Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Limit Order Books and RL:
1. Market orders and limit orders

Igor Halperin

NYU Tandon School of Engineering, 2018

Quote-driven and order-driven markets

- **Quote-driven markets:**
  - A **centralized market**: a market-maker (dealer, or specialist) aggregates all buy and sell orders, and provides liquidity by setting bid and ask quotes
  - The ‘fair price’ of classical financial models is the mid price
  - Example: NYSE specialist system

Standard assumption in Finance

- The Black-Scholes model:
  - A single number (the average cost of hedge) as an option price
  - Infinite liquidity: can buy/sell any number of shares with no impact on price
  - Transaction costs are not taken into account
- The BS model assumptions are especially problematic for:
  - Large trades over short periods
  - High Frequency trading
- Need to pay attention to details of Market Microstructure:
  - Mechanics of execution of buy and sell orders

How NYSE operates today

- **Electronic order-driven markets:**
- On the picture: NYSE NJ data center (http://www.datacenterknowledge.com/closer-look-nyse-euronexts-nj-data-center)

Order-driven markets

- **Order-driven markets:**
  - Electronic platforms that aggregate all available orders in a Limit Order Book (LOB)
  - This is a **de-centralized market**
  - Examples: NYSE, NASDAQ, LSE
  - Order-driven markets account for about 80% of traded stock volumes
  - The same stock is now traded on several ‘venues’.
- In LOB markets:
  - Buyers and sellers have access to real-time microscopic level of description of market dynamics
  - Double-action mechanism
  - In a LOB, complex dynamics arise as a result of local interactions between many heterogeneous agents
  - This makes LOBs an example of complex systems: amenable to modeling using statistical models or ML approaches
  - *See e.g. R. Cont, “Statistical modeling of high frequency financial data: facts, models and challenges” (2011).*

Order-driven markets

- **Order-driven markets:**
  - The average number of orders in 10 sec, and the number of price changes. (Source: R. Cont 2011)
  - *High-frequency data becomes available in real time, shows how the market digests the balance of supply and demand during the price generation.*

	Av. # of orders in 10 sec	Price changes in 1 day
Citigroup	4469	12499
General Electric	2356	7862
General Motors	1275	9016

Hierarchy of time scales

REGIME	Time scale	Issues
Ultra-high frequency (UHF)	< 0.1 sec	Microstructure, latency
High Frequency	1-100 sec	Trade execution
“Daily”	>100 sec	Trading strategies, option hedging

Control question

Select all correct answers

1. In a centralized market, a market-maker (dealer, or specialist) aggregates all buy and sell orders.
2. In an order-driven market, all market-makers follow the same orders from the exchange.
3. The limit order book (LOB) is a database that stores all market transactions. Access to this database is limited but ordered.
4. The limit order book (LOB) aggregates all current outstanding limit orders.

Correct answers: 1, 4

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Limit Order Books and RL: 2. Trades, quotes and order flow

Igor Halperin

NYU Tandon School of Engineering, 2018

Two types of data

- **Transaction data:**
  - Traded quantities of assets with actual traded prices
- **Quotes:**
  - Proposals to buy or sell a certain quantity of an asset at a given price.
  - There is no single price anymore, rather a number of price/quantity pairs

Transaction data

- **Transaction data:**
  - Each trade has a time stamp, a price and a quantity.
  - This is quite different from daily financial time series:
    - **Irregular** time intervals: durations between trades are endogenous and are driven by the behavior of the price and the LOB
    - Price changes are **discrete**: (multiples of a ‘tick’): The price changes are not independent, but autocorrelation quickly decay with the time lag.
    - Trading volumes are strongly **autocorrelated**
    - **Continuous-time** formulation might be preferred in this setting to a discrete-time specification of conventional financial time series.
    - Strong **seasonality** of intraday data: activity of trading is highest at market open and close, lowest around lunch time.

Types of orders

- **Market orders:**
  - Market bid order: buy an amount X of a stock now, for the current market price
  - Market ask order: sell an amount X of stock now, for the current market price
  - Executed immediately via exchange's specialists

- **Limit orders:**
  - Limit bid order: buy an amount X of a stock for price P, eventually
  - Limit ask orders: sell an amount X of a stock for price P, eventually
  - Execution for a prescribed price P can be partial only, or the order may not be executed at all
  - Limit orders can be cancelled
  - HFT play!

Quotes and order flow

- Two types of orders**: market orders and limit orders**

  - **Limit orders** are posted to an electronic trading system
  - State of outstanding limit orders at each price level are summarized by the total number of orders in each bucket - this is known as a **Limit Order Book** (**LOB**)
  - The lowest price on the sell side is called the **ask price**, and the highest price on the buy side is called the **bid price.**
  - A limit order sits in the LOB until it is executed as it matches a marker order on the opposite side of the LOB, or cancelled
  - ***Quote data:*** (Level-I order book): bid and ask prices, and the quantities quoted at these prices. ***Level-II order book data*****:** prices and quantities of first 5 non-empty levels on both sides of the book.

  - **Market orders:**
    - An order to buy quantity X at the best available price in the LOB
    - Upon execution, a quantity available at the bid-ask price is decreased by X.

Using trade and quotes data

- Trades and quotes data are available through a number of providers. The TAQ database maintained by the NYSE is the most well known.
  - Traders and quotes data can be used to reconstruct a sequence of market and limit orders and infer whether a given transaction was a buyer or seller initiated:
  - This is done by comparing the trade price with the bid and ask for the most recent past quote
  - See Lee and Ready, "*Inferring trade direction from intraday data*, Journal of Finance, 46 (1991), pp. 733– 746, R.Cont, A. Kukanov, and S. Stoikov, *The price impact of order book events*, working paper, Columbia University, 2010.

Control question

Select all correct answers

1. Market orders can be cancelled at any point in time.
2. The lowest price on the buy side is called the **ask price**, and the highest price on the sell side is called the **bid price.**
3. The lowest price on the sell side is called the **ask price**, and the highest price on the buy side is called the **bid price.**
4. A limit order sits in the LOB until the dealer changes its status to a marker order.
5. A limit order sits in the LOB until it is executed as it matches a marker order on the opposite side of the LOB, or cancelled

Correct answers: 3, 5

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Limit Order Books and RL: 3. Limit Order Book

Igor Halperin

NYU Tandon School of Engineering, 2018

Limit Order Book Mechanics

- **LOB:**
  - Aggregates all limit orders at both sides at each price level
  - Price priority queue: FIFO (First In-First Out)
  - Matching engine: matches incoming orders to buy or sell against limit orders resting at the opposite side of the LOB
  - A trader should "cross the spread" and pay an ask price (if buying) or get a bid price (if selling)

Limit orders

- **A limit order sits in the order book until it is:**
  - Either executed by getting a matching market order
  - Or it is cancelled

- **When limit order waits for a match:**
  - Transaction cost is known
  - Execution time is uncertain

- Execution timing:
  - Executed very quickly if it is near the bid-ask spread
  - May take a very long time if:
    - the market price moves away from the requested price
    - the requested price is too far from the bid-ask spread
  - Can be cancelled at any time

Market orders

- **A market order** is an order to buy/sell a certain quantity of the asset at the **best available price** in the book**:**
  - Agents put market orders to buy/sell
  - The first share(s) will be traded at the ask/bid price
  - The remaining shares will be traded some ticks up/down
  - This will modify the bid-ask price

