Optimal order placement in limit order markets

Rama Cont and Arseniy Kukanov
Laboratoire de Probabilités & Modèles Aléatoires, CNRS- Université de Paris VI
&
Columbia University, New York

To execute a trade, participants in electronic equity markets may choose to submit limit orders or market orders across various exchanges where a stock is traded. This decision is influenced by the characteristics of the order flow and queue sizes in each limit order book, as well as the structure of transaction fees and rebates across exchanges. We propose a quantitative framework for studying this *order placement* problem by formulating it as a convex optimization problem. This formulation allows to study how the interplay between the state of order books, the fee structure, order flow properties and preferences of a trader determine the optimal placement decision. In the case of a single exchange, we derive an explicit solution for the optimal split between limit and market orders. For the general problem of order placement across multiple exchanges, we propose a stochastic algorithm for computing the optimal policy and study the sensitivity of the solution to various parameters using a numerical implementation of the algorithm.

Key words: limit order markets, optimal order execution, execution risk, order routing, fragmented markets, transaction costs, financial engineering, stochastic approximation, Robbins-Monro algorithm

Contents

1. Introduction

In todays' automated, electronic financial markets, the trading process is divided into several stages, each taking place on a different time horizon: portfolio allocation decisions are usually made on a monthly or daily basis and translate into trades that are executed over time intervals of several minutes to several days. Existing studies on optimal trade execution (Bertsimas and Lo 1998, Almgren and Chriss 2000) have investigated how the execution cost of a large trade may be reduced by splitting it into multiple *orders* spread in time. Once this *order scheduling* decision is taken, one still needs to specify how each individual order should be *placed*: this order placement decision involves the choice of an *order type* (limit order, market order), order size and destination, when multiple trading venues are available. Orders are filled over short time intervals of a few milliseconds to several minutes and the mechanism through which orders are filled in the limit order book are relevant for such order placement decisions. Market participants need to make such decisions thousands of times each day, and their outcomes have a large impact on each participant's transaction cost as well as on aggregate market dynamics.

Early work on optimal trade execution (Bertsimas and Lo 1998, Almgren and Chriss 2000) did not explicitly model the process whereby each order is filled, but more recent formulations have tried to incorporate some elements in this direction. In one stream of literature (see Obizhaeva and Wang (2005), Alfonsi et al. (2010), Predoiu et al. (2011)) a trader is restricted to using market orders whose execution costs are given by an idealized order book shape function. Another approach is to model the process through which an order is filled as a dynamic random process Cont (2011), Cont and De Larrard (2011) and thus formulate the optimal execution problem as a stochastic control problem: this formulation has been studied in various setting with limit orders (Bayraktar and Ludkovski (2011), Gueant and Lehalle (2012)) or limit and market orders (Guilbaud and Pham 2012, Huitema 2012) but its complexity makes it intractable unless restrictive assumptions are made on price and order book dynamics.

In the present work, we adopt a simpler, more tractable approach: assuming that the trade execution schedule has been specified, we focus on the task of filling each order. Decoupling the scheduling problems from the order placement problem leads to a more tractable approach which is closer to market practice and allows us to incorporate some realistic features which matter for order placement decisions, while conserving analytical tractability.

Individual order placement and order routing decisions play an important role in modern financial markets. Brokers are commonly obliged by law to deliver the best execution quality to their clients and empirical evidence confirms that a large percentage of market orders in the U.S. and Europe is sent to trading venues providing lower execution costs or smaller delays (Boehmer and Jennings

2007, Foucault and Menkveld 2008). Market orders gravitate towards exchanges with larger posted quote sizes and low fees, while limit orders are submitted to exchanges with high rebates and lower execution waiting times (see Moallemi et al. (2011)). These studies demonstrate how investors' aggregate order routing decisions have a significant influence on market dynamics, but a systematic study of the order routing problem from the investor's perspective is lacking. A reduced-form model for routing an infinitesimal limit order to a single destination is used by Moallemi et al. (2011), while Ganchev et al. (2010) and Laruelle et al. (2009) propose numerical algorithms to optimize order executions across multiple dark pools, where supply/demand is unobserved. To the best of our knowledge this paper is the first to provide a detailed treatment of investor's order placement decision in a multi-exchange market, unified with the market/limit order choice.

Our key contribution is a quantitative formulation of the order placement problem which takes into account multiple important factors - the size of an order to be executed, lengths of order queues across exchanges, statistical properties of order flows in these exchanges, trader's execution preferences, and the stucture of liquidity rebates across trading venues. Our problem formulation is tractable, intuitive and blends the aforementioned factors into an optimal allocation of limit orders and market orders across available trading venues. Order routing heuristics employed in practice commonly depend on past order fill rates at each exchange and are inherently backward-looking. In contrast, our approach is forward-looking - the optimal order allocation depends on current queue sizes and distributions of future trading volumes across exchanges. When only a single exchange is available for execution, this order placement problem reduces to the problem of choosing an optimal split between market orders and limit orders. We derive an explicit solution for this problem and analyze its sensitivity to the order size, the trader's urgency for filling the order and other factors. Similar results are also established in a case of two trading venues under some assumptions on order flow distributions. Finally, we propose a stochastic approximation method for solving the order placement problem in the general case and demonstrate its efficiency through examples. Our numerical examples demonstrate that the use of our optimal order placement method allows to substantially decreases trading costs with respect to various 'naive' order placement strategies.

An important aspect of our framework is to account for *execution risk*, through the incorporation of a penalty for under- or over-filling an order. This penalty is high for time-sensitive executions or when it is costly to catch up on the unfilled portion of the order. Although market orders are executed at a less favorable price, it becomes optimal to use them when execution risk is a primary concern. Optimal limit order sizes are strongly influenced by total quantities of orders queueing for execution at each exchange and by distributions of order outflows from these queues. For example, if at one of the exchanges the queue size is much smaller than the expected future order outflow, it is optimal to place a larger limit order there. Finally, the total order size to be filled plays an important role - limit orders are used predominantly to execute small order sizes and market orders are used for medium and large orders. The amount that can be realistically filled with a limit order at each exchange is naturally constrained by the corresponding queue size and order outflow distribution, so the share of market orders in the optimal allocation increases as the total order size increases. We find that the optimal order allocation almost always splits the total quantity among all available exchanges, suggesting that there is a benefit in having multiple markets.

Section 2 describes our formulation of the order placement problem and shows that it has a global optimum. In Section 3 we derive an optimal split between market and limit orders for a single exchange. Section 4 analyzes the general case of order placement on multiple trading venues. Section 5 presents a numerical algorithm for solving the order placement problem in a general case and our simulation results, and Section 6 concludes. All proofs are presented in the Appendix.

2. The order placement problem

Consider a trader who has a mandate to buy S shares of a stock within a (short) time interval $[0,T]$. The deadline T may be a fixed horizon (e.g. 1 minute) or a stopping time (triggered by market activity). To gain queue priority the trader may immediately submit K limit orders of sizes L_k to various exchanges $k=1,\ldots,K$ or submit one market order of size M. The trader's *order placement* decision is thus summarized by a vector $X \triangleq (M, L_1, \ldots, L_K) \in \mathbb{R}_+^{K+1}$ whose components are non-negative i.e. only buy orders are allowed. Our objective is to define a meaningful framework in which the trader may choose the various possibilities for this order placement decision.

We focus on limit order placement and execution and assume that a market order of size M can be filled immediately and with certainty[1]. Limit orders with quantities $(L_1, \ldots, L_K)$ join queues of $(Q_1, \ldots, Q_K)$ pre-existing limit orders at the best bids of K exchanges, where $Q_k \geq 0$. To simplify the notation, we make an assumption that all K available bid queues are lined up at the best bid price, but it is easily relaxed. Denote by $(x)_+ \triangleq \max(x,0)$. If L_k is constant within $[0,T]$, the amount purchased with a limit order on exchange k by time T is equal to $(\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+$, where $\xi_k \triangleq C_k + D_k$ is an order outflow from the front of k-th bid queue, consisting of $C_k \in [0, Q_k]$ cancelations of pre-existing orders from that queue and D_k trades with contra-side marketable orders reaching that queue. We specifically note that limit order fill amounts are random, and we allow for partial fills. The *total amount* $A(X,\xi)$ bought by the trader by time T with all of his orders is a function of the order allocation X and an overall bid queue outflow $\xi = (\xi_1, \ldots, \xi_K)$:

[1] This assumption is reasonable if S is small relative to the prevailing market depth. Under the assumption of immediate and certain market order execution it is easy to show that sending market orders to exchanges with high fees is always sub-optimal. We therefore consider a single exchange (with the smallest liquidity fee) for the purpose of sending a single market order.

$$A(X,\xi) = M + \sum_{k=1}^{K} \left((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+\right) \tag{1}$$

The total price of this purchase is divided into a benchmark cost paid regardless of trader's decisions, which may be computed using mid-quote price level, and an *execution cost* given by

$$(s+f)M - \sum_{k=1}^{K}(s+r_k)\left((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+\right), \tag{2}$$

where s is a half of the bid-ask spread at time 0, f is the lowest available liquidity fee and $r_k, k = 1,\dots,K$ are liquidity rebates for all exchanges. The trader can reduce the execution cost by sending more limit orders, but this leads to a risk of underfulfilling the target S because their fills are random. To capture this *execution risk* we include, in the objective function, a penalty for violations of target quantity in both directions:

$$\lambda_u \left(S - A(X,\xi)\right)_+ + \lambda_o \left(A(X,\xi) - S\right)_+, \tag{3}$$

where λ_u, λ_o are marginal penalties for, respectively underfulfilling or overfulfilling the execution target S. These penalties are motivated by a correlation that exists between limit order executions and price movements (so-called adverse selection). If $A(X,\xi) < S$, the trader has to purchase the remaining $S - A(X,\xi)$ shares at time T with market orders. Adverse selection implies that conditionally on the event $\{A(X,\xi) < S\}$ prices have likely moved up and the transaction cost of market orders at time T is higher than their cost at time 0, i.e. $\lambda_u > s + f$. Alternatively, if $A(X,\xi) > S$ the trader experiences buyer's remorse - conditionally on this event prices have likely moved down and he could have achieved a better execution by being more patient. Besides adverse selection, parameters λ_u, λ_o may reflect trader's execution preferences. For example a trader with a positive forecast of short-term returns may prefer to trade early with a market order and set a larger value for λ_u.

