PRESERVATION OF SCALARIZATION OPTIMAL POINTS IN THE EMBEDDING TECHNIQUE FOR CONTINUOUS TIME MEAN VARIANCE OPTIMIZATION *

SHU TONG TSE †, PETER A. FORSYTH ‡, AND YUYING LI §

Abstract. A continuous time mean-variance (MV) problem optimizes the bi-objective criteria $(\mathcal{V}, \mathcal{E})$, respectively representing variance $\mathcal{V}$ and expected value $\mathcal{E}$ of a random variable at the end of a time horizon T. This problem is computationally challenging since the dynamic programming principle cannot be directly applied to the variance criterion. An embedding technique has been proposed in [18, 25] to generate the set of MV scalarization optimal points, which is in general a subset of the mean-variance Pareto optimal points. However, there are a number of complications when we apply the embedding technique in the context of a numerical algorithm. In particular, the frontier generated by the embedding technique may contain spurious points which are not MV optimal. In this paper, we propose a method to eliminate such points, when they exist. We show that the original MV scalarization optimal objective set is preserved if we consider the scalarization optimal points (SOPs) with respect to the MV objective set derived from the embedding technique. Specifically, we establish that these two SOP sets are identical. For illustration, we apply the proposed method to an optimal trade execution problem, which is solved using a numerical Hamilton Jacobi Bellman (HJB) PDE approach.

Key words. Keywords: mean variance, embedding, Pareto optimal, scalarization optimization, optimal trade execution, HJB equation

AMS subject classifications. 90C29,91G80,93C20

Version: July 11, 2013

1. Introduction.

This paper addresses the question of how to determine the mean-variance (MV) Pareto optimal points when applying the embedding technique [18, 25] to solve a continuous time mean-variance optimization problem. For illustration and motivation, we begin with the important optimal trade execution problem [4, 19, 14, 3], which is solved using a numerical Hamilton Jacobi Bellman (HJB) Partial Differential Equation (PDE) approach.

When liquidating a large share position, an investment bank is faced with the following dilemma. If a large sell order is placed on the market, the average execution price obtained per share will be significantly lower than the pre-trade price, due to liquidity or price impact effects. The obvious alternative is to break up the large sell order into a number of small orders, and spread these orders over time. This will minimize the price impact, but expose the bank to the risk that the average price per share will also be less than the pre-trade price, due to the stochastic motion of the stock price.

The conflicting objectives of maximizing trading revenue (minimizing price impacts) and minimizing risk can be naturally formulated as maximizing $\mathcal{E} = E[B(T)]$ and minimizing $\mathcal{V} = Var[B(T)]$, where $B(T)$ is the cash balance at the end of trading

* This work was supported by the Natural Sciences and Engineering Research Council of Canada, and by a Credit Suisse Research Grant. The views expressed herein are solely those of the authors, and not those of any other person or entity, including Credit Suisse. The authors would like to thank P.A.I. Forsyth for many useful discussions

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

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

§ David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 yuying@uwaterloo.ca

horizon T, where $E[\cdot]$ is the expectation operator, and $Var[\cdot]$ is the variance. In line with previous work, we assume that trading takes place continuously at a finite rate [5, 20, 3]. In this approach, risk is measured in terms of variance [12, 11]. Alternatively, risk can be measured in terms of quadratic variation [4] and Value-at-Risk [14].

Using a standard method for multicriteria optimization, a positive scalarization combination of the multiple criteria is optimized to obtain Pareto optimal solutions. Typically, dynamic programming is then applied to solve the resulting optimal stochastic control problem. Unfortunately, in the case of mean-variance criteria, dynamic programming cannot be readily applied due to the variance term. An embedding technique, which uses $\mathcal{Q} = E[B(T)^2]$ instead of the variance $\mathcal{V} = Var[B(T)]$, has been proposed in [18, 25] to overcome this difficulty.

We note that the optimal strategy computed from the embedding technique [12] is a pre-commitment strategy [7], which is not necessarily time consistent. However, as pointed out in [3], the pre-commitment strategy corresponds to the situation where a trading desk optimizes the measured sample mean and variance across a large collection of similar trades. Consequently, the pre-commitment MV strategy optimizes trading effectiveness as measured in practice [24]. In addition, optimal trading strategies typically work on the time scale of one day or less, hence the strategies are essentially pre-commitment in any case.

Using dynamic programming, a combination of the objectives $(\mathcal{Q}, \mathcal{E})$ from the embedding technique is optimized. This optimization problem can be expressed in the form of a nonlinear Hamilton Jacobi Bellman (HJB) partial differential equation. We refer to [12] for the details of the numerical methods used to solve the HJB equation.

In [18, 25], it has been established that an MV scalarization optimal control is also an optimal control for the embedded problem. Let $\mathcal{Y}_P$ denote the set of the original MV scalarization optimal $(\mathcal{V}, \mathcal{E})$ objectives. Assume that $\mathcal{Y}_Q$ denotes the set of the (embedded) mean-variance $(\mathcal{V}, \mathcal{E})$ objective with a suitable combination equal to an optimal value of the embedding problem for a parameter γ. The result in [18, 25] effectively implies that the original MV scalarization optimal set $\mathcal{Y}_P$ is a subset of the (embedded) MV objective set $\mathcal{Y}_Q$ generated by an embedding technique. Points in $\mathcal{Y}_Q$, which do not correspond to points in $\mathcal{Y}_P$, are termed spurious points.

To the best of our knowledge, conditions under which the converse result holds have not been established. Thus we are faced with the problem of determining whether the optimal control computed from the embedding technique is necessarily an optimal control for the original MV scalarization optimization problem. Unfortunately there can be multiple objective points $(\mathcal{V}, \mathcal{E})$ (associated with admissible controls) which yield the single optimal objective of the embedding technique for a specific embedding parameter. For example, there can exist two solutions that optimize the embedding objective function but achieve different MV objectives. Consequently it is not immediately clear how to identify which MV points from the embedding technique belong to the original MV optimal set $\mathcal{Y}_P$.

For the optimal trade execution, this problem is compounded since the optimal control may not be unique. In addition, a numerical algorithm will, in general, compute only a single optimal solution. To see the non-uniqueness of the optimal trade execution strategy, we note that, due to price impact effects, rapid selling will lower the average price obtained for the shares. As a trivial example, consider the case where the desired outcome is a zero variance. This can be achieved by selling all shares at an infinite rate at the initial time $t = 0^+$. This strategy will result in zero expected gain ($\mathcal{E} = 0$) and zero variance ($\mathcal{V} = 0$). Alternatively, the trader could wait until T^-, and then sell all shares at an infinite rate, and achieve the same result. There are infinitely many such strategies which yield the same Pareto point $(\mathcal{E}, \mathcal{V}) = (0, 0)$. For more discussion of this, we refer a reader to [12].

In general the mean-variance frontier generated by the embedding technique may contain spurious points that are not MV Pareto optimal. This gives rise to a number of issues when we use the embedding method to numerically compute the optimal solution based on a non-linear Hamilton Jacobi Bellman (HJB) partial differential equation. Furthermore, it is necessary to devise techniques to identify when the solutions from the embedding formulation yield Pareto mean-variance optimal solutions.

In this paper, we provide a method to identify the MV scalarization optimal points in $\mathcal{Y}_P$ when using the embedding technique. Let $\mathcal{Y}$ denote the MV objective set achievable by admissible strategies. Thus $\mathcal{Y}_P$ is the set of MV points which are scalarization optimal with respect to $\mathcal{Y}$. We address the identification problem by considering the MV scalarization optimal points (SOPs) with respect to the embedded MV objective set $\mathcal{Y}_Q$ from the embedding technique. In the context of the numerical computation, we assume that a numerical algorithm generates a single embedded MV point $(\mathcal{V}, \mathcal{E})$ for each embedding parameter γ. We denote this computed embedded objective set by $\mathcal{Y}_Q^\dagger$. We similarly consider MV SOPs with respect to the computed MV objective set $\mathcal{Y}_Q^\dagger$.

The main contributions of the paper can be summarized as follows.

- We establish that, if an embedded objective point $(\mathcal{V}, \mathcal{E})$ is MV scalarization optimal with respect to the embedded MV objective set $\mathcal{Y}_Q$, it is scalarization optimal with respect to the achievable MV objective set $\mathcal{Y}$ (thus MV Pareto optimal).
- We prove that the set of the MV scalarization optimal points with respect to the computed embedded objective set $\mathcal{Y}_Q^\dagger$ is identical to the scalarization optimal set with respect to the achievable MV objective set $\mathcal{Y}$.
- The above two results allow us to develop a simple technique which can be used to eliminate the potential spurious MV points from the computed embedded objective set $\mathcal{Y}_Q^\dagger$.
- We demonstrate the application of these results to the optimal trade execution problem.

Note that these new mathematical results have a clear geometric interpretation: a scalarization optimal point with respect to a set corresponds to a point at which a supporting hyperplane with a positive slope for the set exists. Hence, for the computed MV set $\mathcal{Y}_Q^\dagger$, a point $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}_Q^\dagger$ is spurious if, at $(\mathcal{V}, \mathcal{E})$, there does not exist a supporting hyperplane for $\mathcal{Y}_Q^\dagger$ with a positive slope. We also emphasize that the results in this paper are not specific to the optimal trade execution problem. Indeed they can be applied to any continuous time MV optimization problem. However, for concreteness, we will first formulate the MV problem specifically for the optimal trade execution problem. The reader should have no difficulty applying our main results to other continuous time MV optimization problems.

