A Mean-Field Control Problem of Optimal Portfolio Liquidation with Semimartingale Strategies

Guanxing Fu[*] Ulrich Horst[†] Xiaonyu Xia[‡]

July 3, 2022

Abstract

We consider a mean-field control problem with càdlàg semimartingale strategies arising in portfolio liquidation models with transient market impact and self-exciting order flow. We show that the value function depends on the state process only through its law, and that it is of linear-quadratic form and that its coefficients satisfy a coupled system of non-standard Riccati-type equations. The Riccati equations are obtained heuristically by passing to the continuous-time limit from a sequence of discrete-time models. A sophisticated transformation shows that the system can be brought into standard Riccati form from which we deduce the existence of a global solution. Our analysis shows that the optimal strategy jumps only at the beginning and the end of the trading period.

AMS Subject Classification: 93E20, 91B70, 60H30

Keywords: mean-field control, semimartingale strategy, portfolio liquidation

1 Introduction

Let $T \in (0, \infty)$ and let W be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be the augmented Brownian filtration. In this paper we consider the mean-field stochastic control problem

$$\min_{Z\in\mathscr{A}} \mathbb{E}\left[\int_0^T \Big(Y_{s-}\, dZ_s + \frac{\gamma_2}{2}\, d[Z]_s + \sigma_s d[Z,W]_s\Big) + \int_0^T \lambda X_s^2\, ds\right] \tag{1.1}$$

subject to the state dynamics

$$\begin{cases} dX_s = -dZ_s \\ dY_s = \Big(-\rho Y_s + \gamma_1 C'_s\Big)\, dt + \gamma_2\, dZ_s + \sigma_s\, dW_s \\ dC_s = -(\beta-\alpha)C_s ds + \alpha(\mathbb{E}[x_0] - \mathbb{E}[X_s])\, ds \\ X_{0-} = x_0;\ Y_{0-} = C_{0-} = 0;\ X_T = 0 \end{cases} \tag{1.2}$$

for $0 \le s \le T$ where the set of admissible controls is given by

$$\mathscr{A} := \{Z : Z \text{ is an } \mathbb{F} \text{ semimartingale with } [Z]_T < \infty\}.$$

[*]The Hong Kong Polytechnic University, Department of Applied Mathematics, Hung Hom, Kowloon, Hong Kong. G. Fu's research is supported by The Hong Kong RGC (ECS No.25215122) and NSFC Grant No. 12101523.

[†]Humboldt University Berlin, Department of Mathematics and School of Business and Economics, Unter den Linden 6, 10099 Berlin

[‡]Wenzhou University, College of Mathematics and Physics, Wenzhou 325035, PR China; corresponding author. X. Xia's research is supported by NSFC Grant No. 12101465.

Control problems of the above form arise in models of optimal portfolio liquidation with instantaneous and transient market impact and self-exciting order flow. In such models the process $(X_t)_{t\in[0,T]}$ describes the portfolio holdings ("inventory") of a large investor, the terminal state constraint $X_T = 0$ reflects the liquidation constraint, $(Z_t)_{t\in[0,T]}$ is the trading strategy, and $(Y_t)_{t\in[0,T]}$ specifies the transient impact of the investor's past trading on future transaction prices. We may think of Y as an additional drift added to an unaffected (martingale) benchmark price process or a randomly fluctuating spread that increases linearly in order flow and recovers from past trading at a constant rate ρ.

The process $(C_t)_{t\in[0,T]}$ can be viewed as describing the expected number of child orders resulting from the large investor's trading activity. There are many reasons why (large) selling orders may trigger child orders. For instance, extensive selling may diminish the pool of counterparties and/or generate herding effects where other market participants start selling in anticipation of further price decreases. Extensive selling may also attract predatory traders that employ front-running strategies; see Brunnermeier and Pedersen [8], Carlin et al [12] and Schied and Schöneborn [37] for an in-depth analysis of predatory trading.

Single and multi-player liquidation models with self-exciting order flow and absolutely continuous controls where $dZ_t = \xi_t dt$ have recently been considered in Chen et al [15] and Fu et al [23], respectively. In both models, the market order dynamics follows a Hawkes process with exponential kernel, and the large investor's trading triggers an additional flow of child orders whose rate increases linearly in the investor's traded volume. Assuming that the drift/spread depends on the aggregate order arrival rate (child order rate plus the large investor's rate) then leads to an equation of the above form for the transient impact factor Y.

While the restriction to absolutely continuous controls is standard in much of the liquidation literature, the assumption seems restrictive; it is often made for mathematical convenience as the resulting control problem is much simpler to analyze. In fact, while abstract existence and characterization of solutions results can be obtained for models allowing for more general classes of admissible strategies (see [29] and references therein), explicit solutions in stochastic settings are rarely available. Retaining the assumption that the large investor's trading activity triggers an absolutely continuous flow of child orders that increases linearly in his/her traded volume, we obtain explicit solutions for both the value function and the optimal strategy when allowing for general càdlàg semimartingale trading strategies.

Allowing for semimartingale strategies, we explicitly allow inventory processes to be of infinite variation.[1] Portfolio liquidation/choice models with inventory processes with infinite variation were considered by several authors before, including [1, 7, 24, 28, 33]. The most important difference between previous work and our model is that we consider a mean field control problem, while the control problems analyzed in the existing literature are standard ones.

Closest to our work are the recent papers by Ackermann et al [1] and Horst and Kivman [28]. As in [1] we analyze the continuous time model by passing to the limit from a sequence of discrete time model; by contrast [28] approximates semimartingale strategies by absolutely continuous ones. At the same time, the main difficulty in [1] arises from allowing a very general filtration while we consider a Brownian filtration but allow for self-exciting order flow which results in the said mean-field control problem.

When working with semimartingale strategies one needs to penalize the (co)variation of the semimartinagle strategy (with the driving Brownian motion); otherwise the optimization problem would not be well posed. This observation goes back at least to Gârleanu and Pedersen [24]. This explains the second and third term in our cost function. Similar cost terms have also been considered in [1, 28, 33]. The first term in the cost function describes the trading cost while the last term captures the market risk. These two terms are standard in the liquidation literature; see [2, 3, 14, 21, 30, 35] and references therein.

[1]Carmona and Webster [13] provide strong evidence that inventories of large traders often do have indeed a non-trivial quadratic variation component.

To the best of our knowledge, we are the first to address mean field control problems with càdlàg semimartingale strategies, which include mean field singular control problems as special cases. Mean field singular control problems have been considered by many authors, including [22, 25, 27, 31]. Among them, using a relaxed approach Fu and Horst [22] established an existence of optimal control result for a general class of mean field singular control problems. Guo et al [25] established a novel Itô's formula for the flow of measures on semimartingales; however, the examples provided in [25] are mean field control problems with either regular or singular controls. Hafayed [27] established a maximum principle for general class of mean field type singular control problems. Using a maximum principle approach, Hu et al [31] studied mean field type singular control games arising in harvesting problems.

Our paper also contributes to the literature on the characterization of non-Markovian singular control problems. One-dimensional models were studied in, e.g. [4, 5, 6, 33]. Multidimensional problems are much more difficult to analyze, even in Markovian settings; see Dianetti and Ferrari [17]. For non-Markovian ones, refer to e.g. [1, 19]. In [1], Ackermann et al solved a two-dimensional problem with random coefficients arising in optimal liquidation problems. In [19], Elie et al studied a multidimensional path-dependent singular control problem arising in utility maximization problems. Our control problem is a three-dimensional non-Markovian one with the non-Markovianity arising from a possibly non-Markovian volatility process σ and the mean field term $\mathbb{E}[X]$. Moreover, our strategy is of infinite variation.

The main challenge when analyzing control problems with semimartingale strategies (with or without mean-field term) is that there are usually no canonical candidates for the value functions and/or the optimal strategies. Even the linear-quadratic case is difficult to analyze; although it is intuitive that the value function is of linear-quadratic form, the dynamics of the coefficients is a priori not clear. This calls for case-by-case approaches when analyzing such problems. We follow the approach taken in [1, 24] and consider a sequence of discrete time models and then pass to a heuristic continuous time limit. Our approach suggests that the value function depends on the state process only through its distribution and that it is of linear quadratic form driven by three deterministic processes and a BSDE.