- A bid or ask queue may be **depleted**
  - Market orders
  - Cancellations
  - In these scenario the price is **updated** to the next level of the order book

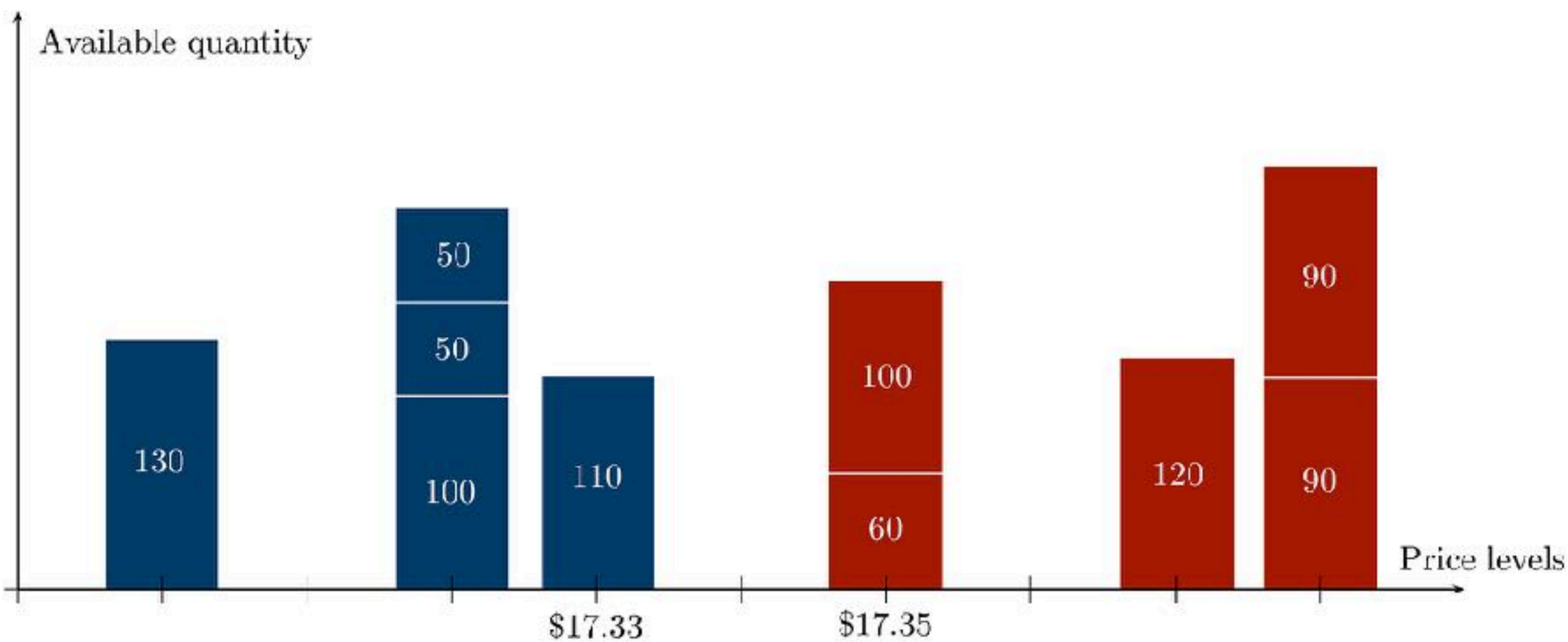
- A market order typically consumes the cheapest limit orders:
  - Immediately executed in full, if the book has sufficient liquidity
    - Uncertainty in the resulting price per share (depends on the order size, and actions of other players)
    - the requested price is too far from the bid-ask spread
  - Rules for market orders:
    - Price priority: best available price is executed first
    - Time priority: first in, first out (FIFO)

Cancellations

- Agents can cancel their orders: cancellation of X orders in a given queue reduces the queue size by X
- Either the bid or ask queue can be depleted by market order and cancellations
- In such cases, the price moves up or down the next level in the order book.
- In intraday trading, number of cancelled orders can be very high, sometimes up to 95%.
- Therefore, study of LOB events may provide a more complete picture of dynamics of supply and demand in order-driven markets, in comparison to studies based only on transaction data and trading volume (R. Cont, A. Kukanov and S. Stoikov, “Price impact of order book events”, 2015)

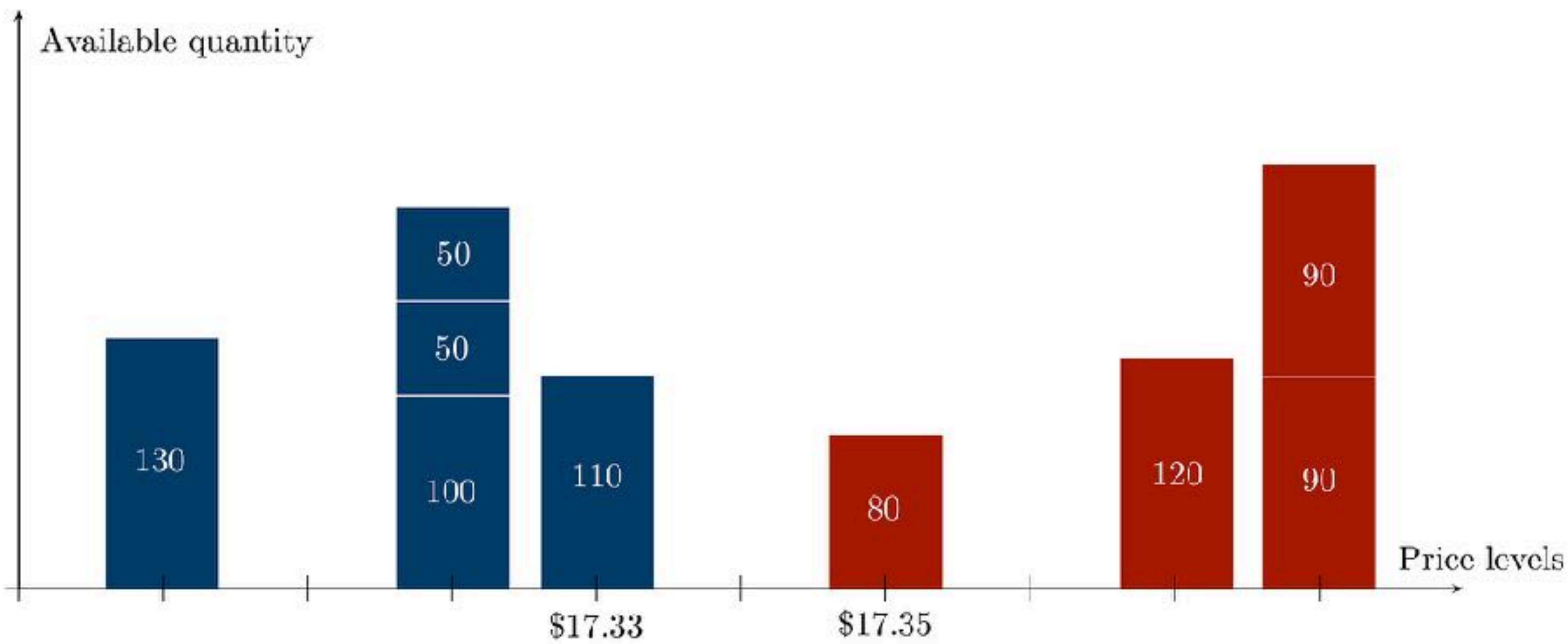
Limit Order Book Example

- Example from https://www.maximemorariu.com/research-topics
  - The current best bid on a stock in $17.33
  - We want to buy 200 stocks



