

Comparison between the Mean Variance optimal and the Mean Quadratic Variation optimal trading strategies *

S. T. Tse [†] P.A. Forsyth [‡] J.S. Kennedy [§] H. Windcliff [¶]

November 2, 2011

Abstract

We compare optimal liquidation policies in continuous time in the presence of trading impact using numerical solutions of Hamilton Jacobi Bellman (HJB) partial differential equations (PDE). In particular, we compare the time-consistent mean-quadratic-variation strategy with the time-inconsistent (pre-commitment) mean-variance strategy. We show that the two different risk measures lead to very different strategies and liquidation profiles. In terms of the optimal trading velocities, the mean-quadratic-variation strategy is much less sensitive to changes in asset price and varies more smoothly. In terms of the liquidation profiles, the mean-variance strategy is much more variable, although the mean liquidation profiles for the two strategies are surprisingly similar. On a numerical note, we show that using an interpolation scheme along a parametric curve in conjunction with the semi-Lagrangian method results in significantly better accuracy than standard axis-aligned linear interpolation. We also demonstrate how a scaled computational grid can improve solution accuracy.

Keywords: optimal trading, mean variance, pre-commitment, mean quadratic variation, time-consistent, arrival price, implementation shortfall, HJB PDE, interpolation, scaled-grid

1 Introduction

Algorithmic trade execution has become a standard technique for institutional market players in recent years, particularly in the equity market where electronic trading is most prevalent. A trade execution algorithm typically seeks to execute a trade decision optimally upon receiving inputs from a human trader.

A common form of optimality criterion seeks to strike a balance between minimizing pricing impact and minimizing timing risk. For example, in the case of selling a large number of shares, a fast liquidation will cause the share price to drop, whereas a slow liquidation will expose the seller to timing risk due to the stochastic nature of the share price.

Several approaches have been suggested in the literature to quantify the minimization of pricing impact and timing risk. The first, and perhaps the most intuitive, approach maximizes the expected revenues while minimizing a risk criterion, for example, variance [15, 14], quadratic variation [2], or value-at-risk (VaR) [18]. Another approach maximizes the expected value of a utility function of revenues, for example, a power-law

*This work was supported by the Natural Sciences and Engineering Research Council of Canada, and by a Morgan Stanley Equity Market Microstructure Research Grant. The views expressed herein are solely those of the authors, and not those of any other person or entity, including Morgan Stanley. Morgan Stanley is not responsible for any errors or omissions. Nothing in this article should be construed as a recommendation by Morgan Stanley to buy or sell any security of any kind.

[†]David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 stse@uwaterloo.ca

[‡]David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 paforsyt@uwaterloo.ca

[§]Morgan Stanley, 1585 Broadway, New York, NY 10036, Shannon.Kennedy@MorganStanley.com

[¶]Morgan Stanley, 1585 Broadway, New York, NY 10036, Heath.Windcliff@MorganStanley.com

31 function or an exponential function [19, 30, 27]. The third approach minimizes the expected execution cost
 32 [8]. All these three approaches model the asset price process in the presence of pricing impact. Yet another
 33 approach, which is somewhat tangential to the above methodologies, minimizes the expected execution cost
 34 by modelling the dynamic distribution of bid and ask orders in a limit order book [28, 1].

35 In this paper we focus on maximizing revenue while minimizing a risk measure. Maximizing revenue is
 36 also know as minimizing implementation shortfall relative to the arrival (pre-trade) price, which is a popular
 37 approach in industry. More specifically, we compare the pre-commitment mean-variance strategy [15, 7, 26]
 38 with the mean-quadratic-variation strategy [16, 12]. We assume that trading takes place continuously at a
 39 finite rate, as in [4, 3]. We note that the risk-criteria in the seminal paper [2] was previously thought to
 40 be variance but is actually quadratic variation, as shown in [16]. This is also discussed in [25]. Therefore,
 41 it is interesting to investigate how suboptimal the mean-quadratic-variation strategy is in terms of mean-
 42 variance efficiency, and, conversely, the question of how suboptimal the mean-variance strategy is in terms
 43 of mean-quadratic-variation efficiency.

44 The main contributions of this article are:

- 45 • We show that for the same variance, the mean-quadratic-variation strategy can have a significantly
 46 suboptimal expected value compared to the mean-variance strategy. For the same quadratic-variation,
 47 the mean-variance strategy can have significantly suboptimal expected value compared to the mean-
 48 quadratic-variation strategy. If one wants to strike the middle ground of balancing both variance and
 49 quadratic-variation, the mean-variance strategy seems to be preferable.
- 50 • We show that the mean-variance strategy is much more sensitive to changes in the asset price than
 51 the mean-quadratic-variation strategy. Consequently, the trading profile of the mean-variance strategy
 52 is much more variable. The mean trading profiles of the two strategies, however, turn out to be
 53 surprisingly similar.
- 54 • We have improved our numerical method in [15] so that results for very challenging parametric cases can
 55 be computed using reasonable time and memory. In particular, we improve the method of interpolation
 56 at the foot of the characteristic in the semi-Lagrangian discretization of HJB PDEs. We also construct a
 57 scaled computational grid so that fewer grid nodes are needed to obtain accurate results. The improved
 58 method also guarantees convergence to the viscosity solution.
- 59 • The mean-variance formulation of the optimal liquidation problem is known to be ill-posed [15]. More
 60 specifically, many similar strategies can give rise to nearly the same efficient frontier. In this paper
 61 we analyze in detail this ill-posedness from both a mathematical and a computational perspective. In
 62 particular, we highlight the numerical challenges created by such ill-posedness and demonstrate that
 63 the choice of interpolation method can be critical.

64 2 Optimal Execution

65 Let

$$\begin{aligned}
 P &= B + AS = \text{Portfolio,} \\
 S &= \text{Price of the underlying risky asset,} \\
 B &= \text{Balance of risk free bank account,} \\
 A &= \text{Number of shares of underlying asset.}
 \end{aligned}$$

66 The optimal execution problem over $t \in [0, T]$ has the initial condition

$$S(0) = s_{init}, B(0) = 0, A(0) = \alpha_{init}. \tag{2.1}$$

67 If $\alpha_{init} > 0$, the trader is liquidating a long position (selling). If $\alpha_{init} < 0$, the trader is liquidating a short
 68 position (buying). In this article, for definiteness, we consider the selling case. At $t = T$,

$$S = S(T), B = B(T), A = A(T) = 0, \tag{2.2}$$

69 where $B(T)$ is the cash generated by selling shares and investing in the risk free bank account B , with a final
70 liquidation at $t = T^-$ to ensure that $A(T) = 0$. The objective of optimal execution is to maximize $B(T)$,
71 while at the same time minimizing a certain risk measure. The two risk measures we consider in this paper,
72 namely variance and quadratic-variation, will be discussed in the next two sections.

73 In this paper, we consider Markovian trading strategies $v(\cdot)$ that specify a trading rate v as a function
74 of the current state, i.e. $v(\cdot) : (S(t), B(t), A(t), t) \mapsto v = v(S(t), B(t), A(t), t)$. Note that in using the
75 shorthand notations $v(\cdot)$ for the mapping, and v for the value $v = v(S(t), B(t), A(t), t)$, the dependence of v
76 on the current state is implicit.

77 By definition,

$$dA(t) = v dt. \quad (2.3)$$

78 We assume that due to temporary price impact, selling shares at the rate v at the market price $S(t)$ gives
79 an execution price $S_{exec}(v, t) \leq S(t)$. It follows that

$$dB(t) = (rB(t) - vS_{exec}(v, t))dt \quad (2.4)$$

80 where r is the risk free rate.

81 We suppose that the market price of the risky asset S follows a Geometric Brownian Motion (GBM),
82 where the drift term is modified due to the permanent price impact of trading [4]:

$$dS(t) = (\eta + g(v))S(t) dt + \sigma S(t) d\mathbb{W}(t),$$

η is the drift rate,
 $g(v)$ is the permanent price impact function,
 σ is the volatility,
 $\mathbb{W}(t)$ is a Wiener process under the real world measure.

(2.5)

83 2.1 Trading impact function

84 We assume the temporary price impact scales linearly with the asset price, i.e.

$$S_{exec}(v, t) = f(v)S(t), \quad (2.6)$$

85 where

$$f(v) = (1 + \kappa_s \operatorname{sgn}(v)) \exp[\kappa_t \operatorname{sgn}(v)|v|^\beta],$$

$\kappa_s =$ the bid-ask spread parameter,
 $\kappa_t =$ the temporary price impact factor,
 $\beta =$ the price impact exponent.

(2.7)

86 Note that we assume $\kappa_s < 1$, so that $S_{exec}(v, t) \geq 0$, regardless of the magnitude of v . For various studies
87 which suggest the form (2.7), see [4, 24, 29].

88 The permanent price impact function $g(v)$ is assumed to be of the form

$$g(v) = \kappa_p v,$$

$\kappa_p =$ the permanent price impact factor.

89 As explained in [16], this form of permanent price impact function eliminates the possibilities of round-trip
90 arbitrage [4, 20].

2.2 Definition of liquidation value

Given the state $(S(T^-), B(T^-), A(T^-))$ at the instant $t = T^-$ before the end of the trading horizon, we have one final liquidation (if necessary) so that the number of shares owned at $t = T$ is $A(T) = 0$. The liquidation value $B(T)$ after this final trade is defined to be

$$\begin{aligned} B(T) &= B(T^-) + \lim_{v \rightarrow -\infty} A(T^-) S_{exec}(v, T^-) \\ &= B(T^-) \end{aligned} \tag{2.8}$$

Definition (2.8) in effect penalizes the strategy if $A(T^-) \neq 0$, so that the optimal algorithm forces the liquidation profile towards $A(T^-) = 0$. In our case, the penalty is such that the shares $A(T^-)$ are simply discarded.¹

3 Mean-Variance Strategy

We review here the pre-commitment mean-variance strategy, as discussed in [15].

3.1 Objective functional and optimal strategy

To simplify notations, we define $x = (s, b, \alpha) = (S(t), B(t), A(t))$ for a space state. Now we specify the pre-commitment mean variance formulation as follows. For all possible states (x, t) and a fixed risk aversion parameter λ , define the family of objective functionals

$$\mathcal{F}_\lambda = \left\{ J_\lambda^{x,t}(v(\cdot)) : v(\cdot) \mapsto E_{v(\cdot)}^{x,t}[B(T)] - \lambda Var_{v(\cdot)}^{x,t}[B(T)] \right\}, \tag{3.1}$$

where $E_{v(\cdot)}^{x,t}[\cdot]$ is the expectation, and $Var_{v(\cdot)}^{x,t}[\cdot]$ is the variance, conditional on the state (x, t) and the control $v(\cdot) : (x, t) \mapsto v = v(x, t)$.

Note that in the notation of (3.1), the members (functionals) in the family \mathcal{F}_λ have different initial states (x, t) but the same λ . For a given initial state (x, t) , we will henceforth use the notation $v_{x,t,\lambda}^*(\cdot)$ to denote the optimal policy that maximizes the corresponding functional, i.e. $J_\lambda^{x,t}(v(\cdot))$.

Let $(x_0, t = 0) = (s_{init}, 0, \alpha_{init}, 0)$ be the initial state. The optimal policy $v_{x_0,0,\lambda}^*(\cdot)$ is termed the pre-commitment mean variance optimal strategy [7, 11].

3.2 Time-inconsistency of optimal strategies

The optimal strategies in the pre-commitment mean variance formulation are time-inconsistent in the following sense. Let (x_1, t_1) be some state at time t_1 and $v_{x_1,t_1,\lambda}^*(\cdot)$ be the corresponding optimal policy. Let (x_2, t_2) be some other state at time $t_2 > t_1$ and $v_{x_2,t_2,\lambda}^*(\cdot)$ be the corresponding optimal policy.

We have time-inconsistency in the sense that

$$v_{x_1,t_1,\lambda}^*(x', t') \neq v_{x_2,t_2,\lambda}^*(x', t') ; t' \geq t_2 . \tag{3.2}$$

The time-inconsistency (3.2) is considered as unnatural by some authors [7] and creates computational difficulties. More specifically, a dynamic programming principle cannot be directly applied to solve this problem.

¹In actual implementation, we would replace $\lim_{v \rightarrow -\infty}$ by a finite $v_{\min} \ll 0$ in the PDE initial condition. Also, in the case of liquidating a short position (buying), which is not considered in this paper, equation (2.8) would be defined as $B(T) = B(T^-) + \lim_{v \rightarrow \infty} A(T^-) S_{exec}(v, T^-)$, and we would replace $\lim_{v \rightarrow \infty}$ by a finite $v_{\max} \gg 0$ in implementation.

119 **3.3 Alternative Formulation**

120 To solve for $v_{x_0,0,\lambda}^*(\cdot)$, we follow the method in [10, 6, 17, 23]. For each fixed initial state (x, t) and risk
 121 aversion parameter λ , the optimal control $v_{x,t,\lambda}^*(\cdot)$ that maximizes (the functionals in the family) (3.1) are
 122 also the optimal controls that minimize

$$\tilde{\mathcal{F}}_\lambda = \left\{ \tilde{J}_\lambda^{x,t}(v(\cdot)) : v(\cdot) \mapsto E_{v(\cdot)}^{x,t} \left[(B(T) - \frac{\gamma(x, t; \lambda)}{2})^2 \right] \right\}, \quad (3.3)$$

123 for some

$$\gamma = \gamma(x, t; \lambda) \geq 0. \quad (3.4)$$

124 Note that under the alternative formulation (3.3), the optimal strategies $v_{x,t,\lambda}^*(\cdot)$ are also time-inconsistent,
 125 due to the dependence of γ on the initial state (x, t) , i.e. (3.4); see [11] for more discussion on this.