2. Optimal Trade Execution Model. Optimal trade execution is concerned with balancing price impact (larger for faster execution) and timing risk (larger for slower execution). In this section we briefly outline our optimal trade execution model. We refer readers interested in optimal trade execution in general to [4, 14, 3, 2, 15] and to [12, 23] for more details about our formulation. Let

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

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

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

In this article, for concreteness, we consider the selling case where $\alpha_{init} > 0$. At $t = T$,

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

where $B(T)$ is the cash generated by selling shares and investing in the risk free bank account B, with a final liquidation at $t = T^-$ to ensure that $A(T) = 0$. The objective of optimal execution is to maximize the expected value of $B(T)$, while at the same time minimizing its variance.

In the following, we only consider feedback control trading strategies $v(\cdot)$ that specify a buying rate v as a function of the current state, i.e. $v(\cdot) : (S(t), B(t), A(t), t) \mapsto v = v(S(t), B(t), A(t), t)$ (i.e. Markovian w.r.t. (S, B, A)). Since v is the buying rate, $v < 0$ will denote selling, which is the example we consider in this paper. Note that in using the shorthand notation $v(\cdot)$ for the mapping, and v for the value $v = v(S(t), B(t), A(t), t)$, the dependence of v on the current state is implicitly assumed.

By definition,

$$dA(t) = v\ dt. \tag{2.3}$$

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

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

where r is the risk free rate.

We suppose that the market price of the risky asset S follows a Geometric Brownian Motion (GBM), where the drift term is modified due to the permanent price impact of trading [5]:

$$\begin{aligned} dS(t) &= \big(\eta + g(v)\big)S(t)\ dt + \sigma S(t)\ d\mathbb{W}(t), \\ &\quad \eta \text{ is the drift rate,} \\ &\quad g(v) \text{ is the permanent price impact function,} \\ &\quad \sigma \text{ is the volatility,} \\ &\quad \mathbb{W}(t) \text{ is a Wiener process under the real world measure.} \end{aligned} \tag{2.5}$$

2.1. Trading Impact Function. We assume that the temporary price impact scales linearly with the asset price, i.e.

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

where

$$
\begin{aligned}
f(v) &= (1+\kappa_s \operatorname{sgn}(v)) \exp[\kappa_t \operatorname{sgn}(v)|v|^{\beta}], \\
\kappa_s &= \text{ the bid-ask spread parameter}, \\
\kappa_t &= \text{ the temporary price impact factor}, \\
\beta &= \text{ the price impact exponent.}
\end{aligned} \tag{2.7}
$$

Here we assume $0 \le \kappa_s < 1$, so that $S_{exec}(v,t) \ge 0$, regardless of the magnitude of v. For various studies which suggest the form (2.7), see [5, 19, 21].

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

$$
\begin{aligned}
g(v) &= \kappa_p v, \\
\kappa_p &= \text{ the permanent price impact factor.}
\end{aligned} \tag{2.8}
$$

As explained in [13], this form of permanent price impact function eliminates the possibility of round-trip price manipulation [5, 17, 2, 15].

2.2. Liquidation Value. Recall that we restrict attention to the selling case in this paper. In this case, we assume that $B(T) = B(T^-)$. Effectively, this penalizes the liquidation strategy if $A(T^-) \ne 0$, since these remaining shares are simply discarded. The optimal strategy should avoid any path where $A(T^-) \ne 0$. This formulation also allows for the (remote) possibility that it may be optimal to simply discard any remaining unsold shares at the end of trading [16].

REMARK 2.1 (Discarding Shares). *Since $B(T) = B(T^-)$, this allows for shares to be discarded at the terminal instant (we do not gain any revenue from these shares). Any strategy which instantaneously discards a finite number of shares at any point in $[0,T)$ cannot be superior to the same strategy which discards the same number of shares at $t = T$. Hence allowing instantaneous discarding of a finite number of shares at the terminal time, produces the same Pareto points as any strategy which allows for discarding shares in $[0,T)$. Effectively this means that the Pareto points computed allowing discarding shares at the terminal time are the same Pareto points as would be computed using the admissible set allowing for discarding shares at any time in $[0,T)$.*

We now introduce some additional notations for subsequent presentation. We use $X(t) = (S(t), B(t), A(t))$ to denote the multi-dimensional process and $x = (s, b, \alpha)$ to denote a state. We will also use the notation $X(t) = x$ as a shorthand for $(S(t), B(t), A(t)) = (s, b, \alpha)$. Let $E^{x,t}_{v(\cdot)}[B(t)]$ be the expectation of $B(T)$ conditional on the initial state (x,t) and on the control $v(\cdot) : (x,t) \mapsto v = v(x,t)$. More specifically, we denote

$$
\begin{aligned}
E[\ \cdot\ ] &: \text{ Expectation operator}, \\
E^{x,t}_{v(\cdot)}[\ \cdot\ ] &: E[\ \cdot\ |X(t) = x] \text{ when observed at time } t \text{ with } v(\cdot) \text{ being the strategy} \\
&\qquad \text{and the stochastic process } X(t) = (S(t), B(t), A(t)) \text{ being given by (2.3-2.5).}
\end{aligned}
$$

Similarly we define $Var^{x,t}_{v(\cdot)}[B(T)]$ as the variance of $B(T)$ conditional on the initial state (x,t) and the control $v(\cdot)$. In addition we introduce the following definitions.

DEFINITION 2.1. *A strategy $v(\cdot) : (x,t) \mapsto v = v(x,t)$ is said to be admissible if $v(x,t) \in [v_{min}, 0]$ and $v(x,t) = 0$ when $A(t) = 0$, where $v_{min} \le 0$. We also require that*

$$
-\int_0^{T^-} v(X(t),t)\ dt \le \alpha_{init}\ . \tag{2.9}
$$

Note that in view of Remark 2.1, since we also permit discarding shares at the terminal time, the Pareto points computed will also be the same Pareto points which allow for instantaneously discarding a finite number of shares at any time in $[0, T]$.

REMARK 2.2 (Admissible Strategies). *The lower bound constraint is not practically restrictive since the continuous trading model is only a proxy for actual discrete trades in the real market practice; the continuous trading rate can be considered to be an averaging of discrete trades over a finite interval. Indeed a continuous model breaks down for extremely small time periods. Our numerical example sets* v_{min} *such that one sixth of the average daily volume is liquidated in* $\simeq 10^{-4}$ *sec. Trading rates this large cannot be observed in practice using any reasonable averaging interval for a continuous trading rate model.*

REMARK 2.3 (Prohibition of Price Manipulation Strategies). *Note that we require* $v \leq 0$ *to prohibit any strategies which involve buying during the course of completing a sell order. Intermediate buying during a sell order is only optimal if the stochastic model admits price manipulation strategies. For the time periods of interest (e.g. less than one day) the drift term* η *in equation (2.5) can be considered negligible. From a mathematical point of view, price-manipulation strategies are possible if a round-trip trade results in positive expected revenues [2, 15] when the drift term* $\eta = 0$. *As pointed out in [2, 15], this is dangerous and unstable in the world of high-frequency trading, and quite possibly illegal. Our requirement that* $v \leq 0$ *for a sell order satisfies one of the regularity conditions for an admissible strategy discussed in [15]. Trading algorithms which violate this condition may result in the following observed unstable market effects due to the interaction of trading algorithms amongst high-frequency traders (HFTs) [10]*

> *"...HFTs began to quickly buy and then resell contracts to each other generating a "hot-potato" volume effect as the same positions were rapidly passed back and forth. Between 2:45:13 and 2:45:27, HFTs traded over 27,000 contracts, which accounted for about 49 percent of the total trading volume, while buying only about 200 additional contracts net."*

3. Mean Variance Pareto Optimal Set.

In this paper, we characterize optimality in terms of the mean and variance values achieved by admissible strategies. We first characterize Pareto optimality and scalarization optimization based on the mean variance objective sets.

DEFINITION 3.1. *Let* $(x_0, 0) = (X(t = 0), t = 0)$ *denote the initial state. Let*

$$\mathcal{Y} = \left\{ (Var_{v(\cdot)}^{x_0,0}[B(T)], E_{v(\cdot)}^{x_0,0}[B(T)]) : v(\cdot) \text{ admissible } \right\} \tag{3.1}$$

denote the **achievable mean-variance objective set** *and* $\bar{\mathcal{Y}}$ *denote its closure.*

DEFINITION 3.2. *An MV point* $(\mathcal{V}_*, \mathcal{E}_*) \in \bar{\mathcal{Y}}$ *is a* **Pareto (optimal) point** *if there exists no admissible strategy* $v(\cdot)$ *such that*

$$\begin{aligned} E_{v(\cdot)}^{x_0,0}[B(T)] &\geq \mathcal{E}_* \\ Var_{v(\cdot)}^{x_0,0}[B(T)] &\leq \mathcal{V}_* \ , \end{aligned} \tag{3.2}$$

and at least one of the inequalities in equation (3.2) is strict. We denote the set of Pareto points by $\mathcal{P} \subseteq \bar{\mathcal{Y}}$. This definition essentially states that the mean-variance tradeoff of a Pareto point cannot be strictly dominated by that of any admissible strategy.