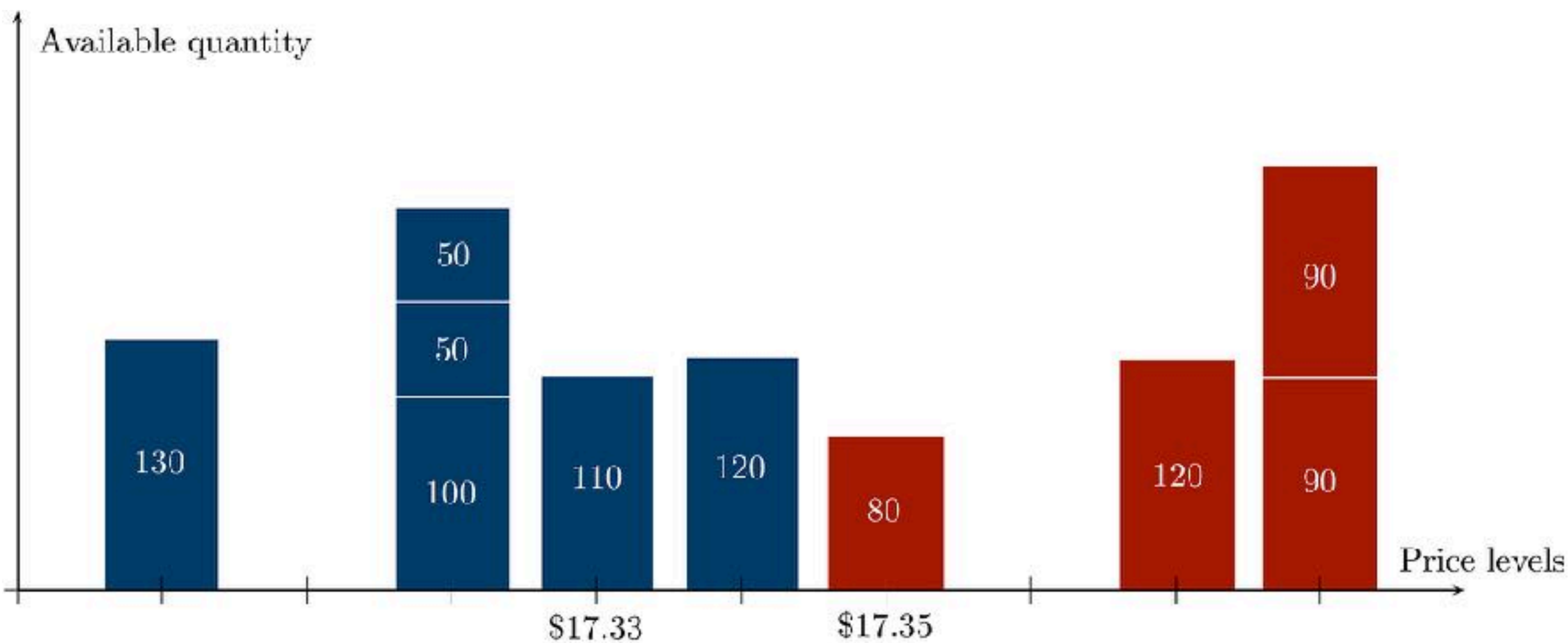
Limit Order Book Example

- The current best on a stock in $17.33
- We want to buy 200 stocks
- Buy 80 shares at $17.35 - immediate execution



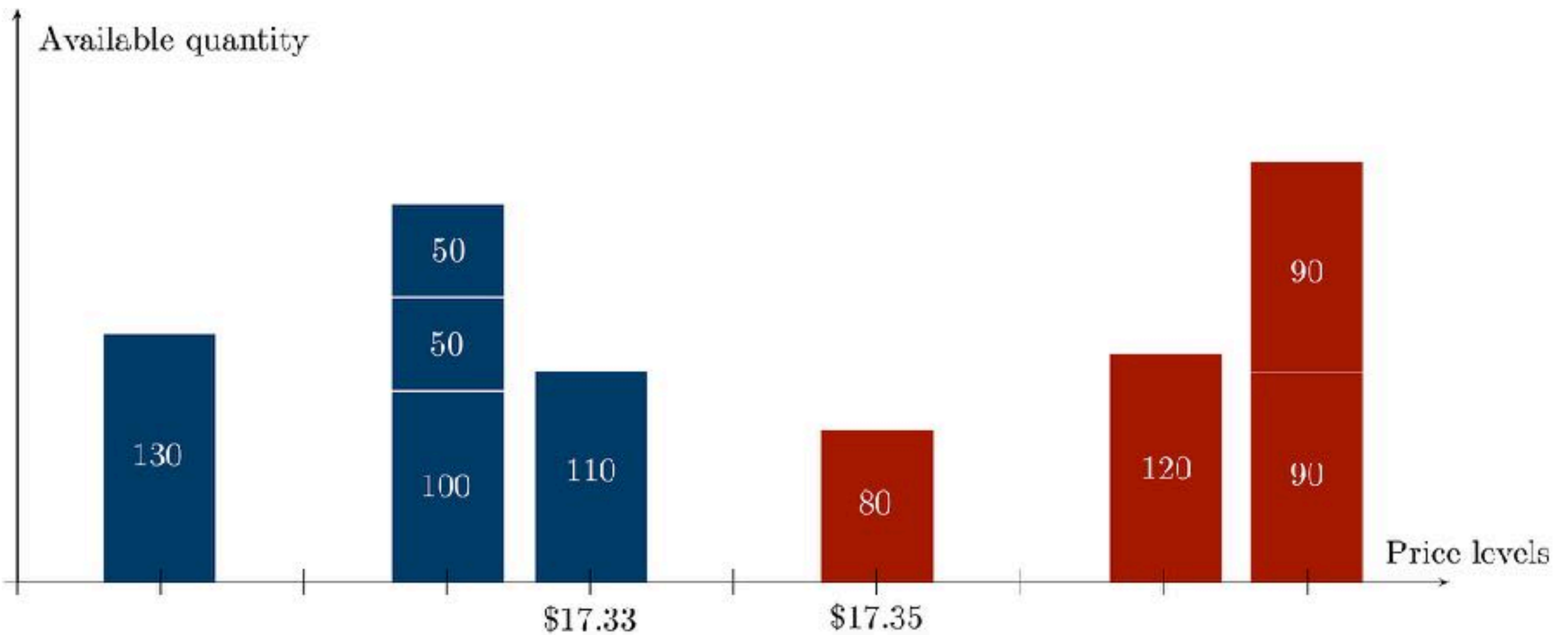
Limit Order Book Example

- The current best on a stock in $17.33
- We want to buy 200 stocks
- Buy order of size 120 at $ 17.34 - becomes a limit order