126 Consequently, for the particular initial state $(x_0, 0)$ and an arbitrary constant $\gamma_0 \geq 0$, the optimal control
 127 $v_{x_0,0,\gamma_0}^*(\cdot)$ that minimizes the member $\hat{J}_{\gamma_0}^{x_0,0}$ in the family

$$\hat{\mathcal{F}}_\gamma = \left\{ \hat{J}_\gamma^{x,t}(v(\cdot)) : v(\cdot) \mapsto E_{v(\cdot)}^{x,t} \left[(B(T) - \frac{\gamma}{2})^2 \right] \right\}, \quad (3.5)$$

128 is also the pre-commitment optimal control $v_{x_0,0,\lambda_0}^*(\cdot)$ for a certain value of λ_0 such that $\gamma_0 = \gamma(x_0, 0, \lambda_0)$.

129 The benefit of reformulating (3.1) as (3.5) is that the dynamic programming principle can be applied to
 130 (3.5) to solve for $v_{x_0,0,\gamma_0}^*(\cdot)$, since γ is a constant in (3.5).

131 Varying γ_0 over $[0, \infty)$ gives the strategies that trace out a variance-minimizing frontier in the expected
 132 value, standard deviation plane.

133 **4 Mean-Quadratic-Variation Strategy**

134 Quadratic variation has been used as an approximation of variance in the algorithmic trading literature
 135 [2, 3, 16]. This approximation, however, can be poor when the trading impact is relatively large, as will
 136 be illustrated in the current paper. Instead of using quadratic variation to approximate variance, it is
 137 conceptually simpler to regard quadratic variation as an alternative risk measure.

138 **4.1 Quadratic variation as a risk measure**

139 Formally, the quadratic variation risk measure is defined as

$$E \left[\int_t^T (A(t') dS(t'))^2 \right]. \quad (4.1)$$

140 Informally, the risk measure definition (4.1) can be interpreted as the quadratic variation of the portfolio
 141 value process as follows: by expanding the square of $dP(t') = dB(t') + d(A(t')S(t'))$ and ignoring higher-order
 142 terms, we have

$$\int_t^T (A(t') dS(t'))^2 = \int_t^T (dP(t'))^2, \quad (4.2)$$

143 when the trading velocity process $v(t)$ is bounded.

144 4.2 Objective functional and value function

145 Now we specify the mean quadratic variation formulation as follows. For a fixed initial point $(s, \alpha, t) =$
 146 $(S(t), A(t), t)$ where $t < T$ with $B(t) = 0$, trading strategy $v(\cdot)$, and risk aversion parameter $\tilde{\lambda}$, we define the
 147 objective functional

$$J(s, \alpha, t, v(\cdot); \tilde{\lambda}) = E_{v(\cdot)}^{s, \alpha, t} \left[B(T) \right] - \tilde{\lambda} E_{v(\cdot)}^{s, \alpha, t} \left[\int_t^T (A(t') dS(t'))^2 \right] \quad (4.3)$$

148 where

$$B(T) = \int_t^{T^-} e^{r(T-t')} (-v S_{exec}(v, t')) dt' + \lim_{v \rightarrow -\infty} A(T^-) S_{exec}(v, T^-) \quad (4.4)$$

149 and $E_{v(\cdot)}^{s, \alpha, t}[\cdot]$ is the conditional expectation at the initial point (s, α, t) using the control $v(\cdot)$.

150 The value function \hat{V}^{MQV} is defined as

$$\hat{V}^{MQV}(s, \alpha, t; \tilde{\lambda}) = \sup_{v(\cdot)} J(s, \alpha, t, v(\cdot); \tilde{\lambda}). \quad (4.5)$$

151 For a given initial state (s, α, t) , we will henceforth use the notation $v_{s, \alpha, t, \tilde{\lambda}}^*(\cdot)$ to denote the optimal policy
 152 that maximizes the corresponding functional, i.e. $J(s, \alpha, t, v(\cdot); \tilde{\lambda})$.

153 4.3 Time Consistency of the optimal strategies

154 Let (s_1, α_1, t_1) be some state at time t_1 and $v_{s_1, \alpha_1, t_1, \tilde{\lambda}}^*(\cdot)$ be the corresponding optimal strategy. Let (s_2, α_2, t_2)
 155 be some other state at time $t_2 > t_1$ and $v_{s_2, \alpha_2, t_2, \tilde{\lambda}}^*(\cdot)$ be the corresponding optimal strategy.²

156 Since the optimal controls satisfy the Bellman's principle of optimality as shown in [16], it follows that
 157 the optimal controls of (4.5) are *time consistent* in the sense that for the same state (s', α', t') at a later time
 158 $t' > t_2$,

$$v_{s_1, \alpha_1, t_1, \tilde{\lambda}}^*(s', \alpha', t') = v_{s_2, \alpha_2, t_2, \tilde{\lambda}}^*(s', \alpha', t'); \quad t' \geq t_2. \quad (4.6)$$

159 Hence, dynamic programming can be directly applied to this problem.

160 In certain special cases, it is known that strategy $v_{s, \alpha, t, \tilde{\lambda}}^*(\cdot)$ is equivalent to a time consistent mean variance
 161 strategy [11, 31]. Hence $v_{s, \alpha, t, \tilde{\lambda}}^*(\cdot)$ can be viewed as a natural time consistent strategy. In addition, as shown
 162 in [16], if arithmetic Brownian motion is assumed, the optimal strategy $v_{s, \alpha, t, \tilde{\lambda}}^*(\cdot)$ is actually identical to the
 163 optimal strategy in [2].

164 Let $V^{MQV}(s, \alpha, \tau; \tilde{\lambda}) = \hat{V}^{MQV}(s, \alpha, t = T - \tau; \tilde{\lambda})$. The derivation in [16] shows that V^{MQV} satisfies the
 165 HJB equation

$$V_\tau^{MQV} = \eta s V_s^{MQV} + \frac{\sigma^2 s^2}{2} V_{ss}^{MQV} - \tilde{\lambda} e^{2r\tau} \alpha^2 s^2 \sigma^2 + \sup_v [e^{r\tau} (-vf(v))s + g(v)s V_s^{MQV} + v V_\alpha^{MQV}]. \quad (4.7)$$

166 We note that both the value function V^{MQV} and strategy $v(\cdot)$ for the mean-quadratic-variation problem
 167 (4.5) is independent of the current bank account balance $B(t)$. In particular $v = v(s, \alpha, \tau)$.

168 The reader is referred to [16] for details about the numerical method used to solve equation (4.7).

²Note that while the initial point is changed from (s_1, α_1, t_1) to (s_2, α_2, t_2) , the risk aversion level $\tilde{\lambda}$ is kept constant.

169 4.4 Arithmetic Brownian Motion

170 Under the arithmetic Brownian motion approximation and the additional assumptions of zero drift, zero
171 interest rate, unbounded control and linear price impact function detailed in [16], the optimal trading strategy
172 has the analytic solution

$$v(\alpha, \tau) = -\alpha K \coth(K\tau), \quad (4.8)$$

173 where $K = \sqrt{\tilde{\lambda}\sigma^2 s_{init}/\kappa_t}$. Note that the optimal strategy is independent of the spot price s .

174 As noted in [16], the strategy (4.8) results in an efficient frontier that is extremely close to the true
175 mean-quadratic-variation efficient frontier computed assuming geometric Brownian motion. We also note
176 that (4.8) is the same strategy as used in [2].

177 The work in [2] is extended in [5] to consider nonlinear trading impacts similar to the form (2.7), and
178 semi-explicit solutions are obtained.

179 5 Comparison between the two strategies

180 5.1 Risk measure

181 The pre-commitment mean variance strategy is optimal in the following sense [25]. Suppose we carry out
182 many thousands of trades. We then examine the post-trade data, and determine the realized mean return
183 and the standard deviation. Assuming that the modeled dynamics very closely match the dynamics in the
184 real world, the optimal pre-commitment strategy would result in the largest realized mean return, for a given
185 standard deviation, compared to any other possible strategy.

186 From the interpretation (4.2), minimizing quadratic variation clearly corresponds to minimizing volatility
187 in the portfolio value process. The definition (4.1) shows that quadratic variation takes into account the
188 trading trajectory $A(t')$ over the whole trading horizon. This is in contrast with using variance ($Var[B(T)]$)
189 as a risk measure, which is independent of the trading trajectory $A(t')$ given the end result $B(T)$. We note
190 that the idea of using quadratic variation as a risk measure was first suggested in [12].

191 5.2 Uniqueness and smoothness

192 In mean-variance optimization, many similar strategies can give rise to almost the same efficient frontier
193 (near ill-posedness). This can be advantageous as it permits more flexibility in executing the trade. On
194 the other hand, this creates difficulties in obtaining a smoothly varying optimal strategy, as demonstrated
195 and explained in [15] and the current paper. In our experience, these issues do not arise in mean-quadratic-
196 variation optimization.

197 6 HJB Formulation: Mean Variance

198 6.1 Change of Variable

199 At first glance it seems necessary to solve the problem (3.5) for each value of γ separately. Fortunately, this
200 can be avoided by a change of variable. Define the new variable \mathcal{B} by

$$\mathcal{B}(0) = \frac{-\gamma e^{-rT}}{2} \leq 0, \quad d\mathcal{B}(t) = r\mathcal{B}(t)dt - vS_{exec}(v, t)dt \quad (6.1)$$

201 it is easy to see that

$$\mathcal{B}(t) = B(t) - \frac{\gamma e^{-r(T-t)}}{2}, \quad \mathcal{B}(T) = B(T) - \frac{\gamma}{2}. \quad (6.2)$$

202 Since equation (6.1) has the same form as equation (2.4), solving for $v_{x,0,\gamma}^*(\cdot)$ in (3.5) is equivalent to solving
 203 for the optimal control in

$$\inf_{v(\cdot)} E_{v(\cdot)}^{s,\hat{b},\alpha,t=0}[\mathcal{B}(T)^2]. \quad (6.3)$$

204 This change of variable is very convenient in the PDE context because the solutions corresponding to different
 205 values of γ can be determined by examining the PDE solutions for different values of \mathcal{B} at $t = 0$. Therefore,
 206 we only need to solve the problem (6.3) once to obtain the entire variance minimizing frontier [15].

207 6.2 Definitions

208 Let $\tau = T - t$ be the backward time. Define the value function V by

$$V = V(s, \hat{b}, \alpha, \tau) = \inf_{v(\cdot)} E_{v(\cdot)}^{s,\hat{b},\alpha,T-\tau}[\mathcal{B}(T)^2]. \quad (6.4)$$

209 We also define the differential operator \mathcal{L} by

$$\mathcal{L}V = \frac{\sigma^2 s^2}{2} V_{ss} + \eta s V_s, \quad (6.5)$$

210 and Lagrangian derivative $\frac{D}{D\tau}(v)$ by

$$\frac{DV}{D\tau}(v) = V_\tau - V_s g(v)s - V_{\hat{b}}(r\hat{b} - vf(v)s) - V_\alpha v, \quad (6.6)$$

211 which is the rate of change of V along the characteristic $s = s(\tau)$, $\hat{b} = \hat{b}(\tau)$, $\alpha = \alpha(\tau)$ defined by the trading
 212 velocity v through

$$\frac{ds}{d\tau} = -g(v)s, \quad \frac{d\hat{b}}{d\tau} = -(r\hat{b} - vf(v)s), \quad \frac{d\alpha}{d\tau} = -v. \quad (6.7)$$

213 6.3 PDE formulation

214 Following standard arguments, the optimal control is given by the solution to the nonlinear HJB equation

$$\mathcal{L}V - \max_{v \leq 0} \frac{DV}{D\tau}(v) = 0. \quad (6.8)$$

215 in the domain $\Omega = \{s \geq 0, \hat{b} \in \mathbf{R}, \alpha \geq 0, \tau > 0\}$.

216 In view of definition (2.8), the initial condition at $\tau = 0$ is

$$V(s, \hat{b}, \alpha, \tau = 0) = \mathcal{B}(T)^2 = (\hat{b} + as \lim_{v \rightarrow -\infty} f(v))^2. \quad (6.9)$$

217 Note that we forbid buying or holding a short position (when liquidating stock) in (6.8), i.e.

$$v(s, \hat{b}, \alpha, \tau) \leq 0, \quad v(s, \hat{b}, \alpha = 0, \tau) = 0, \quad (6.10)$$

218 At $s = 0$, equation (6.8) reduces to

$$\max_{v \leq 0} \left\{ V_\tau - r\hat{b}V_{\hat{b}} - vV_\alpha \right\} = 0. \quad (6.11)$$

219 Therefore, no boundary condition at $s = 0$ is needed.

220 At $\alpha = 0$, (6.10) causes equation (6.8) to reduce to

$$\frac{\sigma^2 s^2}{2} V_{ss} + \eta s V_s - V_\tau + r\hat{b}V_{\hat{b}} = 0. \quad (6.12)$$

221 Therefore, no boundary condition at $\alpha = 0$ is needed.

222 Solving the HJB PDE (6.4) gives the optimal control $v^*(\cdot)$. To obtain an efficient frontier, we proceed as
 223 follows. Define the value function U by

$$U = U(s, \boldsymbol{b}, \alpha, \tau) = E_{v^*(\cdot)}^{s, \boldsymbol{b}, \alpha, T-\tau}[\mathcal{B}(T)], \quad (6.13)$$

224 which satisfies the PDE

$$\mathcal{L}U - \frac{DU}{D\tau}(v^*) = 0, \quad U(s, \boldsymbol{b}, \alpha, \tau = 0) = \mathcal{B}(T) = \boldsymbol{b} + as \lim_{v \rightarrow -\infty} f(v). \quad (6.14)$$

225 Since $v^*(\cdot)$ has been determined, the PDE (6.14) is linear and inexpensive to solve.

226 Having solved for the value functions V and U , the variance-minimizing frontier can be obtained as
 227 described in section B.1. The mean-variance frontier is then obtained by a simple sorting procedure to
 228 eliminate suboptimal points.

