



## THE SELF-COORDINATION MEAN-VARIANCE STRATEGY IN CONTINUOUS TIME

YUN SHI<sup>1</sup>, DUAN LI<sup>2</sup> AND XIANGYU CUI<sup>3,\*</sup>

**Abstract.** The dynamic mean-variance portfolio selection problem is time inconsistent. In the literature, scholars try to derive the pre-committed strategy, the time consistent strategy and the self-coordination strategy. The pre-committed strategy only concern the global investment interest of the investor. The time consistent strategy only concerns the local investment interests of the investor, while the self-coordination strategy balances between the global investment interest and local investment interests of the investor. However, the self-coordination strategy is only studied for the discrete time mean-variance setting. We study the self-coordination strategy for the continuous time mean-variance setting in this paper. With the help of mean-field reformulation, we derive the analytical self-coordination mean-variance strategy and show that the pre-committed strategy and time consistent strategy are special cases of the self-coordination strategy.

**Mathematics Subject Classification.** 49, 91.

Received April 20, 2022. Accepted October 11, 2023.

### 1. INTRODUCTION

In a dynamic mean-variance framework, the investor decides the global optimal strategy according to the following problem,  $(MV_0)$ ,

$$\max_{\{u_t\}_{t=0}^T} \mathbb{E}_{0,x_0}[X^u(T)] - \lambda \text{Var}_{0,x_0}(X^u(T)),$$

where  $\lambda > 0$  is the risk aversion parameter of the investor. The global optimal strategy of  $(MV_0)$ , which is called *pre-committed strategy* in this paper<sup>1</sup>, can guide the investor to the best performance at time  $T$  based on the information and wealth at time 0. When coming to a future time instant  $t(> 0)$ , the investor may reconsider the portfolio selection problem based on the information and wealth at time  $t$ , and would like to choose the

---

*Keywords.* Mean-variance model, time inconsistency, self-coordination strategy, continuous time setting.

<sup>1</sup> School of Statistics and Academy of Statistics and Interdisciplinary Sciences, East China Normal University, Shanghai, P.R. China.

<sup>2</sup> School of Data Science, City University of Hong Kong, Hong Kong, P.R. China.

<sup>3</sup> School of Statistics and Management, Dishui Lake Advanced Finance Institute, Shanghai University of Finance and Economics, Shanghai, P.R. China.

\*Corresponding author: [cui.xiangyu@mail.shufe.edu.cn](mailto:cui.xiangyu@mail.shufe.edu.cn)

<sup>1</sup>It is termed “strategy of pre commitment” in [14], and “pre-committed strategy” in [2, 6].

optimal strategy of the following problem,  $(MV_t)$ ,

$$\max_{\{u_s\}_{s=t}^T} \mathbb{E}_{t,x_t}[X^u(T)] - \lambda \text{Var}_{t,x_t}(X^u(T)).$$

However, as  $\text{Var}_{t,x_t}(X^u(T)) \neq \text{Var}_{t,x_t}(\text{Var}_{s,x_s}(X^u(T)))$  for  $s > t$ , the objectives for different time horizons do not satisfy the *Bellman's principle of optimality* of dynamic programming, and the optimal strategy of  $(MV_t)$  is in general different from the global optimal strategy of  $(MV_0)$ . Such conflict among the investment strategies optimally derived for different time horizons is termed *time inconsistency* in the literature, which represents the conflict among the global investment interest and the local investment interests of the investor (see [2, 6])<sup>2</sup>. Besides the mean-variance model, almost all widely adopted risk measures involving non-expectation operators in the literature, including VaR (Value-at-Risk) and CVaR (Conditional Value-at-risk), are not time consistent (see [1, 5]).

To overcome the time inconsistency issue of the dynamic mean-variance problem, the conflict among the global investment interest and the local investment interests of the investor should be resolved. In the literature, the dynamic mean-variance problem is often reformulated as a sequential game among the investor at different time instants, in which the investor at time  $t$  optimally chooses the equilibrium strategy on the premise that the investors in future time instants have already decided their equilibrium strategies in a similar fashion. The subgame perfect Nash equilibrium strategy is called *time consistent strategy* (see [2, 4, 14, 16]). Although the time consistent strategy achieves equilibrium among the local investment interests, it omits the global investment interest and could lead to poor global investment performance (see [9]).

Different from the sequential game reformulation, [8] proposed a two-tier planner-doer game framework for the discrete time mean-variance setting<sup>3</sup>. Under the framework, the planner, which represents the global interest of the investor, can influence the strategies of the doers, which represents the local interests of the investor at different time instants, *via* a punishment scheme. Thus, there is a sequential game among the doers and a leader-follower game between the planner and the doers. The equilibrium strategy of the two-tier planner-doer game, called *self-coordination strategy*, can achieve the balance between the global interest of the planner and the local interests of the doers.

In this paper, we extend the two-tier planner-doer game framework of [8] to the continuous time setting, and successfully apply the framework to continuous time mean-variance portfolio selection problem. After proposing a particular punishment scheme, which satisfies the assumptions of the game framework, we obtain the portfolio selection problem with self-coordination punishment. Although the new problem is quite difficult, we change it into its mean-field reformulation and derive the explicit self-coordination strategy, which is a closed-loop strategy. Our work pushes forward the state-of-the-art in resolving the long-standing issue of time inconsistency in continuous time mean-variance setting.

The remaining of this paper is organized as follows. In Section 2, we review the pre-committed strategy and time consistent strategy. In Section 3, we extend the two-tier planner-doer game framework to continuous time setting, propose a suitable punishment scheme and derive the explicit self-coordination strategy. Finally, we conclude our paper in Section 4.

## 2. PRE-COMMITTED STRATEGY AND TIME CONSISTENT STRATEGY

We consider a typical Black-Scholes market setting, in which two assets are traded continuously. One of the assets is a risk-free bank account whose value process  $S_0(t)$  is subject to the following ordinary differential equation (ODE):

$$\begin{cases} dS_0(t) = rS_0(t)dt, & t \geq 0, \\ S_0(0) = s_0 > 0, \end{cases}$$

<sup>2</sup>More real life and economical time inconsistent problems can be found in [12, 13, 15].

<sup>3</sup>The two-tier planner-doer game framework is the extension of [10, 11]'s dual self model, which requires the preferences at different time instants be myopic or involve only the expectation operator in the objective functions.

where  $r > 0$  is the interest rate. The other asset is a risky stock, whose price process  $S_1(t)$  satisfies the following stochastic differential equation (SDE):

$$\begin{cases} dS_1(t) = S_1(t)[\mu dt + \sigma dW(t)], & t \geq 0, \\ S_1(0) = s_1 > 0, \end{cases}$$

where  $\mu$  is the appreciation rate,  $\sigma$  is the volatility or dispersion rate of the stock and  $W(t)$  is standard Brownian motion defined on  $(\Omega, \mathcal{F}_T, \{\mathcal{F}_t\}, P)$ . We assume in this paper that all the given market parameters  $r, \mu$  and  $\sigma$  are deterministic constants.

We now consider an investor of a mean-variance type with an initial endowment  $x_0 > 0$  and an investment horizon  $[0, T]$ . Assume that the trading is self-financed and takes place continuously, and that there is no transaction cost and consumptions during the investment process. The wealth process,  $X^u(t)$ , then satisfies

$$\begin{cases} dX^u(t) = [rX^u(t) + (\mu - r)u(t)]dt + \sigma u(t)dW(t), & 0 \leq t \leq T, \\ X^u(0) = x_0, \end{cases} \tag{1}$$

where portfolio  $u(t)$  denotes the total market value of wealth in the stock. We define the infinitesimal generator  $\mathbf{A}^u$  now as follows,

$$\mathbf{A}^u V(t, x) = V_t(t, x) + [rx + (\mu - r)u(t)]V_x(t, x) + \frac{1}{2}\sigma^2 u^2(t)V_{xx}(t, x),$$

where  $V_t(t, x)$ ,  $V_x(t, x)$  and  $V_{xx}(t, x)$  denote the partial derivatives of  $V(t, x)$  with respect to  $t$  and  $x$ , and the second order partial derivative of  $V$  with respect to  $x$ , respectively.

At time 0, the investor is seeking the pre-committed optimal strategy such that the expected value of the terminal wealth is maximized and the variance of the terminal wealth is minimized,

$$\begin{aligned} (MV_0(\lambda)) \quad & \max \mathbb{E}_{0,x_0}[X^u(T)] - \lambda \text{Var}_{0,x_0}(X^u(T)), \\ & \text{s.t. } (X^u(\cdot), u(\cdot)) \text{ satisfies (1),} \end{aligned}$$

where  $u(\cdot)$  is the decision variable,  $\lambda > 0$  is risk aversion parameter, which represents the trade-off between the two conflicting objectives (the larger the  $\lambda$ , the larger the risk aversion). Varying parameter  $\lambda$  in  $(0, +\infty)$  in problem formulation  $(MV_0(\lambda))$  traces out the efficient frontier in the mean-variance space.

The pre-committed strategy of  $(MV_0(\lambda))$  can be found in [17],

$$u^{\text{Com}}(t, x) = -\frac{\mu - r}{\sigma^2} \left[ x - \left( e^{rT} x_0 + \frac{e^{\delta T}}{2\lambda} \right) e^{-r(T-t)} \right], \tag{2}$$

where  $\delta = \frac{(\mu - r)^2}{\sigma^2}$ . Moreover, the corresponding efficient frontier is expressed as

$$\begin{aligned} \text{Var}_{0,x_0}(X^u(T))|_{u^{\text{Com}}} &= \frac{e^{-\delta T}}{1 - e^{-\delta T}} (\mathbb{E}_{0,x_0}[X^u(T)]|_{u^{\text{Com}}} - e^{rT} x_0)^2, \\ &\text{for } \mathbb{E}_{0,x_0}[X^u(T)]|_{u^{\text{Com}}} \geq e^{rT} x_0. \end{aligned}$$

Next, we provide the definition of time consistent strategy and derive the time consistent strategy for  $MV_0(\lambda)$ .