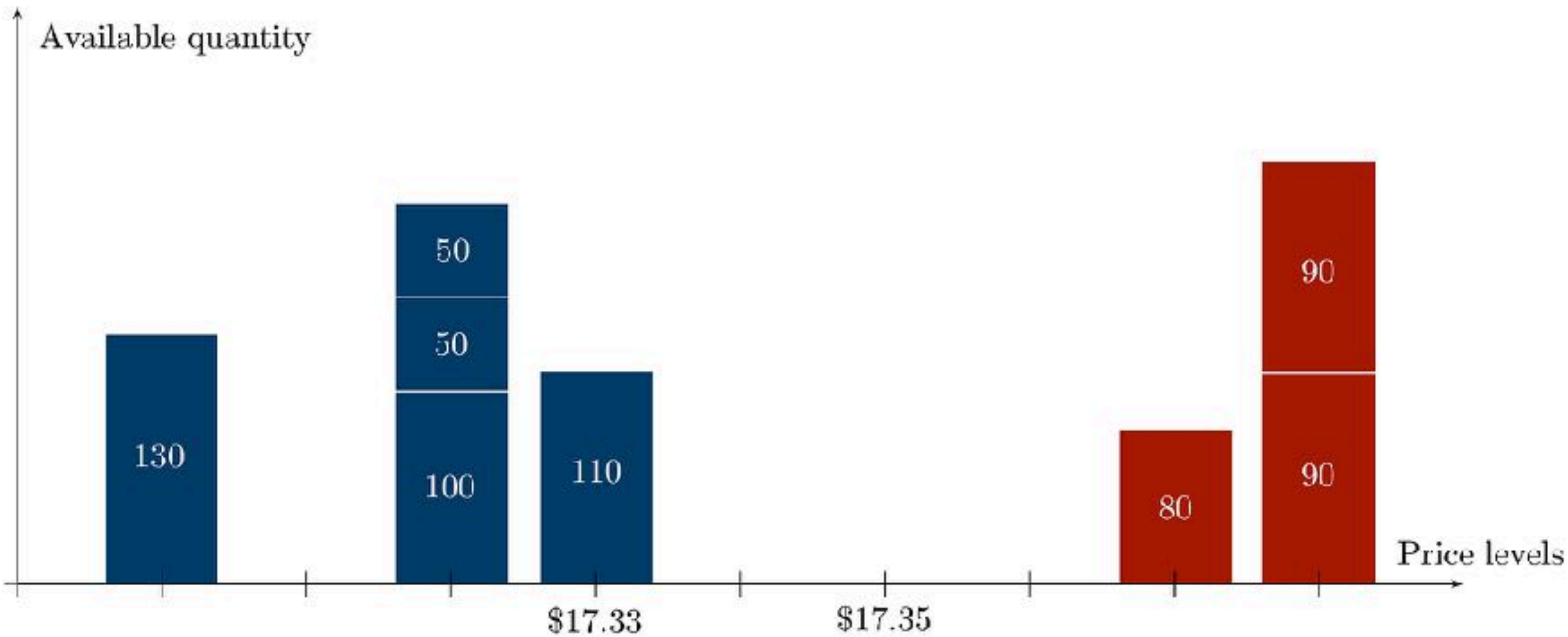
Limit Order Book Example

- The current best on a stock in $17.33
- We want to buy 200 stocks
- Cancellation of buy order of size 120 at $17.34



Limit Order Book Example

- The current best on a stock in $17.33
- We want to buy 200 stocks
- Market order to buy order of size 120: executed 80 stocks for $17.35 and 40 stocks for $17.37 - immediate execution
- The ask queue is now depleted



Control question

Select all correct answers

1. “Crossing the spread” means that your counter-party finally agrees to your price.
2. “Crossing the spread” means neglecting the bid-ask spreads in all of your forecasting and trading models.
3. “Crossing the spread” means an actual transaction in an LOB when you pay an ask price (if buying) or get a bid price (if selling).
4. When a bid or ask queue becomes depleted, the exchange temporarily shuts down. Fortunately, such events are very rare.
5. When a bid or ask queue becomes depleted, the price is updated to the next level of the order book.

Correct answers: 3, 5.

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Limit Order Books and RL: 4. Modeling Limit Order Book

Igor Halperin

NYU Tandon School of Engineering, 2018

Limit Order Book

- **The Limit Order Book (LOB):**
  - Summarizes a multi-level state of supply and demand side
  - The state of a LOB is informative of short-term price movements
  - Therefore, it makes sense to make short-term market forecasts <u>conditional on the state of the LOB.</u>

Why understanding LOB is important

- **Understanding LOB is important for:**
  - Optimal trade execution
  - Market impact minimization
  - Designing better electronic trading algorithms
  - Risk management (due to complex interactions at different time scales)
  - Regulators for assessing market stability: propagation across time scales: the Flash Crash of 2010.

Order placement problem in LOB

- Difference with portfolio allocation decision-making:
  - For portfolio allocation: decisions at the scale of weeks, translates into trades extending over minutes to days
  - This translates into streams of orders placed at high frequency (may be thousands in a minute):
    - Order type (limit order or market order)
    - Order size
    - Destination (when multiple venues are available)

How to model LOB

- **How to model**
  - High dimensionality, discrete-valued data!
  - Complexity of dynamics: many moving parts!
  - Computationally demanding: need analytically tractable models?
  - ML approaches?

Approaches to modeling the LOB

- **Various approaches to modeling LOB:**
  - Economics-based equilibrium approaches: a sequential game of a single agent
  - Statistical and ML approaches: statistical patters of order flows, without trying to explain the ‘mechanics of a LOB’ (replaced by exogenous random processes)
  - Physics-based approaches: order flows are treated as random, driven by stochastic dynamics of complex interactions within a LOB

Applications

- **Applications:**
  - Market making
  - Optimal trade execution
  - Risk management or intraday and ‘daily’ strategies
  - Regulation

Types of LOB data

- Three levels of LOB data, depending on the aggregation level of limit orders:
  - Level-1 data: minimal information, shows only the the best bid and ask prices and the total liquidity at this level
  - Level-2 data: minimally aggregated LOB, assigns total liquidity to each price level in the market
  - Level-3 data: most refined data, provides information about about each individual limit order.

- Level-1 datasets are the most accessible (see e.g. tickdata.com)

Control question

Select all correct answers

1. Modeling of a LOB is hard because it has many moving parts.
2. Modeling of a LOB is hard because it has to deal with high-dimensional irregular discrete-time and discrete-valued data.
3. Level-I data is the most refined data, available only to regulators.
4. Level-1 data shows the the best bid and ask prices and the total liquidity at this level.
5. Level-2 data aggregates the top levels of a LOB, and assigns total liquidity to each price level in the market.