Although the above definition is economically intuitive, solving for $\mathcal{P}$ is a difficult problem since it requires simultaneously optimizing two (conflicting) criteria. A standard scalarization method combines the two criteria into a single objective, using a weighted sum of the two criteria. Specifically, we use a positive weighting parameter $\mu > 0$, and solve the scalarization optimization problem

$$P(x,t;\mu) = \inf_{v(\cdot)} \left\{ \mu Var_{v(\cdot)}^{x,t}[B(T)] - E_{v(\cdot)}^{x,t}[B(T)] \right\}. \tag{3.3}$$

DEFINITION 3.3. *For $\mu > 0$, let*

$$\mathcal{Y}_{P(\mu)} = \left\{ (\mathcal{V}_*, \mathcal{E}_*) \in \bar{\mathcal{Y}} : \mu\mathcal{V}_* - \mathcal{E}_* = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E} \right\} , \tag{3.4}$$

where $\bar{\mathcal{Y}}$ denotes the closure of $\mathcal{Y}$. We denote the **MV scalarization optimal set** *as*

$$\mathcal{Y}_P = \bigcup_{\mu>0} \mathcal{Y}_{P(\mu)}. \tag{3.5}$$

In the context of optimal trade execution, our objective is to determine the set of points $\mathcal{Y}_P$.

REMARK 3.1 (Optimal Strategies). *In practical application, we are also interested in the optimal strategies $v(\cdot)$ which generate $\mathcal{Y}_P$. However, for the purposes of addressing the issues which arise using the embedding technique [25, 18] as part of a numerical algorithm, we define the optimal trade execution problem as determining the set $\mathcal{Y}_P$.*

The original scalarization optimal set $\mathcal{Y}_P$ is with respect to the achievable objective set $\mathcal{Y}$. Since the embedding method and its numerical implementation generate a subset of the achievable MV objectives, we also consider scalarization optimality with respect to a subset.

DEFINITION 3.4. *Let $\mathcal{X}$ be a non-empty subset of $\bar{\mathcal{Y}}$. We define*

$$\mathcal{S}_\mu(\mathcal{X}) = \left\{ (\mathcal{V}_*, \mathcal{E}_*) \in \bar{\mathcal{X}} : \mu\mathcal{V}_* - \mathcal{E}_* = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{X}} \mu\mathcal{V} - \mathcal{E} \right\}. \tag{3.6}$$

where $\bar{\mathcal{X}}$ is the closure of $\mathcal{X}$. We call a point in $\mathcal{S}_\mu(\mathcal{X})$ a scalarization optimal point (SOP) w.r.t. $(\mathcal{X}, \mu)$.

We also define

$$\mathcal{S}(\mathcal{X}) = \left\{ (\mathcal{V}_*, \mathcal{E}_*) : (\mathcal{V}_*, \mathcal{E}_*) \text{ is an SOP w.r.t. } (\mathcal{X}, \mu) \text{ for some } \mu > 0 \right\}. \tag{3.7}$$

We refer to $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{X})$ as **SOP w.r.t.** $\mathcal{X}$.

REMARK 3.2. *Note that Definition 3.4 generalizes Definition 3.3 in the sense that $\mathcal{S}_\mu(\mathcal{Y}) = \mathcal{Y}_{P(\mu)}$ and $\mathcal{S}(\mathcal{Y}) = \mathcal{Y}_P$.*

REMARK 3.3. *A point $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}_\mu(\mathcal{X})$ has the geometric interpretation that, at $(\mathcal{V}_0, \mathcal{E}_0)$, there exists a supporting hyperplane [9] for $\mathcal{X}$ with positive slope μ.*

In general, every point in $\mathcal{Y}_{P(\mu)}$ is in the Pareto optimal set $\mathcal{P}$ but the converse may not hold. If the achievable objective set $\mathcal{Y}$ is convex, however, then every point in $\mathcal{P}$ is in $\mathcal{Y}_{P(\mu)}$ for some $\mu > 0$. This paper is concerned with determining $\bigcup_{\mu>0} \mathcal{Y}_{P(\mu)}$. The more difficult problem of determining the entire set $\mathcal{P}$, in the most general case, is beyond the scope of this paper.

As pointed out in [18, 25], due to the variance term, the value function $P(x, t; \mu)$ is not amenable to solution by means of dynamic programming. To overcome this difficulty, a technique is proposed in [18, 25] to embed the objective in equation (3.3) in the value function below (parameterized by γ)

$$Q(x, t; \gamma) = \inf_{v(\cdot)} \left\{ E^{x,t}_{v(\cdot)}[(B(T) - \gamma/2)^2] \right\}, \tag{3.8}$$

which can be solved by dynamic programming. Note that the strategy $v(\cdot)$ may not be time consistent since $\gamma = \gamma(t, x)$ [8], i.e., γ depends on the initial state.

In [25, 18], it has been shown that an optimal control for the value function $P(x, t; \mu)$ is an optimal solution for the value function $Q(x, t; \gamma)$. We note that an optimal control may not be attained if $\mathcal{Y}$ is not a closed set. In this paper, we discuss optimality with respect to the closed objective set. To be precise, we consider the set of points $\bar{\mathcal{Y}}_P$, which include the limit points of the Pareto optimal points of admissible strategies.

In [1] an alternative to the embedding approach is suggested, which solves Problem 3.4 directly, for a fixed value of μ. The approach in [1] reformulates the problem as a nested minimization problem. The inner minimization requires solution of an HJB equation, with v as the control. This HJB equation contains an additional control, which is the variable for the outer minimization problem. This is perhaps more efficient if it is desired to determine a single point $\mathcal{Y}_{P(\mu)}$. However, the embedding technique described here is undoubtedly more efficient if it is of interest to generate a large number of points in $\mathcal{Y}_P$ (i.e. draw the efficient frontier), which is the objective of this article. This is simply due to the fact that, when using the embedding technique in general, a single point on the frontier is generated with a single HJB equation solve. In addition, if $\mathcal{Y}$ is not convex, then $\mathcal{Y}_{P(\mu)}$ may not be a singleton. In this case, the method in [1] would (apparently) generate only a single point in $\mathcal{Y}_{P(\mu)}$. As a result, varying μ and using the method in [1] may not generate all the points in $\mathcal{Y}_P$. The method we suggest in this paper is theoretically capable of generating all the points in $\mathcal{Y}_P$. In fact, in our particular optimal trade execution application, we can compute the entire efficient frontier using a single HJB solve [12].

4. Preservation of SOP Using the Embedded MV Objective Set. We note that the embedding optimization problem (3.8) is equivalent to

$$\inf_{v(\cdot)} \left\{ E^{x,t}_{v(\cdot)}[(B(T)^2)] - \gamma E^{x,t}_{v(\cdot)}[(B(T))] \right\}.$$

Hence (3.8) is optimization using a scalar combination of the criteria $(\mathcal{Q}, \mathcal{E})$, i.e.,

$$\inf_{v(\cdot)} \left\{ \mathcal{Q} - \gamma \mathcal{E} \right\} \tag{4.1}$$

where $\mathcal{Q} = E^{x,t}_{v(\cdot)}[B(T)^2]$ and $\mathcal{E} = E^{x,t}_{v(\cdot)}[B(T)]$.

To see how the mean and variance are embedded in the scalarization optimization problem (3.8), we note that, from $\mathcal{V} = Var^{x,t}_{v(\cdot)}[B(T)]$,

$$\begin{aligned} &\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} && (4.2)\\ =&Var^{x,t}_{v(\cdot)}[B(T)] + (E^{x,t}_{v(\cdot)}[B(T)])^2 - \gamma E^{x,t}_{v(\cdot)}[B(T)] \\ =&E^{x,t}_{v(\cdot)}[B(T)^2] - (E^{x,t}_{v(\cdot)}[B(T)])^2 + (E^{x,t}_{v(\cdot)}[B(T)])^2 - \gamma E^{x,t}_{v(\cdot)}[B(T)] \\ =&E^{x,t}_{v(\cdot)}[B(T)^2 - \gamma B(T)] \\ =&E^{x,t}_{v(\cdot)}[(B(T) - \gamma/2)^2] - \gamma^2/4. && (4.3) \end{aligned}$$

Since adding a constant term $-\gamma^2/4$ does not change the solution of an optimization problem, the objective in problem (3.8) can be regarded as $Var^{x,t}_{v(\cdot)}[B(T)] + (E^{x,t}_{v(\cdot)}[B(T)])^2 - \gamma E^{x,t}_{v(\cdot)}[B(T)]$. In terms of the mean and variance $(\mathcal{V}, \mathcal{E})$ of $B(T)$, the objective is simply $\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}$. Thus we define the embedded MV objective set to be the set of mean and variance which yields this optimal objective value.

DEFINITION 4.1. *The embedded MV objective set from problem* (3.8) *is*

$$\mathcal{Y}_Q = \bigcup_{-\infty<\gamma<+\infty} \mathcal{Y}_{Q(\gamma)}. \tag{4.4}$$

where

$$\mathcal{Y}_{Q(\gamma)} = \big\{(\mathcal{V}_*, \mathcal{E}_*) \in \bar{\mathcal{Y}} : \mathcal{V}_* + \mathcal{E}_*^2 - \gamma\mathcal{E}_* = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}\big\}. \tag{4.5}$$

For the subsequent analysis, we make the following technical assumption on the achievable objective set $\mathcal{Y}$.

