Optimal execution with rough path signatures

DataSig

A rough path between mathematics and data science

Imanol Perez Arribas

Joint work with Alvaro Cartea, Jasdeep Kalsi, Terry Lyons and Leandro Sanchez Betancourt

University of Oxford · The Alan Turing Institute · Imperial College London · UCL · UKRI Engineering and Physical Sciences Research Council

Optimal execution

- Suppose I would like to **sell 1 million shares** of Apple.
  - How should I proceed?

DataSig

A rough path between mathematics and data science

Optimal execution

- Suppose I would like to **sell 1 million shares** of Apple.
  - How should I proceed?

- A naive approach would be to execute all 1 million shares **at once.**

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Optimal execution

- Suppose I would like to **sell 1 million shares** of Apple.
  - How should I proceed?

- A naive approach would be to execute all 1 million shares **at once.**

- An alternative approach would be to **slice the order** and execute the shares over a period of time.
  - How should I slice the order?

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Optimal execution

- There is a trade-off between **fast execution** (at a cost of obtaining a bad price) and **slow execution** (at a cost of obtaining an uncertain price).

- These trade-offs can be incorporated into an **optimal control problem**.

Market impact

- Let $X : [0, T] \to \mathbb{R}$ be a continuous stochastic process that models the **unaffected midprice** of the asset.

- Assume w.l.o.g. that $X_0 = 1$.

- Denote by $(\theta_t)_{t\in[0,T]}$ the **speed of trading**; i.e. the speed at which the order is executed.

DataSig
A rough path between mathematics and data science

Market impact

- Let $X : [0, T] \to \mathbb{R}$ be a continuous stochastic process that models the **unaffected midprice** of the asset.

- Assume w.l.o.g. that $X_0 = 1$.

- Denote by $(\theta_t)_{t\in[0,T]}$ the **speed of trading**; i.e. the speed at which the order is executed.

- If I follow $(\theta_t)_{t\in[0,T]}$, the trading activity will have an impact on the price.

- The **execution price** will be given by

$$P_t^\theta := X_t - G((\theta_s)_{s\in[0,t]})$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Market impact: examples

- **Temporary market impact**: $G((\theta_s)_{s\in[0,t]]}) := \lambda\theta_t,\ \lambda > 0$
- **Permanent market impact:** $G((\theta_s)_{s\in[0,t]}) := \lambda \int_0^t \theta_s ds,\ \lambda > 0$
- **Transient market impact:** $G((\theta_s)_{s\in[0,t]}) := \lambda \int_0^t e^{-\rho(t-s)}\theta_s ds$
- etc

DataSig
A rough path between mathematics and data science
The Alan Turing Institute Imperial College London UCL

Optimal execution problem

- **Initial inventory:** $q_0 \in \mathbb{R}$ units of stock.
- **Terminal wealth:** $W_T := \int_0^T P_s^\theta \theta_s ds$.
- **Running inventory:** $Q_t := q_0 - \int_0^t \theta_s ds$.

DataSig

A rough path between mathematics and data science

The Alan Turing Institute

Imperial College London

UCL

Optimal execution problem

- **Value function:**

$$V^{\theta} := W_T - \phi \int_0^T Q_t^2 dt + Q_T (P_T^{\theta} - \alpha Q_T).$$

with $\phi, \alpha \geq 0$.

DataSig

A rough path between mathematics and data science

The Alan Turing Institute

Imperial College London

UCL

Optimal execution problem

- **Value function:**

$$V^\theta := W_T - \phi \int_0^T Q_t^2 dt + Q_T(P_T^\theta - \alpha Q_T).$$

with $\phi, \alpha \geq 0$.

- **Optimal execution problem:**

$$\sup_\theta \mathbb{E}[V^\theta]$$

DataSig
A rough path between mathematics and data science

Solving the optimal execution problem

- The **space of trading strategies** is very large and it may be difficult to optimise over it.

