Rebuilding The Limit Order Book: Sequential Bayesian Inference on Hidden States

Hugh L. Christensen[1], Richard E. Turner[2], Simon I. Hill[1], Simon J. Godsill[1]

1: Signal Processing and Communications Laboratory, Engineering Department, Cambridge University, CB2 1PZ, UK

2: Computational and Biological Learning Laboratory, Engineering Department, Cambridge University, CB2 1PZ, UK

{hlc54, ret26, sih22, sjg30}@cam.ac.uk

Introducing the Limit Order Book

The limit order book (LOB) acts as an information store of traders future intentions [1]. The LOB consists of limit orders, each with a specified side, limit price, submission time and size. The resulting high dimensional data structure is a challenge for theoretical modelling and empirical estimation as well as, more practically, for trading.

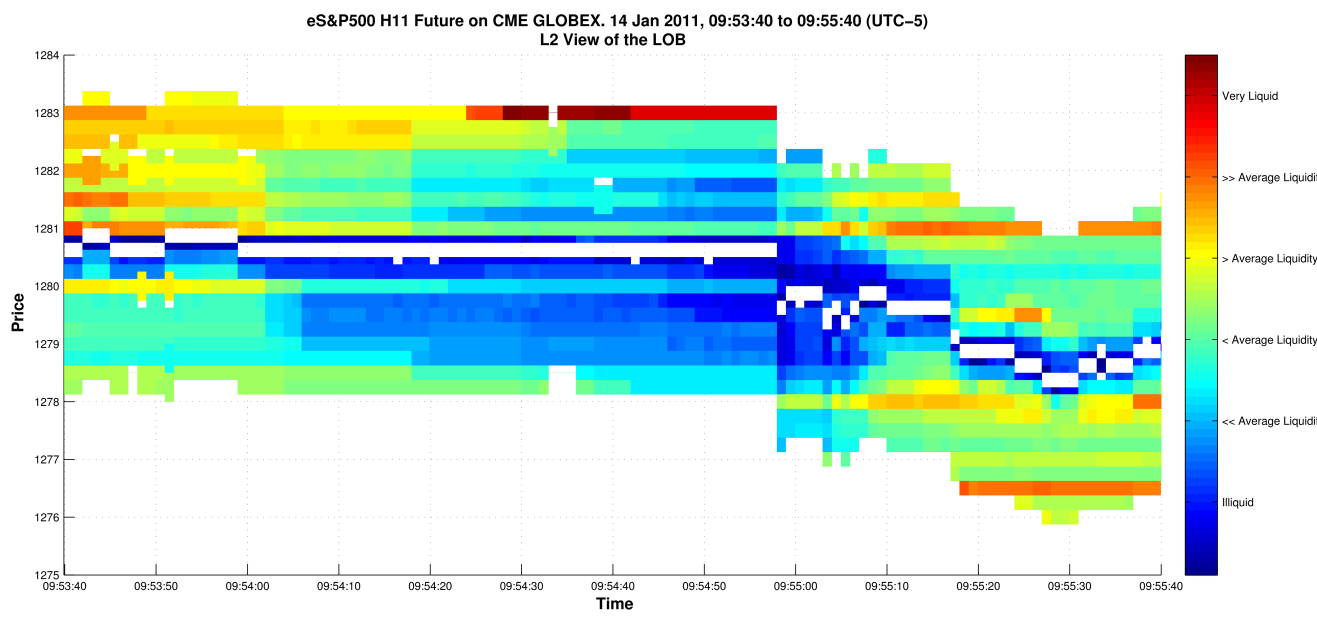
The LOB is an area of high current interest for both academic modelers and practioners,

- Book resiliency
- Buy/sell intensity
- Dynamical behavior
- Shape
- Information content
- High-frequency trading
- Optimally executing large orders
- Hidden volume detection
- Short-term price prediction
- Market making

All these areas of research require the ability to rebuild the LOB. Existing models for rebuilding the LOB are purely *deterministic* implementations of exchange rules. The contribution of this work is to show that the rebuild process can be used to infer additional information which could be beneficially used in any of the above areas of research and thus our approach is a *probabilistic* one.

L2 Structure in 3 Dimensions

The LOB can be rebuilt to the *L2 view* by applying exchange rules to the broadcast data. The broadcast data consists of three dimensions - side (bid, ask), class (price, size) and price level ($m = 1, \ldots, M$), which vary over time. In this data feed the K orders at each price level have been *aggregated* to a single net volume and *individual orders are not visible*. This aggregation represents an *information loss*.



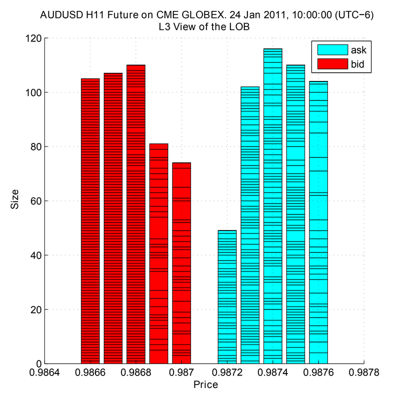
The L2 view of the LOB for the H11 contract of the e-mini S&P500 from CME GLOBEX on January 14th 2011 from 09:53:40 to 09:55:40 (UTC-5) at 100ms sampling frequency.

