

Price dynamics in Limit Order Markets: from multi-scale stochastic models to free-boundary problems

Rama Cont

Dept of Mathematics
Imperial College London

Imperial-ETH Workshop on Mathematical Finance, 2015

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 (2011) Statistical modeling of high frequency data: facts, models and challenges, **IEEE SIGNAL PROCESSING**, Vol 28, No 5, 16–25.
- Rama Cont and Adrien de Larrard (2013) Price dynamics in a Markovian limit order market, **SIAM Journal on Financial Mathematics**, Vol 4, 1–25.
- Rama Cont and Adrien de Larrard (2011) Order book dynamics in liquid markets: limit theorems and diffusion approximations, <http://ssrn.com/abstract=1757861>.
- Rama Cont and Adrien de Larrard (2012) Price dynamics in limit order markets: linking volatility with order flow, Working Paper.
- Rama Cont (2014) High frequency dynamics of limit order markets: multi-scale modeling and asymptotic analysis.

Outline

- 1 Limit order markets
- 2 Stochastic models of the limit order book
- 3 PDE models of the price formation: Lasry & Lions (2007)
- 4 The separation of time scales
- 5 A multiscale model of trading in limit order markets
- 6 Hydrodynamic limit: a two -phase moving boundary problem

At the core of liquidity: the limit order book

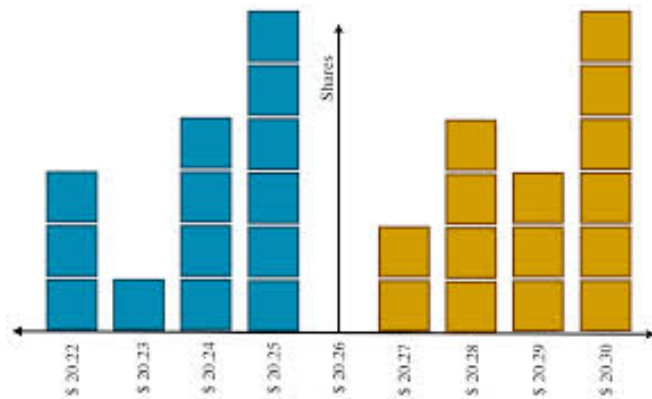


Figure: A limit order book.

Limit orders

A limit order is an order to buy (sell) a certain quantity at a given price. Limit orders queue according to time priority until they are executed against a market order.

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

A market order is an order to buy (sell) a certain quantity at the best available price. Market orders are executed immediately against available limit orders at the best price.

Order Book		
	SHARES	PRICE
↑ ASKS	22	69900
	17	69800
	140	69700
	24	69600
	6	69500
↓ BIDS	42	69300
	42	69200
	41	69100
	32	69000
	21	68900

Order Book		
	SHARES	PRICE
↑ ASKS	22	69900
	17	69800
	140	69700
	24	69600
	6	69500
↓ BIDS	32	69300
	42	69200
	41	69100
	32	69000
	21	68900

Figure: A market sell order of 10.

A cancellation

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 ↑	10	69900
	17	69800
	22	69700
	25	69600
	2	69500
BIDS ↓	28	69300
	33	69200
	29	69100
	27	69000
	25	68900

Figure: Cancellation of 3 sell orders at 69900.

Electronic Limit order markets

- The advent of electronic trading has transformed markets and led to a new market landscape dominated by algorithms which can submit and cancel orders at very high speeds.
- This has enabled the emergency of High Frequency Trading (HFT), a new category of trading strategies operating at millisecond frequency.
- At the same time, there exists a population of market participants submitting orders at lower frequencies.
- The delicate balance between these heterogeneous order flows was seen as the root of the recent Flash Crash(es).

Flash Crash



Figure: The Flash Crash of May 2010 in the US equity markets.

Electronic Limit order markets

Many questions of interest to regulators and market participants:

- What do we understand about market dynamics in such an environment?
- How does order flow interact with price dynamics?
- How does high-frequency market activity affect market dynamics at lower frequencies?
- How does the co-existence of heterogeneous order flows operating at different frequencies affect market depth/liquidity and price dynamics?

Our objective: develop a quantitative modeling framework capable of providing some analytical insight into these complex questions.

Stochastic models of order book dynamics

- Traditional market microstructure theory models the strategic interaction between a small number of agents (informed/non-informed trader, market maker) in a 1 or 2 period game theoretical setting, with an emphasis on information asymmetry and adverse selection.
- These models provide conceptual insights into market design and analysis but are not amenable to quantitative analysis or a realistic comparison with data: need for **quantitative** modeling.
- The recent years have witnessed the emergence of stochastic models for order book dynamics, which aim at incorporating the information in the order flow in view of
 - 1 estimation of intraday risk (volatility, loss distribution)
 - 2 short-term (< second) prediction of order flow and price movements for trading strategies
 - 3 optimal order execution

These applications requires *analytical* tractability and computability.

Limit order books as queueing systems

A limit order book may also 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: Poisson processes for each order type, Self exciting and mutually exciting Hawkes processes (Cont, Jafteson & Vinkovskaya 2010, Bacry et al 2010), Autoregressive Conditional Duration (ACD) model (Engle & Russell 1997, Engle & Lunde 2003, ..), ...

