Dynamics of limit order markets

A journey across time scales

Rama Cont & Adrien de Larrard

Columbia University, New York
&
Laboratoire de Probabilités, CNRS - Université de Paris VI.

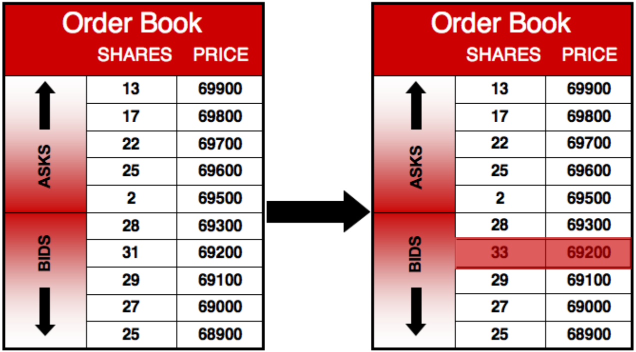
References

- Rama Cont, Sasha Stoikov and Rishi Talreja (2010) A stochastic model for order book dynamics, **Operations Research**, Volume 58, No. 3, 549-563.
- Rama Cont and Adrien de Larrard (2010) Price dynamics in a Markovian limit order market, Working Paper.
- Rama Cont and Adrien de Larrard (2010) Dynamics of limit order markets: linking volatility with order flow, Working Paper.

Outline

1. At the core of liquidity: the Limit order book
2. The separation of time scales
3. A Markovian model for order book dynamics
   - Probability of the price going up
   - Distribution of the first moving time of the price
   - Diffusion limit of the price
   - Link between volatility and order flow
4. A general framework for order book dynamics
5. Heavy-traffic approximation of a general order book model
6. Markovian jump-diffusion approximation for the bid/ask queues
   - Probability of the price going up
   - Distribution of the first moving time of the price
   - Diffusion limit of the price
   - Expression of the volatility of the price
7. Linking volatility with order flow: diffusion limit.

At the core of liquidity: the limit order book



Order Book

	SHARES	PRICE
ASKS	13	69900
	17	69800
	22	69700
	25	69600
	2	69500
BIDS	28	69300
	31	69200
	29	69100
	27	69000
	25	68900

Order Book

	SHARES	PRICE
ASKS	13	69900
	17	69800
	22	69700
	25	69600
	2	69500
BIDS	28	69300
	33	69200
	29	69100
	27	69000
	25	68900

Figure: A limit buy order: Buy 2 at 69200.

A market order

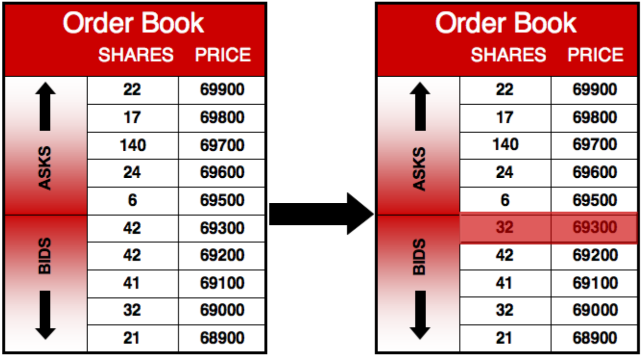


Figure: A market sell order of 10.

A cancellation

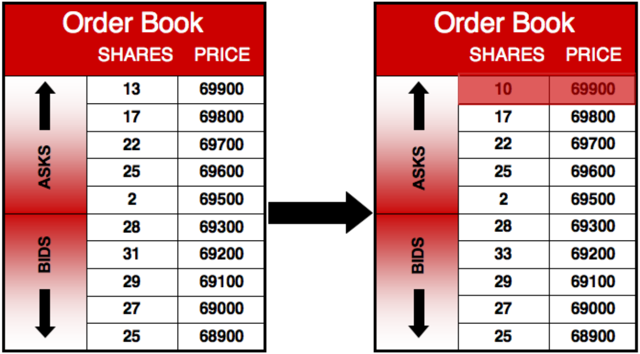


Figure: Cancellation of 3 sell orders at 69900.

Stochastic models of order book dynamics

Stochastic models for order book dynamics have been proposed in order to

- Incorporate the information in
  1. the current state of the order book
  2. statistics on the order flow (arrival rates of market, limit orders and cancellation)

  in view of
  1. optimal order execution
  2. intraday modeling and prediction of price changes and volatility

under statistically realistic assumptions on the order flow.
These applications requires *analytical* tractability and computability.