Problem 1 (Optimal order placement problem) *An optimal order placement is a vector* $X^* \in \mathbb{R}_+^{K+1}$ *solution of*

$$\min_{X \in \mathbb{R}_+^{K+1}} \mathbb{E}[v(X,\xi)] \tag{4}$$

where

$$v(X,\xi) := (s+f)M - \sum_{k=1}^{K}(s+r_k)\left((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+\right) + \lambda_u \left(S - A(X,\xi)\right)_+ + \lambda_o \left(A(X,\xi) - S\right)_+ \tag{5}$$

is the sum of the execution cost and penalty for execution risk.

We will denote $V(X) = E[v(X,\xi)]$. We begin by assuming certain economically reasonable restrictions on parameter values.

Assumptions

A1 $\lambda_u > 0, \lambda_o > 0$: the trader is penalized for under- and over-fulfilling the target size

A2 $\lambda_o > s + \max_k \{r_k\}$ and $\lambda_o > -(s+f)$: it is suboptimal to over-fulfill the target size S regardless of fees and rebates

A3 $\min_k \{r_k\} + s > 0$: possibly negative rebates do not eliminate price improvement from limit order execution

Proposition 1 below shows that it is not optimal to submit limit or market orders that are *a priori* too large or too small (larger than the target size S or whose sum is less than S). Proposition 2 guarantees the existence of an optimal solution.

Proposition 1 *Consider $\mathcal{C}$ - a compact convex subset of $\mathbb{R}_+^{K+1}$ defined by*

$$\mathcal{C} \triangleq \left\{ X \in \mathbb{R}_+^{K+1} \,\middle|\, 0 \le M \le S, \quad 0 \le L_k \le S - M, k = 1, \dots, K, \quad M + \sum_{k=1}^{K} L_k \ge S \right\}$$

Under assumptions A1-A3 for any $\tilde{X} \notin \mathcal{C}$, $\exists \tilde{X}' \in \mathcal{C}$ with $V(\tilde{X}') \le V(\tilde{X})$. Moreover, if $\min_k \{\mathbb{P}(\xi_k > Q_k + S)\} > 0$, the inequality is strict: $V(\tilde{X}') < V(\tilde{X})$.

The penalty function (3) implements a soft constraint for order sizes and effectively focuses the search for an optimal order allocation to the set $\mathcal{C}$. Specific economic or operational considerations could also motivate hard constraints, e.g. $M = 0$ or $\sum_{k=1}^{K} L_k = S$. Such constraints can be easily included in our framework but absent the aforementioned considerations we do not impose them here.

Proposition 2 *Under assumptions A1-A3, $V(X)$ is a convex function on $\mathbb{R}_+^{K+1}$, it is bounded below and has a global minimizer $X^\star \in \mathcal{C}$.*

3. Choice of order type: limit orders vs market orders

To highlight the tradeoff between limit and market order executions in our optimization setup, we first consider a case when the asset is traded on a single exchange, and the trader has to choose an optimal split between limit and market orders. Since $K = 1$, we suppress a subscript 1 throughout this section.

Proposition 3 (Single exchange: optimal split between limit and market orders) *Assume that ξ has a continuous distribution and (A1-A3) hold. Denote*

$$\underline{\lambda_u} \triangleq \frac{2s+f+r}{F(Q+S)} - (s+r), \quad \text{and} \qquad \overline{\lambda_u} \triangleq \frac{2s+f+r}{F(Q)} - (s+r).$$

If $\lambda_u \leq \underline{\lambda_u}$, the optimal allocation is $(M^\star, L^\star) = (0, S)$. If $\lambda_u \geq \overline{\lambda_u}$, the optimal allocation is $(M^\star, L^\star) = (S, 0)$. If $\lambda_u \in (\underline{\lambda_u}, \overline{\lambda_u})$, the optimal allocation is:

$$\begin{cases} M^\star = S - F^{-1}\left(\dfrac{2s+f+r}{\lambda_u + s + r}\right) + Q, \\ L^\star = F^{-1}\left(\dfrac{2s+f+r}{\lambda_u + s + r}\right) - Q, \end{cases} \tag{6}$$

where $F(\cdot)$ is a cumulative distribution function of the bid queue outflow ξ.

In the case of a single exchange, Proposition 1 implies that $M^\star + L^\star = S$, therefore there is no risk of overfullfilling the target size and λ_o does not affect the optimal solution. The trader is only concerned with the risk of underfulfilling the target quantity, and balances this risk with the fee, rebate and other market information. The parameter λ_u can be interpreted as trader's urgency to fill the orders, and higher values of λ_u lead to smaller limit order sizes, as illustrated on Figure 1. In contrast, the optimal market order size increases with λ_u.

The optimal split between market and limit orders depends on the ratio $\frac{2s+f+r}{\lambda_u+(s+r)}$ which balances marginal costs and savings from a market order. It also depends on the distribution F and the queue length Q - keeping all else constant, a trader would submit a larger limit order if its execution is more likely and vice versa. The optimal limit order size decreases with λ_u as order underfulfillments become more expensive and increases with f as market orders become more expensive. Another interesting feature is that $L^\star$ is fully determined by Q, F and pricing parameters s, r, f, λ_u, while $M^\star$ increases with S. As a consequence of this solution feature, as the order size S increases, a larger fraction $\frac{M^\star}{S}$ of that order is executed with a market order. The solution $(M^\star, L^\star)$ depends on the entire distribution of ξ and not just on its mean, as illustrated on Figure 1 for a pair of exponential and Pareto distributions with equal means.

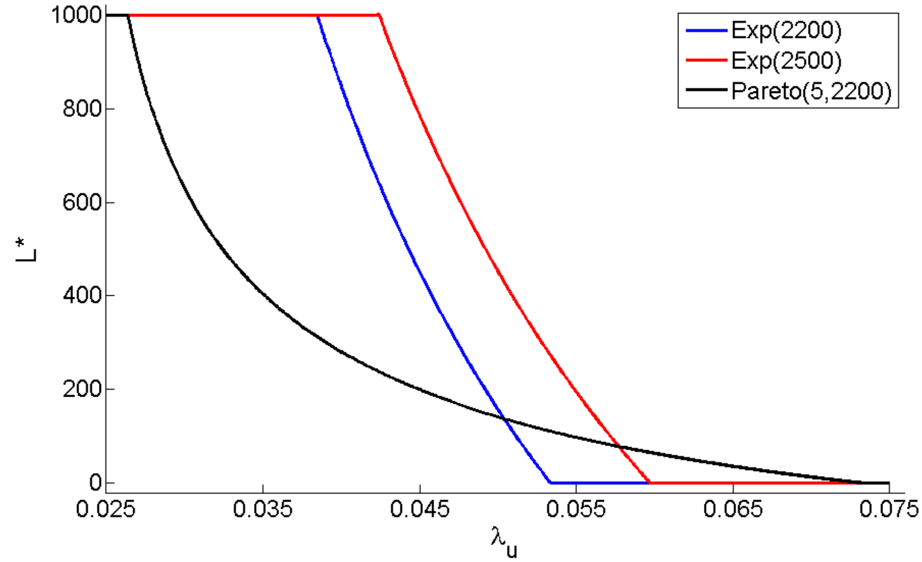

Figure 1 Optimal limit order size $L^{\star}$ for one exchange. The parameters for this figure are: $Q = 2000, S = 1000, s = 0.02, r = 0.002, f = 0.003$. Colors correspond to different order outflow distributions - exponential with means 2200 and 2500 and Pareto with mean 2200 and a tail index 5.

4. Optimal routing of limit orders across multiple exchanges

When multiple trading venues are available, dividing the target quantity among them provides better execution quality by reducing the risk of not filling the order. However, sending too many orders leads to an undesireable possibility of overfulfilling the target size. Proposition 4 gives a criterion for optimality of an order allocation $X^{\star} = (M^{\star}, L_1^{\star}, \ldots, L_K^{\star})$ that balances these risks.