Most of these models are high-dimensional and applications may require heavy simulation/ numerics.

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

Common approach: model separately order flow dynamics and price dynamics through ad-hoc price impact relations/assumptions.

Example: a Markovian limit order book

C. , Stoikov, Talreja (*Operations Research*, 2010) [CST 2010]

State of limit order book $X(t) = (X_i(t))$: $X_i(t)$ = volume of limit order (< 0 for sell, > 0 for buy) at price level i .

Bid / ask price:

$$p_b(t) = \sup\{i = 1..N, X_i(t) > 0\} \leq p_a(t) = \inf\{j = 1..N, X_j(t) < 0\}$$

- Arrival of market orders, limit orders and cancelations at different price levels $i = 1..N$ described by a (spatial) *Poisson point process* with intensity depending on distance from best quote.
- All orders have same size.

→ limit order book $X(t)$ described by a continuous-time Markovian birth-death process \Rightarrow analytical formulas for

- distribution of durations between price changes,
- distribution of time to execution of limit orders,
- probability of price increase conditional on state of the order book.

The limit order book as a measure-valued process

The state of limit order book may be viewed as a signed measure μ on \mathbb{R} :
 $\mu(B) = \text{vol of limit buy orders with prices in } B - \text{vol. of limit sell orders with prices in } B$

The buy/sell side of the book correspond to the Hahn-Jordan decomposition of the measure μ :

$$\mu = \mu_+ - \mu_- \quad a(\mu) = \inf(\text{supp}(\mu_-)) \geq b(\mu) = \sup(\text{supp}(\mu_+)),$$

$$\text{supp}(\mu_+) \subset (-\infty, b(\mu)] \quad \text{supp}(\mu_-) \subset [a(\mu), \infty)$$

We denote \mathbb{L} the set of signed measures whose Hahn-Jordan decomposition is of the form above.

Thus, the limit order book may be viewed in terms of a pair of Radon measures $(\mu_+, \mu_-) \in \mathcal{M}(\mathbb{R})^2$.

In the above example, this leads to a measure-valued Markov process with values in $\mathcal{M}(\mathbb{R})^2$.

PDE models of price formation

Lasry & Lions (2007) proposed a PDE model for the dynamics of the density of buy/sell orders: this model assumes $\mu_t^\pm(dx) = \rho_\pm(t, x)dx$ and postulate that the density ρ is the solution of the following free boundary problem:

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} - \frac{\partial \rho^-}{\partial x}(t, S_t) \delta_{S_t - a} \quad \text{for } x < S_t \quad (1)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} - \frac{\partial \rho^+}{\partial x}(t, S_t) \delta_{S_t + a} \quad \text{for } x > S_t \quad (2)$$

$$\rho^-(t, x) = 0 \quad \text{for } x > S_t, \quad \rho^+(t, x) = 0 \quad \text{for } x \geq S_t \quad (3)$$

Interpretation: after trading takes place at price S_t , buyers become sellers at price $S_t + a$ and sellers become buyers at price $S_t - a$ where $a > 0$ is a 'transaction cost'.

PDE models of price formation

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} - \frac{\partial \rho^-}{\partial x}(t, S_t) \delta_{S_t - a} \quad \text{for } x < S_t \quad (4)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} - \frac{\partial \rho^+}{\partial x}(t, S_t) \delta_{S_t + a} \quad \text{for } x > S_t \quad (5)$$

$$\rho^-(t, x) = 0 \quad \text{for } x > S_t, \quad \rho^+(t, x) = 0 \quad \text{for } x \geq S_t \quad (6)$$

Caffarelli, Markowich & Pietsch (2013), Caffarelli, Markowich & Wolfram (2011)

- there exists a unique smooth solution
- price dynamics is continuous: $S \in C([0, \infty), \mathbb{R})$
- if $\mu_0^+(-\infty, S_0) = M_+ \neq \mu_0^-(S_0, \infty) = M_-$ then

$$S_t \stackrel{t \rightarrow \infty}{\sim} \sqrt{t} \operatorname{erf}^{-1}\left(\frac{M_+}{M_-}\right)$$

Relation between modeling approaches

- What is the relation between the discrete, stochastic models describing high-frequency dynamics of limit order books and the PDE-based price formation models?
- Can the latter be derived as an appropriate scaling limit of the former and if so, under what assumptions?
- How are the parameters of the PDE models related to the parameters of the point processes describing order flow at high frequency?

Tool: asymptotic analysis of the *fluid limit* for stochastic limit order book models

A hierarchy of time scales

Regime	Time scale	Issues
Ultra-high frequency (UHF)	$\sim 10^{-3} - 1$ s	Microstructure, Latency
High Frequency (HF)	$\sim 10 - 10^2$ s	Optimal execution
Low Frequency (minutes - hours)	$\sim 10^3 - 10^4$ s	Trading strategies, Hedging

The separation between these time scales opens the door to the use of asymptotic methods for connecting dynamics at different time scales. Idea: start from a description of the limit order book at the finest scale and derive probabilistic limit theorems for computing quantities at larger time scales.

