

# The Emergence of Continuous Trading

Market Microstructure in Practice 1/3

Charles-Albert Lehalle

Senior Research Advisor, Capital Fund Management, Paris

April 2015, Printed the April 13, 2015



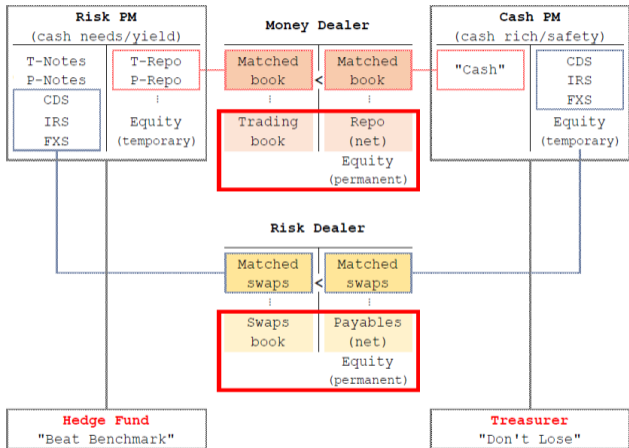
- 1 The Role of Financial Markets Inside the Financial Network
- 2 Recent Evolutions of Financial Markets
- 3 Fragmentation(s)

- 1 The Role of Financial Markets Inside the Financial Network
- 2 Recent Evolutions of Financial Markets
- 3 Fragmentation(s)

If we take the viewpoint of looking at the financial network **from the outside**, we need to understand its inputs and outputs, and deduce the features it provides to the rest of the economy.

- ▶ We can see how the banking network operates a maturity transformation between natural borrowers at different maturities (mid or long term) and natural lenders (short or mid term).
- ▶ Banks are intermediaries: their goal is to have no remaining risk in their inventories. It is the goal of regulators to maintain their risk as low as possible (using capital requirements).
- ▶ The bad cases are when all the banks host risks in the same direction (2008), instead of having a diversification at the scale of the whole system.
- ▶ Nevertheless inflows and outflows in Banks balance sheets (i.e. *transactions*) are not simultaneous, hence regulators need to give them some freedom to wait for a seller once they sold a contract to a buyer (and the reverse).

Let's see this viewpoint is typically a microstructural one: intermediaries, buyers and sellers, inventory risk... More of the buzz words are there...



Source: Pozsar on Shadow Banking (2013)

On the one side (right) you have **Cash PMs** they are *cash rich* but *safety poor* ("fear to loose their money").

On the other side (left) you have **Risk PMs**, they have to beat a benchmark, thus are *securities rich* but *return poor* (the need leverage and non linearities).

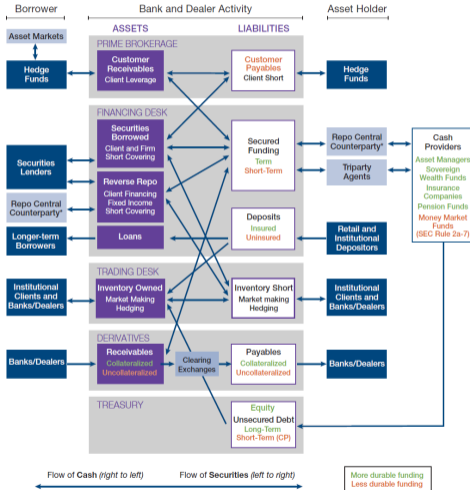
In between (middle) you have **intermediaries**, they match Risk PMs on the asset side of their balance sheet and Cash PMs on the liability side.

This is accountancy, each time a transaction of this kind is made, it has to be marked- to-market, thus all this is pegged to traded prices.

Going back to concepts mathematical finance is more familiar with:

- ▶ you are an investment bank, you sell a structured product or a derivative to clients;
- ▶ you do not hedge each book separately (or at least you shouldn't): you hope to have other clients consuming other products flattening your (risk) inventory.
- ▶ The simplistic traditional example is: you sell a protection to bad/good weather to a farmer, and a protection to good/bad weather (i.e. the reverse) to an electricity producer. You have nothing to hedge, just take your commercial margin.
- ▶ Pricing under risk neutral measure is a way to price under *sellable risk*.
- ▶ Of course you will not succeed in netting 100% of the risk, hence **you have to hedge the remaining book**, in the markets (we hope they use optimal trading algorithms –i.e. continuous trading– to do this).
- ▶ But one step further: if you succeed into hedging continuously on markets (without liquidity, i.e. market impact, issues), it just mean someone has the opposite risk in the market and hedges it on its side: you should / could find it and net both positions (think about the crucial role of CCP here).
- ▶ In this sense **wrong way risk is not good for the liquidity on markets** at all, you cannot believe you really hedge if you impact the price.

