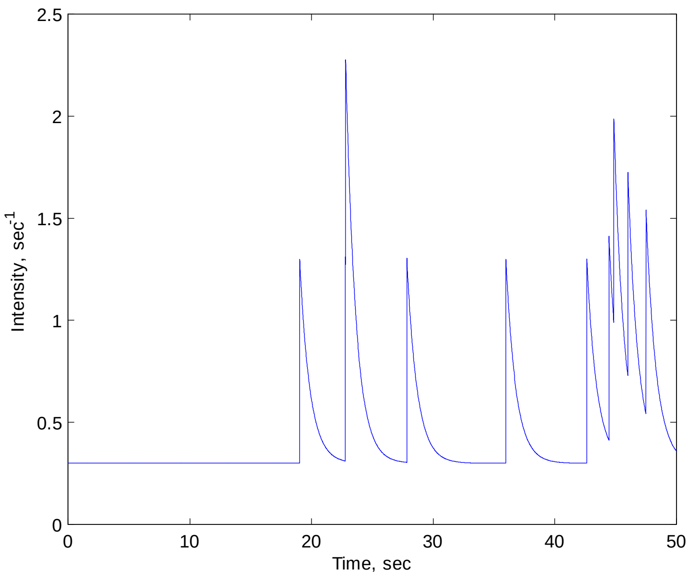
deterministic intensity

$$\begin{aligned} E\left[N_{t+\delta t}-N_t\right|\mathcal{F}_t] &= E\left[A_{t+\delta t}-A_t\right|\mathcal{F}_t] \\ \lim_{\delta t\to 0}\frac{1}{\delta t}E\left[N_{t+\delta t}-N_t\right|\mathcal{F}_t] &= \lim_{\delta t\to 0}\frac{1}{\delta t}E\left[E\left[A_{t+\delta t}-A_t\right|\mathcal{H}_t\vee\mathcal{F}_t]\right|\mathcal{F}_t] \\ &= \lim_{\delta t\to 0}\frac{1}{\delta t}E\left[\left.\int_t^{t+\delta t}\lambda_u du\right|\mathcal{F}_t\right] \\ &= \lambda_t|\,F_t \end{aligned}$$

For a more thorough treatment of doubly stochastic processes, refer to [2].

2.2 Simulation of univariate Hawkes process

We can simulate this self-affected intensity process by the usual thinning method [6]. Below shows part of a simulated univariate intensity process. Note the clustering of intensity as a result of the self-excitation feature of the Hawkes process.

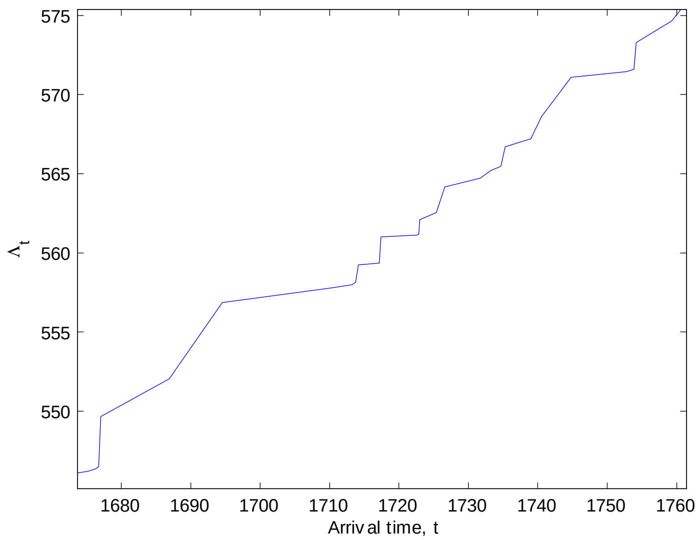


Simulated univariate Hawkes process with $(\mu,\alpha,\beta)=(0.3,0.6,1.0)$

To obtain the compensator Λ, we integrate the intensity piecewise

$$\begin{aligned} \int_0^T \lambda(u)|\,H_u du &= \int_0^T \mu du + \int_0^T \sum_{t_i<u}\alpha e^{-\beta(u-t_i)}dNdu \\ &= \mu T - \frac{\alpha}{\beta}\sum_{i=0}^{m} e^{-\beta(t_{i+1}-t_i)} \end{aligned}$$

Below we plot the compensator for the simulated Hawkes process using the above formula