It turns out that the driving processes follow a system of Riccati equations that does not satisfy the assumptions that are usually required to guarantee the existence and uniqueness of a solution. Our main mathematical contribution is to show - through a sophisticated transformation - that our system can be rewritten in terms of a more standard system that satisfies the assumption in Kohlmann and Tang [32] under a weak interaction condition that bounds the impact of the child order flow on the market dynamics. Subsequently, we employ a non-standard verification argument that shows that the candidate optimal strategy is indeed optimal. The key idea is to rewrite the cost function as a sum of complete squares plus a correction term that turns out to be the value function.

Our analysis shows that the optimal strategy jumps only at the beginning and the end of the trading period. This is consistent with earlier findings in [24, 28, 29, 34]; in the absence of an external "trigger" there is no reason for the optimal strategy to jump, except at the terminal time to close the position and at the initial time.

The remainder of this paper is structured as follows. The main results and assumptions are stated in Section 2. The wellposedness of the system of the Riccati equations that specify the candidate value function and strategy is established in Section 3. The verification argument is given in Section 4. Numerical simulations illustrating the nature of the optimal solution are given in Section 5. Section 6 concludes. The heuristic derivation of the value function is postponed to an appendix.

Notation. For a deterministic function $\mathscr{X}$, denote by $\dot{\mathscr{X}}$ its derivative. For a stochastic process $\mathscr{X}$ satisfying some SDE, we still use the same notation $\dot{\mathscr{X}}$ to denote the drift of $\mathscr{X}$. For a matrix (or vector) $\mathscr{Y}$, $\mathscr{Y}_{ij}$ (or $\mathscr{Y}_i$) denotes its (i,j)- (or i-) component. For a space $\mathscr{D}$, we denote by $L^\infty([0,T];\mathscr{D})$ the space of all $\mathscr{D}$-valued bounded functions. $C([0,T];\mathscr{D})$ is the space of $\mathscr{D}$-valued continuous functions. For an integer n, we denote by $\mathbb{S}^n$ the space of symmetric $n \times n$ matrices. For a measure μ on $\mathbb{R}^n$ we denote by $\bar{\mu}$ the vector of expected values, and for a matrix $A \in \mathbb{R}^{n\times n}$ we put

$$\mathrm{Var}(\mu)(A) := \int_{\mathbb{R}^3} (x-\overline{\mu})^\top A(x-\overline{\mu})\mu(dx) = \int_{\mathbb{R}^3} x^\top Ax\mu(dx) - \overline{\mu}^\top A\overline{\mu}.$$

2 Main results

To solve the control problem (1.1)-(1.2) we consider the following dynamic version of (1.1)-(1.2):

$$\min_{Z\in\mathscr{A}_t} J(t,Z) := \min_{Z\in\mathscr{A}_t} \mathbb{E}\left[\int_t^T \Big(Y_{s-}\, dZ_s + \frac{\gamma_2}{2}\, d[Z]_s + \sigma_s d[Z,W]_s\Big) + \int_t^T \lambda X_s^2\, ds\right] \tag{2.1}$$

subject to the state dynamics

$$\begin{cases} dX_s = -dZ_s \\ dY_s = \Big(-\rho Y_s + \gamma_1 C'_s\Big)\, dt + \gamma_2\, dZ_s + \sigma_s\, dW_s \\ dC_s = -(\beta-\alpha)C_s ds + \alpha(\mathbb{E}[x_0] - \mathbb{E}[X_s])\, ds \\ (X_{t-}, Y_{t-}, C_{t-}) = \mathcal{X};\;\; X_T = 0, \end{cases} \tag{2.2}$$

where

$$\mathscr{A}_t = \{Z : Z \text{ is an } \mathbb{F} \text{ semimartingale starting at } t \text{ with } [Z]_T < \infty\}. \tag{2.3}$$

We denote the value function, given the random initial state $\mathcal{X} = (X_{t-}, Y_{t-}, C_{t-})$ at time $t \in [0,T)$, by

$$V(t,\mathcal{X}) = \inf_{Z\in\mathscr{A}_t} J(t,Z). \tag{2.4}$$

2.1 The value function

Let μ be the law of the initial state $\mathcal{X}$, and let $\bar{\mu}$ be the vector of expected values. Our goal is to represent the value function in terms of two deterministic symmetric $\mathbb{R}^{3\times 3}$-valued processes A, B, an $\mathbb{R}^3$-valued deterministic process D and a real-valued adapted square integrable process F as[2]

$$V(t,\mathcal{X}) = \mathrm{Var}(\mu)(A_t) + \bar{\mu}^\top B_t\bar{\mu} + D_t^\top\bar{\mu} + \mathbb{E}[F_t]. \tag{2.5}$$

The dynamics of the processes A, B, D, F is derived heuristically in Appendix A by first analyzing a discrete time model and then taking the limit as the time difference between two consecutive trading periods tends to zero. It turns out that:

- The process A is symmetric, satisfies $A_{11} = \gamma_2 A_{21}$, $A_{12} = \gamma_2 A_{22} + \frac{1}{2}$, $A_{13} = \gamma_2 A_{23}$, and the ODE

$$\begin{cases} \dot{A}_{11,t} = \left(-\lambda + \dfrac{(\rho A_{11,t}+\lambda)^2}{\gamma_2\rho+\lambda}\right) \\ \dot{A}_{13,t} = \left(\dfrac{\gamma_1(\beta-\alpha)}{\gamma_2}A_{11,t} + (\beta-\alpha)A_{13,t} - \dfrac{(\rho A_{11,t}+\lambda)\,(\gamma_1(\beta-\alpha) - 2\rho A_{13,t})}{2(\gamma_2\rho+\lambda)}\right) \\ \dot{A}_{33,t} = \left(2(\beta-\alpha)A_{33,t} + 2\dfrac{\gamma_1(\beta-\alpha)}{\gamma_2}A_{13,t} + \dfrac{(\gamma_1(\beta-\alpha)-2\rho A_{13,t})^2}{4(\gamma_2\rho+\lambda)}\right) \\ A_{11,T} = \dfrac{\gamma_2}{2}, \quad A_{13,T} = 0, \quad A_{33,T} = 0. \end{cases} \tag{2.6}$$

  The above is a standard ODE system that can be uniquely solved.

[2]The notation $\mathrm{Var}(\mu)(A_t)$ was introduced at the end of the introduction.

- The process B is symmetric, satisfies $B_{11} = \gamma_2 B_{21}$, $B_{12} = \gamma_2 B_{22} + \frac{1}{2}$, $B_{13} = \gamma_2 B_{23}$, and the fully coupled system of Riccati-type equations

$$\begin{cases} \dot{B}_{11,t} = \Bigg(2\dfrac{\gamma_1\alpha}{\gamma_2} B_{11,t} + 2\alpha B_{13,t} - \lambda \\ \qquad + \dfrac{\Big(2(\gamma_1\alpha - \gamma_2\rho)B_{11,t} + \gamma_1\gamma_2\alpha + 2\alpha\gamma_2 B_{13,t} - 2\gamma_2\lambda\Big)^2}{4\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \Bigg) \\ \dot{B}_{33,t} = \Bigg(2(\beta-\alpha)B_{33,t} + 2\dfrac{\gamma_1(\beta-\alpha)}{\gamma_2} B_{13,t} \\ \qquad + \dfrac{\Big(2(\gamma_1\alpha - \gamma_2\rho)B_{13,t} + 2\gamma_2\alpha B_{33,t} + \gamma_1\gamma_2(\beta-\alpha)\Big)^2}{4\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \Bigg) \\ \dot{B}_{13,t} = \Bigg\{ \dfrac{\gamma_1(\beta-\alpha)}{\gamma_2} B_{11,t} + \alpha B_{33,t} + (\beta - \alpha + \dfrac{\gamma_1\alpha}{\gamma_2}) B_{13,t} \\ \qquad + \Big(2(\gamma_1\alpha - \gamma_2\rho)B_{11,t} + \gamma_1\gamma_2\alpha + 2\alpha\gamma_2 B_{13,t} - 2\gamma_2\lambda\Big) \\ \qquad \cdot \dfrac{(2(\gamma_1\alpha - \gamma_2\rho)B_{13,t} + 2\gamma_2\alpha B_{33,t} + \gamma_1\gamma_2(\beta-\alpha))}{4\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \Bigg\} \\ B_{11,T} = \dfrac{\gamma_2}{2}, \quad B_{13,T} = 0, \quad B_{33,T} = 0. \end{cases} \tag{2.7}$$

  This system is complicated to analyze; its analysis is postponed to the next section.