Two good (but stylized) examples in the literature are [Stoikov and Saglam, 2009] and [Carmona and Webster, 2012].



Source: OFR Annual Report 2013

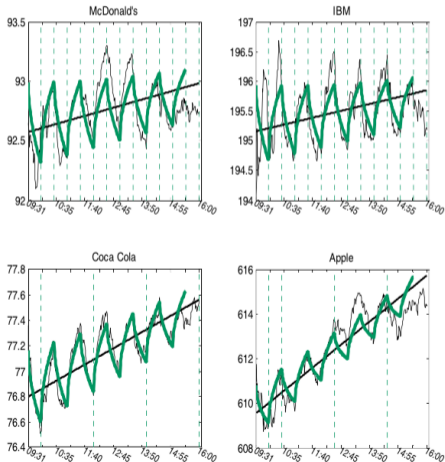
Another (US) view of the same ecosystem.

In this diagram we have CCPs (Central Counterparty), Clearing Exchanges, and Triparty Agents (repo third-parties). They are needed to really net the risk.

Having few identified, regulated *unconflicted intermediaries* is comfortable for regulators and policy makers.

Since the G20 Pittsburgh Summit (2009), the goal is to secure most of these transactions by cash or collateral deposit and netting (positions and contracts).

FIGURE 1: SAWTOOTH PATTERNS ON COCA-COLA, MCDONALD'S, IBM AND APPLE ON 19 JULY 2012



Even on liquid stocks and for vanilla options (close to maturity in this case), hedging can go wrong.

The 19th of July 2012, a trading algorithms bought and sold shares every 30 minutes without any views on its market impact.

For one visible mistake like this on liquid underlyings of vanilla products, how many more sophisticated bad external hedging on less liquid (even OTC) markets...

Anonymous continuous hedging of a remaining position outside of the bank does not mean all is going well. Example of recent literature on this specific issue: [Guéant and Pu, 2013], [Li and Almgren, 2014].

But nothing more generic, for instance the whole process of hedging books in presence of wrong way risk is not studied (as far as I know). One step in this direction is [Schied and Zhang, 2013].



My advices to an investment bank:

- ▶ **Net all your books** , maintain two opposite positions is costly and risky,
- ▶ If you can't it can be because you do not communicate enough internally (they are Chinese walls anyway...), hence **go to the market** ,
- ▶ But **before try to match your small metaorders** : send them to an internal place and cross them as much as possible;
- ▶ You will have synchronization issues (at the level of these metaorders, no reason to be synchronized), ask to your traders to **implement facilitation-like market making** schemes inside the bank.
- ▶ The remaining quantity has to be sent to markets as smoothly as possible, but it does not mean you will have no impact. **Who is your counterpart in the market** should be an obsession: if you trade a one way risk, you will pay it in the future...

- 1 The Role of Financial Markets Inside the Financial Network
- 2 Recent Evolutions of Financial Markets
- 3 Fragmentation(s)

During one of he panel of the last Microstructure conference last Dec. in Paris, we had an Exchange, an Alternative venue, a Market Maker, and a Regulator.

To the question **According to you: what is an Exchange?** I had four different answers:

1. the owner of the **price formation process** ;
2. a place to **finance the economy** ;
3. a place where **risk decreases** ;
4. where **sellers meet buyers** .

Who said what?

During one of he panel of the last Microstructure conference last Dec. in Paris, we had an Exchange, an Alternative venue, a Market Maker, and a Regulator.