Proposition 4 *Assume (A1-A3), also assume that the distribution of ξ is continuous, $\max_k \{F_k(Q_k + S)\} < 1$ and $\lambda_u < \max_k \left\{ \frac{2s + f + r_k}{F_k(Q_k)} - (s + r_k) \right\}$. Then:*

1. *It is optimal to submit a market order $M^{\star} > 0$ if $\lambda_u \geq \frac{2s+f+\max_k\{r_k\}}{\mathbb{P}\left(\bigcap_k \{\xi_k \leq Q_k\}\right)} - (s + \max_k\{r_k\})$.*

2. *It is optimal to submit a limit order $L_j^{\star} > 0$ if $\mathbb{P}\left(\bigcap_{k \neq j} \{\xi_k \leq Q_k\} \middle| \xi_j > Q_j\right) > \frac{\lambda_o - (s+r_j)}{\lambda_u + \lambda_o}$.*

3. *If 1 and 2 hold for all exchanges $j = 1, \ldots, K$, a necessary and sufficient condition for optimality of an order allocation $X^{\star} \in \mathcal{C}$ is that it solves the following equations:*

$$\mathbb{P}\left(M^{\star} + \sum_{k=1}^{K} ((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k^{\star})_+) < S\right) = \frac{s + f + \lambda_o}{\lambda_u + \lambda_o} \tag{7}$$

$$\mathbb{P}\left(M^{\star} + \sum_{k=1}^{K} ((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k^{\star})_+) < S \middle| \xi_j > Q_j + L_j^{\star}\right) = \frac{\lambda_o - (s + r_j)}{\lambda_u + \lambda_o}, \quad j = 1, \ldots, K \tag{8}$$

Equations (7,8) show that an order allocation is optimal as long as it sets the probabilities of underfullfilling the target quantity equal to specific thresholds computed with pricing parameters. When the number of exchanges K is large, the probabilities in (7,8) are difficult to compute in closed-form. However, before turning to numerical procedures we investigate how these equations can be solved in a tractable case of two exchanges.

Corollary *Consider the case of two exchanges with ξ_1, ξ_2 independent with continuous distributions. If*

1. $\max_{k=1,2}\{F_k(Q_k+S)\} < 1,$
2. $\lambda_u < \max_{k=1,2}\left\{\frac{2s+f+r_k}{F_k(Q_k)} - (s+r_k)\right\},\ \lambda_u \geq \frac{2s+f+\max_{k=1,2}\{r_k\}}{F_1(Q_1)F_2(Q_2)} - (s+\max_{k=1,2}\{r_k\}),$ *and*
3. $F_1(Q_1) < 1 - \frac{s+r_2}{\lambda_o}, F_2(Q_2) < 1 - \frac{s+r_1}{\lambda_o}$

then there exists an optimal order allocation $X^\star = (M^\star, L_1^\star, L_2^\star) \in int\{\mathcal{C}\}$ and it verifies

$$L_1^\star = Q_2 + S - M^\star - F_2^{-1}\left(\frac{\lambda_o - (s+r_1)}{\lambda_u + \lambda_o}\right) \tag{9a}$$

$$L_2^\star = Q_1 + S - M^\star - F_1^{-1}\left(\frac{\lambda_o - (s+r_2)}{\lambda_u + \lambda_o}\right) \tag{9b}$$

$$\bar{F}_1(Q_1+L_1^\star)\bar{F}_2(Q_2+S-M^\star-L_1^\star) + \int_{Q_1+S-M^\star-L_2^\star}^{Q_1+L_1^\star} \bar{F}_2(Q_1+Q_2+S-M^\star-x_1)dF_1(x_1) = \frac{\lambda_u-(s+f)}{\lambda_u+\lambda_o}, \tag{9c}$$

where $F_1(\cdot), F_2(\cdot)$ are the cdf of ξ_1, ξ_2 respectively.

In this solution optimal limit order quantities $L_1^\star, L_2^\star$ are linear functions of an optimal market order quantity $M^\star$. When (9a,9b) are substituted into (9c) we obtain a (non-linear) equation for $M^\star$, which can be solved for a given distribution of (ξ_1, ξ_2).

Example If ξ_1, ξ_2 are exponentially distributed with means μ_1, μ_2 respectively, then an optimal order allocation is given by:

$$\begin{cases} M^\star = Q_1 + Q_2 + S - z \\ L_1^\star = z - Q_1 + \mu_2 \log\left(\frac{\lambda_u + s + r_1}{\lambda_u + \lambda_o}\right) \\ L_2^\star = z - Q_2 + \mu_1 \log\left(\frac{\lambda_u + s + r_2}{\lambda_u + \lambda_o}\right), \end{cases} \tag{10}$$

where z is a solution of a transcedental equation:

$$1 + \log\left(\frac{(\lambda_u+s+r_1)(\lambda_u+s+r_2)}{(\lambda_u+\lambda_o)^2}\right) + \frac{z}{\mu} = \frac{\lambda_u-(s+f)}{\lambda_u+\lambda_o}e^{\frac{z}{\mu}}\ ,\ \text{if } \mu_1=\mu_2=\mu \tag{11}$$

or

$$\frac{\mu_1}{\mu_1-\mu_2}e^{-\frac{z}{\mu_1}}\left(\frac{\lambda_u+s+r_1}{\lambda_u+\lambda_o}\right)^{\frac{\mu_1-\mu_2}{\mu_1}} + \frac{\mu_2}{\mu_2-\mu_1}e^{-\frac{z}{\mu_2}}\left(\frac{\lambda_u+s+r_2}{\lambda_u+\lambda_o}\right)^{\frac{\mu_2-\mu_1}{\mu_2}} = \frac{\lambda_u-(s+f)}{\lambda_u+\lambda_o}\ ,\ \text{if } \mu_1 \neq \mu_2 \tag{12}$$

Similarly to the case of one exchange, in this example an optimal market order size $M^\star$ is an increasing linear function of queue sizes Q_1, Q_2 and the target quantity S, while optimal limit order sizes $L_i^\star$ are decreasing functions of the corresponding queue sizes Q_i. As in the case of one exchange, the optimal limit order sizes do not depend on the target quantity, but the optimal market order size increases with it. In addition we note that each $L_i^\star$ depends on the order flow distribution on both exchanges through $\mu_{1,2}$ and z.

5. Numerical solution of the optimization problem

Computing the objective function in the order placement problem (4) or its gradient at any point requires calculating an expectation (a multidimensional integral) which, aside from special examples, is generally not analytically tractable. Stochastic approximation methods, developed specifically for problems where the objective function is an expectation, turn out to be very useful for this problem. We propose in this section a procedure for computing the optimal order placement policy which does not require specifying an order outflow distribution.

We begin by briefly discussing the stochastic approximation approach and the specific method used here. Consider an objective function $V(X) \triangleq \mathbb{E}[f(X,\xi)]$ to be minimized and denote by $g(X,\xi) \triangleq \nabla f(X,\xi)$ where the gradient is taken with respect to X. The Robbins and Monro (1951) stochastic approximation algorithm tackles the problem in the following way:

1: Choose X_0 and a sequence of 'step sizes' $\{\gamma_n\}$;
2: **for** $n = 1, \ldots, N$ **do**
3: Simulate an independent random variable ξ^n with distribution F
4: Set $X_n = X_{n-1} - \gamma_n g(X_{n-1}, \xi^n)$
5: **end**

This algorithm produces an estimate $\hat{X}^\star \triangleq X_N$, which converges to the optimal point $X^\star$ as $N \to \infty$ under some technical assumptions (see e.g. Kushner and Yin (2003)). The sequence of constants $\{\gamma_n\}$ affects the rate of convergence. This sensitivity can be overcome by using for example the robust stochastic approximation of Nemirovski et al. (2009) which follows the same steps as above with a constant step size $\gamma_n = \gamma$ and uses an average of iterates $\hat{X}^\star \triangleq \frac{1}{N}\sum_{n=1}^{N} X_n$ instead of X_N as an estimate of the optimal point. Under some weak assumptions this method achieves a performance bound $V(\hat{X}^\star) - V(X^\star) \leq \frac{DM}{\sqrt{N}}$ for a finite N, where $D = \max\limits_{X,X' \in \mathcal{C}} \|X - X'\|_2$ and $M = \sqrt{\max\limits_{X\in\mathcal{C}} \mathbb{E}\left[\|g(X,\xi)\|^2\right]}$ with $\|\cdot\|$ denoting L_2 norm[2]. Multiplying an optimal step size γ by a constant

[2] The method assumes that $\min\limits_{X\in\mathcal{X}}\{V(X)\}$ is sought, where $V(X)$ is a well-defined and finite-valued expectation for every $X \in \mathcal{X}$ and $\mathcal{X}$ is a non-empty bounded closed convex set. Moreover $V(X)$ needs to be continuous and convex on $\mathcal{X}$. The optimal step size is $\gamma = \frac{D}{\sqrt{N}M}$.