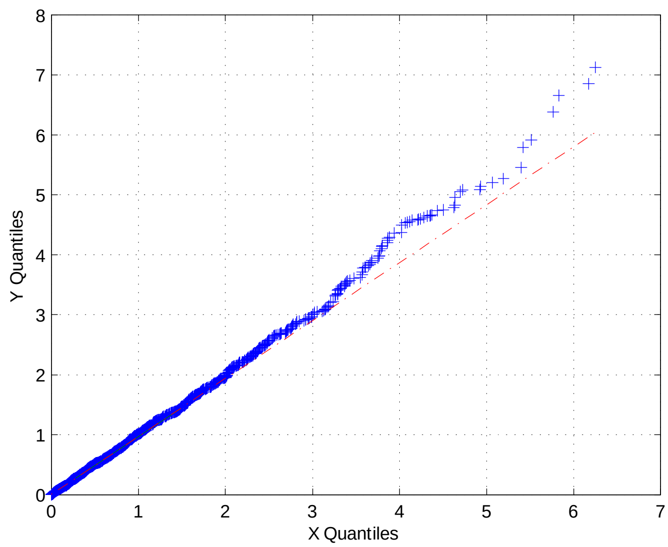
Compenstor for simulated univariate Hawkes process

Theorem 1 *Time Change Theorem. Given a point process with a conditional intensity function* $\lambda_t | H_t$*. Define the time-change*

$$\Lambda = \int_0^T \lambda_u | H_u du$$

where the filtration $H_t = \sigma \{0 < t_1 < t_2, ..., t_i \leq t\}$*. Assume that* $\Lambda_t < \infty$ *a.s.* $\forall t \in (0, T]$*, then* Λ_t *is a standard Poisson process.*

By application of the time change theorem above, we test the goodness of fit of the time-changed simulated process to that of a standard Poisson process. The QQ plot below validates both our compensator and the simulation code.



QQ-plot for time-changed simulated univariate Hawkes process,
with compensator Λ calculated from true parameters

2.3 Bivariate Case

A linear bivariate self-affected process with cross-excitation can be expressed, by modifying (1), to give

$$\begin{cases} \lambda_1(t) = \mu_1 + \int_0^t \upsilon_{11}(t-s)\, dN_1(s) + \int_0^t \upsilon_{12}(t-s)\, dN_2(s) \\ \lambda_2(t) = \mu_2 + \int_0^t \upsilon_{21}(t-s)\, dN_1(s) + \int_0^t \upsilon_{22}(t-s)\, dN_2(s) \end{cases} \tag{2}$$

Consider the parameterization of

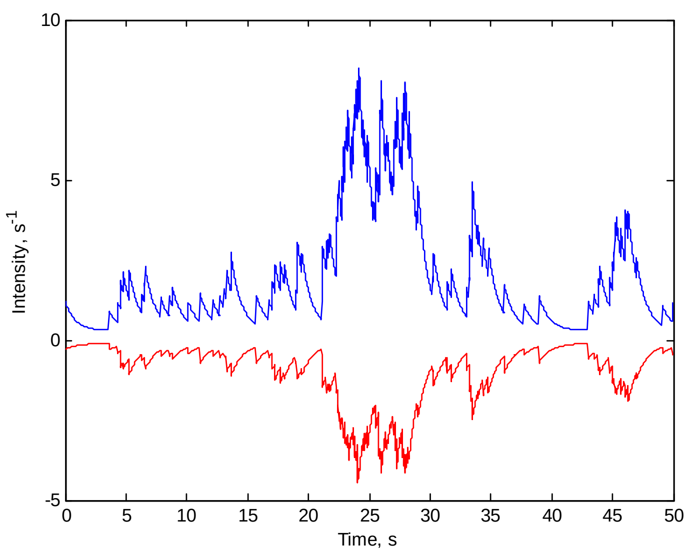
$$\upsilon_{ij}(s) = \alpha_{ij} e^{-\beta_i s}$$

We can then rewrite (2) as

$$\begin{cases} \lambda_1(t) = \mu_1 + \sum_{t_i<t} \alpha_{11} e^{-\beta_1(t-t_i)} + \sum_{t_j<t} \alpha_{12} e^{-\beta_1(t-t_j)} \\ \lambda_2(t) = \mu_2 + \sum_{t_i<t} \alpha_{21} e^{-\beta_2(t-t_i)} + \sum_{t_j<t} \alpha_{22} e^{-\beta_2(t-t_j)} \end{cases}$$

2.4 Simulation of bivariate Hawkes process

We can simulate this cross-affected intensity process again by the usual thinning method [6]. Below shows part of a simulated bivariate intensity process. Note the induced jumps between the two processes and the decay after each jumps.