ASSUMPTION 4.1 (Bounded Properties of $\mathcal{Y}$). *We assume that $\mathcal{Y}$ is a non-empty subset of $\{(\mathcal{V}, \mathcal{E}) \in \mathbf{R}^2 : \mathcal{V} \geq 0, \mathcal{E} \leq C_E\}$ for some constant C_E.*

REMARK 4.1. *In the context of our optimal execution problem it can be easily proven that $0 \leq \mathcal{E} \leq C_E$ [22], which is a natural result of forbidding a short position (Definition 2.1) when selling. The assumption $\mathcal{V} \geq 0$ always holds since the variance is non-negative. Although in our context $0 \leq \mathcal{E} \leq C_E$, we need only require that $\mathcal{E} \leq C_E$ in the following.*

Assumption 4.1 immediately leads to the following technical Lemmas.

LEMMA 4.2. *Suppose Assumption 4.1 holds. For any $\mu > 0$, $\mathcal{Y}_{P(\mu)}$ is non-empty, i.e., there exists $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{P(\mu)} \subseteq \bar{\mathcal{Y}}$ such that*

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E}\ . \tag{4.6}$$

Proof. Since $\mu > 0$, $\mathcal{E} \leq C_E$, and $\mathcal{V} \geq 0$ for any $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}$, the objective function $\mu\mathcal{V} - \mathcal{E}$ is bounded below. Hence the result immediately follows. ∎

REMARK 4.2. *If $\mathcal{X}$ is a nonempty subset of $\bar{\mathcal{Y}}$, then for any $\mu > 0$, $\mathcal{S}_\mu(\mathcal{X})$ is non-empty by a trivial generalization of Lemma 4.2.*

LEMMA 4.3. *Suppose Assumption 4.1 holds. If $(\mathcal{V}', \mathcal{E}') \in \bar{\mathcal{Y}}$, then*

$$\mu\mathcal{V}' - \mathcal{E}' \geq \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E}. \tag{4.7}$$

Similarly,

$$\mathcal{V}' + \mathcal{E}'^2 - \gamma\mathcal{E}' \geq \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}. \tag{4.8}$$

Proof. From Lemma 4.2, $\inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E}$ exists. Similarly, writing equation (4.8) as $\mathcal{V} + (\mathcal{E} - \gamma/2)^2 - \gamma^2/4$, and using Assumption 4.1 implies that $\inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}$ also exists. The results immediately follow since the objective functions are continuous. □

Next we present a characterization of the main property of the embedding technique given in [18, 25] in terms of the achievable objective set.

THEOREM 4.4. *Suppose Assumption 4.1 holds. Let $(\mathcal{V}_0, \mathcal{E}_0) \in \bar{\mathcal{Y}}$ and $\mu > 0$ be such that*

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E}, \quad i.e., \ (\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{P(\mu)}. \tag{4.9}$$

Then

$$\mathcal{V}_0 + \mathcal{E}_0^2 - \gamma\mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}, \quad i.e., \ (\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{Q(\gamma)}, \tag{4.10}$$

where

$$\gamma = \frac{1}{\mu} + 2\mathcal{E}_0. \tag{4.11}$$

We include the proof of this result from [18, 25] below, since we will use some of the similar steps to prove our new results.

Proof. Assume to the contrary that (4.10) does not hold. Then, by Lemma 4.3,

$$\inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} < \mathcal{V}_0 + \mathcal{E}_0^2 - \gamma\mathcal{E}_0. \tag{4.12}$$

Then there exists $(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}$ such that

$$\mathcal{V}_* + \mathcal{E}_*^2 - \gamma\mathcal{E}_* < \mathcal{V}_0 + \mathcal{E}_0^2 - \gamma\mathcal{E}_0.$$

Rearranging and multiplying by $\mu > 0$ gives

$$\mu\big(\mathcal{V}_* + \mathcal{E}_*^2 - (\mathcal{V}_0 + \mathcal{E}_0^2)\big) - \gamma\mu(\mathcal{E}_* - \mathcal{E}_0) < 0 \tag{4.13}$$

Define the function

$$\pi^\mu(v, e) = \mu v - \mu e^2 - e. \tag{4.14}$$

Note that

$$\pi^\mu(v + e^2, e) = \mu v + \mu e^2 - \mu e^2 - e = \mu v - e \ , \tag{4.15}$$

and let

$$\pi_v^\mu = \frac{\partial \pi^\mu}{\partial v}; \qquad \pi_e^\mu = \frac{\partial \pi^\mu}{\partial e} \ . \tag{4.16}$$

Since $\pi^\mu(v, e)$ is a concave quadratic in (v, e), we have,

$$\begin{aligned} \pi^\mu(v + \triangle v, e + \triangle e) &\le \pi^\mu(v, e) + \pi_v^\mu(v, e)\triangle v + \pi_e^\mu(v, e)\triangle e \\ &= \pi^\mu(v, e) + \mu\triangle v - (1 + 2\mu e)\triangle e. \end{aligned} \tag{4.17}$$

A direct application of (4.17) gives

$$\begin{aligned}\pi^{\mu}(\mathcal{V}_* + \mathcal{E}_*^2, \mathcal{E}_*) &\leq \pi^{\mu}(\mathcal{V}_0 + \mathcal{E}_0^2, \mathcal{E}_0) + \mu\big(\mathcal{V}_* + \mathcal{E}_*^2 - (\mathcal{V}_0 + \mathcal{E}_0^2)\big) - (1 + 2\mu\mathcal{E}_0)(\mathcal{E}_* - \mathcal{E}_0) \\ &= \pi^{\mu}(\mathcal{V}_0 + \mathcal{E}_0^2, \mathcal{E}_0) + \mu\big(\mathcal{V}_* + \mathcal{E}_*^2 - (\mathcal{V}_0 + \mathcal{E}_0^2)\big) - \gamma\mu(\mathcal{E}_* - \mathcal{E}_0) \\ &< \pi^{\mu}(\mathcal{V}_0 + \mathcal{E}_0^2, \mathcal{E}_0), \end{aligned} \tag{4.18}$$

where we have used (4.11) in the equality and (4.13) in the last inequality.

By (4.15), the strict inequality (4.18) means that

$$\mu\mathcal{V}_* - \mathcal{E}_* < \mu\mathcal{V}_0 - \mathcal{E}_0,$$

which contradicts equation (4.9). Hence (4.10) holds. ◻

It is immediate that the following holds.

COROLLARY 4.5. *Suppose Assumption 4.1 holds. Then* $\mathcal{Y}_P \subseteq \mathcal{Y}_Q$.

Now we are ready to establish that the embedding technique preserves the scalarization optimal point set $\mathcal{Y}_P$.

LEMMA 4.6. *Assume Assumption 4.1 holds. For any* $\mu > 0$*,*

$$\inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E} = \inf_{(\mathcal{V}',\mathcal{E}')\in\mathcal{Y}_Q} \mu\mathcal{V}' - \mathcal{E}'. \tag{4.19}$$

Proof. Let $(\mathcal{V}_0, \mathcal{E}_0)$ be a SOP w.r.t. $(\mathcal{Y}, \mu)$. By Corollary 4.5, $\mathcal{Y}_P \subseteq \mathcal{Y}_Q$, hence $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_Q$. Consequently,

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E} \geq \inf_{(\mathcal{V}',\mathcal{E}')\in\mathcal{Y}_Q} \mu\mathcal{V}' - \mathcal{E}'. \tag{4.20}$$

Equality follows since the reverse inequality

$$\inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E} \leq \inf_{(\mathcal{V}',\mathcal{E}')\in\mathcal{Y}_Q} \mu\mathcal{V}' - \mathcal{E}' \tag{4.21}$$

holds by $\mathcal{Y}_Q \subseteq \bar{\mathcal{Y}}$. ◻

THEOREM 4.7. *Suppose Assumption 4.1 holds. The SOPs w.r.t.* $\mathcal{Y}_Q$ *are the same as the SOPs w.r.t.* $\mathcal{Y}$*, i.e.,*

$$\mathcal{S}(\mathcal{Y}_Q) = \mathcal{Y}_P = \mathcal{S}(\mathcal{Y}). \tag{4.22}$$

Proof. From Corollary 4.5, we have that $\mathcal{Y}_P \subseteq \mathcal{Y}_Q$. By definition, $\mathcal{Y}_Q \subseteq \bar{\mathcal{Y}}$. Suppose $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y}_Q)$. Hence there exists $\mu > 0$ such that

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}_Q} \mu\mathcal{V} - \mathcal{E}\ . \tag{4.23}$$

Since $(\mathcal{V}_0, \mathcal{E}_0) \in \bar{\mathcal{Y}}$, then from Lemma 4.6,

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E}\ . \tag{4.24}$$

Thus $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y})$. On the other hand, suppose $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y})$. Then

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu\mathcal{V} - \mathcal{E}\ . \tag{4.25}$$

Since $\mathcal{S}(\mathcal{Y}) = \mathcal{Y}_P \subseteq \mathcal{Y}_Q$, we have $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_Q$. From Lemma 4.6,

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}_Q} \mu\mathcal{V} - \mathcal{E} \ , \tag{4.26}$$

hence $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y}_Q)$. ◻

Before concluding this section, we establish a uniqueness property: if $(\mathcal{V}, \mathcal{E})$ is a SOP with respect to $\mathcal{Y}_Q$ for some embedding parameter γ, then $(\mathcal{V}, \mathcal{E})$ is the unique point in $\mathcal{Y}_{Q(\gamma)}$.