$\theta > 0$ leads to a performance bound of the same order of magnitude $\max\{\theta, \theta^{-1}\}\frac{DM}{\sqrt{N}}$, i.e. the method is "robust" to step size misspecifications. For our problem we can further bound

$$D \leq \sqrt{K}S \qquad \text{and} \qquad M \leq \left((s+f+\lambda_u+\lambda_o)^2 + \sum_{k=1}^{K}(s+r_k+\lambda_u+\lambda_o)^2\right)^{1/2}$$

We assume that on each iteration $X_n \in int\{\mathcal{C}\}$ - this is enforced by rescaling X_n when needed and does not affect the convergence - then the stochastic gradient in our problem is given by:

$$g(X_n,\xi) = \begin{pmatrix} s+f-\lambda_u 1_{\{A(X_n,\xi)<S\}} + \lambda_o 1_{\{A(X_n,\xi)>S\}} \\ -(s+r_1)1_{\{\xi_1>Q_1+L_{1,n}\}} - \lambda_u 1_{\{A(X_n,\xi)<S,\xi_1>Q_1+L_{1,n}\}} + \lambda_o 1_{\{A(X_n,\xi)>S,\xi_1>Q_1+L_{1,n}\}} \\ \cdots \\ -(s+r_K)1_{\{\xi_K>Q_K+L_{K,n}\}} - \lambda_u 1_{\{A(X_n,\xi)<S,\xi_K>Q_K+L_{K,n}\}} + \lambda_o 1_{\{A(X_n,\xi)>S,\xi_K>Q_K+L_{K,n}\}} \end{pmatrix}$$

Note that $g(X_n,\xi)$ depends on random variables ξ only through indicator functions, which have simple economic meaning. For example $1_{\{A(X_n,\xi)<S\}} = 1$ if the target quantity was underfulfilled on the n-th step and $1_{\{\xi_k>Q_k+L_{k,n}\}} = 1$ if there was an opportunity to execute a larger limit order at exchange k on that step. This leads to a simple interpretation of numerical iterations - at each step order sizes are increased or decreased depending on whether or not the target quantity was under- or overfullfilled and whether or not a larger limit order could be filled at a given exchange. If a model for ξ is available, one can use it to simulate ξ and compute a numerical solution that takes into account specific order flow assumptions (e.g. forecasts of future trading volume). Alternatively, one can use past order fill data to compute indicator functions in $g(X_n,\xi)$ and obtain a non-parametric numerical solution for the order placement problem (using the empirical distribution of past order fills instead of assuming a functional form for F).

We analyze the numerical stability and convergence of estimates $\hat{X}^\star$ by comparing them with an analytical solution in the case of one exchange. For this computation we use $Q = 2000$ shares, $\xi \sim Pois(\mu T), \mu = 2200$ shares per minute, $T = 1$ minute and $S = 1000$ shares. The pricing parameters (in dollars per share) are $s = 0.02, r = 0.002, f = 0.003$ and fall in a typical value range for US equities. Finally, the penalty costs (also in dollars per share) are set to $\lambda_o = 0.024, \lambda_u = 0.026$. According to (6) the optimal allocation $(M^\star, L^\star) = (728, 272)$ shares. Numerical estimates $\hat{X}^\star$ were computed for five initial points X_0 and different number of iterates N in the stochastic approximation, using a step size $\gamma \triangleq \sqrt{K}S\left(N(s+f+\lambda_u+\lambda_o)^2 + N\sum_{k=1}^{K}(s+r_k+\lambda_u+\lambda_o)^2\right)^{-1/2}$. For each choice of X_0 and N we simulated an additional $L = 1000$ observations of ξ to estimate the objective values $V(X)$ with sample averages $W(X) = \frac{1}{L}\sum_{i=1}^{L} v(X,\xi_i))$. Figures 2 and 3 show that estimates converge to $X^\star$ regardless of X_0. When $X_0 = X^\star$, estimates remain close to that point. Convergence is also quite fast - after as few as 50 samples the algorithm can be within 2%

of the optimal objective value. In the worst case of initial points on the boundary it can take a few thousand samples to converge. It is also worth noting that convergence in terms of the objective value occurs significantly faster than convergence in terms of the order allocation vector.

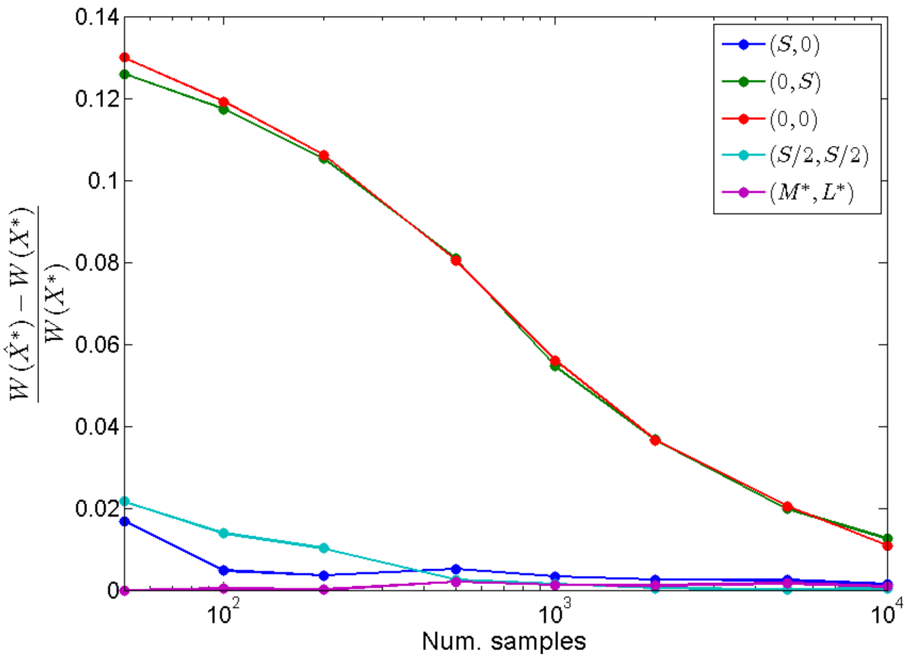

Figure 2 Convergence of objective values to an optimal point for different inital points.

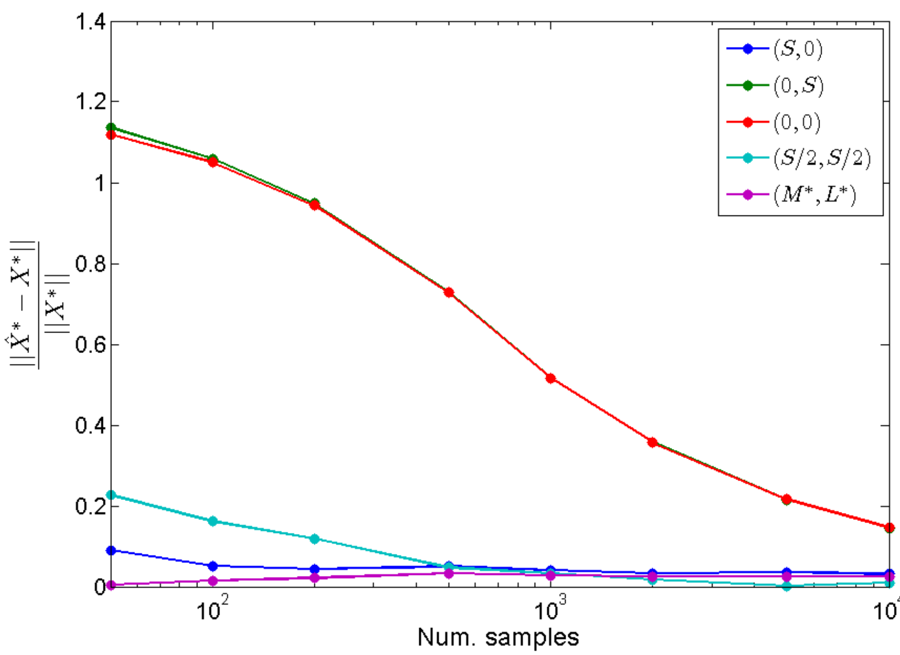

Figure 3 Convergence of order allocation vectors to an optimal point for different inital points.