Simulated bivariate Hawkes process with $(\mu_1, \mu_2, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_1, \beta_2) = (0.3, 0.1, 0.6, 0.9, 0.2, 0.5, 1.2, 1.0)$

To obtain the compensator Λ_1, we integrate the intensity piecewise to give

$$\begin{aligned} \int_0^T \lambda_1(u)|\, H_u du &= \int_0^T \mu_1 du + \int_0^T \sum_{t_i<u} \alpha_{11} e^{-\beta_1(u-t_i)} dN_1 du + \int_0^T \sum_{t_j<u} \alpha_{12} e^{-\beta_1(u-t_j)} dN_2 du \\ &= \mu_1 T - \frac{\alpha_{11}}{\beta_1} \sum_{i=0}^{m} e^{-\beta_1(t_{i+1}-t_i)} - \frac{\alpha_{12}}{\beta_1} \sum_{i=1}^{m} \sum_{\sup\{j<i\}}^{\sup\{j<i+1\}} e^{-\beta_1(t_{i+1}-t_j)} \end{aligned}$$

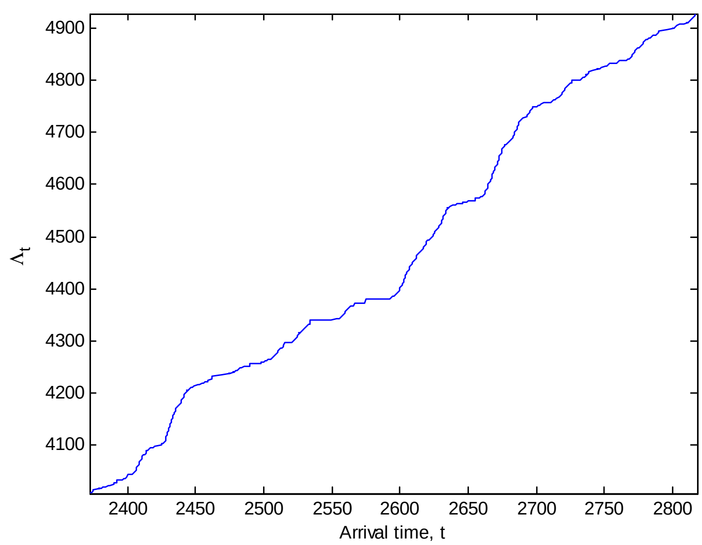
We can express

$$
\begin{aligned}
& \int_0^T \sum_{t_i<t} \alpha_{11} e^{-\beta_1(u-t_i)} dN_1 du \\
= \quad & \int_{\{t_0 \le u < t_1\}} \alpha_{11} e^{-\beta_1(u-t_{i=0})} du + \int_{\{t_1 \le u < t_2\}} \alpha_{11} \left(1 + e^{-\beta_1(t_{i=1}-t_{i=0})}\right) e^{-\beta_1(u-t_{i=1})} du \\
& + \int_{\{t_2 \le u < t_3\}} \alpha_{11} \left(1 + \left(1 + e^{-\beta_1(t_{i=1}-t_0)}\right) e^{-\beta_1(t_{i=2}-t_{i=1})}\right) e^{-\beta_1(u-t_{i=2})} du + ... \\
= \quad & \frac{\alpha_{11}}{-\beta_1} \left[\begin{array}{c} \left(e^{-\beta_1(t_{i=1}-t_{i=0})} - 1\right) + \left(1 + e^{-\beta_1(t_{i=1}-t_{i=0})}\right) \left(e^{-\beta_1(t_{i=2}-t_{i=1})} - 1\right) \\ + \left(1 + \left(1 + e^{-\beta_1(t_{i=1}-t_{i=0})}\right) e^{-\beta_1(t_{i=2}-t_{i=1})}\right) \left(e^{-\beta_1(t_{i=3}-t_{i=2})} - 1\right) + ... \end{array} \right]
\end{aligned}
$$