229 7 Limiting case

230 For illustration purposes, consider a limiting case with extreme parameter values $\sigma = \kappa_t = 0$, and typical
 231 parameter values $r = \kappa_p = \kappa_s = \eta = 0$. Since the asset price is constant, problem (6.3) degenerates to the
 232 deterministic control problem of minimizing $\mathcal{B}(T)^2$. Moreover, since there is no pricing impact, $\mathcal{B}(T) = \alpha s + \boldsymbol{b}$
 233 with certainty. Consequently, the value function V is

$$V(s, \boldsymbol{b}, \alpha, \tau) = \inf_{v(\cdot)} E_{v(\cdot)}^{s, \boldsymbol{b}, \alpha, T-\tau}[\mathcal{B}(T)^2] = \mathcal{B}(T)^2 = (\alpha s + \boldsymbol{b})^2, \quad (7.1)$$

234 which can also be verified by direct substitution into the HJB equation (6.8), (6.9) as follows. First, note
 235 that the initial condition (6.9) is satisfied because $f(v) \equiv 1$. To verify (6.8), note that the parameter values
 236 yield the simplifications

$$\mathcal{L}V = 0, \quad \frac{DV}{D\tau}(v) = V_\tau + vsV_{\boldsymbol{b}} - vV_\alpha. \quad (7.2)$$

237 Substituting (7.1) into (7.2) gives

$$V_\tau = 0, V_{\boldsymbol{b}} = 2(\alpha s + \boldsymbol{b}), V_\alpha = 2s(\alpha s + \boldsymbol{b}) \implies \frac{DV}{D\tau}(v) \equiv 0 \text{ for all } v. \quad (7.3)$$

238 Since any admissible trading velocity v is optimal in this case, the problem of determining the optimal control
 239 v is completely ill-posed.

240 Although the above special case is degenerate, it explains what happens for realistic parametric cases.
 241 Indeed, in practical parametric cases, the values of $\sigma\sqrt{T}$ and κ_t are quite small and r has little effect.
 242 Although the actual trading velocity is highly dependent on $\sigma\sqrt{T}$ and κ_t , we expect that the actual value
 243 function will be only weakly dependent on these parameters.

244 7.1 Motivation for a parametric curve interpolation method

245 In using a semi-Lagrangian method [13, 9] to solve for the optimal velocity v , accurate interpolation at the
 246 foot of the characteristics is essential to achieving high accuracy [13, 9]. Since a monotone discretization
 247 scheme (which allows proof of convergence to the viscosity solution) is at most first-order accurate, we will
 248 deal exclusively with linear interpolation in this paper.

249 Let us consider again the special case in the previous section, where the value function V has the analytic
 250 solution

$$V(s, \boldsymbol{b}, \alpha, \tau) = (\alpha s + \boldsymbol{b})^2. \quad (7.4)$$

251 It is obvious that linear interpolation along each of the three coordinate axes is not exact, since the partial
 252 derivatives $V_{ss}, V_{\boldsymbol{b}\boldsymbol{b}}$ and $V_{\alpha\alpha}$ are all non-zero.

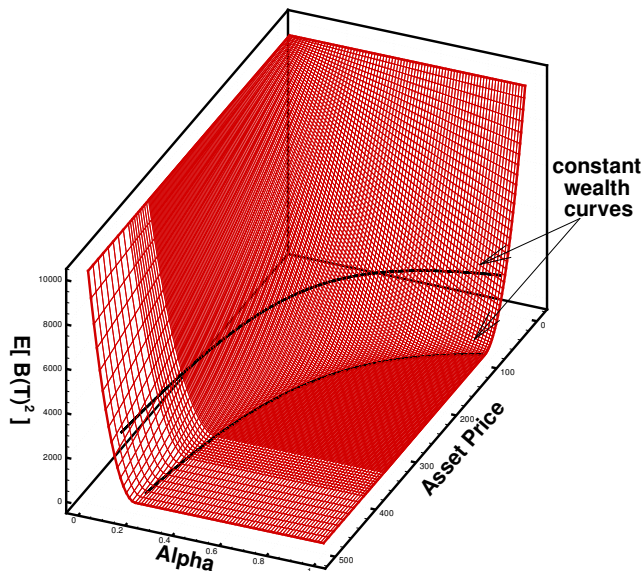


FIGURE 1: Value function at $t = 0$ and $b = -100$ for parametric Case 1 detailed in Table 1 and Table 2. Here we show two curves of approximately constant wealth $\{\alpha s + b = 0\}$ and $\{\alpha s + b = -50\}$, which correspond to the curves $\{\alpha = 100/s\}$ and $\{\alpha = 50/s\}$, respectively, since b is fixed at -100 . Note that the value function V is approximately $V \approx (\alpha + b)^2$ and changes rapidly normal to the lines of constant wealth.

253 Consider linear interpolation along the parametric curve of constant wealth ($\{\alpha s + b = \text{constant}\}$):

$$\frac{ds}{d\zeta} = 0, \quad \frac{db}{d\zeta} = vs, \quad \frac{d\alpha}{d\zeta} = -v, \quad (7.5)$$

254 for any fixed trading velocity v . Since V is constant along this line, linear interpolation along this parametric
 255 curve is exact.

256 As far as the general form for the value value function is concerned, equation (7.4) suggests that the value
 257 function for realistic parameter values should be slowly varying along curves of constant wealth ($\{\alpha s + b =$
 258 $\text{constant}\}$). In Figure 1 we show the value function computed at $t = 0$ (for fixed b) for a typical set of
 259 parameters. The dark curves in Figure 1 are curves of (approximately) constant wealth. Note that the curve
 260 which passes through $(\alpha = 1, s = 100)$ divides the value function into a region which is almost flat; and
 261 a region of rapidly increasing values. Interpolation normal to this curve will result in large errors, while
 262 interpolation along this curve will have relatively small errors.

263 To further improve accuracy, the scheme (7.5) is extended (for general parameters) in section B.4.2 in
 264 the appendix.

265 8 Numerical results

266 Our discussion on numerical results is organized as follows. First, we explain how we arrive at our parametric
 267 cases. Then, we demonstrate convergence by numerical experiments. Finally, we compare the efficient
 268 frontiers using the mean variance strategy and the mean quadratic variation strategy.

T	η	r	s_{init}	α_{init}	κ_p	κ_s	β	Action	v_{min}
1/250	0.0	0.0	100	1.0	0.0	0.0	1.0	Sell	-1000/T

TABLE 1: *Common parameters*

Case	σ	κ_t	Percentage of Daily Volume
1	1.0	2×10^{-6}	16.7%
2	0.2	2.4×10^{-6}	20.0%
3	0.2	6×10^{-7}	5.0%
4	0.2	1.2×10^{-7}	1.0%
5	0.2	2.4×10^{-8}	0.2%

TABLE 2: *Parametric cases*

269 8.1 Parametric cases

270 The parametric cases we consider are listed in Table 1 and Table 2. Case 1 corresponds to a high volatility
271 stock with low liquidity. Cases 2-5 correspond to a low volatility stock with various levels of liquidity. These
272 parameters can be related to typical daily volume traded. As described in Appendix A, we estimate that
273 $\kappa_t = 1.2 \times 10^{-7}$ corresponds to liquidating 1% of the daily volume traded of a typical large-cap liquid stock.
274 In view of our trading model and in particular the temporary trading impact function (2.7) with the choice
275 of $\beta = 1$, we can simulate liquidating $Y\%$ of the daily volume traded by keeping α_{init} unchanged at unity
276 and using $\kappa_t = (1.2 \times 10^{-7})Y$.

277 8.2 Convergence Analysis

278 To demonstrate convergence numerically, we compute the mean-variance frontier from the optimal mean
279 variance strategy using two methods. We explain both methods using the mean variance formulation as an
280 example. The methods for the mean quadratic variation formulation is analogous.

281 Before discussing the two methods, we note that both methods require the following same initial step:
282 solve for the optimal control $v^*(\cdot)$ in the HJB PDE (6.8).

283 8.2.1 The PDE method

284 In the PDE method, the mean variance efficient frontier is obtained from the value functions (6.8) and (6.14).
285 More specifically, the PDE method consists of the following steps.

- 286 1. Note that the value function V is computed when solving for $v^*(\cdot)$.
- 287 2. Compute the value function U by solving the linear PDE (6.14).
- 288 3. Construct the variance-minimizing frontier from U and V as described in Appendix B.1.
- 289 4. Eliminate suboptimal points to obtain the mean variance efficient frontier

290 8.2.2 The Hybrid (PDE-Monte Carlo) method

291 In the Hybrid method, Monte Carlo simulations are carried out using the optimal control $v^*(\cdot)$ to estimate
292 quantities of interest. More specifically, the Hybrid method consists of the following steps.

- 293 1. The optimal control $v^*(\cdot)$ (from solving the HJB PDE (6.8)) is an input.
- 294 2. Quantities of interest are estimated by Monte Carlo simulations, as detailed in Appendix C.

295 An advantage of the Hybrid method is that we can estimate quantities of interest that cannot be obtained
 296 in the PDE method. For example, it is important to understand how liquidation proceeds (in forward time)
 297 on average, which cannot be known from the PDE solutions directly. The Hybrid method also allows us
 298 to estimate both risk measures (variance and quadratic variation) of either the mean variance or the mean
 299 quadratic variation strategies, which may not be obtained directly from the value functions.

300 8.2.3 Computational grid

301 Tables 3 and 4 show the number of nodes and time steps used in the convergence study for the mean-variance
 302 strategy and the mean-quadratic-variation strategy, respectively. Note that only one node is needed in the b
 303 direction, since this variable can be eliminated using a similarity reduction (see section B.2 and [15]).³ The
 304 v node discretization is required in order to carry out a linear search to determine the optimal control [15].

305 Our parametric curve interpolation scheme (see section B.4.2 in the appendix for details) suggests that
 306 the number of s nodes should be significantly more than the number of α nodes, a consideration that is also
 307 confirmed by our numerical experiments.

308 Note also that the same time steps are used in both PDE calculation and Monte Carlo simulations, for
 309 each refinement level. For example, the frontiers labeled with “800 time steps” in Figure 2 use the time steps
 310 as specified as Refinement Level 2 in Table 3. Similarly for the frontiers labeled with “1600 time steps” and
 311 for the frontiers in other figures in the report.

312 8.2.4 Sample size in Monte Carlo simulations

313 To achieve small sampling error in Monte Carlo simulations, 400,000 simulations are performed for parametric
 314 case 1 and 100,000 simulations are performed for each of the other cases. As an example, the standard error
 315 in Figure 2(a) can be estimated as follows. To be more specific, consider a point on the frontier with the
 316 maximum standard deviation, which equals 3. Since this estimate of standard deviation of $B(T)$ is calculated
 317 using 400,000 samples, its standard error is approximately $3/\sqrt{400,000} \approx 0.0047$, which is negligible in Figure
 318 2(a). Similar calculations will show that the standard errors are negligible in other figures as well.

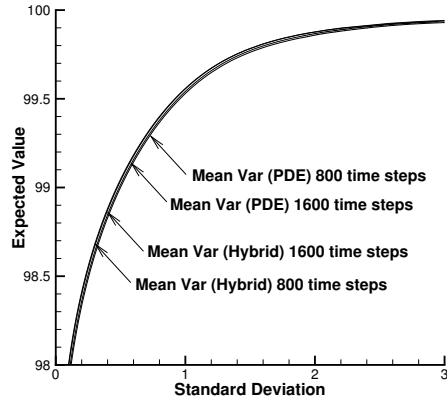
319 8.2.5 Comparison between the PDE method and the Hybrid method

320 For the mean variance optimal strategies, Figure 2 shows that the mean variance efficient frontiers computed
 321 by both the PDE method and the Hybrid method converge to the same frontier as the computational grid
 322 is refined.

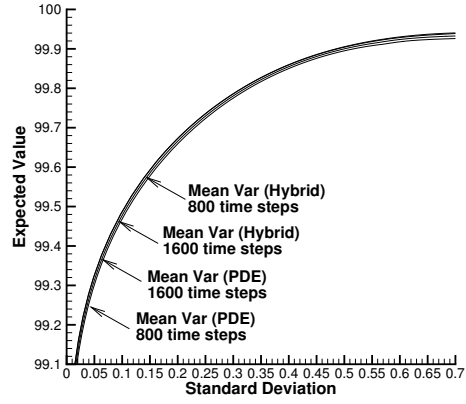
323 Similarly, for the mean quadratic variation optimal strategies, Figure 3 shows that the mean quadratic
 324 variation efficient frontiers computed by both the PDE method and the Hybrid method converge to the same
 325 frontier as the computational grid is refined.