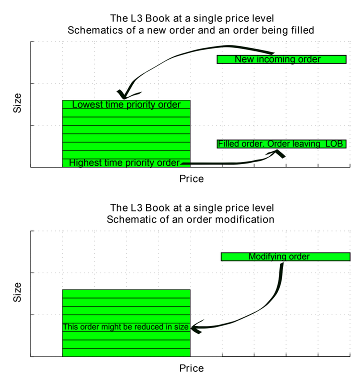
L3 Structure in 4 Dimensions

In the L3 view at each price level the component orders are visible, along with their associated size, order in the queue and time of placement. Generating the L3 view increases the dimensionality of the LOB by one to 4D with the extra dimension being size level ($l = 1, \ldots, L$) for orders.

There are five operations on the LOB; trade, order modification (price change), order modification (size increasing), order modification (size decreasing) and order cancellation.

The first three of these are deterministic operations as the rules of the exchange mean it is known which order has been effected. The last two are stochastic operations, as there is no way of knowing which order the operation has been applied to. Instead conditional probabilities for the operation must be found.





Probabilistic Framework

The problem is formulated in discrete state space by applying the framework of a homogeneous *Hidden Markov Model*, where the relationship between the L2 and L3 states is probabilistic $p(x_t|\boldsymbol{z}_t)$. x_t is the observed variable (s.t. $x_t = \sum_{l=1}^{L} v_{t,l}$ for order set $v_{t,l}$) and $\boldsymbol{z}_t$ is the latent variable, where $\mathbf{X} = \{x_1, \ldots, x_T\}$, $\mathbf{Z} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T\}$ where $\boldsymbol{z}_t = \{z_{t,1}, \ldots, z_{t,K}\}$. L3 is Markovian as $p(\boldsymbol{z}_t|\boldsymbol{z}_{t-1})$.

For parameter set $\boldsymbol{\Theta} = \{\pi, \mathbf{P}, \boldsymbol{\Phi}\}$ the joint probability distribution over latent and observed variables is given by,

$$p(\mathbf{X}, \mathbf{Z}|\boldsymbol{\Theta}) = p(\boldsymbol{z}_1|\pi) \left[\prod_{t=2}^{T} p(\boldsymbol{z}_t|\boldsymbol{z}_{t-1}, \mathbf{P}) \right] \prod_{t=1}^{T} p(x_t|\boldsymbol{z}_t, \boldsymbol{\Phi})$$

Frequentist Learning

Learning finds the state transition matrix $\mathbf{P}$, the probability of transitioning between hidden states. MLE is computationally intractable due to the size of the state space so *kernel density estimation* is used to estimate $\mathbf{P}$ by $P(v_t|v_{t-1}) \approx Q(v_t|v_{t-1}, \boldsymbol{\Psi})$. v_t is the order size and $\boldsymbol{\Psi}$ is a set of five parameters which represent the structure of the LOB $\boldsymbol{\Psi} = \{d, \alpha, \beta, \sum V_{side}, \sum V_l\}$ (distance from mid price, fitted Gamma distribution parameters and volume at given side/ price level).
Kernel density estimators of the joint distribution, $p(v_t, v_{t-1}, \Psi)$ are normalized to get the conditional distribution,

$$p(v_t|v_{t-1}, \Psi) = \frac{p(v_t, v_{t-1}, \Psi)}{\sum_{v_t} p(v_t, v_{t-1}, \Psi)} = \frac{p(v_t, v_{t-1}, \Psi)}{p(v_{t-1}, \Psi)}$$

Bayesian Inference

At each t the most likely L3 structure $\boldsymbol{z}_t$ given the observed L2 structure x_t and $\boldsymbol{\Theta}$ needs to be found. Simple MAP inference is not viable as later v_t may not be reconcilable with the $\boldsymbol{z}_t$ chosen. To overcome this, we retain distributional information at each time-step and not just the most probable state. Paths are killed off when later information shows them to be wrong.