Similarly we have

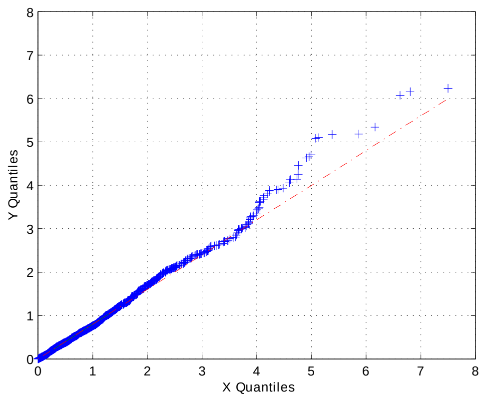
$$
\begin{aligned}
& \int_0^T \sum_{t_j<u} \alpha_{12} e^{-\beta_1(u-t_j)} dN_2 du \\
= \quad & \int_{\{t_{i=0} \le u < t_{i=1}, t_{i=1} \le t_{j=k} < t_{i=2}\}} \left\{ \alpha_{12} e^{-\beta_1(u-t_{j=k})} + \alpha_{12} \left(1 + e^{-\beta_1(t_{j=k+1}-t_{j=k})}\right) e^{-\beta_1(u-t_{i=0})} + ... \right\} du \\
& + \int_{\{t_{i=1} \le u < t_{i=2}, t_{i=1} \le t_{j=k} < t_{i=2}\}} \left\{ \begin{array}{c} \alpha_{12} \left(1 + e^{-\beta_1(t_{j=k+1}-t_{j=k})}\right) e^{-\beta_1(u-t_{i=1})} \\ +\alpha_{12} \left(1 + \left(1 + e^{-\beta_1(t_{k+1}-t_{j=k})}\right) e^{-\beta_1(t_{k+2}-t_{j=k+1})}\right) e^{-\beta_1(u-t_1)} + ... \end{array} \right\} du + ...
\end{aligned}
$$

and similarly for λ_2. Below we plot the compensator for the simulated Hawkes process using the above formula

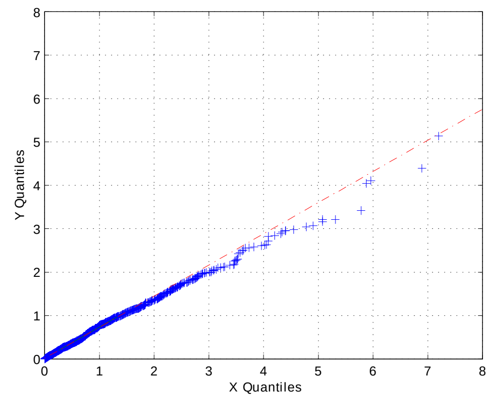


(Zoomed) Compenstor for simulated bivariate Hawkes process
(here shown for type 1 arrival)

By application of the time change theorem, we can again test the goodness of fit of the time-changed simulated process to that of a standard Poisson process. The QQ plot validate both our compensator and the simulation code.



QQ plot for simulated bivariate Hawkes process for arrival of type 1 event



QQ plot for simulated bivariate Hawkes process for arrival of type 2 event

2.5 Maximum likelihood estimation

The log-likelihood function for our bivariate process can be written [5]

$$L_T\left(\mu_1,\mu_2,\beta_1,\beta_2,\alpha_{11},\alpha_{12},\alpha_{21},\alpha_{22}\right) = L_T^{(1)}\left(\mu_1,\beta_1,\alpha_{11},\alpha_{12}\right) + L_T^{(2)}\left(\mu_2,\beta_2,\alpha_{21},\alpha_{22}\right)$$

The first term of RHS can be expressed as

$$\begin{aligned} L_T^{(1)}\left(\mu_1,\beta_1,\alpha_{11},\alpha_{12}\right) &= -\int_0^T \lambda_1(t)\,dt + \int_0^T \log\lambda_1(t)\,dN_1(t) \\ &= -\int_0^T \left(\mu_1 + \sum_{t_i<t}\alpha_{11}e^{-\beta_1(t-t_i)} + \sum_{t_j<t}\alpha_{12}e^{-\beta_1(t-t_j)}\right)dt \\ &\quad + \int_0^T \log\left(\mu_1 + \sum_{t_i<t}\alpha_{11}e^{-\beta_1(t-t_i)} + \sum_{t_j<t}\alpha_{12}e^{-\beta_1(t-t_j)}\right)dN_1(t) \end{aligned}$$

Therefore we have (See Appendix B)

$$\begin{aligned} L_T^{(1)}\left(\mu_1,\beta_1,\alpha_{11},\alpha_{12}\right) &= -\mu_1 T - \frac{\alpha_{11}}{\beta_1}\sum_{i=1}^{n}\left(1-e^{-\beta_1(T-t_i)}\right) - \frac{\alpha_{12}}{\beta_1}\sum_{j=1}^{m}\left(1-e^{-\beta_1(T-t_j)}\right) \\ &\quad + \sum_{i=2}^{n}\log\left(\mu_1 + \alpha_{11}R_{11}(i) + \alpha_{12}R_{12}(i)\right) \end{aligned}$$