- One approach to follow could be to look for strategies in a **restricted** (but **large**) class of trading strategies.
  - $\theta_s = f^{\Theta}\left(t, (X_s)_{s\in[0,t]}\right)$

DataSig
A rough path between mathematics and data science
The Alan Turing Institute
Imperial College London
UCL

Solving the optimal execution problem

- The **space of trading strategies** is very large and it may be difficult to optimise over it.

- One approach to follow could be to look for strategies in a **restricted** (but **large**) class of trading strategies.
  - $\theta_s = f^{\Theta}\left(t, (X_s)_{s\in[0,t]}\right)$

- Example: **neural networks** [1].
  - This class of strategies is very large.
  - We know how to optimise over this class of trading speeds (i.e. by training the neural network).

DataSig
A rough path between mathematics and data science

[1] Leal, L., Laurière, M. and Lehalle, C.A., 2020. Learning a functional control for high-frequency finance. *arXiv: 2006.09611*.

Solving the optimal execution problem

- In this talk, we'll consider a different class of trading strategies instead: **signature trading strategies.**
  - This class is also very large (it can approximate general trading strategies).
  - We can efficiently optimise over this class of strategies.

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Notation and assumptions

- We denote $\widehat{X}_t := (t, X_t) \in \mathbb{R}^2$, where recall that X is the unaffected midprice.
- We assume that $\widehat{X}$ can be lifted to a geometric rough path, whose signature will be denoted by $\widehat{\mathbb{X}}^{<\infty}$ which takes values in $T((\mathbb{R}^2))$.

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Notation and assumptions

- We denote $\widehat{X}_t := (t, X_t) \in \mathbb{R}^2$, where recall that X is the unaffected midprice.
- We assume that $\widehat{X}$ can be lifted to a geometric rough path, whose **signature** will be denoted by $\widehat{\mathbb{X}}^{<\infty}$ which takes values in $T((\mathbb{R}^2))$.
- On the dual space $T((\mathbb{R}^2)^*)$ we make the identification with **words**

$$e_{i_1}^* \otimes \cdots \otimes e_{i_n}^\star \longleftrightarrow \mathbf{i_1} \cdots \mathbf{i_n}$$

  and the empty word is denoted by $\varnothing$.
- Given two words $\mathbf{w}, \mathbf{v}$ denote by $\mathbf{wv}$ their **concatenation** and by $\mathbf{w} \sqcup \mathbf{v}$ their **shuffle product.**

Shuffle product

- Recall the **shuffle product property:**

$$\langle f, \widehat{\mathbb{X}}_{0,t}^{<\infty} \rangle \langle g, \widehat{\mathbb{X}}_{0,t}^{<\infty} \rangle = \langle f \sqcup g, \widehat{\mathbb{X}}_{0,t}^{<\infty} \rangle$$

for all $f, g \in T((\mathbb{R}^2)^*)$.

Signature trading speeds

- We define the space of **signature trading strategies:**

$$\mathcal{T}_{sig} := \{t \mapsto \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle \ : \ \ell \in T((\mathbb{R}^2)^*)\}.$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Signature trading speeds

- We define the space of **signature trading strategies:**

$$\mathcal{T}_{sig} := \{t \mapsto \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle \ : \ \ell \in T((\mathbb{R}^2)^*)\}.$$

- We can **approximate** general trading strategies by such signature trading strategies.

- We will consider the optimal execution problem over signature trading strategies:

$$\sup_{\theta \in \mathcal{T}_{sig}} \mathbb{E}[V^\theta]$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Market impact

For each $\theta \in \mathcal{T}_{sig}$ we consider market impacts of the form

$$G((\theta_s)_{s\in[0,t]})) := \langle g^\theta, \widehat{\mathbb{X}}_{0,t}^{<\infty} \rangle$$

with $g^\theta \in T((\mathbb{R}^2)^*)$.

DataSig

A rough path between mathematics and data science

Market impact

Let $\theta \in \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$, $\ell \in T((\mathbb{R}^2)^*)$.

