2017

$$\hat{f}_b(s) = \left(-\frac{1}{\lambda}\right)^b \left(\prod_{i=1}^{b} \Phi_{k=i}^{\infty} \frac{-\lambda d_k}{\lambda + d_k + s}\right)$$

$$f(t) \approx f(t,M) = \frac{\sigma}{M}\left(\frac{1}{2}\hat{f}(\sigma)e^{\sigma t} + \sum_{k=1}^{M-1} Re\left(e^{ts\theta_k}\hat{f}(s(\theta_k))(1+iw(\theta_k))\right)\right)$$

Stochastic Model of Limit Order Books

RAJAN SUBRAMANIAN

1.0 Introduction

In this paper, we will examine the article “A stochastic model for order book dynamics” by Rama Cont,Sasha Stoikov & Rishi Talreja (henceforth referred to as CST model),February 24, 2010 and attempt to replicate the model that is suggested. The CST model is part of a larger class of Zero-Intelligence (ZI) models and extends the basic premise of the ZI models by assuming a power-law distribution for limit order sizes while at the same time recognizing the Markovian nature of the limit Order Books and utilizes this framework to calculate conditional probabilities.

The Zero-Intelligence (ZI) Model is a statistical model that helps analyze a Continuous Double Auction under assumption of Poisson iid order flows. The most well known ZI model is the SFGK (2002) model that models the behavior NYSE and LSE SETS systems very well (close to best quotes). The model uses dimensional analysis to make predictions about basic properties of the market like price volatility, depth of stored supply & demand vs price (resulting book-shape), bid-ask spread and price Impact function . The disadvantage of using this model however is that deep in the book, the approximation fails and it cannot incorporate strategic behavior of traders. Another disadvantage of using this model is the resulting price signal that slopes upwards on average.

The CST model is a stochastic order placement model that essentially extends the ZI model by incorporating conditional properties on the state of the order book. Like the SFGK, it assumes rates at which orders arrives and cancels are Poisson distributed. It departs from the SFGK in that it assumes a queuing system for limit orders(LOs) where the number of LOs per price level that are awaiting execution is modeled as a Continuous Time Markov Chain (CTMC). It helps explain the behavior of TSE very well. The advantage of this approach is that Laplace Transform techniques can be used to compute various conditional probabilities such as the probability of mid-price to go up vs. down, probability of executing LOs at bid before ask quotes move, probability of executing both buy and sell order at best quotes before price moves. Hence, by passing the need for Monte-Carlo Simulation. The disadvantage of using this approach is that it assumes a constant order size of one when in reality traders can send orders in sizes of multiples of lot size. Another disadvantage is that the model still can’t incorporate strategic interaction among traders nor take into account long memory of order-flow

We will compute the quantities of interest to traders such as the shape of the order-book and conditional probabilities. In addition, Monte-Carlo Simulation is performed to verify the accuracy of the model. By replicating this model, we will try to get a better understanding of the interplay between order flow, liquidity and price dynamics and hopefully use the current state of the order book to predict future behavior.

2.0 Description of Hypothesis

The CST model can help capture the dynamics of a order book for a particular stock close to best quotes very well and should bypass the need to use Monte-Carlo simulations for the Tokyo Stock Exchange

3.0 Literature Review

There are two main approaches to order-book modeling:

-Agent-based approach: traders condition trades based on current state of order-book and consider strategic interactions among another trader.

-Zero-Intelligence approach: stochastic models that assume arrival rates of incoming/cancellations follow some stochastic process.

Traditional limit order book modeling starts with empirical features of the market. Gould et al (2011) describes how different empirical studies in limit order books across exchanges offer conflicting conclusions. However, he notes that there are stylized facts that are common amongst the many exchanges such as power-law distribution of order-sizes and relative prices, the hump shape of the depth profile, long memory in order flow etc. Details are in [3]

Seppi, Parlour, Foucault, & Rosu consider strategic trader models where traders condition their trade based on the book. These models are however hard to estimate and use in applications.

Smith et al (2002) suggest arrival of limit orders, cancellations and market orders follow a independent Poisson process with constant intensities. Dimension analysis is then used to make a forecast of various parameters. Cont et al (2010) extend Smith's model by incorporating empirical features of the order flow by fitting the parameters of arrival limit order intensities through a power law function. The authors then use Birth Death process to make predictions of probabilities that are interesting to traders. This approach however fails to take any clustering effect of order flows

More recently, Large(2007b) focuses on calibrating real data to a Hawkes process-based model. The Hawkes process has the advantage that it takes into account "time clustering" which can reproduce the fact that order arrivals alternate between bursting and quiet periods. Hawkes process also exhibit the property of "mutual excitation" which can reproduce the fact that order flows exhibit non-negligible cross-dependencies.

The rest of this paper attempts to replicate Cont-Stoikov Talreja model using the data supplemented by the authors. Our contribution is as follows:

- An algorithm to compute the continued fractions in Cont et al's paper using modified Lentz's algorithm
- Computation of Inverse Laplace using Fixed-Talbot Algorithm

4.0 CST Model

Statistical model that helps analyze CDA under assumption of Poisson order flows. Makes predictions about price volatility, depth, bid/ask spread, price Impact function and probability of filling orders.