Correct answers: 2, 4, 5

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Limit Order Books and RL: 5. Statistical models for LOBs

Igor Halperin

NYU Tandon School of Engineering, 2018

LOB

- **Models of LOBs**
  - Economics based model of a single agent repetitive games
  - Statistical models
  - Physics-based models
  - Machine Learning approaches

- Needed for:
  - Estimation of intraday risk (loss distributions, volatility, VaR)
  - Short term (about a second) prediction of order flow and price moves for trading strategies
  - Optimal trade execution

  - References:
  - R. Cont, “High Frequency Dynamics of Limit Order Markets: Statistical Modeling and Asymptotic Analysis” (2015).
  - Physics-based approaches: e.g. F. Lillo, S. Mike, J. D. Farmer. A theory for long-memory in supply and demand. Physical Review E, 71, 06-2005, J.P. Bouchaud, J.D. Farmer, and F.Lillo, “How markets slowly digest changes in supply and demand”, https://arxiv.org/abs/0809.0822 (2008).
  -

Statistical models of LOB

- **Statistical models of LOB incorporate the following information**
  - The current state of the LOB
  - Statistics of the order flow: arrival rates of market and limit orders, cancellations)
  - Physics-based models
  - Machine Learning approaches

Statistical models of LOB

- Statistical models model the following components of a LOB
  - Arrival of different order types (market, limit, cancel, buy/sell) through arrival intensities that may depend on the state of the order book
  - Arrival intensity is the probability of event per unit time
  - Execution of market orders through deterministic execution priority rules
  - The price dynamics is obtained from the dynamics of the LOB

- Possible approaches:
  - Model limit orders at all price levels simultaneously (a ‘microscopic’ approach)
  - Focus on the description of the best prices, with the rest of the LOB serving as a source of ‘features’

LOB as a queuing system

- **LOB:**
  - A LOB can be represented as a multi-class queuing system
  - Queuing systems are modeled using arrival rates for orders on both sides
  - If market orders and limit orders at each level are viewed as independent Poisson events, one obtain analytically tractable models

Order flow clustering

- In reality, independence of arrivals (leading to Poisson processes for arrivals) does not hold: orders are clustered in time: (Source: R. Cont 2011)
- Can be described using *self-exciting processes* (e.g. a Hawkes process)

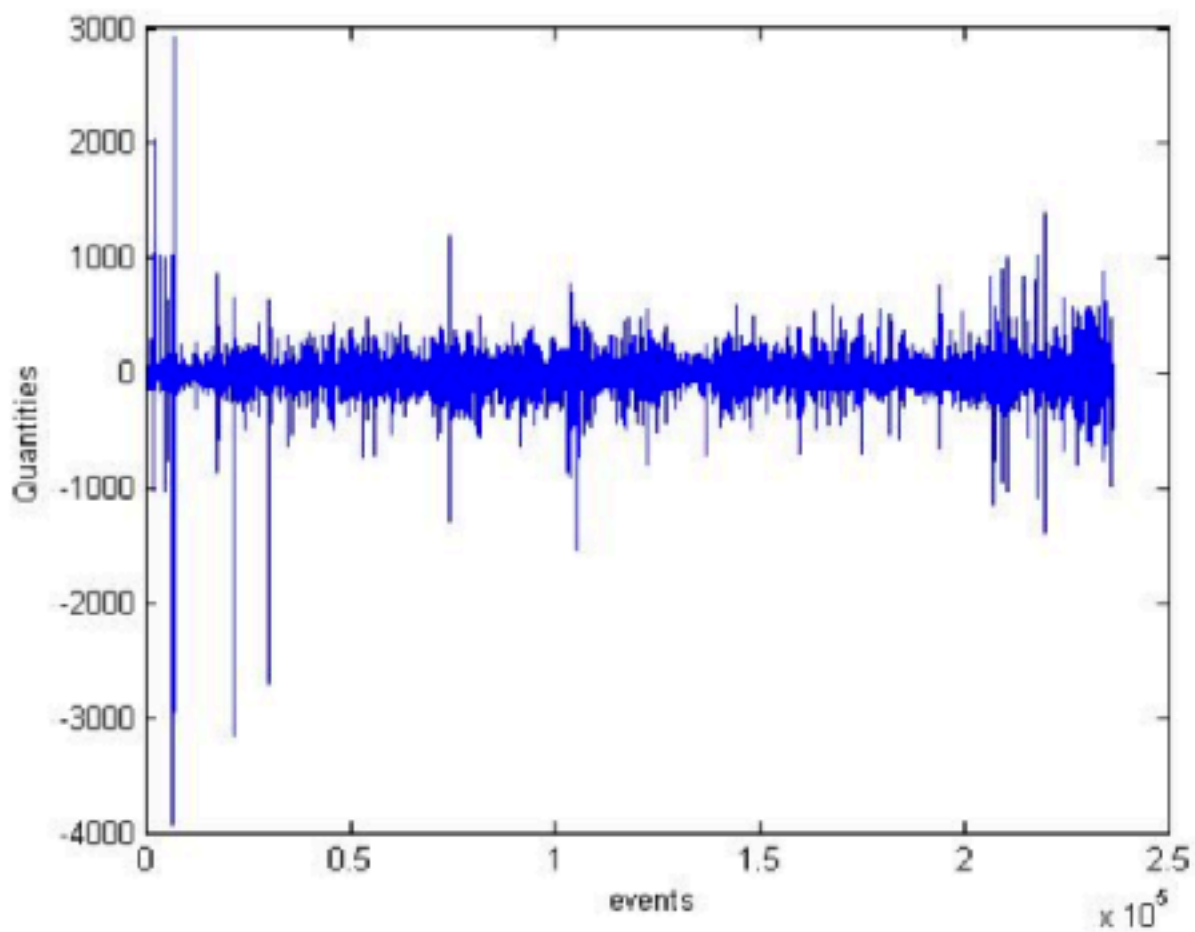
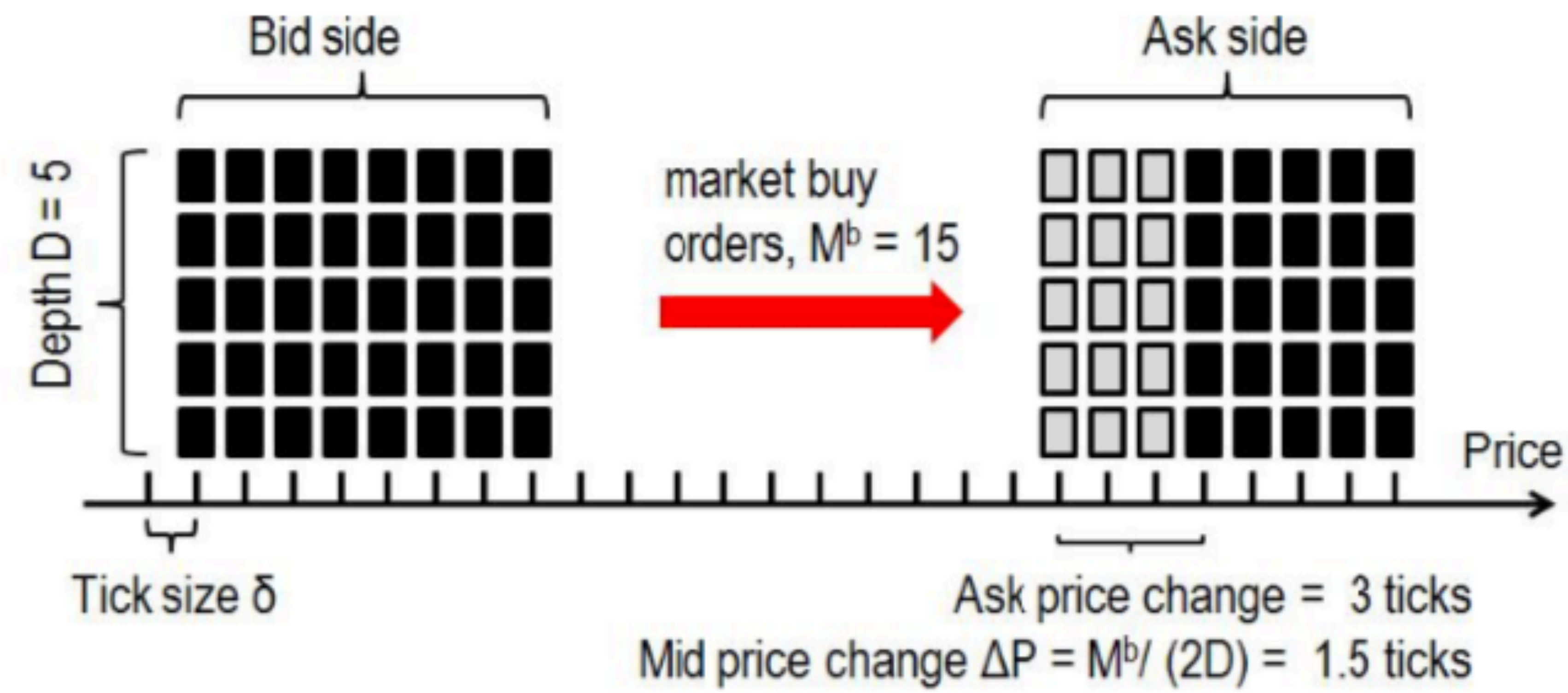