- **Temporary market impact**. Set $g^\ell := \lambda \ell$. Then,

$$\langle g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle = \lambda \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle = \lambda \theta_t .$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Market impact

Let $\theta \in \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$, $\ell \in T((\mathbb{R}^2)^*)$.

- **Temporary market impact**. Set $g^\ell := \lambda \ell$. Then,

$$\langle g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle = \lambda \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle = \lambda \theta_t.$$

- **Permanent market impact**. Set $g^\ell := \lambda \ell \mathbf{1}$. Then,

$$\langle g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle = \lambda \int_0^t \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds = \lambda \int_0^t \theta_s ds.$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

Let $\theta_t := \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$ be a signature trading strategy. We have:

$$W^{\ell}_t =$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

Let $\theta_t := \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$ be a signature trading strategy. We have:

$$W^\ell_t = \int_0^t P^\ell_s \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

Let $\theta_t := \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$ be a signature trading strategy. We have:

$$\begin{aligned} W^\ell_t &= \int_0^t P^\ell_s \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\ &= \int_0^t (X_s - \langle g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle) \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \end{aligned}$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

Let $\theta_t := \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$ be a signature trading strategy. We have:

$$
\begin{aligned}
W_t^\ell &= \int_0^t P_s^\ell \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t (X_s - \langle g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle) \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t \langle \mathbf{2} + \varnothing - g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds
\end{aligned}
$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

Let $\theta_t := \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$ be a signature trading strategy. We have:

$$
\begin{aligned}
W_t^\ell &= \int_0^t P_s^\ell \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t (X_s - \langle g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle) \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t \langle \mathbf{2} + \varnothing - g^\ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t \langle (\mathbf{2} + \varnothing - g^\ell) \sqcup \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds
\end{aligned}
$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

Let $\theta_t := \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,t} \rangle$ be a signature trading strategy. We have:

$$
\begin{aligned}
W^{\ell}_t &= \int_0^t P^{\ell}_s \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t (X_s - \langle g^{\ell}, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle) \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t \langle \mathbf{2} + \varnothing - g^{\ell}, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle \langle \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \int_0^t \langle (\mathbf{2} + \varnothing - g^{\ell}) \sqcup \ell, \widehat{\mathbb{X}}^{<\infty}_{0,s} \rangle ds \\
&= \left\langle \left((\mathbf{2} + \varnothing - g^{\ell}) \sqcup \ell \right) \mathbf{1}, \widehat{\mathbb{X}}^{<\infty}_{0,T} \right\rangle
\end{aligned}
$$

Solving the optimal execution problem

$$W_t^\ell = \left\langle \underbrace{\cdots}_{\substack{\text{depends only}\\ \text{on the control } \ell}}, \underbrace{\widehat{\mathbb{X}}_{0,T}^{<\infty}}_{\substack{\text{depens on the stochastic}\\ \text{component } \widehat{X}}} \right\rangle$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute

Imperial College London

UCL

Solving the optimal execution problem

$$\mathbb{E}[V^\ell] = \Big\langle \underbrace{\cdots}_{\substack{\text{depends only} \\ \text{on the control } \ell}}, \underbrace{\mathbb{E}[\widehat{\mathbb{X}}^{<\infty}_{0,T}]}_{\substack{\text{depens on the stochastic} \\ \text{component } \widehat{X}}} \Big\rangle$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Solving the optimal execution problem

The original problem is written as:

$$\sup_{\ell \in T((\mathbb{R}^2)^*)} \left\langle \left((\mathbf{2} + \varnothing - g^{\ell}) \sqcup \ell \right) \mathbf{1} - (q_0 \varnothing - \ell \mathbf{1})^{\sqcup 2} (\phi \mathbf{1} - \alpha \varnothing) \right.$$
$$\left. + (q_0 \varnothing - \ell \mathbf{1}) \sqcup (\mathbf{2} + \varnothing - g^{\ell}), \mathbb{E}\left[\widehat{\mathbb{X}}^{<\infty}_{0,T} \right] \right\rangle$$