- The vector-valued process D satisfies $D_2 = \gamma_2^{-1} D_1$, and the components D_1 and D_3 satisfy the coupled linear ODE system

$$\begin{cases} \dot{D}_{1,t} = \Bigg\{ - \dfrac{2\gamma_1\alpha\mathbb{E}[x_0]}{\gamma_2} B_{11,t} - 2\alpha\mathbb{E}[x_0] B_{13,t} + \dfrac{\gamma_1\alpha}{\gamma_2} D_{1,t} + \alpha D_{3,t} \\ \qquad + \Big(- 2\lambda\gamma_2 + 2(\gamma_1\alpha - \gamma_2\rho)B_{11,t} + \gamma_1\gamma_2\alpha + \alpha\gamma_2 B_{13,t}\Big) \\ \qquad \cdot \dfrac{\Big(- \gamma_1\gamma_2\alpha\mathbb{E}[x_0] + (\gamma_1\alpha - \gamma_2\rho)D_{1,t} + \alpha\gamma_2 D_{3,t}\Big)}{2\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \Bigg\} \\ \dot{D}_{3,t} = \Bigg\{ - 2\alpha\mathbb{E}[x_0] B_{33,t} - \dfrac{2\gamma_1\alpha\mathbb{E}[x_0]}{\gamma_2} B_{13,t} + \dfrac{\gamma_1(\beta-\alpha)}{\gamma_2} D_{1,t} + (\beta-\alpha) D_{3,t} \\ \qquad + \Big(2(\gamma_1\alpha - \gamma_2\rho)B_{13,t} + 2\gamma_2\alpha B_{33,t} + \gamma_1\gamma_2(\beta-\alpha)\Big) \\ \qquad \cdot \dfrac{\Big(- \gamma_1\gamma_2\alpha\mathbb{E}[x_0] + (\gamma_1\alpha - \gamma_2\rho)D_{1,t} + \alpha\gamma_2 D_{3,t}\Big)}{2\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \Bigg\} \\ D_{1,T} = D_{3,T} = 0 \end{cases} \tag{2.8}$$

- Due to the random volatility process, the process F satisfies a BSDE, namely

$$\begin{cases} -dF_t = \Bigg\{ \sigma_t^2 \dfrac{2A_{11,t} - \gamma_2}{2\gamma_2^2} + \alpha\gamma_1\mathbb{E}[x_0]\dfrac{D_{1,t}}{\gamma_2} + \alpha\mathbb{E}[x_0] D_{3,t} \\ \qquad - \dfrac{1}{4(\lambda + \gamma_2\rho - \alpha\gamma_1)} \left(-\alpha\gamma_1\mathbb{E}[x_0] + (\gamma_1\alpha - \gamma_2\rho)\dfrac{D_{1,t}}{\gamma_2} + \alpha D_{3,t} \right)^2 \Bigg\} dt - Z_t^F \, dW_t \\ F_T = 0. \end{cases} \tag{2.9}$$

We prove in Section 3 that the system (2.7) is well posed and admits a unique global solution if the feedback effect as measured by the constant α is weak enough. In this case, the joint dynamics of the process (A, B, D, F) is well-defined.

2.2 The main result

We assume throughout that the following **standing assumption** holds.

1. The coefficients γ_1, γ_2, α, β, ρ and λ are nonnegative constants.
2. The coefficients satisfy $\beta - \alpha > 0$ and $\gamma_2\rho - \gamma_1\alpha + \lambda > 0$.
3. The initial position x_0 is an integrable r.v. that is independent of the Brownian motion. The volatility process σ is a square integrable progressively measurable process.[3]

It will be convenient to rewrite the state dynamics and the cost function in matrix form as

$$\begin{aligned} d\mathcal{X}_s &= \Big(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G}\Big)\,ds + \mathcal{D}_s\,dW_s + \mathcal{K}\,dZ_s, \quad s \in [t, T), \\ \mathcal{X}_{t-} &= \mathcal{X} \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} \mathcal{H} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\rho & -\gamma_1(\beta-\alpha) \\ 0 & 0 & -(\beta-\alpha) \end{pmatrix}, \quad \overline{\mathcal{H}} = \begin{pmatrix} 0 & 0 & 0 \\ -\alpha\gamma_1 & 0 & 0 \\ -\alpha & 0 & 0 \end{pmatrix}, \\ \mathcal{G} &= \begin{pmatrix} 0 & \alpha\gamma_1\mathbb{E}[x_0] & \alpha\mathbb{E}[x_0] \end{pmatrix}^\top, \quad \mathcal{D}_s = \begin{pmatrix} 0 & \sigma_s & 0 \end{pmatrix}^\top, \\ \mathcal{K} &= \begin{pmatrix} -1 & \gamma_2 & 0 \end{pmatrix}^\top. \end{aligned} \tag{2.11}$$

The following is the main result of this paper. Its proof is given in Sections 3 and 4 below.

Theorem 2.1. *If the standing assumption is satisfied, and if either $\lambda = 0$ or $\lambda\rho\gamma_2 > 0$ and α is small enough, then the following holds.*

i) In terms of the processes A, B, D, F introduced in (2.6)-(2.9) the value function defined in (2.4) is given by

$$V(t, \mathcal{X}) = V(t, \mu) = \mathit{Var}(\mu)(A_t) + \bar{\mu}^\top B_t \bar{\mu} + D_t^\top \bar{\mu} + \mathbb{E}[F_t]. \tag{2.12}$$

ii) The optimal strategy $\widetilde{Z}$ jumps only at the beginning and the end of the trading period where the initial and terminal jump is given by

$$\Delta\widetilde{Z}_t = -\frac{I_t^A}{\tilde{a}}(\mathcal{X}_{t-} - \overline{\mu}) - \frac{I_t^B}{a}\overline{\mu} - \frac{I_t^D}{a} \quad \textit{and} \quad \Delta\widetilde{Z}_T = X_{T-} \tag{2.13}$$

respectively. On the time interval (t, T) the optimal strategy satisfies the dynamics

$$\begin{aligned} d\widetilde{Z}_s = \Bigg(&-\frac{\dot{I}_s^A}{\tilde{a}}(\mathcal{X}_s - \mathbb{E}[\mathcal{X}_s]) - \frac{\dot{I}_s^B}{a}\mathbb{E}[\mathcal{X}_s] - \frac{\dot{I}_s^D}{a} - \frac{I_s^A}{\tilde{a}}\mathcal{H}(\mathcal{X}_s - \mathbb{E}[\mathcal{X}_s]) \\ &- \frac{I_s^B}{a}((\mathcal{H} + \overline{\mathcal{H}})\mathbb{E}[\mathcal{X}_s] + \mathcal{G})\Bigg)\,ds - \frac{I_s^A}{\tilde{a}}\mathcal{D}_s\,dW_s, \qquad s \in (t, T) \end{aligned} \tag{2.14}$$

[3]We emphasize that σ is allowed to be degenerate.

where $\tilde{a} = \gamma_2\rho + \lambda$, $a = \gamma_2\rho - \gamma_1\alpha + \lambda$, *and the processes* I^A, I^B *and* I^D *are given by*

$$I^A = \begin{pmatrix} -\rho A_{11} - \lambda \\ -\rho\frac{A_{11}}{\gamma_2} + \rho \\ \frac{\gamma_1(\beta-\alpha)}{2} - \rho A_{13} \end{pmatrix}^\top, \quad I^B = \begin{pmatrix} \frac{\alpha\gamma_1-\gamma_2\rho}{\gamma_2}B_{11} + \alpha B_{13} - \lambda + \frac{\alpha\gamma_1}{2} \\ \frac{\alpha\gamma_1-\gamma_2\rho}{\gamma_2^2}B_{11} + \alpha\frac{B_{13}}{\gamma_2} + \rho - \frac{\alpha\gamma_1}{2\gamma_2} \\ \frac{\gamma_1(\beta-\alpha)}{2} + (\gamma_1\alpha - \gamma_2\rho)\frac{B_{13}}{\gamma_2} + \alpha B_{33} \end{pmatrix}^\top$$

and

$$I^D = -\frac{\alpha\gamma_1}{2}\mathbb{E}[x_0] + (\gamma_1\alpha - \gamma_2\rho)\frac{D_1}{2\gamma_2} + \frac{\alpha}{2}D_3.$$

3 Wellposedness of the Riccati Equation

In this section, we prove that the system (2.7) is well posed and has a unique global solution. Specifically, we prove the following result.