**Definition 2.1** ([4]). Consider a portfolio strategy  $u^{TC}$ . Choose a fixed  $\bar{u}$ , a fixed real number  $h > 0$ . Also fix an arbitrarily chosen initial point  $(t, x)$ . Define the strategy  $u_h$  by

$$u_h(s, y) = \begin{cases} \bar{u}, & \text{for } t \leq s < t + h, \\ u^{TC}(s, y), & \text{for } t + h \leq s \leq T, \end{cases}$$

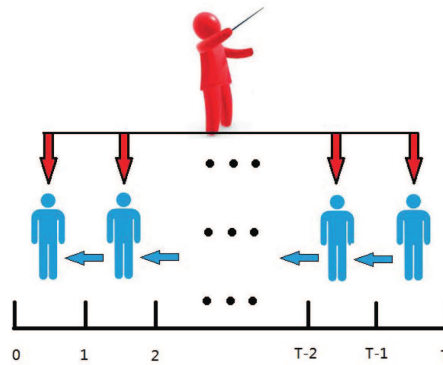


FIGURE 1. Diagram of two-tier planner-doer game framework.

and denote the objectives at time  $t$  of two strategies as  $J(t, x, u^{TC})$  and  $J(t, x, u_h)$ . If

$$\liminf_{h \rightarrow 0} \frac{J(t, x, u^{TC}) - J(t, x, u_h)}{h} \geq 0$$

for all  $\bar{u}$ , we say that  $u^{TC}$  is a subgame perfect Nash equilibrium strategy or time consistent strategy.

Following [2] and [4], we can derive the time consistent strategy as

$$u^{TC}(t, x) = \frac{1}{2\lambda} e^{-r(T-t)} \frac{\mu - r}{\sigma^2}, \tag{3}$$

with the corresponding efficient frontier expressed as

$$\begin{aligned} \text{Var}_{0, x_0}(X^u(T))|_{u^{TC}} &= \frac{1}{\delta T} (\mathbb{E}_{0, x_0}[X^u(T)]|_{u^{TC}} - e^{rT} x_0)^2, \\ &\text{for } \mathbb{E}_{0, x_0}[X^u(T)]|_{u^{TC}} \geq e^{rT} x_0. \end{aligned}$$

### 3. SELF-COORDINATION MEAN-VARIANCE STRATEGY

Before we derive the self-coordination strategy for the continuous time mean-variance setting, we first extend the two-tier planner-doer game framework of [8] to continuous time setting. The two-tier planner-doer game framework in discrete time can be illustrated by Figure 1.

The red person above is the planner, who can induce the blue doers to make decisions by announcing a planned strategy and committing that the deviation from the planned strategy would cause a punishment. And the blue doers in the below would make decisions by considering both the planner’s announced strategy and the future doers’ strategies. Thus, there are a sequential game among the doers and a principle-agent game between the planner and the doers.

Consider a controlled Markovian process  $X^u$  in continuous time within interval  $[0, T]$ , which is defined on a filtered probability space  $(\Omega, \mathcal{F}_T, \{\mathcal{F}_t\}, P)$ . At time  $t$ , after knowing  $X_t^u = x$ , the decision maker wants to decide the  $[t, T]$  optimal strategy by maximizing the following preference over time horizon  $[t, T]$ ,

$$J(t, x, \{u(s, X_s^u)\}_{s \geq t}^T),$$

where  $u(s, X_s^u)$  is the strategy at time  $s$ ,  $J$  is a deterministic function that describes the objective of the decision maker at time  $t$ . In general, these preferences at different time instants do not satisfy Bellman’s principle of optimality, which makes the decision problem under consideration time inconsistent.

To balance the global and local interests of the decision maker, a punishment scheme is proposed. More specifically, the planner, who represents the global interest of the decision maker, decides a planned strategy denoted by  $\{\nu(t, X_t^u)\}_{t \geq 0}^T$ , announces it to the doers, and commits that any derivation by the doer  $t$  from the planned control strategy leads a punishment denoted by

$$\mathbb{E}_{t,x} \left[ \int_t^T g^u(s)(u(s, X_s^u) - \nu(s, X_s^u))^2 ds \right],$$

where  $g^u(s) > 0$  is the weighting function of the punishment, which depends on the applied strategy after time  $s$  but does not depend the strategy at time  $s$ . Apparently, the punishment is always non-negative. And when the strategies  $u$  and  $\nu$  are identical, the punishment is zero, which implies that when the doer follows the planned strategy there is no punishment. Thus, the doer  $t$  faces the following optimization problem,  $(SC^{\text{doer}}(\nu))$ ,

$$\max_{\{u(s, X_s^u)\}_{s \geq t}^T} J(t, x, \{u(s, X_s^u)\}_{s \geq t}^T) - \mathbb{E}_{t,x} \left[ \int_t^T g^u(s)(u(s, X_s^u) - \nu(s, X_s^u))^2 ds \right].$$

As the optimization problems of the doers with punishment are still time inconsistent, the doers tries to find the subgame perfect Nash equilibrium strategy in Definition 2.1, which is denoted as  $\{u^\nu(s, X_s^u)\}_{s=0}^T$ .

At the same time, the total amount of punishment applied by the planner will in turn penalize the planner himself. Thus, the planner faces the following optimization problem,  $(SC^{\text{planner}}(\omega))$

$$\max_{\{\nu(s, X_s^u)\}_{s \geq 0}^T} J(0, x_0, \{u^\nu(s, X_s^u)\}_{s \geq 0}^T) - \omega \mathbb{E}_{0,x_0} \left[ \int_0^T g^u(s)(u^\nu(s, X_s^u) - \nu(s, X_s^u))^2 ds \right],$$

where  $\omega$  is the trade-off parameter between the original global objective and the total expected punishment, which represents the self-coordination willpower of the decision maker. The larger the  $\omega$ , the stronger the planner’s dislike of the difference between the doer’s subgame perfect Nash equilibrium strategy strategy from the planned strategy.

In general, this constructed planner-doe game can be solved by the following procedure: 1. obtain the subgame perfect Nash equilibrium strategy of  $(SC^{\text{doer}}(\nu))$ ,  $\{u^\nu(t, X_t^u)\}_{t \geq 0}^T$ , for any given planned strategy  $\{\nu(t, X_t^u)\}_{t \geq 0}^T$ ; 2. with knowing the doers’ responsive subgame perfect Nash equilibrium strategy, solve  $(SC^{\text{planner}}(\omega))$  to obtain the best planned strategy  $\{\nu^*(t, X_t^u, \omega)\}_{t \geq 0}^T$ ; 3. substitute the best planned strategy  $\{\nu^*(t, X_t^u, \omega)\}_{t \geq 0}^T$  into  $\{u^\nu(t, X_t^u)\}_{t \geq 0}^T$ , and obtain the applied strategy of the doers,  $\{u^{\nu^*}(t, X_t^u, \omega)\}_{t \geq 0}^T$ . Under equilibrium, the applied strategy of the doers is called *self-coordination strategy*, which is denoted  $\{u^{\text{Coor}}(t, X_t^u, \omega)\}_{t \geq 0}^T$ .

In the following theorem, we obtain some useful results for the self-coordination strategy without solving the planner-doe game directly. We give an important assumption of the punishment scheme.

**Assumption 3.1.** *For any planned strategy  $\{\nu(t, X_t^u)\}_{t \geq 0}^T$ , there exists a unique subgame perfect Nash equilibrium strategy of  $(SC^{\text{doer}}(\nu))$ ,  $\{u^\nu(t, X_t^u)\}_{t \geq 0}^T$ . And for any admissible strategy  $\{u(t, X_t^u)\}_{t \geq 0}^T$ , there exists a unique planned strategy  $\{\nu^u(t, X_t^u)\}_{t \geq 0}^T$  such that  $u^{\nu^u}(t, X_t^u) = u(t, X_t^u)$ .*

**Theorem 3.2.** *Under Assumption 3.1, the self-coordination strategy, which is denoted as  $\{u^{\text{Coor}}(t, X_t^u, \omega)\}_{t \geq 0}^T$ , is the pre-committed optimal solution of the following optimization problem,  $(SC(\omega))$ ,*

$$\max_{\{u(t, X_t^u)\}_{t \geq 0}^T} J(0, x, \{u(t, X_t^u)\}_{t \geq 0}^T) - \omega \mathbb{E}_{0,x} \left[ \int_0^T g^u(s)(u(s, X_s^u) - \nu^u(s, X_s^u))^2 ds \right].$$

*Proof.* For a given planned strategy  $\{\nu(t, X_t^u)\}_{t \geq 0}^T$ , the doers adopt the subgame perfect Nash equilibrium strategy  $\{u^\nu(t, X_t^u)\}_{t \geq 0}^T$  by solving the doer’s problem. Then, the planner faces the following problem,

$$\max_{\{\nu(s, X_s^u)\}_{s \geq 0}^T} J(0, x_0, \{u^\nu(s, X_s^u)\}_{s \geq 0}^T) - \omega \mathbb{E}_{0, x_0} \left[ \int_0^T g^u(s) (u^\nu(s, X_s^u) - \nu(s, X_s^u))^2 ds \right].$$

Under Assumption 3.1,  $\{u(t, X_t^u)\}_{t \geq 0}^T$  and  $\{\nu(t, X_t^u)\}_{t \geq 0}^T$  have a one-to-one mapping. Therefore, by changing the decision variable  $\{\nu(t, X_t^u)\}_{t \geq 0}^T$  in the planner’s problem ( $SC^{\text{planner}}(\omega)$ ) into  $\{u(t, X_t^u)\}_{t \geq 0}^T$ , we can obtain the problem ( $SC(\omega)$ ), which determines the self-coordination strategy.  $\square$

Next, we derive the self-coordination strategy for continuous time mean-variance setting. As shown in Section 2, both the pre-committed strategy and the time consistent strategy take linear forms. Thus, we confine ourselves to consider only linear self-coordination strategy.