DataSig
A rough path between mathematics and data science

Solving the optimal execution problem

For a fixed signature degree N, we have to solve:

$$\sup_{\ell \in T^{(N)}((\mathbb{R}^2)^*)} \Big\langle \Big((\mathbf{2} + \varnothing - g^{\ell}) \sqcup \ell \Big) \mathbf{1} - (q_0 \varnothing - \ell \mathbf{1})^{\sqcup 2} (\phi \mathbf{1} - \alpha \varnothing) + (q_0 \varnothing - \ell \mathbf{1}) \sqcup (\mathbf{2} + \varnothing - g^{\ell}), \mathbb{E}\left[\widehat{\mathbb{X}}^{<\infty}_{0,T} \right] \Big\rangle$$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Numerical experiments

- We carry out different experiments for different **midprice models**.

- In each experiment, we:
    - Estimate the **expected signature** with Monte Carlo by sampling 50,000 realisations of the midprice process.

Numerical experiments

- We carry out different experiments for different **midprice models**.

- In each experiment, we:
  - Estimate the **expected signature** with Monte Carlo by sampling 50,000 realisations of the midprice process.
  - We **solve** the signature optimisation problem.

DataSig
A rough path between mathematics and data science

Numerical experiments

- We carry out different experiments for different **midprice models**.

- In each experiment, we:
    - Estimate the **expected signature** with Monte Carlo by sampling 50,000 realisations of the midprice process.
    - We **solve** the signature optimisation problem.
    - We **evaluate** the performance of the signature strategy on 10,000 new samples.

DataSig
A rough path between mathematics and data science

Numerical experiments

- We carry out different experiments for different **midprice models**.

- In each experiment, we:
  - Estimate the **expected signature** with Monte Carlo by sampling 50,000 realisations of the midprice process.
  - We **solve** the signature optimisation problem.
  - We **evaluate** the performance of the signature strategy on 10,000 new samples.
  - If available, we compare the performance of the signature strategy with the performance of the **known closed-form** solution of the problem.

DataSig

A rough path between mathematics and data science

Numerical experiments

- We carry out different experiments for different **midprice models**.

- In each experiment, we:
    - Estimate the **expected signature** with Monte Carlo by sampling 50,000 realisations of the midprice process.
    - We **solve** the signature optimisation problem.
    - We **evaluate** the performance of the signature strategy on 10,000 new samples.
    - If available, we compare the performance of the signature strategy with the performance of the **known closed-form** solution of the problem.
    - In all cases, we use the *time-weighted average price* (TWAP) strategy as a **benchmark**.

DataSig

A rough path between mathematics and data science

Numerical experiment I: incorporating order flow

- In (Cartea and Jaimungal, 2016) the authors incorporate the **order-flow** into the midprice dynamics.

- The midprice is given by

$$X_t := k \int_0^t (\mu_s^+ - \mu_s^-) ds + \sigma W_t$$

DataSig
A rough path between mathematics and data science

Numerical experiment I: incorporating order flow

- In (Cartea and Jaimungal, 2016) the authors incorporate the order-flow into the midprice dynamics.