Theorem 3.1. *In addition to the* **standing assumption**, *let us assume that* α *is small enough and that* $\lambda, \gamma_2, \rho > 0$. *Then the matrix Riccati equation* (2.7) *admits a unique solution*

$$B \in L^\infty([0,T];\mathbb{R}^3) \cap C([0,T];\mathbb{R}^3).$$

To prove Theorem 3.1, it will be convenient to introduce the matrix-valued processes

$$\mathcal{P} = \begin{pmatrix} B_{11} & B_{13} \\ B_{13} & B_{33} \end{pmatrix}, \quad \mathcal{N}_1 = \begin{pmatrix} \frac{\gamma_1\alpha}{\gamma_2} + \frac{(\gamma_1\alpha-\gamma_2\rho)(\gamma_1\alpha-2\lambda)}{2\gamma_2(\gamma_2\rho-\gamma_1\alpha+\lambda)} & \alpha + \frac{\alpha(\gamma_1\alpha-2\lambda)}{2(\gamma_2\rho-\gamma_1\alpha+\lambda)} \\ \frac{\gamma_1(\beta-\alpha)}{\gamma_2} + \frac{(\gamma_1\alpha-\gamma_2\rho)\gamma_1(\beta-\alpha)}{2\gamma_2(\gamma_2\rho-\gamma_1\alpha+\lambda)} & \beta - \alpha + \frac{\gamma_1\alpha(\beta-\alpha)}{2(\gamma_2\rho-\gamma_1\alpha+\lambda)} \end{pmatrix}, \quad \mathcal{N}_2 = \begin{pmatrix} \gamma_1\alpha - \gamma_2\rho \\ \gamma_2\alpha \end{pmatrix},$$

$$\mathcal{N}_0 = \frac{1}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)}, \qquad \mathcal{M} = -\begin{pmatrix} \frac{\gamma_1^2\alpha^2-4\lambda\gamma_2\rho}{4(\gamma_2\rho-\gamma_1\alpha+\lambda)} & \frac{\gamma_1(\beta-\alpha)(\gamma_1\alpha-2\lambda)}{4(\gamma_2\rho-\gamma_1\alpha+\lambda)} \\ \frac{\gamma_1(\beta-\alpha)(\gamma_1\alpha-2\lambda)}{4(\gamma_2\rho-\gamma_1\alpha+\lambda)} & \frac{\gamma_1^2(\beta-\alpha)^2}{4(\gamma_2\rho-\gamma_1\alpha+\lambda)}, \end{pmatrix}, \qquad \mathcal{G} = \begin{pmatrix} \gamma_2 & 0 \\ 0 & 0 \end{pmatrix},$$

so that the system (2.7) can be rewritten in the matrix form as:

$$\begin{cases} \dot{\mathcal{P}}_t = \left(\mathcal{P}_t\mathcal{N}_2\mathcal{N}_0\mathcal{N}_2^\top\mathcal{P}_t + \mathcal{N}_1\mathcal{P}_t + \mathcal{P}_t\mathcal{N}_1^\top - \mathcal{M}\right), \quad t \in [0,T) \\ \mathcal{P}_T = \mathcal{G}. \end{cases} \tag{3.1}$$

The matrix-valued Riccati equation (3.1) does not satisfy the requirements of [32, Proposition 2.1, 2.2] as $\mathcal{M}$ is not positive semi-definite. To overcome this problem we employ a sophisticated transformation to bring the equation into standard Riccati-form. To this end, we define

$$\widetilde{\mathcal{P}} = \mathcal{P} + \widetilde{\Lambda},$$

where the matrix $\widetilde{\Lambda}$ is given by

$$\widetilde{\Lambda} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_3 \end{pmatrix} = \begin{pmatrix} \Lambda\alpha^2 & \Lambda\alpha(\beta-\alpha) \\ \Lambda\alpha(\beta-\alpha) & \Lambda(\beta-\alpha)^2 \end{pmatrix}$$

for some constant Λ that will be determined in what follows. The process $\widetilde{\mathcal{P}}$ satisfies the dynamics

$$\begin{aligned} \dot{\widetilde{\mathcal{P}}}_t &= \widetilde{\mathcal{P}}_t\mathcal{N}_2\mathcal{N}_0\mathcal{N}_2^\top\widetilde{\mathcal{P}}_t + \widetilde{\mathcal{N}_1}\widetilde{\mathcal{P}}_t + \widetilde{\mathcal{P}}_t\widetilde{\mathcal{N}_1}^\top - \widetilde{\mathcal{M}}, \quad t \in [0,T) \\ \widetilde{\mathcal{P}}_T &= \mathcal{G} + \widetilde{\Lambda}, \end{aligned} \tag{3.2}$$

where the matrices $\widetilde{\mathcal{N}_1}$ and $\widetilde{\mathcal{M}}$ are given by, respectively,

$$\widetilde{\mathcal{N}_1} = \mathcal{N}_1 - \widetilde{\Lambda}\mathcal{N}_2\mathcal{N}_0\mathcal{N}_2^\top \qquad \text{and} \qquad \widetilde{\mathcal{M}} = -\widetilde{\Lambda}\mathcal{N}_2\mathcal{N}_0\mathcal{N}_2^\top\widetilde{\Lambda} + \mathcal{N}_1\widetilde{\Lambda} + \widetilde{\Lambda}\mathcal{N}_1^\top + \mathcal{M}.$$

It is enough to prove that the above matrix-valued ODE has a unique solution for a suitable $\Lambda \in \mathbb{R}$.

Proof of Theorem 3.1. We are going to show that the equation (3.2) satisfies the assumptions of of [32, Proposition 2.1, 2.2], that is that the matrix $\widetilde{\mathcal{M}}$ is positive semidefinite for a suitably chosen constant Λ. The entries of $\widetilde{\mathcal{M}}$ are given by, respectively,

$$\begin{aligned}
\widetilde{\mathcal{M}}_{11} =& - \frac{\Big\{\lambda_1(\gamma_1\alpha - \gamma_2\rho) + \lambda_2\gamma_2\alpha\Big\}^2}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} + 2\lambda_1\left(\frac{\gamma_1\alpha}{\gamma_2} + \frac{(\gamma_1\alpha - \gamma_2\rho)(\gamma_1\alpha - 2\lambda)}{2\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) \\
&+ 2\lambda_2\left(\alpha + \frac{\alpha(\gamma_1\alpha - 2\lambda)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) - \frac{\gamma_1^2\alpha^2 - 4\lambda\gamma_2\rho}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)}, \\
\widetilde{\mathcal{M}}_{12} =& \widetilde{\mathcal{M}}_{21} \\
=& - \frac{\Big\{\lambda_1(\gamma_1\alpha - \gamma_2\rho) + \lambda_2\gamma_2\alpha\Big\}\Big\{\lambda_2(\gamma_1\alpha - \gamma_2\rho) + \lambda_3\gamma_2\alpha\Big\}}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \\
&+ \lambda_2\left(\frac{\gamma_1\alpha}{\gamma_2} + \frac{(\gamma_1\alpha - \gamma_2\rho)(\gamma_1\alpha - 2\lambda)}{2\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) + \lambda_3\left(\alpha + \frac{\alpha(\gamma_1\alpha - 2\lambda)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) \\
&+ \lambda_1\left(\frac{\gamma_1(\beta - \alpha)}{\gamma_2} + \frac{(\gamma_1\alpha - \gamma_2\rho)\gamma_1(\beta - \alpha)}{2\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) + \lambda_2\left(\beta - \alpha + \frac{\gamma_1\alpha(\beta - \alpha)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) \\
&- \frac{\gamma_1(\beta - \alpha)(\gamma_1\alpha - 2\lambda)}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)}, \\
\widetilde{\mathcal{M}}_{22} =& - \frac{\Big\{\lambda_2(\gamma_1\alpha - \gamma_2\rho) + \lambda_3\gamma_2\alpha\Big\}^2}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} + 2\lambda_2\left(\frac{\gamma_1(\beta - \alpha)}{\gamma_2} + \frac{(\gamma_1\alpha - \gamma_2\rho)\gamma_1(\beta - \alpha)}{2\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) \\
&+ 2\lambda_3\left(\beta - \alpha + \frac{\gamma_1\alpha(\beta - \alpha)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) - \frac{\gamma_1^2(\beta - \alpha)^2}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)}.
\end{aligned}$$