Denote the doers’ carried out strategy as  $u(t, x) = k^u(t)x + l^u(t)$  and the planner’s planned strategy as  $\nu(t, x) = k^\nu(t)x + l^\nu(t)$ . We propose the following particular form of punishment,

$$\mathbb{E}_{t,x} \left[ \int_t^T \lambda \sigma^2 ([a^u(s)]^2 - A^u(s)) [k^u(s)X^u(s) + l^u(s) - k^\nu(s)X^u(s) - l^\nu(s)]^2 ds \right], \tag{4}$$

where  $a^u(t)$  and  $A^u(t)$  satisfy the following ODEs,

$$\begin{cases} a_t^u(t) + ra^u(t) + (\mu - r)a^u(t)k^u(t) = 0, \\ a^u(T) = 1, \\ A_t^u(t) + 2A^u(t)[r + (\mu - r)k^u(t)] - \sigma^2([a^u(t)]^2 - A^u(t))[k^u(t)]^2 = 0, \\ A^u(T) = 0, \end{cases}$$

where  $a_t^u(t)$  and  $A_t^u(t)$  are the first order derivatives of  $a^u(t)$  and  $A^u(t)$  with respect to  $t$ . It is easy to show that

$$\begin{cases} ([a^u(t)]^2 - A^u(t))_t = -([a^u(t)]^2 - A^u(t))\{2[r + (\mu - r)k^u(t)] + \sigma^2(k^u(t))^2\}, \\ ([a^u(T)]^2 - A^u(T)) = 1, \end{cases}$$

which implies that  $[a^u(t)]^2 - A^u(t) = e^{-\int_t^T 2\{[r+(\mu-r)k^u(s)]+\sigma^2(k^u(s))^2\}ds} > 0$  holds for all  $t$ , and the punishment is non-negative. Furthermore, based on the ODEs, we find that  $[a^u(t)]^2 - A^u(t)$  does not depend on the value of  $k^u(t)$  but depend on the values of  $k^u(s)$ , ( $s > t$ ). Under the assumption of linear strategy, the proposed punishment introduces a running cost into the doers’ and planner’s problems. The running cost takes a quadratic form of the wealth state, which makes it possible to obtain the explicit self-coordination strategy. If both the planner’s planned strategy and the doers’ carried out strategy are general functions of wealth, explicit self-coordination strategy may not be obtained.

The following proposition reveals that the proposed punishment scheme satisfies Assumption 3.1.

**Proposition 3.3.** *Under the proposed form of punishment in (4), the subgame perfect Nash equilibrium strategy of the doers  $\{u^\nu(t, X_t^u)\}_{t \geq 0}^T$  are given as*

$$u^\nu(t, X^u(t)) = k^{u^\nu}(t)X^u(t) + l^{u^\nu}(t) = \frac{k^\nu(t) + \hat{k}_t}{2} X^u(t) + \frac{l^\nu(t) + \hat{l}(t)}{2},$$

where

$$\begin{aligned} \hat{k}(t) &= \frac{(\mu - r)A^{u^\nu}(t)}{\sigma^2([a^{u^\nu}(t)]^2 - A^{u^\nu}(t))}, \quad \hat{l}(t) = \frac{(\mu - r)B^{u^\nu}(t)}{2\sigma^2\lambda([a^{u^\nu}(t)]^2 - A^{u^\nu}(t))}, \\ &\begin{cases} B_t^{u^\nu}(t) + B^{u^\nu}(t)[r + (\mu - r)k^{u^\nu}(t)] \\ \quad + 2\lambda l^{u^\nu}(t)\{(\mu - r)A^{u^\nu}(t) - \sigma^2([a^{u^\nu}(t)]^2 - A^{u^\nu}(t))k^{u^\nu}(t)\} = 0, \\ B^{u^\nu}(T) = 1, \end{cases} \end{aligned} \tag{5}$$

$B_t^{u^\nu}(t)$  is the first order derivative of  $B^{u^\nu}(t)$  with respect to  $t$ . Thus, by setting

$$\nu^u(t, x) = [2k^u(t) - \hat{k}(t)]x + 2l^u(t) - \hat{l}(t),$$

we have  $u^{\nu^u}(t, X^u(t)) = u(t, X^u(t))$ , i.e., Assumption 3.1 is satisfied.

*Proof.* Under the proposed form of punishment in (4), the doers faces the following optimization problem,

$$\begin{aligned} &\max_{\{u(s, X^u(s))\}_{s \geq t}^T} \mathbb{E}_{t,x}[X^u(T)] - \lambda \text{Var}_{t,x}(X^u(T)) \\ &\quad - \lambda \sigma^2 \mathbb{E}_{t,x} \left[ \int_t^T ([a^u(s)]^2 - A^u(s))(k^u(s)X^u(s) + l^u(s) - k^\nu(s)X^u(s) - l^\nu(s))^2 ds \right], \end{aligned}$$

where the wealth dynamics is characterized by

$$\begin{cases} dX^u(t) = [rX^u(t) + (\mu - r)(k^u(t)X^u(t) + l^u(t))]dt \\ \quad + \sigma[k^u(t)X^u(t) + l^u(t)]dW(t), \quad 0 \leq t \leq T, \\ X^u(0) = x_0. \end{cases}$$

Next, we derive the subgame perfect Nash equilibrium strategy  $\{k^{u^\nu}(t)X^u(t) + l^{u^\nu}(t)\}_{s \geq 0}^T$ . Let us define

$$\begin{aligned} g^{u^\nu}(t, x) &= \mathbb{E}_{t,x}[X^u(T)] = \mathbb{E}_{t,x}[g^{u^\nu}(s, X^u(s))], \quad s > t, \\ V^{u^\nu}(t, x) &= \mathbb{E}_{t,x}[X^u(T)] - \lambda \text{Var}_{t,x}(X^u(T)) \\ &= \mathbb{E}_{t,x}[\mathbb{E}_{s, X_s^u}[X^u(T)]] - \lambda \mathbb{E}_{t,x}[\text{Var}_{s, X_s^u}(X^u(T))] - \lambda \text{Var}_{t,x}(\mathbb{E}_{s, X_s^u}[X^u(T)]) \\ &= \mathbb{E}_{t,x}[V^{u^\nu}(s, X^u(s))] - \lambda \text{Var}_{t,x}(g(s, X^u(s))), \quad s > t, \end{aligned}$$

with terminal condition  $g^{u^\nu}(T, x) = x$  and  $V^{u^\nu}(T, x) = x$ .

Applying the analysis method of [3], the extended HJB equation is given as follows,

$$\begin{aligned} V_t^{u^\nu} + \sup_{k^u(t), l^u(t)} &\left\{ [rx + (\mu - r)(k^u(t)x + l^u(t))]V_x^{u^\nu} + \frac{1}{2}\sigma^2(k^u(t)x + l^u(t))^2 V_{xx}^{u^\nu} \right. \\ &\quad \left. - \lambda \sigma^2(k^u(t)x + l^u(t))^2 g_x^2 - \lambda \sigma^2([a^{u^\nu}(t)]^2 - A^{u^\nu}(t)) \right. \\ &\quad \left. \cdot (k^u(t)x + l^u(t) - k^\nu(t)x - l^\nu(t))^2 \right\} = 0, \\ &V^{u^\nu}(T, x) = x, \\ g_t^{u^\nu} + [rx + (\mu - r)(k^{u^\nu}(t)x + l^{u^\nu}(t))]g_x^{u^\nu} &+ \frac{1}{2}\sigma^2(k^{u^\nu}(t)x + l^{u^\nu}(t))^2 g_{xx}^{u^\nu} = 0, \\ &g^{u^\nu}(T, x) = x, \end{aligned}$$

where  $V_t^{u^\nu}$ ,  $V_x^{u^\nu}$ ,  $V_{xx}^{u^\nu}$ ,  $g_t^{u^\nu}$ ,  $g_x^{u^\nu}$ ,  $g_{xx}^{u^\nu}$  denotes the derivatives. As  $[a^{u^\nu}(t)]^2 - A^{u^\nu}(t) = e^{-\int_t^T 2\{[r+(\mu-r)k^{u^\nu}(s)]+\sigma^2(k^{u^\nu}(s))^2\}ds}$  does not depend on the concrete value of  $k^u(t)$ ,  $[a^{u^\nu}(t)]^2 - A^{u^\nu}(t)$  is exogenous when solving the extended HJB equation. It is natural to make the *Ansatz*,

$$\begin{aligned} g^{u^\nu}(t, x) &= a^{u^\nu}(t)x + b^{u^\nu}(t), \\ V^{u^\nu}(t, x) &= \lambda A^{u^\nu}(t)x^2 + B^{u^\nu}(t)x + D^{u^\nu}(t). \end{aligned}$$

Then, we have the subgame perfect Nash equilibrium strategy as follows,

$$k^{u^\nu}(t) = \frac{k^\nu(t) + \hat{k}_t}{2}, \quad l^{u^\nu}(t) = \frac{l^\nu(t) + \hat{l}(t)}{2}, \tag{6}$$

and the parameters satisfy the following ODEs,

$$\begin{cases} a_t^{u^\nu}(t) + a^{u^\nu}(t)[r + (\mu - r)k^{u^\nu}(t)] = 0, \\ a^{u^\nu}(T) = 1, \\ b_t^{u^\nu}(t) + (\mu - r)a^{u^\nu}(t)l^{u^\nu}(t) = 0, \\ b^{u^\nu}(T) = 0, \\ A_t^{u^\nu}(t) + 2A^{u^\nu}(t)[r + (\mu - r)k^{u^\nu}(t)] - \sigma^2([a^{u^\nu}(t)]^2 - A^{u^\nu}(t))(k^{u^\nu}(t))^2 = 0, \\ A^{u^\nu}(T) = 0, \\ B_t^{u^\nu}(t) + B^{u^\nu}(t)[r + (\mu - r)k^{u^\nu}(t)] \\ \quad + 2\lambda l^{u^\nu}(t)\{(\mu - r)A^{u^\nu}(t) - \sigma^2([a^{u^\nu}(t)]^2 - A^{u^\nu}(t))k^{u^\nu}(t)\} = 0, \\ B^{u^\nu}(T) = 1, \\ D_t^{u^\nu}(t) + B^{u^\nu}(t)(\mu - r)l^{u^\nu}(t) - \lambda\sigma^2([a^{u^\nu}(t)]^2 - A^{u^\nu}(t))(l^{u^\nu}(t))^2 = 0, \\ D^{u^\nu}(T) = 0, \end{cases}$$

and  $a_t^{u^\nu}(t)$ ,  $b_t^{u^\nu}(t)$ ,  $A_t^{u^\nu}(t)$ ,  $B_t^{u^\nu}(t)$  and  $D_t^{u^\nu}(t)$  are the first order derivatives.

As equation (6) has constructed the one-to-one mapping between  $u^\nu(t, X_t^u)$  and  $\nu(t, X_t^u)$ , we can simply choose

$$\nu^u(t, x) = [2k^u(t) - \hat{k}(t)]x + 2l^u(t) - \hat{l}(t),$$

which implies  $u^{\nu^u}(t, X^u(t)) = u(t, X^u(t))$ . □

**Remark 3.4.** Proposition 3.3 shows that the subgame perfect Nash equilibrium strategy of the doers  $u^\nu(t, X^u(t))$  depends on the coefficients  $a^{u^\nu}(t)$  and  $A^{u^\nu}(t)$ , which further depends on  $\{u^\nu(s, X^u(s))\}_{s \geq t}^T$ . Therefore, the subgame perfect Nash equilibrium strategy of the doers  $u^\nu(t, X^u(t))$  is not fully explicit. The system should be solved numerically by a backward scheme. When the future strategy after time  $t$ ,  $\{u^\nu(s, X^u(s))\}_{s \geq t + \Delta t}^T$ , are obtained, we can compute  $a^{u^\nu}(t)$  and  $A^{u^\nu}(t)$  by the ODEs, and then determine the current strategy  $u^\nu(t, X^u(t))$ .

