



Understanding a bit of AI in the Capital Markets

Kevin Compher

U.S. Securities and Exchange Commission

Division of Trading and Markets

August 28, 2019



"The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. These slides expresses the author's views and does not necessarily reflect those of the Commission, the Commissioners, or other members of the staff."



Equity Exchanges (Lit)

- NYSE
- NYSE American
- NYSE Arca
- NYSE National
- NASDAQ
- NASDAQ BX
- NASDAQ PSX
- BYX Equities
- BZX Equities
- CHX
- EDGA Equities
- EDGX Equities
- IEX

Market Participants

- Broker/Dealers,
- Banks,
- Market Makers,
- Quant Funds...

Investors

- Retail
- Institutional

Misc.

- Fraudsters
- Cyber..

Dark Pools and ATS'

of which there are < 30

Route to other markets?

- Options, Fixed Income,
- Derivatives, Commodities
- Forex/Foreign Currency or Cryptocurrencies

Dynamics & Perturbations

- (...just a few)
- ... **News & Economic Factors:**
Earnings Announcements, Trade Tensions, Recession, Yield Curve, Funds Rate, etc....
- ... **Algo design,** Alpha factor research, movements of herd, Liquidly, Order Flow...
- ... **Anomalies,** flash crashes, volatility events, circuit breakers, halts...
- ...**Exchange characteristics:**
Execution Quality, Speed
Transaction Costs, Rule changes,
National Best Bid & Offer

- < 1 ns timestamp (gps)
- fiber lines direct & cross connects
- Market Data Centers

- Consolidated processors & Market data providers
- Smart Order Router target < 4 μs windows
- Speed Bump 350 μs

- In 2018,
- Regulators saw a mean of 50 to > 100 billion messages a day ([FINRA](#))
 - Notional value of all trades \$90 trillion market cap of \$30 trillion ([CBOE](#))
 - Over last ten years, 90% of trading has become automated ([JP Morgan](#))

SEC's Mission
protect investors,
maintain **fair,**
orderly, **efficient**
markets & facilitate
capital formation



Finance Is Not Physics *"I can calculate the movement of the stars, but not the madness of men."*
 Isaac Newton, after losing a fortune on his investment in the South Sea Company

All financial models are affected by a Trinity of Errors
 (Deepak Kanungo 2019):

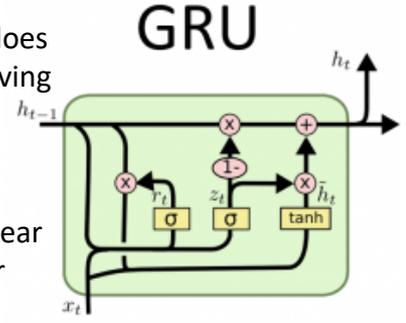
- Errors in...
- model specification
 - model parameter estimates
 - failure of a model to adapt to structural changes



Classical time series mining with normal distributions, does not work well, i.e. Moving Average, Autoregressive Moving Average (ARMA), (ARIMA), (SARIMA), etc.



Turn to Markov Chain Monte Carlo models, Bayesian linear regression, Deep RL, Recurrent Neural Nets in particular
[Long Short-Term Memory](#) and [Gated Recurrent Units](#)



IEX Crumbling Quote signal

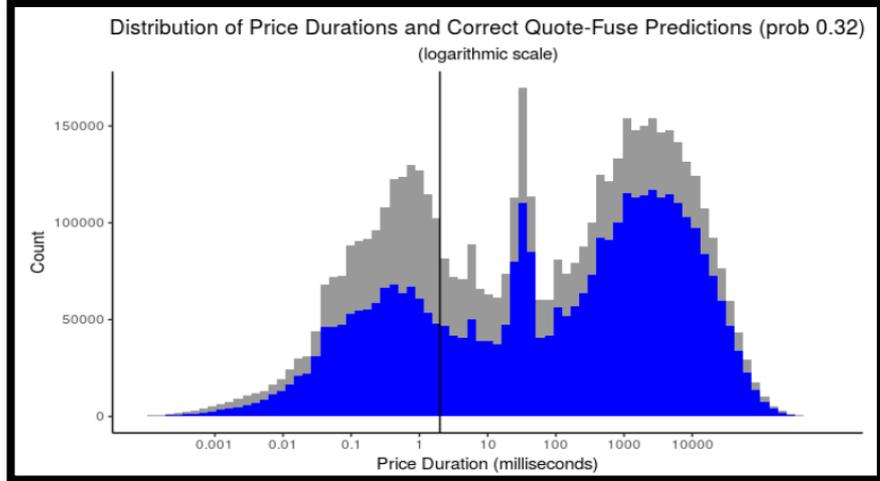
... is a predictive model that forecasts when the price of a stock is likely about to change. The current price of a stock is a compilation of the best prices published by the 13 stock exchanges. Prices don't change as a single event; rather they occur as a **sequence of updates over a sub-second timeframe**, which is only complete when the final exchange's price changes. (Allison Bishop 2017)



$$p := \frac{1}{1 + e^{-(c_0 + c_1 bids + c_2 asks + c_3 bids1 + c_4 asks1)}}$$

Predicting "Crumbles" and Improvements

A key benefit of Quote Fuse is its ability to predict the duration of NBBO quote prices regardless of the direction of the next tick. (Exegy Quote Fuse v1.1 2019)