K	$(\hat{M}^\star$	$\hat{L}_1^\star$	$\hat{L}_2^\star$	$\hat{L}_3^\star$	$\hat{L}_4^\star$	$\hat{L}_5^\star)/S$	$W(X_M)$	$W(X_L)$	$W(X_E)$	$W(\hat{X}^\star)$
					$\bar{S}=500$					
1	0.481	0.519					11.50	3.36	2.79	2.76
2	0.034	0.601	0.615				11.50	3.48	-2.86	-6.00
3	0.003	0.438	0.433	0.421			11.50	3.39	-5.22	-10.56
4	0.002	0.280	0.277	0.273	0.264		11.50	3.52	-6.44	-10.84
5	0.001	0.198	0.215	0.224	0.206	0.214	11.50	3.35	-7.24	-10.91
					$\bar{S}=1000$					
1	0.713	0.287					23.00	16.30	14.80	14.20
2	0.484	0.343	0.338				23.00	16.43	5.88	5.48
3	0.268	0.338	0.334	0.336			23.00	16.29	-3.16	-3.61
4	0.055	0.316	0.313	0.351	0.333		23.00	16.48	-9.27	-12.12
5	0.003	0.309	0.300	0.300	0.309	0.321	23.00	16.44	-12.88	-19.76
					$\bar{S}=5000$					
1	0.839	0.161					115.00	120.33	112.83	107.76
2	0.747	0.192	0.192				115.00	120.44	105.86	99.65
3	0.693	0.189	0.189	0.189			115.00	120.40	97.35	90.71
4	0.650	0.186	0.186	0.186	0.186		115.00	120.37	88.64	81.89
5	0.614	0.167	0.167	0.167	0.167	0.167	115.00	120.40	79.63	72.93

Table 1 Savings from order splitting.

We also estimated savings from optimal order placement and from dividing a limit order among multiple exchanges. Denote a pure market order allocation by $X_M=(S,0,\ldots,0)$, a single limit order allocation by $X_L=(0,S,0,\ldots,0)$ and an equal split allocation by $X_E=(\frac{S}{K+1},\frac{S}{K+1},\ldots,\frac{S}{K+1})$. Table 1 presents outputs from the numerical algorithm with $X_0=X_E, N=1000, L=1000$ for different order sizes S and number of exchanges $K=1,\ldots,5$. The parameters $s,f,r,\lambda_u,\lambda_o$ are same as in the previous simulation and exchanges are identical replicas of each other: $r_k=r, Q_k=Q$ and $\xi_{n,k}\sim Pois(\mu T)$ are i.i.d. copies of each other, where $k=1,\ldots,K,\quad n=1,\ldots,N$. Order allocations produced by stochastic approximation clearly outperform the naive benchmarks, especially when a target quantity S is relatively small and cost savings of limit orders can be fully captured. Comparing $W(X_L)$ and $W(X_E)$ we also see that splitting a limit order across multiple exchanges can be very advantageous when limit order fills are independent. Since multiple exchanges in this example are copies of each other, the algorithm splits the total limit order amount equally among them. This is not the same as the allocation X_E because the latter sets a market order size to $\frac{S}{K+1}$, which may be too large or too small depending on S and the number of available exchanges. Another interesting feature of the numerical solution $\hat{X}^\star$ is a tendency to oversize the total quantity of limit orders, which is clearly observed for $S=1000,5000$ and $K=4,5$. This may be a consequence of assumed independence between ξ_k - by submitting large orders to multiple exchanges the algorithm reduces the probability of underfullfilling the target quanitity with a relatively low probability of overfulfilling it.

To illustrate the structure of a numerical solution we performed a sensitivity analysis with $K=2$ exchanges and parameters $Q_1=Q_2=2000$, $S=1000$, $\xi_{1,2}\sim Pois(\mu_{1,2}T)$, $\mu_1=2600$, $\mu_2=2200$, $T=1$, $s=0.02$, $r_1=r_2=0.002$, $f=0.003$, $\lambda_u=0.26$ and $\lambda_o=0.24$. Varying some of these parameters one at a time we plot the numerical solution $\hat{X}^\star$ after $N=1000$ iterations, together with an analytical solution for a single exchange. The results are presented on Figures 4 and 5. Similarly to the single exchange case, limit order sizes on two exchanges L_1, L_2 decrease and market order size M increases as the penalty λ_u increases. Increasing the half-spread s, the rebate r_1 or the fee f makes a limit order on exchange number one more attractive, so L_1 increases and M decreases. Because the penalty λ_u is large in this example, execution risk is more important than fees and rebates, therefore the queue size Q_1 and the order outflow mean μ_1 have a much stronger effect on the optimal solution than r_1. Both decreasing the Q_1 and increasing μ_1 make a limit order fill more likely at exchange number one and L_1 increases[3]. Finally, as in the case of a single exchange, the target size S has a strong effect on the optimal order allocation. Only limit orders are used while S is small, but as it becomes larger it is difficult to fill that amount solely with limit orders and the optimal market order size begins to grow to limit the execution risk.

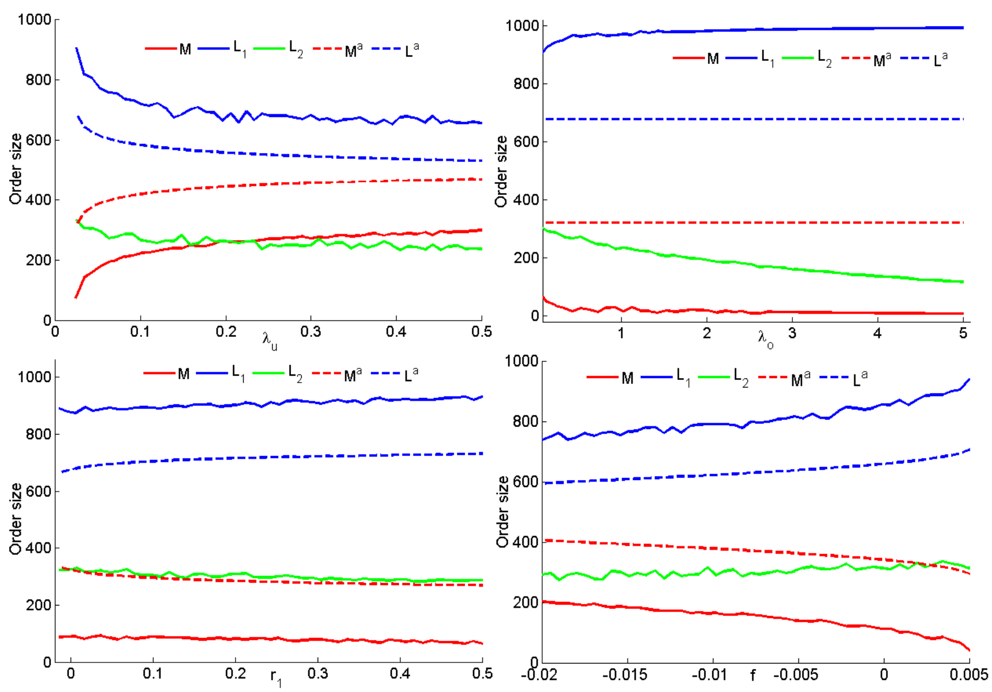

Figure 4 Sensitivity analysis for a numerical solution $\hat{X}^\star=(M,L_1,L_2)$ with two exchanges and an optimal solution (M^a,L^a) with the first exchange only.

[3] The observed drop in L_1 for large μ_1 and small Q_1 is a feature of this example, we were not able to replicate it for other distributions of ξ.

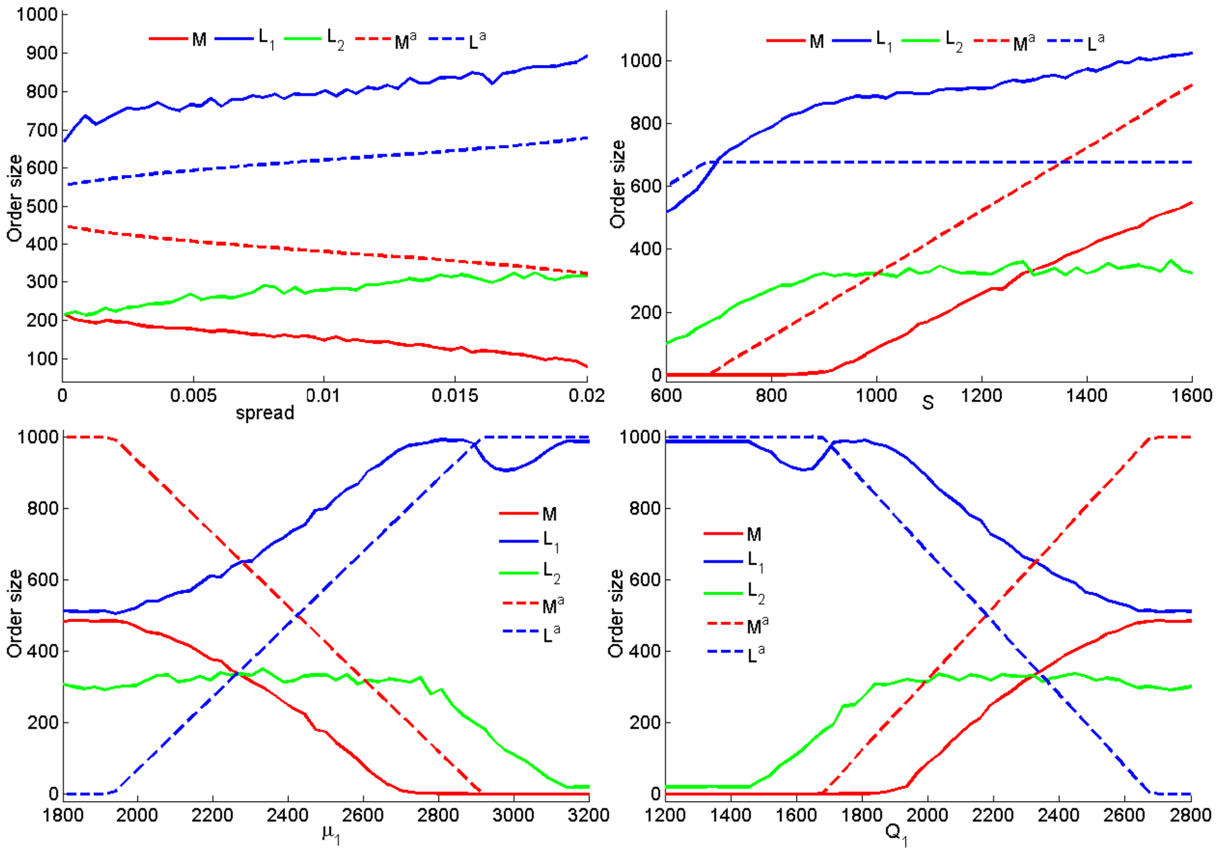

Figure 5 Sensitivity analysis for a numerical solution $\hat{X}^{\star} = (M, L_1, L_2)$ with two exchanges and an optimal solution (M^a, L^a) with the first exchange only.

