

# Summary of the “thesis”

- **Part I:** backtesting will be different than live trading due to micro-structure “games” that can be played (often by high-frequency trading) which affect execution details.
- This might not impact you if you only use aggressive orders (i.e. no passive limit order), and it might not impact you if your trading frequency is low enough that losing a couple of cents per trade won’t affect your bottom line. It could still impact you if you use order book asymmetry/order flow in your strategy.
- **Part II:** proposing a solution for simulating the order book dynamics to try to close the gap between backtesting and live.

# Assumptions

- I will assume that you're already doing everything right:
  - allowing for slippage
  - accounting for dividends
  - transaction costs
  - testing out-of-sample
  - using high quality data
  - using NBBO data
- I will not address further issues that could come from illegal practices such as flickering quotes, etc. I will also assume that your broker is regulated (i.e. not FX, for example), so it isn't front-running you.

# Part I

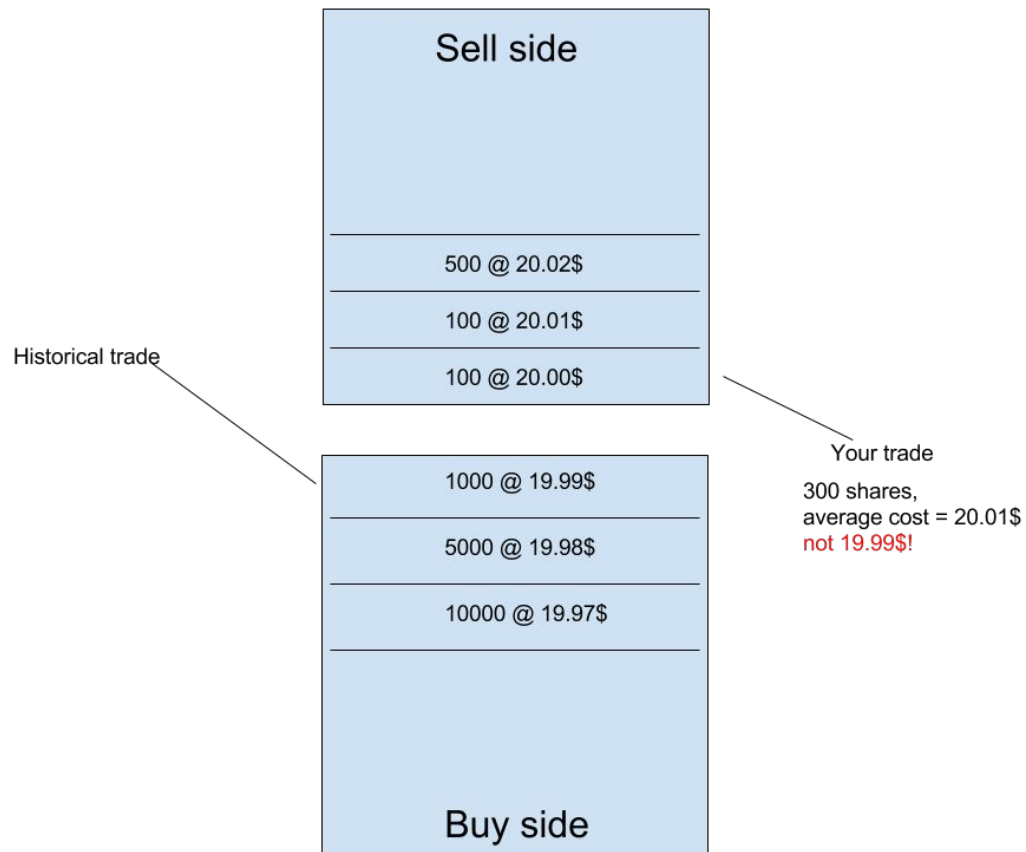
---

Sources of discrepancies between backtesting & live

# Why order book data instead of trades data

1. When buying or selling, you're doing so at the ask/bid price (respectively), not the mid-point or the last historical trade price (which could be either side of the spread).
2. If the size of your order is bigger than the volume available at the top-of-book price, your backtest must account for that, and simulate "walking the book" (i.e. getting the next best price).

# Why order book data instead of trades data



# Cents matter (in intraday trading)

Some math:

Suppose you trade on **average 2 times per day**, with an **average trade size of 1000 shares (and an initial capital of 30,000\$)**.

Then a systematic **1 cent loss** means 1000 shares times  $0.01\$ = 10\$$  per trade,

$10\$ * 2 * 252$  trading days = 5,040\$ differential!

On a capital of 30,000\$ that's a P&L differential of **-16.8%**

# Aggressive vs passive orders

Aggressive orders (Market orders & marketable limit orders):

- You pay the spread
- No rebate

Passive orders (Limit orders that are pending in the order book):

- You earn the spread
- You can rebates (lower execution cost overall)

We **want** to use passive orders, but for the following reasons, it's very difficult.

# Micro-structure plays that affect backtesting

There are 2 components to the main thesis:

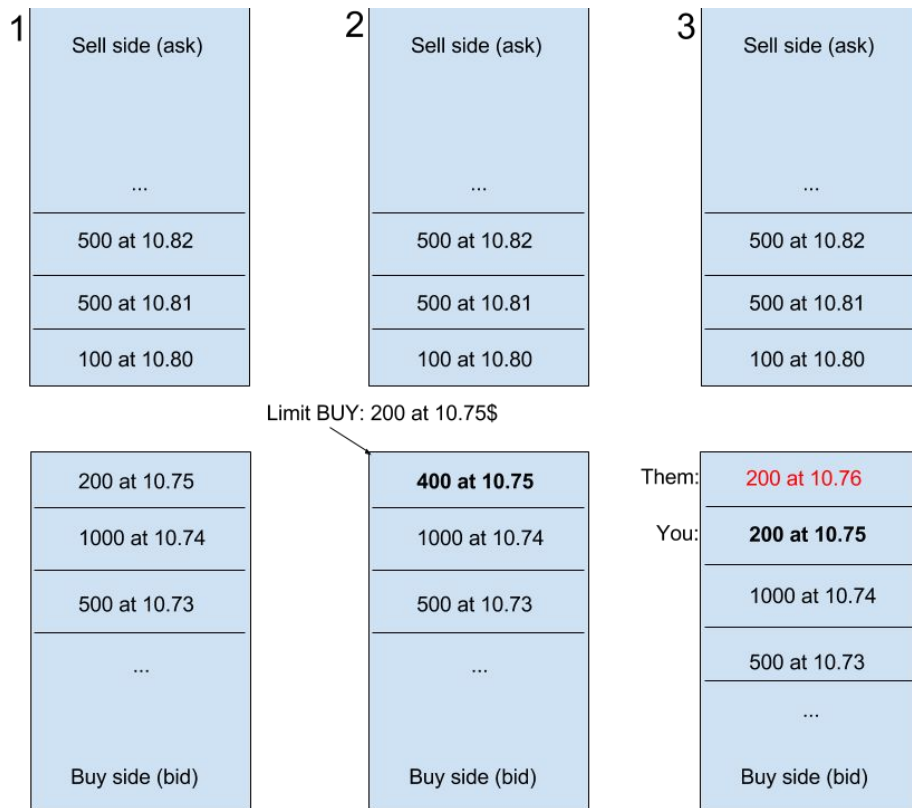
- When you place passive orders, you expose information to the other participants, and they will react to you. In backtesting, you don't have that effect (which is often negative due to adverse selection principles), but in live trading you do.
- The liquidity displayed in the order book does not correspond to the actual liquidity available. So if your backtesting uses the order book to estimate if an order would be filled or not, and at what price, the backtested versus actual values will be different.



# Example #1: automated quote-matching

- As soon as you place a passive limit order, at current best bid/ask, some algorithms will notice that and beat your price by cancelling their current orders and re-posting them at a better price.
- **Conclusion:** your limit orders that get executed in backtesting (because no one beat its price) won't necessarily get executed in live trading (instead, the liquidity consumers will take up the offers from the other one)

# Example #1: automated quote-matching



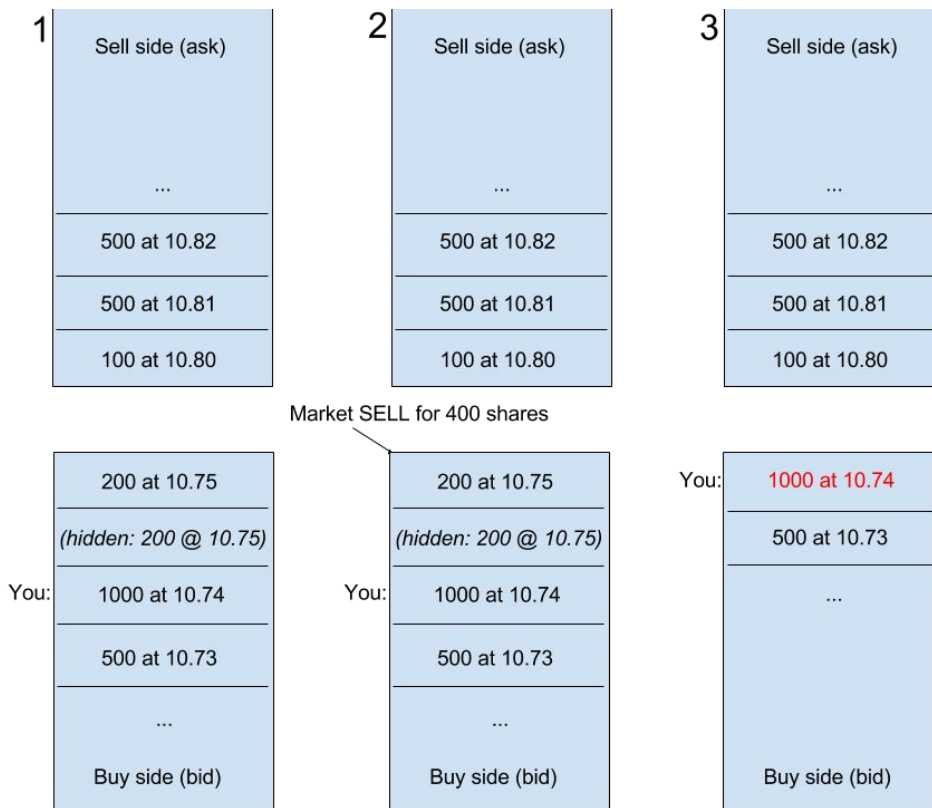
# Example #1: automated quote-matching

- Isn't that front-running, and thus illegal?
- **No:** front-running is when a broker uses private advance knowledge that somebody is going to trade, to place a trade ahead of the trader.
- In this case, the order has already been posted, and is public information. Nothing prevents anyone from reposting at a better price to beat them.

## Example #2: hidden liquidity (icebergs)

- This applies to both passive and aggressive orders
- Some (most?) exchanges offer limit order modifiers to hide volume (often called Hidden Limit Orders, or less formally, iceberg orders). TSX, among others, has that feature.
- 22% of Nasdaq limit orders have hidden volume.
- **Conclusion:** if your limit order is behind that iceberg (invisible wall of liquidity), it will never be taken in live trading, but it backtesting it is taken. For aggressive orders, it implies that you get more orders at that hidden price, instead of going deeper into the book.

# Example #2: hidden liquidity (icebergs)



## Example #2: hidden liquidity (icebergs)

Note: if you use a predictive model based on order book asymmetry, order flow, etc. your model is affected by this.

You don't have direct knowledge of the actual liquidity in the order book.

