Optimal trade execution under stochastic volatility and liquidity

Patrick Cheridito
Princeton University
Princeton, NJ 08544, USA

Tardu Sepin
Princeton University
Princeton, NJ 08544, USA

January 2014

Abstract

We study the problem of optimally liquidating a financial position in a discrete-time model with stochastic volatility and liquidity. We consider the three cases where the objective is to minimize the expectation, an expected exponential and a mean-variance criterion of the implementation cost. In the first case, the optimal solution can be fully characterized by a forward-backward system of stochastic equations depending on conditional expectations of future liquidity. In the other two cases we derive Bellman equations from which the optimal solutions can be obtained numerically by discretizing the control space. In all three cases we compute optimal strategies for different simulated realizations of prices, volatility and liquidity and compare the outcomes to the ones produced by the deterministic strategies of Bertsimas and Lo and Almgren and Chriss.

Key words: Optimal trade execution, implementation cost, discrete-time stochastic control, Bellman equation, stochastic volatility, stochastic liquidity.

1 Introduction

We consider the problem of optimally liquidating a given number of shares of a financial asset in a discrete-time model with stochastic volatility and liquidity. It is assumed that every transaction causes a permanent and temporary impact on the price, and the goal is to find strategies that minimize the expectation, an expected exponential or a mean-variance criterion of the implementation cost.

The optimal trade execution problem was addressed by Bertsimas and Lo (1998) in a model where the asset price process follows a discrete-time random walk and transactions cause a permanent price impact that is linear in the amount of shares traded. In this framework the strategy with the minimal expected execution cost turns out to have constant speed. In a further fundamental contribution, Almgren and Chriss (2001) added a temporary price impact from which the asset recovers in the next period and derived the best deterministic strategy with respect to a mean-variance criterion. Since it takes the variance into account, it is risk averse and executes faster than the constant speed strategy. Almgren and Lorenz (2007, 2011) investigated strategies that are allowed to react to price changes. Schied et al. (2010) studied the optimal execution problem for an agent with expected exponential utility in a continuous-time model with Lévy noise and found the continuous-time analog of the

Almgren–Chriss strategy to be optimal. Forsyth (2011) and Gatheral and Schied (2011) considered a continuous-time model in which the unperturbed price process is a geometric Brownian motion. Resilience of price impact with different decay kernels has been studied in Gatheral et al. (2012), Alfonsi and Schied (2012) and Obizhaeva and Wang (2013). Bayraktar and Ludkovski (2011) and Moazeni et al. (2013) modeled stochastic price impact with a Poisson process. The purpose of this paper is to extend the Almgren–Chriss model by allowing the volatility and liquidity to be stochastic. It is most closely related to Walia (2006) and Almgren (2012). Walia (2006) assumed the volatility and liquidity to follow a Markov chain and minimized one-step mean-variance criteria with dynamic programming methods. Almgren (2012) studied the continuous-time analog in which the volatility and liquidity follow correlated diffusion processes. For a detailed literature review we refer to Gatheral and Schied (2013).

We consider a stochastic volatility and temporary price impact and assume that prices, volatility and liquidity are observable. To keep the model tractable, we make Markov assumptions. If the objective is to minimize the expected implementation cost, we allow volatility and liquidity at each time to depend on the last realization of volatility, liquidity and price innovation. The optimal strategy is given by the solution of a forward-backward system of stochastic equations. It generalizes the constant speed strategy of Bertsimas and Lo (1998) and changes the position in every step by a fraction that depends on conditional expectations of future liquidity terms. The volatility and the price innovations affect the strategy only through their codependence with the liquidity. In the case of an expected exponential or mean-variance criterion, volatility and liquidity are assumed to follow a coupled Markov chain that is independent of the price innovations. In the expected exponential case we derive a Bellman equation that can be solved numerically by discretizing the control space. It turns out that it is enough for the optimal strategy to observe volatility and liquidity. The mean-variance case is more complicated. First, since conditional variances do not iterate, a direct dynamic programing approach does not work. Second, the optimal strategy depends on past realizations of volatility, liquidity and price innovations. To simplify the problem, we first only allow strategies that just observe volatility and liquidity. After that we solve the full problem in which strategies can react to changing volatility, liquidity and prices. In both cases we introduce an auxiliary quadratic cost minimization problem and deduce the associated Bellman equation. In our numerical experiments the optimal restricted strategy has practically the same performance as the fully optimal strategy and the optimal strategy for the expected exponential criterion with appropriately chosen risk aversion parameter. In the presence of stochastic volatility and liquidity they all significantly outperform the deterministic Almgren–Chriss strategy corresponding to the long term time-averages of volatility and liquidity. They are related to the strategies of Walia (2006) and Almgren (2012). The difference is that Walia (2006) and Almgren (2012) minimize a local mean-variance criterion, while in this paper the objectives are global.

The structure of the paper is as follows. In Section 2 we introduce our model of price behavior and trading impact. In Section 3 we study the risk neutral case. Section 4 treats the expected exponential case and Section 5 the mean-variance case. Section 6 concludes. All proofs are given in the Appendix.

2 The model

We consider the problem of liquidating an asset position of $X \in \mathbb{R}_+$ shares until a given time $T \in \mathbb{R}_+$. We divide the interval $[0, T]$ into N subintervals of length Δt and decide at every time $t_{n-1} = (n-1)\Delta t$ how many shares y_n to sell in the interval $(t_{n-1}, t_n]$. We assume that the orders y_n are executed by a tactical program that submits a combination of limit and market orders to obtain an optimal price, and that this results in an execution price of

$$\tilde{S}_n = S_{n-1} - \eta_n y_n, \tag{2.1}$$

where S_n follows the dynamics

$$S_n = S_{n-1} + \sigma_n \xi_n - c y_n. \tag{2.2}$$

We think of S_n as a fixed convex combination of the bid and ask price: $S_n = \lambda S_n^b + (1-\lambda) S_n^a$. For $\lambda = 1/2$, S_n is the mid price. But depending on the particular asset to be traded it could be closer to the bid or ask. (ξ_n) is a sequence of independent normally distributed random variables with mean 0 and variance Δt. (σ_n) is a stochastic volatility. $c \in \mathbb{R}_+$ is a constant describing a permanent price impact and (η_n) a stochastic liquidity process modeling temporary price impact.

We suppose that S_n, σ_n, η_n are observable and denote the filtration they generate by $\mathcal{F}_n$. S_n and $\tilde{S}_n$ can be observed directly. If $y_n > 0$, η_n can be deduced from (2.1). Otherwise, it has to be inferred from the bid-ask spread. σ_n has to be estimated from (2.2) or by a separate statistical procedure using more frequent observations. The proceeds from selling the asset shares are $\sum_{n=1}^N y_n \tilde{S}_n$. We describe a liquidation strategy in terms of remaining shares $x_n = X - \sum_{i=1}^n y_i$ and call it admissible if the following conditions are satisfied:

$$x_0 = X, \quad x_{n-1} \geq x_n, \quad x_N = 0, \quad (x_n) \text{ is predictable with respect to } (\mathcal{F}_n).$$

This means that each y_n must be non-negative and $\mathcal{F}_{n-1}$-measurable such that $\sum_{n=1}^N y_n = X$. An admissible strategy is completely specified by x_n for $1 \leq n \leq N-1$. We denote the set of all admissible strategies by $\mathcal{A}$.

The implementation cost is the difference between the initial value and the proceeds: $C(x) := XS_0 - \sum_{n=1}^N y_n \tilde{S}_n$. Since $x_N = 0$, it can be written as

$$C(x) = \frac{cX^2}{2} + \sum_{n=1}^N (x_{n-1} - x_n)^2 \left(\eta_n - \frac{c}{2}\right) - x_n \sigma_n \xi_n.$$

In the next three sections we try to find strategies that minimize the expectation, expected exponential and a mean-variance criterion of $C(x)$. In all three cases one can drop the constant $cX^2/2$ from $C(x)$ and work with the quadratic form

$$Q(x) = \sum_{n=1}^N (x_{n-1} - x_n)^2 \left(\eta_n - \frac{c}{2}\right) - x_n \sigma_n \xi_n \tag{2.3}$$

without changing the optimal solution.

Parameter	Symbol	Value
Initial stock price	S_0	172 \$/share
Initial position	X	35,000 shares
Duration	T	100 minutes
Number of subintervals	N	100
Length of subintervals	Δt	1 minute
Permanent impact	c	2.5×10^{-7}
Volatility states	σ_{low}, σ_{med}, σ_{high}	3.51×10^{-3}, 3.3×10^{-2}, 1.172×10^{-1}
Liquidity states	η_{low}, η_{med}, η_{high}	10^{-6}, 5×10^{-6}, 25×10^{-6}

Table 1: Parameter values

If the objective is to minimize the expectation of $Q(x)$, the dynamics of (S_n, σ_n, η_n) can be chosen slightly more general than for non-linear criteria. We always assume that ξ_n is independent of $\sigma(\mathcal{F}_{n-1}, \sigma_n, \eta_n)$, (σ_n, η_n) takes values in a finite subset $V \subseteq \mathbb{R}^2_+$, and $\tilde{\eta}_n := \eta_n - c/2 > 0$. In Section 3 we suppose that conditioned on $\mathcal{F}_{n-1}$, the distribution of (σ_n, η_n) just depends on $(\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1})$, where φ_{n-1} is defined as $\varphi_{n-1} := \varphi(\sigma_{n-1}\xi_{n-1})$ for a positive integer k and a measurable function $\varphi : \mathbb{R} \to \{1, \ldots, k\}$ (e.g. $\varphi(s) = 1$ for $s < 0$ and $\varphi(s) = 2$ for $s \geq 0$). It follows that $(\sigma_n, \eta_n, \varphi_n)$ is a Markov chain with finite state space $V^k := V \times \{1, \ldots, k\}$ and time-dependent transition probabilities

$$p^{vw}_{n-1} := \mathbb{P}\left[(\sigma_n, \eta_n, \varphi_n) = w \mid (\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1}) = v\right], \quad v, w \in V^k.$$

This covers models in which (σ_n, η_n) may depend on the last price innovation. In Sections 4 and 5 we assume that (σ_n, η_n) is a Markov chain with state space V which is independent of the sequence of innovations (ξ_n). Its transition probabilities are given by

$$p^{vw}_{n-1} := \mathbb{P}\left[(\sigma_n, \eta_n) = w \mid (\sigma_{n-1}, \eta_{n-1}) = v\right], \quad v, w \in V.$$

A typical liquidation problem has a duration of 1–6 hours. For our numerical simulations we used the parameter values of Table 1 and assumed (σ_n), (η_n), (ξ_n) to be independent such that (σ_n), (η_n) are time-homogeneous Markov chains with transition matrices

$$p^\sigma = \begin{pmatrix} 0.9349 & 0.0434 & 0.0217 \\ 0.7164 & 0.2239 & 0.0597 \\ 0.4400 & 0.4800 & 0.0800 \end{pmatrix} \quad \text{and} \quad p^\eta = \begin{pmatrix} 0.50 & 0.30 & 0.20 \\ 0.15 & 0.80 & 0.05 \\ 0.05 & 0.05 & 0.90 \end{pmatrix}.$$

We estimated the volatility parameters from 20 days of TAQ data for Panera Bread. The liquidity parameters are built around those of Almgren and Chriss (2001). We always used σ_{low} and η_{low} as starting points for the chains (σ_n) and (η_n).