Similarly, we also have

$$\begin{aligned} L_T^{(2)}\left(\mu_2,\beta_2,\alpha_{21},\alpha_{22}\right) &= -\mu_2 T - \frac{\alpha_{21}}{\beta_2}\sum_{i=1}^{n}\left(1-e^{-\beta_2(T-t_i)}\right) - \frac{\alpha_{22}}{\beta_2}\sum_{j=1}^{m}\left(1-e^{-\beta_1(T-t_j)}\right) \\ &\quad + \sum_{j=2}^{m}\log\left(\mu_2 + \alpha_{21}R_{21}(j) + \alpha_{22}R_{22}(j)\right) \end{aligned}$$

where

$$R_{22}(j) = e^{-\beta_2(t_j-t_{j-1})}\left(1+R_{22}(j-1)\right)$$

$$R_{21}(j) = e^{-\beta_2(t_j-t_{j-1})}\left(R_{21}(j-1)\right) + \sum_{\{i':t_{j-1}\le t_{i'}<t_j\}} e^{-\beta_2(t_j-t_{i'})}$$

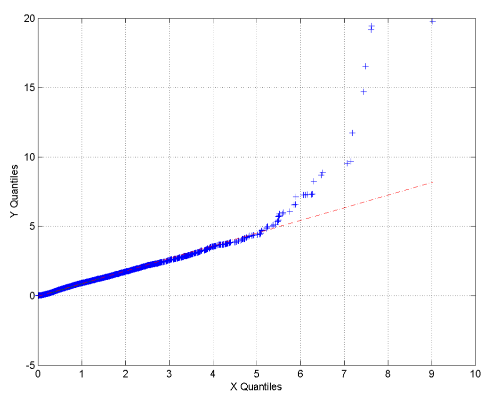
To validate our MLE procedure, we simulate 10,000 time units of the bivariate process. We then feed the resulting arrival times into the MLE program, run using a suitable optimization routine. The following table shows that our MLE routine converges to the true parameters.

parameter	True	Estimate
μ_1	0.3	0.3011
μ_2	0.1	0.0998
β_1	1.2	1.2261
β_2	1.0	1.0547
α_{11}	0.6	0.6006
α_{12}	0.9	0.9266
α_{21}	0.2	0.2089
α_{22}	0.5	0.5377

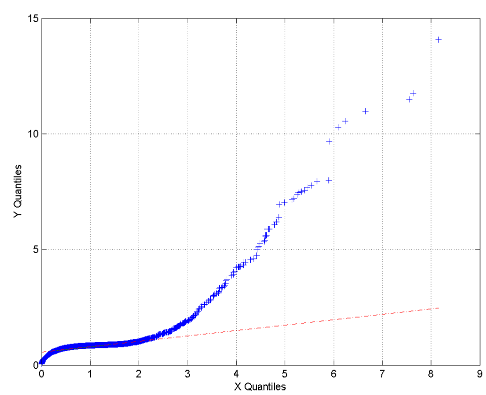
3 Empirical Model Fitting

3.1 Data Classification and Cleaning

Data from TAQ database have two major deficiencies, to which we have to find work around in order to minimize impact on model estimation. The first deficiency is that recorded trade times are descretised in whole seconds, which means that multiple trades within the same second share the same timestamp. One solution we proposed here is to redistribute trades with same timestamps uniformly between recorded. The second deficiency stems from the fact that trades do not come classified into buy and sell orders, only trade prices and volumes are recorded. Following the classic approach to rectify this problem, we first used the Lee and Ready tick test [4] to classify our data. One alternative solution we proposed here is to only use and classify orders that lead to an actual change in traded price, this "thinned classification" seem to yield better model fit, as shown below for DELL



QQ plot for DELL based on thinned classification



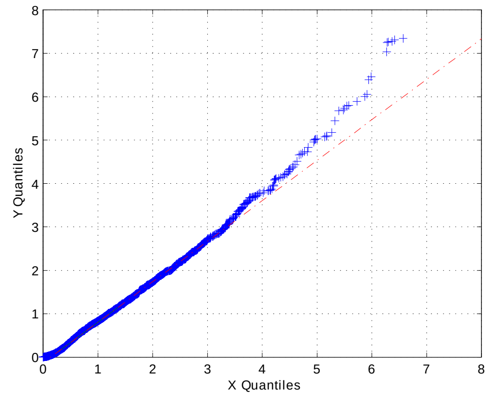
QQ plot for DELL based on Lee&Ready classification

To account for intra-day seasonality, where there is a noticeable change in base intensity at the open and the close, we can either fit a time varying base intensity or simply truncate the data to only look at trades that happened one hour after the open and one hour before the close. Here we adopt the latter approach.