326 Our numerical results demonstrate that the Hybrid frontiers in general converge faster to the limit solution
 327 than the PDE frontiers. This may seem counter-intuitive as the Monte Carlo simulations use the optimal
 328 trading strategies determined by the PDE method. Nevertheless, it is plausible that Monte Carlo simulations
 329 produce a better estimate of the expected value (or standard-deviation/quadratic-variation), which is what
 330 our numerical results suggests. Given their better accuracy, we will use the Hybrid frontiers to compare the
 331 mean-variance strategy with the mean-quadratic-variation strategy. Again, note that the optimal controls
 332 are always computed by solving the HJB PDEs.

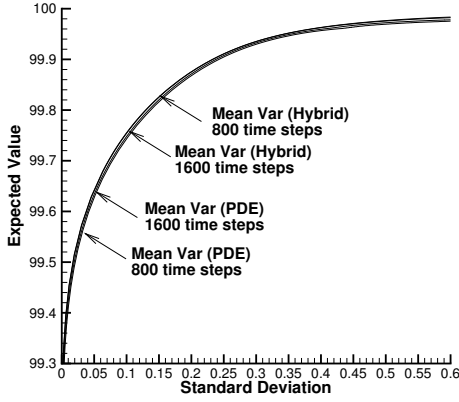
³For the mean-variance strategy, our numerical experiments suggest that the v grid needs to be fine near $t = 0$ (but not when t is larger) to obtain accurate estimate of the optimal v by linear search. In order to have a very fine v grid near $t = 0$ but a coarse v grid elsewhere, we perform 4 additional refinements to the v grid in the last few backward time steps in PDE solve. There is no such concern for the mean- quadratic-variation strategy because the optimal velocity is not determined by linear search (we use a one dimensional optimization method) and hence is not restricted to values in a discrete v grid.



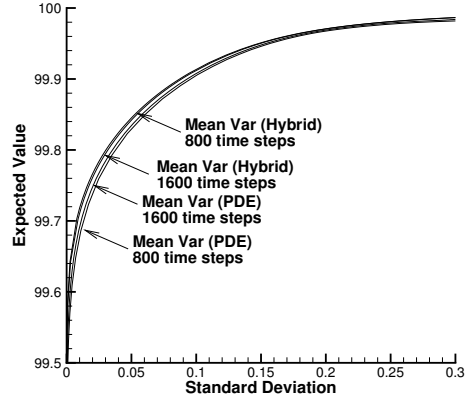
(a) $\sigma=1.0, \kappa_t = 2 \times 10^{-6}$



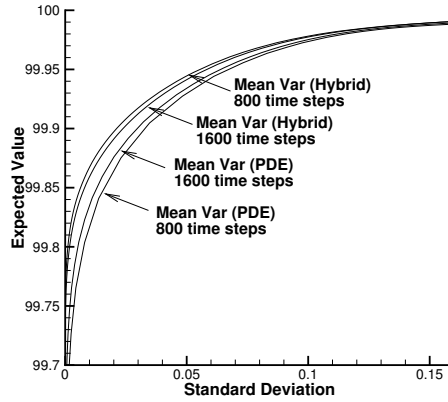
(b) $\sigma=0.2, \kappa_t = 2.4 \times 10^{-6}$



(c) $\sigma=0.2, \kappa_t = 6 \times 10^{-7}$

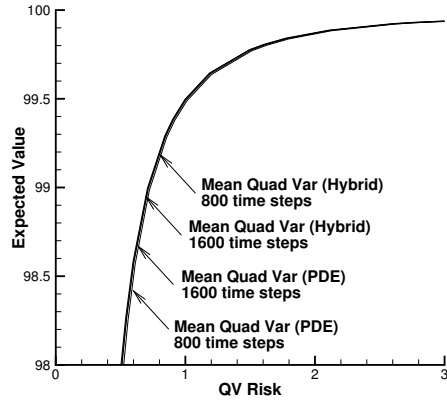


(d) $\sigma=0.2, \kappa_t = 1.2 \times 10^{-7}$

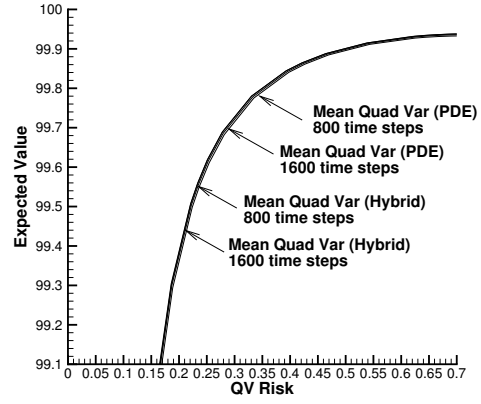


(e) $\sigma=0.2, \kappa_t = 2.4 \times 10^{-8}$

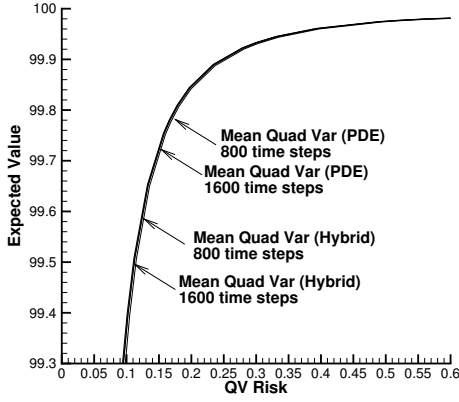
FIGURE 2: Mean variance strategy: convergence of frontiers in the PDE method and the Hybrid method. The frontiers labeled with PDE are obtained from the PDE value functions. The frontiers labeled with Hybrid are obtained from Monte Carlo simulations which use the optimal controls determined by solving the HJB equation (6.8).



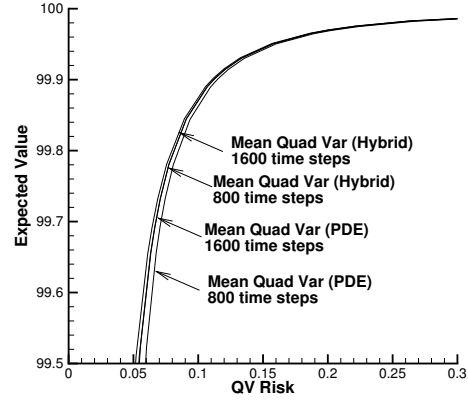
(a) $\sigma=1.0, \kappa_t = 2 \times 10^{-6}$



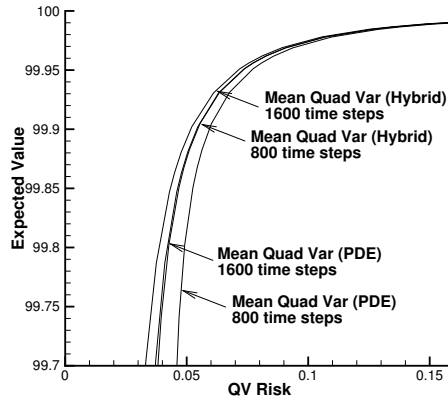
(b) $\sigma=0.2, \kappa_t = 2.4 \times 10^{-6}$



(c) $\sigma=0.2, \kappa_t = 6 \times 10^{-7}$



(d) $\sigma=0.2, \kappa_t = 1.2 \times 10^{-7}$



(e) $\sigma=0.2, \kappa_t = 2.4 \times 10^{-8}$

FIGURE 3: Mean quadratic variation strategy: convergence of frontiers in the PDE method and the Hybrid method. The frontiers labeled with PDE are obtained from the PDE value functions. The frontiers labeled with Hybrid are obtained from Monte Carlo simulations which use the optimal controls determined by solving the HJB equation (4.7).

Refinement Level	Timesteps	s nodes	b node	α nodes	v nodes
0	200	369	1	11	8
1	400	737	1	21	15
2	800	1473	1	41	29
3	1600	2945	1	81	57

TABLE 3: Grid and time step data for convergence studies for the mean variance strategy. The same time steps are used in both PDE calculation and Monte Carlo simulations. Note that there is only one b node because of the use of similarity reduction (see section B.2 and [15]).

Refinement Level	Timesteps	s nodes	α nodes	v nodes
0	800	67	41	30
1	1600	133	81	59

TABLE 4: Grid and time step data for convergence studies for the mean quadratic variation strategy from [16]. The same time steps are used in both PDE calculation and Monte Carlo simulations. The Monte Carlo computations interpolate the optimal control from the PDE grid values.

333 8.3 Comparisons of two risk measures

334 8.3.1 Making comparisons in the same units

335 In our figures, we plot the expected value against the risk measures in the same units. This means standard
336 deviation is plotted instead of variance. Similarly, the square root of quadratic variation is plotted instead
337 of quadratic variation (4.1). In our plots, the terminology $QVRisk$ stands for the square root of quadratic
338 variation, i.e.

$$QVRisk = \sqrt{E \left[\int_0^T (A(t') dS(t'))^2 \right]}. \quad (8.1)$$

339 8.3.2 Summary of comparisons

340 Figures 4 to 8 compare the mean variance trade off and the mean quadratic variation trade off for both the
341 mean variance and the mean quadratic variation strategy. For example, the left plot in Figure 4 compares
342 the results obtained using the mean variance strategy and the mean quadratic variation strategy in terms
343 of using standard deviation as the risk measure. Similarly, the right plot in Figure 4 compares the two
344 strategies in terms of using QV Risk as the risk measure.

345 Several conclusions can be drawn from the comparisons.

- 346 • As one would expect, in terms of using standard deviation as the risk measure, the mean variance
347 optimal strategy dominates the mean quadratic variation optimal strategy.
- 348 • Conversely, in terms of using QV Risk as the risk measure, the mean quadratic variation optimal
349 strategy dominates the mean variance optimal strategy.
- 350 • However, it appears that the mean variance optimal strategy performs reasonably well using either risk
351 measure. The difference between the two strategies is most pronounced at lower risk levels.

352 8.3.3 Remarks

353 Market practitioners may consider expected implementation shortfall (the relative difference between ex-
354 pected value and initial stock price) of 10 basis points to be significant. To achieve small implementation

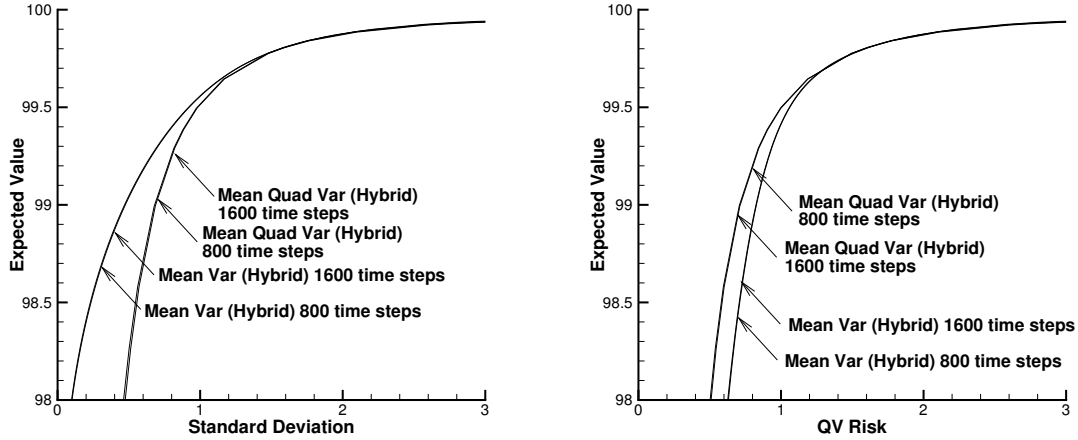


FIGURE 4: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=1.0$, $\kappa_t = 2 \times 10^{-6}$.

355 shortfall, liquidation must be done slowly to reduce trading impact, at the expense of increasing timing risk.
 356 Striking a good balance is important here, as it might not be wise to aim at an expected shortfall of 10 bps
 357 if the risk (as measured by either standard deviation or QV Risk) is several times larger. Our plots show
 358 that risk can be several times of a 10 bps expected shortfall in the parametric cases (a) $\sigma = 1.0$, 16.7% daily
 359 volume; (b) $\sigma = 0.2$, 20% daily volume; and (c) $\sigma = 0.2$, 5% daily volume.

360 The analysis above suggests that one way to choose a risk aversion level on an efficient frontier is to
 361 choose a ratio between the implementation shortfall and risk. Alternatively, a common practice among
 362 market practitioners is to pick the “corner of the frontier”. Our plots show that picking the corner can result
 363 in expected implementation shortfall much larger than 10 bps.

364 8.4 Comparison of strategies for similar expected values

365 In this section we compare the mean-variance strategy with the mean-quadratic- variation strategy when
 366 they give similar expected values. In particular, we focus on the parametric case $\sigma = 1$, $\kappa_t = 2 \times 10^{-6}$ since
 367 the differences are more apparent when volatility and pricing impacts are larger.

368 Figures 9 to 12 correspond to comparisons across four horizontal lines in Figure 4, with four different
 369 expected values chosen to represent the more interesting part of the frontiers. For example, in Figure 9
 370 both strategies give an expected value of around 99.29. For the mean-variance strategy, this corresponds to
 371 $\gamma = 199.82$; for the mean-quadratic-variation strategy, this corresponds to $\lambda = 1$.