3 Risk neutral objective

In the risk neutral case the objective is to find an admissible strategy x that minimizes the expectation of $Q(x)$. In this section $(\sigma_n, \eta_n, \varphi_n)$ is assumed to be a Markov chain with finite state space $V^k :=$

$V \times \{1, \ldots, k\} \subseteq \mathbb{R}_+^2 \times \mathbb{N}$ and transition probabilities

$$p_{n-1}^{vw} := \mathbb{P}\left[(\sigma_n, \eta_n, \varphi_n) = w \mid (\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1}) = v\right], \quad v, w \in V^k.$$

Due to the special form of $Q(x)$, it is enough for an optimal strategy to observe the process $(\sigma_n, \eta_n, \varphi_n)$, $n \geq 0$. Therefore, the goal is to minimize $\mathbb{E}_0^v[Q(x)]$, where for every possible state v of the Markov chain $(\sigma_n, \eta_n, \varphi_n)$, $\mathbb{E}_n^v$ denotes the conditional expectation $\mathbb{E}[. \mid (\sigma_n, \eta_n, \varphi_n) = v]$. The following theorem gives the optimal value and best strategy for this problem.

Theorem 3.1. *One has*

$$\min_{x \in \mathcal{A}} \mathbb{E}_0^v[Q(x)] = X^2 a_0^v, \tag{3.1}$$

and the unique optimal liquidation strategy is given by

$$x_n^*|_{x_{n-1}^*, (\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1}) = v} = x_{n-1}^* \frac{\mathbb{E}_{n-1}^v[\tilde{\eta}_n]}{\mathbb{E}_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^w}, \quad n = 1, \ldots, N-1, \tag{3.2}$$

where the coefficients a_n^v satisfy the backwards recursion:

$$a_{N-1}^v = \mathbb{E}_{N-1}^v[\tilde{\eta}_N], \quad a_{n-1}^v = \frac{\mathbb{E}_{n-1}^v[\tilde{\eta}_n] \sum_{w \in V^k} p_{n-1}^{vw} a_n^w}{\mathbb{E}_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^w}, \quad n \leq N-1. \tag{3.3}$$

The forward-backward system (3.2)–(3.3) defines the optimal strategy as a function of the liquidity available in the market. It decreases the position in every step by a fraction that depends on the expected liquidity in the next period and a weighted sum of future liquidity terms.

Remark 3.2. In the case of constant liquidity $\eta_n \equiv \eta$, Theorem 3.1 yields

$$a_n^v = \frac{\eta - c/2}{N-n}, \quad x_n^* = x_{n-1}^* \frac{N-n}{N-n+1},$$

which gives the deterministic constant speed strategy

$$x_n^* = X \frac{N-n}{N} \tag{3.4}$$

and so is consistent with the findings of Bertsimas and Lo (1998).

Simulation

We calculated the optimal strategy for 50,000 simulated paths of $(\sigma_n, \eta_n, \xi_n)$. Figure 1 shows the average optimal position $\bar{x}_n$ in percent of the initial number of shares X on the left and the corresponding standard deviation s_n on the right. Figure 2 shows the optimal strategy for three particular paths compared to the constant speed strategy (3.4). The histograms of Figure 3 show realizations of the implementation cost $C(x)$. It can be seen, that the optimal admissible strategy produces a significantly lower sample mean than the constant speed strategy. Coincidentally, it also generates a smaller standard deviation.

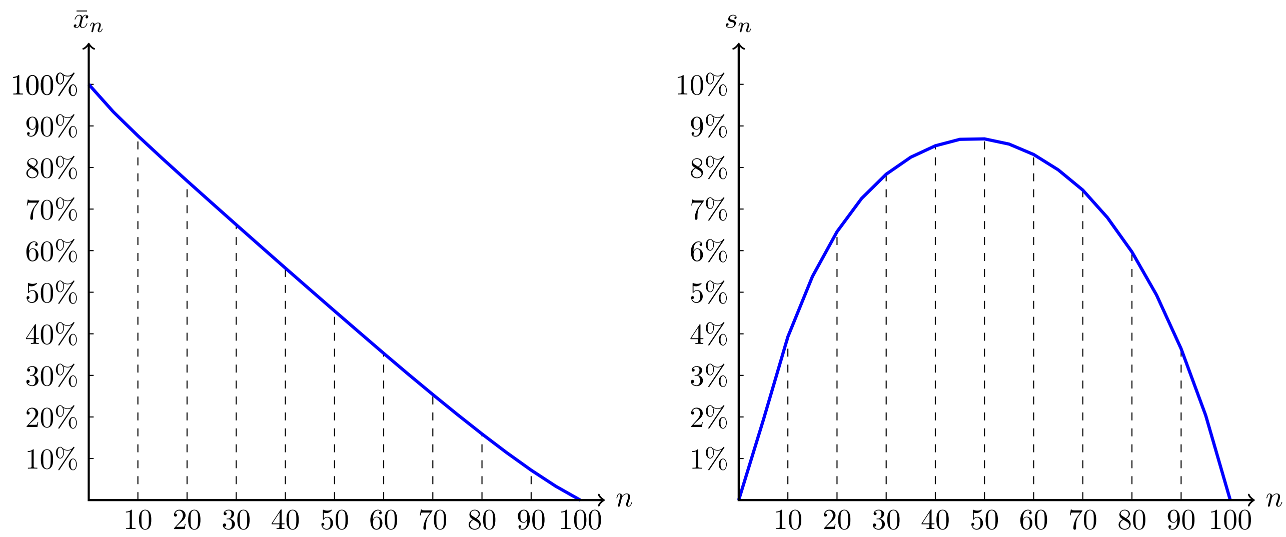

Figure 1: Average and standard deviation of the optimal positions over different scenarios in the risk neutral case.

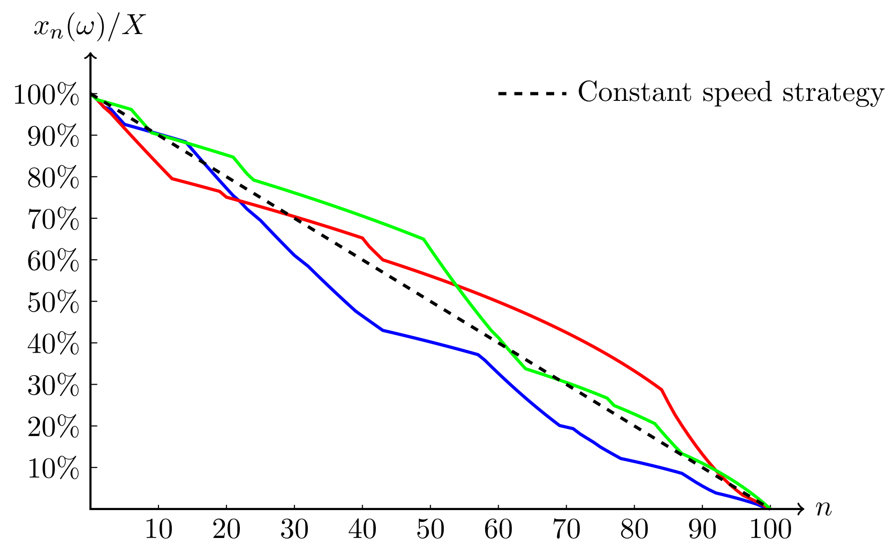

Figure 2: Optimal strategies for three particular scenarios in the risk neutral case compared to the constant speed strategy.