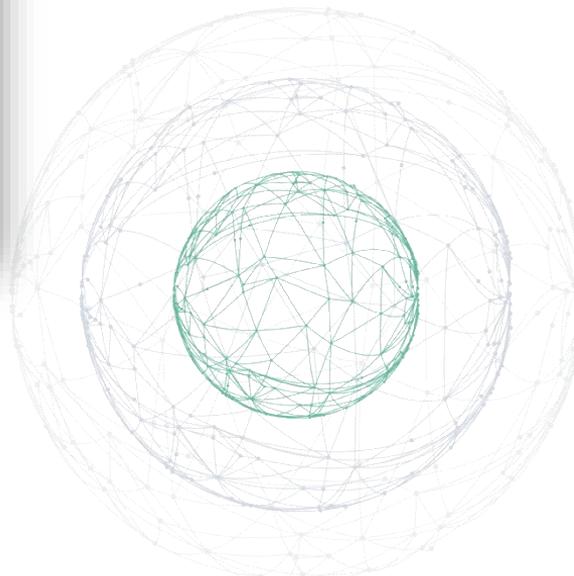
Price duration (grey), opportunity capture (blue)

Backup slide 1: AI-Based Matching Engines



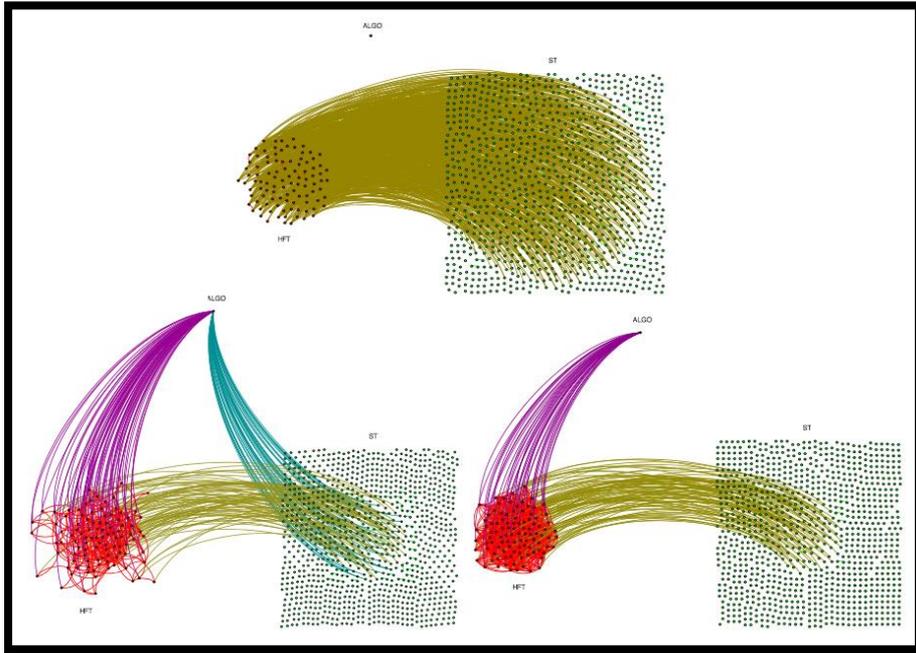
Traditional Limit order book heatmap
(Stefan Jansen 2018)

Intelligent Midpoint: Intelligent Midpoint matches orders at discrete times, milliseconds apart. Its AI system measures the price response after each trade and calibrates the match times to keep it as close to zero as possible. Lower market response leads to lower information leakage, which means less slippage for institutional orders.



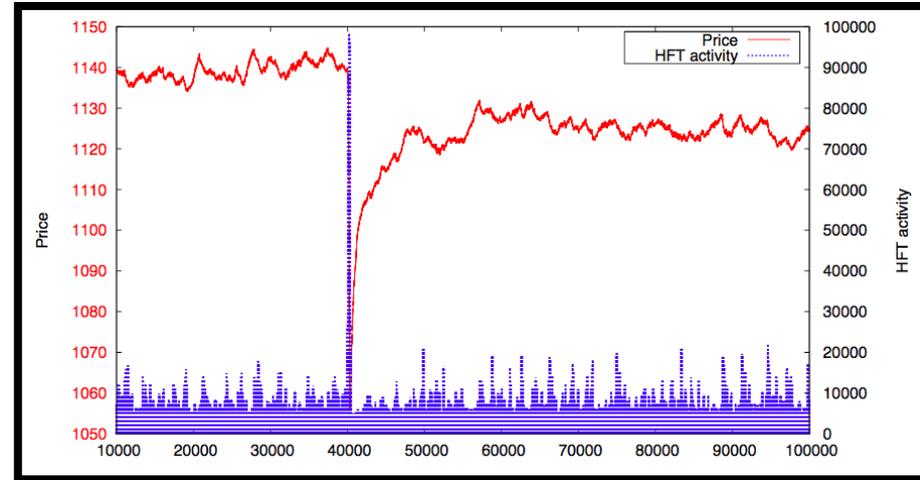
- NBBO midpoint orders
- Discrete matches, milliseconds apart, calibrated per-name for minimal post-trade price response
- Continuous order management—orders can be updated without delays
- Only interacts with other midpoint orders