372 8.4.1 Common observations for each level of expected value

373 In each of Figures 9 to 12, the subplots labeled (a) and (b) compare the optimal trading velocities at $t = 0$,
 374 where we normalized the trading velocities so that a normalized velocity of -1.0 corresponds to the constant
 375 liquidation rate $-\alpha_{init}/T$. It is clear that while both strategies are aggressive in the money (sell faster as
 376 price becomes more favorable), the sensitivity in the mean-variance strategy is much more non-linear. More
 377 specifically, around the initial asset price $s_{init} = 100$, the optimal control for the mean-variance strategy is
 378 a curve with rapidly changing slope whereas that for the mean-quadratic-variation strategy is more or less
 379 a straight line. It is also worth noting that the optimal selling rates at $s_{init} = 100$ for the mean-variance
 380 strategy are close to but slightly larger than those for the mean-variation strategy in Figures 9 to 12.

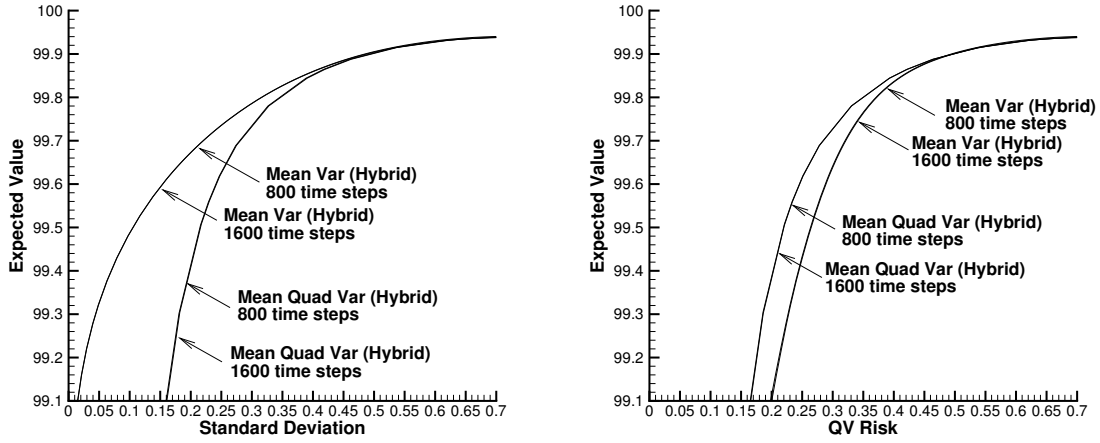


FIGURE 5: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=0.2$, $\kappa_t = 2.4 \times 10^{-6}$.

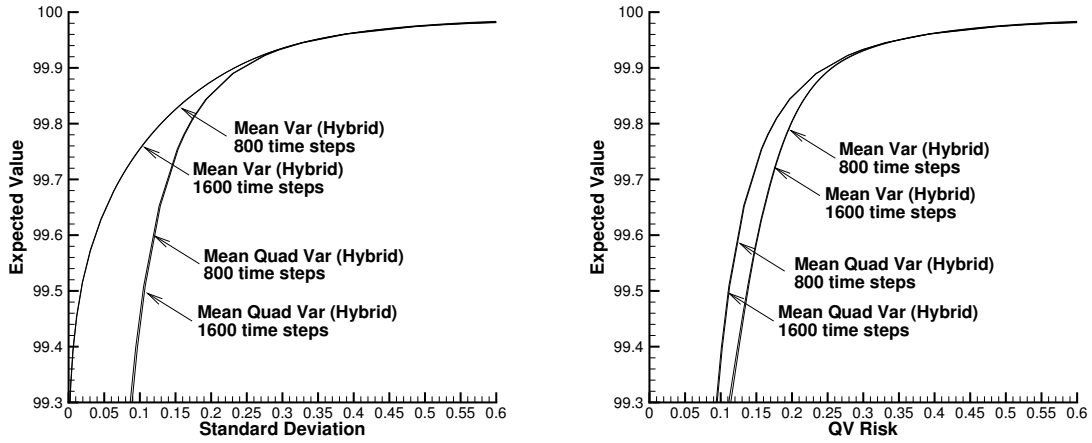


FIGURE 6: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=0.2$, $\kappa_t = 6 \times 10^{-7}$.

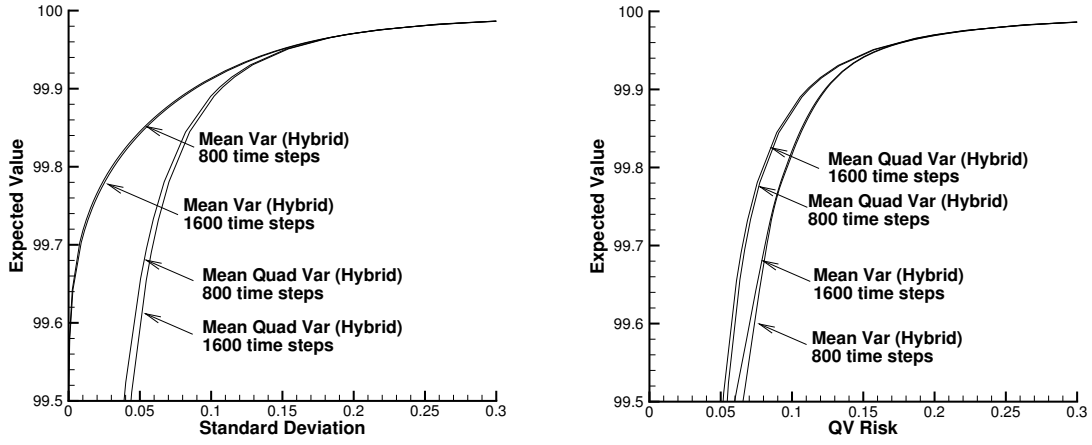


FIGURE 7: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=0.2$, $\kappa_t = 1.2 \times 10^{-7}$.

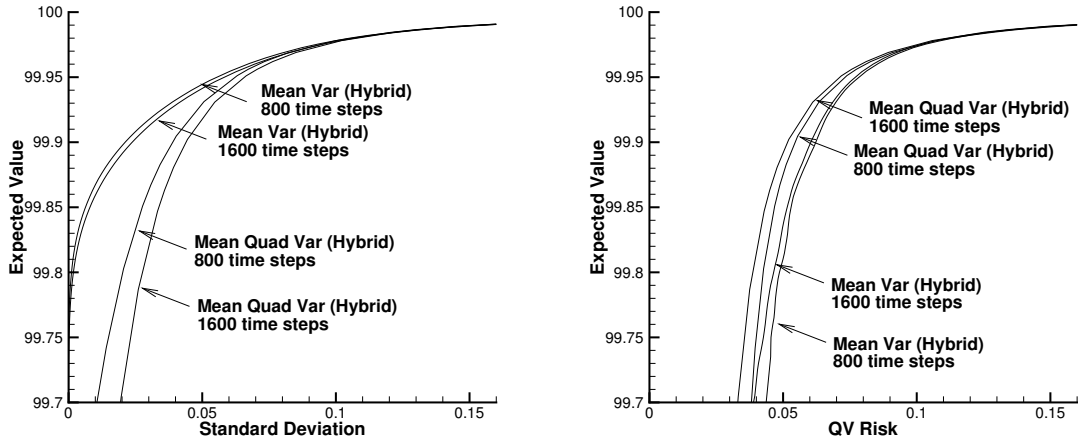


FIGURE 8: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=0.2$, $\kappa_t = 2.4 \times 10^{-8}$.

381 Note that the trading velocity in Figure 9 (b) is nonsmooth for large values of asset price. This appears
382 to be due to the near illposedness of the mean-variance formulation, as discussed in [15].

383 In each of Figures 9 to 12, the subplots labeled (c) and (d) compare the mean and standard deviation,
384 respectively, of the liquidation profiles $\alpha(t)$ of the mean-variance and the mean-quadratic-variation strategies
385 over the trading horizon. It is interesting to note that while the mean profiles are very similar, the standard
386 deviation profile of the mean-variance strategy is much larger than that of the mean-quadratic-variation
387 strategy. This reflects the fact that the mean-variance strategy is much more sensitive to change in asset
388 price during the liquidation, which is also suggested by the strategy subplots. We also note that the mean
389 profiles are convex, so that the mean liquidation rate is always decreasing over time.

390 8.4.2 Differences among different levels of expected value

391 As we move from Figure 9 to 12, the expected value is increasing, and so is the standard deviation and QV
392 Risk. By comparing subplots (a) and (b), we see that the optimal selling rates become slower as expected
393 value increases. Recall that the mean profiles are convex, so that the mean liquidation rate is always
394 decreasing over time. By comparing subplots (c), we observe that the convexities of the mean liquidation
395 profiles diminish as expected value increases, and the mean liquidation profiles approach a straight line. By
396 comparing subplots (d), we observe that the mean-variance strategy becomes less variable as expected value
397 increases.

398 9 Conclusion

399 In this paper, we have compared the optimal trading strategies obtained using two objective functions: mean
400 variance and mean quadratic variation. Recall that the original strategy proposed in [2] is actually a mean
401 quadratic variation strategy [16]. The mean quadratic variation is naturally time-consistent [11, 31]. On
402 the other hand, the pre-commitment mean variance strategy [15, 7, 26] is not time consistent. However,
403 the pre-commitment mean variance strategy is undoubtedly optimal if performance is measured in terms of
404 observed post-trade mean variance data.

405 The mean variance strategy is much more aggressive in the money, and is a highly nonlinear function of
406 the asset price. By contrast, the mean quadratic variation strategy is approximately linear in the asset price.
407 In addition, the mean variance strategy is much more variable than the mean quadratic variation strategy.
408 Nevertheless, both strategies turn out to have very similar mean liquidation profiles.

409 In terms of using both standard deviation and QV Risk as risk measures, the mean variance strategy ap-
410 pears to be, overall, a good strategy. The difference between the two strategies, however, are only significant
411 at low levels of timing risk, or equivalently, high levels of implementation shortfall.

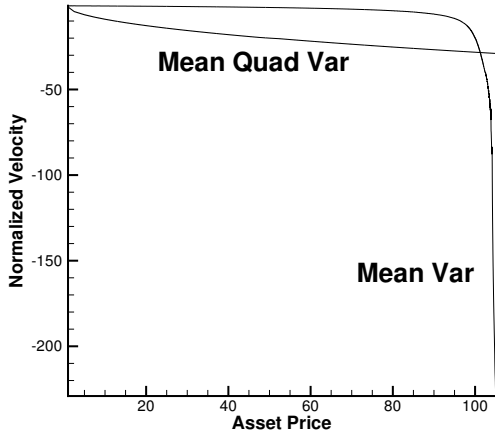
412 Consequently, if a highly variable strategy is acceptable, the mean-variance strategy is perhaps the better
413 choice. Otherwise, the mean quadratic variation strategy should be chosen if less variability in the strategy
414 is desired.

415 We have improved the numerical method in [15] by using a parametric curve interpolation scheme and
416 a scaled computational grid. The parametric curve interpolation accurately approximates the foot of the
417 semi-Lagrangian characteristics, which is essential for obtaining accurate numerical solutions for the optimal
418 control. The scaled computation grid concentrates computational resources on region of interest in the state
419 space so that sufficiently accurate results can be produced using few grid nodes.

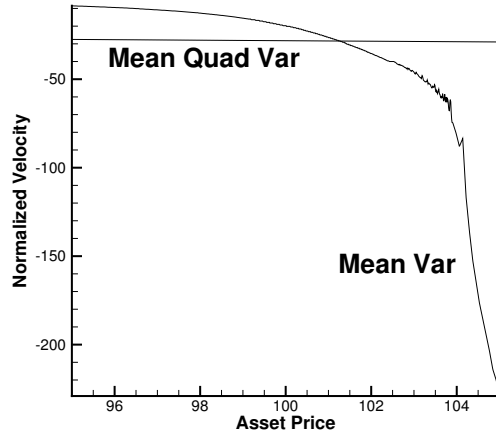
420 A Example Computation for the Temporary Price Impact Factor

421 Here we describe a realistic scenario in which the temporary price impact factor $\kappa_t = 1.2 \times 10^{-7}$ (Case 4 in
422 Table 2) corresponds to 1% of the daily volume of a stock.