In terms of the functions

$$\begin{aligned}
f(\Lambda) :=& - \frac{\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta - \alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\Lambda^2 + 2\left(\frac{\gamma_1\alpha}{\gamma_2} + \frac{\gamma_1\alpha(\gamma_1\alpha - \gamma_2\rho)}{2\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)} + \beta - \alpha\right. \\
&\left. + \frac{\gamma_1\alpha(\beta - \alpha)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right)\Lambda + \frac{\gamma_1^2}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)}, \\
g(\Lambda) :=& \lambda\left(\frac{\gamma_1\alpha - \gamma_2\rho}{\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)} + \frac{\beta - \alpha}{\gamma_2\rho - \gamma_1\alpha + \lambda}\right)\Lambda,
\end{aligned}$$

the matrix $\widetilde{\mathcal{M}}$ can be written as

$$\begin{pmatrix} f(\Lambda)\alpha^2 - 2\alpha^2 g(\Lambda) + \frac{\lambda\gamma_2\rho}{\gamma_2\rho - \gamma_1\alpha + \lambda} & f(\Lambda)\alpha(\beta - \alpha) - \alpha(\beta - \alpha)g(\Lambda) + \frac{\lambda\gamma_1(\beta - \alpha)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \\ f(\Lambda)\alpha(\beta - \alpha) - \alpha(\beta - \alpha)g(\Lambda) + \frac{\lambda\gamma_1(\beta - \alpha)}{2(\gamma_2\rho - \gamma_1\alpha + \lambda)} & f(\Lambda)(\beta - \alpha)^2 \end{pmatrix}.$$

Straightforward calculations show that

$$\begin{aligned}
&\mathrm{Det}[\widetilde{\mathcal{M}}] \\
=& (\beta - \alpha)^2\left(f(\Lambda)\frac{\lambda\gamma_2\rho}{\gamma_2\rho - \gamma_1\alpha + \lambda} - \frac{\lambda^2\gamma_1^2}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)^2} - \alpha^2 g^2(\Lambda) - (f(\Lambda) - g(\Lambda))\frac{\lambda\gamma_1\alpha}{\gamma_2\rho - \gamma_1\alpha + \lambda}\right) \\
=& \frac{\lambda\gamma_2\rho(\beta - \alpha)^2}{\gamma_2\rho - \gamma_1\alpha + \lambda}\left(f(\Lambda) - \frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) + O(\alpha).
\end{aligned}$$

The proof of the positive semi-definiteness of $\widetilde{\mathcal{M}}$ is now split into the following two cases.

- **Case 1.** $\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha) \le 0$. In this case,

$$g(\Lambda) \le 0 \quad \text{for all} \quad \Lambda \ge 0 \tag{3.3}$$

and we put

$$h_1(\Lambda) := f(\Lambda) - \frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho - \gamma_1\alpha + \lambda)}.$$

Case 1.1. $\alpha(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)) = 0$**.** In this case either $\alpha = 0$ or $\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha) = 0$ and hence h_1 is linear with a positive leading coefficient. Thus, choosing $\Lambda = \Lambda_0 > 0$ large enough, $h_1(\Lambda_0) > 0$ and thus $f(\Lambda_0) > 0$ as well. Moreover,

$$\widetilde{\mathcal{M}}_{11} = f(\Lambda_0)\alpha^2 - 2g(\Lambda_0)\alpha^2 + \frac{\lambda\gamma_2\rho}{\gamma_2\rho - \gamma_1\alpha + \lambda} > 0, \quad \widetilde{\mathcal{M}}_{22} = f(\Lambda_0)(\beta-\alpha)^2 > 0$$

and

$$\mathrm{Det}[\widetilde{\mathcal{M}}] = \frac{\lambda\gamma_2\rho(\beta-\alpha)^2}{\gamma_2\rho - \gamma_1\alpha + \lambda}\left(f(\Lambda_0) - \frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) + O(\alpha) > 0$$

by choosing Λ_0 large enough. Hence, in this case our matrix is positive semi-definite.

Case 1.2. $\alpha(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)) < 0$**.** In this case, h_1 is a quadratic function with a maximum point $\Lambda_0 > 0$. If α is small enough, then the discriminant of h_1 is positive since

$$\begin{aligned}
&\left(2(\beta-\alpha) + \frac{\gamma_1\alpha\{\gamma_2\rho - \gamma_1\alpha + \gamma_2(\beta-\alpha) + 2\lambda\}}{\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right)^2 \\
&\quad - 4\frac{\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)} \cdot \left(\frac{\gamma_1^2}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)} + \frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) \\
&= \left(2(\beta-\alpha) + \frac{\gamma_1\alpha\{\gamma_2\rho - \gamma_1\alpha + \gamma_2(\beta-\alpha) + 2\lambda\}}{\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right)^2 \\
&\quad - \frac{\gamma_1^2\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)^2} - \frac{\lambda\gamma_1^2\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{\rho\gamma_2^3(\gamma_2\rho - \gamma_1\alpha + \lambda)^2} \\
&> 4(\beta-\alpha)^2 - o(\alpha) \\
&> 0.
\end{aligned}$$

As a result, $f(\Lambda_0) - \frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho - \gamma_1\alpha + \lambda)} > 0$, which implies that $f(\Lambda_0) > 0$. Thus,

$$\widetilde{\mathcal{M}}_{11} = f(\Lambda_0)\alpha^2 - 2g(\Lambda_0)\alpha^2 + \frac{\lambda\gamma_2\rho}{\gamma_2\rho - \gamma_1\alpha + \lambda} > 0, \quad \widetilde{\mathcal{M}}_{22} = f(\Lambda_0)(\beta-\alpha)^2 > 0$$

and

$$\mathrm{Det}[\widetilde{\mathcal{M}}] = \frac{\lambda\gamma_2\rho(\beta-\alpha)^2}{\gamma_2\rho - \gamma_1\alpha + \lambda}\left(f(\Lambda_0) - \frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right) + O(\alpha) > 0$$

for small α. Hence, in this case, too, the matrix $\widetilde{\mathcal{M}}$ is positive semidefinite.

- **Case 2:** $\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha) > 0$. In this case, we put

$$\begin{aligned}
h_2(\Lambda) &:= f(\Lambda) - 2g(\Lambda) \\
&= -\frac{\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\Lambda^2 \\
&\quad + \left(\frac{2\gamma_2(\gamma_2\rho - \gamma_1\alpha)(\beta-\alpha) + \gamma_1\alpha(\gamma_2\rho - \gamma_1\alpha) + \gamma_1\gamma_2\alpha(\beta-\alpha) + 2\gamma_2\rho\lambda}{\gamma_2(\gamma_2\rho - \gamma_1\alpha + \lambda)}\right)\Lambda \\
&\quad - \frac{\gamma_1^2}{4(\gamma_2\rho - \gamma_1\alpha + \lambda)}.
\end{aligned}$$

The discriminant of the quadratic function h_2 is positive since

$$
\begin{aligned}
&\left(\frac{2\gamma_2(\gamma_2\rho-\gamma_1\alpha)(\beta-\alpha)+\gamma_1\alpha(\gamma_2\rho-\gamma_1\alpha)+\gamma_1\gamma_2\alpha(\beta-\alpha)+2\gamma_2\rho\lambda}{\gamma_2(\gamma_2\rho-\gamma_1\alpha+\lambda)}\right)^2\\
&-4\frac{\alpha^2\Big(\gamma_1\alpha-\gamma_2\rho+\gamma_2(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho-\gamma_1\alpha+\lambda)}\cdot\frac{\gamma_1^2}{4(\gamma_2\rho-\gamma_1\alpha+\lambda)}\\
=&\left(\frac{2\gamma_2(\gamma_2\rho-\gamma_1\alpha)(\beta-\alpha)+\gamma_1\alpha(\gamma_2\rho-\gamma_1\alpha)+\gamma_1\gamma_2\alpha(\beta-\alpha)+2\gamma_2\rho\lambda}{\gamma_2(\gamma_2\rho-\gamma_1\alpha+\lambda)}\right)^2\\
&-\frac{\Big(\gamma_1\alpha(\gamma_2\rho-\gamma_1\alpha)+\gamma_1\gamma_2\alpha(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho-\gamma_1\alpha+\lambda)^2}>0.
\end{aligned}
$$