- The midprice is given by

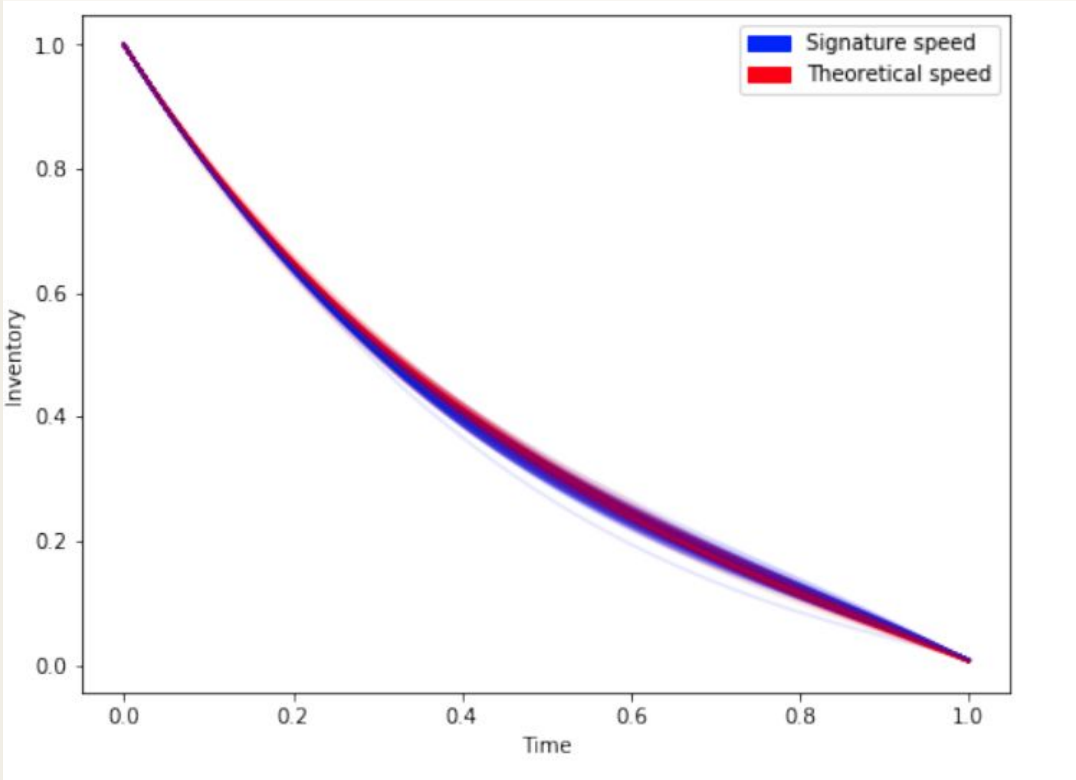
$$X_t := k \int_0^t (\mu_s^+ - \mu_s^-) ds + \sigma W_t$$

buy orders sell orders

Numerical experiment I: incorporating order flow

- We include **temporary** and **permanent** market impacts.
- Signatures of **order 7** are considered.

Numerical experiment I: incorporating order flow



DataSig
A rough path between mathematics and data science

Numerical experiment I: incorporating order flow

Theoretical optimal speed	Signature trading speed	TWAP
$0.995748 \pm 2.12\times10^{-5}$	$0.995697 \pm 3.97\times10^{-5}$	$0.993516 \pm 6.07\times10^{-5}$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Numerical experiment II: incorporating trading signals

- In (Lehalle and Neuman, 2017) the authors include **trading signals** the investor has access to, such as **order imbalance**. The signals predict short-term price movements.

- They model the midprice by

$$X_t := \int_0^t I_s ds + \sigma W_t$$

DataSig

A rough path between mathematics and data science

Numerical experiment II: incorporating trading signals

- In (Lehalle and Neuman, 2017) the authors include **trading signals** the investor has access to, such as **order imbalance**. The signals predict short-term price movements.