Substituting the planned strategy  $\nu^u(t, x) = [2k^u(t) - \hat{k}(t)]x + 2l^u(t) - \hat{l}(t)$  back into the punishment in (4), the punishment becomes

$$\mathbb{E}_{t,x} \left[ \int_t^T \lambda \sigma^2([a^u(s)]^2 - A^u(s))[k^u(s)X^u(s) + l^u(s) - \hat{k}(s)X^u(s) - \hat{l}(s)]^2 ds \right]. \tag{7}$$



This term measures the punishment happened under punishment scheme (4), which can let the doers adopt  $\{k^u(s)X^u(s) + l^u(s)\}_{s=0}^T$ . Based on Theorem 3.2, the self-coordination mean-variance strategy can be obtained by solving the following problem parameterized in  $\omega$ ,  $(MV(\lambda, \omega))$ ,

$$\begin{aligned} & \max \mathbb{E}_{0,x_0}[X^u(T)] - \lambda \text{Var}_{0,x_0}(X^u(T)) \\ & \quad - \omega \mathbb{E}_{0,x_0} \left[ \int_0^T \lambda \sigma^2([a^u(t)]^2 - A^u(t)) [k^u(t)X_t^u + l^u(t) - \hat{k}(t)X_t^u - \hat{l}(t)]^2 dt \right], \\ & \text{s.t. } (X^u(\cdot), u(\cdot)) \text{ satisfies (1),} \\ & \quad u(t, x) = k^u(t)x + l^u(t), \end{aligned}$$

where  $a^u(t)$  and  $A^u(t)$  satisfy the following ODEs,

$$\begin{cases} a_t^u(t) + ra^u(t) + (\mu - r)a^u(t)k^u(t) = 0, \\ a^u(T) = 1, \\ A_t^u(t) + 2A^u(t)[r + (\mu - r)k^u(t)] - \sigma^2([a^u(t)]^2 - A^u(t))[k^u(t)]^2 = 0, \\ A^u(T) = 0. \end{cases}$$

Before deriving the pre-committed strategy of  $(MV(\lambda, \omega))$ , we make more detailed discussion for the punishment scheme in (4). Based on the proof of Proposition 3.3, it is easy to check that any punishment

$$\mathbb{E}_{t,x} \left[ \int_t^T g^u(s) [k^u(s)X_s^u + l^u(s) - k^\nu(s)X_s^u - l^\nu(s)]^2 ds \right],$$

with  $g^u(s) > 0$  can construct a one-to-one mapping between  $u^\nu(t, X^u(t))$  and  $\nu(t, X^u(t))$ . Then, why do we choose a particular form  $\lambda \sigma^2([a^u(t)]^2 - A^u(t))$ ? Because this punishment has an important economic meaning.

Let us first introduce the concept of instantaneous utility loss at time  $t$  for a given policy  $\{u(t, X^u(s))\}_{s \geq t}^T$ . For a fixed real number  $h > 0$ , consider the following dynamic mean-variance problem,

$$\max_{\{\pi(s, X^{\pi_h}(s))\}_{s \geq t}^{t+h}} \mathbb{E}_{t,x}[X^{\pi_h}(T)] - \lambda \text{Var}_{t,x}(X^{\pi_h}(T)),$$

where the strategy  $\pi_h$  is defined as

$$\pi_h(s, y) = \begin{cases} \pi(s, y), & \text{for } t \leq s < t + h, \\ u(t, y), & \text{for } t + h \leq s \leq T. \end{cases}$$

We denote the subgame perfect Nash equilibrium strategy of the problem from  $t$  to  $t + h$  as  $\{\hat{\pi}(s, X^{\hat{\pi}_h}(s))\}_{s \geq t}^{t+h}$ , and denote the objectives of the optimization problem under  $\hat{\pi}_h$  and  $u$  as  $J^{\hat{\pi}_h}(t, x, \hat{\pi}_h)$  and  $J^u(t, x, u)$ .

**Definition 3.5.** The instantaneous utility loss for a given strategy  $\{u(s, X^u(s))\}_{s \geq t}^T$  is

$$\lim_{h \rightarrow 0} (J^{\hat{\pi}_h}(t, x, \hat{\pi}_h) - J^u(t, x, u)).$$

The instantaneous utility loss represents the objective reduction when choosing a local optimal strategy in an instantaneous time slot instead of following the specified strategy  $\{u(s, X^u(s))\}_{s \geq t}^T$ , which is quite similar to the concept of opportunity cost.

**Proposition 3.6.** Under the punishment term in (7),

$$\lambda \sigma^2([a^u(t)]^2 - A^u(t)) [k^u(t)X^u(t) + l^u(t) - \hat{k}(t)X^u(t) - \hat{l}(t)]^2,$$

is just the instantaneous utility loss for the strategy  $\{k^u(s)X^u(s) + l^u(s)\}_{s \geq t}^T$ .

*Proof.* For given strategy  $\{k^u(s)X^u(s) + l^u(s)\}_{s \geq t}^T$ , we have

$$\begin{cases} dX^u(s) = [rX^u(s) + (\mu - r)(k^u(s)X^u(s) + l^u(s))]dt \\ \quad + \sigma[k^u(s)X^u(s) + l^u(s)]dW(t), \quad t \leq s \leq T, \\ X^u(t) = x. \end{cases}$$

Let us define

$$V^u(t, x) = \mathbb{E}_{t,x}[V^u(s, X^u(s))] - \lambda \text{Var}_{t,x}(g^u(s, X^u(s))), \quad s > t,$$

with terminal condition  $g^u(T, x) = x$  and  $V^u(T, x) = x$ . Applying Itô Lemma to  $g^u(t, x)$ ,  $V^u(t, x)$ , and letting  $s \rightarrow t$ , we have

$$\begin{aligned} &g_t^u(t, x) + [rx + (\mu - r)(k^u(t)x + l^u(t))]g_x^u(t, x) \\ &\quad + \frac{1}{2}\sigma^2[k^u(t)x + l^u(t)]^2g_{xx}^u(t, x) = 0, \\ &V_t^u(t, x) + [rx + (\mu - r)(k^u(t)x + l^u(t))]V_x^u(t, x) \\ &+ \frac{1}{2}\sigma^2[k^u(t)x + l^u(t)]^2V_{xx}^u(t, x) - \lambda\sigma^2[k^u(t)x + l^u(t)]^2[g_x^u(t, x)]^2 = 0, \end{aligned}$$

where  $g_t^u(t, x)$ ,  $g_x^u(t, x)$ ,  $V_t^u(t, x)$ ,  $V_x^u(t, x)$ ,  $g_{xx}^u(t, x)$ ,  $V_{xx}^u(t, x)$  are the derivatives. It is easy to verify that

$$\begin{aligned} g^u(t, x) &= a^u(t)x + b^u(t), \\ V^u(t, x) &= \lambda A^u(t)x^2 + B^u(t)x + D^u(t), \end{aligned}$$

with the parameters satisfying the following ODEs,

$$\begin{cases} a_t^u(t) + a^u(t)[r + (\mu - r)k^u(t)] = 0, \\ a^u(T) = 1, \\ b_t^u(t) + (\mu - r)a^u(t)l^u(t) = 0, \\ b^u(T) = 0, \\ A_t^u(t) + 2A^u(t)[r + (\mu - r)k^u(t)] - \sigma^2([a^u(t)]^2 - A^u(t))(k^u(t))^2 = 0, \\ A^u(T) = 0, \\ B_t^u(t) + B^u(t)[r + (\mu - r)k^u(t)] \\ \quad + 2\lambda l^u(t)\{(\mu - r)A^u(t) - \sigma^2([a^u(t)]^2 - A^u(t))k^u(t)\} = 0, \\ B^u(T) = 1, \\ D_t^u(t) + B^u(t)(\mu - r)l^u(t) - \lambda\sigma^2([a^u(t)]^2 - A^u(t))(l^u(t))^2 = 0, \\ D^u(T) = 0, \end{cases}$$

and  $a_t^u(t)$ ,  $b_t^u(t)$ ,  $A_t^u(t)$ ,  $B_t^u(t)$  and  $D_t^u(t)$  are the first order derivatives. We have  $g^u(t+h, x) = a^u(t+h)x + b^u(t+h)$ ,  $V^u(t+h, x) = \lambda A^u(t+h)x^2 + B^u(t+h)x + D^u(t+h)$ , and

$$J^u(t, x, u) = \lambda A^u(t)x^2 + B^u(t)x + D^u(t).$$

The optimization problem

$$\max_{\{\pi(s, X^{\pi_h}(s))\}_{s \geq t}^{t+h}} \mathbb{E}_{t,x}[X^{\pi_h}(T)] - \lambda \text{Var}_{t,x}(X^{\pi_h}(T)),$$

reduces into

$$\begin{aligned} & \max_{\{\pi(s, X^{\pi_h}(s))\}_{s \geq t}^{t+h}} \mathbb{E}_{t,x} [\lambda A^u(t+h)(X_{t+h}^\pi)^2 + B^u(t+h)X_{t+h}^\pi + D^u(t+h)] \\ & \quad - \lambda \text{Var}_{t,x}(a^u(t+h)X_{t+h}^\pi + b^u(t+h)) \end{aligned}$$

The subgame perfect Nash equilibrium strategy  $\{\hat{\pi}(s, X^{\hat{\pi}_h}(s))\}_{s \geq t}^{t+h}$  can be found by solving the extended HJB equations

$$\begin{aligned} V_t^{\hat{\pi}_h} + \sup_{\pi} \left\{ [rx + (\mu - r)\pi]V_x^{\hat{\pi}_h} + \frac{1}{2}\sigma^2\pi^2V_{xx}^{\hat{\pi}_h} - \lambda\sigma^2\pi^2(g_x^{\hat{\pi}_h})^2 \right\} &= 0, \\ V^{\hat{\pi}_h}(t+h, x) &= \lambda A^u(t+h)x^2 + B^u(t+h)x + D^u(t+h), \\ g_t^{\hat{\pi}_h} + [rx + (\mu - r)\hat{\pi}]g_x^{\hat{\pi}_h} + \frac{1}{2}\sigma^2\hat{\pi}^2g_{xx}^{\hat{\pi}_h} &= 0, \\ g^{\hat{\pi}_h}(t+h, x) &= a^u(t+h)x + b^u(t+h). \end{aligned}$$