Limit order books as queueing systems

A limit order book may be viewed as a system of *queues* subject to order book events modeled as a multidimensional **point process**. A variety of stochastic models for dynamics of order book events and/or trade durations at high frequency:

- Independent Poisson processes for each order type (Cont Stoikov Talreja 2010)
- Self exciting and mutually exciting Hawkes processes (Andersen, Cont & Vinkovskaya 2010, Bacry et al 2010)
- Autoregressive Conditional Duration (ACD) model (Engle & Russell 1997, Engle & Lunde 2003, ..)

Aim: reproducing empirical properties (Smith et al, 2003, Bouchaud et al 08) for prediction, trade execution, intraday risk management.

Any of these models implies some dynamics for the (bid/ask) price, but which is difficult to describe explicitly.

In general: price is not Markovian, increments neither independent nor stationary and depend on the state of the order book.

A reduced-form model for the limit order book

- If one is primarily interested in price dynamics, then the 'action' takes place at the best bid/ask levels
- In fact empirical data show that the bulk of orders flow to the queues at the best bid/ask (e.g. Biais, Hillion & Spatt 1995)
- Ask price: best selling price: $s^a = (s^a_t, t \geq 0)$
- Bid price: best buying price $s^b = (s^b_t, t \geq 0)$.
- Reduced modeling framework: state variables= number of orders at the ask:

$$(q^a_t, t \geq 0).$$

  and number of orders at the bid:

$$(q^b_t, t \geq 0)$$

- State variable: $(s^b_t, q^b_t, s^a_t, q^a_t)_{t \geq 0}$

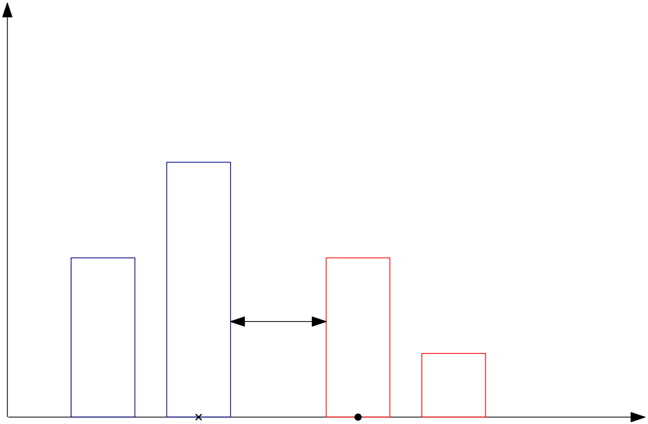


Figure: Reduced-form (Level I) representation of a limit order book

Time scales

Regime	Time scale	Issues
Ultra-high frequency (UHF)	Tick ($\sim 10^{-3} - 1$ s)	Microstructure, Latency
High Frequency (HF)	$\sim 10 - 10^2$ s	Optimal execution
"Daily"	$\sim 10^3 - 10^4$ s	Trading strategies, Option hedging

Table: A hierarchy of time scales.

The relevance of asymptotics

	Average no. of orders in 10s	Price changes in 1 day
Citigroup	4469	12499
General Electric	2356	7862
General Motors	1275	9016

Table: Average number of orders in 10 seconds and number of price changes (June 26th, 2008).

These observations point to the relevance of *asymptotics* when analyzing the dynamics of prices in a limit order market where orders arrivals occur frequently.

Limit order book as reservoir of liquidity

Once the bid (resp. the ask) queue is depleted, the price will move to the queue at the next level, which we assume to be one tick below (resp. above).
The new queue size then corresponds to what was previously the number of orders sitting at the price immediately below (resp. above) the best bid (resp. ask).
Instead of keeping track of these queues (and the corresponding order flow) at *all* price levels we treat the new queue sizes as independent variables drawn from a certain distribution f where $f(x, y)$ represents the probability of observing $(q_t^b, q_t^a) = (x, y)$ right after a price increase. Similarly, after a price decrease (q_t^b, q_t^a) is drawn from a distribution $\tilde{f}(\neq f)$ in general.

- if $q_{t-}^a = 0$ then (q_t^b, q_t^a) is a random variable with distribution f, independent from $\mathcal{F}_{t-}$.
- if $q_{t-}^b = 0$ then (q_t^b, q_t^a) is a random variable with distribution $\tilde{f}$, independent from $\mathcal{F}_{t-}$.

Distribution of queue sizes after a price move

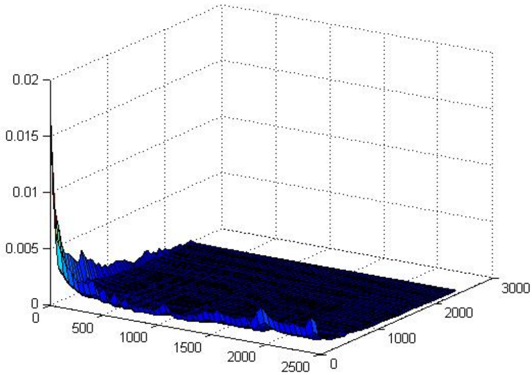


Figure: Joint density of bid and ask queues after a price move.