Analogies with 'hydrodynamic description' of interacting particle systems.

Moving across time scales: fluid and diffusion limits

Idea: study limit of **rescaled** limit order book as

- tick size $\rightarrow 0$
- frequency of order arrivals $\rightarrow \infty$
- order size $\rightarrow 0$

All these quantities are usually parameterized / scaled as a power of a large parameter $n \rightarrow \infty$, which one can think of as number of market participants or frequency of orders.

The limit order book having a natural representation as a (pair of) measures, vague convergence in $D([0, \infty), \mathcal{M}(\mathbb{R})^2)$ is a natural notion of convergence to be considered.

Various combination of scaling assumptions are possible, which may lead to very different limits.

Moving across time scales: fluid and diffusion limits

Various combination of scaling assumptions are possible for the same process, which lead to very different limits.

When scaling assumptions are such that variance vanishes asymptotically, the limit process is deterministic and often described by a PDE or ODE: this is the functional equivalent of a Law of Large Numbers, known as the 'fluid' (or 'hydrodynamic' limit).

Ex: N_t^n Poisson process with intensity λ_n^i .

$$\lambda_n^i \sim n\lambda^i \quad \left(\frac{N_1^n - N_2^n}{n}, t \geq 0 \right) \xrightarrow{n \rightarrow \infty} ((\lambda^1 - \lambda^2)t, t \geq 0)$$

Other scaling assumptions for the same process may lead to a random limit ("diffusion limit"). Example:

$$\lambda_n^i \sim n\lambda, \quad \lambda_n^1 - \lambda_n^2 = \sigma^2 \sqrt{n}, \quad \frac{N_1^n - N_2^n}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} \sigma W$$

'Heavy traffic' asymptotics for limit order books

- Kruk (2003): fluid limit for a simple auction process
- Cont & Larrard (2013): diffusion limit of a reduced-form Poisson order book model → diffusion dynamics for bid/ask queue sizes, price jumps at each exit time of queue from positive orthant
- Cont & Larrard (2012): diffusion limit of a reduced-form order book with general point process dynamics → diffusion limit for price, expression of price volatility in terms of order intensities
- Maglaras & Moallemi (2013): fluid limit for a modified (CST 2010) model → (Average) order book profile and price described by ODE
- Dai et al (2013): fluid limit for (CST 2010) model → constant price, linear ODE with constant coefficients for limit order book
- Horst & Paulsen (2013) : fluid limit for a model with IID Poisson arrivals → ODE for price, 1st order PDE for limit order book
- Lakner et al (2014): (yet another) heavy traffic limit of a one-sided (CST 2010) model → strictly increasing price process, degenerate (single price level) or flat (block) limit for order book.

'Heavy traffic' asymptotics for limit order books

Which scaling limit is the relevant one for real markets?

- Constant, deterministic or strictly increasing price dynamics do not seem realistic, neither at high frequency nor at daily frequency.
- We would like to derive the price process from order book dynamics rather than specify it exogenously.
- Moreover, given the heterogeneity of the order flow (co-existence of high and low frequency traders) it is unlikely that a single time scale/ frequency will give the right asymptotics.
- Finally: all these models are queueing models where limit orders arrive randomly and wait for execution. Is that ALL that is going on in the order flow?
- How to choose the right scaling assumptions? Hint: examine the DATA...

Net order flow at the bid and ask levels displays a diffusion-like behavior over a time scale of seconds or minutes.

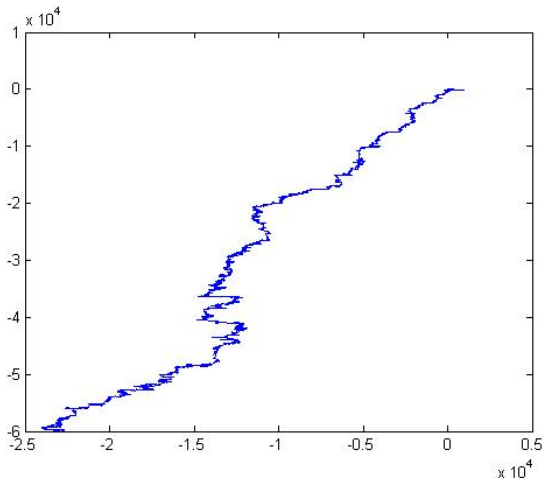


Figure: Intraday dynamics of net order flow at bid and ask: Citigroup, June 26, 2008.

Decomposition of the order flow into components

(Joint with A Kirilenko, A Kukanov, E Vinkovskaya)

Study of detailed database of order flow in one of the most liquid electronic markets: S&P e-mini futures market (CME).

- Electronic limit order market with around 10,000 participants (trade accounts).
- Data: all messages exchanged between market participants and Globex: creation/modification/cancelation of new orders, execution confirmations
- Trader IDs included in data: we can trace order flow of a given account.
- We compute for each trader ID, a range of statistics to describe the characteristics of its sequence of orders/cancels.
- These statistics are then used to classify trader IDs in more or less homogeneous groups with similar characteristic of order flow.