- They model the midprice by

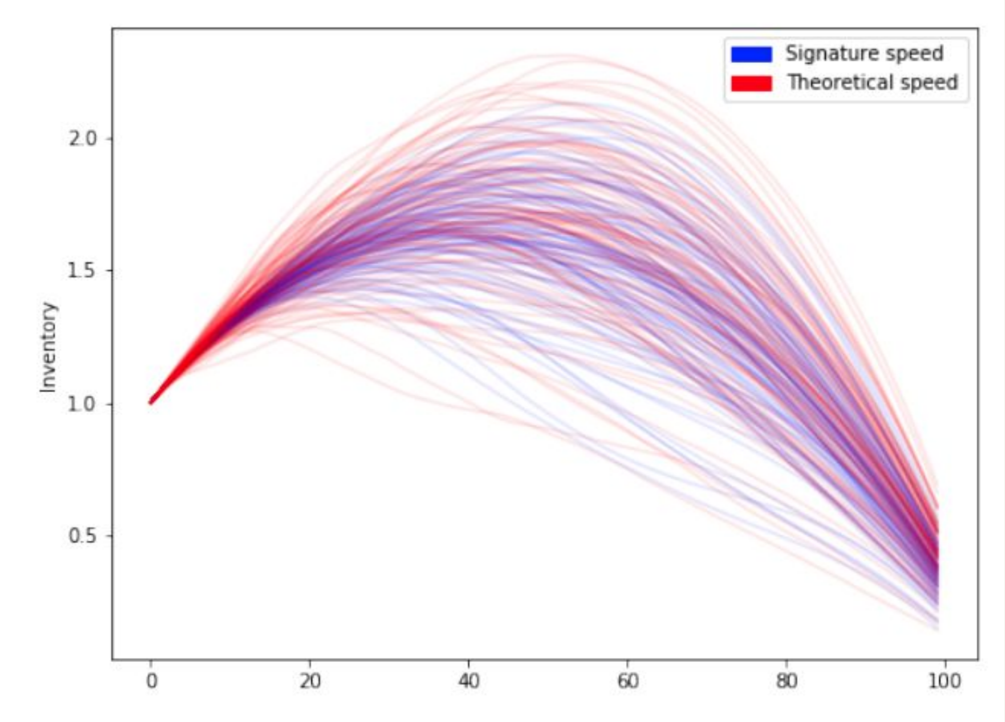
$$X_t := \int_0^t I_s ds + \sigma W_t$$

trading signal

Numerical experiment II: incorporating trading signals

- A **temporary market impact** is included.
- Signatures of **order 9** are considered.

Numerical experiment II: incorporating trading signals



Numerical experiment II: incorporating trading signals

Theoretical optimal speed	Signature trading speed	TWAP
$1.0170735 \pm 4.05\times10^{-5}$	$1.0169903 \pm 4.12\times10^{-5}$	$0.999856 \pm 2.21\times10^{-5}$

DataSig

A rough path between mathematics and data science

The Alan Turing Institute Imperial College London UCL

Numerical experiment III: fractional Brownian motion

- The midprice process is given by

$$X_t := \sigma W_t^H$$

  where H is the Hurst parameter.

- We consider the **rough case**: H < ½.

Numerical experiment III: fractional Brownian motion

- The midprice process is given by

$$X_t := \sigma W_t^H$$

where H is the Hurst parameter.

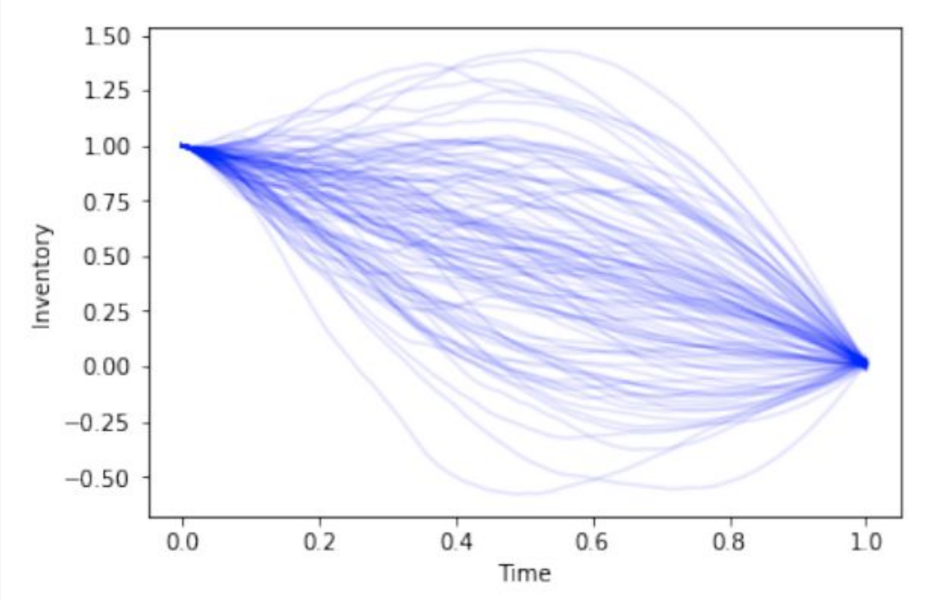
- We consider the **rough** (H < ½) and **smooth** (H > ½) case.
- Signatures of **order 7** are considered.
- We include a temporary market impact.