To the question **According to you: what is an Exchange?** I had four different answers:

1. the owner of the **price formation process** ; Regulator
2. a place to **finance the economy** ; Exchange
3. a place where **risk decreases** ; Market Maker
4. where **sellers meet buyers** . Alternative Venue

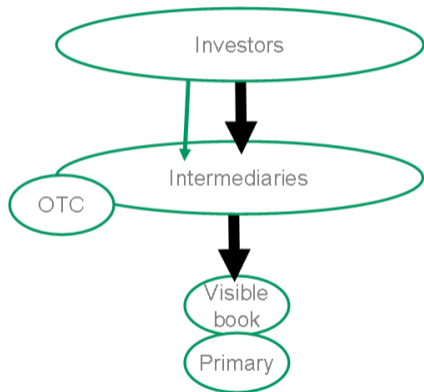
Who said what?

I would add **A firm competing to gain clients and earn money**

- ▶ Batch (Fixing) auctions or Continuous auctions,
- ▶ price driven or order driven logic,
- ▶ bilateral or multilateral trading.



A Limit Order Book (LOB) hosts multilateral, order driven, continuous double auctions.



The traditional viewpoint on microstructure is you have three distinct layers:

- ▶ the **Buy-side**, i.e. Investors
- ▶ the **Sell-side**, i.e. intermediaries (Investment banks and Brokers), selling services to Investors (including liquidity, i.e. access to market makers)
- ▶ **Market Operators**: exchanges, CCPs, etc.

The last layer is not always in the picture (think about [Kyle, 1985]), but more recently it is (see [Foucault and Menkveld, 2008]).

MiFID in Europe and Reg NMS in the US opened the door to more competition between exchanges. Targeting:

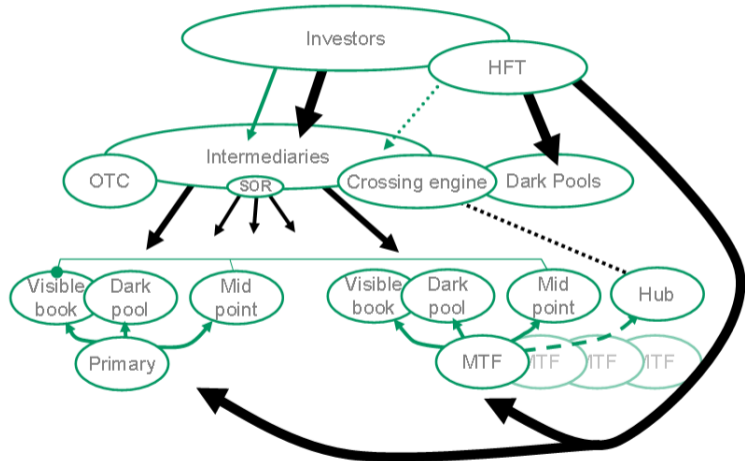
- ▶ increase the **quality of service** ,
- ▶ better, more meaningful and easier to **access information** ,
- ▶ decrease **prices** ,
- ▶ incentive to **innovation** .

MiFID in Europe and Reg NMS in the US opened the door to more competition between exchanges. Targeting:

- ▶ increase the **quality of service** , **more (enough?) reliable and faster matching engines**
- ▶ better, more meaningful and easier to **access information** , **web sites of Chi-X BATS and fidessa**
- ▶ decrease **prices** , **lower trading fees, maker/taker fees**
- ▶ incentive to **innovation** . **Pegged orders, hidden orders, new matching rules**
- ▶ No significant change in **listing**.
- ▶ No competition on **fixing (close) auctions**.

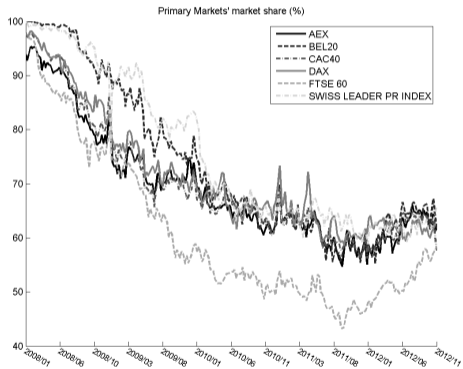


# The emergence of a new market structure



The trading now takes place on a distributed network of heterogenous trading platforms.