6. Conclusion

We formulated and solved the problem of placing a small batch of orders on multiple trading venues. In the case when there is a single exchange we derived an optimal split between a limit and a market order sizes. For more general cases, we proposed and tested a stochastic approximation algorithm that is shown to quickly converge to an optimal point. We explored the properties of an optimal order allocation policy and showed that splitting an order across multiple exchanges can lead to a substantial reduction in transaction costs.

Appendix. Proofs

Note: to avoid extra notation in the following proofs we refer to the function $v(X,\xi)$ by $v(X)$.

Proposition 1 *Consider $\mathcal{C}$ - a compact convex subset of $\mathbb{R}_+^{K+1}$ defined by*

$$\mathcal{C} \triangleq \left\{ X \in \mathbb{R}_+^{K+1} \,\middle|\, 0 \le M \le S, \quad 0 \le L_k \le S - M, k = 1, \dots, K, \quad M + \sum_{k=1}^{K} L_k \ge S \right\}$$

Under assumptions A1-A3 for any $\tilde{X} \notin \mathcal{C}$, $\exists \tilde{X}' \in \mathcal{C}$ with $V(\tilde{X}') \le V(\tilde{X})$. Moreover, if $\min_k \{\mathbb{P}(\xi_k > Q_k + S)\} > 0$, the inequality is strict: $V(\tilde{X}') < V(\tilde{X})$.

Proof: First, for any allocation $\tilde{X}$ that has $\tilde{M} > S$, we automatically have $A(\tilde{X}) > S$ and we can show that the (random) cost and penalty of $\tilde{X}$ is larger than those of $X^{naive} \triangleq (S, 0, \dots, 0) \in \mathcal{C}$:

$$\begin{aligned} v(\tilde{X},\xi) - v(X^{naive},\xi) &= (s+f)(\tilde{M} - S) - \sum_{k=1}^{K} (s + r_k)((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+) + \\ &\quad \lambda_o \left(\tilde{M} - S + \sum_{k=1}^{K} ((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+) \right) = \\ &\quad (\lambda_o + s + f)(\tilde{M} - S) + \sum_{k=1}^{K} (\lambda_o - s - r_k)((\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+) > 0, \end{aligned}$$

which holds for all random ξ. Therefore, $V(\tilde{X}) > V(X^{naive})$. Similarly, for any allocation $\tilde{X}$ with $\tilde{L}_k > S - \tilde{M}$ define a different allocation $\tilde{X}'$ by $\tilde{M}' = \tilde{M}$, $\tilde{L}_j' = \tilde{L}_j, \forall j \ne k$ and $\tilde{L}_k' = S - \tilde{M}$. Then $v(\tilde{X},\xi) - v(\tilde{X}',\xi) = 0$ on the event $B = \{\omega | \xi_k(\omega) < Q_k + S - M\}$. On its complementary event B^c,

$$v(\tilde{X},\xi) - v(\tilde{X}',\xi) = -(s + r_k)((\xi_k - Q_k - S + \tilde{M})_+ - (\xi_k - Q_k - \tilde{L}_k)_+)$$

$$+ \lambda_o((\xi_k - Q_k - S + \tilde{M})_+ - (\xi_k - Q_k - \tilde{L}_k)_+).$$

Therefore

$$\begin{aligned} V(\tilde{X}) - V(\tilde{X}') &= \mathbb{E}\left[v(\tilde{X},\xi) - v(\tilde{X}',\xi) | B \right] \mathbb{P}(B) + \mathbb{E}\left[v(\tilde{X},\xi) - v(\tilde{X}',\xi) | B^c \right] \mathbb{P}(B^c) = \\ &\quad 0 + \mathbb{E}\left[(\lambda_o - (s + r_k))((\xi_k - Q_k - S + \tilde{M})_+ - (\xi_k - Q_k - \tilde{L}_k)_+) | B^c \right] \mathbb{P}(B^c) \ge 0 \end{aligned}$$

with a strict inequality if $\mathbb{P}(B^c) > 0$. If $\tilde{X}' \notin \mathcal{C}$, we can continue truncating limit order sizes $\tilde{L}_j' > S - \tilde{M}'$ following the same argument. Each time the truncation does not increase the objective function and finally we obtain $\tilde{X}'' \in \mathcal{C}$, such that $V(\tilde{X}'') \le V(\tilde{X})$.

Next, if $\tilde{X}$ is such that $\tilde{M} - \sum_{k=1}^{K} \tilde{L}_k < S$ define $s = S - \tilde{M} - \sum_{k=1}^{K} \tilde{L}_k$, take $\tilde{M}' = \tilde{M}, \tilde{L}_k' = \tilde{L}_k, k = 1, \dots, K-1$ and $\tilde{L}_K' = \tilde{L}_k + s$. Then, on the event $B = \left\{\omega | \xi_K(\omega) < Q_K + \tilde{L}_K\right\}$ we have $v(\tilde{X}, \xi) = v(\tilde{X}', \xi)$. However, on the event B^c,

$$v(\tilde{X},\xi) - v(\tilde{X}',\xi) = (s + r_K)((\xi_K - Q_K - \tilde{L}_K)_+ - (\xi_k - Q_k - \tilde{L}_K - s)_+) + \lambda_u((\xi_K - Q_K - \tilde{L}_K)_+ - (\xi_k - Q_k - \tilde{L}_K - s)_+),$$

therefore

$$\begin{aligned} V(\tilde{X}) - V(\tilde{X}') &= \mathbb{E}\left[v(\tilde{X},\xi) - v(\tilde{X}',\xi)|B\right]\mathbb{P}(B) + \mathbb{E}\left[v(\tilde{X},\xi) - v(\tilde{X}',\xi)|B^c\right]\mathbb{P}(B^c) = \\ & 0 + \mathbb{E}\left[(\lambda_u + (s + r_k))((\xi_K - Q_K - \tilde{L}_K)_+ - (\xi_k - Q_k - \tilde{L}_K - s)_+)|B^c\right]\mathbb{P}(B^c) \geq 0 \end{aligned}$$

with a strict inequality if $\mathbb{P}(B^c) > 0$. ■

Proposition 2 *Under assumptions A1-A3, $V(X)$ is a convex function on $\mathbb{R}_+^{K+1}$, bounded from below and admits a global minimizer $X^\star \in \mathcal{C}$.*

Proof: First, note that $(\xi_k - Q_k)_+ - (\xi_k - Q_k - L_k)_+$ are concave functions of L_k. Therefore, $A(X, \xi)$ is concave as a sum of concave functions. Similarly, the cost term in $v(X, \xi)$ is a sum of convex functions, as long as $r_k \geq -s, k = 1, \dots, K$ and is itself a convex function. Second, since $S - A(X, \xi)$ is a convex functon of X, and the function $h(x) \triangleq \lambda_u(x)_+ - \lambda_o(-x)_+$ is convex in x for positive λ_u, λ_o, so the penalty term $h(S - A(X, \xi)))$ is also convex.

If $\lambda_o > s + \max_k\{r_k\}$ the function $V(X)$ is also bounded from below since $v(X, \xi) \geq -(s + \max_k\{r_k\})S$.

