Journal of Computational Finance **26**(4), 1–29
DOI: 10.21314/JCF.2023.003

Research Paper

An optimal control strategy for execution of large stock orders using long short-term memory networks

A. Papanicolaou,[1] H. Fu,[2] P. Krishnamurthy,[2] B. Healy[3] and F. Khorrami[2]

[1]Department of Mathematics, North Carolina State University, Box 8205, Raleigh, NC 27695-8205, USA; email: apapani@ncsu.edu

[2]Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, 2 MetroTech Center, Brooklyn, NY 11201, USA; emails: hf881@nyu.edu, prashanth.krishnamurthy@nyu.edu, khorrami@nyu.edu

[3]Institute of Finance and Technology, University College London, Gower Street, London WC1E 6BT, UK; email: b.healy@ucl.ac.uk

(Received February 9, 2022; revised January 23, 2023; accepted April 26, 2023)

ABSTRACT

We simulate the execution of a large stock order with real data and a general power law in the Almgren and Chriss model. The example we consider is the liquidation of a large position executed over the course of a single trading day in a limit order book. Transaction costs are incurred because large orders walk the order book (that is, they consume order book liquidity beyond the best bid/ask price). We model the order book with a power law that is proportional to trading volume, and thus transaction costs are inversely proportional to a power of the trading volume. We obtain a policy approximation by training a long short-term memory (LSTM) neural network to minimize the transaction costs accumulated when execution is carried out as a sequence of smaller suborders. Using historical Standard & Poor's 100 price and volume data, we evaluate our LSTM strategy relative to strategies based on the time-weighted average price (TWAP) and volume-weighted average price (VWAP).

Corresponding author: A. Papanicolaou

Print ISSN 1460-1559 | Online ISSN 1755-2850

For execution of a single stock, the input to the LSTM is the cross-section of data on all 100 stocks, including prices, volumes, TWAPs and VWAPs. By using this data cross-section, the LSTM should be able to exploit interstock codependence in volume and price movements, thereby reducing transaction costs for the day. Our tests on Standard & Poor's 100 data demonstrate that in fact this is so, as our LSTM strategy consistently outperforms TWAP- and VWAP-based strategies.

Keywords: price impact; order books; optimal execution; long short-term memory (LSTM) networks; trading volume.

1 INTRODUCTION

Institutional investors must consider transaction costs when trading large amounts of stock. For example, a multi-billion-dollar mutual fund may execute several large stock orders each month when rebalancing their stock holdings. In the US stock market, a large order would be to sell 1 million shares of a stock with a typical daily trading volume of 25 million shares. A naive strategy is to place a single, very large market sell order on the exchange. This single order will consume so much of the liquidity available in the limit order book that there will be an average price per share that is considerably lower than the best bid (if the order gets filled at all). A better strategy is to divide the trade into smaller suborders, which then get executed over the course of a fixed time period. In this paper, we cast this subdividing of large trades as an optimal control problem, and then we obtain an optimal policy for execution by training a long short-term memory (LSTM) neural network (Hochreiter and Schmidhuber 1997) to minimize losses to transaction costs.

Stock market liquidity is made available by market makers who submit limit orders at the different price ticks in the order book. Liquidity is consumed when a trader submits a market order to buy or sell. A market order that is not too big will get filled by limit orders at the best available prices. A large market buy (sell) order will "walk the order book" (that is, it will consume all liquidity at multiple ticks). Walking the order book results in an average price per share that is equal to the initial best ask (bid) plus (minus) a transaction cost. Statistical studies of order book data have shown that the depth to which a large order walks the order book is approximately a concave power law of the number of shares (Almgren *et al* 2005; Weber and Rosenow 2005). Simple calculation will show that subdividing a large order into a sequence of suborders will reduce these transaction costs, but optimal subdivision is more complicated because there are several (stochastic) variables to consider when designing a policy.

It is common practice to evaluate a policy in terms of the average price over the entire trade. Average prices to consider are the time-weighted average price (TWAP) and the volume-weighted average price (VWAP). In general, an optimal suborder policy will minimize the expected value of transaction costs. Strategies that aim to achieve TWAP or VWAP can be optimal for executing large orders (Cartea and Jaimungal 2016; Kato 2015).

Our approach to the design of an optimal execution policy is to consider the problem amid the uncertainty of real market data. Machine learning and deep neural networks are very good at learning policies directly from data without the assistance of models. For large-cap US stocks (eg, the constituents of the Standard & Poor's 100 (S&P 100)), there is plenty of data available upon which to train a network and conduct backtests. In this paper we perform these backtests, and conclude that neural networks can be effective for improvement of large-order execution strategies. Our implementation uses LSTM networks, which are a good candidate for constructing a policy function because

(a) they do not require a model for the distribution of the market (we can learn directly from the data),

(b) they can handle Markov or non-Markov states,

(c) they are well suited to learn in the episodic environment of single-day execution, and

(d) they can provide a policy with the required temporal dynamic.

Of particular relevance in this paper is the possible presence of interstock codependence in price and volume movements. The execution problem is posed for a single stock, but the input to the network includes information from all stocks in the data set, which allows the LSTM network to learn interstock dependencies that may provide better prediction of prices and volumes, thereby improving the execution prices achieved by the policy. Inputting such a large amount of data might be problematic for more parsimonious policy functions, but the LSTM is more than capable of handling this high dimensionality; indeed, from our results on S&P 100 stocks, it appears that LSTM does learn useful patterns in this data set.

1.1 Background and literature review

A prototypical model for optimal execution was introduced by Almgren and Chriss (2001). In short, they assume that liquidity consumed by a market order refills quickly in the time between scheduled suborders, and therefore a simple formulation of the problem will consider only the temporary impact on price. The speed of impacted-price reversion is studied by Huberman and Stanzl (2005), who discuss how faster reversion rates affect the aggressiveness of trade execution. Some models, eg, those using Hawkes processes, consider price impact with a slower reversion of impacted price (Alfonsi and Blanc 2016; Amaral and Papanicolaou 2019; Bacry *et al* 2015). Obizhaeva and Wang (2013) show that an optimal execution policy should start with a large order to disrupt the order book's supply and demand balance and then begin trading continuously as the order book refills.

Power laws have been observed in stock markets in the United Kingdom, China and the United States (Bouchaud *et al* 2002; Gould *et al* 2013; Gu *et al* 2008; Maskawa 2007; Potters and Bouchaud 2003; Zovko and Farmer 2002). Almgren *et al* (2005) estimated the power-law exponent to be around 0.67, which they obtained from a large data set from Citibank. Closely related to our paper are the work of Hendricks and Wilcox (2014), which uses a reinforcement approach to approximate the solution to the problem studied by Almgren and Chriss (2001), and the deep Q-Learning approach to optimal execution, studied by Ning *et al* (2021).

Recently, there has been some progress on the development of machine learning algorithms for order book modeling and price execution (Lin and Beling 2021; Nevmyvaka *et al* 2006; Zainal *et al* 2021; Zhang *et al* 2019). In addition, Schnaubelt (2022) implements reinforcement learning for optimal limit order placement in crypto markets.

1.2 Results

The main result in this paper is the improved execution strategies we find by using LSTM. We assume that limit order depth at each tick is proportional to volume and increases by a power law across ticks as we move farther from the best bid/ask. A major advantage of our LSTM approach is that the network's input includes the cross-section of market data, thereby utilizing any interstock codependence that may be present in volume or price changes. To evaluate the efficacy of our approach, we implement LSTM execution on historical minute-by-minute stock market data from January 2020 to July 2022. Our results indicate that, when executing a block trade of S&P 100 stocks, execution with a trained LSTM network can save between 1 and 2 basis points (bps) per stock on a given day compared with the TWAP and VWAP strategies.

2 ORDER BOOK MODEL AND OPTIMAL POLICIES

Let S_t and V_t denote the mid price and volume, respectively, of a stock at time t. Following the model described by Platania *et al* (2018) and Rogers and Singh (2010), the order book has a limit order distribution $\rho(t,s) \geqslant 0$, where the units of s are ticks relative to S_t. Ticks with limit sell orders correspond to $s > 1$, ticks with $s \in (0,1)$ correspond to limit buy orders and the mid price corresponds to $s = 1$. An order of a shares consumes liquidity up to a relative price $r_t(a)$ such that

$$a = \int_1^{r_t(a)} \rho(t, s)\,\mathrm{d}s. \tag{2.1}$$

A simple form for ρ has limit orders distributed continuously in s and proportionally to volume with the following power law:

$$\rho(t, s) = \frac{V_t}{\varepsilon}|s - 1|^{\beta}, \tag{2.2}$$

where $\beta \in [0, \infty)$, $\varepsilon > 0$ is a scaling parameter and V_t is the trading volume at time t. In this paper, we seek to optimize the execution of a large order over the course of a single trading day, in which case each V_t will be the total volume of trades that occurred in the tth minute. The impact function in (2.2) is similar to the power law considered by Almgren (2020a,b).

When the relative price $r_t(a)$ in (2.1) is computed with the distribution $\rho(t, s)$ from (2.2), we see a price that is a concave function of order size divided by volume:

$$r_t(a) = 1 + \mathrm{sgn}(a)\biggl(\frac{\varepsilon(\beta + 1)}{V_t}|a|\biggr)^{1/(\beta+1)}. \tag{2.3}$$

The quantity $S_t r_t(a)$ can be thought of as the impacted price. If the impacted price is assumed to be a linear function of a, then the implication is that $\beta = 0$, so that the order book has equal liquidity at all ticks. The prevailing conclusion in many empirical studies is that impacted price is a sublinear concave function (Almgren *et al* 2005; Bershova and Rakhlin 2013; Bouchaud 2010; Cont *et al* 2014), such as the square root function corresponding to the $\beta = 1$ case. Cases where β is less than zero are not considered because this would imply decreasing liquidity in successive ticks beyond the best bid/ask, which is rarely the case for liquid, large-cap stocks.