1. MTFs offered trading in a **one size fits all** approach,
2. they needed liquidity providers, thus offered rebate to limit orders,
3. statistical arbitrage technology (overnight risk) met automated market makers to produce **High Frequency Traders**.
4. As HFT provided competitive quotes, Brokers had to implement **Smart Order Router** (under regulatory pressure, since MiFID decided to implement a two-layered competition),
5. This sudden **increase of complexity** (and a lack of education) frightened some participants (suspecting that the profit of HFT were their lost),
6. Trading platforms offered **anonymity** features, claiming it will reduce information leakage (birth of European **Dark Pools**),
7. Exchanges provided **internalization** features, and brokers provided crossing capabilities,
8. High fixed costs and high competition conducted to **mergers** (LSE + Turquoise, BATS + Chi-X, ICE + Nyse, Euronext + ?, etc);
9. Exchanges are close to be **technology vendors**.



Fragmentation is mainly visible in terms of trading flows, but it is also significant in terms of

- ▶ communication protocols,
- ▶ order and transaction types.

Hence an investor has more choice.

## A standard optimization paradox

- ▶ The more parameters you add to a situation, the **better the optimum**.
- ▶ But the **more complex** to find.



# A standard optimization paradox

- ▶ The more parameters you add to a situation, the **better the optimum**.
- ▶ But the **more complex** to find.



Quote: BNPPpa.CHI

BNPPpa.CHI	Type:	BNP PARIBAS	FR0000131104	INS /102	EUR	09APR08	11:07
Last	Net.Chng	Trade Vol	Buy Vol	Buy Sell	Sell Vol		
68.5000	-0.9050	50	200	68.4950x68.5250	200		
68.5000	Pct.Chng	Volume	400	68.4900x68.5300	400		
68.5000	-1.3 %	60098	560	68.4850x68.5450	198		
68.5000	Buy-Sell ID	Value	200	68.4800x68.5500	200		
68.5000	- / -	4131007.25	880	68.4700x68.5550	500		
	Moves		380	68.4650x68.5600	200		
	326	15	200	68.4550x68.5650	340		
Open	High	52.Wk.HI:	400	68.4350x68.5700	200		
68.8700	69.1700		1011	68.0800x68.5750	880		
CL: 08APR08	Low	52.Wk.Lo:	1000	68.0500x68.5800	380		
69.4050	68.2900						

Quote: 0#BNPP.PA

0#BNPP.PA	BNP PARIBAS	FR0000131104	BNP	PAR	EUR	09APR08	10:08		
Last	TrdVol	Pct.Chng	Order	Bid Size	Bid 10:08	Ask	Ask Size	Order	
68.5000	100	-1.69 %	11	6524	68.5000	68.5450	359	2	
68.5450	Net.Chng	-1.18	11	6524	68.5000	68.5450	359	2	
68.5450	Volume	744392	2	1074	68.4950	68.5500	396	3	
68.4950			1	200	68.4800	68.5550	344	1	
68.4900	Turnover	News	1	200	68.4600	68.5650	200	1	
VWAP	68.6243	51157744	09:03	1	202	68.4500	68.5700	200	1
Moves	2234								
Up/LoLimit:	72.3650 / 65.4750		RHP	500000	68.4950 / 68.6450	500000	RHP		

- ▶ The more parameters you add to a situation, the **better the optimum**.
- ▶ But the **more complex** to find.
- ▶ simultaneously, a large academic literature emerged to **optimize the trading process** (see [Lehalle, 2013] in the *Handbook on systemic risk*, 2013 for a review):
- ▶ as a result, a large trader can now liquidate a position using a majority of limit (liquidity adding) orders. Optimal schemes to seek liquidity on a large set of pools are known [Pagès et al., 2011] (even the market making problem has been largely studied [Guéant et al., 2013], [Avellaneda and Stoikov, 2008], etc).

**The notion of liquidity changed.** A dynamical and probabilistic approach is now needed.

- ▶ Since the **trading is now pan European** (this is a success of MiFID)
- ▶ market making has to be understood at this scale.
- ▶ Moreover, now that the **dynamics of liquidity** are more important than the usual static measures,
- ▶ the service provided by market makers has to be redefined too.
- ▶ It is important to understand what service they provide when everyone is splitting his orders in so many small pieces...