Backup slide 2: Flash Event Modeling

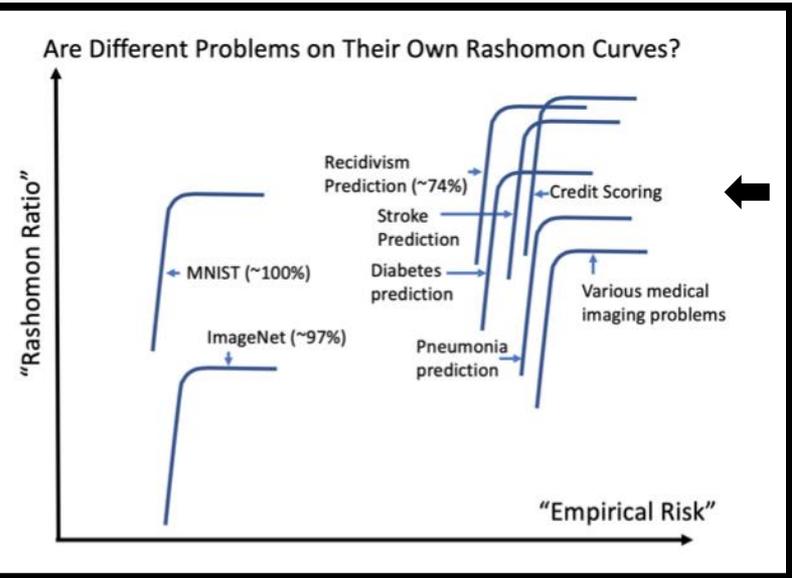


Network snapshots of the market behaving normally (top), when ALGO starts selling and HFTs absorb the initial sell pressure a moment before the hot-potato effect starts (bottom left), and when the price reaches its trough (bottom right). (Vuorenmaa & Wang 2016)

Simulated price and aggregate High Frequency Trading activity showing flash crash

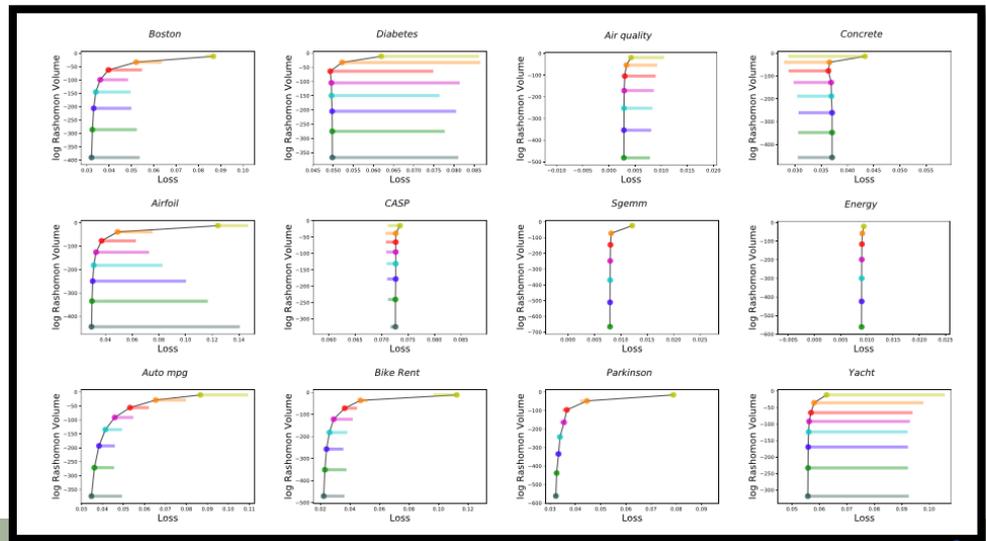


Backup Slide 3 – When is a more complex, black box needed?



A possible perspective on modern machine learning and Rashomon curves: at a given point in time, the state-of-the-art for different datasets, data types, and algorithmic performance may be viewed as a location on that problem's Rashomon curve. The maximal accuracy for each problem is different, as well as the location and shape of the Rashomon elbow. This figure is just an illustration – the values on the axes are not precise.

The generalization ability of the Rashomon elbow for the UCI regression datasets with real features.



A study in Rashomon curves and volumes: A new perspective on generalization and model simplicity in machine learning (Lesia Semenova and Cynthia Rudin 2019)

"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development
(NITRD) Program

Mailing Address: NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

Physical Address: 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674,
Fax: 202-459-9673, Email: nco@nitrd.gov, Website: <https://www.nitrd.gov>

