Project HighLO: identifying anomalies in financial market data using particle physics tools

Axel Naumann, Lorenzo Moneta, Jonas Rembser,

Joost Pennings, Tarek Alskaif, Marjolein Verhulst & Philippe Debie

June 29, 2023

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Financial Markets

For example

- Stocks
- Futures and options

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLo

Challenge For Regulators and Researchers

- Effectively combatting manipulation in financial markets
- So much data, where to begin?
- How to convert *messages* in an understandable and analysable format?
- Irregular data
    - Traditional time series analyses difficult to apply

Project ***High*** *Energy Physics Tools in* ***L****imit* ***O****rder Book Analysis* (HighLO)

- **Goal:**
  Interdisciplinary develop new methodologies to detect/identify manipulation

- **Data (so far):**
  - CME Group
  - All futures and options markets
  - 2015 (**±30 TB**) and 2019-2020 (**±300 TB**)
  - Time in **milli- and nanoseconds**
  - Limit orderbook and order data

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

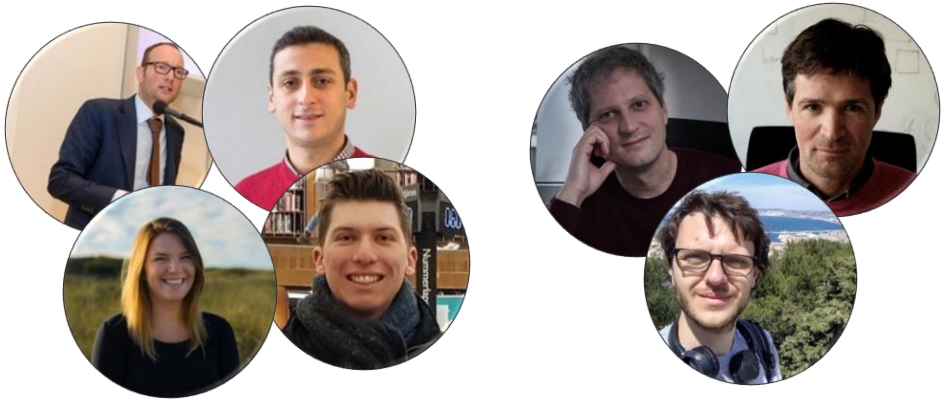
CERN

HighLO

Key Team and Partners Project HighLO

www.highlo.org

Key team



Affiliated Researchers

WAGENINGEN UNIVERSITY & RESEARCH

Maastricht University

BioDynaMo

Autoriteit Consument & Markt

University of Idaho

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Limit orderbook

- Marketplace showcasing all **demand** and **supply** for a specific product

- Limit orders
- Market orders

	Level	Price	Quantity
Ask (sell)	5	$55.00	2
	4	$54.60	10
	3	$53.50	2
	2	$53.00	4
	1	$50.00	9
		Spread	
Bid (buy)	1	$49.00	12
	2	$48.00	4
	3	$47.50	38
	4	$46.00	3
	5	$40.00	1

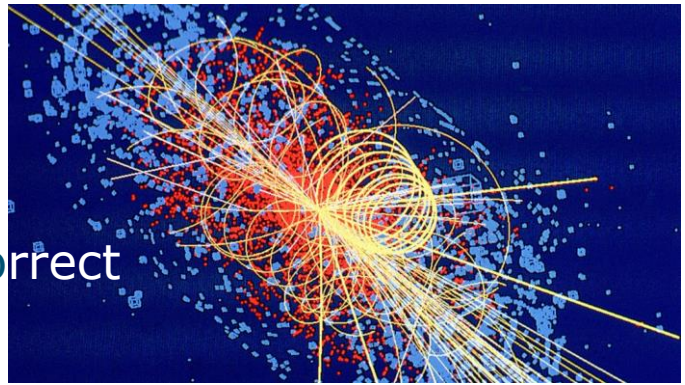
WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLo

Why CERN?



- Data processing framework: ROOT
- 1EB data in ROOT format
- Scientifically analyse data; statistically correct
- Highly efficient and customizable
- Data storage support, machine learning, visualization
- Open source and supported for decades
- Love to look at masses of raw data!

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

CERN vs. Limit Orderbook (LOB)

CERN	LOB
Look for **unusual** particles or for **anomalies** that do not fit the Standard Model	Look for **unusual** trading behavior or for **anomalies** that fit manipulative trading

First step: visualizations

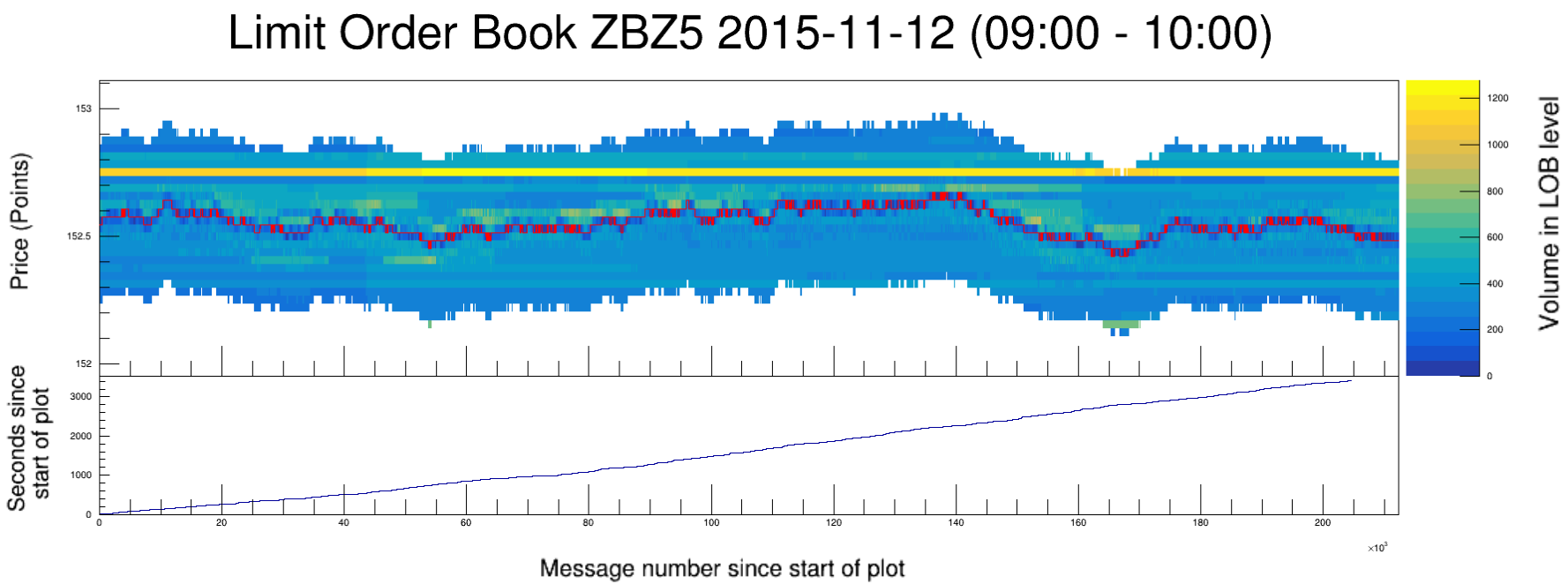
Verhulst, M. E., Debie, P., Hageboeck, S., Pennings, J. M., Gardebroek, C., Naumann, A., ... & Moneta, L. (2021). When two worlds collide: Using particle physics tools to visualize the limit order book. *Journal of Futures Markets*, 41(11), 1715-1734. https://doi.org/10.1002/fut.22251