THEOREM 4.8. *Suppose Assumption 4.1 holds. If $(\mathcal{V}, \mathcal{E}) \in \mathcal{S}(\mathcal{Y}_Q)$, then there exists γ such that $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}_{Q(\gamma)}$ and $\mathcal{Y}_{Q(\gamma)}$ is a singleton.*

Proof. Let $(\mathcal{V}_*, \mathcal{E}_*)$ be a SOP w.r.t. $\mathcal{Y}_Q$ for some μ^*. By Lemma 4.6, $(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}_{P(\mu^*)}$. Hence, following Theorem 4.4, there exists γ^* such that

$$(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}_{Q(\gamma^*)}, \text{ where } \gamma^* = \frac{1}{\mu^*} + 2\mathcal{E}_*. \tag{4.27}$$

Suppose there is another $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{Q(\gamma^*)}$. Since both points are in $\mathcal{Y}_{Q(\gamma^*)}$ we have that

$$\mathcal{V}_* + \mathcal{E}_*^2 - \gamma^*\mathcal{E}_* = \mathcal{V}_0 + \mathcal{E}_0^2 - \gamma^*\mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma^*\mathcal{E} \ , \tag{4.28}$$

Hence

$$\mathcal{V}_0 + \mathcal{E}_0^2 - (\mathcal{V}_* + \mathcal{E}_*^2) - \gamma^*(\mathcal{E}_0 - \mathcal{E}_*) = 0 \ . \tag{4.29}$$

Consider the function $\pi^{\mu^*}(v, e) = \mu^* v - \mu^* e^2 - e$ as in Theorem 4.4. Following similar steps as in the proof of Theorem 4.4, we obtain (using equations (4.27) and (4.29))

$$\begin{aligned} \pi^{\mu^*}(\mathcal{V}_0 + \mathcal{E}_0^2, \mathcal{E}_0) &\le \pi^{\mu^*}(\mathcal{V}_* + \mathcal{E}_*^2, \mathcal{E}_*) + \mu^*(\mathcal{V}_0 + \mathcal{E}_0^2 - (\mathcal{V}_* + \mathcal{E}_*^2)) - (1 + 2\mu^*\mathcal{E}_*)(\mathcal{E}_0 - \mathcal{E}_*) \\ &= \pi^{\mu^*}(\mathcal{V}_* + \mathcal{E}_*^2, \mathcal{E}_*) + \mu^*(\mathcal{V}_0 + \mathcal{E}_0^2 - (\mathcal{V}_* + \mathcal{E}_*^2) - \gamma^*(\mathcal{E}_0 - \mathcal{E}_*)) \\ &= \pi^{\mu^*}(\mathcal{V}_* + \mathcal{E}_*^2, \mathcal{E}_*) \ . \end{aligned} \tag{4.30}$$

Recalling that $\pi^{\mu^*}(v + e^2, e) = \mu^* v - e$, then equation (4.30) yields

$$\mu^*\mathcal{V}_0 - \mathcal{E}_0 \le \mu^*\mathcal{V}_* - \mathcal{E}_*. \tag{4.31}$$

Since $(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}_{P(\mu^*)}$ and $(\mathcal{V}_0, \mathcal{E}_0) \in \bar{\mathcal{Y}}$,

$$\mu^*\mathcal{V}_* - \mathcal{E}_* = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{Y}} \mu^*\mathcal{V} - \mathcal{E} \le \mu^*\mathcal{V}_0 - \mathcal{E}_0. \tag{4.32}$$

Hence

$$\mu^*\mathcal{V}_0 - \mathcal{E}_0 = \mu^*\mathcal{V}_* - \mathcal{E}_* \ . \tag{4.33}$$

Rewrite equations (4.29) and (4.33) as

$$\mu^*(\mathcal{V}_* - \mathcal{V}_0) - (\mathcal{E}_* - \mathcal{E}_0) = 0 \tag{4.34}$$

$$(\mathcal{V}_* - \mathcal{V}_0) + (\mathcal{E}_* - \mathcal{E}_0)(\mathcal{E}_* + \mathcal{E}_0 - \gamma^*) = 0 \ . \tag{4.35}$$

Using equation (4.27), equation (4.35) becomes

$$(\mathcal{V}_* - \mathcal{V}_0) + (\mathcal{E}_* - \mathcal{E}_0)(\mathcal{E}_0 - \mathcal{E}_* - 1/\mu^*) = 0 \ . \tag{4.36}$$

Solving equations (4.34) and equation (4.36) for $(\mathcal{E}_* - \mathcal{E}_0)$ gives the unique solution $\mathcal{E}_* = \mathcal{E}_0$ and $\mathcal{V}_* = \mathcal{V}_0$. ◻

REMARK 4.3 (Properties of $\mathcal{Y}_{Q(\gamma)}$). *For a fixed γ, $\mathcal{Y}_{Q(\gamma)}$ is either*

- *A singleton containing a SOP w.r.t. $\mathcal{Y}_Q$, or*
- *A set which may contain any number of elements. If any of these elements are SOP w.r.t. $\mathcal{Y}_Q$, then these elements are singleton members of $\mathcal{Y}_Q(\gamma')$, $\gamma' \neq \gamma$.*

Moreover, for $(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{S}(\mathcal{Y}_Q)$, given the optimal objective value $(\mathcal{V}_* + \mathcal{E}_*^2 - \gamma\mathcal{E}_*)$, Theorem 4.8 allows us to uniquely reconstruct $\mathcal{V}_*$ given $\mathcal{E}_*$. Note that $\mathcal{E}_*$ can be easily determined from the optimal control of Problem (4.1). We further note that the optimal control $v(\cdot)$ which generates a given point in $\mathcal{S}(\mathcal{Y}_Q)$ may not be unique.

5. Preservation of SOP Using the Computed Embedded MV Set. In §4, we have established that the set $\mathcal{Y}_P$ of MV SOPs is preserved using the embedding method in the sense that the set of the embedded MV points, which yield scalarization optimal values for the embedded optimization problems, is identical to $\mathcal{Y}_P$. This is an interesting theoretical property illustrating the ability of the embedding method to generate the original MV SOP set $\mathcal{Y}_P$. A spurious point corresponds to a point in $\mathcal{Y}_Q$ at which there does not exist a supporting hyperplane with positive slope supporting $\mathcal{Y}_Q$. However, this property does not have immediate practical use in computation since the achievable objective set $\mathcal{Y}_Q$ is not available in the context of computation.

Furthermore, for each embedding parameter $-\infty < \gamma < +\infty$, we can only expect a numerical algorithm to generate a single embedded MV point $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}_Q$. Specifically, a possible computational technique to determine the embedded MV set is as follows:

(a) For each embedding parameter $-\infty < \gamma < +\infty$, solve the embedding optimization problem (3.8) to determine a single optimal control $v_\gamma^*(\cdot)$.
(b) Compute the corresponding MV point $(\mathcal{V}_\gamma^*, \mathcal{E}_\gamma^*)$.
(c) Determine the computed MV set

$$\mathcal{Y}_Q^\dagger = \bigcup_{-\infty<\gamma<+\infty} \{(\mathcal{V}_\gamma^*, \mathcal{E}_\gamma^*)\} \ . \tag{5.1}$$

In general only one out of possibly many optimal controls (which all minimize $\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}$) is selected by the above algorithm. We denote the subset of $\mathcal{Y}_Q$ generated by this algorithm as $\mathcal{Y}_Q^\dagger$. In view of Remark 4.3 we define $\mathcal{Y}_Q^\dagger$ as follows

Definition 5.1 (Numerical $\mathcal{Y}_Q$). *Let $\mathcal{Y}_{Q(\gamma)}^\dagger$ be a singleton subset of $\mathcal{Y}_{Q(\gamma)}$. Specifically $\mathcal{Y}_{Q(\gamma)}^\dagger$ contains either*

- *the unique single point which is SOP w.r.t. $\mathcal{Y}_Q$ if $\mathcal{Y}_{Q(\gamma)}$ is the singleton set containing a point SOP w.r.t. $\mathcal{Y}_Q$, or*
- *an arbitrarily selected single point of $\mathcal{Y}_Q$ otherwise.*

The computed MV objective set is then defined as

$$\mathcal{Y}_Q^\dagger = \bigcup_{-\infty<\gamma<+\infty} \mathcal{Y}_{Q(\gamma)}^\dagger = \bigcup_{-\infty<\gamma<+\infty} \{(\mathcal{V}_\gamma^*, \mathcal{E}_\gamma^*)\}.$$

Following Definition 5.1, we immediately have the following properties for $\mathcal{Y}_Q^\dagger$.

Lemma 5.2. *Suppose Assumption 4.1 holds. Then $\mathcal{Y}_{Q(\gamma)}^\dagger$ has the following properties:*

(a) $\mathcal{Y}_Q^\dagger \subseteq \mathcal{Y}_Q$.
(b) $\mathcal{S}(\mathcal{Y}_Q) \subseteq \mathcal{Y}_Q^\dagger$.
(c) $\mathcal{Y}_P \subseteq \mathcal{Y}_Q^\dagger$.