Distribution of queue sizes after a price move

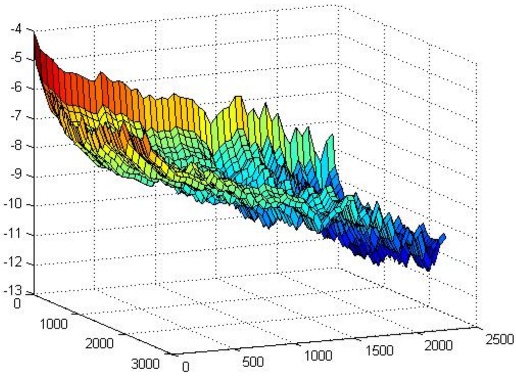


Figure: Joint density of bid and ask queues after a price move: log-scale

A simplifying assumption

Bid-ask spread	1 tick	2 tick	$\geq$ 3 tick
Citigroup	98.82	1.18	0
General Electric	98.80	1.18	0.02
General Motors	98.71	1.15	0.14

Table: Percentage of observations with a given bid-ask spread (June 26th, 2008).

We will assume for simplification that the spread is constant:
$\forall t \geq 0, s_t^a = s_t^b + \delta$.

Summary: dynamics of order book and price

The process $X_t = (s^b_t, q^b_t, q^a_t)$ is thus a continuous-time process with right-continuous, piecewise constant sample paths whose transitions correspond to the order book events at the ask $\{T^a_i, i \geq 1\}$ or the bid $\{T^b_i, i \geq 1\}$ with (random) sizes $(V^a_i)_{i\geq 1}$ and $(V^b_i)_{i\geq 1}$.

- If an order or cancelation arrives on the ask side i.e. $T \in \{T^a_i, i \geq 1\}$:

  $$(s^b_T, q^b_T, q^a_T) = (s^b_{T-}, q^b_{T-}, q^a_{T-} + V^a_i)1_{q^a_{T-} > -V^a_i} + (S^b_{T-} + \delta, R^b_i, R^a_i)1_{q^a_{T-} \leq -V^a_i},$$

  where $(R_i)_{i\geq 1} = (R^a_i, R^b_i)_{i\geq 1}$ is a sequence of IID variables with (joint) distribution f
- If an order or cancelation arrives on the bid side i.e. $T \in \{T^b_i, i \geq 1\}$:

  $$(s^b_T, q^b_T, q^a_T) = (s^b_{T-}, q^b_{T-} + V^b_i, q^a_{T-})1_{q^b_{T-} > -V^b_i} + (s^b_{T-} - \delta, \tilde{R}^a_i, \tilde{R}^b_i)1_{q^b_{T-} \leq -V^b_i},$$

  where $(\tilde{R}_i)_{i\geq 1} = (\tilde{R}^a_i, \tilde{R}^b_i)_{i\geq 1}$ is a sequence of IID variables with (joint) distribution $\tilde{f}$

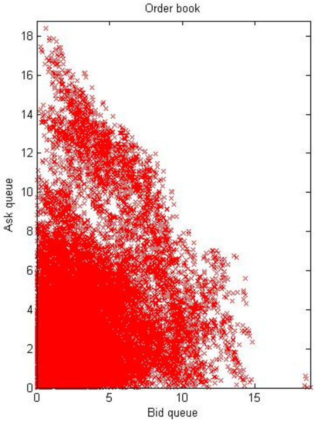


Figure: Dynamics of bid and ask queues and dynamics of the mid-price.

Example: a Markovian limit order book

Cont & de Larrard (2010) Price dynamics in a Markovian limit order market, SSRN.
Simplified version of Cont, Stoikov, Talreja (*Operations Research*, 2010)

- Market buy (resp. sell) orders arrive at independent, exponential times with rate μ,
- Limit buy (resp. sell) orders arrive at independent, exponential times with rate λ,
- Cancellations orders arrive at independent, exponential times with rate θ.
- The above events are mutually independent.
- All orders sizes are constant.