However, there's one problem...

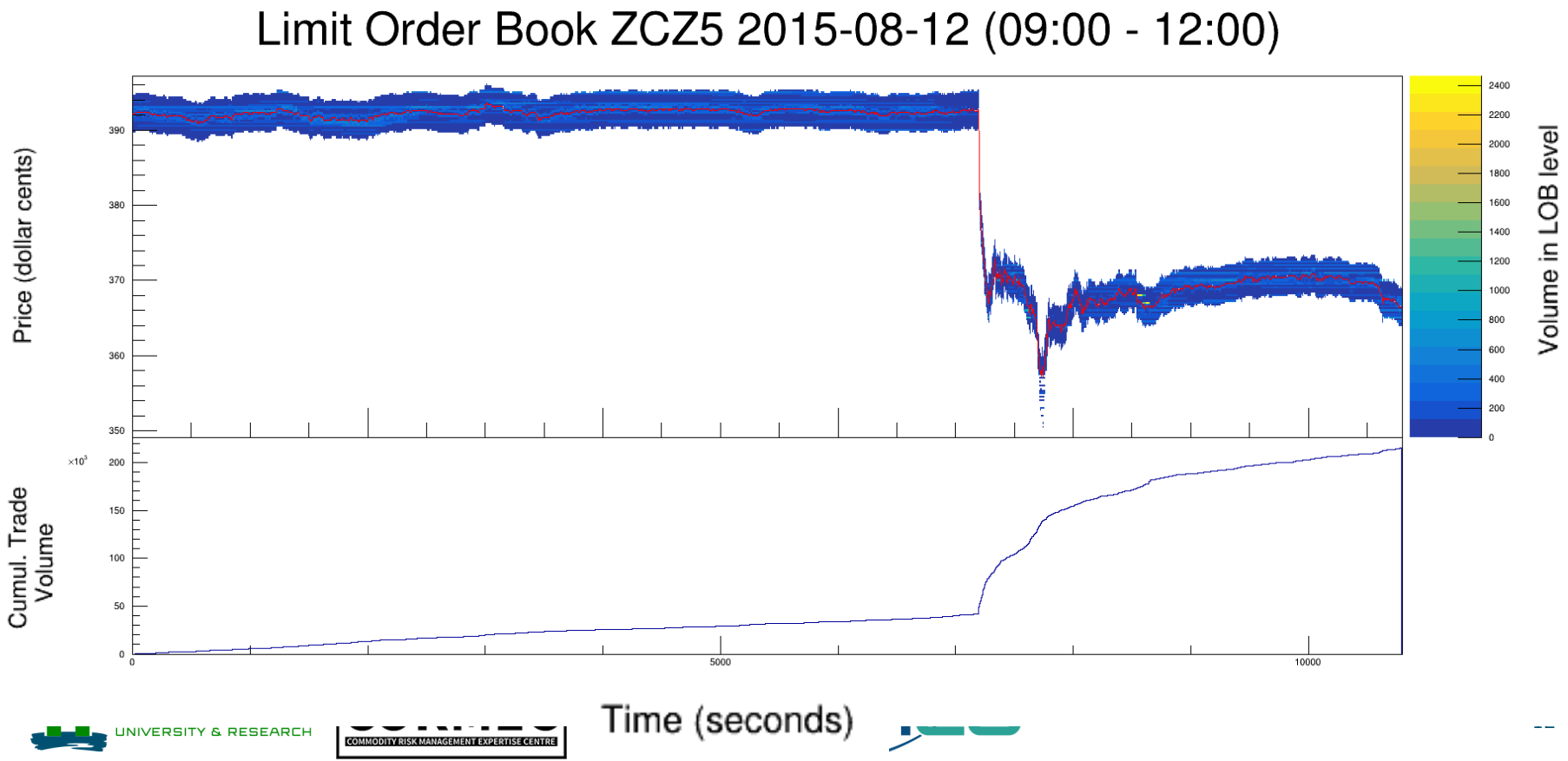
- Traditional timeseries analysis uses **regular** intervals in data
  - E.g. seconds, minutes, hours, days, etc.
- **But** our data is not regular
  - Market activity is irregular
  - Sometimes 20 updates in the same **millisecond!**

→ Traditional timeseries research aggregates to make it regular: lose a lot of information (data)!