DataSig

A rough path between mathematics and data science

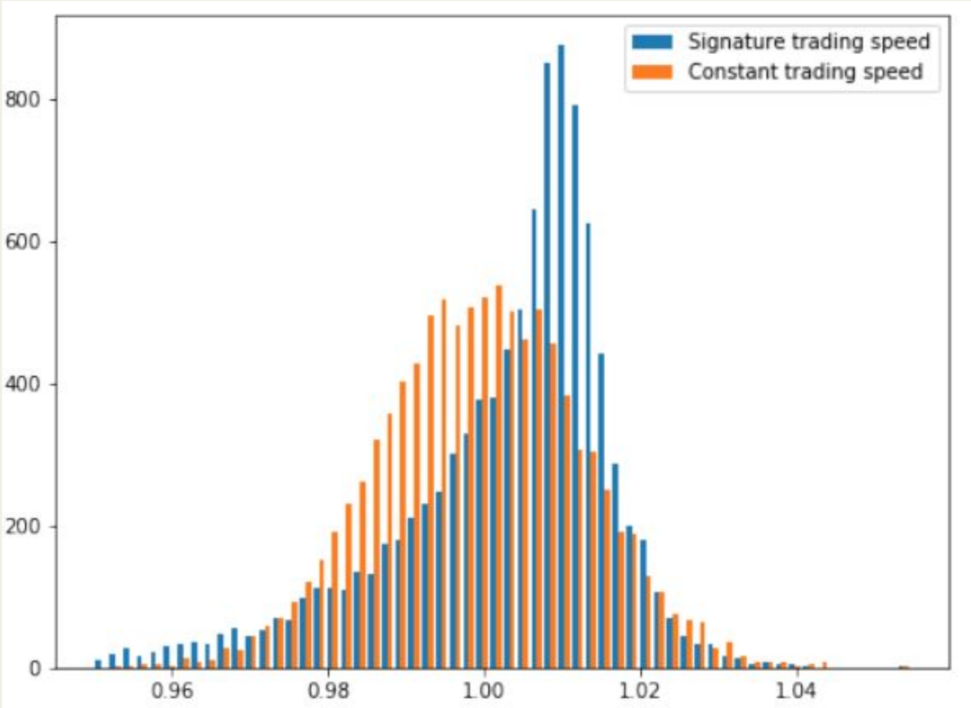
The Alan Turing Institute Imperial College London UCL

Numerical experiment III: fractional Brownian motion



H = 1/3

Numerical experiment III: fractional Brownian motion



H = 1/3

A rough path between mathematics and data science

Numerical experiment II: incorporating trading signals

H	Signature optimal speed	Signature trading speed
1/3	$1.0031498 \pm 1.38\times10^{-4}$	$0.9991785 \pm 5.72\times10^{-4}$
0.7	$1.0203925 \pm 2.82\times10^{-6}$	$0.99921470 \pm 1.08\times10^{-6}$

Numerical experiment IV: double-execution

- We want to sell a block of shares in a **foreign** stock market (e.g. Apple).
- The proceeds are exchanged into the investor's **domestic currency** (e.g. GBP).