Market making is easy to understand when it is described in a bilateral context (cf. [Gabaix et al., 2006]):

- ▶ Each **market maker has a bilateral relationship with investors**,
- ▶ he answers to specific solicitations.
- ▶ This information asymmetry gives him a protection against adverse selection.
- ▶ In a multilateral context, Nyse's specialists inherited from such features, implementing a two-layered market.

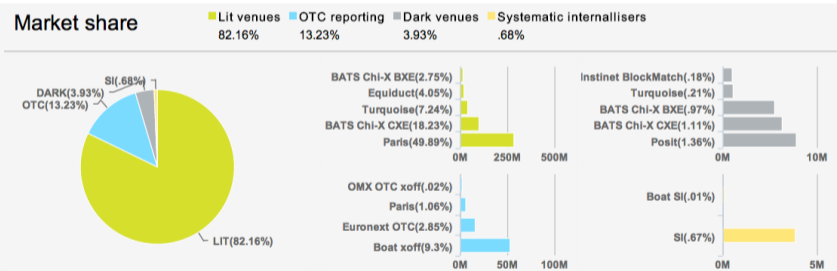
Today's market structure is in favour of **multilateral trading**. To be protected against adverse selection when counterparts are anonymous and having to disclose his prices and quantities to everyone is a challenge...



- 1 The Role of Financial Markets Inside the Financial Network
- 2 Recent Evolutions of Financial Markets
- 3 Fragmentation(s)

## CREDIT AGRICOLE

Monday, 2013 November 25 to Friday, 2013 November 29. View this in [Fragulator Live](#)

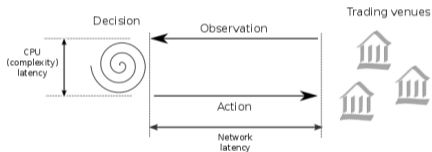


These numbers add all the transactions: they are good to estimate exchanges revenues.

Regulations (Reg NMS and MiFID) implemented competition across trading venues: Exchange / Regulated Markets; ECN / MTF (Multilateral Trading Facilities).

### In the US

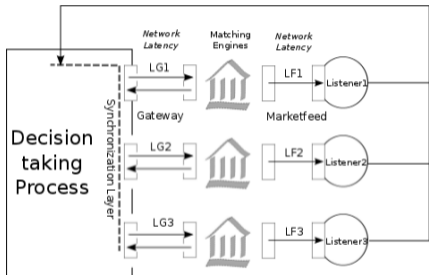
- ▶ The “trade through rule” demand to trading venues to re-route orders that can obtain a better price “somewhere”;
- ▶ That for, a “National Bid and Offer” (NBBO) is available to venues;
- ▶ 3 SIP (centralized Securities Information Processor) aggregate quotes coming from all venues to build a **Consolidated Tape**. They are maintained by the **The Consolidated Tape Association (CTA)**.
- ▶ They are then sent back to venues so that the NBBO is known by every one.
- ▶ Thanks to this mechanism an order can be send “blindly” to any venue and take profit of competition (on the paper...).



In Europe, no similar mechanism have been implemented

- ▶ because when you compare two best bids (or asks) in the US, if one is larger than the other, you can be sure you will have more selling on this venue at the end of the day, even if you count the post trading costs (clearing and settlement costs mainly: there is only one silo).
- ▶ In Europe post trading costs can be different from one venue to the other (because of post trade costs),
- ▶ hence a consolidated tape will not tell a trader a price is more attractive on one venue or on the other.
- ▶ The Regulator demanded to broker to implement Smart Order Routers on their side.

A SOR is parametrized to implement a “Best Execution Policy” (choice of the venues, priority / ranking rules, etc).



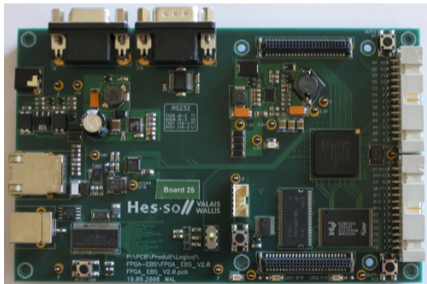
The latency complexifies the story:

- ▶ If a SOR is at 1 ms from venue *A*, 2 ms from venue *B* and 3 ms from venue *C*,
- ▶ when he look at the three orderbooks, they are **not synchronized**,
- ▶ since the SOR has an SLE connection, it can have some insight about the accuracy of its snapshots.
- ▶ If the SOR see 20 shares at the offer at 10.00 on *A*, 30 at the same price on *B* and 50 on *C*, and if the SOR want to buy 100 shares, it can split it in  $20(A)+30(B)+50(C)$  and send the three orders (i.e. 3 insert IOC Buy orders at 10.00) at  $t_0$
- ▶ but the 3 orders will reach the venues at  $t_0 + 1$ ,  $t_0 + 2$  and  $t_0 + 3$  ms (respectively).