We have

$$\hat{\pi}(s, x) = k^\pi(s)x + l^\pi(s) = \frac{(\mu - r)A^\pi(s)}{\sigma^2([a^\pi(s)]^2 - A^\pi(s))}x + \frac{(\mu - r)B^\pi(s)}{2\sigma^2\lambda([a^\pi(s)]^2 - A^\pi(s))}$$

and

$$J^{\hat{\pi}_h}(t, x, \hat{\pi}) = \lambda A^\pi(t)x^2 + B^\pi(t)x + D^\pi(t),$$

where the parameters satisfy the following ODEs,

$$\begin{cases} a_t^\pi(t) + a^\pi(t)[r + (\mu - r)k^\pi(t)] = 0, \\ a^\pi(t+h) = a^u(t+h), \\ b_t^\pi(t) + (\mu - r)a^\pi(t)l^\pi(t) = 0, \\ b^\pi(t+h) = b^u(t+h), \\ A_t^\pi(t) + 2A^\pi(t)[r + (\mu - r)k^\pi(t)] - \sigma^2([a^\pi(t)]^2 - A^\pi(t))(k^\pi(t))^2 = 0, \\ A^\pi(t+h) = A^u(t+h), \\ \begin{cases} B_t^\pi(t) + B^\pi(t)[r + (\mu - r)k^\pi(t)] \\ + 2\lambda l^\pi(t)\{(\mu - r)A^\pi(t) - \sigma^2([a^\pi(t)]^2 - A^\pi(t))k^\pi(t)\} = 0, \\ B^\pi(t+h) = B^u(t+h), \end{cases} \\ \begin{cases} D_t^\pi(t) + B^\pi(t)(\mu - r)l^\pi(t) - \lambda\sigma^2([a^\pi(t)]^2 - A^\pi(t))(l^\pi(t))^2 = 0, \\ D^\pi(t+h) = D^u(t+h). \end{cases} \end{cases}$$

Then, we have  $\lim_{h \rightarrow 0} A^\pi(t) = A^u(t)$ ,  $\lim_{h \rightarrow 0} a^\pi(t) = a^u(t)$ ,  $\lim_{h \rightarrow 0} b^\pi(t) = b^u(t)$ ,  $\lim_{h \rightarrow 0} B^\pi(t) = B^u(t)$ ,  $\lim_{h \rightarrow 0} D^\pi(t) = D^u(t)$ ,  $\lim_{h \rightarrow 0} k^\pi(t) = \hat{k}(t)$ ,  $\lim_{h \rightarrow 0} l^\pi(t) = \hat{l}(t)$ , which implies that the instantaneous utility

loss is

$$\begin{aligned}
 & \lim_{h \rightarrow 0} (J^{\hat{\pi}_h}(t, x, \hat{\pi}_h) - J^u(t, x, u)) \\
 &= \lambda(2A^u(t)[r + (\mu - r)\hat{k}(t)] - \sigma^2([a^u(t)]^2 - A^u(t))(\hat{k}(t))^2)x^2 \\
 &+ (B^u(t)[r + (\mu - r)\hat{k}(t)] + 2\lambda\hat{l}(t)\{(\mu - r)A^u(t) - \sigma^2([a^u(t)]^2 - A^u(t))\hat{k}(t)\})x \\
 &+ B^u(t)(\mu - r)\hat{l}(t) - \lambda\sigma^2([a^u(t)]^2 - A^u(t))(\hat{l}(t))^2 \\
 &- \lambda(2A^u(t)[r + (\mu - r)k^u(t)] - \sigma^2([a^u(t)]^2 - A^u(t))(k^u(t))^2)x^2 \\
 &- (B^u(t)[r + (\mu - r)k^u(t)] + 2\lambda l^u(t)\{(\mu - r)A^u(t) - \sigma^2([a^u(t)]^2 - A^u(t))k^u(t)\})x \\
 &- B^u(t)(\mu - r)l^u(t) + \lambda\sigma^2([a^u(t)]^2 - A^u(t))(l^u(t))^2 \\
 &= \lambda\sigma^2([a^u(t)]^2 - A^u(t))[2(\hat{k}(t))^2x^2 - (\hat{k}(t))^2x^2 + 2\hat{l}(t)\hat{k}(t)x + (\hat{l}(t))^2 - 2\hat{k}(t)k^u(t)x^2 \\
 &+ (k^u(t))^2x^2 - 2\hat{l}(t)k^u(t)x - 2l^u(t)\hat{k}(t)x + 2k^u(t)l^u(t)x - 2\hat{l}(t)l^u(t) + (l^u(t))^2] \\
 &= \lambda\sigma^2([a^u(t)]^2 - A^u(t))[k^u(t)x + l^u(t) - \hat{k}(t)x - \hat{l}(t)]^2.
 \end{aligned}$$

□

Next, we derive the self-coordination mean-variance strategy by solving  $(MV(\lambda, \omega))$ . Due to the non smoothness of variance term, *i.e.*,  $\text{Var}_t(\cdot) \neq \text{Var}_t(\text{Var}_s(\cdot))$  for  $(t \leq s)$ , we are unable to solve problem  $(MV(\lambda, \omega))$  directly by dynamic programming. We, instead, employ the mean-field reformulation proposed by [7] to derive the explicit strategy of self-coordination.

The key idea of mean-field reformulation is to enlarge the state space  $(X^u(t))$  into  $(\mathbb{E}_{0,x_0}[X^u(t)], X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)])$  and the control space  $(u(t))$ , or equivalently  $(k^u(t)X^u(t) + l^u(t))$ , into

$$(\mathbb{E}_{0,x_0}[u(t)], u(t) - \mathbb{E}_{0,x_0}[u(t)]) = (\mathbb{E}_{0,x_0}[u(t)], k^u(t)(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]).$$

The evolution of the unconditional expectation of the wealth dynamics specified in (1) can be presented as

$$\begin{cases} d\mathbb{E}_{0,x_0}[X^u(t)] = [r\mathbb{E}_{0,x_0}[X^u(t)] + (\mu - r)\mathbb{E}_{0,x_0}[u(t)]]dt, & 0 \leq t \leq T, \\ \mathbb{E}_{0,x_0}[X^u(0)] = x_0. \end{cases} \tag{8}$$

with  $\mathbb{E}_{0,x_0}[u(t)] = k^u(t)\mathbb{E}_{0,x_0}[X^u(t)] + l^u(t)$ . Combining the wealth dynamics in (1) and the expected wealth dynamics in (8) yields the following,

$$\begin{cases} d(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]) = [r(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]) \\ \quad + (\mu - r)(u(t) - \mathbb{E}_{0,x_0}[u(t)])]dt \\ \quad + \sigma(u(t) - \mathbb{E}_{0,x_0}[u(t)] + \mathbb{E}_{0,x_0}[u(t)])dW(t), & 0 \leq t \leq T, \\ X^u(0) - \mathbb{E}_{0,x_0}[X^u(0)] = 0, \end{cases} \tag{9}$$

with  $u(t) - \mathbb{E}_{0,x_0}[u(t)] = k^u(t)(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)])$ . The punishment term  $\lambda\sigma^2([a^u(t)]^2 - A^u(t))[k^u(t)X^u(t) + l^u(t) - \hat{k}(t)X^u(t) - \hat{l}(t)]^2$  can be thus rewritten as

$$\begin{aligned}
 & \lambda\sigma^2([a^u(t)]^2 - A^u(t))\{[k^u(t) - \hat{k}(t)]^2(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)])^2 \\
 &+ 2[k^u(t) - \hat{k}(t)](\mathbb{E}_{0,x_0}[u(t)] - \mathbb{E}_{0,x_0}[\hat{u}(t)])(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]) \\
 &+ (\mathbb{E}_{0,x_0}[u(t)] - \mathbb{E}_{0,x_0}[\hat{u}(t)])^2\}, \tag{10}
 \end{aligned}$$

where  $\mathbb{E}_{0,x_0}[\hat{u}(t)] = \hat{k}(t)\mathbb{E}_{0,x_0}[X^u(t)] + \hat{l}(t)$ . Although the two control vectors  $\mathbb{E}_{0,x_0}[u(t)]$  and  $u(t) - \mathbb{E}_{0,x_0}[u(t)]$  can be decided independently, they should be chosen such that

$$\mathbb{E}_{0,x_0}[u(t) - \mathbb{E}_{0,x_0}[u(t)]] = k^u(t)\mathbb{E}_{0,x_0}[X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]] = 0. \tag{11}$$

With the help of the system dynamics in (9), it is easy to see that  $\mathbb{E}_{0,x_0}[X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]] = 0$  holds. Thus, (11) would be satisfied automatically.

Therefore, we now can transfer the problem  $(MV(\lambda, \omega))$  into the following mean-field reformulation  $(MF(\lambda, \omega))$ ,

$$\begin{aligned} & \max \mathbb{E}_{0,x_0}[X^u(T)] - \lambda \mathbb{E}_{0,x_0}[X^u(T) - \mathbb{E}_{0,x_0}[X^u(T)]]^2 \\ & - \omega \mathbb{E}_{0,x_0} \left[ \int_0^T \lambda \sigma^2 ([a^u(t)]^2 - A^u(t)) \left\{ [k^u(t) - \hat{k}(t)]^2 (X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)])^2 \right. \right. \\ & + 2[k^u(t) - \hat{k}(t)](\mathbb{E}_{0,x_0}[u(t)] - \mathbb{E}_{0,x_0}[\hat{u}(t)])(X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]) \\ & \left. \left. + (\mathbb{E}_{0,x_0}[u(t)] - \mathbb{E}_{0,x_0}[\hat{u}(t)])^2 \right\} dt \right], \\ & \text{s.t. } (E_{0,x_0}[X^u(\cdot)], E_{0,x_0}[u(\cdot)]) \text{ satisfies (8),} \\ & (X^u(\cdot) - E_{0,x_0}[X^u(\cdot)], \mathbb{E}_{0,x_0}[u(\cdot)], u(\cdot) - \mathbb{E}_{0,x_0}[u(\cdot)]) \text{ satisfies (9).} \end{aligned}$$

The mean-field reformulation is a separable two-dimensional linear quadratic control problem, which can be readily solved by dynamic programming.