Proof. From Definition 5.1, $\mathcal{Y}_Q^\dagger \subseteq \mathcal{Y}_Q$ clearly holds.

Assume that $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{S}(\mathcal{Y}_Q)$, applying Theorem 4.8, then

$$\exists \gamma, \quad \text{such that} \quad (\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{Y}_{Q(\gamma)},$$

which contains a single point. Using Definition 5.1, $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{Y}_Q^\dagger$. Thus $\mathcal{S}(\mathcal{Y}_Q) \subseteq \mathcal{Y}_Q^\dagger$.

In addition, Theorem 4.7 implies that $\mathcal{Y}_P = \mathcal{S}(\mathcal{Y}_Q)$. Using (b), we conclude (c) holds. ◻

We now show that it is possible to identify and remove spurious points from only the computed MV points $\mathcal{Y}_Q^\dagger$. Similar to the approach with respect to $\mathcal{Y}_Q$, the main idea is to consider SOP with respect to the computed MV set $\mathcal{Y}_Q^\dagger$. We first establish an auxiliary Lemma.

LEMMA 5.3. *Suppose Assumption 4.1 holds. For any $\mu > 0$,*

$$\inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}_Q} \mu\mathcal{V} - \mathcal{E} = \inf_{(\mathcal{V}',\mathcal{E}') \in \mathcal{Y}_Q^\dagger} \mu\mathcal{V}' - \mathcal{E}'. \tag{5.2}$$

Proof. Let $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_Q$ be SOP w.r.t. $(\mathcal{Y}_Q, \mu)$. By Theorem 4.8, there exists γ, such that $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{Q(\gamma)}$ and $\mathcal{Y}_{Q(\gamma)}$ is a singleton. Hence $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_Q^\dagger$ by Lemma 5.2.

This implies that

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}_Q} \mu\mathcal{V} - \mathcal{E} \geq \inf_{(\mathcal{V}',\mathcal{E}') \in \mathcal{Y}_Q^\dagger} \mu\mathcal{V}' - \mathcal{E}'.$$

The reverse inequality holds since $\mathcal{Y}_Q^\dagger \subseteq \mathcal{Y}_Q$. ◻

Next we establish that SOP with respect to $\mathcal{Y}_Q^\dagger$ preserves SOP with respect to $\mathcal{Y}$.

THEOREM 5.4. *Suppose Assumption 4.1 holds. Then*

$$\mathcal{S}(\mathcal{Y}_Q^\dagger) = \mathcal{Y}_P = \mathcal{S}(\mathcal{Y}) \ . \tag{5.3}$$

Proof. By Theorem 4.7, we know that $\mathcal{S}(\mathcal{Y}_Q) = \mathcal{Y}_P$. Hence we need only show that $\mathcal{S}(\mathcal{Y}_Q^\dagger) = \mathcal{S}(\mathcal{Y}_Q)$.

Suppose $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y}_Q)$. Hence there exists $\mu > 0$ such that

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}_Q} \mu\mathcal{V} - \mathcal{E} \ . \tag{5.4}$$

From Lemma 5.2, $\mathcal{S}(\mathcal{Y}_Q) \subseteq \mathcal{Y}_Q^\dagger$, hence $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_Q^\dagger$. From Lemma 5.3,

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}_Q^\dagger} \mu\mathcal{V} - \mathcal{E} \ , \tag{5.5}$$

and $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y}_Q^\dagger)$.

On the other hand, suppose $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y}_Q^\dagger)$. Then

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}_Q^\dagger} \mu\mathcal{V} - \mathcal{E} \ . \tag{5.6}$$

From Lemma 5.2, $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_Q$. Following Lemma 5.3,

$$\mu\mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}_Q} \mu\mathcal{V} - \mathcal{E} \ , \tag{5.7}$$

hence $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{S}(\mathcal{Y}_Q)$. □

Theorem 5.4 implies that the set $\mathcal{S}(\mathcal{Y}_Q^\dagger)$ is identical to $\mathcal{Y}_P$ which contains all the MV Pareto points that can be obtained by scalarization of the original MV Pareto problem. This is, of course, the best that can be done, given that the embedded technique is designed to solve the scalarization optimization problem for the MV Pareto problem. Following Theorem 5.4, a MV point $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}_Q^\dagger$ is spurious if there exists no supporting hyperplane at $(\mathcal{V}, \mathcal{E})$ with a positive slope for $\mathcal{Y}_Q^\dagger$.

REMARK 5.1 (Significance of Theorem 5.4). *A numerical algorithm can be used to generate $\mathcal{Y}_Q^\dagger$. The set of points $\mathcal{S}(\mathcal{Y}_Q^\dagger)$ is thus identical to the the set of Pareto points $\mathcal{Y}_P$ that is obtained by scalarization of the original MV Pareto problem.*

6. SOP for a Finite Set.

Finally we establish a property for SOPs for a set containing a finite number of points. This result will be used in the postprocessing technique described in §7.2 to identify SOPs w.r.t. $\mathcal{Y}$ based on an approximate MV set, which has only a finite number of points.

Assume that $\mathcal{A} = \{(\mathcal{V}_i, \mathcal{E}_i) : i = 1, \cdots, N\}$ is a finite set and **conv** $\mathcal{A}$ denotes the convex hull of $\mathcal{A}$. Define $\mathcal{C}^*(\mathcal{A})$ as the upper-left boundary of **conv** $\mathcal{A}$, i.e.,

$$\mathcal{C}^*(\mathcal{A}) = \{(\mathcal{V}^*, \mathcal{E}^*) : (\mathcal{V}^*, -\mathcal{E}^*) \text{ is a minimal element of } \mathbf{conv}\{(\mathcal{V}, -\mathcal{E}) : \ (\mathcal{V}, \mathcal{E}) \in \mathcal{A}\}\}. \tag{6.1}$$

Here a minimal element is with respect to the componentwise inequality, see, e.g., [9]. We show next that $\mathcal{S}(\mathcal{A})$ can be obtained from the upper-left boundary $\mathcal{C}^*(\mathcal{A})$ of **conv** $\mathcal{A}$.

THEOREM 6.1. *Assume that the set $\mathcal{A} = \{(\mathcal{V}_i, \mathcal{E}_i) : i = 1, \cdots, N\}$ has a finite number of points. Let $\mathcal{C}^*(\mathcal{A})$ be the upper-left boundary of* **conv** $\mathcal{A}$ *defined in* (6.1). *Then*

$$\mathcal{S}(\mathcal{A}) = \mathcal{C}^*(\mathcal{A}) \cap \mathcal{A}. \tag{6.2}$$

Proof. Assume $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{C}^*(\mathcal{A}) \cap \mathcal{A}$. Since the set $\mathbf{conv}\{(\mathcal{V}, -\mathcal{E}) : \ (\mathcal{V}, \mathcal{E}) \in \ \mathcal{A}\}$ is convex, and $\mathcal{A}$ has a finite number of points, following a dual characterization of minimal elements, see, e.g., [9], there exists $\mu > 0$, such that $(\mathcal{V}^*, \mathcal{E}^*)$ solves

$$\inf_{(\mathcal{V}, -\mathcal{E}) \in \mathbf{conv}\{(\mathcal{V}, -\mathcal{E}) : \ (\mathcal{V}, \mathcal{E}) \in \ \mathcal{A}\}} \mu\mathcal{V} - \mathcal{E}. \tag{6.3}$$

Since $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{A}$, this implies that $(\mathcal{V}^*, \mathcal{E}^*)$ solves

$$\inf_{(\mathcal{V}, \mathcal{E}) \in \mathcal{A}} \mu\mathcal{V} - \mathcal{E}.$$

Consequently $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{S}(\mathcal{A})$.

Conversely, let $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{S}(\mathcal{A})$. This implies that $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{A}$ and $(\mathcal{V}^*, \mathcal{E}^*)$ solves,

$$\inf_{(\mathcal{V}, \mathcal{E}) \in \mathcal{A}} \mu\mathcal{V} - \mathcal{E}$$

for some $\mu > 0$. Since $\mathcal{A}$ has a finite number of points, it can be easily shown (by contradiction) that $(\mathcal{V}^*, \mathcal{E}^*)$ solves

$$\inf_{(\mathcal{V}, -\mathcal{E}) \in \mathbf{conv}\{(\mathcal{V}, -\mathcal{E}) : \ (\mathcal{V}, \mathcal{E}) \in \ \mathcal{A}\}} \mu\mathcal{V} - \mathcal{E}.$$

Following the sufficient condition for a minimal element of a set, as given in [9], $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{C}^*(\mathcal{A})$. Consequently $(\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{C}^*(\mathcal{A}) \cap \mathcal{A}$. This completes the proof. □

7. Optimal Trade Execution: a Computational Example. In this section we use the optimal trade execution problem, with the objective of determining $\mathcal{Y}_P$, to illustrate how to use the mathematical properties established in §4 and §5 to post-process an efficient frontier computed using the embedding technique. This problem is introduced in §2.