## Example #3: phantom liquidity (?)

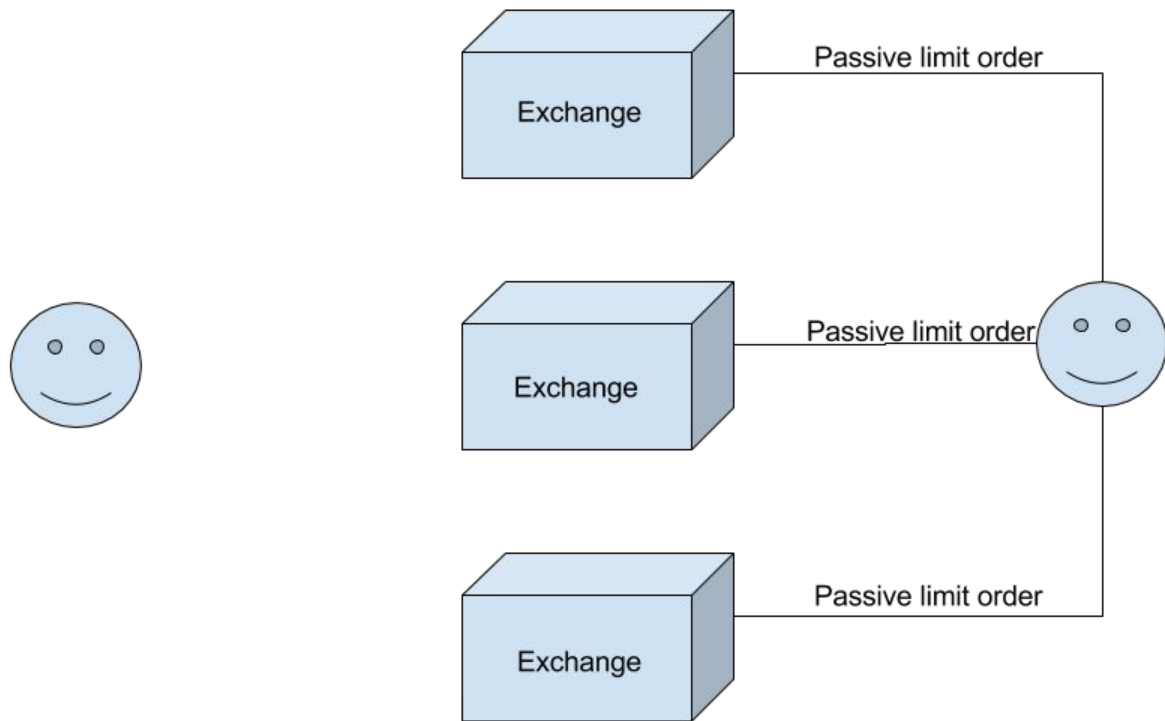
- Disclaimer: not sure if this is still an issue. It definitely was a few years ago. Due to the secretive nature of HFT tactics, it's difficult to tell if they're still able to do something like this. It is possible that legal and technical advances over the last few years have eradicated this risk.
- There were many things that could cause vanishing liquidity. Most of these are now illegal.
- I will provide a recent example of how HFTs could cancel their orders just ahead of your aggressive orders. May or may not still apply.

## Example #3: phantom liquidity (?)

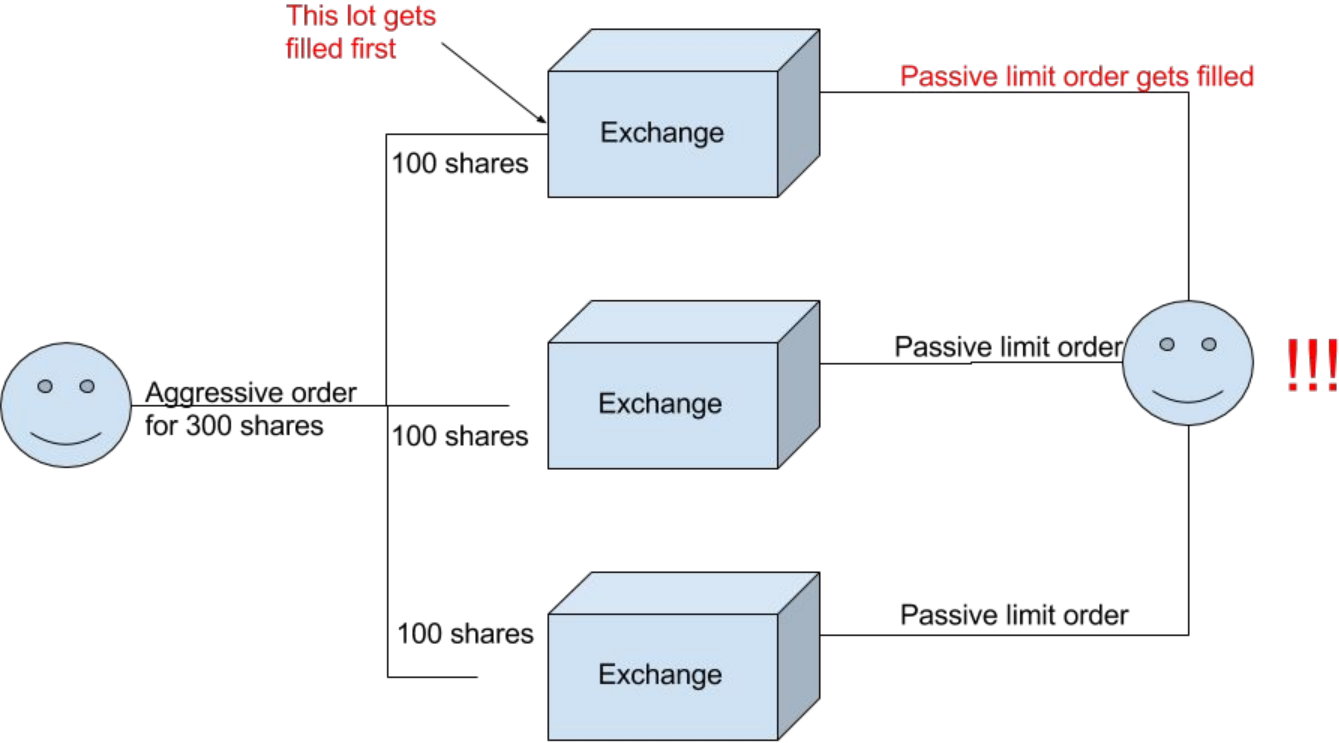
- When you submit an aggressive order, it can get broken down into smaller lots and be executed across different exchanges (to be able to give you best bid/ask price execution without walking the book).
- Unfortunately, your order won't reach all exchanges at exactly the same time. It might come in at some exchanges a few microseconds before another exchange.
- HFTs could see an incoming order on one exchange, and you will get filled at the current bid/ask on that exchange, but then they would cancel their offers on the other exchanges before you, causing liquidity to vanish on those other exchanges and giving you a worse fill price.