Fig. 5. Changes in the size of the ask queue: each data point represents one order or cancellation. Positive values represent sizes of limit orders, negative values represent market orders or cancellation. Citigroup, June 26, 2008.

Modeling price impact

- Price impact modeling: Bertsimas and Lo, Almgren and Chriss, etc.
- Price impact: linear/nonlinear/square root, permanent or transient, etc.
- Classical impact models focus on the relation between the direction of the price and the size of trades



- Order flow imbalance = (L - M - C)/D
- Stock price changes driven by order flow imbalances: (Cont et. al. 2010)

Control question

Select all correct answers

1. Arrival intensity is the probability of a trade per unit time.
2. Arrival intensity is the probability of a LOB event per unit time.
3. Machine Learning approaches analyze limit orders at all price levels simultaneously.
4. A self-exciting process is a process that ensures monotonicity of a wealth growth as a function of intensity of trading.
5. The order flow imbalance measure that partially drive market changes is equal to Limit orders minus Market orders minus Cancellation, and divided by the depth of the book.

Correct answers: 2, 5.

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Limit Order Books and RL: 6. Using ML and RL for modeling LOBs

Igor Halperin

NYU Tandon School of Engineering, 2018

Are ML and RL relevant for modeling of the LOB?

- **Of course:**
  - All statistical methods can be viewed as parametric (‘model-based’) ML
  - Linear Regression is ML, too!

- **Real questions:**
  - What ML method to choose?
  - Is it better to focus on predictions (ML/Supervised Learning), or directly on action optimization (RL)?
  - What prior knowledge to embed in your model:
    - Model architectures/parametric function families (ML)
    - Regularization (ML)
    - Bayesian priors (Bayesian approach)
    - Physics-based arguments (invariances, analytical properties, etc.)

ML methods: shallow architectures

- **Shallow architectures:**
  - Support Vector Machines (SVM) (Kercheval, Alec N., and Zhang, Yuan. ”Modeling high-frequency limit order book dynamics with support vector machines.” (2014).
  - Random Forest, Boosted Trees, etc. (see e.g. this report for some analysis: http://jcyhong.github.io/assets/machine-learning-price-movements.pdf)

ML methods: deep architectures

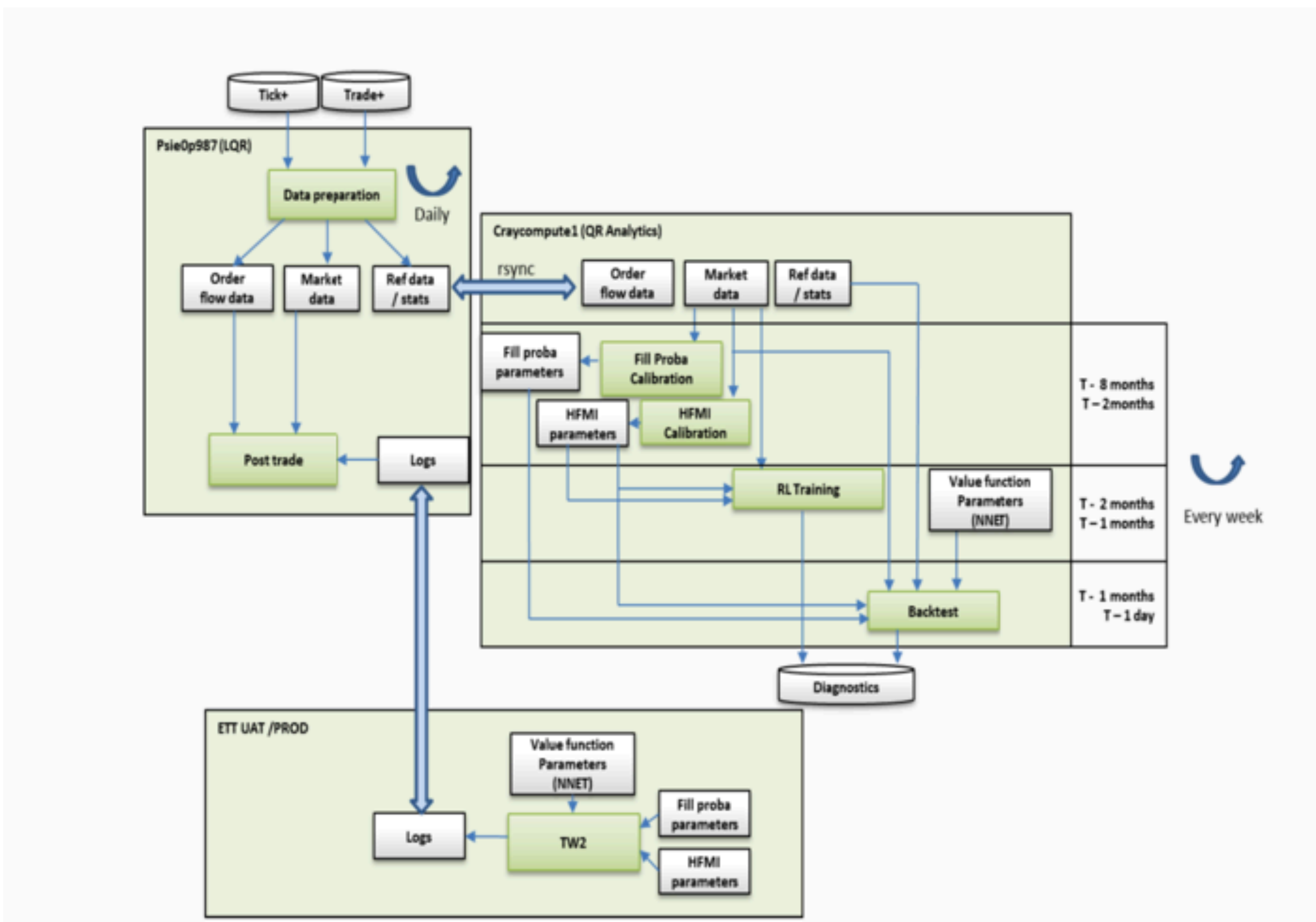
- **Deep Learning architectures:**
  - “Spatial neural networks” (J. Sirignano, Deep Learning for Limit Order Books”, 2016, https://arxiv.org/abs/1601.01987, http://jasirign.github.io/pdf/Extended_Abstract.pdf)
  - Recurrent Neural Networks (M. Dixon, “Sequence Classification of the Limit Order Book using Recurrent Neural Networks”, https://arxiv.org/pdf/1707.05642.pdf)
  - Convolutional Neural Networks (A. Tsantekidis et. al. “Forecasting Stock Prices from the Limit Order Book using Convolutional Neural Networks”, http://poseidon.csd.auth.gr/papers/PUBLISHED/CONFERENCE/pdf/2017/2017_CBI_CNNLOB.pdf)
  - …