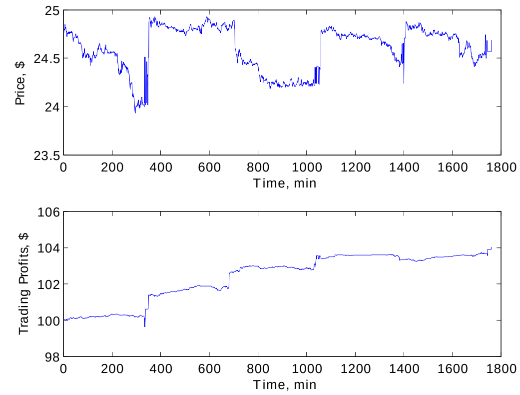
4 Example Trading Strategy

4.1 Simple Buy Sell Signal Based on Intensity Ratio[3]

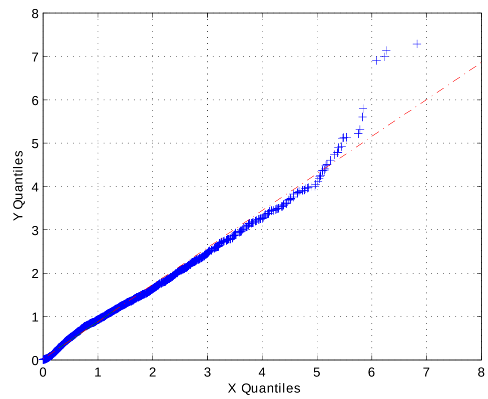
We tested, on a number of stocks, the naive strategy of holding long one share of the stock if the ratio of the buy vs sell intensity reaches a threshold of 8 and shorting one share of the stock if the ratio drops below 1/8, hold the position for 10 seconds, then liquidate. Below, we plot the stock price together with trading P&L using this strategy



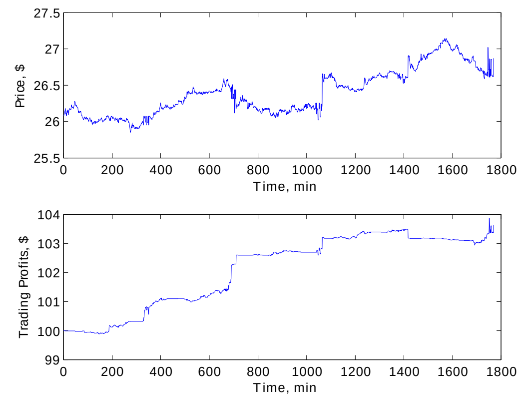
QQ-plot for DELL with MLE parameters



DELL trading performance from 1-Nov-2006 to 7-Nov-2006

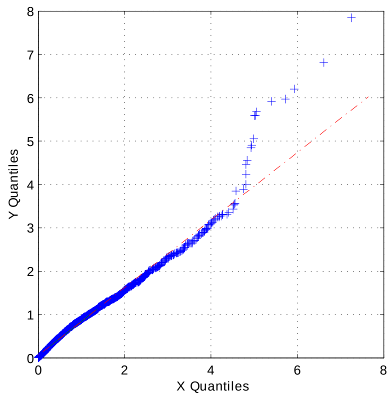


QQ-plot for YHOO with MLE parameters

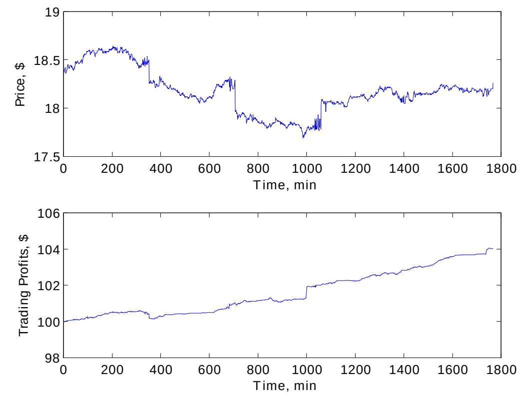


YHOO trading performance from 1-Nov-2006 to 7-Nov-2006

[3] Note this strategy is based on thinned classification



QQ-plot for ORCL with MLE parameters



ORCL trading performance from 1-Nov-2006 to 7-Nov-2006

The table below summarizes this simple trading strategy. We see that for the three stocks we investigated, all showed significant gain vs. a simple buy and hold strategy.

	YHOO	DELL	ORCL
Total Number of trades	2765	3569	3843
Number of buys	1535	2262	1856
Number of sells	1230	1327	1987
Buy and hold return	2.00%	-0.57%	-0.70%
Strategy return	3.39%	4.03%	4.02%

5 Conclusion

In this work, we looked at the goodness of fit of a bivariate Hawkes model to classified tick-data from TAQ database for a number of liquid stocks. We have shown that, at least for the names we studied, the model seems to describe the underlying buy and sell order arrival times well. Then based on the MLE fitted model, we tested a naive trading strategy which showed significant return compared to just a simple buy and hold strategy.