Decomposition of the order flow into components

- Market order flow is a superposition of heterogeneous order flows operating at widely different frequencies.
- The vast majority of accounts are "low frequency traders" who submit infrequent, small orders, cancel very few of them, trade directionally, and accumulate inventory. These are the main contribution to the volume of the order book at deeper levels .
- A very small number of HFTs account for around 50% of volume of orders and trades. Their order flow is concentrated close to the bid/ask with the vast majority of orders being placed at the best or second-best levels. Cross section distribution of order arrivals and sizes appear to be random.
- HFT order flow is NOT simply an accelerated version of the order flow of other participants: in particular, HFTs do not accumulate inventory, contribute zero net volume to the book on average and shift orders across different levels close to the best bid/ask.

A multi-component order flow model

Motivated by these observations, we model the order flow as a superposition of two distinct components.

- The 1st component is a large population of "low frequency traders" who submit infrequent, small orders (order size $\rightarrow 0$) at all price levels, cancel very few of them. Their order flow is modeled as a Poisson point process as in (CST 2010). The heavy traffic regime arises here due to their large number $n \rightarrow \infty$, but the volume of orders at each level remains finite .
- If the price submitted by such a trader is better than the best available price, it is executed as a market order, at the best quote of the opposite side: this leads to an intensity of market orders at the best bid/ask levels which is ALSO of order n .
- Order arrival intensities can be allowed to depend on distance to best bid/ask and more generally, on the state of the order book.

A multi-component order flow model

2nd component: HFT order flow

- The 2nd component ("HFT order flow") is a constantly balanced order flow occurring at high frequency ($\sim n$): each buy order is followed by a sell order after a very short time ($\sim 1/n$) and vice versa. Thus, the net result is that this components shifts orders in the order book from one level to a neighboring one.
- A very small number of HFTs account for around 50% of volume of orders and trades. Their order flow is concentrated close to the bid/ask with the vast majority of orders being placed at the best or second-best levels. Cross section distribution of order arrivals and sizes appear to be random.
- At the best bid/ask, we thus have submission/deletion of orders at rates $\sim n$: to account for the fact that best/bid ask \sqrt{n}

A multi-component model of order flow

Mathematical model: a multi-scale Markov model for the limit order book

- State space

$$\mathcal{L} = \{\eta : \mathbb{Z} \rightarrow \mathbb{Z}, \exists p \in \mathbb{Z}, \quad \eta \mathbf{1}_{x < p} \geq 0, \quad \eta \mathbf{1}_{x > p} \leq 0\}$$

- For $\eta \in \mathcal{L}$ define
 Bid price : $b(\eta) = \sup\{x \in \mathbb{Z}, \eta(x) > 0\}$
 Ask price : $a(\eta) = \inf\{x \in \mathbb{Z}, \eta(x) < 0\}$
- We will now describe the evolution of the order book in \mathcal{L} through elementary 'order book events' and their occurrence rates.

A multi-component model of order flow

i Arrival of a new limit order at price $x \in \mathbb{N}$ at rate $\lambda_+(x, \eta)$:

$$(i) \quad x \leq b(\eta) \quad \eta \mapsto \eta + \mathbf{1}_x \quad \text{at rate} \quad \lambda_+^b(x - b(\eta), \eta)$$

$$x \geq a(\eta) \quad \eta \mapsto \eta - \mathbf{1}_x \quad \text{at rate} \quad \lambda_+^a(x - a(\eta), \eta)$$

ii Cancellation of a limit order without replacement, at rate λ_- :

$$(ii) \quad x \leq b(\eta) \quad \eta \mapsto \eta - \mathbf{1}_x \quad \text{at rate} \quad \lambda_-^b(x - b(\eta), \eta)$$

$$x \geq a(\eta) \quad \eta \mapsto \eta + \mathbf{1}_x \quad \text{at rate} \quad \lambda_-^a(x - a(\eta), \eta)$$

iii Cancellation of an order and its replacement by an order closer to the bid/ask:

$$(iii) \quad x < b(\eta) : \quad \eta \mapsto \eta + \mathbf{1}_{x+1} - \mathbf{1}_x \quad \text{at rate} \quad r^b(x - b(\eta), \eta)$$

$$x > a(\eta) : \quad \eta \mapsto \eta - \mathbf{1}_{x-1} + \mathbf{1}_x \quad \text{at rate} \quad r^a(x - a(\eta), \eta)$$