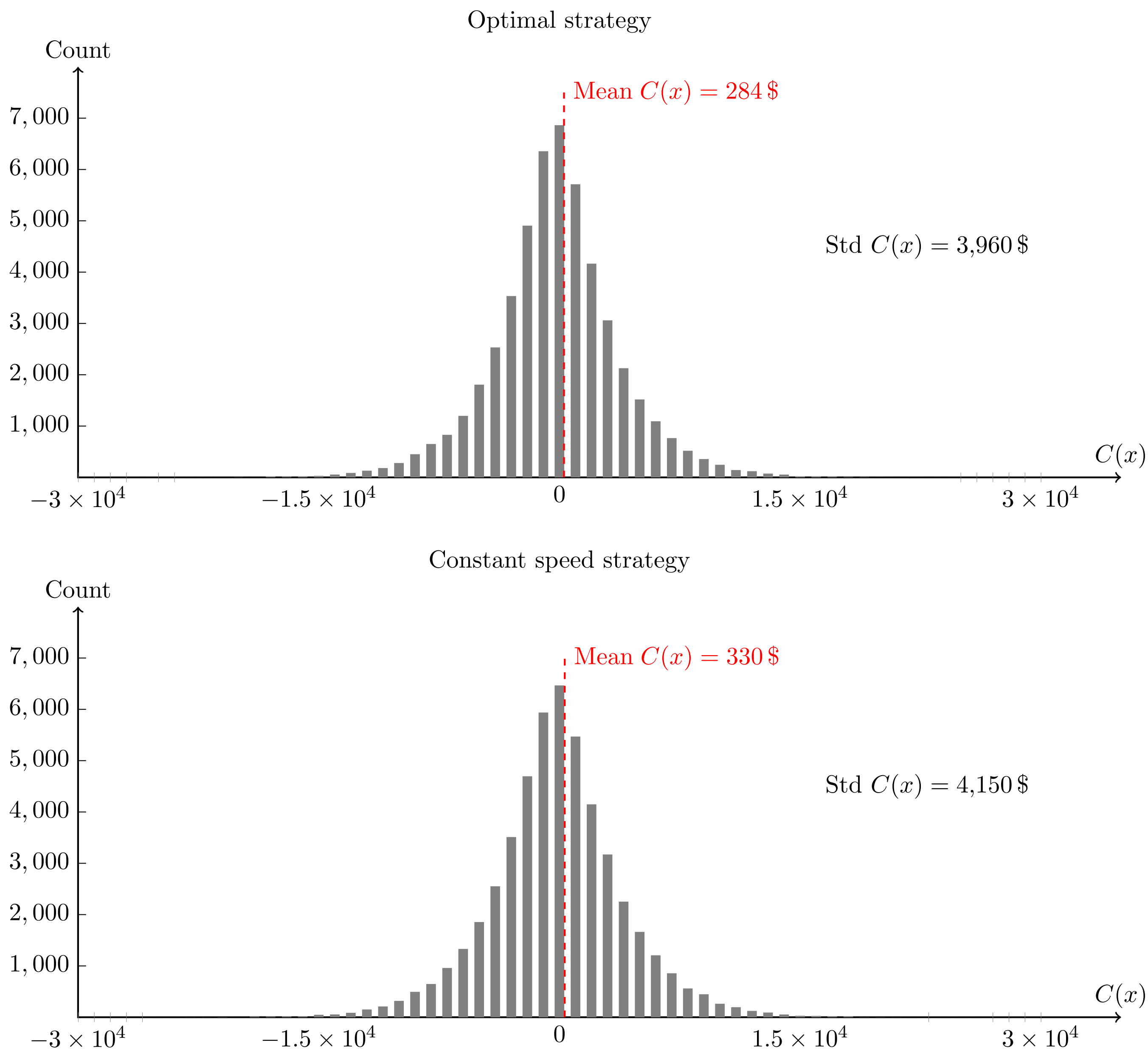

Figure 3: Histograms of $C(x)$ for the optimal strategy and the constant speed strategy in the risk neutral case.

4 Expected exponential cost

In this section the goal is to minimize the expected value of $\exp(\alpha Q(x))$ for a given risk aversion parameter $\alpha > 0$. It is assumed that (σ_n, η_n) is a Markov chain with finite state space $V \subseteq \mathbb{R}^2_+$ and transition probabilities

$$p^{vw}_{n-1} := \mathbb{P}\left[(\sigma_n, \eta_n) = w \mid (\sigma_{n-1}, \eta_{n-1}) = v\right], \quad v, w \in V,$$

which is independent of the sequence of innovations (ξ_n). Denote by $\mathcal{A}_n(z)$ the set of $(\mathcal{F}_n)$-predictable strategies $(x_i)_{i=n}^N$ satisfying $x_n = z$, $x_{i-1} \geq x_i$, $x_N = 0$. Since $Q(x)$ is a sum of local terms and the exponential function factorizes, it is enough for the optimal strategy to observe (σ_n, η_n), $n \geq 0$. So the problem is to minimize the expected exponential

$$\mathbb{E}^v_0[\exp(\alpha(Q(x))] \tag{4.1}$$

for all possible states $v \in V$, where $\mathbb{E}^v_n$ denotes the conditional expectation $\mathbb{E}[. \mid (\sigma_n, \eta_n) = v]$. Consider the value function

$$J^v_n(z) := \min_{x \in \mathcal{A}_n(z)} \mathbb{E}^v_n\left[\exp\left(\alpha Q_n(x)\right)\right],$$

where $Q_n(x) := \sum_{i=n+1}^N (x_{i-1} - x_i)^2 \tilde{\eta}_i - x_i \sigma_i \xi_i$. Then one has the following

Theorem 4.1. *The value function J satisfies the Bellman equation*

$$\begin{aligned} J^v_{N-1}(x_{N-1}) &= \sum_{w \in V} p^{vw}_{N-1} \exp\left(\alpha x^2_{N-1}(w_2 - c/2)\right) \\ J^v_{n-1}(x_{n-1}) &= \min_{0 \leq x_n \leq x_{n-1}} \sum_{w \in V} p^{vw}_{n-1} \exp\left(\alpha(x_{n-1} - x_n)^2(w_2 - c/2) + \frac{1}{2}x^2_n \alpha^2 w^2_1 \Delta t\right) J^w_n(x_n) \end{aligned} \tag{4.2}$$

for $n \leq N-1$ and the minimizing x^*_n form the unique optimal strategy for problem (4.1).

Remark 4.2. It can be seen from Theorem 4.1 that for constant volatility $\sigma_n \equiv \sigma$ and liquidity $\eta_n \equiv \eta$, the minimization of the expected exponential (4.1) reduces to the deterministic problem

$$\min_{0 \leq x_n \leq x_{n-1}} \sum_{n=1}^N (x_n - x_{n-1})^2(\eta - c/2) + \frac{1}{2}x^2_n \alpha \sigma^2 \Delta t. \tag{4.3}$$

This observation is in line with Schied et al. (2010). Since (4.3) is a convex problem, it can be reduced to the first order condition

$$x_n \alpha \sigma^2 \Delta t = (x_{n-1} - 2x_n + x_{n+1})(2\eta - c), \quad n = 1, \ldots, N-1.$$

This is the equation derived by Almgren and Chriss (2001) for the deterministic strategy x minimizing the mean-variance criterion $\mathbb{E}[Q(x)] + (\alpha/2)\mathrm{Var}(Q(x))$. Its solution is given by

$$x^*_n = X \frac{\sinh(\kappa(T - n\Delta t))}{\sinh(\kappa T)} \quad \text{for the unique } \kappa > 0 \text{ satisfying } \cosh(\kappa \Delta t) - 1 = \frac{\alpha \sigma^2 \Delta t}{4\eta - 2c}. \tag{4.4}$$

We point out that while in the case of constant volatility and liquidity, (4.4) is the best $(\mathcal{F}_n)$-predictable strategy for the exponential problem, the optimal $(\mathcal{F}_n)$-predictable mean-variance strategy is not deterministic; see Almgren and Lorenz (2007), Lorenz and Almgren (2011) and Section 5.

If the volatility and liquidity are stochastic, the Bellman equation of Theorem 4.1 can be solved numerically by discretizing the position space. If the asset can be sold in single units, the complexity of the algorithm is $\mathcal{O}(X^2)$. This is because in every step, $J^v_{n-1}(x_{n-1})$ has to be computed for all $0 \le x_{n-1} \le X$, and for fixed x_{n-1}, the calculation of $J^v_{n-1}(x_{n-1})$ requires the evaluation of the right side of (4.2) for $x_n = 0, \ldots, x_{n-1}$. For large X, the complexity can be reduced by selling the asset in larger lots.

Simulation
In our numerical experiment we allowed trading in lots of 350 shares, which amounts to 1% of the initial position of 35,000 shares. We calculated optimal strategies along 50,000 simulated realizations of $(\sigma_n, \eta_n, \xi_n)$. Figure 4 shows the average $\bar{x}_n$ of the optimal positions and the standard deviations s_n for three different values of absolute risk aversion α. It can be seen that the average liquidation speed is increasing in α. Figure 5 shows three realizations of the optimal admissible strategy compared to the Almgren–Chriss strategy (4.4) corresponding to the mean volatility $\bar{\sigma}$ and liquidity $\bar{\eta}$ with respect to the steady state distributions of (σ_n) and (η_n). Figure 6 shows histograms of realized implementation costs $C(x)$ produced by the optimal strategy of Theorem 4.1 and the Almgren–Chriss strategy (4.4). It can be seen that the sample mean and variance of the optimal admissible strategy are significantly lower.

5 Mean-variance criterion

In this section we try to find a strategy that minimizes a mean-variance criterion of the form

$$\mathbb{E}[Q(x)] + \lambda \mathrm{Var}(Q(x)) \tag{5.1}$$

for a given trade-off parameter $\lambda > 0$. As in Section 4, we assume that (σ_n, η_n) is a Markov chain with finite state space $V \subseteq \mathbb{R}^2_+$ and transition probabilities

$$p^{vw}_{n-1} := \mathbb{P}\left[(\sigma_n, \eta_n) = w \mid (\sigma_{n-1}, \eta_{n-1}) = v\right], \quad v, w \in V,$$

which is independent of the sequence of innovations (ξ_n). Problem (5.1) is more difficult than the exponential problem (4.1). First, the criterion (5.1) does not factorize. So it is no longer enough for the optimal admissible strategy to only observe (σ_n, η_n). Second, since conditional variances do not iterate, problem (5.1) cannot be treated recursively. We solve problem (5.1) in two different cases. In Subsection 5.1 we restrict the class of admissible trading strategies and find the best strategy among those which only observe (σ_n, η_n). This is equivalent to solving for the best deterministic strategy in the case where volatility and liquidity are constant. In Subsection 5.2 we compute the fully optimal admissible strategy. Our simulation results suggest that the restricted solution is very close to being fully optimal. Moreover it has virtually the same performance as the the optimal strategy for the exponential criterion (4.1) with $\alpha = 2\lambda$.