Finally, since $V(X)$ is convex, it is also continous and reaches a local minimum V_{min} on the compact set $\mathcal{C}$ at some point $X^\star \in \mathcal{C}$. By convexity, V_{min} is a global minimum of $V(X)$ on $\mathcal{C}$. Moreover, since $\lambda_o > s + \max_k\{r_k\}$, Proposition 1 guarantees that $V_{min} < V(\tilde{X})$ for any $\tilde{X} \notin \mathcal{C}$, so V_{min} is also a global minimum of $V(X)$ on $\mathbb{R}_+^{K+1}$. ■

Proposition 3 *Consider the case of a single exchange where ξ has a continuous distribution. Denote by $\underline{\lambda_u} \triangleq \dfrac{2s+f+r}{F(Q+S)} - (s+r)$ and $\overline{\lambda_u} \triangleq \dfrac{2s+f+r}{F(Q)} - (s+r)$.*

Then, under Assumptions (A1-A3):

- *If $\lambda_u \leq \underline{\lambda_u}$, the optimal allocation is $(M^\star, L^\star) = (0, S)$. If $\lambda_u \geq \overline{\lambda_u}$, the optimal allocation is $(M^\star, L^\star) = (S, 0)$.*
- *If $\lambda_u \in (\underline{\lambda_u}, \overline{\lambda_u})$, the optimal allocation is given by (6).*

Proof: By Proposition 1 there exists an optimal split $(M^\star, L^\star) \in \mathcal{C}$ between limit and market orders. Moreover for $K = 1$ the set $\mathcal{C}$ reduces to a line $M^\star + L^\star = S$ so it is sufficient to find $M^\star$.

Restricting $L = S - M$ implies that $\{A(X,\xi) > S\} = \varnothing$, $\{A(X,\xi) < S, \xi > Q + L\} = \varnothing$, and we can rewrite the objective function as

$$V(M) = \mathbb{E}\Big[(s+f)M - (s+r)((\xi - Q)_+ - (\xi - Q - S + M)_+) \;\; + \\ \lambda_u\,(S - M - ((\xi - Q)_+ - (\xi - Q - S + M)_+))))_+\Big]. \quad (13)$$

For $M \in (0,S)$ the expression under the expectation in (13) is bounded for all ξ and differentiable with respect to M for almost all ξ, so we can compute $V'(M) = \frac{dV(M)}{dM}$ by interchanging the order of differentiation and integration (see e.g. Aliprantis and Burkinshaw (1998), Theorem 24.5):

$$V'(M) = \mathbb{E}\Big[s + f + (s+r)1_{\{\xi > Q+S-M\}} - \lambda_u 1_{\{\xi < Q+S-M\}}\Big] = 2s + f + r - (s + r + \lambda_u)F(Q + S - M) \quad (14)$$

Note that if $\lambda_u \le \dfrac{2s+f+r}{F(Q+S)} - (s+r)$, then $V'(M) \ge 0$ for $M \in (0,S)$ and therefore V is non-decreasing at these points. Checking that $V(S) - V(0) \ge (s + f - \lambda_u)S + (\lambda_u + s + r)S(1 - F(Q+S)) \ge 0$ we conclude that $M^\star = 0$. Similarly, if $\lambda_u \ge \dfrac{2s+f+r}{F(Q)} - (s+r)$, then $v(M) \le 0$ for all $M \in (0,S)$ and $V(M)$ is non-increasing at these points. Checking that $V(S) - V(0) \le (s + f - \lambda_u)S + (\lambda_u + s + r)S(1 - F(Q)) \le 0$ we conclude that $M^\star = S$. Finally, if λ_u is between these two values, $\exists \epsilon > 0$, such that $V'(\epsilon) < 0, V'(S - \epsilon) > 0$ and by continuity of V' there is a point where $V'(M^\star) = 0$. This $M^\star$ is optimal by convexity of $V(M)$ and (6) solves equations $v(M^\star, \xi) = 0, L^\star = S - M^\star$. ■

Proposition 4 *Assume (A1-A3), also assume that the distribution of ξ is continuous,* $\max_k \{F_k(Q_k + S)\} < 1$ *and* $\lambda_u < \max_k \left\{\dfrac{2s+f+r_k}{F_k(Q_k)} - (s + r_k)\right\}$. *Then:*

1. *It is optimal to submit a market order $M^\star > 0$ if* $\lambda_u \ge \dfrac{2s+f+\max_k\{r_k\}}{\mathbb{P}\left(\bigcap_k \{\xi_k \le Q_k\}\right)} - (s + \max_k\{r_k\})$.

2. *It is optimal to submit a limit order $L_j^\star > 0$ if* $\mathbb{P}\left(\bigcap_{k \ne j} \{\xi_k \le Q_k\} \,\middle|\, \xi_j > Q_j\right) > \frac{\lambda_o - (s + r_j)}{\lambda_u + \lambda_o}$.

3. *If 1 and 2 hold for all exchanges $j = 1, \ldots, K$, a necessary and sufficient condition for optimality of an order allocation $X^\star \in \mathcal{C}$ is that it solves equations* (7)–(8).

Proof: Proposition 2 implies the existence of an optimal order allocation $X^\star \in \mathcal{C}$. First, we define $X_M \triangleq (S, 0, \ldots, 0)$ and prove that $X^\star \ne X_M$ by contradiction. If X_M were optimal in the problem (4) it would also be optimal in the same problem with a constraint $L_k = 0, k \ne j$, for any one j. In other words, the solution $(S,0)$ would be optimal for any one-exchange problem, defined by using only exchange j. But by our assumption, there exists$J$ such that $\lambda_u < \dfrac{2s+f+r_J}{F_J(Q_J)} - (s + r_J)$ and Proposition 3 implies that $(S,0)$ is not optimal for the J-th single-exchange subproblem, leading to a contradiction.

The function $v(X,\xi)$ is bounded for $X \in \mathcal{C}$ and for all ξ, differentiable with respect to M and $L_k, k=1,\ldots,K$ for $X \in \mathcal{C}\setminus\{X_M\}$ for almost all ξ. Applying the same theorem as in the proof of Proposition 3 we conclude that $V(X)$ is differentiable for $X \in \mathcal{C}\setminus\{X_M\}$ and we can compute all of its partial derivatives by interchanging the order of differentiation and integration. The KKT conditions for problem (4) and $X \in \mathcal{C}\setminus\{X_M\}$ are

$$s+f-\lambda_u\mathbb{P}(A(X^\star,\xi)<S)+\lambda_o\mathbb{P}(A(X^\star,\xi)>S)-\mu_0=0 \tag{15}$$

$$\begin{aligned}&-(s+r_k)\mathbb{P}(\xi_k>Q_k+L_k^\star)-\lambda_u\mathbb{P}(A(X^\star,\xi)<S,\xi_k>Q_k+L_k^\star)+\\ &\qquad\lambda_o\mathbb{P}(A(X^\star,\xi)>S,\xi_k>Q_k+L_k^\star)-\mu_k=0,\quad k=1,\ldots,K\end{aligned} \tag{16}$$

$$M\geq 0,\quad L_k\geq 0,\quad \mu_0\geq 0,\quad \mu_k\geq 0,\quad \mu_0 M=0,\quad \mu_k L_k=0,\quad k=1,\ldots,K \tag{17}$$

Since the objective function $V(.)$ is convex, conditions (15)–(17) are both necessary and sufficient for optimality. The first result of this proposition follows from considering any $\tilde{X}$ with $\tilde{M}=0$: $V(\tilde{X}) \geq \lambda_u S\mathbb{P}\left(\bigcap_k\{\xi_k\leq Q_k\}\right) - (s+\max_k\{r_k\})S\mathbb{P}\left(\overline{\bigcap_k\{\xi_k\leq Q_k\}}\right) \geq (s+f)S = V(X_M)$ and we already argued that $\exists X^\star$ with $V(X^\star)<V(X_M)$, so $X^\star \neq \tilde{X}$ and therefore $M^\star>0$. Rearranging terms in a j-th equality (16) we obtain

$$\mathbb{P}(\xi_j>Q_j+L_j^\star)\left[\lambda_o-(s+r_j)-(\lambda_u+\lambda_o)\mathbb{P}(A(X^\star,\xi)<S|\xi_j>Q_j+L_j^\star)\right]-\mu_j=0 \tag{18}$$

The term in square brackets in (18) is negative for any $X \in \mathcal{C}\setminus\{X_M\}$ with $L_j=0$, because $\mathbb{P}(A(X,\xi)<S|\xi_j>Q_j+L_j) > \mathbb{P}\left(\bigcap_{k\neq j}\{\xi_k\leq Q_k\}\middle|\xi_j>Q_j\right) > \frac{\lambda_o-(s+r_j)}{\lambda_u+\lambda_o}$ by assumption and since $\mu_j\geq 0$ the condition (16) cannot be satisfied with $L_j^\star=0$. We showed that $M^\star>0, L_j^\star>0$ for all $j=1,\ldots,K$ and therefore, $\mu_0=\mu_1=\cdots=\mu_K=0$ by complimentary slackness. Then the KKT conditions (15)–(17) reduce to (7)–(8). ■

Corollary *Assume that two exchanges are available for execution, ξ_1 is independent of ξ_2 and the distribution of (ξ_1,ξ_2) is continuous. Also assume that:*

1. $\max\limits_{k=1,2}\{F_k(Q_k+S)\}<1$
2. $\lambda_u<\max\limits_{k=1,2}\left\{\dfrac{2s+f+r_k}{F_k(Q_k)}-(s+r_k)\right\}$, $\lambda_u\geq\dfrac{2s+f+\max\limits_{k=1,2}\{r_k\}}{F_1(Q_1)F_2(Q_2)}-(s+\max\limits_{k=1,2}\{r_k\})$
3. $F_1(Q_1)<1-\dfrac{s+r_2}{\lambda_o}, F_2(Q_2)<1-\dfrac{s+r_1}{\lambda_o}$

Then an optimal order allocation $X^\star=(M^\star,L_1^\star,L_2^\star)\in int\{\mathcal{C}\}$ and it solves the equations (9a-9c).