**Theorem 3.7.** *The self-coordination mean-variance strategy which solves  $(MV(\lambda, \omega))$ ,  $\{u^{\text{Coor}}(t, x)\}_{t \geq 0}^T$ , is given as*

$$u^{\text{Coor}}(t, x) = k^{\text{Coor}}(t)x + l^{\text{Coor}}(t), \tag{12}$$

with

$$\begin{aligned} k^{\text{Coor}}(t) &= \hat{\beta}^{\text{Coor}}(t) \cdot \frac{\mu - r}{\sigma^2}, \\ l^{\text{Coor}}(t) &= \left[ \lambda^{-1} \hat{\gamma}^{\text{Coor}}(t) - \hat{\beta}^{\text{Coor}}(t) \left( e^{rt} x_0 + \lambda^{-1} \delta \int_0^t e^{r(t-s)} \hat{\gamma}^{\text{Coor}}(s) ds \right) \right] \cdot \frac{\mu - r}{\sigma^2}, \end{aligned}$$

where the parameters  $\hat{\beta}^{\text{Coor}}(t)$  and  $\hat{\gamma}^{\text{Coor}}(t)$  are defined as follows,

$$\hat{\beta}^{\text{Coor}}(t) = \frac{\omega A^u(t) - \zeta^u(t)}{\omega([a^u(t)]^2 - A^u(t)) + \zeta^u(t)}, \quad \hat{\gamma}^{\text{Coor}}(t) = \frac{\omega \gamma^u(t) + e^{r(T-t)}}{2\omega([a^u(t)]^2 - A^u(t)) + 2\zeta^u(t)},$$

other parameters are governed by the following ODEs,

$$\begin{cases} \zeta_t^u(t) + 2r\zeta^u(t) - \delta[\omega([a^u(t)]^2 - A^u(t)) + \zeta^u(t)][\hat{\beta}^{\text{Coor}}(t)]^2 + \frac{\omega\delta[A^u(t)]^2}{[a^u(t)]^2 - A^u(t)} = 0, \\ \zeta^u(T) = 1, \\ a_t^u(t) + ra^u(t) + (\mu - r)a^u(t)k^{\text{Coor}}(t) = 0, \\ a^u(T) = 1, \\ A_t^u(t) + 2A^u(t)[r + (\mu - r)k^{\text{Coor}}(t)] - \sigma^2([a^u(t)]^2 - A^u(t))(k^{\text{Coor}}(t))^2 = 0, \\ A^u(T) = 0, \\ \beta_t^u(t) + r\beta^u(t) + [r + (\mu - r)k^{\text{Coor}}(t)]\beta^u(t) - (\mu - r)A^u(t)k^{\text{Coor}}(t) \\ \quad + \sigma^2([a^u(t)]^2 - A^u(t))(k^{\text{Coor}}(t))^2 = 0, \\ \beta^u(T) = 0, \\ \gamma_t^u(t) + 2\delta\beta^u(t)\hat{\gamma}^{\text{Coor}}(t) + [r + (\mu - r)k^{\text{Coor}}(t)]\gamma^u(t) + 2\delta A^u(t)\hat{\gamma}^{\text{Coor}}(t) \\ \quad - 2(\mu - r)([a^u(t)]^2 - A^u(t))k^{\text{Coor}}(t)\hat{\gamma}^{\text{Coor}}(t) = 0, \\ \gamma^u(T) = 1, \end{cases}$$

and  $\zeta_t^u(t)$ ,  $a_t^u(t)$ ,  $A_t^u(t)$ ,  $\beta_t^u(t)$  and  $\gamma_t^u(t)$  are the first order derivatives.

Finally, the efficient frontier is expressed as

$$\begin{aligned} & \text{Var}_{0,x_0}(X^u(T))|_{u^{\text{Coor}}} \\ &= \frac{\int_0^T e^{\int_s^T [2r + 2\delta\hat{\beta}^{\text{Coor}}(\tau) + \delta(\hat{\beta}^{\text{Coor}}(\tau))^2]d\tau} (\hat{\gamma}^{\text{Coor}}(s))^2 ds}{\delta \left( \int_0^T e^{r(T-s)} \hat{\gamma}^{\text{Coor}}(s) ds \right)^2} \left[ \mathbb{E}_{0,x_0}[X^u(T)]|_{u^{\text{Coor}}} - e^{rT}x_0 \right]^2, \\ & \text{for } \mathbb{E}_{0,x_0}[X^u(T)]|_{u^{\text{Coor}}} \geq e^{rT}x_0. \end{aligned}$$

*Proof.* Before solving the mean-field reformulation, we need to compute the value of  $\mathbb{E}_{0,x_0}[\hat{u}(t)]$ . Parameter  $B^u(t)$  in ODE (5) can be written into the following form,

$$B^u(t) = 2\lambda\beta^u(t)\mathbb{E}_{0,x_0}[X^u(t)] + \gamma^u(t),$$

where  $\beta^u(t)$  and  $\gamma^u(t)$  are new parameters. By noticing  $l^u(t) = \mathbb{E}_{0,x_0}[u(t)] - k^u(t)\mathbb{E}_{0,x_0}[X^u(t)]$ , we can transfer ODE (5) into

$$\begin{cases} 2\lambda\beta_t^u(t)\mathbb{E}_{0,x_0}[X^u(t)] + \gamma_t^u(t) + 2\lambda\beta^u(t)[r\mathbb{E}_{0,x_0}[X^u(t)] + (\mu - r)\mathbb{E}_{0,x_0}[u(t)]] \\ \quad + [r + (\mu - r)k^u](2\lambda\beta^u(t)\mathbb{E}_{0,x_0}[X^u(t)] + \gamma^u(t)) \\ \quad + 2\lambda(\mu - r)A^u(t)(\mathbb{E}_{0,x_0}[u(t)] - k^u\mathbb{E}_{0,x_0}[X^u(t)]) \\ \quad - 2\lambda\sigma^2((a^u(t))^2 - A^u(t))k^u(\mathbb{E}_{0,x_0}[u(t)] - k^u\mathbb{E}_{0,x_0}[X^u(t)]) = 0, \\ \beta^u(T)\mathbb{E}_{0,x_0}[X^u(T)] + \gamma^u(T) = 1, \end{cases}$$

which implies the dynamics of  $\beta^u(t)$  and  $\gamma^u(t)$ . It is easy to check that

$$\begin{cases} (A_t^u(t) + \beta_t^u(t)) + [2r + (\mu - r)k^u(t)](A^u(t) + \beta^u(t)) = 0, \\ A^u(T) + \beta^u(T) = 0. \end{cases}$$

Thus,  $A^u(t) + \beta^u(t) = 0$  for all  $t$ , which further implies

$$\mathbb{E}_{0,x_0}[\hat{u}(t)] = \hat{k}(t)\mathbb{E}_{0,x_0}[X^u(t)] + \hat{l}(t) = \frac{(\mu - r)\gamma^u(t)}{2\sigma^2\lambda([a^u(t)]^2 - A^u(t))}.$$

Now, we solve the problem  $(MF(\lambda, \omega))$ . To simplify the derivation, we denote the portfolio strategy and the state, respectively, as

$$\begin{aligned} \pi(t) &= (\pi^1(t), \pi^2(t)) := (k^u(t), \mathbb{E}_{0,x_0}[u(t)]), \\ Z(t) &= (Z^1(t), Z^2(t)) := (X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)], \mathbb{E}_{0,x_0}[X^u(t)]). \end{aligned}$$

As  $\mathbb{E}_{0,x_0}[Z^1(t)] = 0$ , we have

$$\mathbb{E}_{0,x_0} \left[ \int_0^T \left\{ -2\omega\lambda\sigma^2([a^u(t)]^2 - A^u(t))(\pi^1(t) - \hat{k}(t))(\pi^2(t) - \mathbb{E}_{0,x_0}[\hat{u}(t)])Z^1(t) - 2\lambda\sigma^2\zeta(t)\pi^1(t)\pi^2(t)Z^1(t) \right\} dt \right] = 0,$$

holds for any deterministic function  $\zeta^u(t)$ . Thus, the following auxiliary optimization problem  $(\mathcal{A})$ ,

$$\begin{aligned} \max \mathbb{E}_{0,Z^1(0),Z^2(0)} & \left[ Z^2(T) - \lambda(Z^1(T))^2 - \lambda\sigma^2 \int_0^T \zeta(t)\pi^1(t)\pi^2(t)Z^1(t)dt \right. \\ & \left. - \omega\lambda\sigma^2 \int_0^T ([a^u(t)]^2 - A^u(t)) \left\{ [\pi^1(t) - \hat{k}(t)]^2(Z^1(t))^2 + (\pi^2(t) - \mathbb{E}_{0,x_0}[\hat{u}(t)])^2 \right\} dt \right], \\ \text{s.t. } & (Z^2(\cdot), \pi^2(\cdot)) \text{ satisfies (8),} \\ & (Z^1(\cdot), \pi^1(\cdot), \pi^2(\cdot)) \text{ satisfies (9),} \end{aligned}$$

has the same optimal strategy as problem  $(MF(\lambda, \omega))$ .

Then, the value function of  $(\mathcal{A})$  at time  $t$  can be defined as

$$\begin{aligned} & V(t, z^1, z^2) \\ &= \max_{\{\pi(s)\}_{s \geq t}^T} \mathbb{E}_{t,z^1,z^2} \left[ Z^2(T) - \lambda(Z^1(T))^2 - \lambda\sigma^2 \int_t^T \zeta(s)\pi^1(s)\pi^2(s)Z^1(s)ds \right. \\ & \left. - \omega\lambda\sigma^2 \int_t^T ([a^u(s)]^2 - A^u(s)) \left\{ [\pi^1(s) - \hat{k}(s)]^2(Z^1(s))^2 + (\pi^2(s) - \mathbb{E}_{0,x_0}[\hat{u}(s)])^2 \right\} ds \right], \end{aligned}$$

where  $\mathbb{E}_{t,z^1,z^2}[\cdot] = \mathbb{E}[\cdot | Z^1(t) = z^1, Z^2(t) = z^2]$ . Obviously,  $V(T, z^1, z^2) = (z^1)^2 - \lambda(z^2)$ . We assume that  $V(t, z^1, z^2) \in \mathcal{C}^{1,2,2}$  (once continuously differentiable on  $t$  and twice continuously differentiable on  $z^1$  and  $z^2$ ).

Without confusion, we omit the time index in the following analysis.