The transaction costs incurred by walking the order book, as described by (2.1)–(2.3), will be a convex function of trade size. The US dollar amount of trading loss due to the price impact is computed as follows:

$$\begin{aligned} \mathrm{loss}(t, a) &= S_t \biggl| \int_1^{r_t(a)} s\rho(t, s)\,\mathrm{d}s - a \biggr| \\ &= C_{\varepsilon,\beta} S_t (V_t)^{-1/(\beta+1)} |a|^{(\beta+2)/(\beta+1)}, \end{aligned} \tag{2.4}$$

where

$$C_{\varepsilon,\beta} = \frac{1}{\beta + 2}(\varepsilon(\beta + 1))^{(\beta+2)/(\beta+1)}.$$

From the convexity of (2.4) with respect to $|a|$ it is clear that very large orders should be divided into suborders to reduce transaction costs.

REMARK 2.1 (Cost of paying the spread) This order book model ignores the bid–ask spread. For liquid stocks the bid–ask spread is usually one tick, or equivalently 1¢. For such stocks the transaction costs for market orders filled at the best bid/ask can be proxied by US$0.005. This amounts to a flat fee for execution of an order and will remain constant for all the strategies that we test in this paper. Therefore, we omit the cost of the spread.

REMARK 2.2 (Exchange fees) Typically, there are exchange fees that may be proportional to the US dollar amount traded for smaller orders. In this paper we do not consider these fees because the problem we are considering is from the perspective of a large institutional investor, for whom fees are structured differently, usually decreasing percentagewise as the trade size increases.

REMARK 2.3 (Permanent impact) We do not consider permanent impact in this paper. The assumption is that we are trading in highly liquid stocks for which the order book is replenished very quickly after a suborder. It would certainly be interesting to consider execution with permanent price impact, but in this paper we focus our effort on finding policies that are able to optimize amid the stochasticity and uncertainty in real-life historical price and volume data.

2.1 Optimal execution policy

Let us work on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0,1,2,\dots}, \mathbb{P})$, where $\mathcal{F}_t$ denotes the σ-algebra representing all information known to us at time t. Assume that times $t = 0, 1, 2, 3, \dots$ are equally spaced. Let the initial inventory be a number of shares x, and consider the situation where this inventory needs to be completely liquidated by terminal time T; in the example we present, the times $t = 0$ and $t = T$ are the open and close of the trading day, respectively. Let X_t denote the number of remaining unexecuted shares at time t for $t = 1, 2, 3, \dots, T$. Initially, we have $X_0 = x$. An execution policy is a sequence of $\mathcal{F}_t$-adapted suborders a_t such that $X_t = X_{t-1} + a_{t-1}$ (that is, a_{t-1} is this execution policy's suborder placed at time $t-1$ and executed at time t); these suborders are chosen so that $X_T = 0$. Using the loss function given in (2.4), an optimal execution policy is the minimizer of the expected loss,

$$\min_a \mathbb{E} \sum_{t=1}^{T} \operatorname{loss}(t, a_{t-1}) \quad \text{such that } X_t = X_{t-1} + a_{t-1},\ X_T = 0,\ X_0 = x. \tag{2.5}$$

where the minimization is carried out over the family of $\mathcal{F}_t$-measurable policies a_t. In (2.5) we are allowing for broad generality of the processes (S_t, V_t), aside from them being well defined on probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0,1,2,\dots}, \mathbb{P})$, and also assuming that our trading does not affect them. Later on we will narrow the family of policies to those a_t that are given by an LSTM network, but still this narrowing of policies will permit broad generality in the distribution of (S_t, V_t), such as being non-Markovian and having nonlinear dependence on volume.

The loss function in (2.5) is a risk-neutral optimization in the sense that there is no penalty on variance or risk. The optimization in (2.5) results in a policy that is on the frontier constructed by Almgren and Chriss (2001), which is formed by minimizing implementation shortfall with a penalty on its variance. The objective in (2.5) is similar to the objective of Kato (2015), which is risk neutral and also volume dependent.

2.2 TWAP and VWAP strategies

Before we approach solving (2.5) with full generality in (S_t, V_t), we first discuss the two industry-standard benchmarks for large order execution (namely TWAP and VWAP) and how they are related to (2.5).

DEFINITION 2.4 (Time-weighted average price) The TWAP is

$$\bar{S}_T = \frac{1}{T}\sum_{t=1}^{T} S_t.$$

The TWAP is a target for some execution policies because it is often used as an average execution price benchmark for large orders. A common execution strategy is the so-called TWAP strategy, wherein the policy is to subdivide the order into equally sized deterministic suborders:

$$a_t = -\frac{x}{T} \quad \text{for } t = 0, 1, \dots, T-1. \tag{2.6}$$

The TWAP is easy to implement, as it is guaranteed to satisfy the terminal condition $X_T = 0$ and does not require any parameter estimation. However, the TWAP strategy ignores volume and any other pertinent information acquired during the trading period. Indeed, volume is often a concern when it comes to evaluating trades, which is why execution policies often target the VWAP (see Cartea and Jaimungal (2016) for more on VWAP targeting).

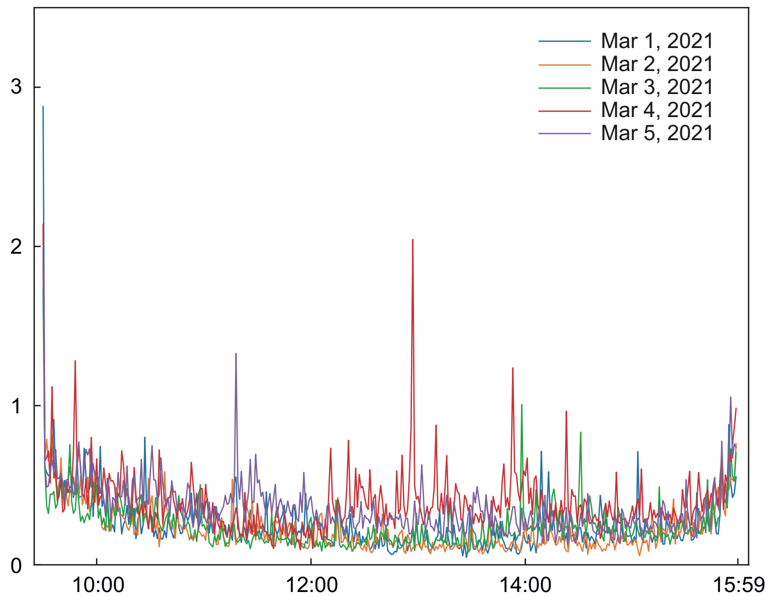
DEFINITION 2.5 (Volume-weighted average price) The VWAP is

$$\bar{S}_V = \frac{\sum_{t=1}^{T} V_t S_t}{\sum_{t=1}^{T} V_t}.$$

A VWAP strategy is to subdivide proportionally to moments of the volume:

$$a_t = -\frac{x\bar{V}_{t+1}}{\sum_{t=1}^{T} \bar{V}_t} \quad \text{for } t = 0, 1, 2 \dots, T-1, \tag{2.7}$$

FIGURE 1 The U-shaped pattern of daily volume for Apple Inc. (AAPL).



Source: data obtained from Reuters.com.

where $\bar{V}_t = (\mathbb{E}(V_t)^{-1/(\beta+1)})^{-(\beta+1)}$ (see Kato 2015). This VWAP strategy can be effective for single-day execution because volume follows a somewhat predictable U-shaped curve (see Figure 1). In practice, the building of a VWAP strategy requires some prior data to determine the typical evolution of volume over a trading period. For example, if we observe a history of volumes from past trading days, then we can use historically estimated $\bar{V}_t$ values in (2.7).

In general, the TWAP strategy (2.6) and the VWAP strategy (2.7) do not minimize the loss in (2.5). However, with some simplification of the volume process and minimal assumptions on S_t, we can show that VWAP is optimal if V_t is a deterministic function of t, and we can show that TWAP is optimal if V_t is deterministic and constant in t.

PROPOSITION 2.6 (Deterministic volume) *Suppose $V_t = \mathbb{E}V_t = \bar{V}_t$ for all t (that is, volume is a deterministic function of t). Also assume that S_t is a martingale with respect to the filtration $(\mathcal{F}_t)_{t=0,1,2,\dots}$. Then the VWAP strategy (2.7) is optimal, and if $\bar{V}_t$ is constant in t, then the TWAP strategy (2.6) is optimal.*

PROOF For a deterministic volume, and omitting constant $C_{\varepsilon,\beta}$, the optimization in (2.5) for stock i can be posed as

$$\min_a \mathbb{E}\bigg[\sum_{t=1}^{T} S_t(\bar{V}_t)^{-1/(\beta+1)}|a_{t-1}|^{(\beta+2)/(\beta+1)}\bigg]$$
$$\text{such that } X_t = X_{t-1} + a_{t-1}, \quad X_T = 0, \quad X_0 = x.$$

This optimization can be written as a Lagrangian,

$$\begin{aligned}
&\mathbb{E}\sum_{t=0}^{T-1}\Big(S_{t+1}(\bar{V}_{t+1})^{-1/(\beta+1)}|a_t|^{(\beta+2)/(\beta+1)} + \lambda_t(X_{t+1} - X_t - a_t)\Big) \\
&= \mathbb{E}\bigg[\sum_{t=0}^{T-1}\Big(S_{t+1}(\bar{V}_{t+1})^{-1/(\beta+1)}|a_t|^{(\beta+2)/(\beta+1)} - (\lambda_{t+1} - \lambda_t)X_{t+1} - \lambda_t a_t\Big) \\
&\qquad\qquad + \lambda_T X_T - \lambda_0 X_0\bigg],
\end{aligned}$$

with terminal condition $X_T = 0$, initial condition $X_0 = x$ and where λ_t is an $\mathcal{F}_t$-adapted Lagrange multiplier process. First-order conditions in a_t and X_{t+1} yield the following co-state equations:

$$\begin{aligned}
\frac{\beta+2}{\beta+1}\mathbb{E}_t\bigg[\operatorname{sgn}(a_t)S_{t+1}\bigg(\frac{|a_t|}{\bar{V}_{t+1}}\bigg)^{1/(\beta+1)}\bigg] - \lambda_t &= 0 \quad \text{for } t = 0, 1, 2 \dots, T-1, \\
\lambda_t - \mathbb{E}_t\lambda_{t+1} &= 0 \quad \text{for } t = 0, 1, 2 \dots, T-2,
\end{aligned}$$

with $\mathbb{E}_t$ denoting expectation conditional on $\mathcal{F}_t$. For $x > 0$ the optimal policy is the VWAP strategy,

$$a_t = -\frac{x\bar{V}_{t+1}}{\sum_{t=1}^{T}\bar{V}_t}, \qquad \lambda_t = -\frac{\beta+2}{\beta+1}\bigg(\frac{x}{\sum_{t=1}^{T}\bar{V}_t}\bigg)^{1/(\beta+1)}\mathbb{E}_t S_{t+1},$$

where the martingale property $S_t = \mathbb{E}_t S_{t+1}$ ensures that $\lambda_t = \mathbb{E}_t\lambda_{t+1}$. Further, this optimal policy is the TWAP strategy if $\bar{V}_t$ is constant in t. An analogous proof holds for $x < 0$. □

Kato (2015) presents results similar to Proposition 2.6 and also a theorem suggesting that, under certain Markovian assumptions, a deterministic VWAP is the optimal $\mathcal{F}_t$-adapted policy for stochastic volume. The following proposition shows how the VWAP strategy (2.7) can be optimal under certain assumptions on S_t and V_t.