The model

- LBOs are placed with uniform probability in range $(-\infty, A_t]$ and LSOs are placed in $[B_t, \infty)$
- Limit Orders can be placed inside spread
- Market Orders arrive at chunks of σ shares at rate of μ per unit time
- Limit Orders arrive at chunks of σ shares, a distance of L ticks (L≥ 1) from opposite best quote at rate of $\lambda_p(L) = kL^{-\alpha}$ shares per unit price, per unit time.
- Cancellations existing limit orders is proportional to number of shares at that level. If number of shares at a price level is x, then cancellation rate is θx per unit time

- Above events occur at independent exponential times.
- Assume $\sigma = 1$. Hence, at any given time, only 1 share can arrive or leave.

Market Dimensional Analysis

Using dimension analysis, we can guess basic properties of the limit Order Book using our simple model above:

Parameter	**Description**	**Dimensions**
λ	Limit Order Rate	$\frac{shares}{dp * T}$
μ	Market Order Rate	$\frac{shares}{T}$
θ	Cancellation Rate	$\frac{1}{T}$
dp	Tick Size	$price$
σ	Characteristic Order Size	shares

- There are three order-flow parameters and 3 fundamental dimensions (shares,price and time)
- Discrete parameters are tick size dp and σ

- Ignore the Discrete parameters and invert the above table to solve for 3 fundamental dimensions:

Dimensions	**Parameter Combinations**
Shares	$\frac{\mu}{\theta}$
T	$\frac{1}{\theta}$
Ticks	$\frac{\mu}{\lambda}$

➢ Note, dimensions of inversion table is called characteristic dimensions since its characteristic of dimensional analysis, i.e, shares means characteristic shares := N_C, characteristic time:= t_C, characteristic price interval or ticks:= p_C.

Non-Dimensional Parameters

Since σ is the lot size, measured in shares, non-dimensional scale parameter based on order size is constructed by dividing order size by characteristic shares, i.e:

1) $\epsilon = \frac{\sigma}{N_c} = \frac{2\theta\sigma}{\mu}$

- ε represents chunkiness of orders stored in the LOB. If $\varepsilon \rightarrow 0$, price diffusion cannot occur since there are no shares at that level. Its unrealistic assumption & poor approximation to market. This is because if we assume a fixed depth and $\varepsilon \rightarrow 0$, this happens if market orders arrive too fast and suppose cancellations tend to 0, then # of individual orders would be infinite. A contradiction!

Note, in non-dimensional terms, # of shares can be expressed as $\widehat{N} := \frac{n}{N_c}$, and using 1), we get:

2) $\widehat{N} := \frac{n\epsilon}{\sigma}$

• If $\theta = 0$, which implies ε = 0, then there is no decay of orders and depending on σ or μ, either depth will accumulate without bound or spread will become infinite. As long as $\theta > 0$, i.e as long as we allow cancellations to occur,this is not a problem.

Now, we will use dimensional scaling relationships to determine various properties of the market which does depend on order flow rates

Spread

Inside spread, LOs are removed due to incoming market orders or cancellation. In a steady state situation, total arrival rate of limit orders is $2s\lambda dt$ which must balance arrival rate of market orders and limit orders leaving due to cancellation $2(\mu + \theta x)dt$. Hence we must have $2s\lambda dt = 2(\mu + \theta x)dt$ and solving for s, we find:

3) $s \sim \frac{\mu+\theta x}{\lambda}$

• inside spread, when $\theta = 0$, we see this reduces to the characteristic price interval p_C which is exact.

Variance & Volatility

Variance has dimensions of $\frac{Tick^2}{T}$. According to the table above, this implies $\frac{\mu^2}{\lambda^2} * \theta$, i.e

4) $v^2 \sim \frac{\mu^2}{\lambda^2} * \theta$

Hence, volatility must scale as. . .

5) $v \sim \frac{\mu}{\lambda} * \sqrt{\theta}$

Asymptotic Depth

Refers to depth of shares deep in the books. We see that asymptotic depth has dimensions of shares/tick, hence, using above inversion table we see:

6) $d \sim \frac{\mu}{\theta} * \frac{\lambda}{\mu} = \frac{\lambda}{\theta}$

➢ Intuitively, deep in the books, LOs mostly are cancelled. Execution due to market orders is low. Hence total arrival rate of LO is $2\lambda dt$ and departure rate of LOs being cancelled is $2d\theta dt$, where d is number of shares at that price level. For these to balance, we must have $\lambda dt = d\theta dt$, or solving for d, we get. . ., $d = \frac{\lambda}{\theta}$

Slope of Order Book

This is the depth profile near mid-point. Its an important determinant of liquidity since it effects price response when MO arrives. This has dimensions $\frac{shares}{price^2}$ i.e derivative wrt price of the book density. Now, since $N_c = \frac{\mu}{\theta}, price^2 = \frac{\mu^2}{\lambda^2}, \rightarrow \frac{N_C}{price^2} \sim \frac{\lambda^2}{\mu\theta}$. Hence, slope scales as…

7) $\xi \sim \frac{\lambda^2}{\mu\theta} = \frac{d\lambda}{\mu} = \frac{d}{s}$

Hence we see, the slope is roughly asymptotic depth to spread ratio

Market Impact

provides measure of liquidity for executing market orders. For a given number of shares n, we want to see how much impact buying/selling will have on price. Price has dimensions of ticks. This has parameter combination $\frac{\mu}{\lambda}$. Hence a general form of market impact is:

8) $I(n) \sim \frac{\mu}{\lambda} f\left(\frac{n}{N_c}\right) = \frac{\mu}{\lambda} f\left(\frac{2\theta n}{\mu}\right)$

for some function $f(\cdot)$

Quantity at given Price Level as a Birth-Death Process

WLOG, lets focus on the ask as the bid side is symmetric. Let X_t^a represent the number of orders at ask at a particular price level at time t. Over time $[t, t + dt]$, only one event can take place. Hence X_t^a makes a transition:

$X_{t+1}^a \rightarrow X_t^a + 1$ if limit Order occurs w/ $P(birth) = \frac{\lambda}{2\mu+2\lambda+\theta x_t^a+\theta x_t^b}$

$X_{t+1}^a \rightarrow X_t^a - 1$ if market or Cancel occurs w/ P(death) = $\frac{\mu+\theta x^a}{2\mu+2\lambda+\theta x_t^a+\theta x_t^b}$

Hence, X_t is a Birth-Death Process which is a CTMC with state space $\{0,1,2 \dots\}$ where transition from state i (there are i orders at time t) can go to state $j = i + 1$ or $j = i - 1$. Hence for any price level, quantities in limit order book behave like this:

- Quantity remains in state i for a period of time exponentially distributed w/mean $-\frac{1}{2\mu+2\lambda+\theta i+\theta x_t^b}$
- It then jumps to another state. This state, is state j. Hence, if we are in state i, probability of going to j = i+1 occurs if birth occurs before death. This is same as exponential random variable with birth rate λ occurs before independent exponential random variable with death rate μ or cancellation occurring. This is same as $\frac{\lambda}{2\mu+2\lambda+\theta i+\theta x_t^b}$
- It then stays at this state j for a period of time exponentially distributed with mean $-\frac{1}{2\mu+2\lambda+\theta j+\theta x_t^b}$ and so on…

4.2 Laplace Transformation (LT)

We want to predict short term dynamics of 3 quantities of interest which is useful for traders. The three probabilities of interest are

a) Probability of mid-price increase
b) Probability of executing my order before mid-price moves (spread = 1)
c) Probability of executing both my orders before mid-price moves (spread = 1)

These quantities can be obtained by computing the Inverse Laplace Transform (LT) of the probability distribution. Monte-Carlo simulations of these probabilities are computed to verify the accuracy of the model

Laplace Transforms of Birth-Death Process

Consider a birth-death process with birth rate λ and death rate $d_i = \mu + \theta_i$ at state $i, i \geq 1$. Let τ_b denote the first passage time of this process to state 0 given it starts at state b. The first passage time can be written as:

$$\tau_b = \tau_{b,b-1} + \tau_{b-1,b-2} + \ldots + \tau_{1,0}$$

where $\tau_{i,i-1}$ is first passage time of this process from state i to $i-1$. Let $\hat{f}_b$ be Laplace Transform of τ_b and $\hat{f}_{i,i-1}$ be Laplace Transform of $\tau_{i,i-1}$. Now for $i = 1, \ldots, b$, since above first passage times are independent, we have…:

$$\hat{f}_b(s) = \prod_{i=1}^{b} \hat{f}_{i,i-1}(s)$$

To compute single Laplace transforms, recall that $\hat{f}(s) = \int_{-\infty}^{\infty} e^{-st} f(t)$ is the Laplace transform of function $f(t)$. When $f(t)$ is a pdf of a random variable X, we can also rewrite this as: $\hat{f}(s) = E(e^{-sX})$. Hence:

$$\hat{f}_{i,i-1}(s) = E(e^{-s\tau_{i,i-1}})$$

Now, given we are in state i, two events can happen next. Either a birth occurs first before death, $T_\lambda < T_d$ in which case we go to state $i+1$ and need to calculate remaining time to go to state $i-1$, or a death occurs before birth $T_d < T_\lambda$ in which case we go to state $i-1$ directly from state i. Hence:

$$\begin{cases} t_{i,i-1} = T_\lambda + t_{i+1,i-1} & if \min(T_\lambda, T_{d_i}) = T_\lambda \\ t_{i,i-1} = T_{d_i} & if \min(T_\lambda, T_{d_i}) = T_{d_i} \end{cases}$$

Plugging the formula into the expectation above we see…

$$E(e^{-s\tau_{i,i-1}}) = E\left(e^{-s(T_\lambda+\tau_{i+1,i-1})} I_{T_\lambda<T_{d_i}}\right) + E(e^{-sT_d} I_{T_{d_i}<T_\lambda})$$

The first expectation can be rewritten as…

$$E\left(e^{-s(T_\lambda+\tau_{i+1,i-1})} I_{T_\lambda<T_{d_i}}\right) = E\left(e^{-s(T_\lambda+\tau_{i+1,i}+\tau_{i,i-1})} I_{T_\lambda<T_{d_i}}\right)$$

$$= E\left(e^{-(sT_\lambda)} I_{T_\lambda<T_{d_i}}\right) \hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s) = \frac{\lambda}{s+\lambda+d_i} \hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s)$$

since the join density $g_{T_\lambda,T_{d_i}} = \lambda e^{-\lambda t} d_i e^{-d_i t}$ is independent.