The value function of  $(\mathcal{A})$  at time 0 is the solution of following HJB equation,

$$\begin{cases} \max_{\pi_t} \{ f(t, \pi^1, \pi^2, z^1, z^2) + \mathbf{A}^\pi v(t, z^1, z^2) \} = 0, \\ v(T, z^1, z^2) = z^2 - \lambda(z^1)^2, \end{cases}$$

where

$$\begin{aligned} & f(t, \pi^1, \pi^2, z^1, z^2) \\ &= -\lambda\sigma^2\zeta^u(t)\pi^1\pi^2z^1 - \omega\lambda\sigma^2([a^u(t)]^2 - A^u(t)) \left\{ (\pi^1 - \hat{k}(t))^2(z^1)^2 + (\pi^2 - \mathbb{E}_{0,x_0}[\hat{u}(t)])^2 \right\}, \end{aligned}$$

and  $\mathbf{A}^\pi v(t, z^1, z^2)$  is the infinitesimal generator for function  $v(t, z^1, z^2) \in \mathcal{C}^{1,2,2}$ ,

$$\begin{aligned} \mathbf{A}^\pi v(t, z^1, z^2) &= v_t + v_{z^2}[rz^2 + (\mu - r)\pi^2] + v_{z^1}[rz^1 + (\mu - r)\pi^1z^1] \\ & \quad + \frac{1}{2}v_{z^1z^1}\sigma^2(\pi^1z^1 + \pi^2)^2. \end{aligned}$$

We conjecture the solution as the following quadratic form,

$$V(t, z^1, z^2) = -\lambda\zeta^u(t)(z^1)^2 + e^{r(T-t)}z^2 + \xi^u(t),$$

with  $\zeta^u(T) = 1, \xi^u(T) = 0$ . It is easy to see that

$$\begin{aligned} & f(t, \pi^1, \pi^2, z^1, z^2) + \mathbf{A}^\pi V(t, z^1, z^2) \\ &= -\lambda(z^1)^2 \left\{ \omega\sigma^2([a^u(t)]^2 - A^u(t))(\pi^1 - \hat{k}(t))^2 + \zeta_t^u(t) + 2\zeta^u(t)[r + (\mu - r)\pi^1] \right. \\ & \quad \left. + \zeta^u(t)\sigma^2(\pi^1)^2 \right\} - \omega\lambda\sigma^2([a^u(t)]^2 - A^u(t))(\pi^2 - \mathbb{E}_{0,x_0}[\hat{u}(t)])^2 + e^{r(T-t)}(\mu - r)\pi^2 \\ & \quad - \lambda\zeta^u(t)\sigma^2(\pi^2)^2 + \xi_t^u(t). \end{aligned}$$

Thus, the optimal strategy is given as

$$\begin{aligned} (\pi^1)^{\text{Coor}}(t) &= \frac{\omega A^u(t) - \zeta^u(t)}{\omega([a^u(t)]^2 - A^u(t)) + \zeta^u(t)} \cdot \frac{\mu - r}{\sigma^2} = \hat{\beta}^{\text{Coor}}(t) \cdot \frac{\mu - r}{\sigma^2}, \\ (\pi^2)^{\text{Coor}} &= \frac{1}{\lambda} \cdot \frac{\omega\gamma^u(t) + e^{r(T-t)}}{2\omega([a^u(t)]^2 - A^u(t)) + 2\zeta^u(t)} \cdot \frac{\mu - r}{\sigma^2} = \hat{\gamma}^{\text{Coor}}(t) \cdot \frac{1}{\lambda} \cdot \frac{\mu - r}{\sigma^2}. \end{aligned}$$

Substituting them back to  $\max_\pi \{f(t, \pi^1, \pi^2, z^1, z^2) + \mathbf{A}^\pi V(t, z^1, z^2)\} = 0$  yields

$$\begin{cases} \zeta_t^u(t) + 2r\zeta^u(t) - \delta[\omega([a^u(t)]^2 - A^u(t)) + \zeta^u(t)](\hat{\beta}^{\text{Coor}}(t))^2 + \frac{\omega\delta[A^u(t)]^2}{[a^u(t)]^2 - A^u(t)} = 0, \\ \zeta^u(T) = 1, \\ \xi_t^u(t) + \lambda^{-1}\delta \left[ \omega([a^u(t)]^2 - A^u(t)) + \zeta^u(t) \right] (\hat{\gamma}^{\text{Coor}}(t))^2 - \frac{\omega\gamma^2}{4([a^u(t)]^2 - A^u(t))} = 0, \\ \xi^u(T) = 0. \end{cases}$$

Therefore, we have

$$k^{\text{Coor}}(t) = \hat{\beta}^{\text{Coor}}(t) \cdot \frac{\mu - r}{\sigma^2}, \quad \mathbb{E}_{0,x_0}[u^{\text{Coor}}(t)] = \hat{\gamma}^{\text{Coor}}(t) \cdot \frac{1}{\lambda} \cdot \frac{\mu - r}{\sigma^2}.$$

According to the dynamics of  $\mathbb{E}_{0,x_0}[X^u(t)]$  specified in (8), we have

$$\mathbb{E}_{0,x_0}[X^u(t)] = e^{rt}x_0 + \lambda^{-1}\delta \int_0^t e^{r(t-s)}\hat{\gamma}^{\text{Coor}}(s)ds,$$

which implies

$$\begin{aligned} & l^{\text{Coor}}(t) \\ &= \mathbb{E}_{0,x_0}[u^{\text{Coor}}(t)] - k^{\text{Coor}}(t)\mathbb{E}_{0,x_0}[X^u(t)] \\ &= \left[ \lambda^{-1}\hat{\gamma}^{\text{Coor}}(t) - \hat{\beta}^{\text{Coor}}(t) \left( e^{rt}x_0 + \lambda^{-1}\delta \int_0^t e^{r(t-s)}\hat{\gamma}^{\text{Coor}}(s)ds \right) \right] \cdot \frac{\mu - r}{\sigma^2}. \end{aligned}$$

Applying Itô Lemma to (9) and taking expectation, we obtain

$$\begin{cases} d\mathbb{E}_{0,x_0}[X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]]^2 = \left\{ [2r + 2(\mu - r)k^u(t) + \sigma^2(k^u(t))^2] \right. \\ \quad \left. \cdot \mathbb{E}_{0,x_0}[X^u(t) - \mathbb{E}_{0,x_0}[X^u(t)]]^2 + \sigma^2(\mathbb{E}_{0,x_0}[u(t)])^2 \right\} dt, \\ \mathbb{E}_{0,x_0}[X^u(0) - \mathbb{E}_{0,x_0}[X^u(0)]]^2 = 0. \end{cases} \tag{13}$$



Substituting the equilibrium strategy in (13), we have

$$\mathbb{E}_{0,x_0} [X^u(t) - \mathbb{E}_{0,x_0} [X^u(t)]]^2 = \lambda^{-2} \delta \int_0^t e^{\int_s^t [2r+2\delta\hat{\beta}^{\text{Coor}}(\tau)+\delta(\hat{\beta}^{\text{Coor}}(\tau))^2]d\tau} (\hat{\gamma}^{\text{Coor}}(s))^2 ds.$$

Therefore, we have the variance at the terminal as

$$\begin{aligned} & \text{Var}_{0,x_0}(X^u(T))|_{u^{\text{Coor}}} \\ &= \frac{\int_0^T e^{\int_s^T [2r+2\delta\hat{\beta}^{\text{Coor}}(\tau)+\delta(\hat{\beta}^{\text{Coor}}(\tau))^2]d\tau} (\hat{\gamma}^{\text{Coor}}(s))^2 ds}{\delta \left( \int_0^T e^{r(T-s)} \hat{\gamma}^{\text{Coor}}(s) ds \right)^2} \left[ \mathbb{E}_{0,x_0} [X^u(T)]|_{u^{\text{Coor}}} - e^{rT} x_0 \right]^2. \end{aligned}$$

□

**Remark 3.8.** Similar to Proposition 3.3, the self-coordination strategy is not explicitly given, which depends on a circular connected system of  $k^{\text{Coor}}$ ,  $\hat{\beta}^{\text{Coor}}$  and  $A^u(t)$  and can be obtained by a backward numerical method.

**Remark 3.9.** The solution schemes in Proposition 3.6 and Theorem 3.7 are also applicable for the punishment term,

$$\mathbb{E}_{t,x} \left[ \int_t^T \lambda g(s) [k^u(s)X^u(s) + l^u(s) - k^\nu(s)X^u(s) - l^\nu(s)]^2 ds \right],$$

where  $g(s)$  is a deterministic function of time  $s$ . Furthermore, when the parameters of the market,  $\mu$ ,  $r$ , and  $\omega$ , are deterministic functions of time, our results still hold.

**Remark 3.10.** Consider a linear combination strategy of the pre-committed strategy and time consistent strategy as follows,

$$u^{\text{Comb}}(t, x) = \omega_1 \frac{\mu - r}{\sigma^2} x - \omega_1 \left( e^{rt} x_0 + \frac{e^{\delta T - r(T-t)}}{2\lambda} \right) \frac{\mu - r}{\sigma^2} + \omega_2 \frac{e^{-r(T-t)}}{2\lambda} \frac{\mu - r}{\sigma^2},$$

where  $\omega_1$  and  $\omega_2$  are the weights on the pre-committed strategy and the time consistent strategy, respectively. The self-coordination strategy can be rewritten as

$$\begin{aligned} u^{\text{Coor}}(t, x) &= \hat{\beta}^{\text{Coor}}(t) \frac{\mu - r}{\sigma^2} x - \hat{\beta}^{\text{Coor}}(t) \left( e^{rt} x_0 + \frac{\delta \int_0^t e^{r(t-s)} \hat{\gamma}^{\text{Coor}}(s) ds}{\lambda} \right) \frac{\mu - r}{\sigma^2} \\ &\quad + \hat{\gamma}^{\text{Coor}}(t) \frac{1}{\lambda} \frac{\mu - r}{\sigma^2}. \end{aligned}$$

Comparing the two strategies, we can see that they are quite different. Even with a time-dependent weights, the combination strategy can not generate the self-coordination strategy.

Theorem 3.7 has shown that the self-coordination mean-variance strategy in continuous-time setting also takes a linear feedback form, and the corresponding efficient frontier is also a line in mean-standard deviation plane. Although the calculation of parameters in the self-coordination mean-variance strategy seems complicated, it does not depend on risk attitude parameter  $\lambda$  and can be computed off-line easily by, for example, Euler method.