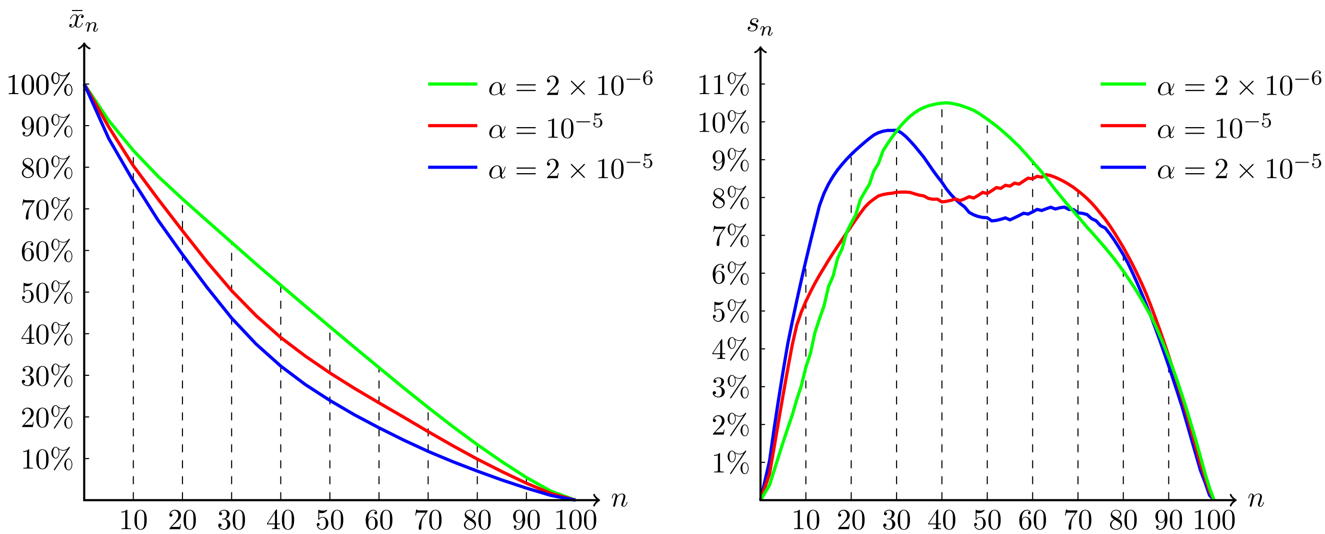

Figure 4: Average and standard deviation of optimal positions for the expected exponential cost criterion with different absolute risk aversions α.

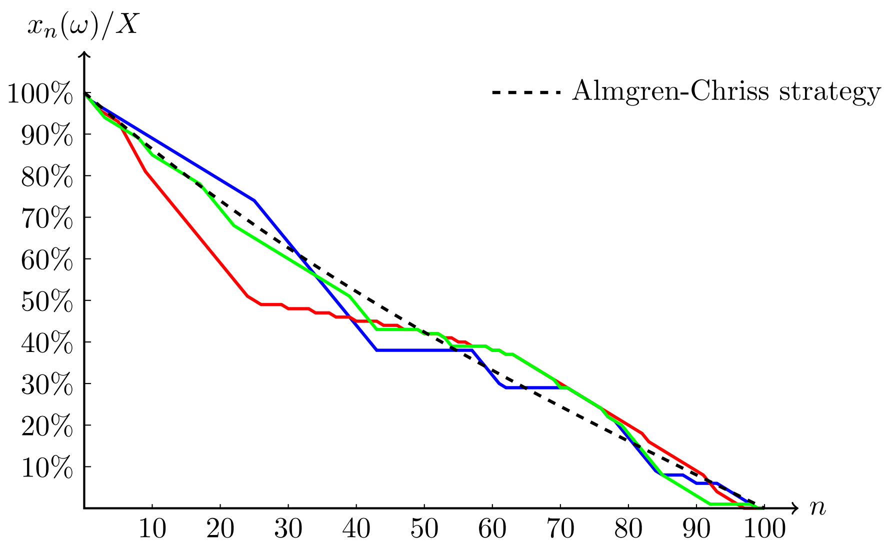

Figure 5: Three realizations of the optimal strategy for the expected exponential criterion with parameter $\alpha = 2 \times 10^{-5}$ compared to the Almgren–Chriss strategy (4.4) corresponding to $\bar{\sigma}$ and $\bar{\eta}$.

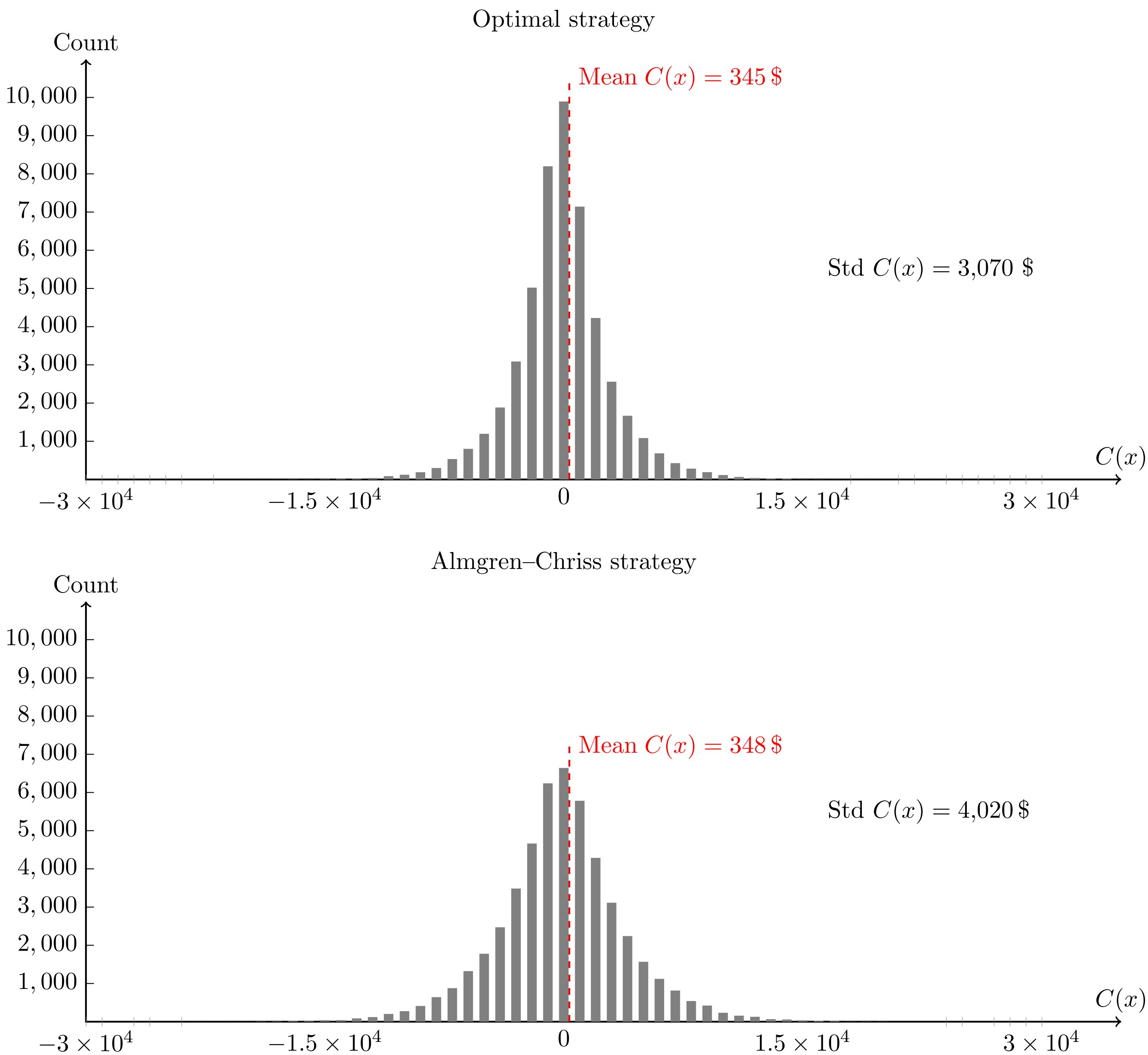

Figure 6: Histogram of $C(x)$ for the optimal strategy and the Almgren–Chriss strategy (4.4) corresponding to $\bar{\sigma}$ and $\bar{\eta}$ in the expected exponential cost case with absolute risk aversion $\alpha = 2 \times 10^{-5}$.

5.1 The restricted mean-variance problem

Define the filtration $\mathcal{G}_n := \sigma(\sigma_i, \eta_i : 0 \le i \le n)$, and denote by $\mathcal{R}$ the subset of admissible strategies $x \in \mathcal{A}$ that are predictable with respect to $(\mathcal{G}_n)$. The restricted mean-variance problem is

$$P(\lambda) \quad \min_{x \in \mathcal{R}} \mathbb{E}[Q(x) \mid \mathcal{G}_0] + \lambda \mathrm{Var}(Q(x) \mid \mathcal{G}_0)$$

for given $\lambda > 0$. Since the filtration $(\mathcal{G}_n)$ is discrete, $\mathcal{R}$ is a compact subset of a finite-dimensional space. Moreover, the objective function of $P(\lambda)$ is continuous in x. It follows that $P(\lambda)$ admits an optimal solution $x^* \in \mathcal{R}$. However, since conditional variances do not iterate, it cannot be solved recursively. But it is well-known (see Proposition A.1 in the appendix for a simple proof) that a solution x^* of $P(\lambda)$ also solves the problem

$$P(\lambda, \mu) \quad \min_{x \in \mathcal{R}} \mathbb{E}\left[\mu Q(x) + \lambda Q(x)^2 \mid \mathcal{G}_0\right]$$

for $\mu = 1 - 2\lambda \mathbb{E}\left[Q(x^*)\right]$. It is possible to derive a Bellman equation for problem $P(\lambda, \mu)$. However, μ is not known before x^*. So we compute solutions to $P(\lambda, \mu)$ for different values of μ and check which one minimizes $P(\lambda)$. To solve $P(\lambda, \mu)$, we denote by $\mathbb{E}_n^v$ the conditional expectation $\mathbb{E}[. \mid (\sigma_n, \eta_n) = v]$ and by $\mathcal{R}_n(z)$ the set of $(\mathcal{G}_n)$-predictable strategies $(x_i)_{i=n}^N$ satisfying $x_n = z$, $x_{i-1} \ge x_i$, $x_N = 0$. Define

$$J_n^v(z) := \min_{x \in \mathcal{R}_n(z)} \mathbb{E}_n^v\left[\mu Q_n(x) + \lambda Q_n(x)^2\right]$$

for $Q_n(x) := \sum_{i=n+1}^N (x_{i-1} - x_i)^2 \tilde{\eta}_i - x_i \sigma_i \xi_i$. Then $J_0^v(X)$ gives the optimal value of problem $P(\mu, \lambda)$ if $(\sigma_0, \eta_0) = v$, and the following holds:

Theorem 5.1. *The value function J satisfies the Bellman equation*

$$\begin{aligned}
J_{N-1}^v(x_{N-1}) &= \mu x_{N-1}^2 \mathbb{E}_{N-1}^v[\tilde{\eta}_N] + \lambda x_{N-1}^4 \mathbb{E}_{N-1}^v[\tilde{\eta}_N^2], \\
J_{n-1}^v(x_{n-1}) &= \min_{0 \le x_n \le x_{n-1}} \mu (x_{n-1} - x_n)^2 \mathbb{E}_{n-1}^v\left[\tilde{\eta}_n\right] + \lambda (x_{n-1} - x_n)^4 \mathbb{E}_{n-1}^v\left[\tilde{\eta}_n^2\right] \\
&\quad + \lambda x_n^2 \Delta t \mathbb{E}_{n-1}^v\left[\sigma_n^2\right] + 2\lambda (x_{n-1} - x_n)^2 \mathbb{E}_{n-1}^v\left[\tilde{\eta}_n \sum_{i=n+1}^N (x_{i-1} - x_i)^2 \tilde{\eta}_i\right] \\
&\quad + \sum_{w \in V} p_{n-1}^{vw} J_n^w(x_n), \quad n \le N-1,
\end{aligned} \tag{5.2}$$

and a strategy $x^* \in \mathcal{R}$ solves problem $P(\lambda, \mu)$ if for every $n = 1, \ldots, N-1$, x_n^* minimizes (5.2) *for $x_{n-1} = x_{n-1}^*$. Moreover, if $\mu \ge 0$, the optimal strategy $x^* \in \mathcal{R}$ is unique.*

Remark 5.2. For constant volatility $\sigma_n \equiv \sigma$ and liquidity $\eta_n \equiv \eta$, problem $P(\lambda)$ becomes the Almgren–Chriss problem (4.3) with $\lambda = \alpha/2$. (4.4) is the optimal deterministic solution, but it is suboptimal among all admissible strategies $\mathcal{A}$; see Almgren and Chriss (2001), Almgren and Lorenz (2007), Lorenz and Almgren (2011) and Subsection 5.2.

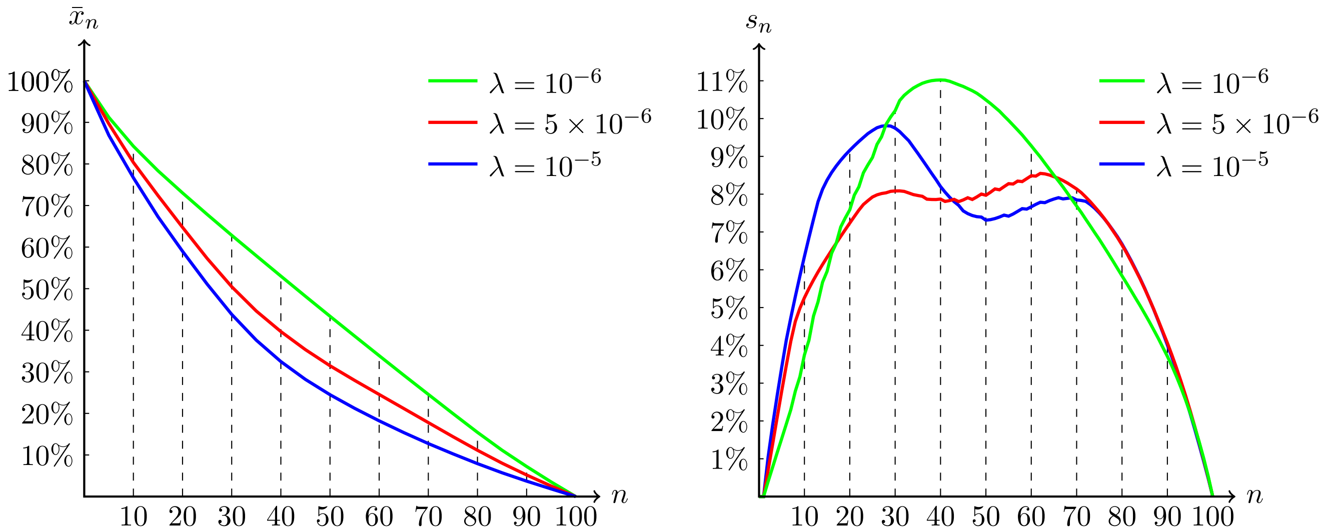

Figure 7: Average and standard deviation of the $\mathcal{R}$-optimal positions for the mean-variance criterion with different values of λ.

For stochastic volatility and liquidity, the Bellman equation of Theorem 5.1 can be solved numerically on a discrete grid of controls. As in the exponential case, if shares can be sold in quantities of single shares, the computational cost is of order $\mathcal{O}(X^2)$, and for big X it can be kept down by selling the asset in larger lots.

Simulation

In the numerical simulation we allowed shares to be sold in lots of 350 shares, which is 1% of the initial position, and we calculated optimal strategies along 50,000 simulated realizations of $(\sigma_n, \eta_n, \xi_n)$. Figure 7 shows the average optimal positions $\bar{x}_n$ and the standard deviations s_n for three different values of the mean-variance tradeoff parameter λ. It can be seen that the average speed of liquidation increases for larger values of λ. Figure 8 shows three realizations of the optimal restricted strategy compared to the Almgren–Chriss strategy (4.4) corresponding to the mean values $\bar{\sigma}$ and $\bar{\eta}$ under the steady state distributions of the Markov chains (σ_n) and (η_n). In Figure 9 the mean-variance criterion of the optimal solutions x^μ to problem $P(\lambda, \mu)$ are plotted for $\lambda = 10^{-5}$ and different values of μ. The minimum is attained for $\mu^* = 0.968$, which is very close to $1 - 2\lambda \mathbb{E}[Q(x^{\mu^*})] = 0.997$. This is consistent with the theoretical result of Proposition A.1. The histogram in Figure 10 shows realized implementation costs $C(x)$ generated by the best restricted strategy. It was computed along the same sample paths as the histograms of Figure 6. Note that the optimal strategy for the exponential criterion (4.1) with $\alpha = 2\lambda$ produced practically the same histogram.

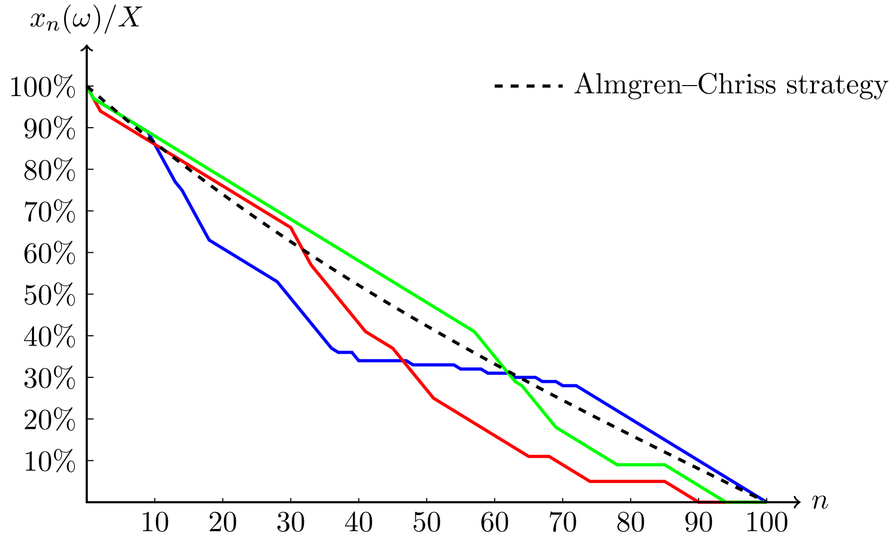

Figure 8: Three realizations of the optimal restricted strategy for the mean-variance criterion with parameter $\lambda = 10^{-5}$ compared to the Almgren–Chriss strategy (4.4) corresponding to $\bar{\sigma}$ and $\bar{\eta}$.

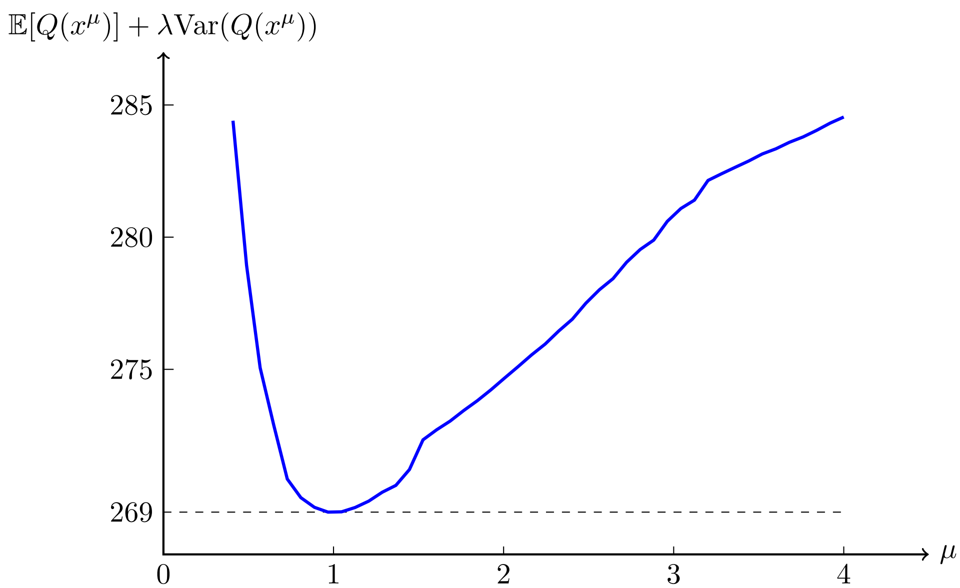

Figure 9: Mean-variance criterion for optimal solutions of $P(\lambda, \mu)$ for $\lambda = 10^{-5}$ and different values of μ. The minimum is attained for $\mu^* = 0.968$, and $1 - 2\lambda \mathbb{E}[Q(x^{\mu^*})] = 0.997$.