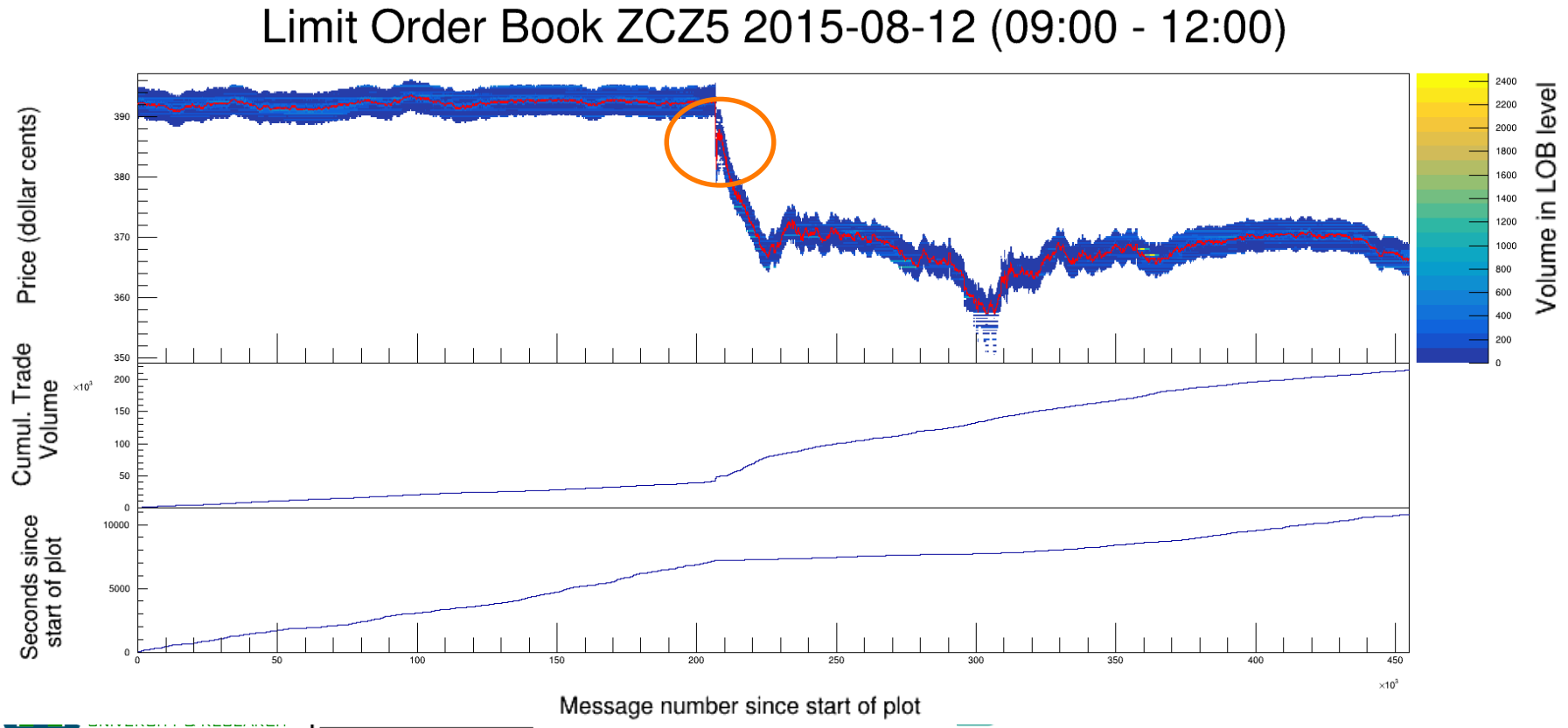
The new visualization: Dec 2015 T-Bond



Traditional vs. new visualization



... the new visualization



Visualizing Manipulation: the JPMorgan Case

Debie, P., Gardebroek, C., Hageboeck, S., van Leeuwen, P., Moneta, L., Naumann, A., ... & Verhulst, M. E. (2022). Unravelling the JPMorgan Spoofing Case Using Particle Physics Visualization Methods. *European Financial Management*. Early View. **https://doi.org/10.1111/eufm.12353**

JPMorgan Market Manipulation

- 2008 – 2016
- Spoofing
- Benefitted $172,034,790
- $311,737,008 market damages
- Settlement $920 million

FINANCE • JPMORGAN CHASE

JP Morgan will pay record settlement to resolve 'spoofing' case against 15 traders

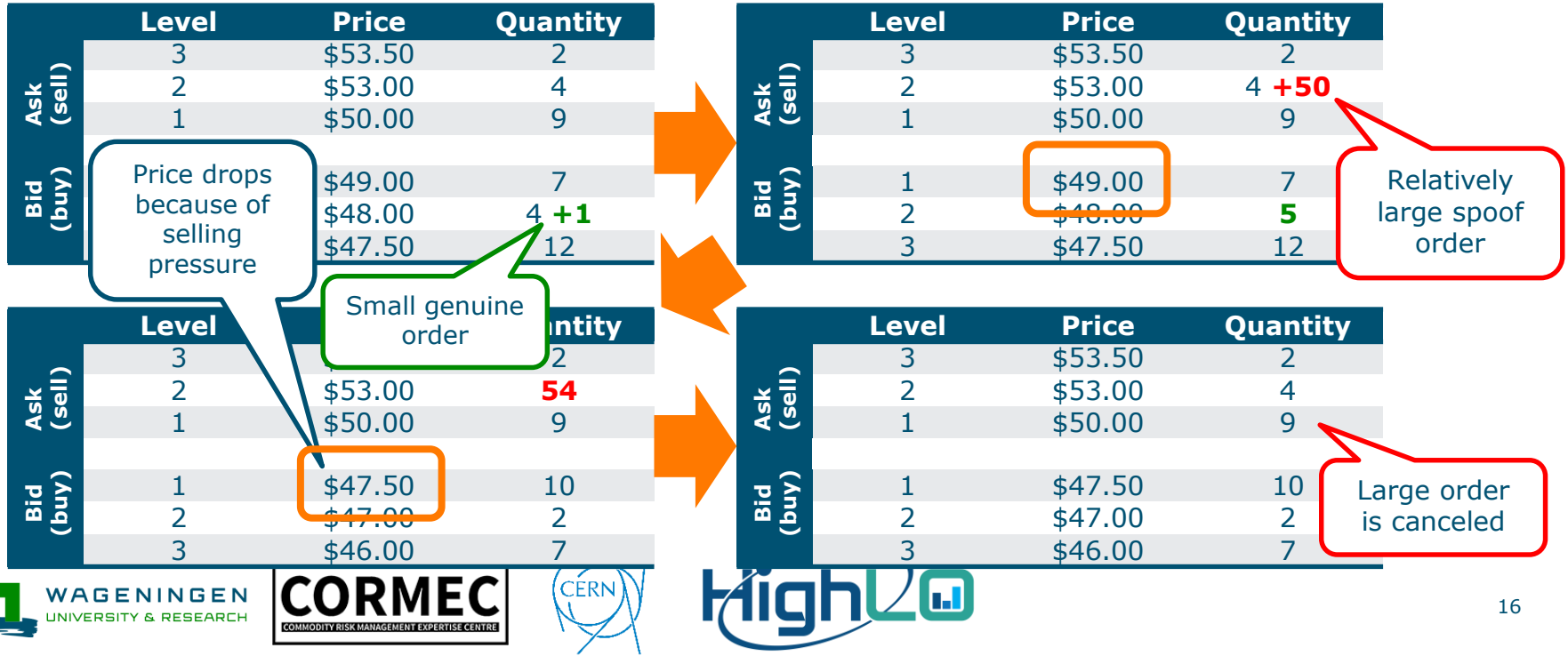
BY TOM SCHOENBERG, MATT ROBINSON, AND BLOOMBERG
September 30, 2020 11:18 AM GMT-2