6 Appendix

6.1 Appendix A: Univariate Hawkes Process

$$d\lambda_t = \kappa\left(\rho(t) - \lambda_t\right)dt + \delta dN_t$$

The solution for λ_t takes the form

$$\lambda_t = c(t) + \int_0^t \delta e^{-\kappa(t-u)} dN_u$$

where

$$c(t) = c(0)\, e^{-\kappa t} + \kappa \int_0^t e^{-\kappa(t-u)} \rho(u)\, du$$

Verify by Ito formula on $e^{\kappa t}\lambda_t$

$$\begin{aligned} e^{\kappa t}\lambda_t &= c(0) + \kappa \int_0^t e^{\kappa u} \rho(u)\, du + \int_0^t \delta e^{\kappa u} dN_u \\ \kappa e^{\kappa t}\lambda_t dt + e^{\kappa t} d\lambda_t &= \kappa e^{\kappa t} \rho(t)\, dt + \delta e^{\kappa t} dN_t \\ \kappa \lambda_t dt + d\lambda_t &= \kappa \rho(t)\, dt + \delta dN_t \\ d\lambda_t &= \kappa\left(\rho(t) - \lambda_t\right) dt + \delta dN_t \end{aligned}$$

consider the limit $\lim_{t\longrightarrow\infty} c(t)$

$$\begin{aligned}
\lim_{t\to\infty} c(t) &= \lim_{t\to\infty}\left\{c(0)\,e^{-\kappa t}+\varkappa\int_0^t e^{-\kappa(t-u)}\rho(u)\,du\right\}\\
&= \lim_{t\to\infty}\kappa\int_0^t e^{-\kappa(t-u)}\rho(u)\,du\\
&= \lim_{t\to\infty}\kappa e^{-\kappa t}\int_0^t e^{\kappa u}\rho(u)\,du\\
&= \lim_{t\to\infty}\kappa\frac{\int_0^t e^{\kappa u}\rho(u)\,du}{e^{\kappa t}}\\
\text{(apply L'Hospital)} &= \lim_{t\to\infty}\kappa\frac{e^{\kappa t}\rho(t)}{\kappa e^{\kappa t}}\\
&= \lim_{t\to\infty}\rho(t)\\
&= \lambda_\infty
\end{aligned}$$

Treating $\rho(t)$ as a constant $\rho(t)=\lambda_\infty$, then we have

$$\begin{aligned}
c(t) &= c(0)\,e^{-\kappa t}+\kappa\int_0^t e^{-\kappa(t-u)}\rho(u)\,du\\
&= c(0)\,e^{-\kappa t}+\kappa\lambda_\infty e^{-\kappa t}\int_0^t e^{\kappa u}du\\
&= c(0)\,e^{-\kappa t}+\lambda_\infty e^{-\kappa t}\left(e^{\kappa t}-1\right)\\
&= \lambda_\infty+e^{-\kappa t}\left(c(0)-\lambda_\infty\right)
\end{aligned}$$

Notice that if we set $c(0)=\lambda_\infty$ then the process is simply

$$\lambda_t=\lambda_\infty+\delta\int_0^t e^{-\kappa(t-u)}dN_u$$

where we can think of λ_∞ as the long run "base" intensity, i.e. the intensity if there have been no past arrival.

6.2 Appendix B: Bivariate MLE

Since the parameters are bounded, so by Fubini's theorem we have

$$\begin{aligned}
L_T^{(1)}(\mu_1,\beta_1,\alpha_{11},\alpha_{12}) &= -\left(\int_0^T\mu_1 dt+\sum_{t_i<t}\int_0^T\alpha_{11}e^{-\beta_1(t-t_i)}dt+\sum_{t_j<t}\int_0^T\alpha_{12}e^{-\beta_1(t-t_j)}dt\right)\\
&\quad+\int_0^T\log\left(\mu_1+\sum_{t_i<t}\alpha_{11}e^{-\beta_1(t-t_i)}+\sum_{t_j<t}\alpha_{12}e^{-\beta_1(t-t_j)}\right)dN_1(t)\\
&= -\mu_1 T-\frac{\alpha_{11}}{\beta_1}\sum_{i=1}^n\left(1-e^{-\beta_1(T-t_i)}\right)-\frac{\alpha_{12}}{\beta_1}\sum_{j=1}^m\left(1-e^{-\beta_1(T-t_j)}\right)\\
&\quad+\sum_{i=2}^n\log\left(\mu_1+\alpha_{11}\sum_{i'=1}^i e^{-\beta_1(t_i-t_{i'})}+\alpha_{12}\sum_{j'=1}^i e^{-\beta_1\left(t_i-t_{j'}\right)}\right)
\end{aligned}$$