If 20 of the 50 shares at *C* are duplicates of the 20 at *A* owned by a trader hosted at *A* and with a latency of 1 ms to *C*, he can cancel his shares before the SOR order reaches *C*.

In such a case the SOR will obtain  $70=20(A)+20(B)+30(C)$  instead of the 100 it expected.

One solution is to delay the order to *A* by 2 ms and to *B* by 1 ms, so that the three orders reach the venues simultaneously. But the delays are not that deterministic, and waiting more means having a largest probability of an exogenous orderbook change.



Hardware can be useful.

- ▶ GPU is not that good since it is slow to transfer data from a motherboard to the GPU, and in trading we talk about being able to deal with a fast flow of data.
- ▶ The flow to work on is made of the SLC.
- ▶ FPGA is better since it can embed physical IP layers, hence you can put on an FPGA: (1) the reading of the flow, (2) its enrichment with indicators. Without paying that more latency.



Low latency means to be able to react first to external information. It is not mandatory for this external information to be frequent. To implement such strategies you must be fast on long distances.

High frequency means to be able to (re)adjust your orders 100 times per second. To implement this you need to be very close to the orderbooks you are working in.

You cannot be simultaneously close to 2 venues that are not located in the same datacenter.

Remark on Timestamping and Event Ordering:

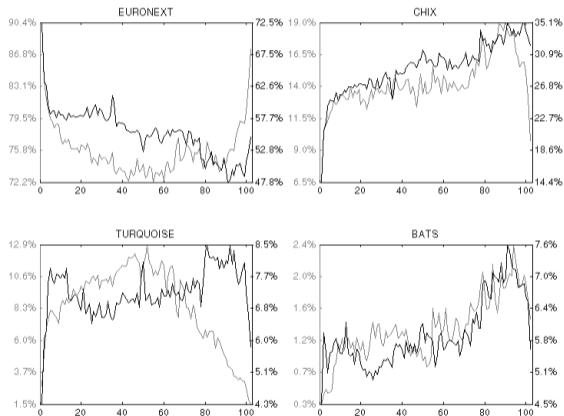
- ▶ Two events occurring simultaneously at points  $A$  and  $B$  in space are perceived in a different orders with respect to the position of the observer.
- ▶ When an observer suffers from a delay, he can choose between (1) taking decisions on estimators of the state of the orderbook, or (2) implementing a simultaneous optimization on its strategy.

As an example, **to trade in Dark Pools** :

- ▶ [Ganchev et al., 2010] estimates the liquidity in each pool and implements a deterministic optimization;
- ▶ [Agarwal et al., 2010] uses a minimum regret approach;
- ▶ [Pagès et al., 2011] implements a stochastic implementation of an optimal trading scheme.

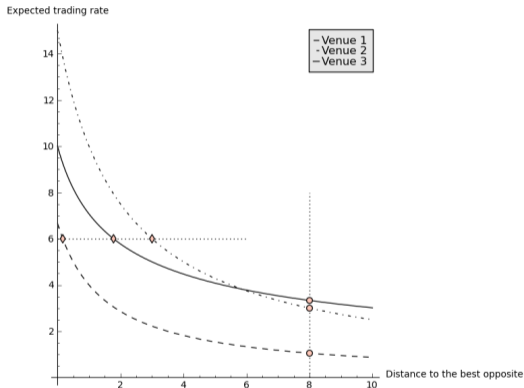
The first approach is good for opportunistic trading (hedge fund), the second for a rare and not really flexible flow (investor), the last one is good for very large systematic flow (broker).





- ▶ The probability that a trade typically occurred on a given venue should take into account the time,
- ▶ the probability to find liquidity on a venue? (given a trade occurred)
- ▶ The Average Trade Size on each venue (is in fact the average size of a limit order..).
- ▶ Cost of a “roundtrip” for a given amount of Euros on each venue.