7.1. Computing $\mathcal{Y}^{\dagger}_{Q(\gamma)}$. The computed MV set $\mathcal{Y}^{\dagger}_{Q(\gamma)}$ can be determined by solving an HJB PDE and then running Monte Carlo simulations as follows. We refer readers to [23] for more detail. Let $\tau = T - t$ and

$$V(s,b,\alpha,\tau;\gamma) = \inf_{v(\cdot)} \left\{ E^{x,t}_{v(\cdot)} \left[\left(B(T) - \frac{\gamma}{2} \right)^2 \right] \right\} \tag{7.1}$$

Using standard dynamic programming, $V(s,b,\alpha,\tau;\gamma)$ is the viscosity solution to the Hamilton Jacobi Bellman PDE:

$$V_\tau = \frac{\sigma^2 s^2}{2} V_{ss} + \eta s V_s + rbV_b + \inf_{v \in [v_{\min},0]} \left\{ -vsf(v)V_b + vV_\alpha + g(v)V_s \right\}, \tag{7.2}$$

with the initial condition

$$V(s,b,\alpha,0;\gamma) = \left(b - \frac{\gamma}{2} \right)^2 . \tag{7.3}$$

We solve equation (7.2) using a finite difference method as described in [12].

Let the optimal control of problem (7.2) be denoted by $v^*_\gamma(\cdot)$. Once we have determined $v^*_\gamma(\cdot)$, we can use Monte Carlo simulations to compute the embedded MV points:

$$(\mathcal{V}^*_\gamma, \mathcal{E}^*_\gamma) = \left(Var^{x_0,0}_{v^*_\gamma(\cdot)}\left[B(T)\right], E^{x_0,0}_{v^*_\gamma(\cdot)}\left[B(T)\right] \right). \tag{7.4}$$

REMARK 7.1 (Non-unique controls in optimal trade execution). *We can see immediately from equation (7.2) that if $V_b = V_\alpha = V_s = 0$ at (s,b,α,τ), then the optimal control can be arbitrary at this point. In fact, the numerical results in [12] demonstrate existence of large regions where the value function is flat, suggesting non-unique optimal strategies which give essentially the same value of the objective function.*

7.2. Numerical Estimates of $\mathcal{Y}^{\dagger}_{Q}$. Our theoretical result in Theorem 5.4 establishes that $\mathcal{S}(\mathcal{Y}^{\dagger}_{Q}) = \mathcal{S}(\mathcal{Y})$. This implies that we can determine whether a MV point $(\mathcal{V},\mathcal{E})$ is in $\mathcal{Y}_P$ by checking whether it is an SOP with respect to $\mathcal{Y}^{\dagger}_{Q}$. This requires that the entire set $\mathcal{Y}^{\dagger}_{Q}$ is available. In practice, however, we can compute $(\mathcal{V}^*_\gamma, \mathcal{E}^*_\gamma)$ only for a finite number of of $\gamma \in (-\infty, \infty)$ values.

More precisely, $\mathcal{Y}^{\dagger}_{Q}$ needs to be approximated in two aspects:

(a) $\mathcal{Y}^{\dagger}_{Q(\gamma)}$ can be computed for only a finite number of γ values, giving rise to a *finite set error*. In other words, we compute only a (finite) subset of $\mathcal{Y}^{\dagger}_{Q}$.

(b) For a fixed γ, $\mathcal{Y}^{\dagger}_{Q(\gamma)}$ needs to be approximated by a sequence converging to $\mathcal{Y}^{\dagger}_{Q(\gamma)}$, due to PDE discretization errors, Monte Carlo sampling error, and timestepping errors.

We denote by $(\mathcal{Y}_Q^\dagger)^k$ an approximation computed with a fixed grid size for γ, a fixed mesh size (for the numerical PDE solve), and a fixed number of Monte Carlo simulations (using a fixed timestep to solve the SDEs). The solution for $(\mathcal{Y}_Q^\dagger)^{k+1}$ uses a finer grid for γ, a finer mesh for PDE, increased Monte Carlo simulations, and timesteps.

In practice, we can compute a sequence of approximations $(\mathcal{Y}_Q^\dagger)^k$ and generate $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big)$. We expect convergence to occur, in the sense that the difference between $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^{k+1}\big)$ and $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big)$ is sufficiently small for sufficiently large k. We also assume that for k sufficiently large, $\mathcal{S}(\mathcal{Y}_Q^\dagger)$ can be arbitrarily well approximated by $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big)$. Provided we use a convergent method to solve the HJB PDE, and an increasing number of Monte Carlo simulations on each refinement level k, we do not expect that the PDE discretization and the Monte Carlo errors are of much concern. More importantly, it is not obvious what properties are required to ensure that the finite sampling of the γ values (the finite set error) will produce a good approximation to $\mathcal{S}(\mathcal{Y}_Q^\dagger)$ as k becomes large. This will require a precise definition of convergence in terms of sets, and a precise requirement on the sampling method for the set of γ values. We conjecture that any reasonable sampling method (i.e. a systematically refined uniform grid) for γ will produce a good approximation as k becomes large, but we have no proof of this. We leave these important questions to future work.

7.3. Relevant Range for γ. For the case of optimal trade execution with buying rate $v \le 0$, we have $0 \le \mathcal{E} \le C_E$, with C_E a positive constant and $\gamma \in (0, \infty)$. In practice we are only interested in a subrange of γ. Recall that for MV scalarization optimal points with respect to $\mathcal{Y}$,

$$\gamma = \frac{1}{\mu} + 2E_{v^*(\cdot)}\left[B(T)\right], \tag{7.5}$$

where $1/\mu$ can be regarded as a risk aversion coefficient. If μ is large, then $\gamma \simeq 2E_{v^*(\cdot)}\left[B(T)\right]$. Large μ would correspond to the lower end of the MV frontier. In practice, we are only interested in strategies which do not have too large a price impact. This roughly corresponds to strategies with the expected cash flow per share at least 99% of the arrival price. If S_0 is the arrival price, then the smallest value of γ of interest would thus typically be $\gamma \simeq 2 * (.99S_0)$. The upper end of the MV frontier corresponds to trading at a constant rate which completes the trade. This constant trading strategy is known to maximize the expected cash flow under typical assumptions. Since we are only interested in practically relevant strategies, we consider $\gamma < 4 * S_0$. Hence a range of $\gamma \in [2 * (.99S_0), 4 * S_0]$ is a good estimate for the useful section of the efficient frontier.

7.4. Computing $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big)$. As discussed in Section 7.2, we assume that $\mathcal{S}\big(\mathcal{Y}_Q^\dagger\big)$ can be identified by $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big)$ as $k \to \infty$. Note however that $(\mathcal{Y}_Q^\dagger)^k$ is a discrete set. We now describe an approach for determining $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big)$, based on the result in Theorem 6.1.

Standard algorithms exist for generating the (vertices of the) convex hull of a finite set of points, see, e.g., [6]. Consequently the upper-left boundary $\mathcal{C}^*(\mathcal{A})$ of the convex hull $\textbf{conv}\mathcal{A}$ can be determined by starting from the left-most vertex of $\textbf{conv}\mathcal{A}$ and ending at the top-most vertex of $\textbf{conv}\mathcal{A}$ by going clockwise. This process is illustrated in Figure 7.1. (If there are multiple left-most/top-most vertices, the upper-left convex hull starts with the top-most left-most vertex and ends with the right-most top-most vertex).

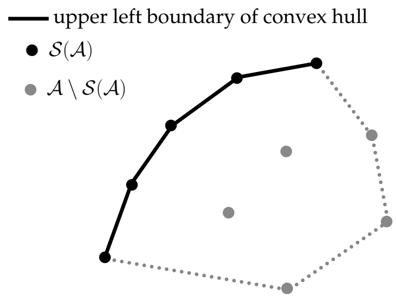


Fig. 7.1: *Given a finite set of points $\mathcal{A}$, the solid line shows the upper left boundary of the convex hull of $\mathcal{A}$. The solid dots are the scalerization optimal points $\mathcal{S}(\mathcal{A})$.*

σ	η	r	κ_p	κ_t	κ_s	s_{init}	α_{init}	β	T	v_{min}
0.2	0.0	0.0	0.0	3×10^{-5}	0.0	100	1.0	0.5	1/250	-10^6/T

Table 7.1: *Parameter Values for the Optimal Trade Execution Example*

7.5. Numerical Results. Recall that $(\mathcal{Y}_Q^\dagger)^k$ denotes an approximation to $\mathcal{Y}_Q^\dagger$, computed using a finite number of values of γ and a finite mesh size, where k is the refinement index. We then apply the post-processing step (described in §7.4) to $(\mathcal{Y}_Q^\dagger)^k$. This leads to determination of $\mathcal{S}((\mathcal{Y}_Q^\dagger)^k)$, using Theorem 6.1. If convergence occurs, this will provide an increasingly accurate estimate of $\mathcal{S}(\mathcal{Y})$ as k increases. We illustrate this by a numerical example.