$\rightarrow$ Poisson point process $\Rightarrow$ explicit computations possible

- Between price changes, (q_t^a, q_t^b) are independent birth and death process with birth rate λ and death rate $\mu + \theta$.
- Let σ^a (resp. σ^b be the first time the size of the ask (resp bid) queue reaches zero. Duration until next price move: $\tau = \sigma^a \wedge \sigma^b$
- These are hitting times of a birth and death process so conditional Laplace transform of σ^a solves:

$$\mathcal{L}(s,x) = \mathbb{E}[e^{-s\sigma^a}|q_0^a = x] = \frac{\lambda\mathcal{L}(s,x+1) + (\mu+\theta)\mathcal{L}(s,x-1)}{\lambda+\mu+\theta+s},$$

- We obtain the following expression for the (conditional) Laplace transform of σ_a:

$$\mathcal{L}(s,x) = (\frac{(\lambda+\mu+\theta+s) - \sqrt{((\lambda+\mu+\theta+s))^2 - 4\lambda(\mu+\theta)}}{2\lambda})^x.$$

Duration until the next price change

- The duration τ until the next price change is given by:

$$\tau = \sigma^a \wedge \sigma^b.$$

- The distribution of τ conditional on the current queue sizes is

$$\mathbb{P}[\tau > t | q_0^a = x, q_0^b = y] = \mathbb{P}[\sigma^a > t | q_0^a = x]\mathbb{P}[\sigma^b > t | q_0^b = y].$$

- Inverting the Laplace transforms of σ^a, σ^b we obtain

$$\mathbb{P}[\tau > t | q_0^a = x, q_0^b = y] = \int_t^\infty \hat{\mathcal{L}}(u, x) du \int_t^\infty \hat{\mathcal{L}}(u, y) du,$$

where

$$\hat{\mathcal{L}}(t, x) = \sqrt{(\frac{\mu + \theta}{\lambda})^x} \frac{x}{t} I_x(2\sqrt{\lambda(\theta + \mu)} t) e^{-t(\lambda + \theta + \mu)}.$$

Duration until next price move

- Littlewood & Karamata's Tauberian theorems links the tail behavior of τ to the behavior of the conditional Laplace transforms of σ^a and σ^b at zero.
- When $\lambda < \theta + \mu$
  - $\mathbb{P}[\sigma^a > t | q_0^a = x] \sim_{t\to\infty} \dfrac{x(\lambda+\mu+\theta)}{2\lambda(\mu+\theta-\lambda)} \dfrac{1}{t}$
  - $\mathbb{P}[\tau > t | q_0^a = x, q_0^b = y] \sim_{t\to\infty} \dfrac{xy(\lambda+\mu+\theta)^2}{\lambda^2(\mu+\theta-\lambda)^2} \dfrac{1}{4t^2}$.
  - Tail index of order 2
- When $\lambda = \theta + \mu$
  - $\mathbb{P}[\sigma^a > t | q_0^a = x] \sim_{t\to\infty} \dfrac{x}{\sqrt{\pi\lambda}} \dfrac{1}{\sqrt{t}}$
  - $\mathbb{P}[\tau > t | q_0^a = x, q_0^b = y] \sim_{t\to\infty} \dfrac{x}{\sqrt{\pi\lambda}} \dfrac{1}{\sqrt{t}}$
  - Tail index of order 1: the mean between two consecutive moves of the price is infinite.

Probability of the price moving up given the current order book

Proposition

When $\lambda = \theta + \mu$, the probability $\phi(n, p)$ that the next price move is an increase, conditioned on having the n orders on the bid side and p orders on the ask side is:

$$\phi(n, p) = \frac{1}{\pi} \int_0^{\pi} (2 - \cos(t) - \sqrt{(2 - \cos(t))^2 - 1})^p \frac{\sin(nt)\cos(t/2)}{\sin(t/2)} dt. \tag{1}$$

Interestingly: this quantity does not depend on the arrival rates λ, θ, μ as long as $\lambda = \theta + \mu$!

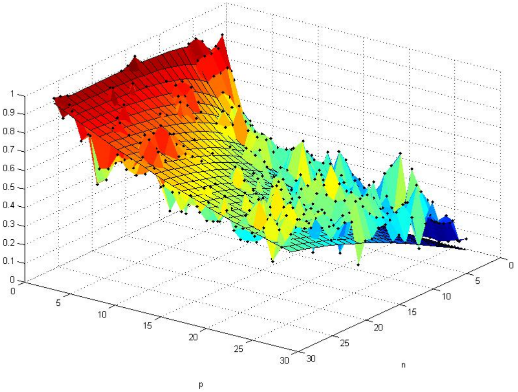


Figure: Conditional probability of a price increase, as a function of the bid and ask queue size (solid curve) compared with transition frequencies for CitiGroup tick-by-tick data on June 26, 2008 (points).

Diffusion limit of the price

At a *tick* time scale the price is a piecewise constant, discrete process. But over larger time scales, prices are observed to have "diffusive" dynamics and modeled as such.