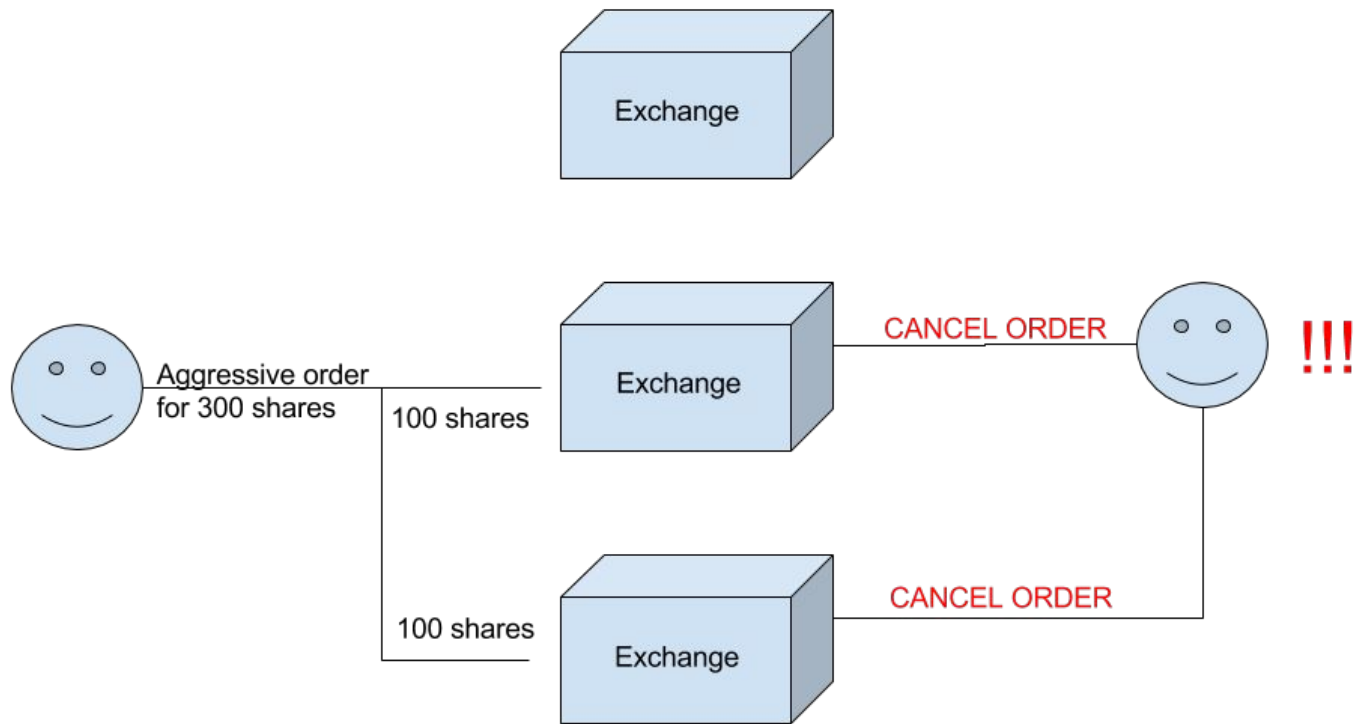
## Example #3: phantom liquidity (?)



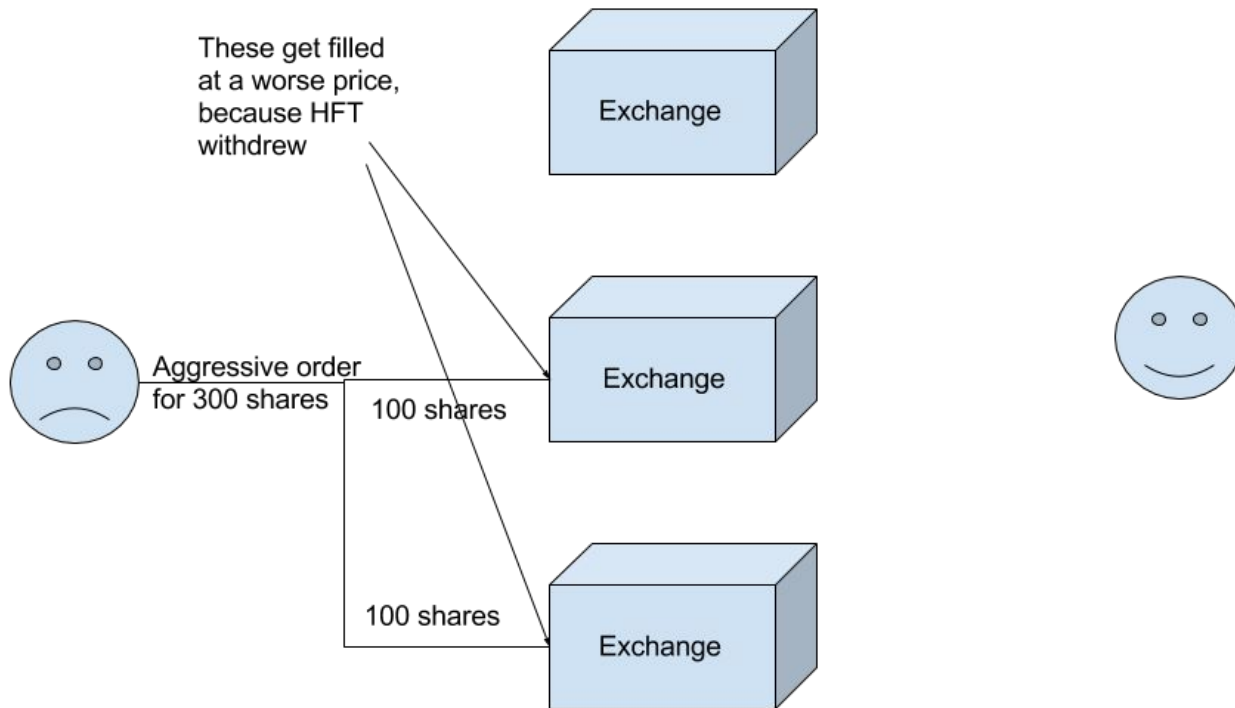
# Example #3: phantom liquidity (?)