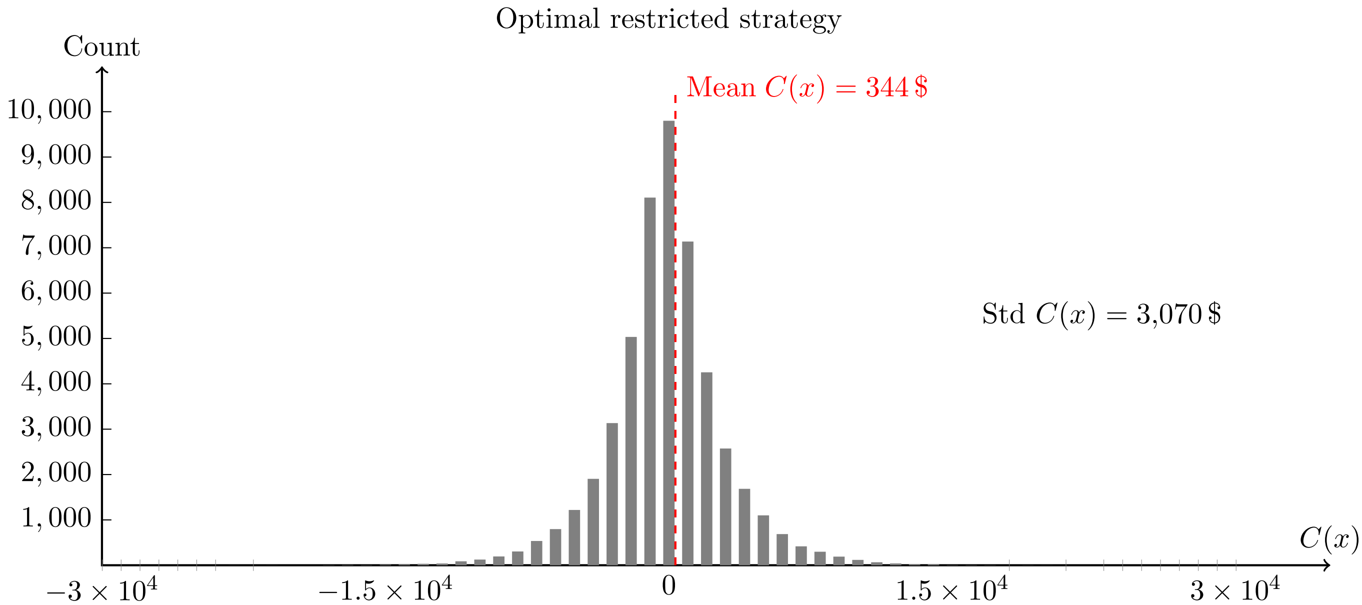

Figure 10: Histogram of $C(x)$ for the optimal restricted strategy in the mean-variance case for $\lambda = 10^{-5}$.

5.2 The full mean-variance problem

In this subsection we consider the problem

$$\hat{P}(\lambda) \quad \min_{x\in\mathcal{A}} \mathbb{E}[Q(x) \mid \mathcal{F}_0] + \lambda \mathrm{Var}(Q(x) \mid \mathcal{F}_0)$$

of finding a mean-variance optimal strategy in the set of all admissible strategies $\mathcal{A}$. $\mathcal{A}$ can be viewed as a bounded, closed, convex subset of a Hilbert space. Therefore, it is weakly compact, and it follows from standard arguments that $\hat{P}(\lambda)$ has an optimal solution $x^* \in \mathcal{A}$. As in Subsection 5.1, we derive a Bellman equation for the auxiliary problem

$$\hat{P}(\lambda, \mu) \quad \min_{x\in\mathcal{A}} \mathbb{E}\left[\mu Q(x) + \lambda Q(x)^2 \mid \mathcal{F}_0\right]$$

and try to solve it for different μ. To do this we introduce the running cost

$$h_n(x) := \sum_{i=1}^{n} (x_{i-1} - x_i)^2 \tilde{\eta}_i - x_i \sigma_i \xi_i,$$

and the value function

$$J_n^v(h, z) := \min_{x\in\mathcal{A}_n(z)} \mathbb{E}_n^v\left[(\mu + 2\lambda h) Q_n(x) + \lambda Q_n(x)^2\right],$$

where $\mathbb{E}_n^v$, $\mathcal{A}_n(z)$ and $Q_n(x)$ are defined as in Section 4. $J_0^v(0, X)$ is the optimal value of problem $\hat{P}(\mu, \lambda)$ given that $(\sigma_0, \eta_0) = v$, and the following theorem gives a Bellman equation for J.

Theorem 5.3. *The value function J satisfies the Bellman equation*

$$J_{N-1}^v(h, x_{N-1}) = (\mu + 2\lambda h)x_{N-1}^2 \mathbb{E}_{N-1}^v[\tilde{\eta}_N] + \lambda x_{N-1}^4 \mathbb{E}_{N-1}^v[\tilde{\eta}_N^2],$$

$$J_{n-1}^v(h, x_{n-1}) = \min_{0 \le x_n \le x_{n-1}} (\mu + 2\lambda h)(x_{n-1} - x_n)^2 \mathbb{E}_{n-1}^v\left[\tilde{\eta}_n\right] + \lambda (x_{n-1} - x_n)^4 \mathbb{E}_{n-1}^v\left[\tilde{\eta}_n^2\right]$$
$$+ \lambda x_n^2 \Delta t \mathbb{E}_{n-1}^v\left[\sigma_n^2\right] + \sum_{w \in V} p_{n-1}^{vw} \int_{\mathbb{R}} J_n^w\left(h + (x_{n-1} - x_n)^2 (w_2 - c/2) - x_n w_1 \sqrt{\Delta t}\xi, x_n\right) \rho(\xi) d\xi,$$

where ρ is the standard normal density. A strategy $x^* \in \mathcal{A}$ solves problem $\hat{P}(\mu, \lambda)$ if in every step, x_n^* realizes the minimum for $h = h_{n-1}(x^*)$ and $x_{n-1} = x_{n-1}^*$. Moreover, if $\mu \ge 0$, the optimal strategy $x^* \in \mathcal{A}$ is unique.

Remark 5.4. $\hat{P}(\lambda)$ extends the problem of Lorenz and Almgren (2011) to the case of stochastic volatility and liquidity. To solve it, one has to find solutions x^μ to the auxiliary problem $\hat{P}(\mu, \lambda)$ for different values of μ and check which one minimizes $\hat{P}(\lambda)$. However, the additional variable h makes the numerical evaluation of J much more complex than in Section 4 and Subsection 5.1. Therefore, we did the computation in an Almgren–Chriss model with constant volatility and liquidity and only five time steps. The numerical results indicate that the optimal solution is only insignificantly better than the Almgren–Chriss strategy (4.4).

Simulation
Since the numerical evaluation of the value function J of Theorem 5.3 is more complex than in Section 4 and Subsection 5.1, we restricted the simulation to a model with constant volatility and liquidity and only five time steps. We simulated $50,000$ realizations of $(\xi_n)_{n=1}^5$ and set the volatility and liquidity equal to the same mean values $\bar{\sigma}$ and $\bar{\eta}$ as in Section 4. To compute the optimal strategy we discretized the state space of running cost using 250 evenly spaced grid points in the interval $[150 - 2 \times 1500, 150 + 2 \times 1500]$. The values 150 and 1500 were inspired by the mean and standard deviation of $Q(x)$ generated by the Almgren–Chriss strategy. We allowed trading in lots of 350 shares, which corresponds to 1% of the initial position. The integral in Theorem 5.3 was computed with Monte Carlo. Figure 11 shows the μ-value producing the lowest mean-variance criterion. Figure 12 contains the histograms of the optimal and the Almgren–Chriss strategy for the case $\lambda = 10^{-5}$. The optimal strategy improves the mean-variance criterion slightly over the Almgren–Chriss strategy. But while it gives a better sample mean, its standard deviation is worse.