Consider a time scale $t_n = t\zeta(n)$ over which the average number of order book events is of order n: $\zeta(n)$ is chosen such that

$$\frac{\tau_1 + ... + \tau_n}{\zeta(n)}$$

has a well-defined limit. We will then show that

$$(s_t^n := \frac{s_{t_n}}{\sqrt{n}}, t \geq 0)_{n\geq 1}$$

indeed behaves as a diffusion as for n large and compute its volatility in terms of order flow statistics i.e. a **functional central limit theorem** for $(s_t^n)_{n\geq 1}$.

Rem: diffusion limits of queues have been widely studied (Harrison, Reiman, Williams, Iglehart & Whitt,..) but the *price* process has no analogue in queueing theory.

Diffusion limit of the price: balanced order flow

Theorem (C & De Larrard (2010)

When $\lambda = \theta + \mu$ the price behaves as a diffusion at the time scale $\sim \zeta(n) = n\log(n)$:

$$s^n = \left(\frac{s(n\log n \quad t)}{\sqrt{n}}\right)_{t\geq 0} \overset{\mathcal{D}}{\Rightarrow} \sqrt{\frac{\pi\lambda\delta^2}{m(f)}}B$$

where B is a Brownian motion, $m(f) = \int_{\mathbb{R}_+^2} xydF(x,y)$.

Diffusion limit of the price: balanced order flow

Theorem (C & De Larrard (2010)

When $\lambda = \theta + \mu$ the price behaves as a diffusion at the time scale $\sim \zeta(n) = n\log(n)$:

$$s^n = \left(\frac{s(n\log n \quad t)}{\sqrt{n}}\right)_{t\geq 0} \overset{\mathcal{D}}{\Rightarrow} \sqrt{\frac{\pi\lambda\delta^2}{m(f)}}B$$

where B is a Brownian motion, $m(f) = \int_{\mathbb{R}_+^2} xydF(x,y)$.

Remark

If τ_0 is the 'tick' time scale and $\tau_2 \gg \tau_0$ the variance of the price increments at time scale τ_2 is:

$$\sigma^2 = \delta^2 \frac{\tau_2}{\tau_0} \frac{\pi\lambda}{m(f)}$$

Diffusion limit of the price

Theorem

When $\lambda < \theta + \mu$ (market orders/ cancelations dominate limit orders),

$$\left(\frac{s(nt)}{\sqrt{n}}\right)_{t\geq 0} \stackrel{\mathcal{D}}{\Rightarrow} \delta\sqrt{\frac{1}{m(f,\theta+\mu,\lambda)}}B$$

where B is a Brownian motion and $m(f,\theta+\mu,\lambda) = \mathbb{E}[\tau_f]$ is the average time between two consecutive prices moves.

Diffusion limit of the price

Theorem

When $\lambda < \theta + \mu$ (market orders/ cancelations dominate limit orders),

$$(\frac{s(nt)}{\sqrt{n}})_{t\geq 0} \stackrel{\mathcal{D}}{\Rightarrow} \delta\sqrt{\frac{1}{m(f,\theta+\mu,\lambda)}}B$$

where B is a Brownian motion and $m(f,\theta+\mu,\lambda) = \mathbb{E}[\tau_f]$ is the average time between two consecutive prices moves.

Remark

If τ_0 is the (UHF) time scale of incoming orders and $\tau_2 >> \tau_0$ tthe variance of the price increments at time scale τ_2 is

$$\sigma^2 = \frac{\tau_2}{\tau_0}\frac{\delta^2}{m(f,\lambda+\mu,\theta)}$$

Durations are not exponentially distributed..

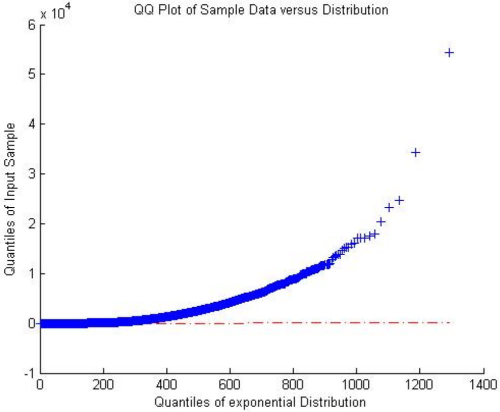


Figure: Quantile-Plot for inter-event durations, referenced against an exponential distribution. Citigroup June 2008.