- iv Execution of HFT orders: limit orders at the best bid/ask prices may get executed against incoming market orders of the opposite sign. If a market order is executed against a limit order posted by a high-frequency trader, the trader posts a new limit order on the opposite side of the book. If a limit buy order is executed at $b(\eta)$, the trader posts a limit sell order at a slightly higher price $b(\eta) + \xi$ where ξ is modeled as a positive random variable with distribution g . If a limit sell order is executed at $a(\eta)$, the trader posts a limit buy order at a slightly lower price $a(\eta) - \xi$. Denoting by q the proportion of limit orders posted by high-frequency traders, this gives

$$(iv) \quad \eta \mapsto \eta - \mathbf{1}_{b(\eta)} + \mathbf{1}_{a(\eta)+x} \quad \text{at rate} \quad q\mu^b \quad g(x)$$

$$\eta \mapsto \eta - \mathbf{1}_{a(\eta)} + \mathbf{1}_{b(\eta)-x} \quad \text{at rate} \quad q\mu^a \quad g(x)$$

where μ^b, μ^a the rate of arrival of market orders

- v Execution of market order against non-HFT limit orders

$$(v) \quad \eta \mapsto \eta - \mathbf{1}_{b(\eta)} \quad \text{at rate} \quad (1 - q)\mu^b$$

$$\eta \mapsto \eta - \mathbf{1}_{a(\eta)} \quad \text{at rate} \quad (1 - q)\mu^a$$

The limit order book as a measure-valued Markov process

$(\eta_t, t \geq 0)$ is a \mathcal{L} -valued Markov process with infinitesimal generator

$$\begin{aligned}
 Af(\eta) &= \sum_{x \leq b(\eta)} \lambda_+^b(x, \eta) [f(\eta + \mathbf{1}_x) - f(\eta)] + \lambda_-^b(x, \eta) [f(\eta - \mathbf{1}_x) - f(\eta)] \\
 &+ \sum_{x < b(\eta)} r^b(x, \eta) [f(\eta - \mathbf{1}_x + \mathbf{1}_{x+1}) - f(\eta)] \\
 &+ \sum_{x > a(\eta)} r^a(x, \eta) [f(\eta + \mathbf{1}_x - \mathbf{1}_{x-1}) - f(\eta)] \\
 &+ \sum_{x \geq a(\eta)} \lambda_+^a(x, \eta) [f(\eta - \mathbf{1}_x) - f(\eta)] + \lambda_-^a(x, \eta) [f(\eta + \mathbf{1}_x) - f(\eta)] \\
 &+ (1 - q)\mu^b [f(\eta - \mathbf{1}_{b(\eta)}) - f(\eta)] \\
 &+ (1 - q)\mu^a [f(\eta + \mathbf{1}_{a(\eta)}) - f(\eta)] \\
 &+ q\mu^b \sum_{x \geq a(\eta)} g(x - a(\eta)) [f(\eta - \mathbf{1}_{b(\eta)} - \mathbf{1}_x) - f(\eta)] \\
 &+ q\mu^a \sum_{x \leq b(\eta)} g(b(\eta) - x) [f(\eta + \mathbf{1}_{a(\eta)} - \mathbf{1}_x) - f(\eta)].
 \end{aligned}$$

for any cylindrical function $f : \mathcal{L} \mapsto \mathbb{R}$.

High-frequency dynamics of the limit order book

High-frequency traders submit orders very frequently, cancel a high percentage of their orders before execution (up to 95%) and maintain a low inventory (i.e. they do not accumulate a large number of buy or sell orders). This is only possible if high frequency traders primarily submit or cancel orders through procedures (ii) and (iv). Thus, in a market where order flow is dominated by HFTs, we expect

$$r^b, r^a \gg |\lambda_+^a - \lambda_-^b|, |\lambda_+^b - \lambda_-^a| \quad \text{and} \quad (1 - q) \ll 1 \quad (7)$$

We translate this into a scaling regime where:

$$r^b, r^a \sim N, \quad |\lambda_+^a - \lambda_-^b|, |\lambda_+^b - \lambda_-^a| \sim \sqrt{N} \quad \text{and} \quad (1 - q) \sim 1/\sqrt{N} \quad (8)$$

High-frequency dynamics of the limit order book

Assumptions on parameters and scaling of initial condition η_N^0

- 1 Assumption 1: the distribution g has compact support.
Interpretation: HFTs post orders near the bid/ask.
- 2 Assumption 2: centering on initial price $\forall N \geq 1, b(\eta_N^0) = S_0$.
- 3 Assumption 3: There exists $\rho_{0,+}, \rho_{0,-} \in C(\mathbb{R}, \mathbb{R}_+) \cap L^1(\mathbb{R})$ with

$$\text{supp}(\rho_{0,+}) \subset (-\infty, 0], \quad \text{supp}(\rho_{0,-}) \subset [0, \infty)$$

and $\forall f \in C_K^0(\mathbb{R}_-), \forall g \in C_K^0(\mathbb{R}_+), \forall \epsilon > 0$

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{N} \sum_{x \in \mathbb{N}} f\left(\frac{x}{N}\right) \eta_N^0(x) - \int_0^\infty f(u) \rho_{0,-}(u) du \right| \geq \epsilon \right) = 0.$$

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{N} \sum_{x \in \mathbb{N}} g\left(\frac{x}{N}\right) \eta_N^0(x) - \int_0^\infty g(u) \rho_{0,+}(u) du \right| \geq \epsilon \right) = 0.$$

A fluid limit for the multicomponent limit order book

Assumption **(F)**: market orders, cancels and limit orders balance each other at scale N : $N(\mu^a(x) - \lambda^a(x)) \rightarrow 0$, $N(\mu^b(x) - \lambda^b(x)) \rightarrow 0$.