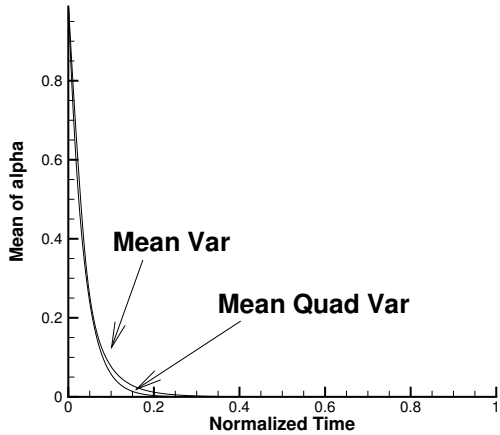
423 Suppose that the initial stock price $s_{init} = 100$ dollars, buy rate = 1,000 shares/min, corresponding
424 temporary price impact = 3 dollars/min, daily trading time = 420 minutes, and daily volume = 42,000,000



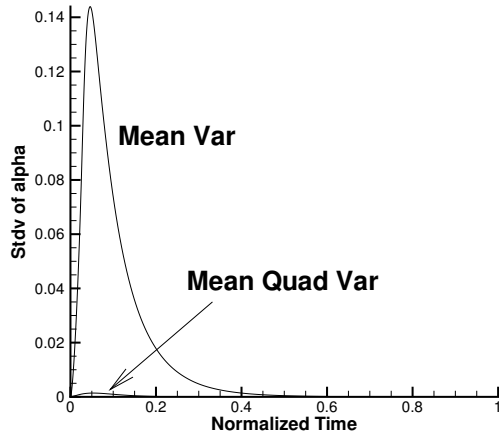
(a) strategy - complete view



(b) strategy - zoom in

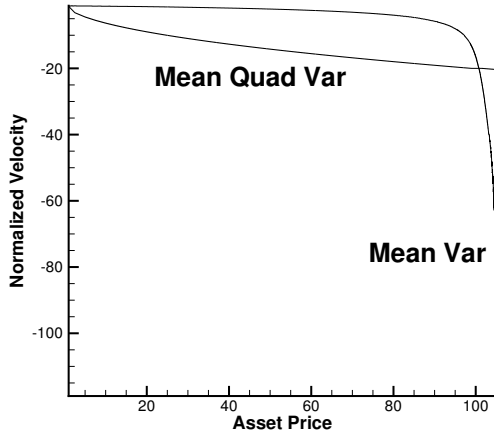


(c) mean of liquidation profile

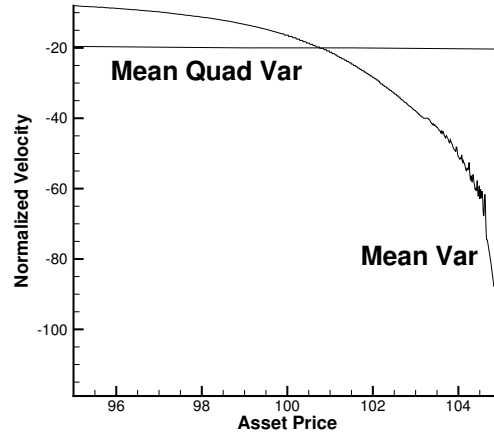


(d) standard deviation of liquidation profile

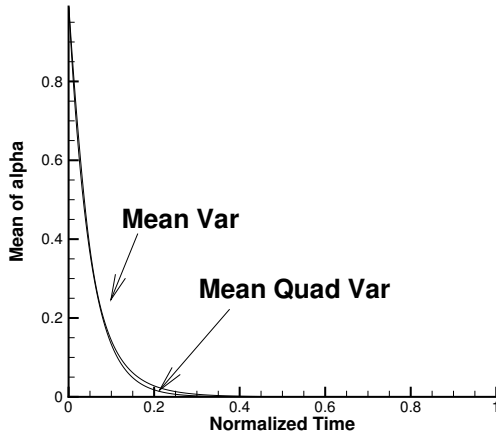
FIGURE 9: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=1.0$, $\kappa_t = 2 \times 10^{-6}$. The mean- variance strategy plotted has mean 99.29, standard deviation 0.68, QV Risk 0.93, and corresponds to $\gamma=199.82$. The mean-quadratic-variation strategy plotted has mean 99.29, standard deviation 0.82, QV Risk 0.84, and corresponds to $\lambda=1$. 1600 time steps are used to compute the results.



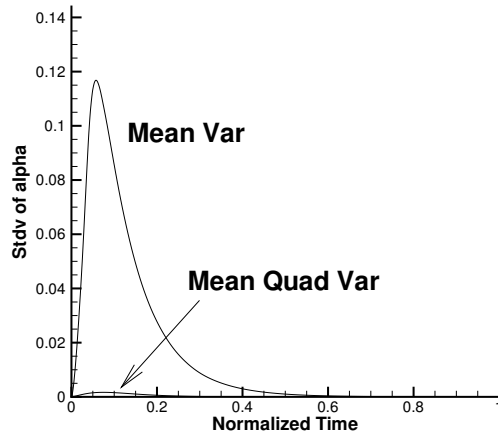
(a) strategy - complete view



(b) strategy - zoom in

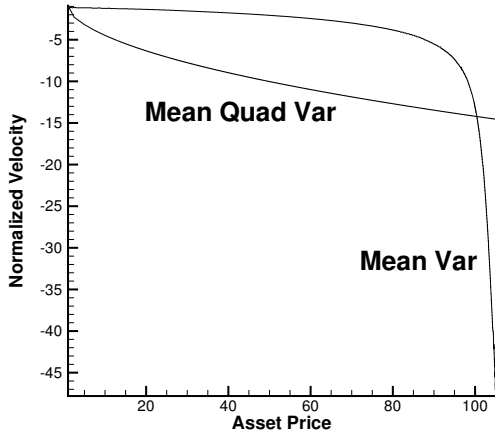


(c) mean of liquidation profile

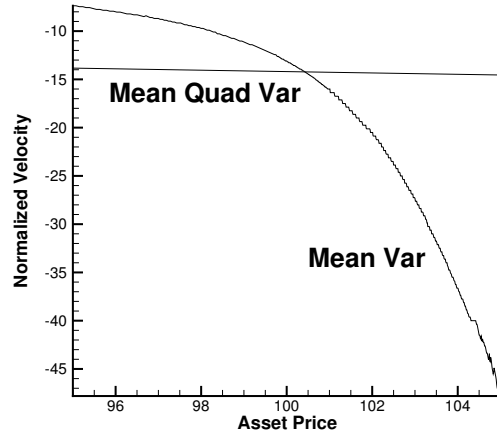


(d) standard deviation of liquidation profile

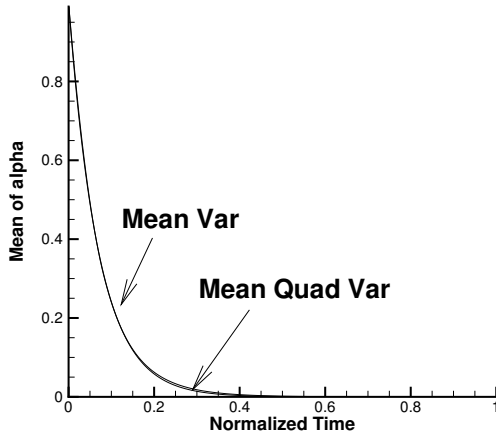
FIGURE 10: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=1.0$, $\kappa_t = 2 \times 10^{-6}$. The mean- variance strategy plotted has mean 99.50, standard deviation 0.90, QV Risk 1.05, and corresponds to $\gamma=201.30$. The mean-quadratic-variation strategy plotted has mean 99.50, standard deviation 0.98, QV Risk 1.00, and corresponds to $\lambda=0.5$. 1600 time steps are used to compute the results.



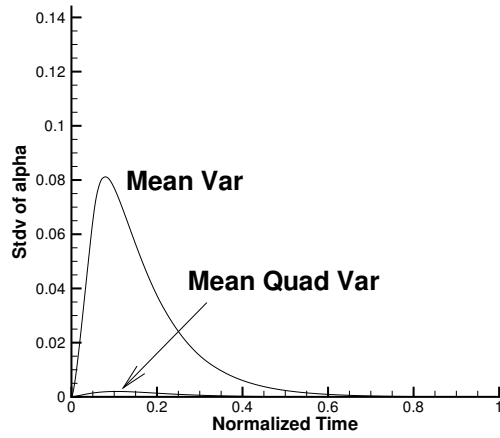
(a) strategy - complete view



(b) strategy - zoom in

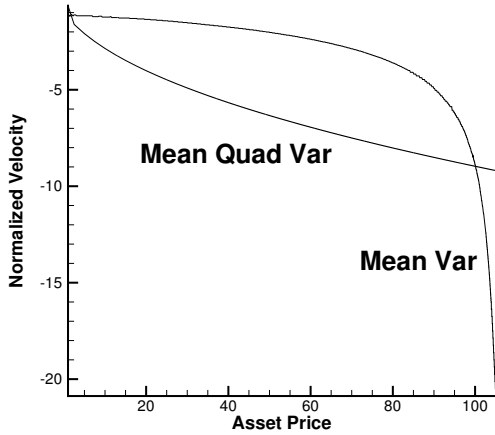


(c) mean of liquidation profile

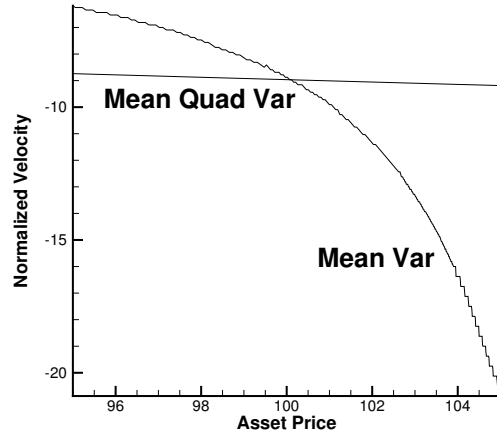


(d) standard deviation of liquidation profile

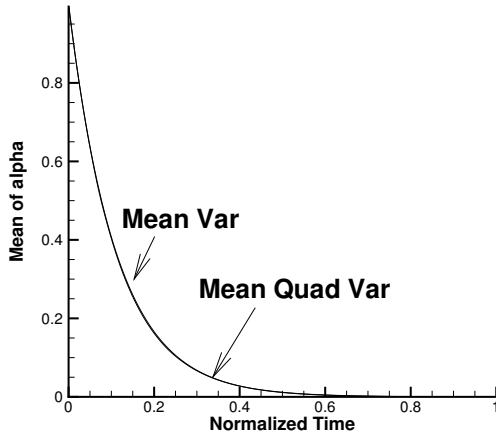
FIGURE 11: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=1.0$, $\kappa_t = 2 \times 10^{-6}$. The mean- variance strategy plotted has mean 99.65, standard deviation 1.13, QV Risk 1.21, and corresponds to $\gamma=203.50$. The mean-quadratic-variation strategy plotted has mean 99.65, standard deviation 1.17, QV Risk 1.19, and corresponds to $\lambda=0.25$. 1600 time steps are used to compute the results.



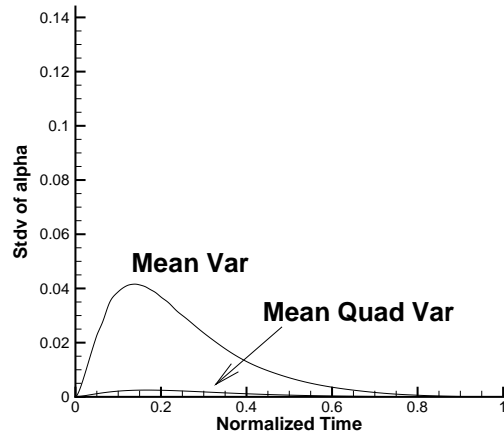
(a) strategy - complete view



(b) strategy - zoom in



(c) mean of liquidation profile



(d) standard deviation of liquidation profile

FIGURE 12: Comparison between mean variance strategy and mean quadratic variation strategy for the case $\sigma=1.0$, $\kappa_t = 2 \times 10^{-6}$. The mean- variance strategy plotted has mean 99.78, standard deviation 1.46, QV Risk 1.49, and corresponds to $\gamma=209.42$. The mean-quadratic-variation strategy plotted has mean 99.78, standard deviation 1.48, QV Risk 1.49, and corresponds to $\lambda=0.1$. 1600 time steps are used to compute the results.

425 shares. For such a scenario, our trading corresponds to 1% of the daily volume, and the daily market turnover
 426 for the stock is 4.2 billion dollars, corresponding to that typical of a large-cap stock.

427 Assuming a constant trading rate over one day ($T = 1/250$), then the total price impact is 3×420 . The
 428 ratio of total price impact to total initial value of stock is then given by

$$R = \frac{\text{total price impact}}{\text{total initial value}} = \frac{3 \times 420}{420 \times 1000 \times 100} = 3 \times 10^{-5} \quad (\text{A.1})$$

429 From the trading model (2.4) and (2.7), the captured price is $s_{init}f(v) = s_{init} \exp(-\kappa_t v) \approx s_{init}(1 - \kappa_t v)$.
 430 Therefore, the ratio R is approximately $\kappa_t |v|$.

431 Since $\alpha_{init} = 1$ and $T = 1/250$, the constant trading rate is $v = -250$. Substituting $v = -250$ into $\kappa_t |v|$
 432 $= 3 \times 10^{-5}$ gives $\kappa_t = 1.2 \times 10^{-7}$.

433 B Details of numerical method for solving PDEs (6.8) and (6.14)

434 The numerical method used in this paper for solving equations (6.8) and (6.14) is essentially the method
 435 used in [15] with two improvements: the use of a parametric curve linear interpolation scheme at the foot of
 436 the semi-Lagrangian characteristics and a scaled computational grid. In order to highlight the differences,
 437 we provide the discretization details only for these two improvements. Readers are referred to [15] for details
 438 on other aspects of the numerical method.