Similarly, the second expression is given by…

$$E\left(e^{-sT_{d_i}} I_{T_{d_i}<T_\lambda}\right) = \frac{d_i}{s+d_i+\lambda}$$

Combining above results, we see…

$$\hat{f}_{i,i-1}(s) = \frac{\lambda}{s+\lambda+d_i} \hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s) + \frac{d_i}{s+d_i+\lambda}$$

Rewriting above we get…

$$\hat{f}_{i,i-1}(s) \left(1 - \frac{\lambda \hat{f}_{i+1,i}(s)}{s+\lambda+d_i}\right) = \frac{d_i}{s+d_i+\lambda}$$

Solving for $\hat{f}_{i,i-1}(s)$, we obtain:

$$\hat{f}_{i,i-1}(s) = \frac{d_i}{s+\lambda+d_i-\lambda\hat{f}_{i+1,i}(s)}$$

Iterating the above expression, we get the continued fraction of the form:

$$\hat{f}_{i,i-1}(s) = -\frac{1}{\lambda} \Phi_{\mathrm{k=i}}^{\infty} \frac{-\lambda d_k}{\lambda+d_k+s}$$

Now, since $\hat{f}_b(s) = \prod_{i=1}^{b} \hat{f}_{i,i-1}(s)$, we get using above, laplace transform from state b to state 0:

$$\hat{f}_b(s) = \left(-\frac{1}{\lambda}\right)^b \left(\prod_{i=1}^{b} \Phi_{\mathrm{k=i}}^{\infty} \frac{-\lambda d_k}{\lambda+d_k+s}\right)$$

where $d_k = \mu + k\theta$

Probability of Mid-Price Increase (spread = 1)

The mid-price increases when the quantity at the ask goes to 0 before quantity at bid goes to 0. Let $X_A \equiv X_A(t=0)$ and $X_B \equiv X_B(t=0)$ be independent BD process that denotes the number of shares at the ask and bid respectively at time 0. Let T_A be first time X_A goes to 0 and T_B be first time X_B goes to 0. We need $P(T_A < T_B)$.

Let $T = T_A - T_B$ and define $G_T(t) := P(T \leq t)$ as CDF of T.

Now, Laplace Transform of pdf $g_T(t)$ of random variable T is given by $\hat{g}_T(s)$ and this is…

$$\hat{g}_T(s) = E(e^{-sT}) = E\left(e^{-s(T_A - T_B)}\right) = \hat{f}_A(s)\hat{f}_B(-s)$$

Define $\hat{G}_t(s) := \int_{-\infty}^{\infty} e^{-st} G_T(t)dt$ as Laplace transform of $G_T(t)$. We can solve this integral by

using integration by parts $\int u dv = uv - \int v du$ where

$$dv = e^{-st}dt\,, \qquad u = G(t)\,, \quad v = -\frac{1}{s}e^{-st}\,, \qquad du = g(t)dt$$

Substituting above into our equation, we see…

$$\int_{-\infty}^{\infty} e^{-st} G_T(t)dt = \left(-\frac{G_T(t)}{s}e^{-st}\right)\Bigg|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{s} e^{-st} g_T(t)dt = \frac{\hat{g}_T(s)}{s}$$

Hence the Laplace transform of CDF $G_T(t)$ is given by…

$$\hat{G}_t(s) = \frac{\hat{f}_A(s)\hat{f}_B(-s)}{s}$$

Therefore, taking Inverse Laplace transform of above equation will recover $G_T(t)$.

We can then calculate $G_T(0) = P(T_A - T_B < 0) = P(T_A < T_B)$

Probability my order at the bid is executed before mid-price moves (Spread = 1)

Let W_A be the number of orders remaining at time t of initial $X_A(0)$ shares and W_B be defined similarly. My order at the bid is at the last position in W_B. Let ϵ_A (ϵ_B) be defined as the first time W_A (W_B) goes to 0. The mid-price changes before time ϵ_B if and only if event $\{T_A < \epsilon_B\}$ occurs. We need $P(\epsilon_B < T_A)$

Note, that since ϵ_B is first passage time of pure death process to 0. Using equation ( ), we see by letting $\lambda = 0$, the Laplace transform of ϵ_B:

$$\hat{f}_{W_B}(s) = \prod_{i=1}^{W_B} \hat{f}_{i,i-1}(s) = \prod_{i=1}^{W_B} \frac{d_{i-1}}{d_{i-1} + s}$$

Let $T = \epsilon_B - T_A$ and define $G_T(t) := P(T \leq t)$ be the CDF of T. Using same argument as previous section, we get the Laplace transform of CDF of $G_T(t)$:

$$\hat{G}_t(s) = \frac{\hat{f}_{W_B}(s)\hat{f}_A(-s)}{s}$$

Hence, taking the inverse Laplace of above recovers $G_T(t)$. We can then calculate $G_T(0)$

Probability of making the Spread (Spread = 1)

Let W_A be the number of orders remaining at time t of initial X_A(t=0) shares and W_B be defined similarly. My order at the bid is at the last position in W_B and last position in W_A. Let ϵ_A (ϵ_B) be defined as the first time W_A (W_B) goes to 0. We find probability that both our orders are executed before mid-price moves given it is not cancelled. This can be represented as:

$$P(\max(\epsilon_B, \epsilon_A) < \min(T_B, T_A)) = P(\epsilon_A < T_B, \epsilon_B < \epsilon_A) + P(\epsilon_B < T_A, \epsilon_A < \epsilon_B)$$