APAC SEPTEMBER 29, 2020 / 6:55 PM / UPDATED 2 YEARS AGO

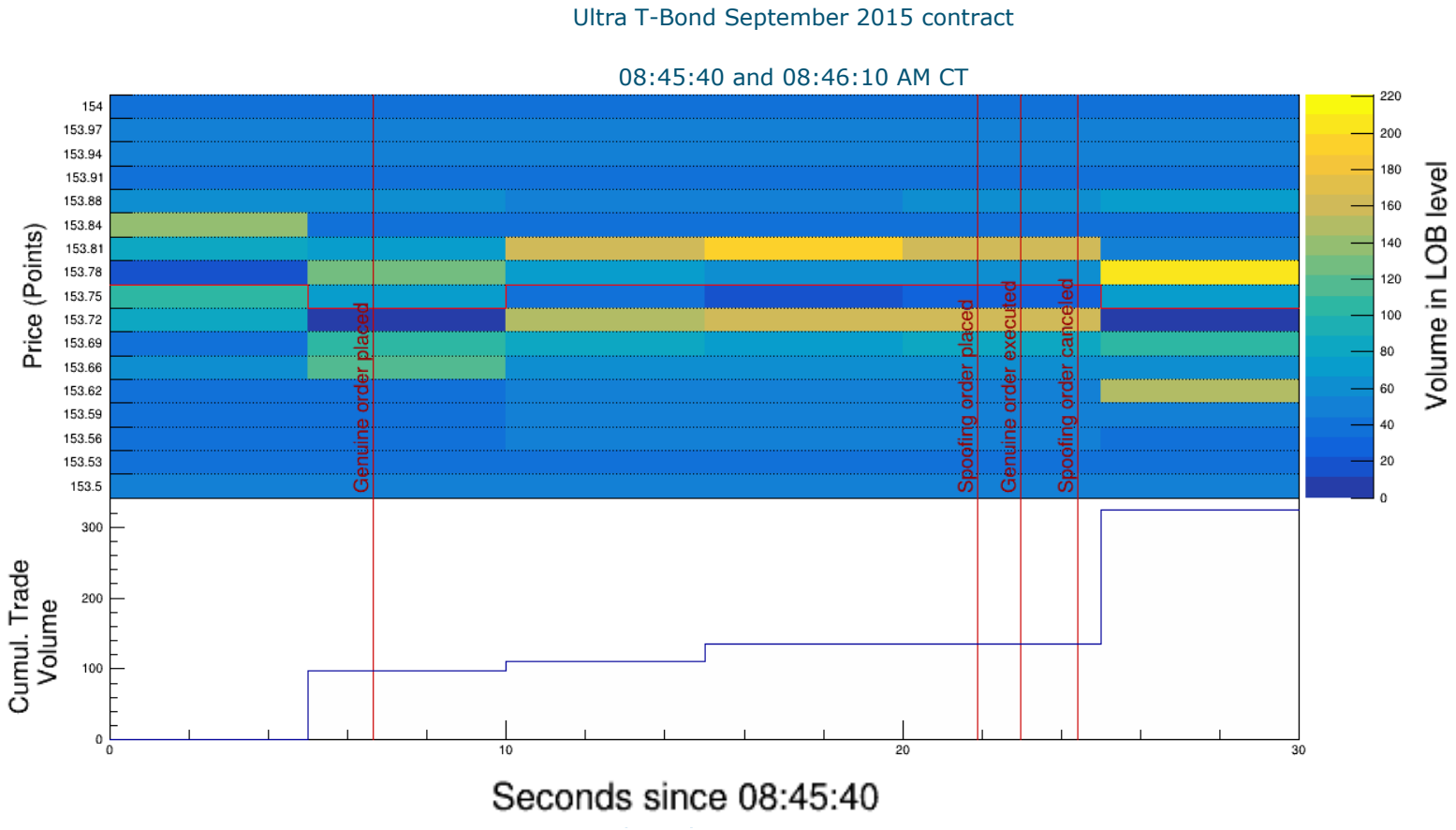
JPMorgan to pay $920 million for manipulating precious metals, treasury market

Market manipulation: spoofing

- *"The illegal practice of bidding or offering with **intent** to cancel before execution."* (Dodd-Frank Act, 2010)