439 B.1 Construction of efficient frontier

440 Having solved (6.8) and (6.14), the variance minimizing frontier can be obtained as follows. Let $s = s_{init}$
 441 and $\alpha = \alpha_{init}$ be the initial values of s and α in forward time. For each value of γ , the corresponding point
 442 on the variance minimizing frontier can be shown to be given by the formulae

$$E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)] = U_0(\hat{b}) + \frac{\gamma}{2}, \quad (\text{B.1})$$

$$Var_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)] = V_0(\hat{b}) - (U_0(\hat{b}))^2, \quad (\text{B.2})$$

443 where $\hat{b} = \mathcal{B}(0) = -\gamma e^{-rT}/2$ is computed using equation (6.1), and $U_0(\hat{b})$ and $V_0(\hat{b})$ are shorthand notations
 444 for

$$V_0(\hat{b}) \equiv V(s, \hat{b}, \alpha, \tau = T) = E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[\mathcal{B}(T)^2], \quad (\text{B.3})$$

$$U_0(\hat{b}) \equiv U(s, \hat{b}, \alpha, \tau = T) = E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[\mathcal{B}(T)], \quad (\text{B.4})$$

445 which are obtained from solving the PDEs (6.8) and (6.14). The formulae (B.1) and (B.2) are obtained by
 446 solving for $E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)]$ and $E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)^2]$ from the linear system

$$E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)^2] - \gamma E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)] + \frac{\gamma^2}{4} = E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[\mathcal{B}(T)^2] \quad (\text{B.5})$$

$$E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[B(T)] - \frac{\gamma}{2} = E_{v^*(\cdot)}^{s, \hat{b}, \alpha, t=0}[\mathcal{B}(T)] \quad (\text{B.6})$$

447 The whole variance minimizing frontier is then obtained by varying γ .

448 B.2 Similarity Reduction

449 The assumption of Geometric Brownian Motion (2.5), the form of the price impact functions (2.7), (2.8),
 450 and the initial conditions (6.8), (6.14) imply the homogeneity properties

$$\begin{aligned} V(\xi s, \xi \hat{b}, \alpha, \tau) &= \xi^2 V(s, \hat{b}, \alpha, \tau), \\ U(\xi s, \xi \hat{b}, \alpha, \tau) &= \xi U(s, \hat{b}, \alpha, \tau), \\ v^*(\xi s, \xi \hat{b}, \alpha, \tau) &= v^*(s, \hat{b}, \alpha, \tau). \end{aligned} \quad (\text{B.7})$$

451 Therefore we can use similarity reduction to reduce the original three dimensional problem to a two dimen-
 452 sional problem, in which we only need to solve for one fixed value of \hat{b} .

453 B.3 Semi-Lagrangian discretization

454 In this section we demonstrate how equation (6.8) can be discretized by a semi-Lagrangian method. Equation
 455 (6.14) can be discretized in a similar fashion. For more details concerning semi-Lagrangian methods for HJB
 456 equations, the reader is referred to the references in [13].

457 Define a set of nodes $\{s_i\}$, $\{\hat{b}_j\}$, $\{\alpha_k\}$ and $\{\tau^n\}$, where $0 \leq i \leq i_{max}$, $\hat{b}_j \equiv \hat{b}^* < 0$, $0 \leq k \leq k_{max}$, and
 458 $0 \leq n \leq n_{max}$. We order the nodes in ascending order and make $s_0 = 0$, $\alpha_0 = 0$, $\alpha_{k_{max}} = \alpha_{init}$, $\tau_0 = 0$, and
 459 $\tau^{n_{max}} = T$. Note that there is only one node in the \hat{b} grid because of the use of similarity reduction. We
 460 denote the discrete approximation to V at the point $(s_i, \hat{b}_j, \alpha_k, \tau^n)$ by $V_{i,j,k}^n$ to distinguish it from the exact
 461 value $V(s_i, \hat{b}_j, \alpha_k, \tau^n)$. We also specify that the set of admissible control Z is of the form $[v_{min}, 0]$, where
 462 $v_{min} < 0$ is the fastest liquidation rate allowed.

463 Since the Lagrangian derivative $\frac{DV}{D\tau}(v_{i,j,k}^{n+1})$ at the node $(s_i, \hat{b}_j, \alpha_k, \tau^{n+1})$ is the derivative of V along the
 464 trajectory defined by (6.7). Solving equations (6.7) backwards in time from τ^{n+1} to τ^n , for a fixed $v_{i,j,k}^{n+1}$,
 465 gives the the foot of the characteristics $(s_{\hat{i}}, \hat{b}_j, \alpha_{\hat{k}}, \tau^n)$, which in general is not on the PDE mesh. We use the
 466 notation $V_{i,\hat{j},\hat{k}}^n$ to denote an approximation of $V(s_{\hat{i}}, \hat{b}_j, \alpha_{\hat{k}}, \tau^n)$ obtained by interpolation.

467 B.3.1 Local optimization

468 Denote the discrete form of \mathcal{L} by \mathcal{L}_h . By using an implicit discretization of $\mathcal{L}V$ and the semi-Lagrangian
 469 discretization on equation (6.8), we obtain

$$V_{i,j,k}^{n+1} = \min_{v_{i,j,k}^{n+1} \in Z_k^{n+1}} V_{i,\hat{j},\hat{k}}^n + (\tau^{n+1} - \tau^n)(\mathcal{L}_h V)_{i,j,k}^{n+1}, \quad (\text{B.8})$$

470 with the initial condition

$$V_{i,j,k}^0 = \hat{b}_j^2, \quad (\text{B.9})$$

471 where we restrict the admissible velocities to Z_k^{n+1} so that $\alpha_{\hat{k}} \geq 0$.

472 Once the optimal control $(v^*)_{i,j,k}^{n+1}$ is determined, equation (6.14) can be solved by

$$U_{i,j,k}^{n+1} = U_{i,\hat{j},\hat{k}}^n|_{v=(v^*)_{i,j,k}^{n+1}} + (\tau^{n+1} - \tau^n)(\mathcal{L}_h V)_{i,j,k}^{n+1}, \quad (\text{B.10})$$

473 with the initial condition

$$U_{i,j,k}^0 = \hat{b}_j. \quad (\text{B.11})$$

474 Since no analytical expression is available for the local objective function, we find the optimal $v_{i,j,k}^{n+1}$
 475 by discretizing the control space Z_k^{n+1} and look for the optimal value using a linear search. This has the
 476 advantage of not making any assumptions about the local objective function, at the expense of a higher
 477 computational cost. Numerical experiments demonstrate that accurate results can be obtained by a rather
 478 coarse discretization of the control space.

479 B.4 Computational challenges and solutions

480 B.4.1 Difficulties in determining optimal velocity numerically

481 Recall that in the special case considered in section 7, the Lagrangian derivative is identically zero for any
 482 admissible trading velocities. In terms of equation (B.8), this means that $V_{i,\hat{j},\hat{k}}^n(v_{i,j,k}^{n+1})$ as a function of $v_{i,j,k}^{n+1}$
 483 is completely flat. In the parametric cases we consider, both $\sigma\sqrt{T}$ and κ_t are quite small, therefore these
 484 realistic cases are indeed similar to the completely ill-posed special case. Consequently $V_{i,\hat{j},\hat{k}}^n(v_{i,j,k}^{n+1})$ as a
 485 function of $v_{i,j,k}^{n+1}$ can be very flat, which means determining the true minimizer demands extremely high

486 accuracy. It is also obvious that even small interpolation error can significantly alter the estimated trading
 487 velocities.

488 Similar computational issues also arise when ordinary finite-differencing is used instead of the semi-
 489 Lagrangian method: since the optimal velocity v will be a function that depends on the ratio of the partial
 490 derivatives in $\frac{DV}{D\tau}$. Any error in approximating the partial derivatives can also significantly alter the estimated
 491 trading velocities.

492 B.4.2 Parametric curve linear interpolation

493 The previous section has explained the importance of accurate interpolation at the foot of the semi-
 494 Lagrangian characteristics. In [15], a standard axis-aligned linear interpolation is used, which turns out
 495 to be too inaccurate. This is not surprising given the quadratic nature of the value function V and the
 496 analysis in section 7.

497 In section 7 we have shown the benefit of performing a parametric curve linear interpolation for the
 498 special case considered. Here we extend the idea to general cases. In essence, the parametric line (7.5) is
 499 generalized to the line L defined by

$$L = (s_{\hat{i}}, b_{\hat{j}}, \alpha_{\hat{k}}) + \zeta \left(\frac{ds}{d\zeta}, \frac{db}{d\zeta}, \frac{d\alpha}{d\zeta} \right),$$

$$\frac{ds}{d\zeta} = -g(v_{i,j,k}^{n+1})s_{\hat{i}}, \quad \frac{db}{d\zeta} = -(rb_{\hat{j}} - v_{i,j,k}^{n+1}f(v_{i,j,k}^{n+1})s_{\hat{i}}), \quad \frac{d\alpha}{d\zeta} = -v_{i,j,k}^{n+1}, \quad (\text{B.12})$$

500 where $v_{i,j,k}^{n+1}$ is a candidate control value.

501 Since equations (B.12) express how changes in α lead to changes in s and b through both trading impact
 502 and pricing impact (through the terms $g(v_{i,j,k}^{n+1})$ and $f(v_{i,j,k}^{n+1})$), interpolating along L can be seen as an
 503 extension to interpolation along (7.5), which takes into account trading but not pricing impact.

504 Figure 13 compares the standard axis-aligned linear interpolation and the parametric curve linear inter-
 505 polation along the line L . For simplicity in illustration, linear interpolation along the s coordinate axis is
 506 not shown, and there is a actual b grid, i.e. no similarity reduction.

507 Note that the parametric curve linear interpolation as shown in Figure 13 does not require interpolation
 508 along the α coordinate axis but still requires linear interpolation along the other axes. When linear interpo-
 509 lation along the s direction or the b direction is performed, a fine grid is still needed to reduce interpolation
 510 error. In other words, when we treat α specially as in Figure 13, we avoid the need of a fine α grid, but a
 511 fine s grid (and a fine b grid when no similarity reduction is used) is still necessary.

512 We also note that the parametric curve linear interpolation is similar to the edge-directed interpolation
 513 method [22] in the image processing literature. Our method is similar in the sense that the direction of the
 514 line L is not fixed but adapts to the candidate control velocity $v_{i,j,k}^{n+1}$. Our method is different from that in
 515 [22] in that the parametric curve linear interpolation method does not necessarily use the neighboring grid
 516 nodes (See Figure 13(b)).

517 B.4.3 Remark on convergence proof

518 Having changed the interpolation scheme in the semi-Lagrangian method, it is important that the conver-
 519 gence proof in [15] is still valid. Linear interpolation is obviously consistent. To demonstrate stability, we
 520 need the following easy observation:

521 $V_{\hat{i},\hat{j},\hat{k}}^n$ as approximated in the new scheme takes the form

$$V_{\hat{i},\hat{j},\hat{k}}^n = \sum_p w_p V_p^n, \quad (\text{B.13})$$

522 where $w_p \geq 0$, $\sum_p w_p = 1$, and V_p^n are grid node values. Therefore, we have $|V_{\hat{i},\hat{j},\hat{k}}^n| \leq \|V^n\|$. This property
 523 allows the proof of stability in [15] to go through without change. In addition, since $U_{\hat{i},\hat{j},\hat{k}}^n$ as approximated

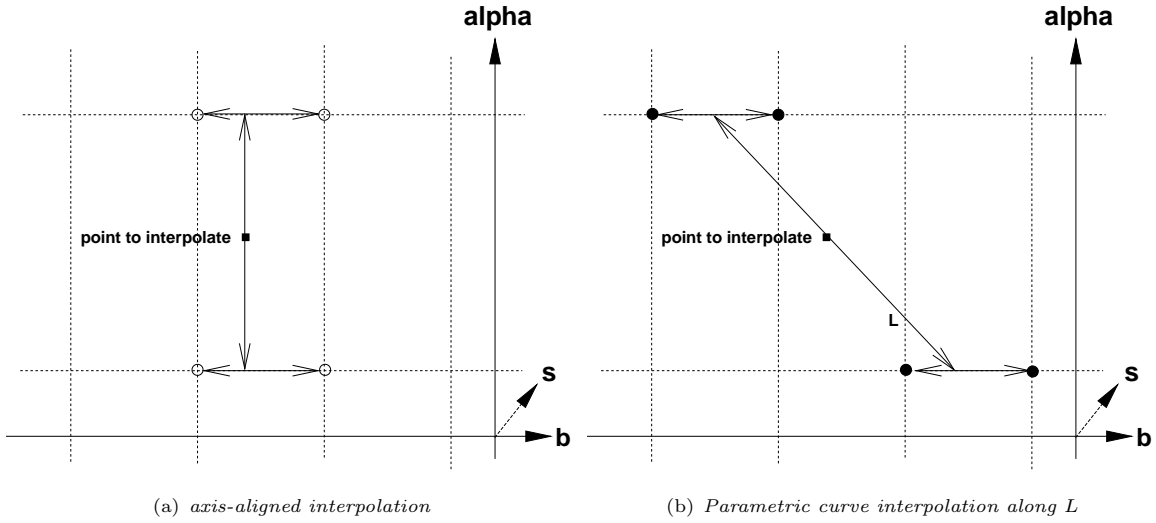


FIGURE 13: Comparing the two methods of interpolation at the foot of characteristics, shown as “point to interpolate” in the diagram. The dashed lines correspond to the computational grid and the dots are interpolation nodes.

524 in the scheme also takes the form

$$U_{\hat{i},\hat{j},\hat{k}}^n = \sum_p w_p U_p^n, \quad (\text{B.14})$$

525 where the w_p 's are the same as those in equation (B.13). Therefore, the proof in [15] which shows $(U_{i,j,k}^n)^2 \leq$
 526 $V_{i,j,k}^n$ is also valid for our scheme.

527 B.4.4 Computational grid consideration

528 Consider again the analytical solution for the special case derived in section 7:

$$V(s, \mathfrak{b}, \alpha, \tau) = (\alpha s + \mathfrak{b})^2. \quad (\text{B.15})$$

529 As can be seen in Figure 1, (B.15) is a good approximation for realistic parametric cases. Both Figure 1
 530 and (B.15) show that V does not change a significantly along a constant line of wealth $\{(\alpha s + \mathfrak{b}) = \text{const}\}$.
 531 This observation suggests constructing a computational grid with taking into account constant lines of wealth.