# Example #3: phantom liquidity (?)



# Example #3: phantom liquidity (?)



# Summary

- Other players react in real time to your placed orders. You're not a passive observer -- yet in backtesting you act as if you were.
- You don't really know what actual liquidity exists just by looking at the order book. This skews your backtesting estimates of how much liquidity is available at what prices, which in turn skews your P&L calculations.
- Small mistakes in P&L calculations can accumulate quite quickly when you trade intraday. Example: once I forgot to account for transaction costs in my backtesting. This is a difference of 1.5 cent per trade. Yet, this was enough to make my strategy go from Sharpe ratio 3.5 to -8, once I fixed the bug.

# Part II

---

A proposal for micro-structure dynamics simulation

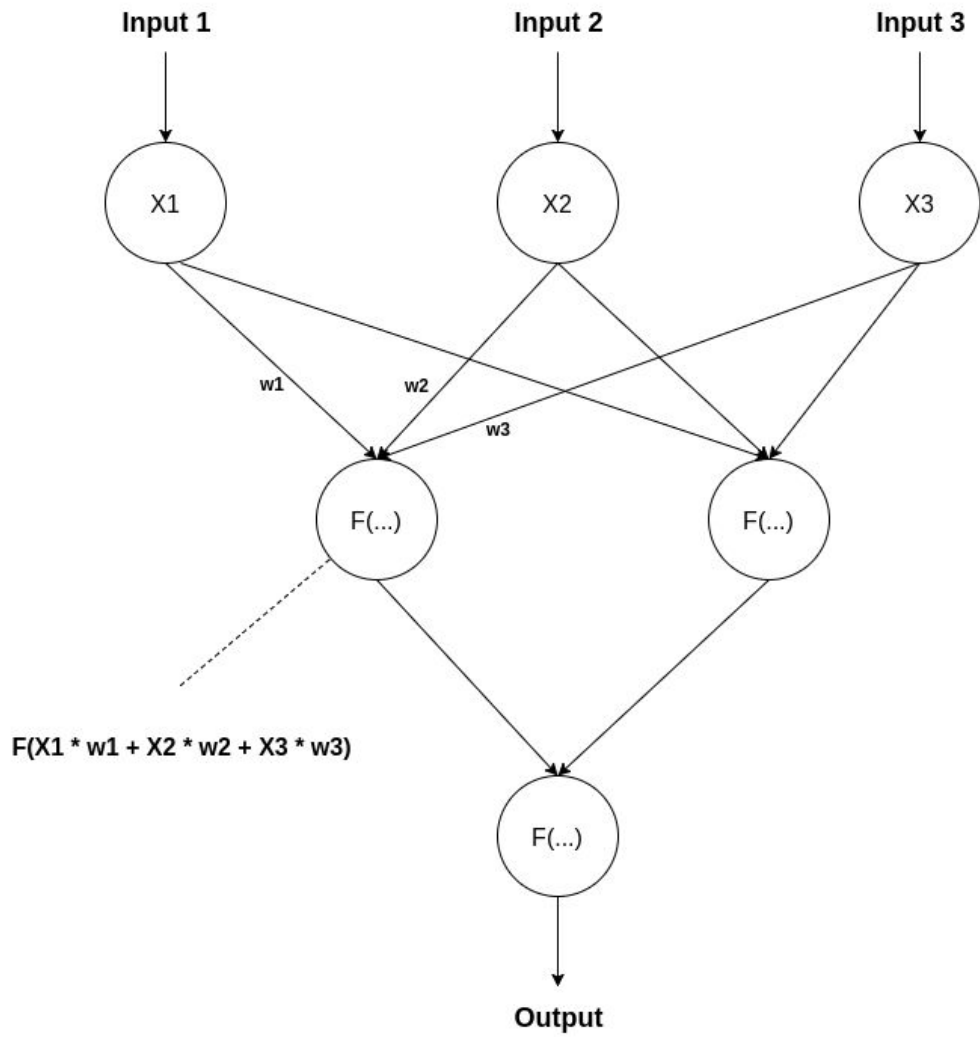
# An untested proposal

- If backtesting simulation can't rely on static historical order book data, why not model the micro-structure behaviour and use a “dynamic”, simulated order book?
- Generative deep learning models seem like good candidates for this kind of thing.
- This relates to a deep learning/reinforcement learning concept called “Imagination”

# Deep learning & generative models

- Class of unsupervised neural networks that can be used to generate new data that is similar to the data that was trained upon.
- For example, LSTMs (Long-Short Term Memory) are neural networks that you can train on a time series, and then you use them to “predict” the rest of the time series. It is generative because you can predict more than 1 step ahead, thereby generating an entire predictive time series if desired.