Proof: Solutions on the boundary of $\mathcal{C}$ are sub-optimal: $M^\star=0$ and $M^\star=S$ are ruled out by assumption 2, $L_1^\star=S-M$ and $L_2^\star=S-M$ are ruled out by assumption 3 and (16). Solutions with $M^\star+\sum\limits_{k=1}^{K}L_k^\star=S$ are ruled out by directly checking (16). Finally, $L_1^\star=0$ and $L_2^\star=0$ are also ruled out by (16). For example if $L_1^\star=0$, then by Proposition 1 $M^\star+L_2^\star=S$ and in (16) $\mu_2=0$ by complimentary slackness, $\mathbb{P}(A(X^\star,\xi)<S,\xi_2>Q_2+L_2^\star)=\mathbb{P}(A(X^\star,\xi)>S,\xi_2>Q_2+L_2^\star)=0$. But then (16) cannot hold because $\mathbb{P}(\xi_2>Q_2+L_2^\star)>0$.

For any $X\in int\{\mathcal{C}\}$, $A(X,\xi)>S$ if and only if all the following three inequalities are satisfied:

$$\xi_1>Q_1+S-M-L_2 \tag{19a}$$

$$\xi_2>Q_2+S-M-L_1 \tag{19b}$$

$$\xi_1+\xi_2>Q_1+Q_2+S-M \tag{19c}$$

These inequalities give a simple characterization of the event $\{A(X,\xi)>S\}$ and their equivalence is directly verified by considering subsets of (ξ_1,ξ_2) forming a complete partition of $\mathbb{R}_+^2$.

Case 1: $\xi_1>Q_1+L_1,\xi_2>Q_2+L_2$. Since $L_1+L_2+M>S$, we have $A(X,\xi)=L_1+L_2+M>S$ and at the same time all of the inequalities (19a-19c) are satisfied, so they are trivially equivalent in this case.

Case 2: $\xi_1>Q_1+L_1,Q_2\leq\xi_2\leq Q_2+L_2$. Because of the condition $\xi_1>Q_1+L_1$, (19a) is satisfied. We have in this case that $A(X,\xi)=L_1+\xi_2-Q_2+M$ and thus $A(X,\xi)>S$ if and only if (19b) is satisfied. Finally, $\xi_1>Q_1+L_1$ together with (19b) imply (19c), so $A(X,\xi)>S$ and (19a-19c) are equivalent in this case.

Case 3: $\xi_2>Q_2+L_2,Q_1\leq\xi_1\leq Q_1+L_1$. Similarly to Case 2 we can show that inequalities (19a-19c) are satisfied if and only if $A(X,\xi)>S$.

Case 4: $Q_1+S-M-L_2<\xi_1\leq Q_1+L_1,Q_2+S-M-L_1<\xi_2\leq Q_2+L_2$. This set is non-empty because $0<S-M-L_1<L_2$ and similarly for L_1,L_2 reversed. Inequalities (19a)–(19b) hold trivially, only (19c) needs to be checked. We can write $A(X,\xi)=\xi_1-Q_1+\xi_2-Q_2+M>S$ if and only if (19c) holds, so $A(X,\xi)>S$ is equivalent to (19a-19c).

Case 5: Outside of Cases 1-4, either (19a) or (19b) is not satisfied. If $\xi_1 \le Q_1 + S - M - L_2, \xi_2 \le Q_2 + L_2$, then $A(X,\xi) \le S - M - L_2 + L_2 + M = S$. The case $\xi_2 \le Q_2 + S - M - L_1, \xi_1 \le Q_1 + L_1$ is completely symmetric, and it shows that neither $A(X,\xi) > S$ nor (19a-19c) hold in this case.

Next, we use inequalities (19a-19c) to characterize the set $\{A(X,\xi) > S\}$ in the first-order conditions (7)–(8). We observe that in the two-exchange case

$$\{A(X,\xi) > S, \xi_1 > Q_1 + L_1\} = \{\xi_1 > Q_1 + L_1, \xi_2 > Q_2 + S - M - L_1\}$$
$$\{A(X,\xi) > S, \xi_2 > Q_2 + L_2\} = \{\xi_2 > Q_2 + L_2, \xi_1 > Q_1 + S - M - L_2\},$$

and then use the independence of ξ_1 and ξ_2 to compute

$$\mathbb{P}(A(X,\xi) > S|\xi_1 > Q_1 + L_1) = \bar{F}_2(Q_2 + S - M - L_1)$$
$$\mathbb{P}(A(X,\xi) > S|\xi_2 > Q_2 + L_2) = \bar{F}_1(Q_1 + S - M - L_2)$$

Together with (7) and (8), this leads to a pair of equations for limit orders sizes:

$$\bar{F}_2(Q_2 + S - M - L_1) = \frac{\lambda_u + s + r_1}{\lambda_u + \lambda_o} \quad \bar{F}_1(Q_1 + S - M - L_2) = \frac{\lambda_u + s + r_2}{\lambda_u + \lambda_o}$$

whose solution is given by $L_1^\star, L_2^\star$ from (9a,9b). To obtain the equation (9c), we rewrite the first equation in (7,8) using the inequalities (19a-19c). Then $P(A(X,\xi) > S)$ may be computed as the integral of the product measure $F_1 \otimes F_2$ over the region defined by

$$U(Q,S,M,L_1,L_2) = \{(x_1,x_2) \in \mathbb{R}^2, \quad x_1 > Q_1 + S - M - L_2, \quad x_2 > Q_2 + S - M - L_1, \quad x_1 + x_2 > Q_1 + Q_2 + S - M\}.$$

This integral is given by

$$P(A(X,\xi) > S) = F_1 \otimes F_2\,(U(Q,S,M,L_1,L_2))$$
$$= \bar{F}_1(Q_1 + L_1)\bar{F}_2(Q_2 + S - M - L_1) + \int_{Q_1+S-M-L_2}^{Q_1+L_1} \bar{F}_2(Q_1 + Q_2 + S - M - x_1)dF_1(x_1) = \frac{\lambda_u - (s+f)}{\lambda_u + \lambda_o}$$

■

References

Alfonsi, Aurélien, Antje Fruth, Alexander Schied. 2010. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance* **10**(2) 143–157.

Aliprantis, Charalambos, Owen Burkinshaw. 1998. *Principles of Real Analysis*. 3rd ed. Academic Press.

Almgren, R, N Chriss. 2000. Optimal execution of portfolio transactions. *Journal of Risk* **3** 5–39.

Bayraktar, Erhan, Michael Ludkovski. 2011. Liquidation in limit order books with controlled intensity.

Bertsimas, D, Andrew W Lo. 1998. Optimal control of execution costs. *Journal of Financial Markets* **1**(1) 1–50.

Boehmer, Ekkehart, Robert Jennings. 2007. Public disclosure and private decisions: Equity market execution quality and order routing. *Review of Financial Studies* **20**(2) 315–358.

Cont, R. 2011. Statistical modeling of high-frequency financial data. *IEEE Signal Processing* **28**(5) 16–25.

Cont, R., A. De Larrard. 2011. Price dynamics in a markovian limit order market. Working paper. URL `http://ssrn.com/abstract=1735338`.

Foucault, T, A.J. Menkveld. 2008. Competition for Order Flow and Smart Order Routing Systems. *Journal of Finance* **63**(1) 112.

Ganchev, Kuzman, Yuriy Nevmyvaka, Michael Kearns. 2010. Censored exploration and the dark pool problem. *Communications of the ACM* **53**(5) 99.

Gueant, Olivier, Charles-Albert Lehalle. 2012. General Intensity Shapes in Optimal Liquidation.

Guilbaud, Fabien, Huyen Pham. 2012. Optimal high frequency trading in a pro-rata microstructure with predictive information.

Huitema, Robert. 2012. Optimal Portfolio Execution using Market and Limit Orders.

Kushner, Harold, George Yin. 2003. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, New York.

Laruelle, Sophie, Charles-Albert Lehalle, Gilles Pagès. 2009. Optimal split of orders across liquidity pools: a stochastic algorithm approach.

Moallemi, Ciamac, Constantinos Maglaras, Hua Zheng. 2011. Optimal Order Flow Routing, Exchange Competition, and the Effect of Make/Take Fees.

Nemirovski, A, A Juditsky, G Lan, A Shapiro. 2009. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization* **19**(4) 1574–1609.

Obizhaeva, Anna, Jiang Wang. 2005. Optimal Trading Strategy and Supply/Demand Dynamics.

Predoiu, Silviu, Gennady Shaikhet, Steven Shreve. 2011. Optimal Execution in a General One-Sided Limit-Order Book. *SIAM Journal on Financial Mathematics* **2**(1) 183–212.

Robbins, H, S Monro. 1951. A stochastic approximation method. *The Annals of Mathematical Statistics* **22**(3).