- 532 • For $\alpha > 0$, scale the s grid by $\{s_i\} \rightarrow \{s_i\}/\alpha$.
- 533 • For $\alpha = 0$, no scaling is performed, i.e. the original s grid $\{s_i\}$ is used.

534 Figure 14(a) and 14(b) illustrate the scaled computational grid, and the value function under the scaled
 535 grid, respectively. Compare the lines of constant wealth in Figure 1 and Figure 14(b).

536 Note that the shape of the value function is simpler in Figure 14(b) than in Figure 1. Note also that the
 537 mesh is always dense in the region $V \approx 0$ in the scaled grid in Figure 14(b), which is not the case in Figure
 538 1. In our experience, accurate values of V in the region $V \approx 0$ are important for computing the efficient
 539 frontiers. Thus, the scaled grid is also computationally more effective.

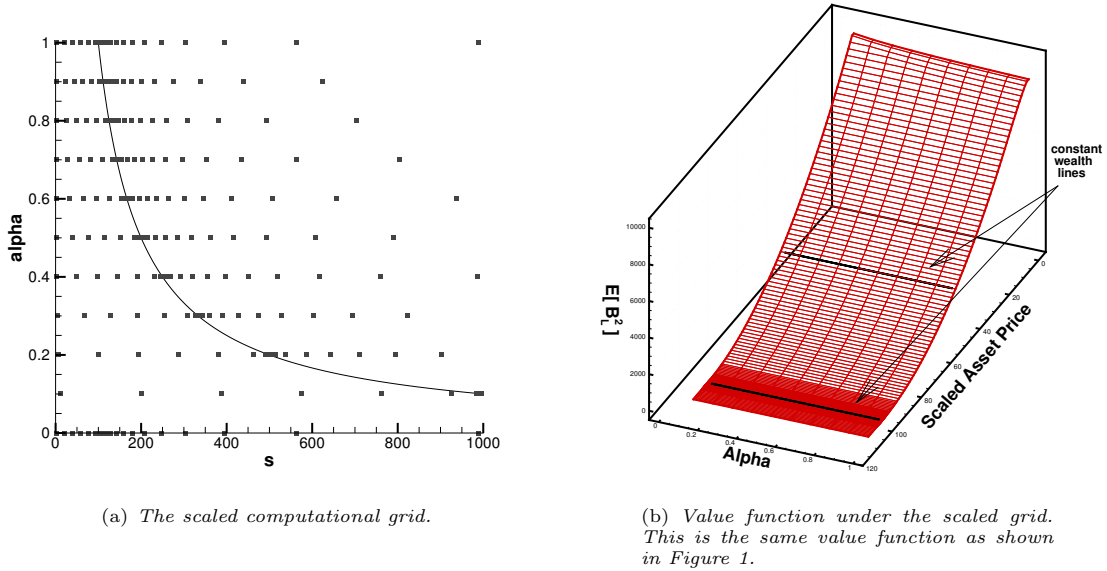


FIGURE 14: Illustration of scaled computational grid.

C Details of Monte Carlo simulations

540

541 Numerically solving the mean-variance HJB equation (6.8) gives us the optimal strategy on the discrete
 542 computational mesh, i.e. $v_{MV}(s_i, \hat{\mathbf{b}}^*, \alpha_k, t^n)$ ⁴. By using $v_{MV}(s_i, \hat{\mathbf{b}}^*, \alpha_k, t^n)$ as input for Monte Carlo simu-
 543 lations, we can obtain information about the trading strategy which is not necessarily available in the PDE
 544 solutions. For example, we can estimate the probability distribution of $B(T)$, the quadratic variation, and
 545 the mean and the standard deviation of the liquidation profile (plot of $A(t)$ against t).

546 Similarly, numerically solving the mean-quadratic-variation HJB equation (4.7), for each fixed value of
 547 λ , gives us the discrete optimal strategy $v_{MQV}(s_{\tilde{i}}, \alpha_{\tilde{k}}, t^n; \lambda)$ ⁵, which can be used as input for Monte Carlo
 548 simulations.

549 In particular, Monte Carlo simulations enable us to compute the quadratic variation of the mean-variance
 550 strategy, and conversely, the variance of the mean-quadratic-variation strategy. These allow us to compare
 551 the two strategies in terms of either variance or quadratic variation, given the same level of expected return.

552 The Monte Carlo simulations also provide a verification of the PDE solutions, in the sense that given the
 553 optimal control, we can obtain independent estimates of mean, variance and quadratic variation.

554 In the following we detail how the Monte Carlo method is conducted for the mean-variance strategy.
 555 Simulations of the mean-quadratic-variation strategy are performed in the same way, except that the mean
 556 quadratic variation strategy $v_{MQV}(s_{\tilde{i}}, \alpha_{\tilde{k}}, t^n; \lambda)$ is explicitly indexed by λ , and standard axis-aligned linear
 557 interpolation suffices.

C.1 Change of variable

558

559 Suppose that the optimal strategy $v_{MV}(s_i, \hat{\mathbf{b}}^*, \alpha_k, t^n)$ is obtained from the PDE solve. For a fixed value of
 560 γ , each Monte Carlo simulation starts with the initial values $S(0), B(0), A(0)$ at time $t = 0$ and is updated
 561 at the discrete times $\{t^n\}$, i.e. the time grid nodes in the PDE solve. Below we give a full specification of

⁴Note the use of forward time notation.

⁵We use the notations \tilde{i} and \tilde{k} to emphasize that the s grid and α grid for solving the mean-quadratic-variation HJB equation (4.7) is not necessarily the same as that for solving the mean-variance HJB equation (6.8). The time grid $\{t^n\}$, however, is chosen to be the same.

562 the simulation procedure by detailing the simulation from time point t_{old} to the immediate next time point
563 t_{new} .

564 At t_{old} , the state is $(S_{old}, B_{old}, A_{old})$. To look up the optimal trading velocity, we first need to change
565 the variable from B to \mathcal{B} . For the fixed value of γ , we have $\mathcal{B}_{old} = B_{old} - \gamma e^{-r(T-t_{old})}/2$ from equation
566 (6.2). Now the optimal trading velocity $v(S_{old}, \mathcal{B}_{old}, A_{old}, t_{old})$ needs to be interpolated from the discrete
567 $v_{MV}(s_i, \hat{\mathbf{b}}^*, \alpha_k, t^n)$.

568 C.1.1 Interpolation

569 Our numerical study shows that it is more accurate to linearly interpolate $v_{MV}(s_i, \hat{\mathbf{b}}^*, \alpha_k, t^n)$ along a
570 constant line of wealth $\{\alpha s + \hat{\mathbf{b}} = \text{constant}\}$ than along the coordinate axes. Therefore, we interpolate
571 $v_{MV}(s_i, \hat{\mathbf{b}}^*, \alpha_k, t^n)$ as in Figure 13(b) with L given by the constant line of wealth $\{\alpha s + \hat{\mathbf{b}} = \text{constant}\}$. This
572 is the same as the interpolation for the limiting parametric case in section 7.1. Note that the form of the
573 line L as defined by equation (B.12) is not applicable in the current context because there is no candidate
574 control $v_{i,j,k}^{n+1}$.

575 C.1.2 Updating state variables

576 Let $\Delta t = t_{new} - t_{old}$, we update the state variable as follows:

$$A_{opt} = A_{old} + v(S_{old}, \mathcal{B}_{old}, A_{old}, t_{old})\Delta t, \quad (\text{C.1})$$

$$A_{new} = \max(A_{opt}, 0), \quad (\text{C.2})$$

$$v_{opt} = (A_{new} - A_{old})/\Delta t, \quad (\text{C.3})$$

$$S_{new} = S_{old} \exp\{(\eta + g(v_{opt}) - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}\mathcal{N}(0, 1)\}, \quad (\text{C.4})$$

$$B_{new} = B_{old} \exp\{r\Delta t\} - v_{opt}f(v_{opt})S_{old}\Delta t, \quad (\text{C.5})$$

$$QV_{new} = QV_{old} + (A_{old}(S_{new} - S_{old}))^2, \quad (\text{C.6})$$

577 where $\mathcal{N}(0, 1)$ is a standard normal variate and QV_{new} is an approximation of $\int_0^{t_{new}} (A(t') dS(t'))^2$.

578 References

- 579 [1] A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general
580 shape functions. *Quantitative Finance*, 10(2):143–157, 2010.
- 581 [2] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.
- 582 [3] R. Almgren and J. Lorenz. Adaptive arrival price. *Trading*, 2007(1):59–66, 2007.
- 583 [4] R. Almgren, C. Thum, E. Hauptmann, and H. Li. Direct estimation of equity market impact. *Risk*,
584 18:57–62, 2005.
- 585 [5] R.F. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied*
586 *Mathematical Finance*, 10(1):1–18, 2003.
- 587 [6] L. Bai and H. Zhang. Dynamic mean-variance problem with constrained risk control for the insurers.
588 *Mathematical Methods of Operations Research*, 68(1):181–205, 2008.
- 589 [7] S. Basak and G. Chabakauri. Dynamic mean-variance asset allocation. *Review of Financial Studies*,
590 2010.
- 591 [8] D. Bertsimas and A. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50,
592 1998.

- 593 [9] N. Besse and E. Sonnendrucker. Semi-Lagrangian schemes for the Vlasov equation on an unstructured
594 mesh of phase space. *Journal of Computational Physics*, 191(2):341–376, 2003.
- 595 [10] T.R. Bielecki, H. Jin, S.R. Pliska, and X.Y. Zhou. continuous-time mean-variance portfolio selection
596 with bankruptcy prohibition. *Mathematical Finance*, 15(2):213–244, 2005.
- 597 [11] Tomas Bjork and Agatha Murgoci. A General Theory of Markovian Time Inconsistent Stochastic
598 Control Problems. *SSRN eLibrary*, 2010.
- 599 [12] P. Brugiére. Optimal portfolio and optimal trading in a dynamic continuous time framework. 6th AFIR
600 Colloquium. *Nuremberg, Germany*, 12:89, 1996.
- 601 [13] Z. Chen and P.A. Forsyth. A semi-Lagrangian approach for natural gas storage valuation and optimal
602 operation. *SIAM Journal on Scientific Computing*, 30(1):339–368, 2007.
- 603 [14] R.F. Engle and R. Ferstenberg. Execution Risk. *The Journal of Trading*, 2(2):10–20, 2007.
- 604 [15] P.A. Forsyth. A Hamilton Jacobi Bellman approach to optimal trade execution. *Applied Numerical
605 Mathematics*, 61:241–265, 2011.
- 606 [16] PA Forsyth, JS Kennedy, ST Tse, and H. Windcliff. Optimal trade execution: a mean quadratic
607 variation approach. *Submitted to Journal of Economic Dynamics and Control*, 2011.
- 608 [17] C. Fu, A. Lari-Lavassani, and X. Li. Dynamic mean-variance portfolio selection with borrowing con-
609 straint. *European Journal of Operational Research*, 200(1):312–319, 2010.
- 610 [18] J. Gatheral and A. Schied. Optimal Trade Execution under Geometric Brownian Motion in the Almgren
611 and Chriss Framework. *Forthcoming in International Journal of Theoretical and Applied Finance*, 2011.
- 612 [19] H. He and H. Mamaysky. Dynamic trading policies with price impact. *Journal of Economic Dynamics
613 and Control*, 29(5):891–930, 2005.
- 614 [20] G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275,
615 2004.
- 616 [21] B. Johnson. *Algorithmic Trading & DMA: An introduction to direct access trading strategies*. 4Myeloma
617 Press, 2010.
- 618 [22] X. Li and M.T. Orchard. New edge-directed interpolation. *Image Processing, IEEE Transactions on*,
619 10(10):1521–1527, 2001.
- 620 [23] X. Li and X.Y. Zhou. Continuous-time mean-variance efficiency: the 80% rule. *The Annals of Applied
621 Probability*, 16(4):1751–1763, 2006.
- 622 [24] F. Lillo, J.D. Farmer, and R.N. Mantegna. Econophysics: Master curve for price-impact function.
623 *Nature*, 421(6919):129–130, 2003.
- 624 [25] J. Lorenz and R. Almgren. Mean-variance optimal adaptive execution. *Applied Mathematical Finance*,
625 2011. forthcoming.
- 626 [26] J.M. Lorenz. *Optimal trading algorithms: Portfolio transactions, multiperiod portfolio selection, and
627 competitive online search*. PhD thesis, ETH Zurich, 2008.
- 628 [27] V. Ly Vath, M. Mnif, and H. Pham. A model of optimal portfolio selection under liquidity risk and
629 price impact. *Finance and Stochastics*, 11(1):51–90, 2007.
- 630 [28] A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *NBER Working
631 Paper*, 2005.

- 632 [29] M. Potters and J.P. Bouchaud. More statistical properties of order books and price impact. *Physica A:*
633 *Statistical Mechanics and its Applications*, 324(1-2):133–140, 2003.
- 634 [30] A. Schied and T. Schoeneborn. Risk aversion and the dynamics of optimal liquidation strategies in
635 illiquid markets. *Finance and Stochastics*, 13(2):181–204, 2009.
- 636 [31] J. Wang and PA Forsyth. Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous
637 time mean variance asset allocation. *Journal of Economic Dynamics and Control*, 34(2):207–230, 2010.