Theorem: Under the scaling assumptions 1, 2, 3, **(F)** the measure-valued process

$$\eta_t^N = \frac{1}{N} \left(\eta_t^+ \left(\frac{x}{N} \right), \eta_t^- \left(\frac{x}{N} \right) \right) \quad (9)$$

converges weakly as $N \rightarrow \infty$, in $D([0, \infty), \mathcal{M}(\mathbb{R})^2)$ equipped with the Skorokhod topology, to a measure (μ_t^+, μ_t^-) whose density $(\rho^-(t, x), \rho^+(t, x), t \geq 0)$ is a weak solution of

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} + b_+ \frac{\partial \rho^+}{\partial x}(t, x) + \lambda \rho^+(t, x) \quad \text{for } x < S_t \quad (10)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} + b_- \frac{\partial \rho^-}{\partial x}(t, x) + \lambda \rho^-(t, x) \quad \text{for } x > S_t \quad (11)$$

$$\rho^-(t, x) = 0 \quad \text{for } x > S_t, \quad \rho^+(t, x) = 0 \quad \text{for } x \geq S_t \quad (12)$$

$$\dot{S}(t) = \frac{1}{\theta} \left(\frac{\partial \rho^+}{\partial x}(t, S_t^-) - \frac{\partial \rho^-}{\partial x}(t, S_t^+) \right) \quad (13)$$

Fluid limit: two phase Stefan problem

In particular we have: $\forall t \geq 0, \forall f \in C_K^0(\mathbb{R}), \forall \epsilon > 0$

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{N} \sum_{x \in \mathbb{N}} f\left(\frac{x}{N}\right) \eta_t^N(x) - \int_0^\infty f(u) \rho(t, u) du \right| \geq \epsilon \right) = 0.$$

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{N} b(\eta_t^N) - S_t \right| \geq \epsilon \right) = 0.$$

Fluid limit: two phase Stefan problem

- This is a 'moving boundary' problem, known in physics as the Stefan problem, which captures the average evolution of the limit order book profile.
- The evolution of the price is driven by order flow imbalance as in the Kyle model.
- This is consistent with empirical studies (C., Kukanov, Stoikov 2013) which show evidence for linear impact of small orders on price.
- Different from Lasry & Lions (2007) price formation model.
- This result provides a micro-foundation for these models and relates their parameters to arrival rates and variances of order flows.
- Price moves are generated purely by the market maker/HFT order flow: the other agents are pure liquidity providers but their flow is not directional and perfectly equilibrates at $1/N$ scale in the fluid limit.

Fluid limit: Lasry-Lions (2007) model

Recall that HFTs, once their orders are executed, place a new order of opposite sign at a distance $S(t) + X$ where X is a random variable with distribution g . Here we have assumed

g has compact support

We can recover the Lasry-Lions (2007) models if instead we assume

$$g_N(\cdot) = g(\cdot/N) \text{ weakly converges to } \delta_a$$

This assumption is less natural if the tick/price unit is scaled to zero as $1/N$ since it would imply a 'macroscopic' transaction cost (fixed cost) rather than a proportional cost.

More generally if $g_N \Rightarrow G$ we obtain a variant of the Lasry-Lions (2007) models with a 'smoothed' integral source term delocalized over $S_t \pm \text{supp}(G)$.

Fluid limit: Lasry-Lions (2007) model

Assume that $g_N(\cdot) = g(\cdot/N) \rightarrow \delta_a$ where $a > 0$. Then under the scaling assumptions 2, 3 the measure-valued process

$$\eta_t^N = \frac{1}{N} \left(\eta_t^+ \left(\frac{x}{N} \right), \eta_t^- \left(\frac{x}{N} \right) \right) \quad (14)$$

converges weakly as $N \rightarrow \infty$, in $D([0, \infty), \mathcal{M}(\mathbb{R})^2)$ equipped with the Skorokhod topology, to a measure (μ_t^+, μ_t^-) whose density $(\rho^-(t, x), \rho^+(t, x), t \geq 0)$ is a weak solution of

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} - \frac{\partial \rho^-}{\partial x}(t, S_t) \delta_{S_t - a} \quad \text{for } x < S_t \quad (15)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} - \frac{\partial \rho^+}{\partial x}(t, S_t) \delta_{S_t + a} \quad \text{for } x > S_t \quad (16)$$

$$\rho^-(t, x) = 0 \quad \text{for } x > S_t, \quad \rho^+(t, x) = 0 \quad \text{for } x \geq S_t \quad (17)$$

with initial condition $\rho(0, \cdot) = \lim_N \eta_0^N$.

Homogeneous vs inhomogeneous scaling

In this result, as in the previous work on fluid limits of limit order book models (and all the literature on hydrodynamic limits of particle systems) we use scaling assumptions that are **uniform** in space (price variable). However, it is empirically observed that the intensity of order submissions and cancellations is an order of magnitude higher at the best bid/ask price levels.