Reinforcement Learning for LOB

- **Reinforcement Learning:**
  - Focuses on the main task (order placement optimization)
  - Reinforcement Learning for optimized trade execution (Y. Nevmyvaka, Y. Feng, M. Kearns, “Reinforcement Learning for optimized trade execution”, ICML 2006 23th International Conference on Machine Learning”, “Machine Learning for Market Microstructure and high frequency trading”, 2013)
    - First large-scale implementation of RL for optimal trading.
    - Features constructed from the LOB characteristics (bid-ask volume imbalance, signed transaction volume = #shares bought in the last 15 sec minus #shares sold, etc.)
    - Performance is measured by implementation shortfall (how much the average price paid is larger than the bid-ask spread at the beginning of the trading period
    - Online learning similar to Q-learning

Deep Reinforcement Learning for LOB

- **Deep Reinforcement Learning:**
  - JPM implementation of DRL for optimal execution
  - See e.g. description here: https://medium.com/@ranko.mosic/reinforcement-learning-based-trading-application-at-jp-morgan-chase-f829b8ec54f2
  - Uses temporal difference methods (Q-learning)
  - Not too many details, proprietor algorithm.



Control question

Select all correct answers

1. The ML approach to modeling the LOB focuses on predictions, and operates with conventional ML methods such as out-of-sample tests etc.
2. RL-based approaches to modeling the LOB focus on reinforcing compliance rules in electronic trading.
3. RL-based approaches to modeling the LOB focus on optimization of execution strategy.

Correct answers: 1, 3.

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 1: Other topics RL: Market making, P2P lending, cryptocurrencies, perception-action cycles.

Igor Halperin

NYU Tandon School of Engineering, 2018

Market maker problem

- **Problems of decision makers in the market:**
  - Investor
  - Dealer
  - Market maker

- **Market maker optimization:**
  - Dynamic programming approaches:
  - Reinforcement Learning
  - TradingGym based implementation of Q-learning for Market Making: M.Dixon, https://github.com/mfrdixon/dq-MM, https://www.researchgate.net/publication/310951905_High_Frequency_Market_Making_with_Machine_Learning

Credit portfolios in P2P lending

- **Peer-to-peer (P2P lending):**
  - P2P platforms for online lending (Lending Club, Prosper, OnDeck)
  - $26B total notional on about 2.2 million loans by Lending Club in Q1 2017
  - Funded by hedge funds, banks, family offices
  - Portfolio optimization problems (optimization of risk-return)!

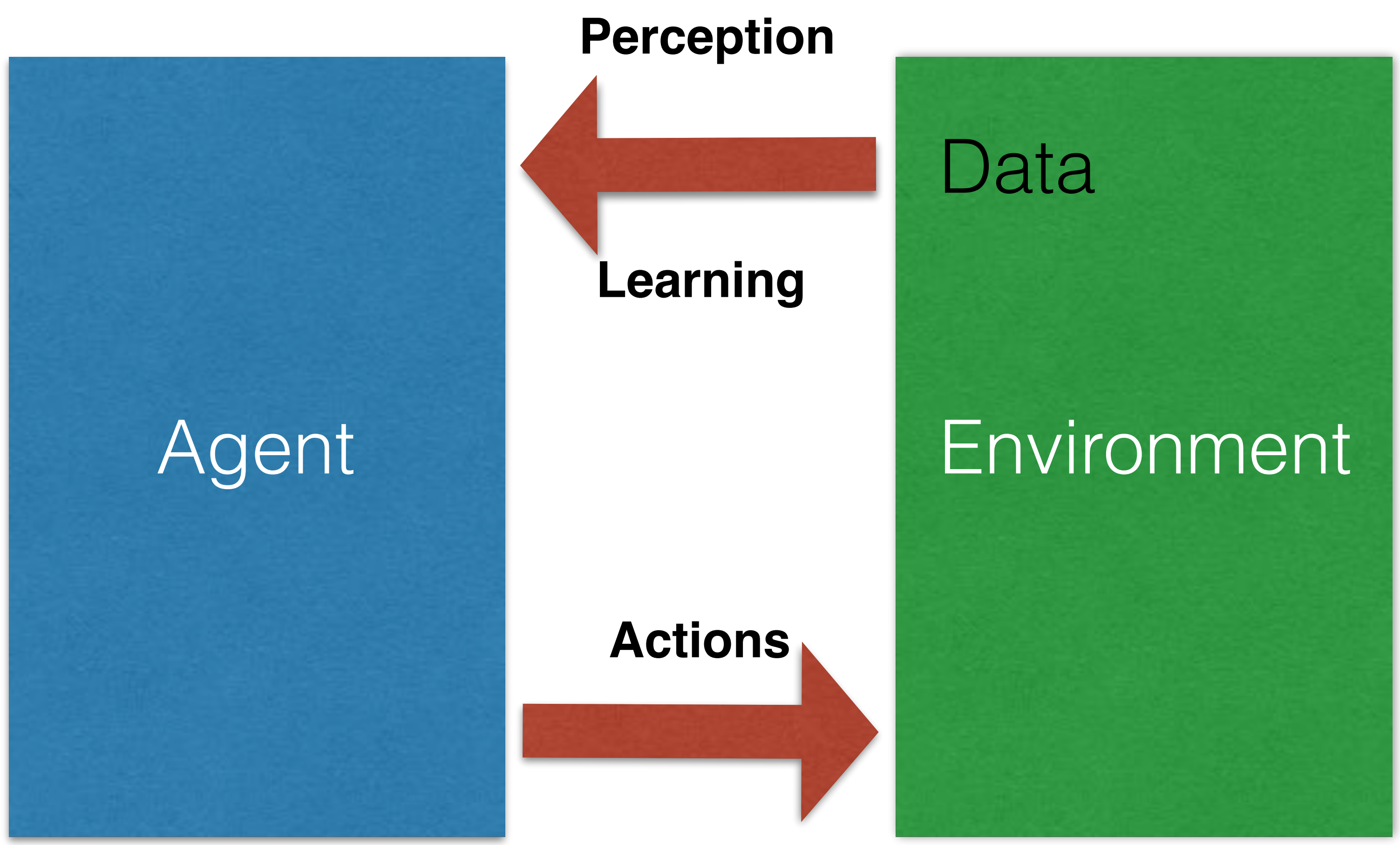
- **Reinforcement Learning for P2P credit portfolio:**
  - Sequential decision making
  - Can be formulated as RL with convex objective, similar to how we did it with stocks and options
  - Some differences in the problem formulation, e.g. you can only take long positions

RL for cryptocurrency portfolio management

- **Cryptocurrencies:**
  - Bitcoin, Etherium, etc - more than 100 cryptocurrencies
  - 2017: $120Bn market value ($40bn in Bitcoin)
  - Very similar to classical financial markets in may ways, but not all