PROPOSITION 2.7 (Stochastic volume) *Define*

$$M_t = \frac{(\mathbb{E}V_t^{-1/(\beta+1)})^{-1}}{V_t^{1/(\beta+1)}},$$

and assume M_t is a martingale with respect to the filtration $(\mathcal{F}_t)_{t=0,1,2,\dots}$. Also assume that S_t is a martingale with respect to the filtration $(\mathcal{F}_t)_{t=0,1,2,\dots}$, independent of M_t. Then the VWAP strategy (2.7) is optimal.

PROOF Taking the Lagrangian approach as we did in the proof of Proposition 2.6, we arrive at the following co-state equations:

$$\frac{\beta+2}{\beta+1}\mathbb{E}_t\biggl[\operatorname{sgn}(a_t)S_{t+1}\biggl(\frac{|a_t|}{V^i_{t+1}}\biggr)^{1/(\beta+1)}\biggr]-\lambda_t=0,$$
$$\lambda_t-\mathbb{E}_t\lambda_{t+1}=0.$$

For $x>0$ we make the ansatz that $a_t<0$, which gives us

$$\lambda_t=-\frac{\beta+2}{\beta+1}\mathbb{E}_t\biggl[S_{t+1}\biggl(\frac{|a_t|}{V_{t+1}}\biggr)^{1/(\beta+1)}\biggr],$$

and if we insert the VWAP strategy (2.7) we see that

$$\begin{aligned}
\mathbb{E}_t\biggl[S_{t+1}\biggl(\frac{|a_t|}{V_{t+1}}\biggr)^{1/(\beta+1)}\biggr]
&=S_t\mathbb{E}_t\biggl[\biggl(\frac{|a_t|}{V_{t+1}}\biggr)^{1/(\beta+1)}\biggr]\\
&=\biggl(\frac{x}{\mathcal{K}}\biggr)^{1/(\beta+1)}S_t\mathbb{E}_t\biggl[\biggl(\frac{(\mathbb{E}(V_{t+1})^{-1/(\beta+1)})^{-(\beta+1)}}{V_{t+1}}\biggr)^{1/(\beta+1)}\biggr]\\
&=\biggl(\frac{x}{\mathcal{K}}\biggr)^{1/(\beta+1)}S_t\mathbb{E}_tM_{t+1}\\
&=\biggl(\frac{x}{\mathcal{K}}\biggr)^{1/(\beta+1)}S_tM_t,
\end{aligned}$$

where $\mathcal{K}$ is the denominator of a_t in (2.7). Thus, when VWAP strategy (2.7) is used we have

$$\begin{aligned}
\lambda_t&=-\biggl(\frac{x}{\mathcal{K}}\biggr)^{1/(\beta+1)}\frac{\beta+2}{\beta+1}S_tM_t\\
&=-\biggl(\frac{x}{\mathcal{K}}\biggr)^{1/(\beta+1)}\frac{\beta+2}{\beta+1}\mathbb{E}_tS_{t+1}M_{t+1}\\
&=\mathbb{E}_t\lambda_{t+1},
\end{aligned}$$

thereby confirming is an optimal policy. An analogous proof holds for $x<0$. □

An example of the martingale M_t in Proposition 2.7 is the lognormal volume $\log(V_{t+1}/V_t)=\mu_t+\sigma_tZ_{t+1}$ (see, for example, Kato 2015), where $(Z_t)_{t=1,2,\dots}$

is a sequence of standard normals independent of the past and where μ_t and σ_t are deterministic functions of t calibrated so that V_t adheres to a U-shape over the course of a trading day; it is straightforward to check if parameters μ_t and σ_t allow for M_t to be a martingale. At this point, however, rather than pursue the specification of the underlying stochastic processes, we instead choose to leave the data distribution unspecified and then proceed to train an LSTM to trade optimally based on historical observations of financial data.

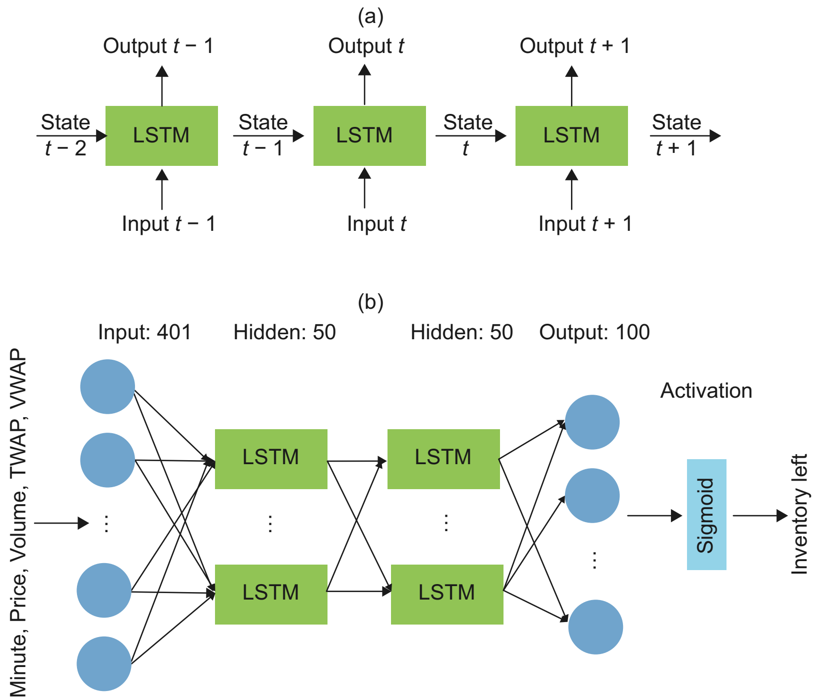
3 LONG SHORT-TERM MEMORY EXECUTION POLICY: EXPERIMENTAL SETUP

A policy approximation for the optimal solution of (2.5) can be obtained by training a neural network on historical market data. Our approach is to train an LSTM neural network to minimize the objective in (2.5), and then compare this strategy with the TWAP strategy (2.6) and the VWAP strategy (2.7). The choice of LSTM rather than a convolutional neural network or a recurrent neural network is based on the following two considerations. First, the problem in (2.5) is time dependent, requiring the neural network to memorize prior information. A convolutional neural network is static and thus cannot memorize prior information for use in backpropagation, whereas a recurrent neural network can memorize prior information but suffers from the gradient vanishing problem (Hochreiter 1998). The LSTM is able to handle both of these considerations.

The backtests we conduct will have a fixed number of suborders and a fixed submission time. We will submit suborders every 5 minutes, yielding a total of 78 suborders over the 390 minutes of the trading day. Each of the strategies we test (namely the TWAP, VWAP and LSTM strategies) will execute in the same five-minute intervals, thus ensuring a fair comparison. To be as realistic as possible, we also assume the LSTM lags one minute behind the real-time market (ie, the input of the LSTM network only includes the data up to and including the prior minute when the suborder is executed) to prevent the LSTM from having any foresight bias.

Figure 2(a) shows the structure of a single LSTM unit. We refer to the internal parameters of the LSTM unit as the state. At a particular time t, the LSTM unit updates its state to state_t using the old state, state_{t-1}, and the new input, input_t. The LSTM unit also generates an output, output_t. While state_t is then used to update state_{t+1}, output_t is used for other calculation purposes. Figure 2(b) shows the structure of the LSTM network used: it has two LSTM layers with 50 LSTM units in each. With fewer or smaller LSTM layers the network tends to underfit, whereas the performance of larger, more complex LSTM networks is comparable with that of our architecture but with increased computational complexity. The input has length 401, which comprises the current minute, the S&P 100 stocks' prices, their volumes and

FIGURE 2 The structure of (a) the LSTM unit and (b) the LSTM network used.



the remaining inventories in each stock under both TWAP and VWAP strategies. By including the current inventories remaining under TWAP and VWAP strategies in the input, the LSTM strategy performs no worse than TWAP or VWAP, as the LSTM can simply replicate them. The LSTM state also contains the current level of inventory in stock i, and after passing the results through a sigmoid activation function, the network outputs the updated inventory remaining in each stock i for the current time. The total number of network parameters is 116 100.

3.1 Data and parameter estimation

Our data was obtained from FirstRate Data and consists of minute-by-minute prices and volumes from January 2, 2020 through July 1, 2022, for the S&P 100 stocks as listed on March 21, 2022. The data was split into nine groups for training and testing the LSTM networks, as shown in Table 1. The S&P 100 contains the largest 100 companies in the US stock market by market capitalization. The limit order book for each of these stocks is extremely deep at all times, meaning that there

TABLE 1 The data sets and their sizes.