Let $h_{a,b} := P(\epsilon_A < T_B, \epsilon_B < \epsilon_A) = P(\epsilon_B < \epsilon_A < T_B)$ be defined as the probability the order placed at the bid is executed before order placed at the ask and the order placed at the ask is executed before bid process goes to 0. Conditioning on ϵ_B, we obtain:

$$h_{a,b} = \int_0^\infty P(\epsilon_B < \epsilon_A < T_B \mid \epsilon_B = t)\, f_{W_B}(t)dt$$

where $f_{W_B}(t)$ is the inverse Laplace Transform of $\hat{f}_{W_B}(s)$

The integrand in the above equation can be rewritten as:

$$P(\epsilon_B < \epsilon_A < T_B \mid \epsilon_B = t) = \sum_{i=1}^{\infty}\sum_{j=1}^{A} P(\epsilon_j < T_i)P(X_B(t) = i \mid \epsilon_B = t)P(W_A = j)$$

The authors estimate the quantity $P(X_B(t) = i \mid \epsilon_B = t)$ using a $MM\infty$ queue example and this is represented as…

$$P(X_B(t) = i \mid \epsilon_B = t) = \frac{e^{-\lambda^X(t)}\lambda^X(t)^i}{i!}, \quad \lambda^X(t) \equiv \frac{\lambda}{\theta}\left(1 - e^{-\theta t}\right)$$

and since W_A is a pure death process, its represented by an infinitesimal generator matrix:

$$P(W_A = j) = \left(e^{Q_A^W t}\right)_{A,j} \equiv \left(\sum_{k=1}^{\infty} \frac{t^k}{k!}\left(Q_A^W\right)^k\right)_{A,j}$$

Substituting everything above and using tonnelli's theorem, we see our probability can be represented as…

$P(\max(\epsilon_B, \epsilon_A) < \min(T_B, T_A)) = h_{a,b} + h_{b,a}$

$$h_{a,b} = \sum_{i=1}^{\infty} \sum_{j=1}^{A} \int_0^{\infty} P(\epsilon_j < T_i)\, P(X_B(t) = i \mid \epsilon_B = t) P(W_A = j) dt$$

5.0 Monte-Carlo Simulation of the Order Book

Assume orders can be placed at band of width L=30 around best prices. Define the following intensities for 1,2,3…,L ticks from opposite best quote:

$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_L)$ $\Lambda_\lambda = \sum_{i=1}^{L} \lambda_i$ $\theta(a) = (\theta_1 x_1^a, \theta_2 x_2^a, \dots, \theta_L x_L^a)$ $\theta(b) = (\theta_1 x_1^b, \theta_2 x_2^b, \dots, \theta_L x_L^b)$

$X^a = (x_1^a, x_2^a, \dots, x_L^a)$ $X^b = (x_1^b, x_2^b, \dots, x_L^b)$ $\Lambda_\theta(b) = \sum_{i=1}^{L} \theta_i * x_i^b$ $\Lambda_\theta(a) = \sum_{i=1}^{L} \theta_i * x_i^a$

$\Lambda(a, b) = 2(\mu + \Lambda_\lambda) + \Lambda_\theta(a) + \Lambda_\theta(b)$

Event Type:

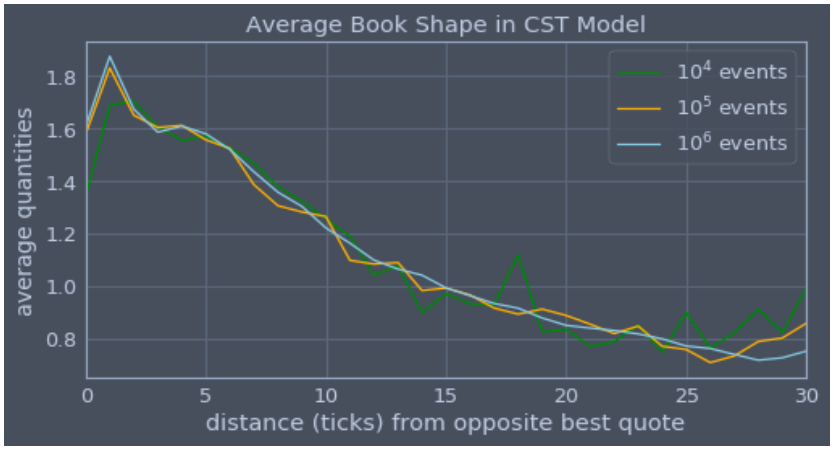
MB: Market Buy; MS: Market Sell; LB: Limit Buy; LS: Limit Sell; CB: Cancel Buy; CS: Cancel Sell