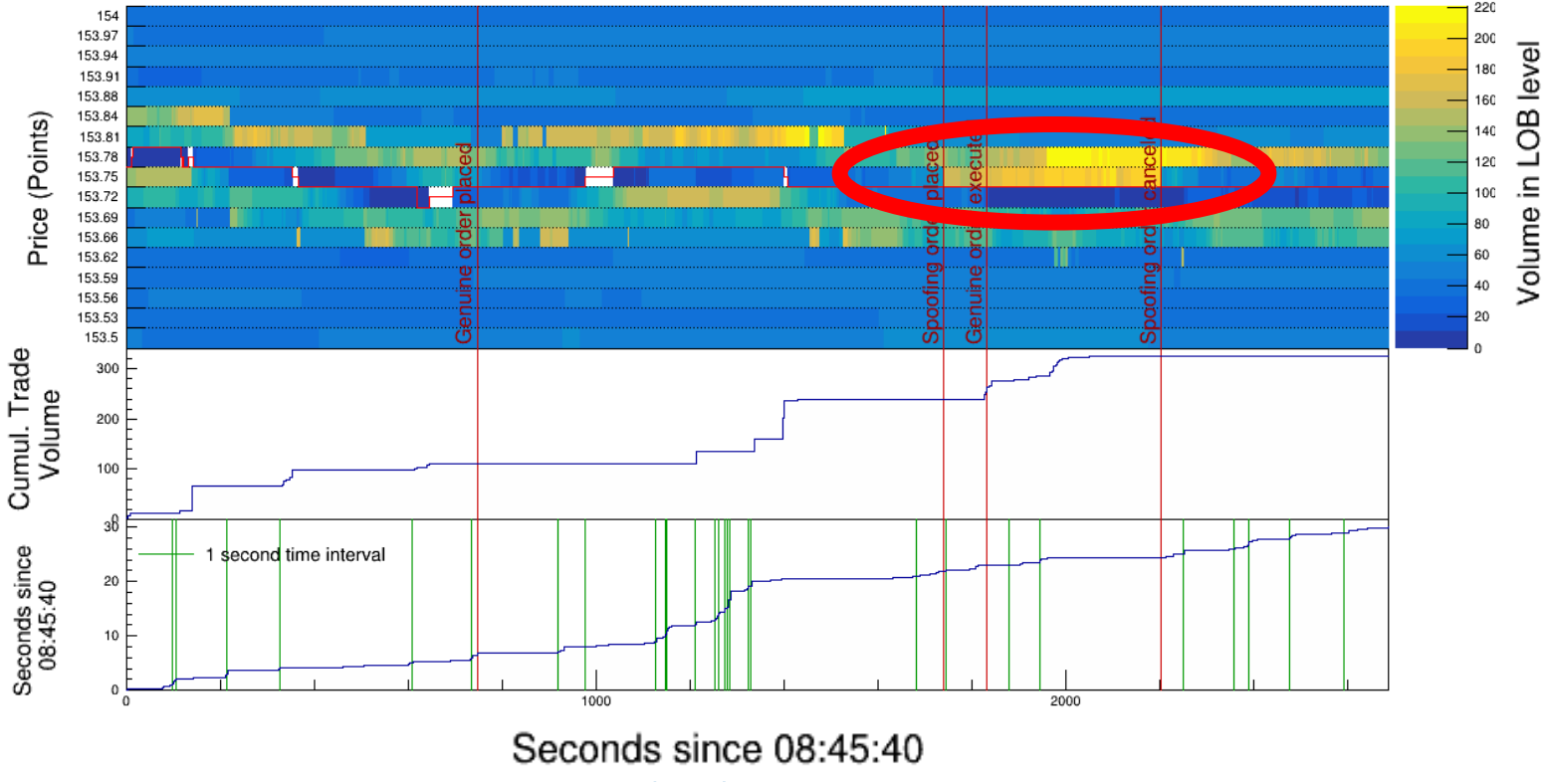
JPMorgan Case: 5-second interval



JPMorgan Case: all market activity

Ultra T-Bond September 2015 contract

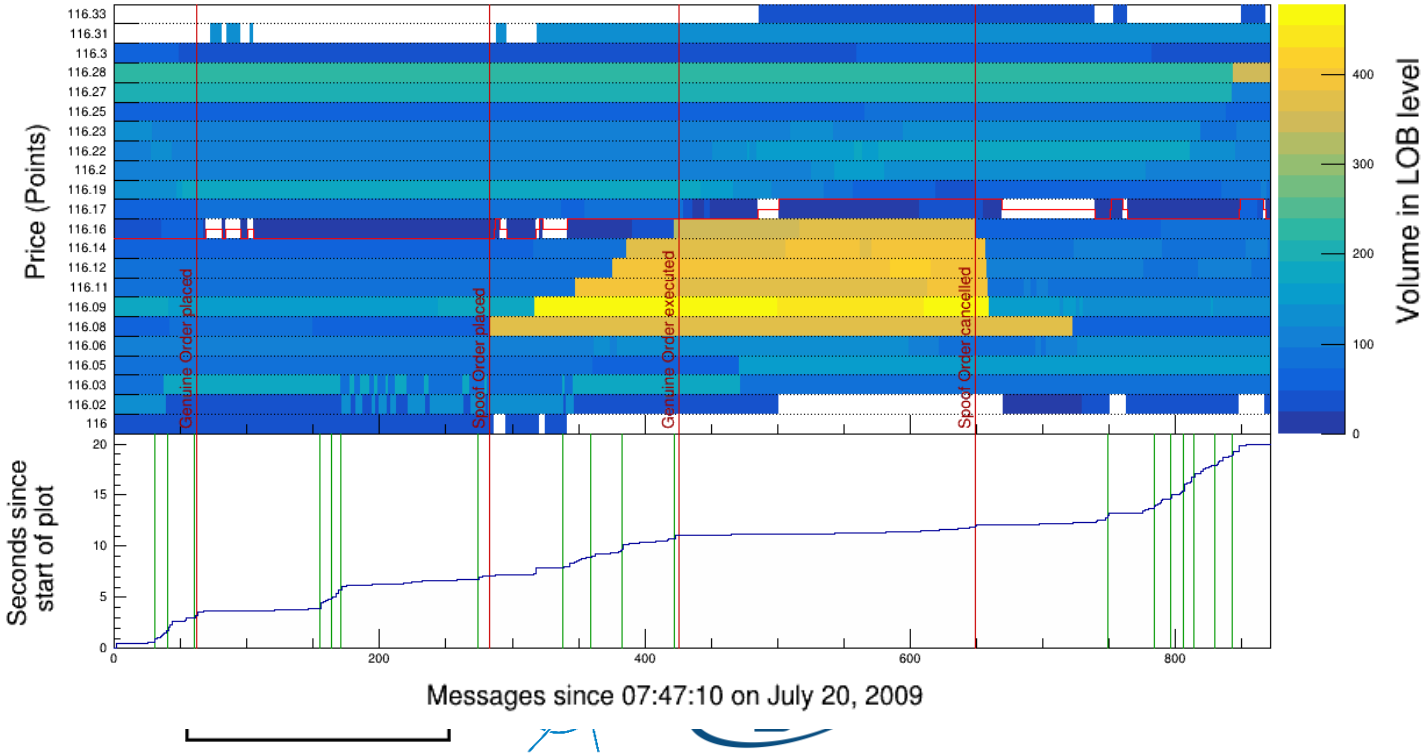
08:45:40 and 08:46:10 AM CT



Layered Spoofing

U.S. Treasury Bond September 2009 contract

07:47:10 – 07:47:30 AM CT



High-Frequency Data Analysis using the Tools from High-Energy Physics

Research Questions

1. How to measure the impact of an event

   on a time series

   using knowledge and tools from particle physics?

2. What is the impact of large limit order submissions

   on the transaction price?

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Impact Plot – The Concept

Trigger

Impact value

Idea:

Measure the changes in a time series
at fixed intervals before or after an event