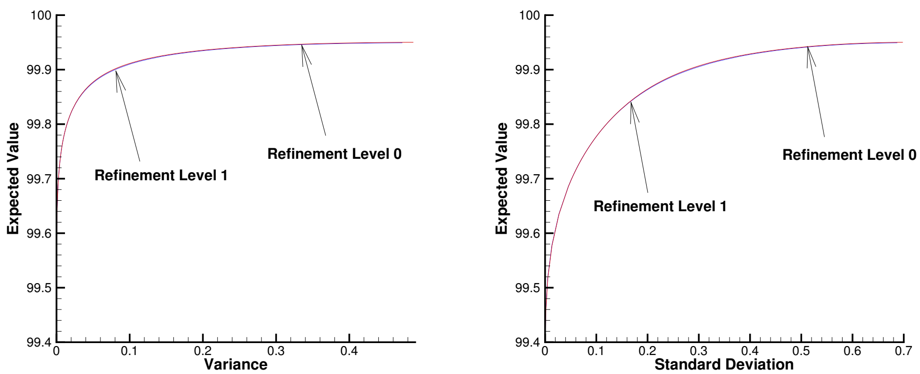
Table 7.1 summarizes the parameter values in our example. The price impact factor κ_t corresponds roughly to liquidating one-sixth of the daily trading volume of a large-cap stock; see [23] for how this estimation is done. The estimate in [23] uses $\beta = 1$. For our value of $\beta = 0.5$, the calculation needs to be modified slightly.

Subplot (a) in Figure 7.2 graphs $(\mathcal{Y}_Q^\dagger)^k$ for two grid refinement levels (corresponding to parameters in Table 7.2). We note that the efficient frontier for refinement level zero visually coincides that for refinement level one; this suggests convergence of the numerical solution and the frontier. Subplot (b) shows a curve of expected cash flow versus standard deviation, which is a more practically meaningful display of the results because standard deviation and expected value have the same units. Since the number of γ values is quite large, the computed efficient frontiers appear smooth. Note that the method used in [12] generates an arbitrary number of points along the efficient frontier (i.e. many different values of γ) from a single solution of the HJB PDE.

From Figure 7.2, we see that for this example every point of $(\mathcal{Y}_Q^\dagger)^k$ lies on the upper-left boundary of the convex hull of $(\mathcal{Y}_Q^\dagger)^k$. Therefore, every point in $(\mathcal{Y}_Q^\dagger)^k$ is in its MV scalarization optimal set, following Theorem 5.4. This suggests that, in this case, the scalarization formulation generates all the Pareto points. Of course this will not be true in general due to the fact that the achievable objective set $\mathcal{Y}$ can be non-convex, since $B(T)$ is a nonlinear function of the control v in equation (2.7). We expect that, in some cases, the post-processing algorithm will generate *gaps* in the efficient frontier, corresponding to cases where the scalarization formulation does not generate all the Pareto points.

Refine Level	Timesteps	s nodes	b nodes	α nodes	v nodes	MC sample size	γ nodes
0	2000	369	1	11	8	10,000	65
1	4000	737	1	21	15	40,000	129

TABLE 7.2: *Computational grid for both solving HJB PDE equation* (7.1) *and running Monte Carlo simulations. There is only one node in the b direction since we use a similarity reduction to eliminate a variable. We refer an interested reader to [23] for more details.*



(a) *Expected value against variance.*

(b) *Expected value against standard deviation.*

FIG. 7.2: *Plot of* $(\mathcal{Y}_Q^\dagger)^k$ *for parameters in Table 7.1. Expected value refers to the average execution price per share. The initial share price is* 100*. Subplot (a) shows* $\mathcal{S}((\mathcal{Y}_Q^\dagger)^k)$ *for two different refinement levels. Subplot (b) shows the same Pareto points plotted as expected value versus standard deviation. Note that the efficient frontier for refinement level zero visually coincides that for refinement level one. Further refinement steps show a negligible change. This suggests convergence of the numerical solution and the frontier.*

In this particular example, Figure 7.2 shows that $\mathcal{S}\big((\mathcal{Y}_Q^\dagger)^k\big) = (\mathcal{Y}_Q^\dagger)^k$, which provides strong evidence that $\mathcal{S}\big(\mathcal{Y}_Q^\dagger\big) = \mathcal{Y}_Q^\dagger$. In addition, $(\mathcal{Y}_Q^\dagger)^k$ also suggests that $\mathcal{Y}_Q^\dagger$ is a continuous monotone increasing curve. This indicates that a simple uniform sampling of γ will produce a convergent method (see the discussion in Section 7.2) for this particular example.

8. Conclusion. Many problems in finance can be reduced to a multi-period MV optimization. The standard scalarization optimal method for the multi-objective optimization yields a subset of MV Pareto optimal points. Using the embedding technique of [18, 25], an embedded MV set is determined. This embedded problem can be solved using dynamic programming. In the context of the optimal trade execution, the optimal strategy is determined by solving an HJB equation.

However, when using a numerical method to solve the HJB equation, several issues arise. This technique generates embedded MV points $(\mathcal{V}, \mathcal{E})$ indirectly and these can be a superset of the MV Pareto points. In addition, there may be more than one optimal strategy, given by the solution of the HJB equation, which generates the same value of the objective function. In practice, any numerical algorithm used to solve the embedded problem will generate only one such strategy. This raises the question of whether this strategy corresponds to a MV Pareto optimal point. In addition, it is important to determine which embedded MV points are MV scalarization optimal for the achievable MV objective value set $\mathcal{Y}$.

In this paper, we establish that, if an embedded objective point $(\mathcal{V}, \mathcal{E})$ is MV scalarization optimal with respect to the embedded MV objective set, it is scalarization optimal with respect to the achievable MV objective set (thus MV Pareto optimal). In addition, we prove that the set of the MV scalarization optimal points with respect to the computed embedded objective set $\mathcal{Y}_Q^\dagger$ is identical to the scalarization optimal set with respect to the achievable MV objective set. These two results allow us to develop a simple post processing technique which can be used to eliminate spurious points in the (computed) embedded objective set.

In practical application, we can only obtain an approximation to the solution of the embedded problem. In particular, we can only compute a finite set of optimal points for the embedded problem. Assuming that this finite set approximates the complete solution set sufficiently well, we can apply our post-processing algorithm to obtain the Pareto points of the original MV problem.

It remains to determine the characteristics of the MV problem with which finite sampling of the solution of the embedded problem can be shown to approximate (arbitrarily well) the complete solution of the embedded problem. We leave this to future work.

REFERENCES

[1] G. Aivaliotis and J. Palczewski. Investment strategies and compensation of a mean-variance optimizing fund manager. Working paper, School of Mathemaics, University of Leeds, 2012.

[2] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM Journal on Financial Mathematics*, 3:511–533, 2012.

[3] R. Almgren. Optimal Trading with Stochastic Liquidity and Volatility. *SIAM Journal on Financial Mathematics*, 3:163–181, 2012.

[4] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[5] R. Almgren, C. Thum, E. Hauptmann, and H. Li. Direct estimation of equity market impact. *Risk*, 18:57–62, 2005.

[6] A. M. Andrew. Another Efficient Algorithm for Convex Hulls in Two Dimensions. *Information Processing Letters*, 9:216–219, 1979.

[7] S. Basak and G. Chabakauri. Dynamic mean-variance asset allocation. *Review of Financial Studies*, 23:2970–3016, 2011.

[8] Tomas Bjork and Agatha Murgoci. A General Theory of Markovian Time Inconsistent Stochastic Control Problems. *SSRN eLibrary*, 2010.

[9] S.P. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge Univ Press, 2004.

[10] CFTC-SEC. Findings regarding the market events of May 6, 2010. Report of the staffs of the CFTC and the SEC to the Joint Advisory Committee on Emerging Regulatory Issues, 2010.

[11] R.F. Engle and R. Ferstenberg. Execution Risk. *The Journal of Trading*, 2(2):10–20, 2007.

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

[13] P.A. Forsyth, J.S. Kennedy, S.T. Tse, and H. Windcliff. Optimal trade execution: a mean quadratic variation approach. *Journal of Economic Dynamics and Control*, 36:1971–1991, 2012.

[14] J. Gatheral and A. Schied. Optimal Trade Execution under Geometric Brownian Motion in

the Almgren and Chriss Framework. *International Journal of Theoretical and Applied Finance*, 14:353–368, 2011.

[15] J. Gatheral and A. Schied. Dynamical models for market impact and algorithms for optimal order execution. In J.-P. Fouque and J. Langsam, editors, *Handbook on Systematic Risk*. Cambridge University Press, Cambridge, 2013. to appear.

[16] F. Guibaud, M. Mnif, and H. Pham. Numerical methods for an optimal order execution problem. *Journal of Computational Finance*, 2013. forthcoming.

[17] G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275, 2004.

[18] D. Li and W.L. Ng. Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation. *Mathematical Finance*, 10(3):387–406, 2000.

[19] F. Lillo, J.D. Farmer, and R.N. Mantegna. Econophysics: Master curve for price-impact function. *Nature*, 421(6919):129–130, 2003.

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

[21] M. Potters and J.P. Bouchaud. More statistical properties of order books and price impact. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):133–140, 2003.

[22] S.T. Tse. *Numerical Methods for Optimal Trade Execution.* PhD thesis, Cheriton School of Computer Science, University of Waterloo, 2012.

[23] S.T. Tse, P.A. Forsyth, J. S. Kennedy, and H. Windcliff. Comparison between the mean variance optimal and the mean quadratic variation optimal trading strategies. *Applied Mathematical Finance*, 2013. Forthcoming.

[24] H. Windcliff. Private Communication, 2012.

[25] X.Y. Zhou and D. Li. Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42(1):19–33, 2000.