Let us denote the maximum point by Λ_1. Then $\Lambda_1>0$ and so $h_2(\Lambda_1)>0$. To show that the determinant of $\widetilde{M}$ is positive we first show that

$$f(\Lambda_1)-\frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho-\gamma_1\alpha+\lambda)}>0$$

for α small enough. Indeed,

$$
\begin{aligned}
&f(\Lambda_1)-\frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho-\gamma_1\alpha+\lambda)}\\
=&\frac{1}{4\alpha^2\Big(\gamma_1\alpha-\gamma_2\rho+\gamma_2(\beta-\alpha)\Big)^2(\gamma_2\rho-\gamma_1\alpha+\lambda)}\cdot\\
&\Bigg(\Big\{2\gamma_2(\beta-\alpha)(\gamma_2\rho-\gamma_1\alpha)+2\lambda(\gamma_1\alpha-\gamma_2\rho+\gamma_2(\beta-\alpha))+\gamma_1\alpha(\gamma_2\rho-\gamma_1\alpha+\gamma_2(\beta-\alpha))+2\gamma_2\rho\lambda\Big\}^2\\
&\quad-4\lambda^2(\gamma_1\alpha-\gamma_2\rho+\gamma_2(\beta-\alpha))^2-\alpha^2\gamma_1^2\Big(\gamma_1\alpha-\gamma_2\rho+\gamma_2(\beta-\alpha)\Big)^2\left(1+\frac{\lambda}{\gamma_2\rho}\right)\Bigg)\\
>&\frac{1}{4\alpha^2\Big(\gamma_1\alpha-\gamma_2\rho+\gamma_2(\beta-\alpha)\Big)^2(\gamma_2\rho-\gamma_1\alpha+\lambda)}\cdot\big(4\gamma_2^2\rho^2\lambda^2-o(\alpha)\big)>0
\end{aligned}
$$

for α small enough. In this case,

$$\widetilde{\mathcal{M}}_{11}=h_2(\Lambda_1)\alpha^2+\frac{\lambda\gamma_2\rho}{\gamma_2\rho-\gamma_1\alpha+\lambda}>0,\quad \widetilde{\mathcal{M}}_{22}=f(\Lambda_1)(\beta-\alpha)^2>0$$

and

$$\mathrm{Det}[\widetilde{\mathcal{M}}]=\frac{\lambda\gamma_2\rho(\beta-\alpha)^2}{\gamma_2\rho-\gamma_1\alpha+\lambda}\left(f(\Lambda_1)-\frac{\lambda\gamma_1^2}{4\gamma_2\rho(\gamma_2\rho-\gamma_1\alpha+\lambda)}\right)+O(\alpha)>0,$$

and so $\widetilde{\mathcal{M}}$ is positive semidefinite.

In conclusion, we can always find a constant $\Lambda>0$ such that $\widetilde{\mathcal{M}}$ is positive semidefinite. Since the terminal value

$$\begin{pmatrix}\frac{\gamma_2}{2}+\Lambda\alpha^2 & \Lambda\alpha(\beta-\alpha)\\ \Lambda\alpha(\beta-\alpha) & \Lambda(\beta-\alpha)^2\end{pmatrix}$$

is positive definite, all the coefficients in (3.2) satisfy the requirements in [32, Proposition 2.1, 2.2]. As a result, the system (3.2) has a unique solution in $L^\infty([0,T];\mathbb{S}^2)\cap C([0,T];\mathbb{S}^2)$. □

Remark 3.2. In the case of risk-neutral investors, i.e. if $\lambda=0$, the assumption that α is small enough can be dropped. Indeed, in the risk-neutral case, the matrix $\widetilde{\mathcal{M}}$ can be written as

$$\begin{pmatrix}f(\Lambda)\alpha^2 & f(\Lambda)\alpha(\beta-\alpha)\\ f(\Lambda)\alpha(\beta-\alpha) & f(\Lambda)(\beta-\alpha)^2\end{pmatrix},$$

where

$$f(\Lambda) := -\frac{\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{2\gamma_2^2(\gamma_2\rho - \gamma_1\alpha)}\Lambda^2 + \left(2(\beta-\alpha) + \frac{\gamma_1\alpha}{\gamma_2} + \frac{\gamma_1\alpha(\beta-\alpha)}{(\gamma_2\rho-\gamma_1\alpha)}\right)\Lambda - \frac{\gamma_1^2}{2(\gamma_2\rho-\gamma_1\alpha)}.$$

Since the determinant of $\widetilde{\mathcal{M}}$ is zero, it is sufficient to prove that $f(\Lambda) > 0$, for some Λ. This can be seen as follows. If $\alpha(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)) = 0$, then $f(\Lambda) > 0$ by choosing a Λ large enough. If $\alpha(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)) \neq 0$, then the maximum point of the quadratic function f is strictly positive and the discriminant of f is positive since

$$\begin{aligned}
&\left(2(\beta-\alpha) + \frac{\gamma_1\alpha}{\gamma_2} + \frac{\gamma_1\alpha(\beta-\alpha)}{(\gamma_2\rho-\gamma_1\alpha)}\right)^2 - 4\frac{\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{2\gamma_2^2(\gamma_2\rho-\gamma_1\alpha)}\cdot\frac{\gamma_1^2}{2(\gamma_2\rho-\gamma_1\alpha)}\\
&= \left(2(\beta-\alpha) + \frac{\gamma_1\alpha\Big(\gamma_2\rho - \gamma_1\alpha + \gamma_2(\beta-\alpha)\Big)}{\gamma_2(\gamma_2\rho-\gamma_1\alpha)}\right)^2 - \frac{\gamma_1^2\alpha^2\Big(\gamma_1\alpha - \gamma_2\rho + \gamma_2(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho-\gamma_1\alpha)^2}\\
&> \left(2(\beta-\alpha) + \frac{\gamma_1^2\alpha^2\Big(\gamma_2\rho - \gamma_1\alpha + \gamma_2(\beta-\alpha)\Big)^2}{\gamma_2(\gamma_2\rho-\gamma_1\alpha)}\right) - \frac{\gamma_1^2\alpha^2\Big(\gamma_2\rho - \gamma_1\alpha + \gamma_2(\beta-\alpha)\Big)^2}{\gamma_2^2(\gamma_2\rho-\gamma_1\alpha)^2}\\
&> 0.
\end{aligned}$$

Hence there exists a $\Lambda > 0$ such that $f(\Lambda) > 0$.

4 The verification theorem

To verify that the strategy given by equations (2.13) and (2.14) is indeed optimal we first prove that the cost functional can be written as a sum of two complete square terms and a correction term that we will identify as the value function.

Proposition 4.1. *Let the* ***standing assumption*** *hold. Then the cost functional can be rewritten as*

$$\begin{aligned}
J(t,Z) = \mathbb{E}\Big[\int_t^T \frac{1}{\tilde{a}}\left(I_s^A(\mathcal{X}_s - \bar{\mu}_s)\right)^2 ds + \int_t^T \frac{1}{a}\left(I_s^B\bar{\mu}_s + I_s^D\right)^2 ds\Big]\\
+ \mathrm{Var}(\mu_{t-})(A_t) + \bar{\mu}_{t-}^\top B_t\bar{\mu}_{t-} + D_t^\top\bar{\mu}_{t-} + \mathbb{E}[F_t],
\end{aligned}$$

where $\mu_\cdot$ is the law of $\mathcal{X}_\cdot$. In particular, the cost functional reaches its global minimum if

$$\int_t^T \left(I_s^A(\mathcal{X}_s - \bar{\mu}_s)\right)^2 ds = \int_t^T \left(I_s^B\bar{\mu}_s + I_s^D\right)^2 ds = 0, \quad a.s.$$