Example:

What is the average price change 50ms after a limit order submission

Data

	E-mini S&P 500 futures	Crude oil Futures	US T-Bond Futures
CME Globex code	ES	CL	ZB
Data range	2019/07/01 – 2020/07/01		
Number of messages	4.9B	14.8B	1.0B
Raw data	656 GB	1 140 GB	144 GB
In ROOT	46 GB	80 GB	9 GB

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Impact plot – The Construction

1. Collect a set of triggers

   → Large limit order submissions

2. For each trigger, extract a time series

   → 200ms before to 200ms after the trigger

3. Overlap these time series (align the triggers)

4. For each time delay, build a distribution

   → 401 distributions, sequence of distributions

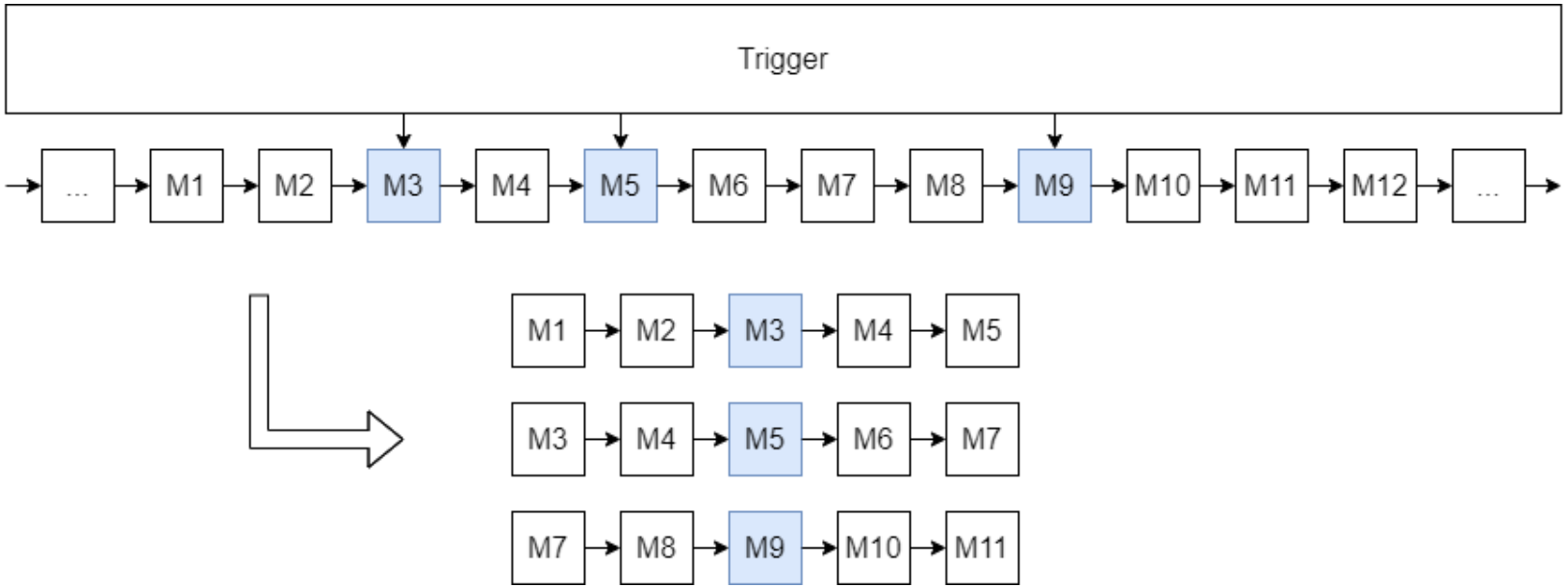
WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