DataSig

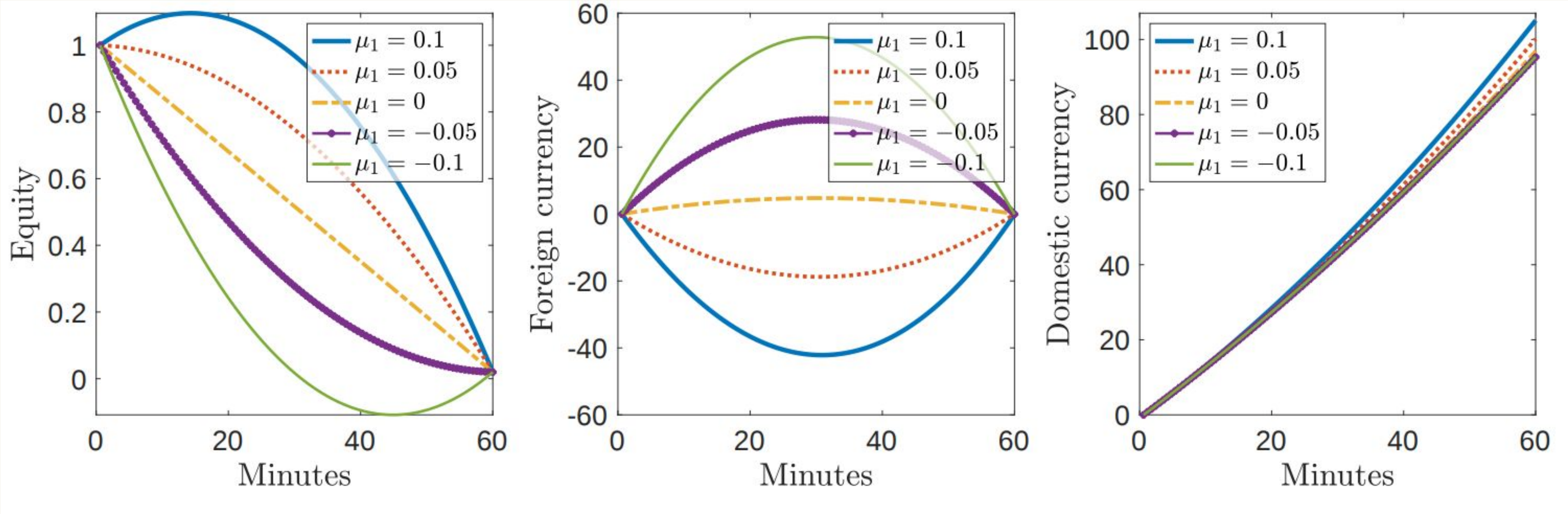
A rough path between mathematics and data science

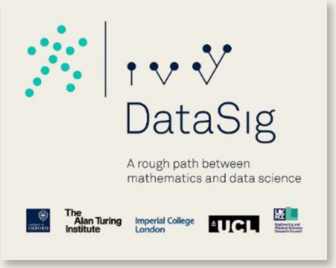
Numerical experiment IV: double-execution

- We want to sell a block of shares in a **foreign** stock market (e.g. Apple).
- The proceeds are exchanged into the investor's **domestic currency** (e.g. GBP).
- The foreign stock S^1 and exchange rate S^2 follow geometric Brownian motions:

$$dS_t^1 = \mu_1 S_t^1 dt + \sigma_1 S_t^1 dW_t^1$$
$$dS_t^2 = \mu_2 S_t^2 dt + \sigma_2 S_t^2 dW_t^2$$

Numerical experiment IV: double-execution





Thank you

Kalsi, J., Lyons, T. and Arribas, I.P., 2020. **Optimal execution with rough path signatures**. *SIAM Journal on Financial Mathematics, 11*(2), pp.470-493.

Cartea, Á., Perez Arribas, I. and Sánchez-Betancourt, L., 2020. **Optimal Execution of Foreign Securities: A Double-Execution Problem with Signatures and Machine Learning**. *SSRN 3562251*.

DataSig

A rough path between mathematics and data science