- **Reinforcement Learning for portfolios of cryptocurrencies:**
  - See e.g. Z. Jian et. al. "A deep Reinforcement Learning Framework for the Financial Portfolio Management Problem" (arXiv: 1706.10059)
  - Deep RL with CNN, RNN, LSTM

Perception-action cycles



“Perception tasks” and action tasks

- Distinct tasks in traditional ML/RL approaches
- Become a part of the same perception-action loop in information-theory based approaches (Ortega, Braun, et. al. “Information-Theoretic Bounded Rationality” (2015))
- Can be important for better feature selection

Control question

Select all correct answers

1. Perception-action cycles are patterns in human and machine learning, where at each stage, an agent tends to keep focusing on one of these tasks, producing cyclicality effects.
2. Perception-action cycles is a learning paradigm for an agent that considers its both types of interaction with the environment (respectively, perception and action) as parts of the same information loop using tools from information theory.
3. The advantage of the perception-action cycle-based approach is that both perception and actions can be modeled in this framework as deep neural networks.
4. One of the advantages of the perception-action cycle-based approach is that feature selection becomes an integral part of the main task of action optimization.

Correct answers: 2, 4.

Advanced Topics on Reinforcement Learning in Finance

Week 4 Lesson 2: Other topics on RL
4.Summary: Universality and Unification: RL and beyond

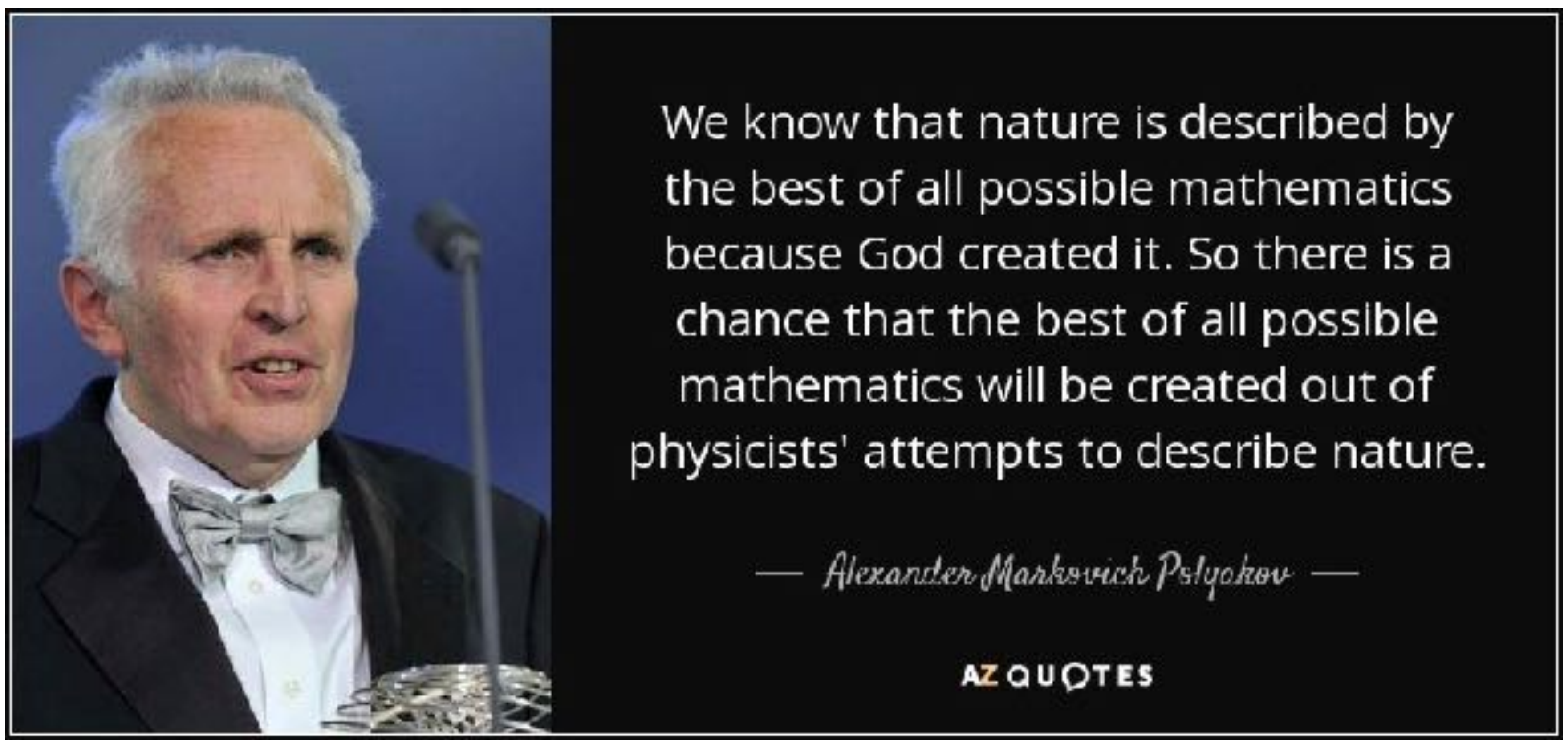
Igor Halperin

NYU Tandon School of Engineering, 2018

Generalization, deductive methods, universality

- **Generalization:**
  - A goal for Machine Learning
  - A goal for Human Learning
  - Deductive methods
  - Universality…

- **Universality of…**
  - “The true value of a physical theory is its degree of universality” (A. Polyakov. Fundamental Physics Prize 2013. http://www.math.columbia.edu/~woit/wordpress/?p=5679)



Universality of language

- **You say ‘tomato’, I say ‘tomato’!**

Universality and unification of methods

- Stochastic dynamics (Brownian motion, GBM, diffusion) = Quantum Mechanics
- Singular perturbation theory = WKB approximation of Quantum Mechanics
- Regularization = Bayesian priors
- Parametric vs non-parametric
- Bayesian vs frequentist methods
- Deep (… insert your favorite method…) for (…insert your favorite topic…)
- Universality of price formation across assets from Deep Learning: J. Sirignano and R. Cont, “Universal features of price formation in financial markets: perspectives from Deep Learning” (2018), http://lanl.arxiv.org/pdf/1803.06917.
- RL or/and ML as a preferred paradigm for problems in quantitative finance?
- RL can be used not only to compute some quantities, but to design new models, for example models of market dynamics.
- Optimize what?

Universalities and insights from physics

- "There is nothing more practical than a good theory" (V. Vapnik)
- Many methods from the physical sciences are used in both Finance and ML/RL
- Financial models (Brownian motion, GBM, diffusion models) have origins in physics
- The Langevin approach generalizes a linear GBM model
- Leads to potentially far richer dynamics
- Entropy: Statistical Physics -> ML -> General AI?
- Free energy: Statistical Physics -> ML -> RL -> self-organization -> biology?
- Symmetries: drive universality of behavior (phase transitions, etc.)
- Analytical properties in the complex plane…
- Thermodynamics and Quantum Mechanics as tools for ML and RL
- (see P. Ortega and D. Braun, "Thermodynamics as a theory of decision-making with information-processing costs", Proceeding of the Royal Society A, 2013, http://rspa.royalsocietypublishing.org/content/469/2153/20120683)
- Perception-action cycles and feature selection

The end

Thank you!