Order sizes are heterogeneous

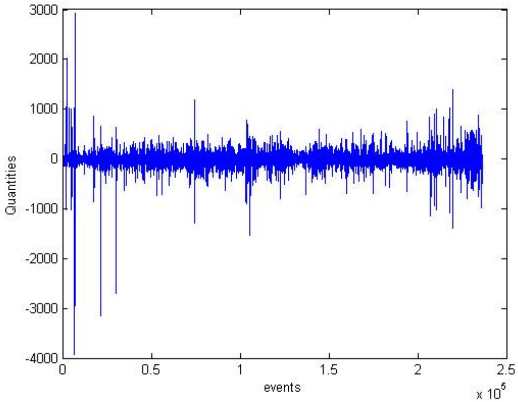


Figure: Number of shares per event for events affecting the ask. Citigroup stock, June 26, 2008.

Beyond Markovian models

This Markovian model is analytically tractable because:

- All orders have the same size ($\rightarrow$ queue)
- The time between two orders is exponential
- The orders arrive at independent times
- The dynamics of the bid is independent from the ask

Beyond Markovian models

This Markovian model is analytically tractable because:

- All orders have the same size ($\rightarrow$ queue)
- The time between two orders is exponential
- The orders arrive at independent times
- The dynamics of the bid is independent from the ask

Are the results of this Markovian model robust to these assumptions?

Beyond Markovian models

This Markovian model is analytically tractable because:

- All orders have the same size ($\rightarrow$ queue)
- The time between two orders is exponential
- The orders arrive at independent times
- The dynamics of the bid is independent from the ask

Are the results of this Markovian model robust to these assumptions?

Answer: in a liquid market, YES.

A general framework for order book dynamics

The dynamics of the order book may be described in terms of

- T_i^a: durations between order book events at the ask
- V_i^a size of the ith event at the ask. If the ith event is a market order or a cancelation, $V_i^a < 0$; if it is a limit order $V_i^a \geq 0$.
- T_i^b the time between the $i-1$th and the ith order coming on the bid side
- V_i^b the size of the ith event at the bid

- We do *not* assume these random variables to be independent!
- For general sequences $(T_i^a, V_i^a)_{i \geq 0}$ and $(T_i^b, V_i^b)_{i \geq 0}$, the order book $q = (q^a, q^b)$ is not a Markov process.
- It is not possible, for general sequences $(T_i^a, V_i^a)_{i \geq 0}$ and $(T_i^b, V_i^b)_{i \geq 0}$, to compute the probability transitions of the price and the distribution of the moving times of the price

From micro- to meso-structure: heavy traffic approximation

- Let τ_0 be the time scale of order arrivals (the millisecond)
- At the time scale $\tau_1 >> \tau_0$, the impact of one order is 'very small' compared to the total number of orders q^a and q^b.
- It is reasonable to approximate $q = (q^a, q^b)$ by a process whose state space is continuous ($\mathbb{R}_+^2$)
- More precisely we will show that the *rescaled* order book

$$Q_n(t) = (\frac{q^a(tn)}{\sqrt{n}}, \frac{q^b(tn)}{\sqrt{n}})_{t\geq 0}$$

converges in distribution to a limit *Heavy traffic approximation* of $q = (q^a, q^b) =$ limit (in distribution) $Q = (Q^a, Q^b)$ of Q_n

Assumptions on the order flow

- We assume that $(T_i^a, i \geq 1)$ and $(T_i^b, i \geq 1)$ are two stationary sequences and $\mathbb{E}[T_1^a] = \mathbb{E}[T_1^b] = \dfrac{1}{\lambda} < \infty$
- Examples verifying these assumptions:
  - Independent Poisson processes for each order type (Cont Stoikov Talreja 2010)
  - Self exciting and mutually exciting Hawkes processes (Andersen, Cont & Vinkovskaya 2010)
  - Autoregressive Conditional Duration (ACD) model (Engle & Russell 1997)
- The event sizes $(V_i^a, i \geq 1)$ and $(V_i^b, i \geq 1)$ are stationary, weakly dependent (e.g. uniformly mixing) with mean zero and verify:
  - $\mathbb{E}[(V_1^a)^2] + 2\sum_{i=2}^{\infty} Cov(V_1^a, V_i^a) = \mu^2$
  - $\mathbb{E}[(V_1^b)^2] + 2\sum_{i=2}^{\infty} Cov(V_1^b, V_i^b) = \mu^2$
  - $\dfrac{1}{\sigma^2}\left(\mathbb{E}[V_1^a V_1^b] + 2\sum_{i=1}^{\infty} \mathbb{E}[V_1^a V_i^b] + \mathbb{E}[V_1^b V_i^a]\right) = \rho < \infty$