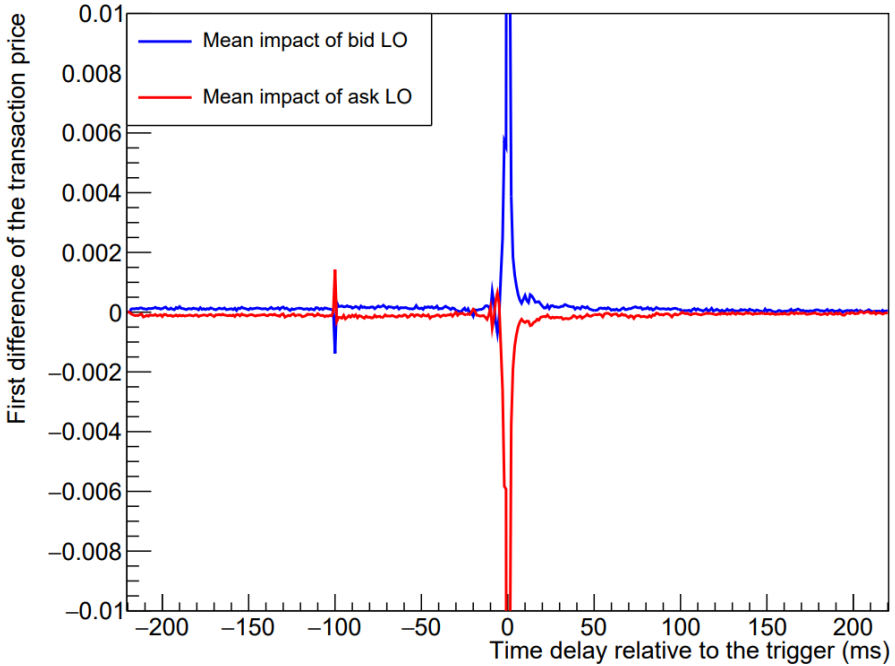
CERN

HighLo

Impact plot – A Diagram



Impact plot – E-mini S&P 500



Statistical Analysis

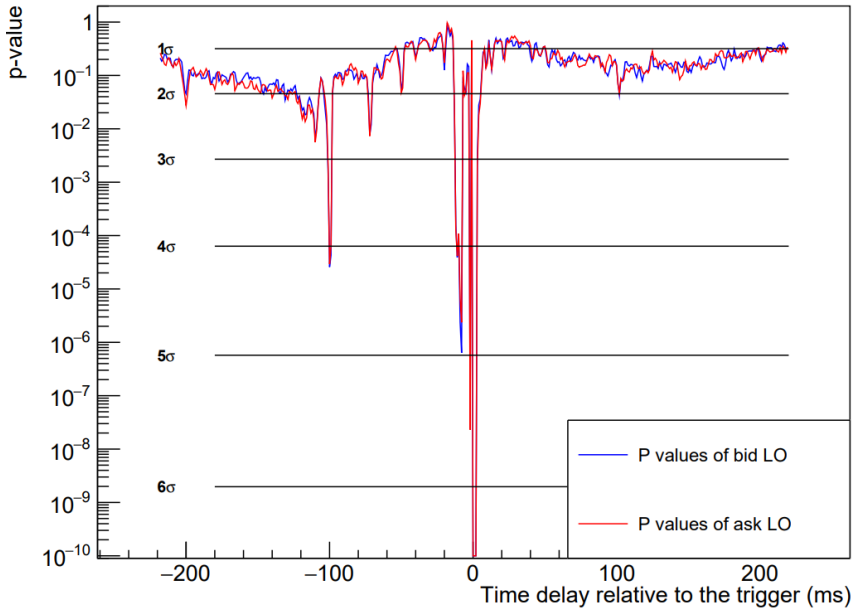
Compare the impact plot against average market conditions

- The signal impact plot: As described
- The background impact plot: Randomly drawn message

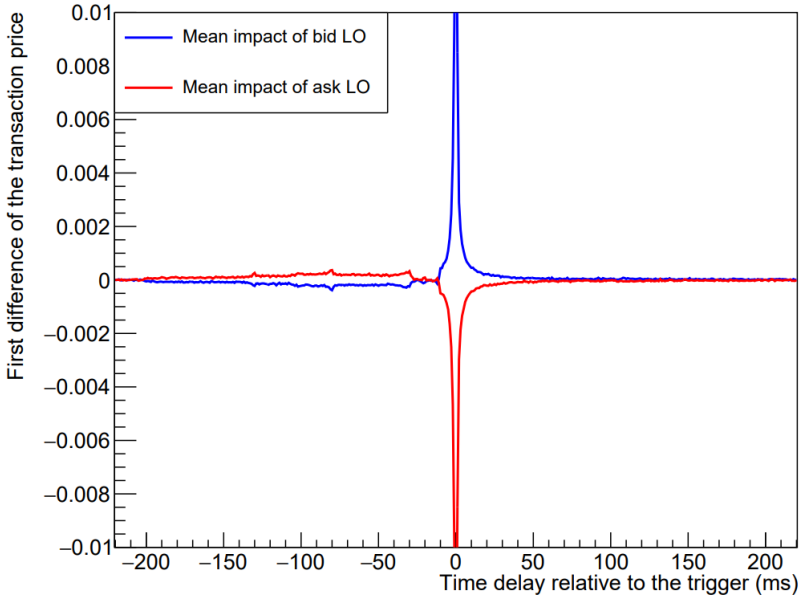
→ Pairwise statistical testing, per time slice

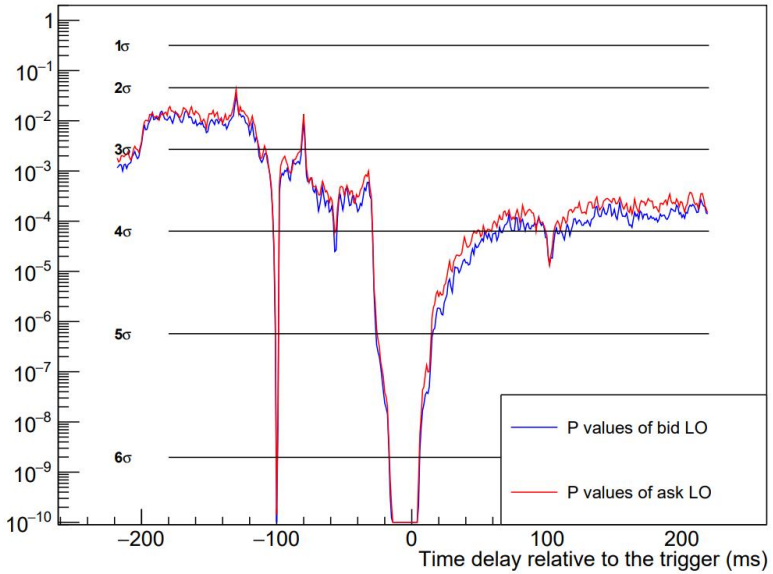
→ A sequence of p-values, one for each delay