Fold	Training data	Days	Testing data	Days
1	Jan 2, 2020 to Mar 27, 2020	60	Mar 30, 2020 to Jun 3, 2020	45
2	Mar 30, 2020 to Jun 23, 2020	60	Jun 24, 2020 to Aug 27, 2020	45
3	Jun 24, 2020 to Sep 17, 2020	60	Sep 18, 2020 to Nov 20, 2020	45
4	Sep 18, 2020 to Dec 11, 2020	60	Dec 14, 2020 to Feb 19, 2021	45
5	Jan 4, 2021 to Mar 30, 2021	60	Mar 31, 2021 to Jun 4, 2021	45
6	Mar 31, 2021 to Jun 24, 2021	60	Jun 25, 2021 to Aug 30, 2021	45
7	Jun 25, 2021 to Sep 20, 2021	60	Jun 21, 2021 to Nov 23, 2022	45
8	Sep 21, 2021 to Dec 14, 2021	60	Dec 15, 2021 to Feb 18, 2022	45
9	Jan 3, 2022 to Mar 29, 2022	60	Mar 30, 2022 to Jun 3, 2022	45

is plenty of liquidity and it is very unlikely that a suborder will not get filled. For these liquid stocks, the simulated performance of the LSTM policy will be a realistic characterization of how it will perform in real-life trading.

For the power law in (2.2), we take $\beta = 0.67$ so that the impact in (2.3) has a power of 0.6, as suggested by Almgren *et al* (2005). However, we will conduct backtests with both constant β and stochastically fluctuating β, the latter being a more realistic description of real-life order books. We set ε so that the transaction cost equals 0.01–0.02% (ie, 1–2bps) of the value traded, which is realistic for S&P 100 stocks. For example, when trading 1 million shares, an appropriate ε would be 0.003. Then,

$$C_{\varepsilon,\beta} = \frac{1}{\beta + 2}(\varepsilon(\beta + 1))^{(\beta+2)/(\beta+1)} \approx 7.87 \times 10^{-5}. \tag{3.1}$$

For $i = 1, 2, \dots, 100$ let S_t^i and V_t^i denote the price and volume for the ith stock in the data set. The σ-algebra $\mathcal{F}_t$ is generated by $\bigcup_i \{(S_u^i, V_u^i)_{u=0,1,\dots,t}\}$. We separately train the neural network to execute optimally for each individual stock, but the inputs to the network include the vectors of all prices and volumes at time t, which we denote by $\boldsymbol{S}_t = (S_t^1, S_t^2, \dots, S_t^{100})$ and $\boldsymbol{V}_t = (V_t^1, V_t^2, \dots, V_t^{100})$, respectively.

3.2 Algorithm and LSTM training

Because suborders are executed every five minutes, the total number of executions during the trading day is $390/5 = 78$. The output of the LSTM is X_t^i, which represents the remaining inventory for stock i at time t. We train the LSTM network on each of the nine folds for each of the S&P 100 stocks. The total number of trained LSTM networks for all nine folds is $100 \times 9 = 900$. The loss function used to train the LSTM network for stock i is the following empirical approximation of (2.5):

$$\begin{aligned} L^i &= C_{\varepsilon,\beta} \sum_{\ell=1}^{78} S^i_{5\ell}(V^i_{5\ell})^{-1/(\beta+1)} |a^i_{5\ell-1}|^{(\beta+2)/(\beta+1)} \\ &= C_{\varepsilon,\beta} \sum_{\ell=1}^{78} S^i_{5\ell}(V^i_{5\ell})^{-1/(\beta+1)} |X^i_{5\ell} - X^i_{5\ell-5}|^{(\beta+2)/(\beta+1)}. \end{aligned} \tag{3.2}$$

Define the LSTM network weights as w. Then the training problem becomes

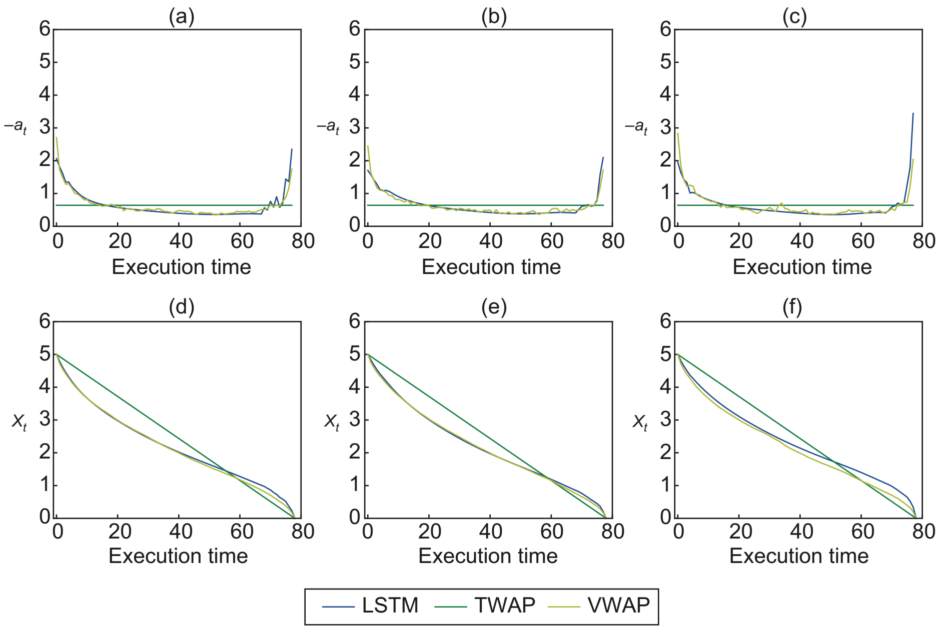
$$\begin{aligned} w^* &= \arg\min_w L^i(w) \\ &= \arg\min_w C_{\varepsilon,\beta} \sum_{\ell=1}^{78} S^i_{5\ell}(V^i_{5\ell})^{-1/(\beta+1)} |X^i_{5\ell}(w) - X^i_{5\ell-5}(w)|^{(\beta+2)/(\beta+1)}. \end{aligned}$$

ALGORITHM 3.1 (LSTM architecture training for optimal execution every 5 minutes in a 390-minute trading day for stock i)

–Initialize parameters of LSTM units w
for $k = 1$ to NUM_EPOCH **do**
  $X^i_0 = x^i_0$, $L^i = 0$, $t = 1$, $\beta = 0.67$, $\text{lr} = 0.001$, $h^i_0 = \text{None}$
  while $t < 390$ **do**
    $X^i_t, h^i_t = \text{LSTM}(h^i_{t-1}, t, \boldsymbol{S}_t, \boldsymbol{V}_t, \boldsymbol{X}^V_t, \boldsymbol{X}^T_t)$
    if $\text{mod}(t, 5) = 0$ **then**
      $L^i \mathrel{+}= C_{\varepsilon,\beta} S^i_t (V^i_t)^{-1/(\beta+1)} |X^i_t - X^i_{t-5}|^{(\beta+2)/(\beta+1)}$
    end if
    $t+ = 1$
  end while
  #### Close all the positions ####
  $X^i_{390} = 0$
  $L^i \mathrel{+}= C_{\varepsilon,\beta} S^i_{390} (V^i_{390})^{-1/(\beta+1)} |X^i_{390} - X^i_{385}|^{(\beta+2)/(\beta+1)}$
  #### Update the LSTM weights ####
  $w = \text{Adam}(L^i, lr, w)$
end for

Algorithm 3.1 shows the training procedure for stock i's LSTM network. For each training we loop for 10 000 epochs, which allows us to train a single LSTM in less than 20 minutes using graphics processing units (using the central processing unit, the runtime is around 10 hours). We initialize the LSTM state h^i as "None". We also initialize h^i with random numbers but observe little difference. Adam is used as the optimization algorithm, with a learning rate (lr) of 0.001. The initial inventory is 5% of the stock's average daily volume: that is, $x^i_0 = 0.05A^i$, where A^i is the sample mean of daily volume for stock i.

FIGURE 3 Graphs of $-a_t$ (in units of 10^4 shares) and X_t (in units of 10^5 shares) for selected stocks.

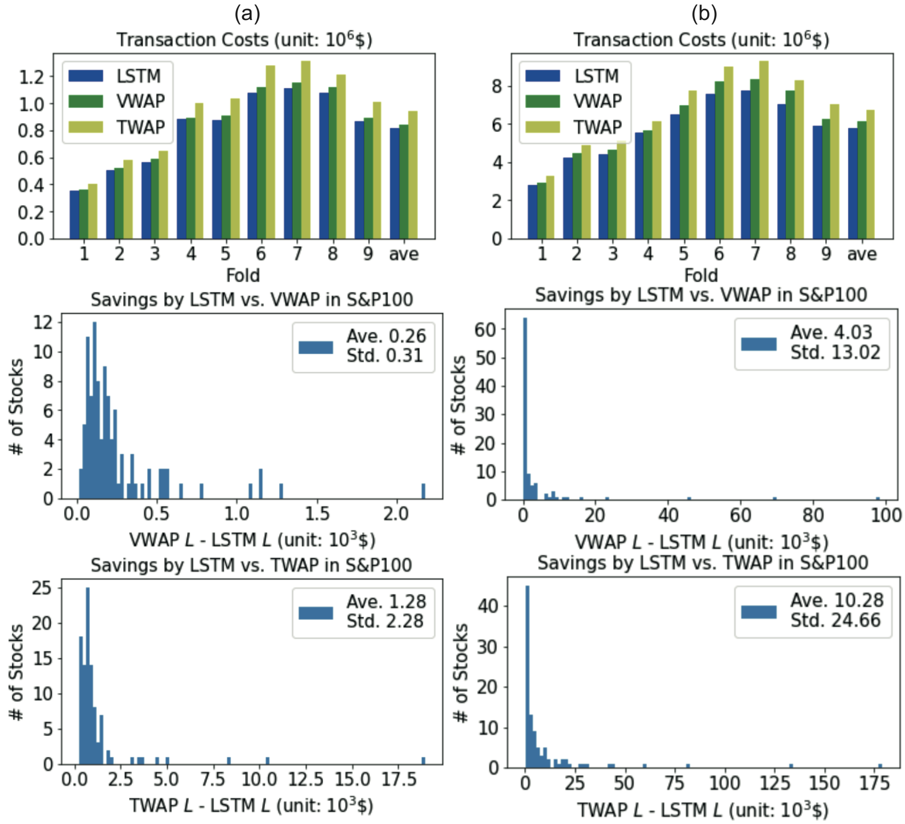


Parts (a) and (d) show MRNA. Parts (b) and (e) show JNJ. Parts (c) and (f) show NVDA.