Theorem: Heavy traffic limit of order book dynamics

Under the previous assumptions on the order flow:

$$Q_n = \left(\frac{q^a(tn)}{\sqrt{n}}, \frac{q^b(tn)}{\sqrt{n}}\right)_{t\geq 0} \overset{\mathcal{D}}{\Rightarrow} B^f \quad on \quad (\mathcal{D}, J_1),$$

where B^f is a Markov process on $\mathbb{R}_+^2$ with generator:

$$on\ x>0,\ y>0,\ \mathcal{G}h(x,y) = \frac{\delta^2}{2}\mu^2\lambda\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + 2\rho\frac{\partial^2 h}{\partial x \partial y}\right),$$

$$\mathcal{G}h(x,0) = \int_{\mathbb{R}_+^2} h(x,y)dF(x,y), \quad and \quad \mathcal{G}h(0,y) = \int_{\mathbb{R}_+^2} h(y,x)dF(x,y)$$

whose domain $\mathcal{D}$ is the set of functions $h \in \mathcal{C}^2(\mathbb{R}_+)^2$ with

$$\int_{\mathbb{R}_+^2} \left(|h(x,y)| + |h(x,y)|\right) dF(x,y) < \infty.$$

Heavy traffic limit of order book: description

The limit Markov process B^f

- behaves like planar Brownian motion with covariance matrix

$$\delta^2\mu^2\lambda \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{2}$$

  on the orthant $\{x > 0\} \cap \{y > 0\}$
- is reinitialized according to F each time it hits the x-axis
- is reinitialized according to $\tilde{F}$ each time it hits the y-axis

It is an example of 'regulated Brownian motion' (Harrison 1990). B^f is a planar Brownian motion in the orthant “regulated by F”.

Heavy traffic limit: description

Let τ_0 the time scale of incoming orders and $\tau_1 >> \tau_0$. Under the previous assumptions we can approximate the dynamics of the order book $q = (q^a, q^b)$ by the process B^f with covariance matrix

$$\Sigma = \frac{\tau_1}{\tau_0}\delta^2 v^2 \lambda \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \qquad \text{with} \tag{3}$$

- $\mathbb{E}[T_1^a] = \mathbb{E}[T_1^b] = 1/\lambda$: average duration between events
- $v^2 = \mathbb{E}[(V_1^a)^2] + 2\sum_{i=2}^{\infty} Cov(V_1^a, V_i^a)$: variance of order sizes
- $\rho = \frac{1}{\sigma^2}\left(\mathbb{E}[V_1^a V_1^b] + 2\sum_{i=1}^{\infty}\mathbb{E}[V_1^a V_i^b] + \mathbb{E}[V_1^b V_i^a]\right)$ measures the "correlation" between the order sizes at the bid and at the ask.

If order sizes at bid and ask are symmetric and uncorrelated then $\rho = 0$.
Empirically we find that $\rho < 0$ for all data sets examined.

Proposition (R C & Larrard, 2010

Let τ be duration until the next price change. The distribution of τ given x orders at the ask and y orders at the bid is given by:

$$\mathbb{P}[\tau > t | q_0^a = x, q_0^b = y] = \frac{2r_0}{\sqrt{2\pi\sigma_Q^2 t}} e^{-\frac{3r_0^2}{4\sigma_Q^2 t}} \sum_{n=1}^{\infty} \frac{1}{n} \sin \frac{n\pi\theta_0}{\alpha} (I_{(\nu-1)/2}(r_0^2/4\sigma_Q^2 t)$$

$$+ I_{(\nu+1)/2}(r_0^2/4\sigma_Q^2 t)),$$

where $\nu = n\pi/\alpha$ and I_n is the modified Bessel function of order n.

$$\alpha = \arctan(-\frac{\sqrt{1-\rho^2}}{\rho}) \qquad r_0 = \sqrt{\frac{x^2 + y^2 - 2\rho xy}{\lambda\sigma_V^2(1-\rho)}}$$

$$\theta_0 = \begin{cases} \pi + \tan^{-1}(-\frac{y\sqrt{1-\rho^2}}{x - \rho y}) & x < \rho y \\ \frac{\pi}{2} & x = \rho y \end{cases} \qquad (4)$$

Duration between price changes

- Spitzer (1958) computed the tail index of τ as a function of the correlation coefficient ρ
- The tail index of τ given x orders at the ask and y orders at the bid is $\dfrac{\pi}{2(\pi/2 + \sin^{-1}(\rho))}$
  - If $\rho = 0$, the tail index is 1. The tail index was the same for the Markovian order book when $\lambda = \theta + \mu$
  - If $\rho > 0$, the tail index is strictly less than one
  - If $\rho < 0$, the tail index is more than one. The duration between two consecutive moves of the price has a finite first moment