Impact plot – E-mini S&P 500

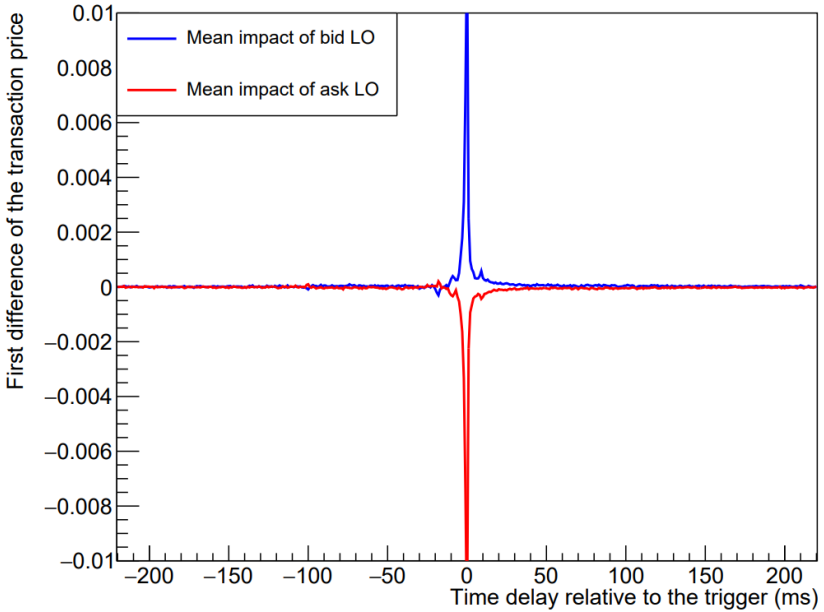


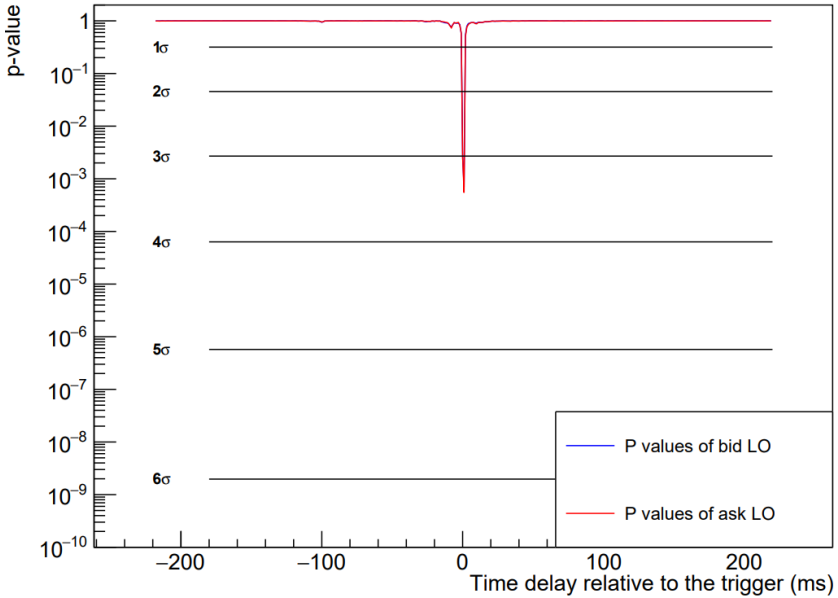
Impact plot – Crude Oil





Impact plot – US T-Bonds





WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Summary

Advantages compared to an impact analysis using VAR models

1. No data fitting or user-chosen modelling parameters
2. Better interpretation and visualisation
3. Linear computational cost and trivial multi-threading
4. Measure past and future time correlations

→ Interdisciplinary collaboration led to new techniques

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLo

An Agent-Based Financial Simulation of the Futures Commodity Market

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Manipulation Detector Verification

Possible verification methods:

- Testing on proven cases (= labels) → Bias to existing work
- Spoof the market to generate labels → Not feasible
- Spoof in a simulation

→ Agent-based modelling using BioDynaMo

Research Questions

1. What is the accuracy of our detection method?
2. Who is hurt due to the spoofing?

"The economy needs agent-based modelling"

Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. Nature, 460(7256), 685-686.

Next Steps

Next Steps for Project HighLO

- Interdisciplinary overview of spoofing
- Market manipulation/anomaly detection:
  - Spoofing: Expert vs. Machine
  - Detection tool & real-time monitoring
- Agent-based simulation:
  - BioDynaMo
  - Simulate financial markets
- Energy market collaboration
  - Dutch Authority for Consumers and Markets
  - ACER (EU)

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Next Steps for Project HighLO (2)

Need:

- Interaction academics and practitioners
- Feedback and input from regulators, exchanges, surveillance agencies
- Reactive and proactive

Goal: International Expert Group on Market Surveillance

International Expert Group on Market Surveillance (IMS Group)

- **18** regulatory agencies across the world
- Combine science with industry through **co-creation**

Goal

- Combine academic research with expert knowledge to:
    - Tackle market surveillance **challenges**
    - Create **consistency** in regulation, definitions and interpretations of the law
    - Develop new surveillance **methods and tools**

Kick-off at CERN: March 15-16, 2023

Topics:

- Past and current surveillance **systems**
- Surveillance in **energy** markets
- Spoofing **case**
- Spoofing **identification** tools and metrics
- **Challenges** of algorithmic trading surveillance
- Research done by **Project HighLO**

WAGENINGEN UNIVERSITY & RESEARCH

CORMEC COMMODITY RISK MANAGEMENT EXPERTISE CENTRE

CERN

HighLO

Questions?

axel.naumann@cern.ch

marjolein.verhulst@wur.nl

philippe.debie@wur.nl

www.highlo.org

References

Verhulst, M.E., Debie, P., Hageboeck, S., Pennings, J.M.E., Gardebroek, C., Naumann, A., ... & Moneta, L. (2021). When two worlds collide: Using particle physics tools to visualize the limit order book. *Journal of Futures Markets*, 41(11), 1715-1734. https://doi.org/10.1002/fut.22251

Debie, P., Gardebroek, C., Hageboeck, S., van Leeuwen, P., Moneta, L., Naumann, A., ... & Verhulst, M. E. (2022). Unravelling the JPMorgan Spoofing Case Using Particle Physics Visualization Methods. *European Financial Management*. Early View. https://doi.org/10.1111/eufm.12353

Debie, P., Verhulst, M.E., Rembser, J., Pennings, J.M.E., Naumann, A., Tekinerdogan, B., Moneta, L., Catal, C., Alskaif, T. & Demirel, S. (2022). *A Novel Approach to Examine the Impact of Behavior in Financial Markets using Particle Physics Methods: The Impact of Limit Order Submissions* (Working Paper).