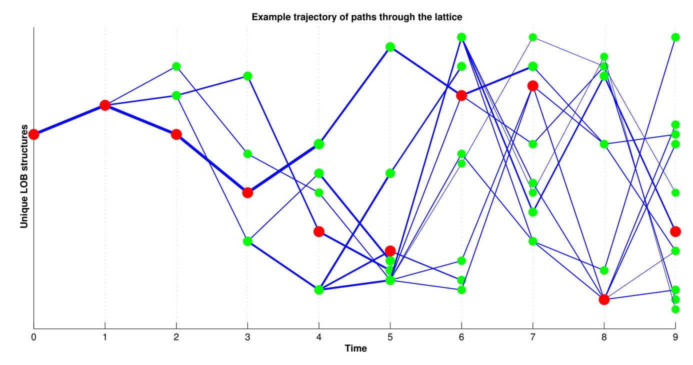
The forward algorithm is used to find the filtering distribution $\boldsymbol{z}_t^* = \underset{\boldsymbol{z}_t}{\operatorname{argmax}}\, p(\boldsymbol{z}_t|x_{1:t})$, allowing linear complexity wrt time K^2T. Arithmetic underflow is avoided by using the "soft-max trick" to re-scale probabilities. $\mathbf{Q}$ is defined as a $K \times T$ matrix of unnormalized probabilities $\mathbf{Q} = \{q_1, \ldots, q_{t-1}\}$ corresponding to $p(\boldsymbol{z}_t|x_{1:t})$.

Alg. 1 Soft-Max Forward Algorithm. $p(\mathbf{Z}|\mathbf{X}, \boldsymbol{\Theta}) = \text{SMF}(\mathbf{X}, \mathbf{P}, \boldsymbol{\Phi}, \pi)$;

for $t = 1$ to T **do**

- $E_{t-1,k} = \log(\mathbf{Q})$ {Calculate the "energies"}
- $p(\boldsymbol{z}_t = k|x_{1:t-1}) = \frac{exp(E_{t-1,k})}{\sum_k e^{E_{t-1,k}}}$ {Normalize}
- $E_{t,k} = \log(\boldsymbol{\Phi}_{\mathbf{ik}}) + \log\left(\sum_i \mathbf{P}_{\mathbf{ik}} \times e^{E_{t-1,i}}\right)$ {Propagation}
- $\Gamma = \underset{k}{\operatorname{argmax}}\,(E_{t,k})$ {Calculate scaling factor}
- $E_{t,k} \leftarrow E_{t,k} - \Gamma$ {Rescale}
- $q_t = e^{E_{t,k}}$ {Access probability}
- $\boldsymbol{z}_t^* = \underset{k}{\operatorname{argmax}}\,(q_{t,k})$ {MAP. Most likely state of the LOB.}

end for

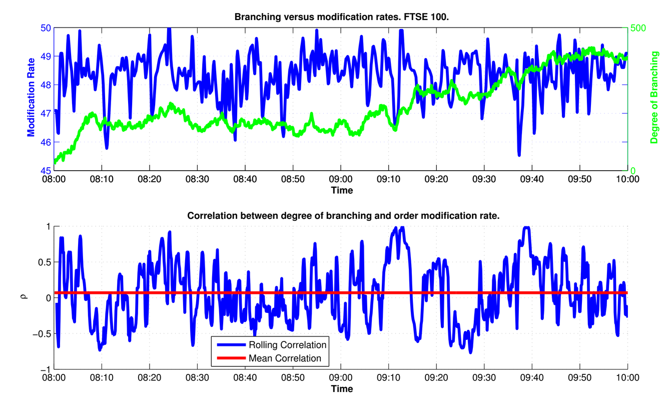


Simplified plot of LOB trajectories after a ten time steps. Line width is proportional to $p(\boldsymbol{z}_t|x_{1:t})$, such that the line can only get thinner over time, unless non-unique structures are merged. At each t LOB structures which are consistent with the observations are shown in green and the most likely of these $\boldsymbol{z}_t^*$ is shown in red.

Experimental Results

As the hidden state is never known, synthetic data is created by a generative model allowing the true L3 and L2 structures to be seen. Monte Carlo simulations comparing the true L3 state to the inferred L3 state $\boldsymbol{z}_{1:t}^*$ find statistically significant improvements over randomly generated L3 states, with R^2 values of $> 45\%$.

For the NYSE Liffe FTSE 100 future we generate a moving average of the rate of stochastic LOB updates (i.e. size reducing modifications) and a moving average of the number of branches present in the (unpruned) lattice of possible LOB structures.



Applications

The informational advantage the L3 structure gives has many applications, for example,

- Hidden volume [2]. The L3 structure of the LOB would allow the existence of iceberg orders to be probabilistically detected.
- VWAP tracking [3]. Volume participation algorithms systematically interact with the LOB and in doing so leave "footprints". Detection and prediction of such activity would allow large trades to be "front run".
- Market making for pro-rata futures [4]. The successful liquidity supplier in these securities will need to submit orders in such a way that it maximizes his matching-engine allocation and this is conditional on the L3 structure.

References

[1] M. Gould, M. Porter, S. Williams, M. McDonald, D. Fenn, and S. Howison. Limit order books. Technical report, Oxford University, 2011.

[2] S. Frey and P. Sandas. The impact of iceberg orders in limit order books. In *American Finance Association 2009 San Francisco Meetings Papers*, 2009.

[3] QSG. Beware of the VWAP trap. Technical report, QSG, 2009.

[4] Karel Janecek and Martin Kabrhel. Matching algorithms of international exchanges. Technical report, Faculty of Mathematics and Physics of Charles University in Prague, 2007.