Probability of the price moving up

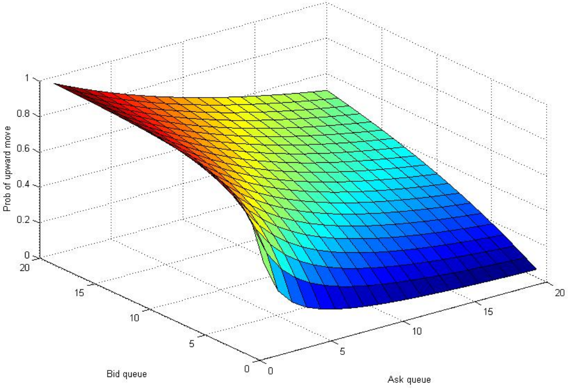
Proposition

(R C & Larrard, 2010): The probability $p_{\mathrm{up}}(x,y)$ that the next price move is an increase, given a queue of x shares on the bid side and y shares on the ask side is

$$p_{\mathrm{up}}(x,y) = \frac{1}{2} - \frac{\arctan\left(\sqrt{\frac{1+\rho}{1-\rho}}\frac{y-x}{y+x}\right)}{2\arctan\left(\sqrt{\frac{1+\rho}{1-\rho}}\right)}$$

Avellaneda, Stoikov & Reed (2010) have computed this for the special case $\rho = -1$.
When $\rho = 0$ (independent flows at bid and ask)
$p_{\mathrm{up}}(x,y) = 2\arctan(y/x)/\pi$.

Probability of price moving up given queue sizes



Diffusion limit of the price

At a *tick* time scale the price is a piecewise constant, discrete process.
But over larger time scales, prices are observed to have "diffusive" dynamics and modeled as such.
Consider a time scale $t_n = t\zeta(n)$ where $\zeta(n) \to \infty$.
Over which time scales does the rescaled price process

$$s_t^n = \frac{s_{t^n}}{\sqrt{n}}$$

behave like a diffusion?
What is this diffusion limit?
How is the (low frequency) volatility of the price related to order flow statistics?
Approach: derive a functional Central Limit theorem for the price process $(s_t^n, t \geq 0)$ as $n \to \infty$

Diffusion limit of the price

Theorem (R.C, & de Larrard, 2010)

- *When* $\rho = 0$,

$$\left(\frac{s(n \log n \quad t)}{\sqrt{n}}\right)_{t \geq 0} \overset{\mathcal{D}}{\Rightarrow} \sigma B \quad \text{on } (\mathcal{D}, J_1),$$

where

$$\sigma^2 = \frac{\pi \delta^2 v^2 \lambda}{m(f)} \qquad m(f) = \int_{\mathbb{R}_+^2} xydF(x,y).$$

- *When* $\rho < 0$,

$$\left(\frac{s(n \quad t)}{\sqrt{n}}\right)_{t \geq 0} \overset{\mathcal{D}}{\Rightarrow} \sigma_p B$$

where

$$\sigma^2 = \frac{\delta^2}{m(f, \sigma_Q, \rho)}, \qquad m(f, \sigma_Q, \rho) = \mathbb{E}[\tau_f]$$

Link between volatility and order flow

The variance of price increments at time scale $\tau_2 \gg \tau_1$ is thus given by

$$\sigma^2 = \frac{\tau_2}{\tau_1} \frac{\pi \delta^2 v^2 \lambda}{m(f)} \qquad m(f) = \int_{\mathbb{R}_+^2} xy dF(x,y).$$

Up to a factor v^2 (reflecting variance/heterogeneity of order sizes), same expression as when orders are independent Poisson processes!

Conclusion

- Analytically tractable approximation of the dynamics of a limit order book by a Markovian jump-diffusion process
- Rigorous analysis of behavior across time scales using Functional CLT.
- General assumptions: finite second moment of order sizes, finite first moment of quote durations and weak dependence, allows for dependence in order arrival times and sizes
- Allows for **dependent** order durations, dependence between order size and durations, autocorrelation, ...
- Explicit expression of probability transitions of the price
- Distribution of the duration between consecutive price moves
- Different regimes for price behavior depending on the correlation between buy and sell order sizes
- Expression of the volatility of the price as a function of orders flow statistics

A journey across time scales

Regime	Time scale	Order book	Price
UHF	$\sim 10^{-3} - 1$ s	Pair of queues	Discrete/ pure jump process
HF	$\sim 10 - 10^{2}$ s	Jump-diffusion in pos. orthant	Discrete/ pure jump process
Daily	$\sim 10^{3} - 10^{4}$ s	—	Diffusion or Lévy process

Table: Dynamics at different time scales.