We also train for a fixed amount of shares for each stock (ie, $x_0^i \equiv 10^6$). Note that, although time increments of 5 minutes are used, the data in between trade times is still seen by the LSTM, which means the strategy makes full use of the available information. The trading day has exactly 390 minutes, and so the first trade occurs at time $t = 5$, and the final trade occurs at minute 390 when the market closes. The LSTM networks are trained on nine folds, each comprising 60 days of one-minute data, as described in Table 1 for the S&P 100 stocks. We consider the days to be independent of each other. Therefore, the shape of the LSTM training input is $(60, 390, 401)$. The ratio of training data size $(60 \times 390 \times 401 = 9\,383\,400)$ to trained parameter size (116 100) is over 80. Therefore, over-fitting is highly unlikely to occur.

The $\boldsymbol{X}_t^V$ and $\boldsymbol{X}_t^T$ in Algorithm 3.1 represent the remaining inventories for the

FIGURE 4 Performance of LSTM in the noiseless order book case.



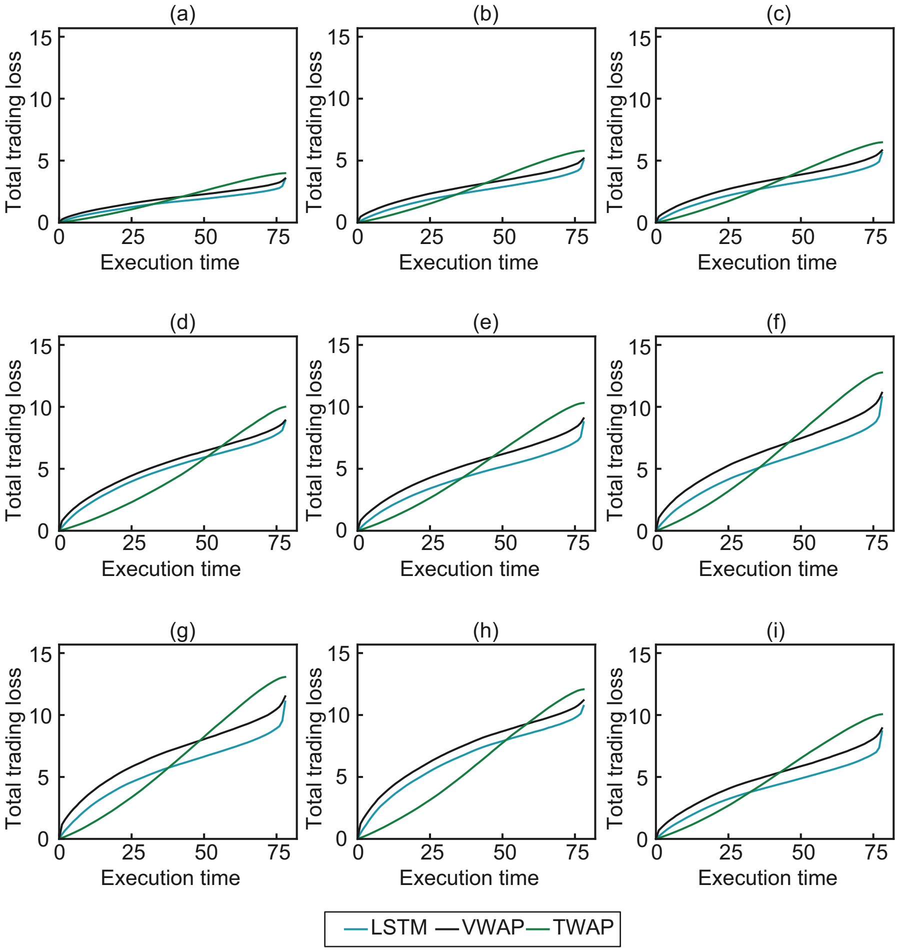
(a) $x_0 = 0.05A$. (b) $x_0 = 10^6$.

S&P 100 stocks by using VWAP and TWAP strategies at minute t; ie,

$$\left.\begin{aligned} X_t^T &= \left(x_0^1 - \frac{t}{T}x_0^1, \dots, x_0^{100} - \frac{t}{T}x_0^{100}\right), \\ X_t^V &= \left(x_0^1 - \frac{\sum_{j=1}^t \bar{V}_j^1}{\sum_{t=1}^T \bar{V}_t^1}x_0^1, \dots, x_0^{100} - \frac{\sum_{j=1}^t \bar{V}_j^{100}}{\sum_{t=1}^T \bar{V}_t^{100}}x_0^{100}\right). \end{aligned}\right\} \tag{3.3}$$

The TWAP strategy does not require the estimation of any parameters. As for the VWAP strategy, we used the 60 days' training volume data to estimate $\bar{V}_t^i$ for each fold and stock i as described in (2.7), which is $\bar{V}_t^i = (\mathbb{E}(V_t^i)^{-1/(\beta+1)})^{-(\beta+1)}$. As an example, Figure 3 shows the evolution of the actions and remaining inventories for

FIGURE 5 Transaction cost (in units of 10^5US$) with execution time for the noiseless order book case with $x_0 = 0.05A$.



(a) Fold 1. (b) Fold 2. (c) Fold 3. (d) Fold 4. (e) Fold 5. (f) Fold 6. (g) Fold 7. (h) Fold 8. (i) Fold 9.

the three strategies for three different stocks. During testing, the LSTM strategy is compared with the TWAP and VWAP strategies in (3.3).

TABLE 2 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the VWAP strategy for the noiseless order book case with $x_0 = 0.05A$ (5% of average volume) and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	VWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
AMZN	42.65	329.26	55.98 (1.70)	53.81 (1.63)	2.18 (0.07)
TSLA	28.20	909.89	162.77 (1.79)	161.49 (1.77)	1.29 (0.01)
MSFT	20.61	256.82	35.02 (1.36)	33.87 (1.32)	1.16 (0.04)
AAPL	87.88	570.61	77.39 (1.36)	76.24 (1.34)	1.14 (0.02)
GOOG	0.62	66.42	12.20 (1.84)	11.11 (1.67)	1.08 (0.16)
NFLX	3.85	93.59	18.52 (1.98)	17.73 (1.89)	0.78 (0.08)
UNH	2.07	39.53	6.82 (1.73)	6.17 (1.56)	0.65 (0.16)
TMO	0.93	22.87	4.33 (1.89)	3.75 (1.64)	0.57 (0.25)
NVDA	32.57	274.19	39.69 (1.45)	39.12 (1.43)	0.56 (0.02)
META	15.87	219.53	31.12 (1.42)	30.59 (1.39)	0.54 (0.02)
⋮	⋮	⋮	⋮	⋮	⋮
PM	3.24	13.50	2.09 (1.55)	1.93 (1.43)	0.17 (0.12)
⋮	⋮	⋮	⋮	⋮	⋮
AIG	4.07	8.88	1.43 (1.61)	1.37 (1.54)	0.06 (0.07)
DUK	1.93	8.90	1.33 (1.50)	1.27 (1.43)	0.06 (0.07)
COP	6.52	18.26	2.61 (1.43)	2.55 (1.40)	0.06 (0.03)
SO	3.12	9.28	1.32 (1.42)	1.27 (1.37)	0.05 (0.06)
CL	2.94	11.28	1.61 (1.43)	1.56 (1.38)	0.05 (0.05)
EXC	5.72	9.33	1.35 (1.45)	1.30 (1.40)	0.04 (0.05)
WBA	4.62	10.07	1.49 (1.48)	1.45 (1.44)	0.04 (0.04)
PG	4.69	31.84	4.29 (1.35)	4.25 (1.34)	0.04 (0.01)
DD	3.24	10.72	1.90 (1.77)	1.87 (1.75)	0.03 (0.02)
BK	3.88	8.69	1.29 (1.48)	1.27 (1.46)	0.01 (0.02)

4 LONG SHORT-TERM MEMORY EXECUTION POLICY: EXPERIMENTAL RESULTS

4.1 Evaluation metrics

The metric used to compare our LSTM strategy with the VWAP and TWAP strategies is the transaction cost L^i. As in (3.2), a smaller L^i means better performance of the selected execution strategy for stock i. We consider the following scenarios: the

TABLE 3 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the TWAP strategy for the noiseless order book case with $x_0 = 0.05A$ (5% of average volume) and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	TWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
TSLA	28.20	909.89	180.43 (1.98)	161.49 (1.77)	18.94 (0.21)
AMZN	42.65	329.26	64.19 (1.95)	53.81 (1.63)	10.38 (0.32)
AAPL	87.88	570.61	84.61 (1.48)	76.24 (1.34)	8.37 (0.15)
NVDA	32.57	274.19	44.06 (1.61)	39.12 (1.43)	4.94 (0.18)
MSFT	20.61	256.82	38.40 (1.50)	33.87 (1.32)	4.53 (0.18)
NFLX	3.85	93.59	21.47 (2.29)	17.73 (1.89)	3.73 (0.40)
META	15.87	219.53	34.09 (1.55)	30.59 (1.39)	3.50 (0.16)
BA	14.42	139.82	26.20 (1.87)	23.01 (1.65)	3.19 (0.23)
GOOG	0.62	66.42	13.14 (1.98)	11.11 (1.67)	2.02 (0.30)
PYPL	7.05	70.38	13.17 (1.87)	11.34 (1.61)	1.83 (0.26)
⋮	⋮	⋮	⋮	⋮	⋮
ORCL	8.61	30.58	5.21 (1.70)	4.50 (1.47)	0.71 (0.23)
⋮	⋮	⋮	⋮	⋮	⋮
AIG	4.07	8.88	1.71 (1.92)	1.37 (1.54)	0.34 (0.38)
EMR	2.12	8.54	1.61 (1.89)	1.28 (1.50)	0.33 (0.39)
MET	3.72	9.63	1.70 (1.76)	1.38 (1.44)	0.31 (0.32)
DUK	1.93	8.90	1.58 (1.78)	1.27 (1.43)	0.31 (0.35)
SO	3.12	9.28	1.57 (1.70)	1.27 (1.37)	0.31 (0.33)
CL	2.94	11.28	1.86 (1.65)	1.56 (1.38)	0.29 (0.26)
DOW	3.72	9.69	1.62 (1.67)	1.36 (1.41)	0.25 (0.26)
BKNG	0.22	22.54	3.96 (1.75)	3.70 (1.64)	0.25 (0.11)
WBA	4.62	10.07	1.70 (1.68)	1.45 (1.44)	0.25 (0.25)
BK	3.88	8.69	1.51 (1.73)	1.27 (1.46)	0.23 (0.27)

noiseless order book case with $x_0^i = 0.05A^i$ and a fixed amount of shares $x_0^i = 10^6$; and the noisy order book case (ie, β is stochastic) with $x_0^i = 0.05A^i$ and $x_0^i = 10^6$.[1]

4.2 Noiseless order book case

We test the trained LSTM models on the data sets shown in Table 1. The parameter ε in (3.1) is set to 0.006 for $x_0 = 0.05A$ and is set to 0.003 for $x_0 = 10^6$ so that

[1] For notational simplicity, from here onward we drop the superscript i except where it is necessary to show our results.