This can be modeled by assuming a **different scaling** behavior of intensities of events at the best price level:

$N(\mu^a(x) - \lambda^a(x)) \rightarrow 0$, $N(\mu^b(x) - \lambda^b(x)) \rightarrow 0$ for $x \neq b(\eta)$, $a(\eta)$
while $(\mu^a - \lambda^a(0)) \sim 1/\sqrt{N}$, $(\mu^b(x) - \lambda^b(x)) \sim 1/\sqrt{N}$
at the best bid/ask.

This allows us to take into account temporary (random) imbalances at the best bid/ask, which is a realistic feature of intraday dynamics of supply and demand

A stochastic PDE model for evolution of the limit order book

Intuitively, if the order flow at the best price levels is much higher than other price levels, there exists a scaling regime in which fluctuations vanish away from the interface but not **at** the interface.

Then one expects to obtain a stochastic version of the price dynamics:

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} + b_+ \frac{\partial \rho^+}{\partial x}(t, x) + \lambda \rho^+(t, x) \quad \text{for } x < S_t \quad (18)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} + b_- \frac{\partial \rho^-}{\partial x} + \lambda \rho^-(t, x) \quad \text{for } x > S_t \quad (19)$$

$$\rho^+(t, x) = 0 \quad \text{for } x \geq S_t, \quad \rho^-(t, x) = 0 \quad \text{for } x > S_t \quad (20)$$

$$dS_t = \frac{1}{\theta} \underbrace{\left(\frac{\partial \rho^+}{\partial x}(t, S_t-) - \frac{\partial \rho^-}{\partial x}(t, S_t+) \right)}_{\text{Order flow imbalance}} dt + \sigma dW_t \quad (21)$$

Assume we have a measure-valued process $\rho_t = (\rho_t^+, \rho_t^-)$ with values in $\mathcal{M}(\mathbb{R})^2$ and a process S verifying

$$\frac{\partial \rho^+}{\partial t} = L\rho^+(t, x) \quad \text{for } x < S_t \quad \frac{\partial \rho^-}{\partial t} = L\rho^-(t, x) \quad \text{for } x > S_t \quad (22)$$

$$\rho^+(t, x) = 0 \quad \text{for } x \geq S_t, \quad \rho^-(t, x) = 0 \quad \text{for } x \leq S_t \quad (23)$$

$$dS_t = \frac{1}{\theta} \left(\frac{\partial \rho^+}{\partial x}(t, S_{t-}) - \frac{\partial \rho^-}{\partial x}(t, S_{t+}) \right) dt + \sigma dW_t \quad (24)$$

Then the process ρ is characterized by the property that, for any test function $\varphi \in C_0^\infty([0, T] \times \mathbb{R})$,

$$\begin{aligned} & \langle \rho_t^+, \varphi \rangle = \langle \rho_0^+, \varphi \rangle + \\ & \int_0^t du \left(\langle \rho_u^+, \left(\frac{\partial \varphi}{\partial t} + L^* \varphi \right) \rangle + \frac{\sigma^2}{2} \varphi(u, S_u) \frac{\partial \rho^-}{\partial x}(u, S_{u-}) \right. \\ & \left. \langle \rho_t^-, \varphi \rangle = \langle \rho_0^-, \varphi \rangle + \right. \\ & \left. \int_0^t du \left(\langle \rho_u^-, \left(\frac{\partial \varphi}{\partial t} + L^* \varphi \right) \rangle + \frac{\sigma^2}{2} \varphi(u, S_u) \frac{\partial \rho^-}{\partial x}(u, S_{u-}) \right) \end{aligned}$$

Definition: A weak solution on $[0, \tau)$ of

$$\frac{\partial \rho_{\pm}}{\partial t} = L\rho_{\pm}(t, x), \quad \rho^+(t, x) = 0 \text{ for } x \geq S_t, \quad \rho^-(t, x) = 0 \text{ for } x \leq S_t \quad (25)$$

$$dS_t = \frac{1}{\theta} \left(\frac{\partial \rho^+}{\partial x}(t, S_{t-}) - \frac{\partial \rho^-}{\partial x}(t, S_{t+}) \right) dt + \sigma dW_t \quad (26)$$

is a pair (ρ, S) where S is a semimartingale and $\rho = (\rho_t^+, \rho_t^-, t \in [0, \tau))$ a measure-valued process in $\mathcal{M}(\mathbb{R})^2$ such that

- $\forall t \leq \tau, \rho^+([S_t, \infty)) = 0, \quad \rho^-((-\infty, S_t]) = 0,$
- $\forall t \leq \tau,$ the following limits exist \mathbb{P} – a.s.:

$$\lim_{\epsilon \downarrow 0} \frac{\rho^+[S_t - \epsilon, S_t]}{\epsilon} = \frac{\partial \rho^+}{\partial x}(t, S_{t-}) \quad \lim_{\epsilon \downarrow 0} \frac{\rho^+(S_t, S_t + \epsilon]}{\epsilon} = \frac{\partial \rho^-}{\partial x}(t, S_{t+})$$