In this case, the value function is indeed given by (2.12)*.*

Proof. For any strategy $Z \in \mathscr{A}_t$, we first separate the cost of the jump at the terminal time from the cost functional. To this end, we write the cost functional as

$$
\begin{aligned}
J(t,Z) &= \mathbb{E}\left[\int_{[t,T)} \Big(Y_{s-}\,dZ_s + \frac{\gamma_2}{2}\,d[Z]_s + \sigma_s\,d[Z,W]_s\Big) + \int_t^T \lambda X_s^2\,ds\right] \\
&\quad + \mathbb{E}\left[-\Big(Y_{T-} - \frac{\gamma_2}{2}\Delta X_T\Big)\Delta X_T\right] \\
&= \mathbb{E}\left[\int_{[t,T)} \Big(Y_{s-}\,dZ_s + \frac{\gamma_2}{2}\,d[Z]_s + \sigma_s\,d[Z,W]_s\Big) + \int_t^T \lambda X_s^2\,ds\right] \\
&\quad + \mathbb{E}\left[\frac{\gamma_2}{2}X_{T-}^2 + X_{T-}Y_{T-}\right] \quad (\text{ since } X_T = 0) \\
&= \mathbb{E}\left[\int_{[t,T)} \Big(Y_{s-}\,dZ_s + \frac{\gamma_2}{2}\,d[Z]_s + \sigma_s\,d[Z,W]_s\Big) + \int_t^T \lambda X_s^2\,ds\right] \\
&\quad + \mathbb{E}\left[(\mathcal{X}_{T-} - \bar{\mu}_{T-})^\top A_T(\mathcal{X}_{T-} - \bar{\mu}_{T-}) + \bar{\mu}_{T-}^\top B_T\bar{\mu}_{T-} + D_T^\top\bar{\mu}_{T-} + F_T\right].
\end{aligned} \tag{4.1}
$$

Next, we are going to analyze the expected jump cost term-by-term. From this we will see that many terms cancel and then arrive at the desired representation of the cost functional.

- We start with the term $\mathbb{E}\left[\mathcal{X}_{T-}A_T\mathcal{X}_{T-}\right]$. Using Itô's formula in [36, Theorem 36],

$$
\begin{aligned}
\int_{\mathbb{R}^3} x^\top A_T x\mu_{T-}(dx) = \mathbb{E}\Big[&\int_t^{T-} 2(A_s\mathcal{X}_{s-})^\top d\mathcal{X}_s + Tr(A_s d[\mathcal{X},\mathcal{X}]_s^c) \\
&+ \sum_{t\le s<T} \big(\mathcal{X}_s^\top A_s\mathcal{X}_s - \mathcal{X}_{s-}^\top A_s\mathcal{X}_{s-} - 2(A_s\mathcal{X}_{s-})^\top\Delta\mathcal{X}_s\big)\Big] \\
&+ \int_{\mathbb{R}^3} x^\top A_t x\mu_{t-}(dx) + \int_t^T\int_{\mathbb{R}^3} x^\top \dot{A}_s x\mu_s(dx)\,ds.
\end{aligned}
$$

Note that

$$
d[\mathcal{X},\mathcal{X}]_s^c = \begin{pmatrix} d[Z^c,Z^c]_s & -\sigma_s\,d[Z^c,W]_s + \gamma_2 d[Z^c,Z^c]_s & 0 \\ -\sigma_s d[Z^c,W]_s + \gamma_s d[Z^c,Z^c]_s & \sigma_s^2\,ds + \gamma_2^2\,d[Z^c,Z^c]_s + 2\gamma_2\sigma_s\,d[Z^c,W]_s & 0 \\ 0 & 0 & 0 \end{pmatrix},
$$

which implies by the relationship between the entries of the matrix A (cf. the statement above (2.6)) that

$$
\int_t^{T-} \mathrm{Tr}(A_s\,d[\mathcal{X}^c,\mathcal{X}^c]_s)\,ds = \int_t^T \Big(-\frac{\gamma_2}{2}\,d[Z^c,Z^c]_s - \sigma_s\,d[Z^c,W]_s + \sigma_s^2 A_{22,s}\,ds\Big).
$$

Taking this back into the above equation shows that

$$
\begin{aligned}
&\int_{\mathbb{R}^3} x^\top A_T x\mu_{T-}(dx) \\
&= \mathbb{E}\left[\int_t^T 2(A_s\mathcal{X}_s)^\top\big(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G}\big)\,ds + \int_t^{T-} 2(A_s\mathcal{X}_{s-})^\top\mathcal{K}\,dZ_s\right] \\
&\quad + \mathbb{E}\left[\int_t^T \Big(-\frac{\gamma_2}{2}\,d[Z^c,Z^c]_s - \sigma_s\,d[Z^c,W]_s + \sigma_s^2 A_{22,s}\,ds\Big)\right] \\
&\quad + \mathbb{E}\left[\sum_{t\le s<T}\big(\mathcal{X}_s^\top A_s\mathcal{X}_s - \mathcal{X}_{s-}^\top A_s\mathcal{X}_{s-} - 2(A_s\mathcal{X}_{s-})^\top\mathcal{K}\Delta Z_s\big)\right] \\
&\quad + \int_{\mathbb{R}^3} x^\top A_t x\mu_{t-}(dx) + \mathbb{E}\left[\int_t^T \mathcal{X}_s^\top\dot{A}_s\mathcal{X}_s\,ds\right].
\end{aligned} \tag{4.2}
$$

- Next, we consider the term $\bar{\mu}_{T-}^\top A_T \bar{\mu}_{T-}$. In view of (2.10) the expected value $\overline{\mu}$ follows the dynamics

$$d\overline{\mu}_s = \Big((\mathcal{H} + \overline{\mathcal{H}})\overline{\mu}_s + \mathcal{G} \Big)\, ds + \mathcal{K}\, d\mathbb{E}[Z_s]. \tag{4.3}$$

Applying the chain rule to $\overline{\mu}^\top A \overline{\mu}$ from $t-$ to $T-$, it follows that

$$\begin{aligned} \bar{\mu}_{T-}^\top A_T \bar{\mu}_{T-} = &\int_t^{T-} 2(A_s \bar{\mu}_{s-})^\top d\overline{\mu}_s + \sum_{t\le s<T} \big(\bar{\mu}_s^\top A_s \bar{\mu}_s - \bar{\mu}_{s-}^\top A_s \bar{\mu}_{s-} - 2(A_s \bar{\mu}_{s-})^\top \Delta \overline{\mu}_s \big) \\ &+ \bar{\mu}_{t-}^\top A_t \bar{\mu}_{t-} + \int_t^T \bar{\mu}_s^\top \dot{A}_s \bar{\mu}_s\, ds. \end{aligned}$$

Taking (4.3) into the expression of $\overline{\mu}_{T-}^\top A_T \overline{\mu}_{T-}$ we arrive at

$$\begin{aligned} \overline{\mu}_{T-}^\top A_T \overline{\mu}_{T-} = &\bar{\mu}_{t-}^\top A_t \bar{\mu}_{t-} + \int_t^T 2(A_s \bar{\mu}_s)^\top \big(\mathcal{H}\overline{\mu}_s + \overline{\mathcal{H}}\overline{\mu}_s + \mathcal{G} \big)\, ds + \int_t^{T-} 2(A_s \bar{\mu}_{s-})^\top \mathcal{K}\, d\mathbb{E}[Z_s] \\ &+ \sum_{t\le s<T} \Big(\bar{\mu}_s^\top A_s \bar{\mu}_s - \bar{\mu}_{s-}^\top A_s \bar{\mu}_{s-} - 2(A_s \bar{\mu}_{s-})^\top \Delta \overline{\mu}_s \Big) + \int_t^T \bar{\mu}_s^\top \dot{A}_s \bar{\mu}_s\, ds. \end{aligned} \tag{4.4}$$

- Similarly to the last step it holds that

$$\begin{aligned} \overline{\mu}_{T-}^\top B_T \overline{\mu}_{T-} = &\bar{\mu}_{t-}^\top B_t \bar{\mu}_{t-} + \int_t^T 2(B_s \bar{\mu}_s)^\top \big(\mathcal{H}\overline{\mu}_s + \overline{\mathcal{H}}\overline{\mu}_s + \mathcal{G} \big)\, ds + \int_t^{T-} 2(B_s \bar{\mu}_{s-})^\top \mathcal{K}\, d\mathbb{E}[Z_s] \\ &+ \sum_{t\le s<T} \Big(\bar{\mu}_s^\top B_s \bar{\mu}_s - \bar{\mu}_{s-}^\top B_s \bar{\mu}_{s-} - 2(B_s \bar{\mu}_{s-})^\top \Delta \overline{\mu}_s \Big) + \int_t^T \bar{\mu}_s^\top \dot{B}_s \bar{\mu}_s\, ds. \end{aligned} \tag{4.5}$$