The self-coordination strategy is the pre-committed optimal strategy of problem  $(MV(\lambda, \omega))$  (i.e.,  $MF(\lambda, \omega)$ ). It is easy to check that the pre-committed strategy and the time consistent strategy are the special cases of self-coordination strategies with  $\omega = 0$  and  $\omega = +\infty$ , respectively. Although the pre-committed strategy and the time consistent strategy are derived by very different approaches, they can be unified in the self-coordination framework. The objective of problem  $(MV(\lambda, \omega))$  contains two terms, one is the original mean-variance objective

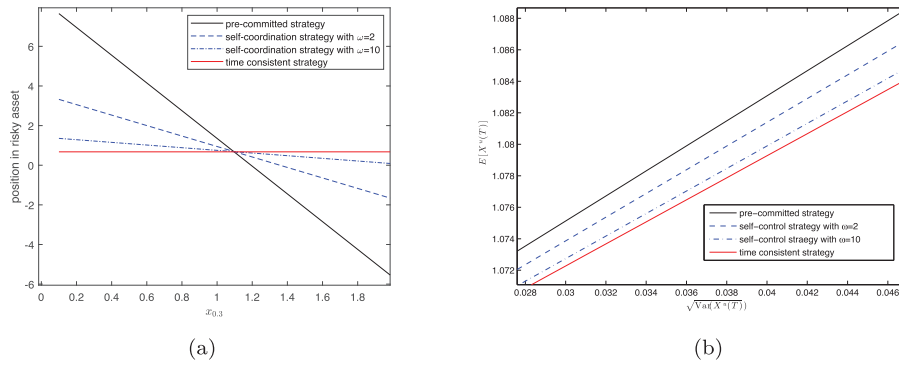


FIGURE 2. Comparisons among strategies. (a) Strategies with  $\lambda = 5$ . (b) Efficient frontiers.

and the other is the punishment term. On one hand, when  $\omega = 0$ , the punishment term vanishes. Problem  $(MV(\lambda, \omega))$  reduces into the original  $MV_0(\lambda)$  and the self-coordination strategy is just the pre-committed strategy of  $MV_0(\lambda)$ . The self-coordination strategy includes the pre-committed strategy because  $MV_0(\lambda)$  is a special case of  $(MV(\lambda, \omega))$ . On the other hand, when  $\omega = +\infty$ , the original mean-variance objective vanishes. The self-coordination strategy of  $(MV(\lambda, \omega))$  can be simply obtained as  $\hat{k}(t)x_t + \hat{l}(t)$ , which has the zero punishment. It is easy to check that  $\hat{k}(t) = 0$  and  $\hat{l}(t) = \frac{e^{-r(T-t)}(\mu-r)}{2\lambda\sigma^2}$ , which implies that the self-coordination strategy is just the time consistent strategy. The self-coordination strategy includes the time consistent strategy because the punishment term forces the investors not deviate from the time consistent strategy.

Next, we would like to use an example to illustrate the influences of trade-off parameter  $\omega$  on the investment performance.

**Example 3.11.** We assume the market parameters set as  $r = 0.05$ ,  $\mu = 0.12$ ,  $\sigma = 0.10$  and  $T = 1$ . Figure 2a shows the pre-committed strategy, two self-coordination strategies and time consistent strategy as functions of wealth level at time 0.3,  $x_{0.3}$ . These strategies have the same risk aversion  $\lambda = 5$ . We can see that the pre-committed strategy and self-coordination strategies have negative slopes, and the time consistent strategy is a constant. Comparing to the self-coordination strategy with  $\omega = 10$ , the self-coordination strategy with  $\omega = 2$  has larger negative slope, which implies that it changes more as long as the wealth level changes. Furthermore, the self-coordination strategy with  $\omega = 2$  is also much closer to pre-committed strategy. The efficient frontiers achieved by these strategies are reported in Figure 2b. We can see that the efficient frontiers achieved by self-coordination strategies lie in between the two achieved by the pre-committed strategy and the time consistent strategy.

For different  $\omega$ s, we also compute the corresponding Sharpe ratios, which are depicted in Figure 3. It is evident that the smaller the  $\omega$ , the larger the Sharpe ratio achieved by the self-coordination strategy. Please note that the largest Sharpe ratio and smallest Sharpe ratio are achieved by the pre-committed strategy and the time consistent strategy, respectively.

Figure 4 reveals the following interesting findings for the total punishment: 1) the stronger the willpower to conduct self-coordination (the smaller the value of  $\omega$ ), the better the long-term investment performance, and the larger the total punishment, *i.e.*, to achieve better mean-variance efficient frontier of the terminal wealth, the investor needs to bear larger punishment; 2) the less risk aversion of the investor (the smaller the value of  $\lambda$ ), the larger the intention of holding risky assets, and the harder to conduct self-coordination; 3) when  $\lambda = +\infty$ , the total punishment is zero, *i.e.*, when the investor is fully risk averse and only invests in riskless asset, there is no need to conduct self-coordination.

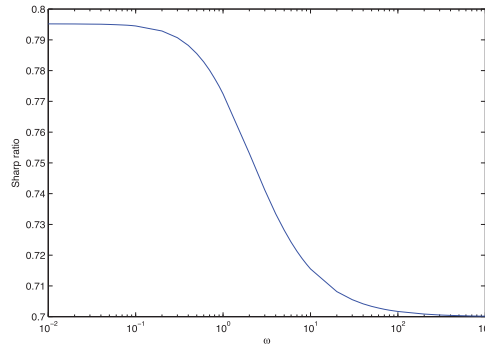


FIGURE 3. Sharpe ratio *vs.*  $\omega$ .

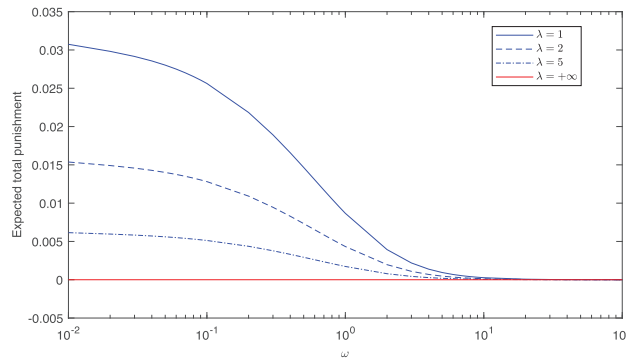


FIGURE 4. Total punishment *vs.*  $\omega$ .

### 4. CONCLUSION

We have extended the two-tier planner-doer game framework of [8] to continuous time setting in this paper and applied the framework to analyze the continuous time mean-variance portfolio selection problem. The two-tier planner-doer game framework adds a punishment term into the objective, which makes the portfolio selection problem quite challenging. We have changed the problem into its equivalent mean-filed reformulation and derived the analytical self-coordination strategy. Our work provides a different way to resolve time inconsistency of continuous time mean-variance portfolio selection problem.

While we have demonstrated that the continuous-time two-tier planner-doer game framework exhibits promising theoretical properties, our study falls short in providing an in-depth exploration of the selection of an appropriate punishment term that aligns with Assumption 3.1. The punishment term we’ve examined in our paper possesses sound economic significance but introduces computational complexities when addressing the associated problem.

Therefore, an avenue for future research in this domain is to identify a punishment term that not only retains its economic relevance but also lends itself to straightforward mathematical handling. Additionally, we suggest extending the applicability of this framework to address other time-inconsistent portfolio selection challenges within a continuous-time context, such as the mean-conditional value-at-risk problem. These research directions hold the potential to enhance the practicality and effectiveness of our proposed framework.

*Acknowledgements.* This work was supported by National Natural Science Foundation of China [Grants 71971083, 71931004, 72171138]; Program for Innovative Research Team of Shanghai University of Finance and Economics [Grant 2020110930]; Open Research Fund of Key Laboratory of Advanced Theory and Application in Statistics and Data Science-Ministry of Education, East China Normal University.

## REFERENCES

- [1] P. Artzner, F. Delbaen, J.M. Eber, D. Heath and H. Ku, Coherent multiperiod risk adjusted values and Bellman's principle. *Ann. Oper. Res.* **152** (2007) 5–22.
- [2] S. Basak and G. Chabakauri, Dynamic mean-variance asset allocation. *Rev. Financ. Stud.* **23** (2010) 2970–3016.
- [3] T. Björk and A. Murgoci, A general theory of Markovian time inconsistent stochastic control problems. *Available at SSRN 1694759* (2010).
- [4] T. Björk, A. Murgoci and X.Y. Zhou, Mean-variance portfolio optimization with state dependent risk aversion. *Math. Financ.* **24** (2014) 1–24.
- [5] K. Boda and J.A. Filar, Time consistent dynamic risk measures. *Math. Methods Oper. Res.* **63** (2006) 169–186.
- [6] X.Y. Cui, D. Li, S.Y. Wang and S.S. Zhu, Better than dynamic mean-variance: Time inconsistency and free cash flow stream. *Math. Financ.* **22** (2012) 346–378.
- [7] X.Y. Cui, X. Li and D. Li, Unified framework of mean-field formulations for optimal multi-period mean-variance portfolio selection. *IEEE Trans. Automat. Contr.* **59** (2014) 1833–1844.
- [8] X.Y. Cui, D. Li and Y. Shi, Self-coordination in time inconsistent stochastic decision problems: A planner-doer game framework. *J. Econ. Dyn. Control* **75** (2017) 91–113.
- [9] D. Dang and P.A. Forsyth, Better than pre-commitment mean-variance portfolio allocation strategies: A semi-self-financing Hamilton–Jacobi–Bellman equation approach. *Eur. J. Oper. Res.* **250** (2016) 827–841.
- [10] D. Fudenberg and D.K. Levine, A dual-self model of impulse control. *Am. Econ. Rev.* **96** (2006) 1449–1476.
- [11] D. Fudenberg and D.K. Levine, Timing and self-control. *Econometrica* **80** (2012) 1–42.
- [12] S.R. Grenadier and N. Wang, Investment under uncertainty and time-inconsistent preferences. *J. Financ. Econ.* **84** (2007) 2–39.
- [13] F. Gul and W. Pesendorfer, Temptation and self-control. *Econometrica* **69** (2001) 1403–1435.
- [14] R.H. Strotz, Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* **23** (1956) 165–180.
- [15] R.H. Thaler, Some empirical evidence on dynamic inconsistency. *Econ. Lett.* **8** (1981) 201–207.
- [16] J. Wang and P.A. Forsyth, Continuous time mean variance asset allocation: A time-consistent strategy. *Eur. J. Oper. Res.* **209** (2011) 184–201.
- [17] X.Y. Zhou and D. Li, Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Appl. Math. Optim.* **42** (2000) 19–33.

### Please help to maintain this journal in open access!



This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting [subscribers@edpsciences.org](mailto:subscribers@edpsciences.org).

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.