TABLE 4 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the VWAP strategy for the noiseless order book case with $x_0 = 10^6$ and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	VWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
GOOG	0.62	2150.54	1046.09 (4.86)	947.77 (4.41)	98.32 (0.46)
BKNG	0.22	2076.48	1741.56 (8.39)	1671.94 (8.05)	69.62 (0.34)
BLK	0.46	707.64	439.79 (6.21)	394.23 (5.57)	45.57 (0.64)
TMO	2.98	493.96	194.35 (3.93)	171.52 (3.47)	22.82 (0.46)
AVGO	1.08	435.77	149.75 (3.44)	133.44 (3.06)	16.32 (0.37)
CHTR	0.61	616.99	289.98 (4.70)	278.15 (4.51)	11.83 (0.19)
ADBE	1.43	494.80	140.58 (2.84)	129.16 (2.61)	11.42 (0.23)
LMT	0.96	365.99	126.62 (3.46)	117.34 (3.21)	9.28 (0.25)
UNH	2.89	381.23	84.43 (2.21)	75.98 (1.99)	8.45 (0.22)
ACN	1.17	276.23	81.30 (2.94)	72.95 (2.64)	8.35 (0.30)
⋮	⋮	⋮	⋮	⋮	⋮
ABT	4.05	111.71	14.13 (1.27)	13.58 (1.22)	0.55 (0.05)
⋮	⋮	⋮	⋮	⋮	⋮
BA	14.42	193.90	12.90 (0.67)	12.82 (0.66)	0.08 (0.00)
BK	3.88	44.78	5.83 (1.30)	5.75 (1.28)	0.08 (0.02)
GM	13.78	44.10	2.48 (0.56)	2.43 (0.55)	0.05 (0.01)
C	18.34	56.97	2.74 (0.48)	2.69 (0.47)	0.05 (0.01)
AAPL	87.88	129.86	2.40 (0.18)	2.36 (0.18)	0.04 (0.00)
WFC	14.04	37.81	1.46 (0.39)	1.43 (0.38)	0.03 (0.01)
XOM	27.13	53.12	2.25 (0.42)	2.23 (0.42)	0.03 (0.01)
PFE	25.04	40.06	1.59 (0.40)	1.57 (0.39)	0.02 (0.01)
BAC	47.20	34.29	0.95 (0.28)	0.93 (0.27)	0.02 (0.01)
F	70.81	11.81	0.26 (0.22)	0.26 (0.22)	0.00 (0.00)

the transaction cost is approximately 1–2bps of the traded equity values. The top row of Figure 4 shows the average daily transaction cost for each fold. On average, over the nine folds, the daily transaction cost is US$811 433 for the LSTM strategy, US$843 248 for the VWAP strategy and US$939 695 for the TWAP strategy for the $x_0 = 0.05A$ case. For the fixed initial shares case, the daily transaction cost is US$5 748 601 for the LSTM, US$6 135 958 for the VWAP and US$6 760 903 for the TWAP. Figure 5 shows how the daily transaction cost evolves with execution time for the $x_0 = 0.05A$ case. The LSTM strategy tends to incur large transaction costs toward the end of the trading session, whereas the VWAP strategy tends to

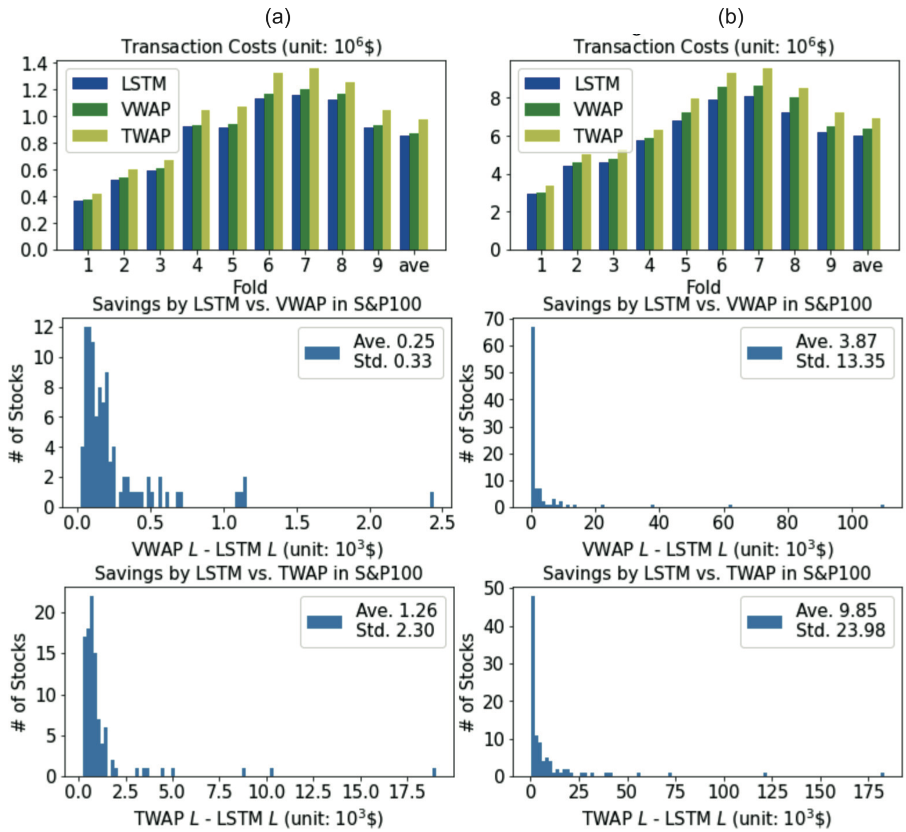
TABLE 5 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the TWAP strategy for the noiseless order book case with $x_0 = 10^6$ and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	TWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
GOOG	0.62	2150.54	1126.70 (5.24)	947.77 (4.41)	178.93 (0.83)
BKNG	0.22	2076.48	1805.40 (8.69)	1671.94 (8.05)	133.46 (0.64)
BLK	0.46	707.64	476.08 (6.73)	394.23 (5.57)	81.85 (1.16)
CHTR	0.61	616.99	337.32 (5.47)	278.15 (4.51)	59.17 (0.96)
AVGO	1.08	435.77	177.81 (4.08)	133.44 (3.06)	44.37 (1.02)
TMO	0.93	493.96	214.06 (4.33)	171.52 (3.47)	42.53 (0.86)
ADBE	1.43	494.80	161.14 (3.26)	129.16 (2.61)	31.97 (0.65)
LIN	1.07	270.08	109.40 (4.05)	79.80 (2.95)	29.59 (1.10)
LMT	0.96	365.99	144.76 (3.96)	117.34 (3.21)	27.41 (0.75)
COST	1.45	404.42	124.43 (3.08)	102.18 (2.53)	22.24 (0.55)
⋮	⋮	⋮	⋮	⋮	⋮
JNJ	4.49	156.07	19.59 (1.26)	17.23 (1.10)	2.36 (0.15)
⋮	⋮	⋮	⋮	⋮	⋮
C	18.34	56.97	3.08 (0.54)	2.69 (0.47)	0.39 (0.07)
INTC	22.63	51.50	2.50 (0.48)	2.11 (0.41)	0.38 (0.07)
GM	13.78	44.10	2.78 (0.63)	2.43 (0.55)	0.35 (0.08)
XOM	5.87	53.12	2.52 (0.47)	2.23 (0.42)	0.29 (0.05)
AAPL	87.88	129.86	2.62 (0.20)	2.36 (0.18)	0.26 (0.02)
PFE	25.04	40.06	1.82 (0.46)	1.57 (0.39)	0.25 (0.06)
WFC	4.62	37.81	1.65 (0.44)	1.43 (0.38)	0.22 (0.06)
BAC	47.20	34.29	1.08 (0.32)	0.93 (0.27)	0.15 (0.04)
T	45.78	19.94	0.61 (0.30)	0.55 (0.27)	0.06 (0.03)
F	70.81	11.81	0.30 (0.26)	0.26 (0.22)	0.04 (0.03)

incur large transaction costs in the morning and the TWAP strategy incurs transaction costs fairly consistently throughout the day. However, the LSTM has an overall lower transaction cost. Similar execution behavior for the three strategies is observed in the fixed initial shares case.

The second row in Figure 4 shows the transaction costs saved by using the LSTM strategy compared with the VWAP strategy for each stock over the nine testing folds. The LSTM strategy outperforms the VWAP by having a smaller transaction cost for all the stocks in $x_0 = 0.05A$ case and the average saving by LSTM is US$260 for each stock. As for the fixed initial shares case, the average savings become US$4030.

FIGURE 6 Performance of LSTM in the noisy order book case.



(a) $x_0 = 0.05A$. (b) $x_0 = 10^6$.

The third row shows the transaction costs saved by using the LSTM strategy compared with the TWAP strategy for each stock over the nine testing folds. Similarly, the LSTM is able to save US\$1280 in the percentage of daily volume case and US\$10 280 in the fixed initial shares case.