Algorithm: CST Order Book Simulation

1. **Initialization:**
    a. Initialize price levels from [-*l*,*l*], *l* = 1000 with increments of one tick (dp=1)
    b. Note: spread = 2 ticks. Choose L = 30, depth = 5 shares.
    c. Fit a power law function using non-linear least Squares on empirical limit order intensities for ticks 1:5. For $i > 5$, extrapolate using power-law parameters.
    d. Use empirical frequencies for cancellation intensities θ_i, $i = 1{:}5$ ticks. Let $\theta_i = \theta_5$ for $i > 5$ ticks
2. For $n = 1, \dots, N$: do
    a. Update Cancellation Intensities: $\Lambda_\theta(a) = \sum_{i=1}^{L} \theta_i * x_i^a$, $\Lambda_\theta(b) = \sum_{i=1}^{L} \theta_i * x_i^b$
    b. **Event Type:**
        I. Draw an event type {MB,MS,LB,LS,CB,CS} according to probability vector $\frac{(\mu, \mu, \Lambda_\lambda, \Lambda_\lambda, \Lambda_\theta(b), \Lambda_\theta(a))}{\Lambda(a,b)}$
    c. **Event Price:**
        I. If selected event is a market order, then draw from best quotes
        II. If selected event is a Limit order, then draw from relative price level $[1, 2, \dots, L]$ according to probability vector $(\lambda_1, \lambda_2, \dots, \lambda_L)/\Lambda_\lambda$
        III. If selected event is cancel order, then draw from relative price level $[1, 2, \dots, L]$ according to probability vector $\frac{X^a}{\Lambda_\theta(a)}$ for ask or $\frac{X^b}{\Lambda_\theta(b)}$ for bid
    d. Update order-book according to the new event.

5.1 CST Average Book Shape

We let our model run for $10^4, 10^5$ *and* 10^6 simulations. Graph below shows the order-book shape in the CST model for Sky-Perfect communications.



The above "steady state profile" of order book describes the average market impact of trading 1 share for various simulations.

-Note the hump shape (2 ticks from the best ask) observed in empirical studies.

5.2 Implementation of Analytical Probabilities

In order to compute the probabilities, we first need to compute the continued fraction to obtain the Laplace transform, then invert the Laplace transform and evaluate at $t = 0$. The steps are described below in the following algorithm:

Step 1: Compute the Laplace transform of first passage time: $\hat{f}_{i,i-1}(s) = -\frac{1}{\lambda} \Phi_{\mathrm{k=i}}^{\infty} \frac{-\lambda d_k}{\lambda + d_k + s}$ via modified Lentz algorithm

Step 2: Compute $\hat{f}_x(s) = \prod_{i=1}^{x} \hat{f}_{i,i-1}(s)$, to obtain Laplace transform from state x to 0

Step 3: Obtain $G_T(t)$ via Fixed-Talbot method by shifting the distribution to the right

- In order to compute the Laplace Transform of first passage time, we use the modified Lentz algorithm. The following algorithm is due to Numerical Recipes in C and slightly modified to compute the second step in the above algorithm.

Algorithm: Continued Fractions via modified Lentz

1. **Initialize :**
   a. Set $tiny = 10^{-30}, f_0 = tiny; C_0 = f_0; D_0 = 0$
   b. Def $d_k = \mu + k\theta$; Def $a_j = -\lambda\, d_{k+j-1}$; Def $b_j := b(j,s) = \lambda + d_{k+j-1} + s$
   c. Declare D, C, delta and f as vectors of size n initialized to 0.
   d. Set $C[0] = C_0, f[0] = f_0$
2. For $j = 1, \ldots, n$: do
   a. $D_j = b_j + a_j D_{j-1}$
   b. $D_j = tiny$ if $D_j = 0, else\ D_j$
   c. $C_j = b_j + \frac{a_j}{C_{j-1}}$
   d. $C_j = tiny$ if $C_j = 0, else\ C_j$
   e. $D_j = \frac{1}{D_j}$
   f. $\Delta_j = C_j D_j$
   g. $f_j = f_{j-1}\Delta_j$
3. Return $f[n] * -\frac{1}{\lambda}$ to get **Step 1**
4. Generate $i = 1,2, \ldots, x$ orders and repeat 1-3 for each i, then take the product to get **Step 2**

Inverse Laplace Transform

The inverse Laplace Transform is defined as the Bromwich contour integral,

$$\mathcal{L}^{-1}\{\hat{f}(s)\} = f(t) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{st}\hat{f}(s)ds$$

where the abscissa of convergence $\sigma > 0$ is a real constant and chosen to put the contour to the right of all singularities in $\hat{f}(s)$.

The authors use the Euler Method for inverting the Laplace Transform which is based on formula from Abate and Whitt (1995). The algorithm is specified below:

Algorithm: Euler Method

1. Rewrite the Bromwich contour as:

$$f(t) = \frac{2e^{at}}{\pi}\int_0^\infty \cos(ut)\, Re\left(\hat{f}(\sigma + iu)\right) du$$

where $i = \sqrt{-1}$ and Re(s) is real part of s.

2. Use trapezoidal rule to with step-size h to approximate the function:

$$f(t) \approx f_h(t) = \frac{e^{\frac{A}{2}}}{2t} Re\left(\hat{f}\left(\frac{A}{2t}\right)\right) + \frac{e^{\frac{A}{2}}}{t}\sum_{k=1}^{\infty}(-1)^k\, Re\left(\hat{f}\left(\frac{A+2k\pi i}{2t}\right)\right)$$

3. Use Poisson summation formula to identify the discretization error associated with above formula and use Euler's summation to accelerate convergence:

$$E(m,n,t) = \sum_{k=0}^{m}\binom{m}{k} 2^{-m} s_{n+k}(t)$$

where

$$s_n(t) = \frac{e^{\frac{A}{2}}}{2t} Re\left(\hat{f}\left(\frac{A}{2t}\right)\right) + \frac{e^{\frac{A}{2}}}{t}\sum_{k=1}^{\infty}(-1)^k a_k(t) \quad \text{and} \quad a_k(t) = Re\left(\hat{f}\left(\frac{A+2k\pi i}{2t}\right)\right)$$

- The authors recommend to choose A = 18.4, m = 50, n = 25

For this project, we will instead choose the Fixed-Talbot Method. This method has the advantage that it is highly accurate and very fast.