We can recursively express

$$
\begin{aligned}
R_{11}(i) &= \sum_{i'=1}^{i} e^{-\beta_1 (t_i - t_{i'})} \\
&= e^{-\beta_1(t_i-t_0)} + e^{-\beta_1(t_i-t_1)} + ... + e^{-\beta_1(t_i-t_{i-1})} \\
&= e^{-\beta_1(t_i-t_{i-1})} e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_1-t_0)} + e^{-\beta_1(t_i-t_{i-1})} e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_2-t_1)} + ... + e^{-\beta_1(t_i-t_{i-1})} \\
&= e^{-\beta_1(t_i-t_{i-1})} \left(e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_1-t_0)} + e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_2-t_1)} + ... + 1 \right) \\
&= e^{-\beta_1(t_i-t_{i-1})} \left(1 + \sum_{i'=1}^{i-1} e^{-\beta_1(t-t_{i'})} \right) \\
&= e^{-\beta_1(t_i-t_{i-1})} \left(1 + R_{11}(i-1) \right)
\end{aligned}
$$

Now let $j^* = \sup \{ j' : t_{j'} < t_i \}$, again we can recursively express

$$
\begin{aligned}
R_{12}(i) &= \sum_{j'=1}^{i} e^{-\beta_1 \left(t_i - t_{j'}\right)} \\
&= e^{-\beta_1(t_i-t_0)} + e^{-\beta_1(t_i-t_1)} + ... + e^{-\beta_1\left(t_i-t_{j^*-1}\right)} + e^{-\beta_1\left(t_i-t_{j^*}\right)} \\
&= e^{-\beta_1(t_i-t_0)} + e^{-\beta_1(t_i-t_1)} + ... + e^{-\beta_1\left(t_i-t_{j^*-1}\right)} + \sum_{\left\{j' : t_{i-1} \le t_{j'} < t_i\right\}} e^{-\beta_1\left(t_i-t_{j'}\right)} \\
&= e^{-\beta_1(t_i-t_{i-1})} e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_1-t_0)} + e^{-\beta_1(t_i-t_{i-1})} e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_2-t_1)} \\
&\quad + e^{-\beta_1(t_i-t_{i-1})} e^{-\beta_1\left(t_{i-1}-t_{j^*-1}\right)} + \sum_{\left\{j' : t_{i-1} \le t_{j'} < t_i\right\}} e^{-\beta_1\left(t_i-t_{j'}\right)} \\
&= e^{-\beta_1(t_i-t_{i-1})} \left(e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_1-t_0)} + e^{-\beta_1(t_{i-1}-t_{i-2})} ... e^{-\beta_1(t_2-t_1)} + e^{-\beta_1\left(t_{i-1}-t_{j^*-1}\right)} \right) \\
&\quad + \sum_{\left\{j' : t_{i-1} \le t_{j'} < t_i\right\}} e^{-\beta_1\left(t_i-t_{j'}\right)} \\
&= e^{-\beta_1(t_i-t_{i-1})} \left(R_{12}(i-1) \right) + \sum_{\left\{j' : t_{i-1} \le t_{j'} < t_i\right\}} e^{-\beta_1\left(t_i-t_{j'}\right)}
\end{aligned}
$$

References

[1] Giesecke, Kay. *Credit Risk Modeling and Management.* Lecture Note Part II: 46-48, Stanford University, 2007.

[2] Duffie, Darrell. *Credit Risk Modeling with Affine Processes.* 2002 Cattedra Galileana Lectures (revised: June 2004)

[3] Hewlett, Patrick. Clustering of Order Arrivals, Price Impact and Trade Path Optimisation, Internal Report, Oxford University, May 6, 2006

[4] Lee, Charles M. C., Makr J. Ready. Inferring Trade Direction From Intraday Data, The Journal of Finance, Vol. 46, No. 2. (Jun., 1991), pp. 733-746

[5] Ogata, Yosihiko. The Asymptotic Behavour of Maximum Likelihood Estimators for Stationary Point Processes, Ann, Inst, Sttist. Math. 30 (1978), Part A, 243-261

[6] Ogata, Yosihiko, On Lewis' Simulation Method for Point Processes, IEEE Transactions on Information Theory, Vol. IT-27, No. 1, January 1981