Tables 2 and 3 list the top 10 stock cases, for which the LSTM saves the most, and the bottom 10 stock cases, for which LSTM saves the least compared with the VWAP and TWAP strategies, respectively, for $x_0 = 0.05A$. Most of the stocks in which the LSTM saves the most are large capitalization tech companies. The median savings are approximately US\$170 if LSTM is used rather than the VWAP strategy, and US\$710 if LSTM is used rather than the TWAP strategy. Tables 4 and 5 list the

TABLE 6 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the VWAP strategy for the noisy order book case with $x_0 = 0.05A$ (5% of average volume) and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	VWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
AMZN	42.65	329.26	59.85 (1.82)	57.41 (1.74)	2.44 (0.07)
TSLA	0.93	909.89	171.62 (1.89)	170.47 (1.87)	1.15 (0.01)
GOOG	13.78	66.42	12.25 (1.84)	11.11 (1.67)	1.14 (0.17)
AAPL	87.88	570.61	83.39 (1.46)	82.29 (1.44)	1.11 (0.02)
MSFT	8.54	256.82	36.59 (1.42)	35.48 (1.38)	1.11 (0.04)
NFLX	6.03	93.59	18.77 (2.00)	18.05 (1.93)	0.71 (0.08)
UNH	28.20	39.53	6.92 (1.75)	6.24 (1.58)	0.68 (0.17)
NVDA	4.60	274.19	42.02 (1.53)	41.41 (1.51)	0.62 (0.02)
META	3.72	219.53	32.52 (1.48)	31.96 (1.46)	0.57 (0.03)
TMO	45.78	22.87	4.36 (1.91)	3.80 (1.66)	0.56 (0.25)
⋮	⋮	⋮	⋮	⋮	⋮
GD	1.88	7.79	1.45 (1.86)	1.29 (1.66)	0.16 (0.20)
⋮	⋮	⋮	⋮	⋮	⋮
KHC	10.87	8.63	1.46 (1.70)	1.41 (1.63)	0.06 (0.07)
TGT	2.96	28.09	4.82 (1.71)	4.76 (1.69)	0.06 (0.02)
MS	8.03	31.26	4.76 (1.52)	4.71 (1.51)	0.05 (0.02)
DIS	1.66	66.71	10.72 (1.61)	10.67 (1.60)	0.05 (0.01)
COP	2.31	18.26	2.67 (1.46)	2.62 (1.43)	0.05 (0.03)
AIG	4.07	8.88	1.46 (1.64)	1.40 (1.58)	0.05 (0.06)
DUK	3.72	8.90	1.35 (1.52)	1.30 (1.46)	0.05 (0.05)
MO	1.69	14.41	2.24 (1.56)	2.20 (1.52)	0.05 (0.03)
WBA	5.69	10.07	1.52 (1.51)	1.48 (1.47)	0.04 (0.04)
SO	5.93	9.28	1.34 (1.44)	1.30 (1.40)	0.04 (0.04)

top (bottom) 10 stocks in which the LSTM saves the most (least) compared with the VWAP and TWAP strategies, respectively, for $x_0 = 10^6$. A median saving of approximately US$550 is achieved by using LSTM rather than the VWAP strategy, and US$2360 by using LSTM rather than the TWAP strategy.

4.3 Noisy order book shape

The $\beta = 0.67$ estimated by Almgren *et al* (2005) is an average. In any given minute the limit order books may have a different power law near to but not equal to 0.67.

TABLE 7 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the TWAP strategy for the noisy order book case with $x_0 = 0.05A$ (5% of average volume) and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	TWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
TSLA	28.20	909.89	189.50 (2.08)	170.47 (1.87)	19.03 (0.21)
AMZN	42.65	329.26	67.78 (2.06)	57.41 (1.74)	10.37 (0.31)
AAPL	87.88	570.61	91.02 (1.60)	82.29 (1.44)	8.73 (0.15)
NVDA	32.57	274.19	46.33 (1.69)	41.41 (1.51)	4.92 (0.18)
MSFT	20.61	256.82	40.02 (1.56)	35.48 (1.38)	4.53 (0.18)
NFLX	3.85	93.59	21.74 (2.32)	18.05 (1.93)	3.69 (0.39)
META	15.87	219.53	35.39 (1.61)	31.96 (1.46)	3.43 (0.16)
BA	14.42	139.82	27.06 (1.94)	23.84 (1.70)	3.22 (0.23)
GOOG	0.62	66.42	13.09 (1.97)	11.11 (1.67)	1.97 (0.30)
BAC	47.20	80.93	13.73 (1.70)	11.88 (1.47)	1.85 (0.23)
⋮	⋮	⋮	⋮	⋮	⋮
FDX	1.88	21.55	4.20 (1.95)	3.51 (1.63)	0.69 (0.32)
⋮	⋮	⋮	⋮	⋮	⋮
COP	6.52	18.26	2.95 (1.62)	2.62 (1.43)	0.33 (0.18)
KHC	5.07	8.63	1.74 (2.01)	1.41 (1.63)	0.33 (0.39)
AIG	4.07	8.88	1.73 (1.95)	1.40 (1.58)	0.32 (0.36)
EMR	2.12	8.54	1.62 (1.90)	1.31 (1.53)	0.32 (0.37)
CL	2.94	11.28	1.87 (1.66)	1.57 (1.40)	0.30 (0.27)
MET	3.72	9.63	1.72 (1.78)	1.42 (1.47)	0.30 (0.31)
SO	3.12	9.28	1.59 (1.72)	1.30 (1.40)	0.29 (0.31)
DUK	1.93	8.90	1.59 (1.79)	1.30 (1.46)	0.29 (0.32)
DOW	3.72	9.69	1.64 (1.69)	1.39 (1.43)	0.25 (0.26)
WBA	4.62	10.07	1.72 (1.71)	1.48 (1.47)	0.24 (0.24)

To model this variation, we consider β to be stochastic: that is,

$$\beta_t = 0.67 + \eta_t, \tag{4.1}$$

where η_t is a random variable with a uniform distribution on $(-0.3, 0.3)$. The range of noise is $0.6/0.67 \approx 0.9$. The trading loss then becomes

$$L = \sum_{\ell=1}^{78} C_{\varepsilon,\beta_{5\ell}} S_{5\ell}(V_{5\ell})^{-1/(\beta_{5\ell}+1)} |X_{5\ell} - X_{5\ell-5}|^{(\beta_{5\ell}+2)/(\beta_{5\ell}+1)}, \tag{4.2}$$

where $C_{\varepsilon,\beta_{5\ell}}$ is the stochastic version of $C_{\varepsilon,\beta}$ obtained by substituting $\beta_{5\ell}$ into (3.1).

TABLE 8 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the VWAP strategy for the noisy order book case with $x_0 = 10^6$ and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	VWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
GOOG	0.62	2150.54	1084.43 (5.04)	974.45 (4.53)	109.98 (0.51)
BKNG	0.22	2076.48	1795.88 (8.65)	1734.05 (8.35)	61.83 (0.30)
BLK	0.46	707.64	454.80 (6.43)	417.02 (5.89)	37.78 (0.53)
TMO	45.78	493.96	201.56 (4.08)	179.50 (3.63)	22.06 (0.45)
AVGO	1.08	435.77	155.41 (3.57)	141.31 (3.24)	14.09 (0.32)
ADBE	1.43	494.80	145.13 (2.93)	133.28 (2.69)	11.85 (0.24)
CHTR	0.61	616.99	299.32 (4.85)	290.13 (4.70)	9.19 (0.15)
LMT	0.96	365.99	130.78 (3.57)	121.84 (3.33)	8.94 (0.24)
UNH	28.20	381.23	87.05 (2.28)	79.21 (2.08)	7.84 (0.21)
GD	0.90	173.48	70.89 (4.09)	63.23 (3.64)	7.65 (0.44)
⋮	⋮	⋮	⋮	⋮	⋮
CVS	5.27	77.55	8.99 (1.16)	8.59 (1.11)	0.40 (0.05)
⋮	⋮	⋮	⋮	⋮	⋮
BA	14.42	193.90	13.29 (0.69)	13.22 (0.68)	0.07 (0.00)
EXC	5.72	32.63	3.41 (1.05)	3.35 (1.03)	0.06 (0.02)
KHC	5.07	34.06	4.41 (1.30)	4.36 (1.28)	0.06 (0.02)
GM	13.78	44.10	2.55 (0.58)	2.50 (0.57)	0.05 (0.01)
AAPL	87.88	129.86	2.47 (0.19)	2.43 (0.19)	0.04 (0.00)
C	18.34	56.97	2.82 (0.50)	2.79 (0.49)	0.03 (0.01)
BK	3.88	44.78	6.02 (1.34)	5.99 (1.34)	0.03 (0.01)
WFC	5.69	37.81	1.51 (0.40)	1.48 (0.39)	0.03 (0.01)
XOM	4.62	53.12	2.32 (0.44)	2.29 (0.43)	0.02 (0.00)
BAC	47.20	34.29	0.99 (0.29)	0.97 (0.28)	0.02 (0.01)

The first row in Figure 6 shows the average daily transaction cost L for each fold in the noisy order book case. On average over the nine folds, the daily transaction cost is US$847 328 for the LSTM strategy, US$871 606 for the VWAP strategy and US$973 718 for the TWAP strategy for $x_0 = 0.05A$. As for $x_0 = 10^6$, the daily transaction cost is US$5 975 930 for the LSTM, US$6 344 318 for the VWAP, and US$6 946 579 for the TWAP. Compared with the noiseless order book shape case, the transaction costs increase for all the strategies. However, the LSTM strategy has a consistently smaller transaction cost than the VWAP and TWAP strategies. Note that the LSTM networks used are the same as in the noiseless order book shape

TABLE 9 The 10 stocks with highest savings and the 10 stocks with lowest savings under LSTM compared with the TWAP strategy for the noisy order book case with $x_0 = 10^6$ and the stock with the median saving.