6 Conclusion

This paper studies the optimal asset liquidation problem in a discrete-time model with stochastic volatility and liquidity. We computed optimal liquidation strategies for risk neutral, expected exponential and mean-variance cost objectives. Our formulation allows volatility and liquidity to follow a

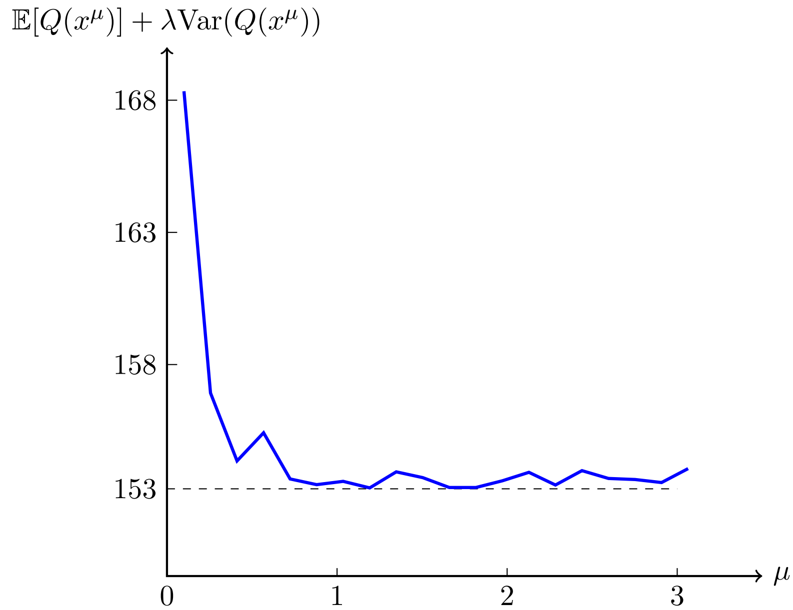

Figure 11: Mean-variance criterion for optimal solutions to $\hat{P}(\lambda, \mu)$ for $\lambda = 10^{-5}$ and different values of μ. The minimum is attained for $\mu^* = 1.192$, and $1 - 2\lambda\mathbb{E}[Q(x^{\mu^*})] = 0.998$.

coupled Markov chain, which in the risk neutral case, can also depend on past price innovations. The transition probabilities can be time-dependent. This makes the model flexible enough to describe situations in which price dynamics depend on the time of the day. For example, some small or medium cap stocks are traded less liquidly at certain times of the day, and volatility typically increases around scheduled news announcements. In the risk neutral case the optimal strategy is given as the solution of a forward-backward system of stochastic equations and generalizes the constant speed strategy of Bertsimas and Lo (1998). In the expected exponential and mean-variance case we derived Bellman equations than can be solved numerically by discretizing the control space. We found that expected exponential optimal strategies require less computing time than mean-variance optimal strategies and produce practically the same outcomes if the risk aversion parameters are chosen appropriately. In the presence of stochastic volatility and liquidity they significantly outperform strategies which assume that volatility and liquidity are constant. We conclude that in the kind of models studied in this paper it is most efficient to compute risk averse execution strategies by minimizing an expected exponential criterion even if the goal is to find mean-variance optimal strategies. To apply the model in practice, the parameters have to be estimated from real world data.

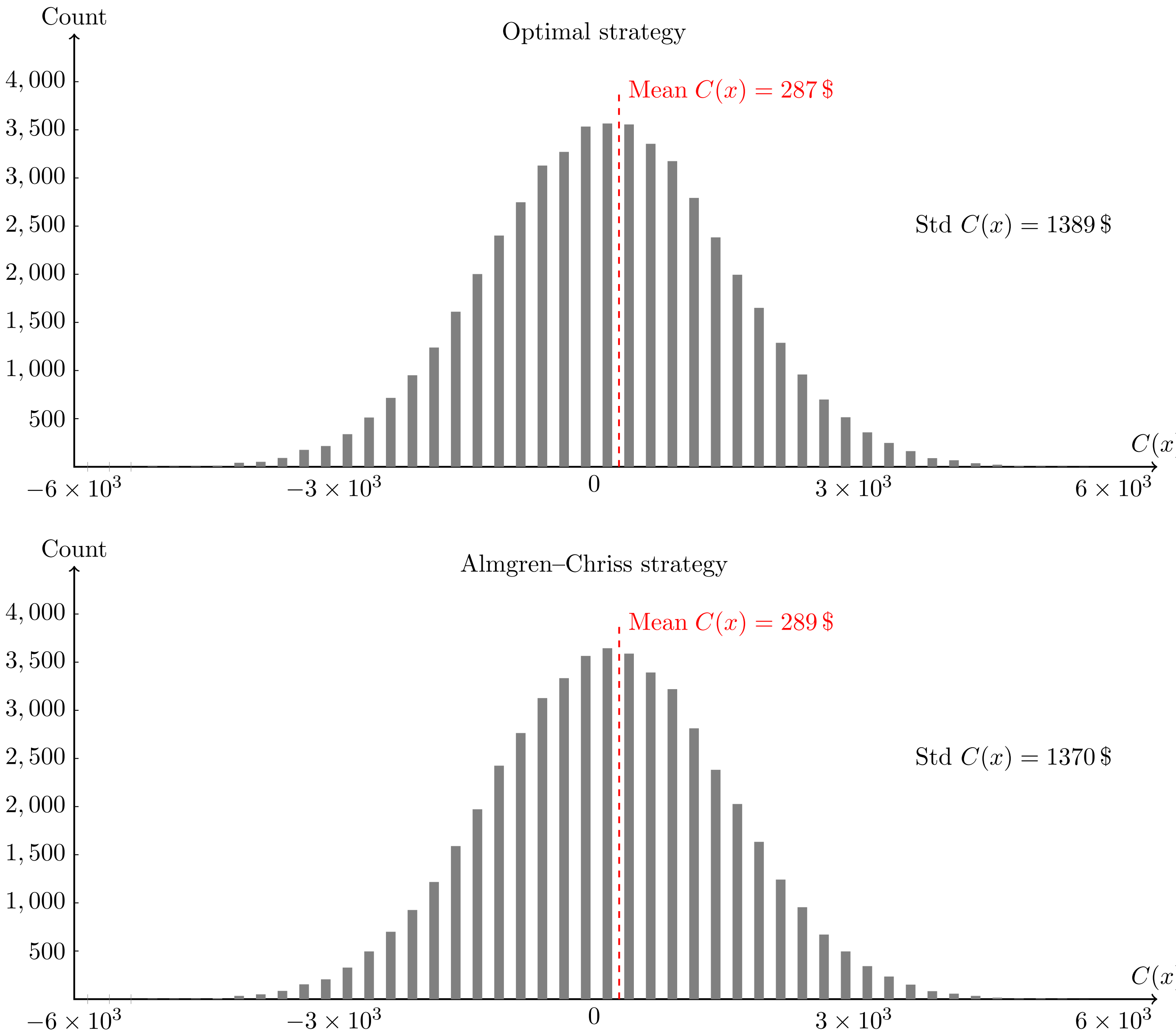

Figure 12: Histograms of $C(x)$ for the optimal strategy and the Almgren–Chriss strategy (4.4) in the mean-variance case with constant volatility $\bar{\sigma}$ and liquidity $\bar{\eta}$ for $\lambda = 10^{-5}$.

A Proofs

Proof of Theorem 3.1
We prove the theorem by backwards induction. Since $\mathbb{E}_0^v[Q(x)] = \mathbb{E}_0^v[R_0(x)]$ for

$$R_n(x) = \sum_{i=n+1}^{N} (x_{i-1} - x_i)^2 \tilde{\eta}_i,$$

we denote by $\mathcal{A}_n(z)$ the set of $(\mathcal{F}_n)$-predictable strategies $(x_i)_{i=n}^N$ satisfying $x_n = z$, $x_{i-1} \geq x_i$, $x_N = 0$, and define

$$J_n^v(z) := \min_{x \in \mathcal{A}_n(z)} \mathbb{E}_n^v[R_n(x)].$$

Then

$$J_{N-1}^v(x_{N-1}) = x_{N-1}^2 \mathbb{E}_{N-1}^v[\tilde{\eta}_N] = x_{N-1}^2 a_{N-1}^v,$$

and inductively,

$$\begin{aligned}
J_{n-1}^v(x_{n-1}) &= \min_{x \in \mathcal{A}_{n-1}(x_{n-1})} \mathbb{E}_{n-1}^v[R_{n-1}(x)] \\
&= \min_{x \in \mathcal{A}_{n-1}(x_{n-1})} (x_{n-1} - x_n)^2 \mathbb{E}_{n-1}^v[\tilde{\eta}_n] + \mathbb{E}_{n-1}^v[R_n(x_n)] \\
&= \min_{x_n} (x_{n-1} - x_n)^2 \mathbb{E}_{n-1}^v[\tilde{\eta}_n] + x_n^2 \sum_{w \in V^k} p_{n-1}^{vw} a_n^w, \quad n \leq N-1.
\end{aligned}$$

It follows that the unique optimal strategy is given by

$$x_n^* = x_{n-1}^* \frac{\mathbb{E}_{n-1}^v[\tilde{\eta}_n]}{\mathbb{E}_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^w},$$

and $J_{n-1}^v(x_{n-1}^*)$ becomes

$$(x_{n-1}^*)^2 \frac{\mathbb{E}_{n-1}^v[\tilde{\eta}_n] \sum_{w \in V^k} p_{n-1}^{vw} a_n^w}{\mathbb{E}_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^w} = (x_{n-1}^*)^2 a_{n-1}^v.$$

In particular, $J_0^v(X) = X^2 a_0^v$. □

Proof of Theorem 4.1
Since (σ_n, η_n) is independent of (ξ_n), one has $\mathbb{E}_n^v[\exp(\alpha Q_n(x))] = \mathbb{E}_n^v[R_n(x)]$, where

$$R_n(x) := \exp\left(\sum_{i=n+1}^{N} \alpha (x_{i-1} - x_i)^2 \tilde{\eta}_i + \frac{1}{2} x_i^2 \alpha^2 \sigma_i^2 \Delta t \right).$$

So

$$J_{N-1}^v(x_{N-1}) = \sum_{w \in V} p_{N-1}^{vw} \exp\left(\alpha x_{N-1}^2 (w_2 - c/2)\right)$$

and

$$
\begin{aligned}
J_{n-1}^{v}(x_{n-1}) &= \min_{x \in \mathcal{A}_{n-1}(x_{n-1})} \mathbb{E}_{n-1}^{v}\left[\exp\left(\alpha(x_{n-1}-x_n)^2\tilde{\eta}_n + \frac{1}{2}x_n^2\alpha^2\sigma_n^2\Delta t\right) R_n(x_n)\right] \\
&= \min_{0 \le x_n \le x_{n-1}} \sum_{w \in V} p_{n-1}^{vw} \exp\left(\alpha(x_{n-1}-x_n)^2(w_2 - c/2) + \frac{1}{2}x_n^2\alpha^2 w_1^2\Delta t\right) J_n^w(x_n),
\end{aligned}
$$

$n \le N-1$. Since $R_0(x)$ is strictly convex in x, the optimal strategy x^* is unique. □

Proposition A.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and E a non-empty subset of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Denote*

$$
V^{\lambda}(Y) := \mathbb{E}Y + \lambda \mathrm{Var}(Y), \quad V^{\lambda,\mu}(Y) := \mathbb{E}(\mu Y + \lambda Y^2),
$$

and assume V^{λ} attains a minimum over E at $Y^* \in E$. Then Y^* minimizes $V^{\lambda,\mu}$ over E for $\mu = 1 - 2\lambda\mathbb{E}Y^*$.