# The data & the problem to solve

- We have the historical order book data.
- We have orders that we place on the order book.
- We want to predict how the order book will “adjust” to your order, up to the point where it will be filled. In other words: we will want to generate a variant on the historical order book data that accounts for the fact that we placed an order on the book which wasn't there in the original data.

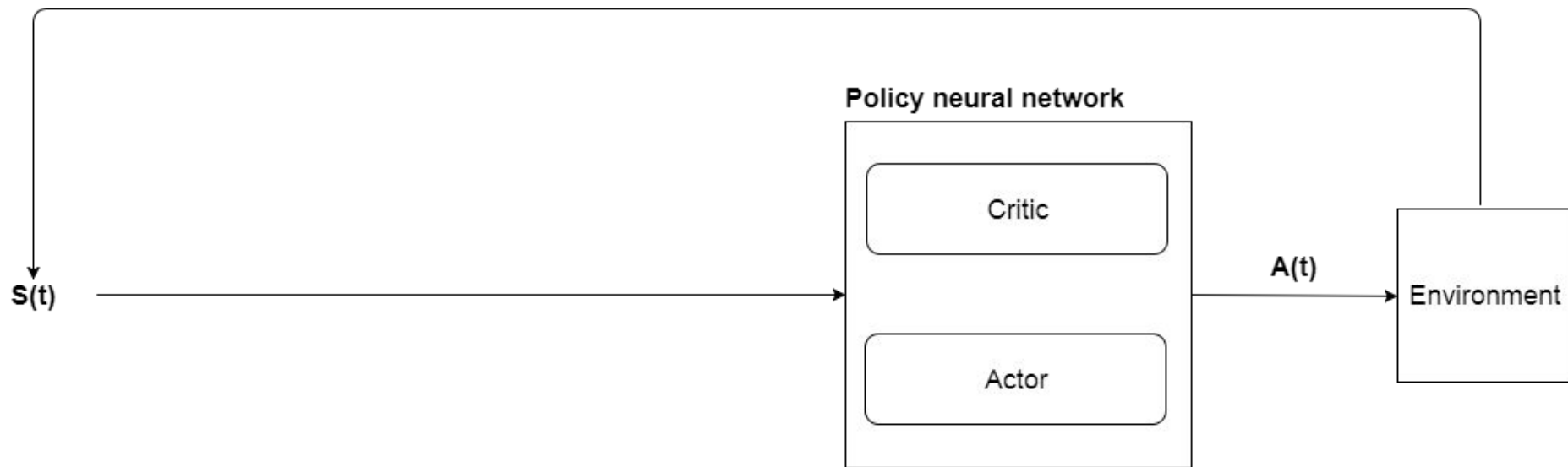
# “Imagination” and the problem of imperfect models

In paper: **“Imagination-Augmented Agents for Deep Reinforcement Learning”** (*July 2017*), the authors improve the performance of a deep reinforcement learning agent by adding a model that learns to generate more data from the input data it is using.

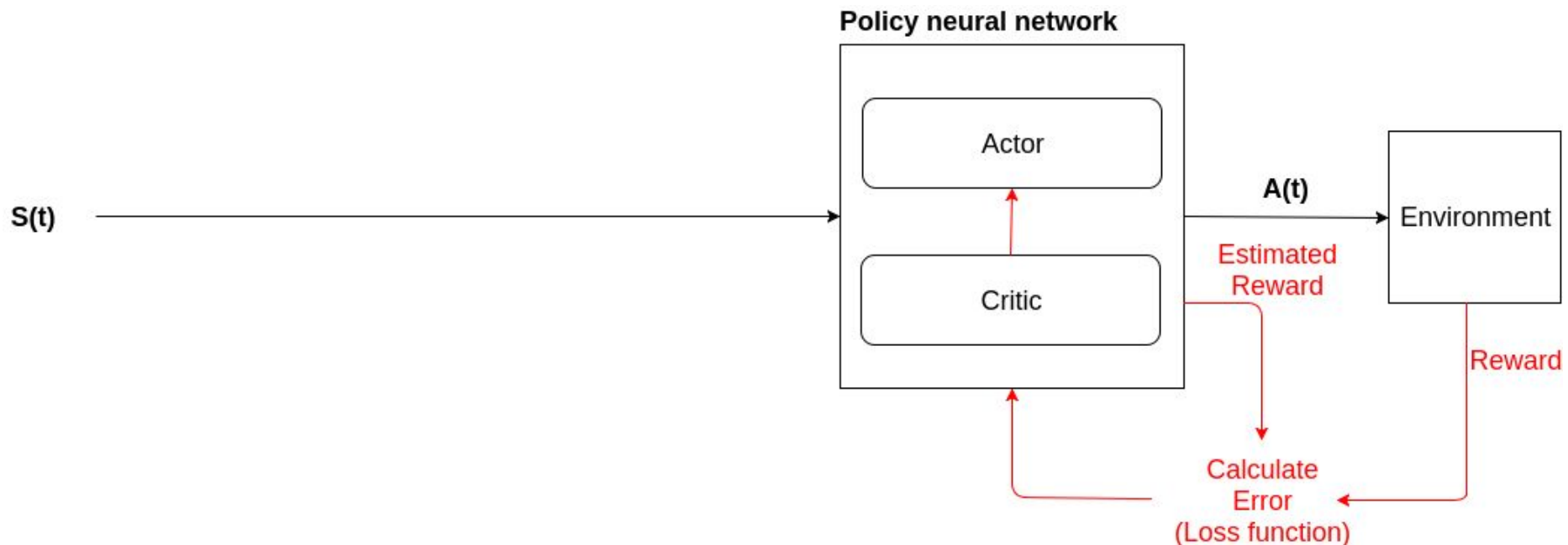
**The imperfection of the generated data is mitigated by another module that learns to interpret the imagined data, resulting in an overall performance improvement.**