Ticker↓	Daily volume (millions)	Equity value of traded (US$m)	TWAP (10^3US$) (bps)	LSTM (10^3US$) (bps)	Savings (10^3US$) (bps)
GOOG	0.62	2150.54	1157.61 (5.38)	974.45 (4.53)	183.16 (0.85)
BKNG	0.22	2076.48	1856.14 (8.94)	1734.05 (8.35)	122.09 (0.59)
BLK	0.46	707.64	488.79 (6.91)	417.02 (5.89)	71.77 (1.01)
CHTR	0.61	616.99	346.58 (5.62)	290.13 (4.70)	56.45 (0.91)
AVGO	1.08	435.77	182.79 (4.19)	141.31 (3.24)	41.47 (0.95)
TMO	0.93	493.96	219.81 (4.45)	179.50 (3.63)	40.31 (0.82)
ADBE	1.43	494.80	165.55 (3.35)	133.28 (2.69)	32.27 (0.65)
LIN	1.07	270.08	112.32 (4.16)	83.64 (3.10)	28.68 (1.06)
LMT	0.96	365.99	148.46 (4.06)	121.84 (3.33)	26.62 (0.73)
COST	1.45	404.42	127.93 (3.16)	107.02 (2.65)	20.91 (0.52)
⋮	⋮	⋮	⋮	⋮	⋮
QCOM	6.20	130.65	14.88 (1.14)	12.82 (0.98)	2.07 (0.16)
⋮	⋮	⋮	⋮	⋮	⋮
KO	12.37	51.75	3.78 (0.73)	3.25 (0.63)	0.53 (0.10)
VZ	5.69	52.80	3.35 (0.63)	2.93 (0.55)	0.42 (0.08)
C	18.34	56.97	3.16 (0.55)	2.79 (0.49)	0.37 (0.07)
INTC	22.63	51.50	2.56 (0.50)	2.19 (0.43)	0.37 (0.07)
GM	13.78	44.10	2.85 (0.65)	2.50 (0.57)	0.35 (0.08)
XOM	5.87	53.12	2.58 (0.49)	2.29 (0.43)	0.29 (0.05)
AAPL	87.88	129.86	2.69 (0.21)	2.43 (0.19)	0.26 (0.02)
PFE	25.04	40.06	1.87 (0.47)	1.63 (0.41)	0.24 (0.06)
WFC	4.62	37.81	1.70 (0.45)	1.48 (0.39)	0.22 (0.06)
BAC	47.20	34.29	1.11 (0.32)	0.97 (0.28)	0.14 (0.04)

case. Therefore, the LSTM strategy is robust to noise perturbations in β. We observe empirically that the difference in daily transaction cost between LSTM and VWAP decreases with the increase in noise intensity, as expected. The second row in Figure 6 shows the transaction costs saved by using the LSTM strategy rather than the VWAP and TWAP strategies for each stock: the average savings become smaller than in the noiseless order book case. Similarly, Tables 6 and 7 show the top (bottom) 10 stock cases in which the LSTM saves the most (least) transaction costs compared with the VWAP and TWAP strategies, respectively, for $x_0 = 0.05A$. The median savings are around US$160 for the VWAP case and US$690 for the TWAP case.

Tables 8 and 9 show the top 10 and bottom 10 stock cases for $x_0 = 10^6$. The median savings are around US$400 for the VWAP case and US$2070 for the TWAP case.

5 CONCLUSION

We showed how LSTM can be used for the optimal execution of large stock orders in a limit order book. Our backtests demonstrated that LSTM can outperform TWAP- and VWAP-based strategies in order book models with both a noiseless and a noisy power-law parameter. It is possible that the improved performance of the LSTM is due to its ability to aggregate across multiple stocks and to exploit codependence in the price and volume time series. There are a variety of future avenues for continuing this work. One such avenue is to include the permanent price impact and see how LSTM can adjust to early suborders adversely affecting the price. Another avenue would be to consider the optimal length of the trading period and the frequency of trading, both of which were static hyperparameters in this paper.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

This work was supported in part by National Science Foundation (NSF) grant DMS-1907518 and in part by the New York University Abu Dhabi (NYUAD) Center for Artificial Intelligence and Robotics, funded by Tamkeen under the NYUAD Research Institute Award CG010. We are grateful to the associate editor and anonymous referees, each of whom provided invaluable feedback on this work.

REFERENCES

Alfonsi, A., and Blanc, P. (2016). Dynamic optimal execution in a mixed-market-impact Hawkes price model. *Finance and Stochastics* **20**(1), 183–218 (https://doi.org/10.1007/s00780-015-0282-y).

Almgren, R. (2020a). Futures cost model. Technical Report, Quantitative Brokers, New York.

Almgren, R. (2020b). Transaction cost model in March 2020. Technical Report, Quantitative Brokers, New York.

Almgren, R., and Chriss, N. (2001). Optimal execution of portfolio transactions. *The Journal of Risk* **3**(2), 5–40 (https://doi.org/10.21314/JOR.2001.041).

Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk* **18**(7), 58–62.

Amaral, L., and Papanicolaou, A. (2019). Price impact of large orders using Hawkes processes. *ANZIAM Journal* **61**(2), 161–194 (https://doi.org/10.1017/S1446181119000038).

Bacry, E., Mastromatteo, I., and Muzy, J.-F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity* **1**(1), Paper 1550005 (https://doi.org/10.1142/S2382626615500057).

Bershova, N., and Rakhlin, D. (2013). The non-linear market impact of large trades: evidence from buy-side order flow. *Quantitative Finance* **13**(11), 1759–1778 (https://doi.org/10.1080/14697688.2013.861076).

Bouchaud, J.-P. (2010). Price impact. In *Encyclopedia of Quantitative Finance*, R. Cont (ed). Wiley (https://doi.org/10.1002/9780470061602.eqf18006).

Bouchaud, J.-P., Mézard, M., and Potters, M. (2002). Statistical properties of stock order books: empirical results and models. *Quantitative Finance* **2**(4), 251 (https://doi.org/10.1088/1469-7688/2/4/301).

Cartea, Á., and Jaimungal, S. (2016). A closed-form execution strategy to target volume weighted average price. *SIAM Journal on Financial Mathematics* **7**(1), 760–785 (https://doi.org/10.1137/16M1058406).

Cont, R., Kukanov, A., and Stoikov, S. (2014). The price impact of order book events. *Journal of Financial Econometrics* **12**(1), 47–88 (https://doi.org/10.1093/jjfinec/nbt003).

Gould, M. D., Porter, M. A., Williams, S., McDonald, M., Fenn, D. J., and Howison, S. D. (2013). Limit order books. *Quantitative Finance* **13**(11), 1709–1742 (https://doi.org/10.1080/14697688.2013.803148).

Gu, G.-F., Chen, W., and Zhou, W.-X. (2008). Empirical regularities of order placement in the chinese stock market. *Physica* A **387**(13), 3173–3182 (https://doi.org/10.1016/j.physa.2008.01.114).

Hendricks, D., and Wilcox, D. (2014). A reinforcement learning extension to the almgren-chriss framework for optimal trade execution. In *2014 IEEE Conference on Computational Intelligence for Financial Engineering and Economics*, pp. 457–464. IEEE Press, Piscataway, NJ (https://doi.org/10.1109/CIFEr.2014.6924109).

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **6**(2), 107–116 (https://doi.org/10.1142/S0218488598000094).

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation* **9**(8), 1735–1780 (https://doi.org/10.1162/neco.1997.9.8.1735).

Huberman, G., and Stanzl, W. (2005). Optimal liquidity trading. *Review of Finance* **9**(2), 165–200 (https://doi.org/10.1007/s10679-005-7591-5).

Kato, T. (2015). VWAP execution as an optimal strategy. *JSIAM Letters* **7**, 33–36 (https://doi.org/10.14495/jsiaml.7.33).

Lin, S., and Beling, P. A. (2021). An end-to-end optimal trade execution framework based on proximal policy optimization. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pp. 4548–4554. ACM Press, New York (https://doi.org/10.24963/ijcai.2020/627).

Maskawa, J. (2007). Correlation of coming limit price with order book in stock markets. *Physica* A **383**(1), 90–95 (https://doi.org/10.1016/j.physa.2007.04.091).

Nevmyvaka, Y., Feng, Y., and Kearns, M. (2006). Reinforcement learning for optimized trade execution. In *Proceedings of the 23rd International Conference on Machine Learning*, pp. 673–680. ACM Press, New York (https://doi.org/10.1145/1143844.1143929).

Ning, B., Lin, F. H. T., and Jaimungal, S. (2021). Double deep Q-learning for optimal execution. *Applied Mathematical Finance* **28**(4), 361–380 (https://doi.org/10.1080/1350486X.2022.2077783).

Obizhaeva, A. A., and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets* **16**(1), 1–32 (https://doi.org/10.1016/j.finmar.2012.09.001).

Platania, F., Serrano, P., and Tapia, M. (2018). Modelling the shape of the limit order book. *Quantitative Finance* **18**(9), 1575–1597 (https://doi.org/10.1080/14697688.2018.1433312).

Potters, M., and Bouchaud, J.-P. (2003). More statistical properties of order books and price impact. *Physica* A **324**(1), 133–140 (https://doi.org/10.1016/S0378-4371(02)01896-4).

Rogers, L. C., and Singh, S. (2010). The cost of illiquidity and its effects on hedging. *Mathematical Finance* **20**(4), 597–615 (https://doi.org/10.1111/j.1467-9965.2010.00413.x).

Schnaubelt, M. (2022). Deep reinforcement learning for the optimal placement of cryptocurrency limit orders. *European Journal of Operational Research* **296**(3), 993–1006 (https://doi.org/10.1016/j.ejor.2021.04.050).

Weber, P., and Rosenow, B. (2005). Order book approach to price impact. *Quantitative Finance* **5**(4), 357–364 (https://doi.org/10.1080/14697680500244411).

Zainal, M., Gad, I., and AlQaheri, H. (2021). An optimal limit order book prediction analysis based on deep learning and pigeon-inspired optimizer. *Journal of System and Management Sciences* **11**(3), 75–100.

Zhang, Z., Zohren, S., and Roberts, S. (2019). Deeplob: deep convolutional neural networks for limit order books. *IEEE Transactions on Signal Processing* **67**(11), 3001–3012 (https://doi.org/10.1109/TSP.2019.2907260).

Zovko, I., and Farmer, J. D. (2002). The power of patience: a behavioural regularity in limit-order placement. *Quantitative Finance* **2**(5), Paper 387 (https://doi.org/10.1088/1469-7688/2/5/308).