- **Note:** For all probabilities, we multiply the Laplace Transforms by e^{-s100} to shift the distribution to the right prior to performing the inverse Laplace Transform.

Algorithm: Fixed-Talbot Algorithm

1. Rewrite the Bromwich integral as:

$$f(t) = \frac{1}{2\pi i}\int_{-\pi}^{\pi} e^{ts(\theta)}\hat{f}\big(s(\theta)\big)s'(\theta)d\theta$$

where $i = \sqrt{-1}$ and $s(\theta) = \sigma\theta(\cot(\theta) + i)$ and $-\pi < \theta < +\pi$

2. Since $s'(\theta) = i\sigma\big(1 + i\sigma w(\theta)\big)$ where $w(\theta) = \theta + (\theta\cot(\theta) - 1)\cot(\theta)$, we substitute above and find…

$$f(t) = \frac{\sigma}{\pi}\int_{0}^{\pi} Re\left(e^{ts(\theta)}\hat{f}\big(s(\theta)\big)\big(1 + iw(\theta)\big)\right)d\theta$$

3. Use trapezoidal rule with step-size $\frac{\pi}{M}$ and $\theta_k = \frac{k\pi}{M}$ to approximate the function:

$$f(t) \approx f(t,M) = \frac{\sigma}{M}\left(\frac{1}{2}\hat{f}(\sigma)e^{\sigma t} + \sum_{k=1}^{M-1} Re\left(e^{ts\theta_k}\hat{f}\big(s(\theta_k)\big)\big(1 + iw(\theta_k)\big)\right)\right)$$

where $\sigma = \frac{2M}{5t}$ and M is number of precision decimal digits

5.3 Monte-Carlo Simulation of Conditional Probabilities

The following two algorithms uses Monte-Carlo Simulation to verify the accuracy of the analytical formulas above.

Monte-Carlo Algo: Probability that mid-price increases (Spread=1)

1. **Initialization:**
   a. Initialize n_A and n_B as quantity at the Ask and Bid. Order size = 1
   b. Set count = 0
2. For $i = 1,2,\ldots,n$:
   a. $Set\ n_{Aold} = n_A, n_{Bold} = n_B$
   While True:
   b. Initialize death intensities for buy : $D_B = \mu + \theta_1 n_B$ and sell: $D_A = \mu + \theta_1 n_A$
   c. Initialize $\Lambda(a,b) = 2(\mu + \lambda_1) + \theta_1 n_A + \theta_1 n_B$
   d. Draw an event type {LB,LS,Bid_Down,Ask_Down} according to probability vector $\frac{(\lambda,\lambda,D_B,D_A)}{\Lambda(a,b)}$ and update $n_{A_{old}}\ and\ n_{Bold}$
   e. $if\ \ n_{Bnew} = 0: break$ #mid-price decreased
   elseif: $n_{Anew} = 0: count = += 1$; break #mid-price increased
   f. Update n_{Aold} with n_{Anew} and $n_{Bold} = n_{Bnew}$
3. Return count/n

For the next probability, assume my order is the last order at the Bid. The ask side has n_A orders. We want to find the probability that my order at the bid is executed before the mid-price moves, given my order is not cancelled.

The algorithm for making the spread is very similar and hence not included.

Monte-Carlo Algo: Probability my order at Ask is executed before mid-price moves (Spread=1)

1. **Initialization:**
    a. Initialize n_A as quantity at Ask. n_B is my B^{th} order at the bid. Order size = 1
    b. d_pos = n_B (keep track of my current position in the queue)
    c. count = 0
2. For $i = 1,2,\ldots,n$:
    a. Set: $n_{A_{old}} = n_A, n_{B_{old}} = n_B, dpos_{old} = d\_pos$ #initialize after each simulation

    While True:

    b. Initialize death intensities sell: $D_{A\_old} = \mu + \theta_1 n_{A\_old}$ to denote event E= $\{MS \vee CS\}$ $occurs$
    c. Initialize $\Lambda(a,b) = 2(\mu + \lambda_1) + \theta_1 n_{A\_old} + \theta_1(n_{B\_old} - 1)$
    d. **Event Type:**
        i. Draw an event type {LB,LS,MS,CB,E} according to probability vector $\frac{(\lambda,\lambda,\mu,\theta_1(n_{B\_old}-1),D_{A\_old})}{\Lambda(a,b)}$ and update $n_{A_{old}}$ and $n_{B_{old}}$
    e. **Event Quantity**
        i. If selected event is a MS, set $dpos_{old} -= 1$ if $dpos_{old} > 0, else$ 0 (my position moves up)
        ii. If selected event is a CB, generate $U \sim Uniform(0,1)$. Then:

            if $U > \left(\frac{n_{B_{old}} - dpos_{old}}{n_{B_{old}-1}}\right)$:

                Update $d_{pos_{old}} -= 1$; #cancellation occurred in front of me, decrement my position by 1

            Update $n_{B_{old}} -= 1$
    f. Set $n_{A_{new}} = n_{A_{old}}$; $n_{B_{new}} = n_{B_{old}}$; $d_{new} = d_{pos_{old}}$
    g. if $d_{new} == 0$:

        Set $count += 1$ #my order has been executed

        break;
    h. if $n_{A_{new}} == 0$ and $d_{new} > 0$: #mid-price moved

        break;
    i. Set: $n_{Aold} = n_{A_{new}}$ and $n_{Bold} = n_{B_{new}}$ and $dpos_{old} = d\_new$ and repeat from step b until break
3. Return count/n

5.4 Numerical Results

We run our Monte-Carlo Simulation for n=10000 iterations. Following table shows the Monte-Carlo Simulation results and Laplace Transformation results for the case when spread = 1. Each row represents quantity of orders at the best bid and columns represents quantity of orders at best ask.

Probability of mid-price Increase (Laplace)

	1	2	3	4	5
1	0.4932	0.3325	0.2796	0.2100	0.1397
2	0.6624	0.4997	0.4120	0.3460	0.2948
3	0.7400	0.5930	0.5015	0.4359	0.3869
4	0.7837	0.6521	0.5640	0.4997	0.4512
5	0.8119	0.6927	0.6096	0.5476	0.4997

Probability of mid-price Increase (Monte-Carlo)

	1	2	3	4	5
1	0.5048	0.3333	0.2542	0.2135	0.1973
2	0.6678	0.4941	0.4097	0.3473	0.3085
3	0.7469	0.5797	0.4989	0.4359	0.3847
4	0.7834	0.6531	0.5635	0.4940	0.4586
5	0.8110	0.6963	0.6078	0.5623	0.5076

Probability of bid-order execution (Laplace)

	1	2	3	4	5
1	0.5025	0.6980	0.7939	0.8483	0.8824
2	0.3589	0.5508	0.6643	0.7372	0.7869
3	0.2911	0.4653	0.5784	0.6564	0.7128
4	0.2555	0.4107	0.5180	0.5958	0.6544
5	0.2643	0.3829	0.4773	0.5507	0.6085

Probability of bid-order execution (Monte-Carlo)

	1	2	3	4	5
1	0.5046	0.7022	0.7916	0.8478	0.8824
2	0.3592	0.5418	0.6657	0.7415	0.7877
3	0.2919	0.4714	0.5792	0.6584	0.7096
4	0.2435	0.4090	0.5167	0.5949	0.6514
5	0.2261	0.3682	0.4624	0.5450	0.6080

Probability making spread(Laplace)

	1	2	3	4	5
1	0.2604	0.2955	0.3088	0.2998	0.2899
2	0.2955	0.3856	0.4056	0.4063	0.4007
3	0.3088	0.4056	0.4406	0.4518	0.4527
4	0.2998	0.4063	0.4518	0.4714	0.4787
5	0.2899	0.4007	0.4527	0.4787	0.4910

Probability of making spread (Monte-Carlo)

	1	2	3	4	5
1	0.2679	0.3049	0.3109	0.3005	0.2889
2	0.3051	0.3901	0.4089	0.4099	0.3944
3	0.3077	0.4044	0.4337	0.4512	0.4573
4	0.3013	0.4184	0.4458	0.4713	0.4855
5	0.2923	0.3957	0.4557	0.4732	0.4898

6.0 Conclusion

We can see that the CST model can predict conditional probabilities for traders very quickly without the need to use Monte-Carlo simulations. By fitting the power law function to the arrival rates, we were able to better capture the distribution of limit orders and use it successfully in the CST model.

Although this paper concerned with perfect replication, the next step would be to use this model for various securities and commodities and see if the model is able to describe the behavior for a particular stock at an exchange well.

I would like to thank Dr. Jim Gatheral for introducing me to the area of market microstructure. Without his notes, this research wouldn’t be possible.

I would also like to thank Dr. Sasha Stoikov for useful discussions on this paper.

References

1 A Stochastic model for order book dynamics (2010)-Cont,Stoikov,Talreja
2 Multi-precision Laplace transform inversion (2004) J.Abate and P.P. Valko
3 Computing Laplace Transforms for Numerical Inversion Via Continued Fractions (1999)-J.Abate,W.Whitt
4 An Introduction to Numerical Transform Inversion and its application to Probability Models(1999)-J.Abate, G.Choudhury, W.Whitt
5 Limit Order Books-Abergel,Anane,Chakraborti,Jedidi,Toke
6 Statistical properties of stock order books: empirical results and models-Bouchaud, Mezard,Potters
7 Baruch Market MicroStructure Lecture Notes: Jim Gatheral
8 Statistical Theory of Continuous Double Auction-Eric Smith,J Doyne Farmer, Laszlo Gillemot & Supriya Krishnamurthy
9 A Random Order Placement Model of Price Formation in the Continuous Double Auction-Farmer,Gillemot,Lori,Krishnamurthy,Smith,Daniels
10 The Microstructure of Financial Markets-Frank de Jong & Barbara Rindi
11 Numerical Recipes in C: The Art of Scientific Computing-B Flannery,S Teukolsky, W H. Press,W T. Vetterling