Market Share of venues on a typical Fench Stock in 2012



Given that you know in advance the flow  $\Phi_k$  of liquidity consumption on each venue  $k$ , you can choose any combination (to focus on price or speed). The estimation of the flow is difficult, since the picture you see is outdated at the rhythm of the updates of the orderbook: if the updating rate  $\rho$  is 250Hz, then the time to react is:

$$\tau_{\text{reaction}} = \frac{1}{2\rho} = 2ms$$

In [Pagès et al., 2011] and [Laruelle et al., 2013], we show how to adjust the limit price and the quantities to send to each venue for an arbitrary criterion (in a stochastic world). And we provide uncertainty bounds on the result.

We have seen

- ▶ Intermediation does not take place on electronic markets only, it is one of the main mechanism supporting the functioning of the financial system. Intermediaries guarantee anonymity and price diffusion (cf [Merton, 1995]), not only on electronic markets, but for all their inflows and outflows. They **hedge their remaining risk on electronic markets**.
- ▶ With recent evolutions, they can do it more continuously than ever, and they have more occasions than ever to **cross their flows** (less sophisticated products means more flows to maintain incomes).
- ▶ The **fragmentation comes with competition accros trading venues**, could trading be more simple? we are in a transition phase. Automation can help.
- ▶ **Trading simultaneously** on several platforms is complex, and latency is of importance.

We have two simultaneous problems:

- ▶ planification / scheduling (inside each intermediary, or *just outside it*),
- ▶ synchronization of small remaining parts of the metaorders in anonymous pools.

In terms of methodologies: I would say **stochastic control** is good for the former, and **statistical learning** (stochastic algorithms) for the later.



Agarwal, A., Bartlett, P. L., and Dama, M. (2010).

Optimal Allocation Strategies for the Dark Pool Problem.

In Teh, Y. W. and Titterton, M., editors, *Proceedings of The Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 9, pages 9–16.



Avellaneda, M. and Stoikov, S. (2008).

High-frequency trading in a limit order book.

*Quantitative Finance*, 8(3):217–224.



Carmona, R. and Webster, K. (2012).

High Frequency Market Making.



Foucault, T. and Menkveld, A. J. (2008).

Competition for Order Flow and Smart Order Routing Systems.

*The Journal of Finance*, 63(1):119–158.



Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2006).

Institutional Investors and Stock Market Volatility.

*The Quarterly Journal of Economics*, 121(2):461–504.



Ganchev, K., Nevmyvaka, Y., Kearns, M., and Vaughan, J. W. (2010).

Censored exploration and the dark pool problem.

*Commun. ACM*, 53(5):99–107.



Guéant, O., Lehalle, C.-A., and Fernandez-Tapia, J. (2013).

Dealing with the inventory risk: a solution to the market making problem.

*Mathematics and Financial Economics*, 4(7):477–507.



Guéant, O. and Pu, J. (2013).

Option pricing and hedging with execution costs and market impact.



Kyle, A. P. (1985).

Continuous Auctions and Insider Trading.  
*Econometrica*, 53(6):1315–1335.



Laruelle, S., Lehalle, C.-A., and Pagès, G. (2013).

Optimal posting price of limit orders: learning by trading.  
*Mathematics and Financial Economics*, 7(3):359–403.



Lehalle, C.-A. (2013).

Market Microstructure knowledge needed to control an intra-day trading process.  
In Fouque, J.-P. and Langsam, J., editors, *Handbook on Systemic Risk*. Cambridge University Press.



Li, T. M. and Almgren, R. (2014).

Option Hedging with Smooth Market Impact.  
Technical report.



Merton, R. C. (1995).

A Functional Perspective of Financial Intermediation.  
*Financial Management*, 24(2):23+.



Pagès, G., Laruelle, S., and Lehalle, C.-A. (2011).

Optimal split of orders across liquidity pools: a stochastic algorithm approach.  
*SIAM Journal on Financial Mathematics*, 2:1042–1076.



Schied, A. and Zhang, T. (2013).

A state-constrained differential game arising in optimal portfolio liquidation.



Stoikov, S. and Saglam, M. (2009).

*Option market making under inventory risk*.