- For any test function $\varphi \in C_0^\infty([0, T] \times \mathbb{R}),$

$$\begin{aligned} \langle \rho_t^+, \varphi \rangle &= \langle \rho_0^+, \varphi \rangle + \int_0^t du \langle \rho_u^+, \left(\frac{\partial \varphi}{\partial t} + L^* \varphi \right) \rangle + \frac{\sigma^2}{2} \varphi(u, S_u) \frac{\partial \rho^-}{\partial x}(u, S_u) \\ \langle \rho_t^-, \varphi \rangle &= \langle \rho_0^-, \varphi \rangle + \int_0^t du \langle \rho_u^-, \left(\frac{\partial \varphi}{\partial t} + L^* \varphi \right) \rangle + \frac{\sigma^2}{2} \varphi(u, S_u) \frac{\partial \rho^+}{\partial x}(u, S_u) \end{aligned}$$

High-frequency dynamics of the limit order book

Assumption **(D)**: imbalance of order \sqrt{N} at the best bid/ask :
 $\sqrt{N}(\mu^a - \lambda^a(0)) \rightarrow \sigma^2$, $\sqrt{N}(\mu^b - \lambda^b(0)) \rightarrow \sigma^2$

Theorem: Under Assumptions 1, 2, 3, **(D)** η_t^N converges weakly as $N \rightarrow \infty$ to a measure-valued process (μ_t^+, μ_t^-) whose density $(\rho^-(t, x), \rho^+(t, x), t \geq 0)$ is a weak solution of the stochastic partial differential equation

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} + b_+ \frac{\partial \rho^+}{\partial x}(t, x) + \lambda \rho^+(t, x) \quad \text{for } x < S_t \quad (27)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} + b_- \frac{\partial \rho^-}{\partial x} + \lambda \rho^-(t, x) \quad \text{for } x > S_t \quad (28)$$

$$\rho^+(t, x) = 0 \quad \text{for } x \geq S_t, \quad \rho^-(t, x) = 0 \quad \text{for } x > S_t \quad (29)$$

$$dS_t = \underbrace{\frac{1}{\theta} \left(\frac{\partial \rho^+}{\partial x}(t, S_{t-}) - \frac{\partial \rho^-}{\partial x}(t, S_{t+}) \right)}_{\text{Order flow imbalance}} dt + \sigma dW_t \quad (30)$$

with initial condition $\rho(0, \cdot) = \lim_N \eta_0^N$.

A stochastic 2-phase Stefan problem

Proposition: the "stochastic 2-phase Stefan problem"

$$\frac{\partial \rho^+}{\partial t} = \frac{\sigma_+^2}{2} \frac{\partial^2 \rho^+}{\partial x^2} + b_+ \frac{\partial \rho^+}{\partial x}(t, x) + \lambda \rho^+(t, x) \quad \text{for } x < S_t \quad (31)$$

$$\frac{\partial \rho^-}{\partial t} = \frac{\sigma_-^2}{2} \frac{\partial^2 \rho^-}{\partial x^2} + b_- \frac{\partial \rho^-}{\partial x} + \lambda \rho^-(t, x) \quad \text{for } x > S_t \quad (32)$$

$$\rho^+(t, x) = 0 \quad \text{for } x \geq S_t, \quad \rho^-(t, x) = 0 \quad \text{for } x > S_t \quad (33)$$

$$dS_t = \frac{1}{\theta} \underbrace{\left(\frac{\partial \rho^+}{\partial x}(t, S_t^-) - \frac{\partial \rho^-}{\partial x}(t, S_t^+) \right)}_{\text{Order flow imbalance}} dt + \sigma dW_t \quad (34)$$

admits a weak solution (ρ, S) where

- ρ is a random field on $[0, \tau] \times \mathbb{R}$ with continuous sample paths
- S is a semimartingale with decomposition (34)
- $\tau = \{t > 0, \sup(|\Delta \partial_x \rho(t, S_t)|) < \infty\}$

Diffusion limit: a stochastic two-phase Stefan problem

- Two-phase moving-boundary problem with "random forcing" at the boundary: stochastic version of the two-phase Stefan problem.
- Similar (but not identical) stochastic moving boundary problem studied recently by Mueller et al (2013), Keller-Ressel & Müller (2014).
- Price follows a "diffusion in a random environment" defined by the limit order books:
 - its drift is generated by price impact of the HFT/market maker order flow at the bid/ask
 - its volatility is driven by stochastic imbalance between the market and limit orders at the best-bid ask.
- Perpetual competition between
 - the random drift term (action of 'market marker') which stabilizes the price and
 - the white noise term which prevents the price from settling down.

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 (2011) Statistical modeling of high frequency data: facts, models and challenges, **IEEE SIGNAL PROCESSING**, Vol 28, No 5, 16–25.
- Rama Cont and Adrien de Larrard (2013) Price dynamics in a Markovian limit order market, **SIAM Journal on Financial Mathematics**, Vol 4, 1–25.
- Rama Cont and Adrien de Larrard (2011) Order book dynamics in liquid markets: limit theorems and diffusion approximations, <http://ssrn.com/abstract=1757861>.
- Rama Cont and Adrien de Larrard (2012) Price dynamics in limit order markets: linking volatility with order flow, Working Paper.
- Rama Cont (2014) High frequency dynamics of limit order markets: an asymptotic analysis.