Proof. One can write $V^{\lambda}(Y) = f\left(\mathbb{E}Y, \mathbb{E}Y^2\right)$ for $f(y) = y_1 - \lambda y_1^2 + \lambda y_2$. Since f is concave, one has

$$
f(y) \le f(y^*) + \nabla f(y^*) \cdot (y - y^*) = f(y^*) + (1 - 2\lambda y_1^*)(y_1 - y_1^*) + \lambda(y_2 - y_2^*).
$$

So if $Y^* \in E$ is a minimizer of V^{λ} and $\mu = 1 - 2\lambda\mathbb{E}Y^*$, then

$$
V^{\lambda}(Y^*) \le V^{\lambda}(Y) \le V^{\lambda}(Y^*) + \mu\left(\mathbb{E}Y - \mathbb{E}Y^*\right) + \lambda\left(\mathbb{E}Y^2 - \mathbb{E}(Y^*)^2\right)
$$

for all $Y \in E$, and therefore, $V^{\lambda,\mu}(Y^*) \le V^{\lambda,\mu}(Y)$. □

Proof of Theorem 5.1
Define

$$
R_n(x) := \sum_{i=n+1}^{N} \mu(x_{i-1}-x_i)^2\tilde{\eta}_i + \lambda(x_{i-1}-x_i)^4\tilde{\eta}_i^2 + \lambda x_i^2\Delta t\sigma_i^2 + 2\lambda(x_{i-1}-x_i)^2\tilde{\eta}_i \sum_{j>i}(x_{j-1}-x_j)^2\tilde{\eta}_j
$$

and note that

$$
\mathbb{E}_n^v[R_n(x)] = \mathbb{E}_n^v[\mu Q_n(x) + \lambda Q_n^2(x)].
$$

So one has

$$
J_{N-1}^{v}(x_{N-1}) = \mu x_{N-1}^2 \mathbb{E}_{N-1}^{v}[\tilde{\eta}_N] + \lambda x_{N-1}^4 \mathbb{E}_{N-1}^{v}[\tilde{\eta}_N^2]
$$

and

$$
\begin{aligned}
J^v_{n-1}(x_{n-1}) = & \min_{x\in\mathcal{R}_{n-1}(x_{n-1})} \mu(x_{n-1}-x_n)^2\mathbb{E}^v_{n-1}\left[\tilde\eta_n\right] + \lambda(x_{n-1}-x_n)^4\mathbb{E}^v_{n-1}\left[\tilde\eta_n^2\right] \\
& +\lambda x_n^2\Delta t\mathbb{E}^v_{n-1}\left[\sigma_n^2\right] + 2\lambda(x_{n-1}-x_n)^2\mathbb{E}^v_{n-1}\left[\tilde\eta_n\sum_{j=n+1}^{N}(x_{j-1}-x_j)^2\tilde\eta_j\right] + \mathbb{E}^v_{n-1}\left[R_n(x)\right] \\
= & \min_{0\le x_n\le x_{n-1}} \mu(x_{n-1}-x_n)^2\mathbb{E}^v_{n-1}\left[\tilde\eta_n\right] + \lambda(x_{n-1}-x_n)^4\mathbb{E}^v_{n-1}\left[\tilde\eta_n^2\right] \\
& +\lambda x_n^2\Delta t\mathbb{E}^v_{n-1}\left[\sigma_n^2\right] + 2\lambda(x_{n-1}-x_n)^2\mathbb{E}^v_{n-1}\left[\tilde\eta_n\sum_{j=n+1}^{N}(x_{j-1}-x_j)^2\tilde\eta_j\right] \\
& +\sum_{w\in V} p^{vw}_{n-1}J^w_n(x_n), \quad n\le N-1.
\end{aligned}
$$

$x^*\in\mathcal{R}$ is optimal if in every step x^*_n realizes the minimum for $x_{n-1}=x^*_{n-1}$. Moreover, if $\mu\ge 0$, then $\mathbb{E}^v_0[\mu Q(x)+\lambda Q(x)^2]$ is strictly convex in x, and the optimal strategy $x^*\in\mathcal{R}$ is unique. $\square$

Proof of Theorem 5.3
It is clear that

$$
J^v_{N-1}(h,x_{N-1}) = (\mu+2\lambda h)x^2_{N-1}\mathbb{E}^v_{N-1}[\tilde\eta_N] + \lambda x^4_{N-1}\mathbb{E}^v_{N-1}[\tilde\eta^2_N].
$$

Moreover, for $n\le N-1$,

$$
\begin{aligned}
J^v_{n-1}(h,x_{n-1}) &= \min_{x\in\mathcal{A}_{n-1}(x_{n-1})}\mathbb{E}^v_{n-1}\left[(\mu+2\lambda h)Q_{n-1}(x)+\lambda Q^2_{n-1}(x)\right] \\
&= \min_{x\in\mathcal{A}_{n-1}(x_{n-1})}\mathbb{E}^v_{n-1}\left[(\mu+2\lambda h)(u_n+Q_n(x))+\lambda u_n^2+2\lambda u_nQ_n(x)+\lambda Q^2_n(x)\right],
\end{aligned}
$$

where $u_n=(x_{n-1}-x_n)^2\tilde\eta_n - x_n\sigma_n\xi_n$. So

$$
\begin{aligned}
J^v_{n-1}(h,x_{n-1}) = & \min_{0\le x_n\le x_{n-1}}(\mu+2\lambda h)(x_{n-1}-x_n)^2\mathbb{E}^v_{n-1}[\tilde\eta_n] + \lambda(x_{n-1}-x_n)^4\mathbb{E}^v_{n-1}[\tilde\eta^2_n] \\
& +\lambda x_n^2\Delta t\mathbb{E}^v_{n-1}[\sigma_n^2] + \mathbb{E}^v_{n-1}\left[(\mu+2\lambda h+2\lambda u_n)Q_n(x)+\lambda Q_n^2(x)\right] \\
= & \min_{0\le x_n\le x_{n-1}}(\mu+2\lambda h)(x_{n-1}-x_n)^2\mathbb{E}^v_{n-1}[\tilde\eta_n] + \lambda(x_{n-1}-x_n)^4\mathbb{E}^v_{n-1}[\tilde\eta^2_n] \\
& +\lambda x_n^2\Delta t\mathbb{E}^v_{n-1}[\sigma_n^2] + \sum_{w\in V}p^{vw}_{n-1}\int_{\mathbb{R}} J^w_n\left(h+(x_{n-1}-x_n)^2(w_2-c/2)-x_nw_1\sqrt{\Delta t}\xi, x_n\right)\rho(\xi)d\xi,
\end{aligned}
$$

and $x^*\in\mathcal{A}$ is optimal if in every step, x^*_n realizes the minimum for $h=h_{n-1}(x^*)$ and $x_{n-1}=x^*_{n-1}$. Finally, for $\mu\ge 0$, $\mathbb{E}^v_0[\mu Q(x)+\lambda Q(x)^2]$ is strictly convex in x, and the optimal strategy $x^*\in\mathcal{A}$ is unique. $\square$

Acknowledgements

We thank an anonymous referee for helpful comments. The first author gratefully acknowledges financial support from NSF Grant DMS-0642361.

References

[1] Alfonsi, A. and Schied, A. (2012). Capacitary measures for completely monotone kernels via singular control. *SIAM Journal on Control and Optimization* 51(2), 1758–1780.

[2] Almgren, R. (2012). Optimal trading with stochastic liquidity and volatility. *SIAM Journal on Financial Mathematics* 3(1), 163–181.

[3] Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk* 3, 5–39.

[4] Almgren, R. and Lorenz, J. (2007). Adaptive arrival price. *Trading* Vol. 2007, No. 1, 59–66.

[5] Bayraktar, E. and Ludkovski, M. (2011). Optimal trade execution in illiquid markets. *Mathematical Finance* 21, 681–701.

[6] Bertsimas, D. and Lo, A. (1998). Optimal control of execution costs. *Journal of Financial Markets* 1, 1–50.

[7] Forsyth, P.A. (2011). A Hamilton Jacobi Bellman approach to optimal trade execution. *Applied Numerical Mathematics* 61, 241–265.

[8] Gatheral, J. and Schied, A. (2011). Optimal trade execution under geometric brownian motion in the Almgren and Chriss framework. *International Journal of Theoretical and Applied Finance* 14(3), 353–368.

[9] Gatheral, J. and Schied, A. (2013). Dynamical models of market impact and algorithms for order execution. *In: Handbook on Systemic Risk (eds.: J.P. Fouque and J.A. Langsam), Cambridge University Press,* 579–602.

[10] Gatheral, J., Schied, A., and Slynko, A. (2012). Transient linear price impact and Fredholm integral equations. *Mathematical Finance* 22, 445–474.

[11] Lorenz, J. and Almgren, R. (2011). Mean-variance optimal adaptive execution. *Applied Mathematical Finance* 18(5), 395–422.

[12] Moazeni, S., Coleman, T., Li, Y. (2013). Optimal execution under jump models for uncertain price impact. *Journal of Computational Finance* 4, 35–78.

[13] Obizhaeva, A. and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets* 16, 1–32.

[14] Schied, A., Schöneborn, T., Tehranchi, M. (2010). Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance* 17(6), 471–489.

[15] Walia, N. (2006). *Optimal Trading: Dynamic Stock Liquidation Strategies.* Senior Thesis. Princeton University. Princeton, NJ.