- The paper: <https://arxiv.org/pdf/1707.06203.pdf>
- YouTube explanation: <https://www.youtube.com/watch?v=agXIYMCICcc>

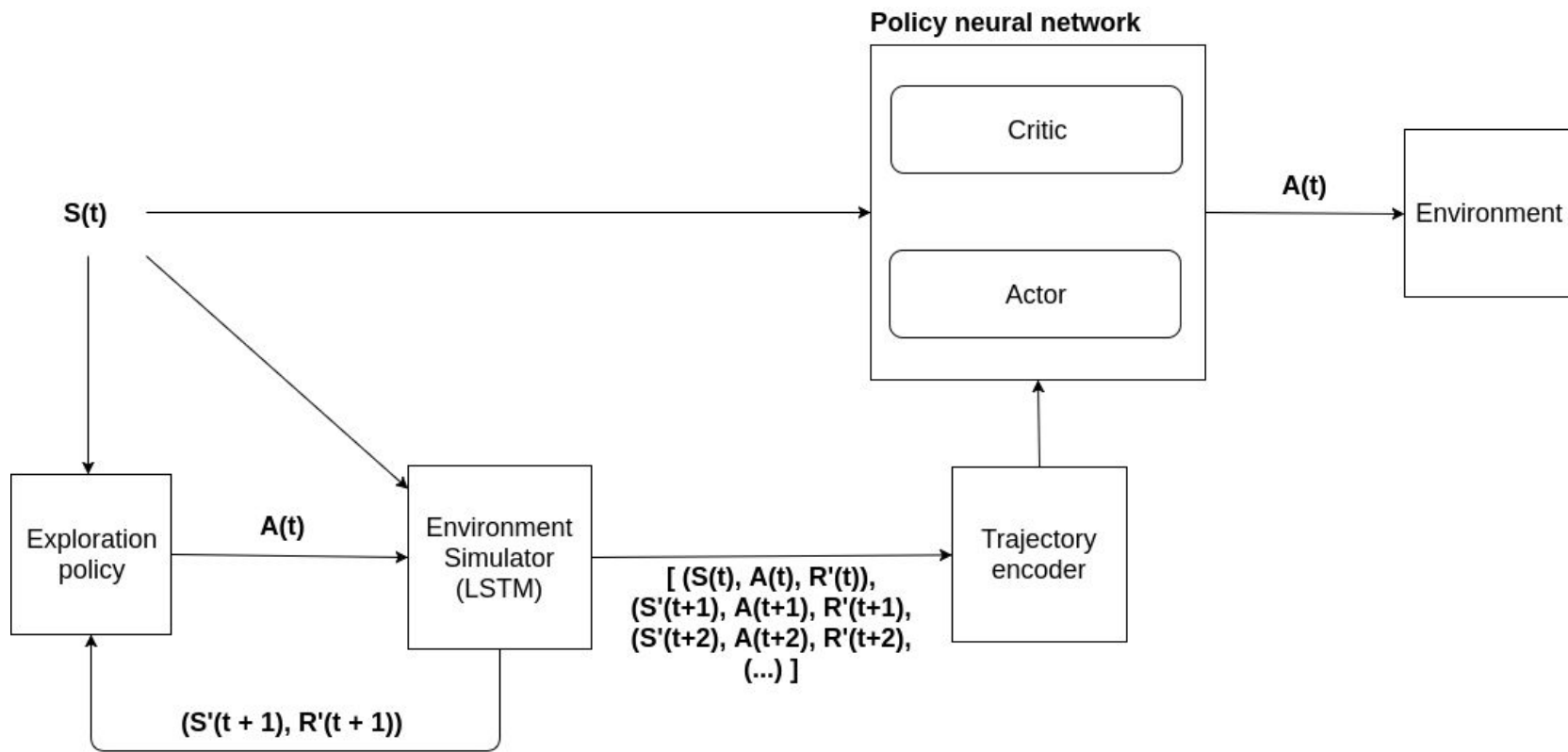
# Basic RL agents



# Basic RL agents (training phase)



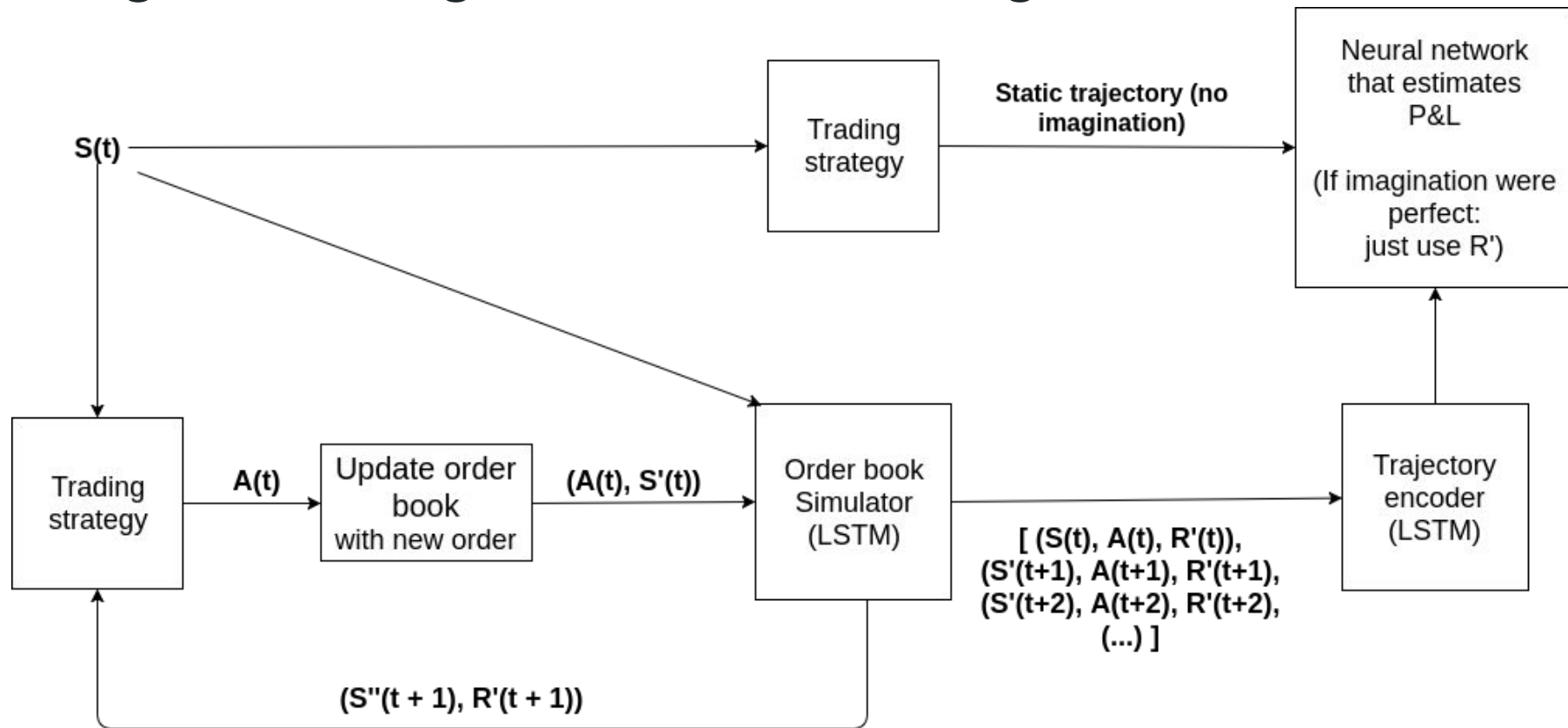
# Imagination-augmented RL agents



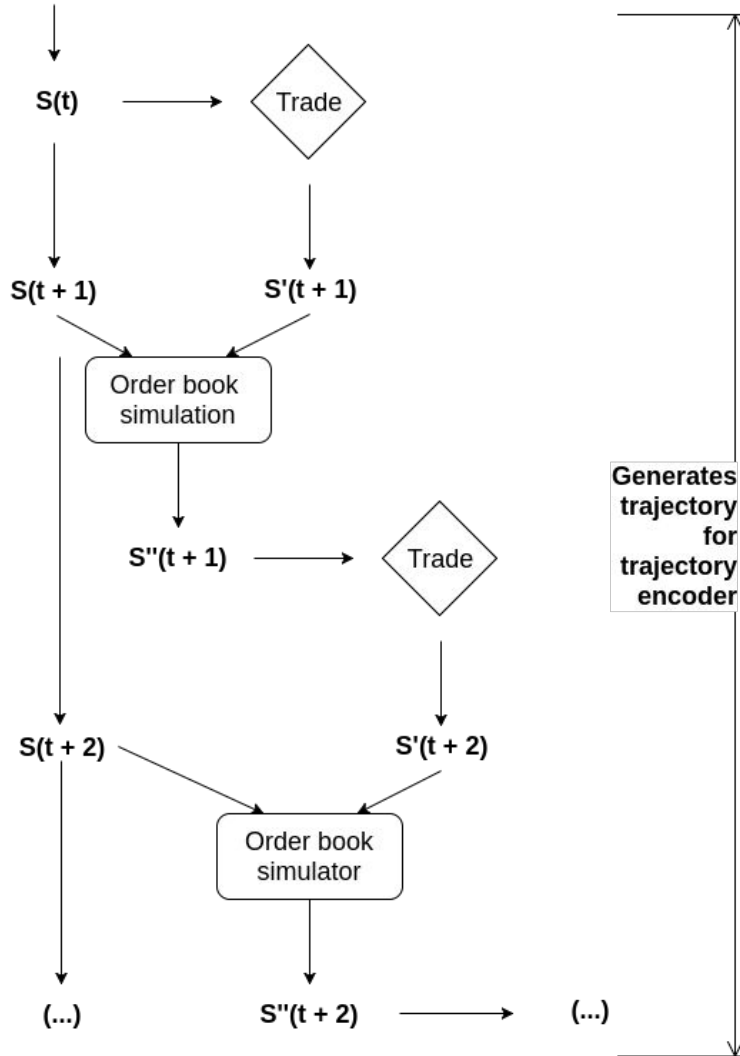
# Now back to the task of order book simulation

- We drop the RL-related part, keep the imagination module, replacing the agent policy with a neural-network based P&L estimator.
- Analogy: in reinforcement learning, the goal of the optimization process is to increase the fitness of the agent. In backtesting, the goal of the optimization process is to increase the accuracy of evaluation of the agent's fitness.
- So it's the same idea, but the optimized module is the P&L estimator, instead of the agent policy.

# Imagination-augmented backtesting module







# What could the model learn? (plausibility)

- For the quote matching issue, it could learn on what types of assets (based on levels of liquidity, spread, etc.) the likelihood of being “shaved” is high. If no spread, no “shaving” can occur. Typically this happens on medium to low liquidity assets with a spread of at least a few cents.
- For the hidden volume issue, maybe some displayed volume levels are more likely than others to have hidden volume attached (e.g. some exchanges required to display at least 200 shares from their total offer -- some perhaps offers at 200 are more likely to contain hidden volume)? It could remember that some broker IDs have a tendency to use, or not to use, hidden volume?

# References

Slides and references will be available on my blog:

**<http://www.simonouellette.com>**