- Applying the chain rule to $D^\top \overline{\mu}$, we see that

$$\begin{aligned} D_T^\top \bar{\mu}_{T-} &= \mathbb{E}\Big[\int_t^T D_s^\top \big(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G} \big)\, ds + \int_t^{T-} D_s^\top \mathcal{K}\, dZ_s \\ &\quad + \sum_{t\le s<T} \big((D_s^\top \mathcal{X}_s - D_s^\top \mathcal{X}_{s-}) - D_s^\top \Delta \mathcal{X}_s \big) \Big] + D_t^\top \bar{\mu}_{t-} + \int_t^T \dot{D}_s^\top \bar{\mu}_s\, ds \\ &= \mathbb{E}\left[\int_t^T D_s^\top \big(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G} \big)\, ds + \int_t^{T-} D_s^\top \mathcal{K}\, dZ_s + D_t^\top \bar{\mu}_{t-} + \int_t^T \dot{D}_s^\top \bar{\mu}_s\, ds \right]. \end{aligned} \tag{4.6}$$

Next, we collect all the terms in (4.2), (4.4), (4.5) and (4.6) involving jumps. Their sum equals

$$\begin{aligned} &\mathbb{E}\Big[\int_t^{T-} 2(A_s \mathcal{X}_{s-})^\top \mathcal{K}\, dZ_s + \sum_{t\le s<T} \big(\mathcal{X}_s^\top A_s \mathcal{X}_s - \mathcal{X}_{s-}^\top A_s \mathcal{X}_{s-} - 2(A_s \mathcal{X}_{s-})^\top \mathcal{K} \Delta Z_s \big) \Big] \\ &- \mathbb{E}\Big[\int_t^{T-} 2(A_s \bar{\mu}_{s-})^\top \mathcal{K}\, dZ_s + \sum_{t\le s<T} \big(\bar{\mu}_s^\top A_s \bar{\mu}_s - \bar{\mu}_{s-}^\top A_s \bar{\mu}_{s-} - 2(A_s \bar{\mu}_{s-})^\top \Delta \mathcal{X}_s \big) \Big] \\ &+ \mathbb{E}\Big[\int_t^{T-} 2(B_s \bar{\mu}_{s-})^\top \mathcal{K}\, dZ_s + \sum_{t\le s<T} \big(\bar{\mu}_s^\top B_s \bar{\mu}_s - \bar{\mu}_{s-}^\top B_s \bar{\mu}_{s-} - 2(B_s \bar{\mu}_{s-})^\top \Delta \mathcal{X}_s \big) \Big] \\ &+ \mathbb{E}\left[\int_t^{T-} D_s^\top \mathcal{K}\, dZ_s \right]. \end{aligned} \tag{4.7}$$

Since

$$\mathcal{K}^\top A = \begin{pmatrix} -1 & \gamma_2 & 0 \end{pmatrix} A = \begin{pmatrix} 0 & -\tfrac{1}{2} & 0 \end{pmatrix} = \mathcal{K}^\top B, \tag{4.8}$$

we obtain that

$$-\mathbb{E}\Big[\int_t^{T-} 2(A_s\bar{\mu}_{s-})^\top \mathcal{K}\, dZ_s\Big] + \mathbb{E}\Big[\int_t^{T-} 2(B_s\bar{\mu}_{s-})^\top \mathcal{K}\, dZ_s\Big] = 0.$$

Since $\Delta\bar{\mu} = \mathbb{E}[\mathcal{K}\Delta Z]$ we also obtain that

$$\begin{aligned}
\mathbb{E}\left[(\bar{\mu}_s^\top B_s\bar{\mu}_s - \bar{\mu}_{s-}^\top B_s\bar{\mu}_{s-}) - 2(B_s\bar{\mu}_{s-})^\top \Delta\mathcal{X}_s\right] &= \Delta\bar{\mu}_s^\top B_s \Delta\bar{\mu}_s \\
&= \mathbb{E}[\Delta Z_s \mathcal{K}^\top B_s \Delta\bar{\mu}_s] \\
&= \mathbb{E}[\Delta Z_s \mathcal{K}^\top A_s \Delta\bar{\mu}_s] \\
&= \Delta\bar{\mu}_s^\top A_s \Delta\bar{\mu}_s \\
&= \mathbb{E}\left[(\bar{\mu}_s^\top A_s\bar{\mu}_s - \bar{\mu}_{s-}^\top A_s\bar{\mu}_{s-}) - 2(A_s\bar{\mu}_{s-})^\top \Delta\mathcal{X}_s\right].
\end{aligned}$$

Using (4.8) again, we have that

$$\begin{aligned}
2(A_s\mathcal{X}_{s-})^\top \mathcal{K}\, dZ_s &= 2\mathcal{K}^\top A_s \mathcal{X}_{s-}\, dZ_s \\
&= \begin{pmatrix} 0 & -1 & 0 \end{pmatrix} \mathcal{X}_{s-}\, dZ_s \\
&= -Y_{s-}\, dZ_s.
\end{aligned}$$

Moreover, the definition of $\mathcal{K}$ implies that

$$\begin{aligned}
\mathcal{X}_s^\top A_s \mathcal{X}_s - \mathcal{X}_{s-}^\top A_s \mathcal{X}_{s-} - 2(A_s\mathcal{X}_{s-})^\top \mathcal{K}\Delta Z_s &= \Delta\mathcal{X}_s^\top A_s \Delta\mathcal{X}_s \\
&= \Delta Z_s \mathcal{K}^\top A_s \Delta\mathcal{X}_s \\
&= -\frac{1}{2}\Delta Z_s \Delta Y_s \\
&= -\frac{\gamma_2}{2}(\Delta Z_s)^2,
\end{aligned}$$

and that

$$D^\top \mathcal{K} = \begin{pmatrix} D_1 & D_2 & D_3 \end{pmatrix} \begin{pmatrix} -1 \\ \gamma_2 \\ 0 \end{pmatrix} = -D_1 + \gamma_2 D_2 = 0.$$

As a result, the jump terms (4.7) together with the term $\mathbb{E}\left[\int_t^T \left(-\frac{\gamma_2}{2}\, d[Z^c, Z^c]_s - \sigma_s\, d[Z^c, W]_s\right)\right]$ in (4.2) cancel with the first three terms in (4.1). Hence, taking (4.2), (4.4), (4.5) and (4.6) into (4.1) yields that

$$\begin{aligned}
&J(t, Z) \\
&= \mathbb{E}\left[\int_t^T 2(A_s\mathcal{X}_s)^\top \left(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G}\right) ds + \int_t^T \mathcal{D}_s^\top A_s \mathcal{D}_s\, ds + \int_t^T \mathcal{X}_s^\top \dot{A}_s \mathcal{X}_s\, ds + \int_t^T \mathcal{X}_s^\top \mathcal{Q}\mathcal{X}_s\, ds\right] \\
&\quad - \mathbb{E}\left[\int_t^T 2(A_s\bar{\mu}_s)^\top \left(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G}\right) ds\right] - \int_t^T \bar{\mu}_s^\top \dot{A}_s \bar{\mu}_s\, ds \\
&\quad + \mathbb{E}\left[\int_t^T 2(B_s\bar{\mu}_s)^\top \left(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G}\right) ds\right] + \int_t^T \bar{\mu}_s^\top \dot{B}_s \bar{\mu}_s\, ds \\
&\quad + \mathbb{E}\left[\int_t^T D_s^\top \left(\mathcal{H}\mathcal{X}_s + \overline{\mathcal{H}}\mathbb{E}[\mathcal{X}_s] + \mathcal{G}\right) ds\right] + \int_t^T \dot{D}_s^\top \bar{\mu}_s\, ds + \mathbb{E}\left[\int_t^T \dot{F}_s\, ds\right] \\
&\quad + \operatorname{Var}(\mu_{t-})(A_t) + \bar{\mu}_{t-}^\top B_t \bar{\mu}_{t-} + D_t^\top \bar{\mu}_{t-} + \mathbb{E}[F_t],
\end{aligned}$$

where

$$\mathcal{Q} := \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Recalling that $\overline{\mu} = \mathbb{E}[\mathcal{X}]$ and collecting the terms $\mathcal{X}^\top(\cdots)\mathcal{X}$, $\overline{\mu}^\top(\cdots)\overline{\mu}$, $(\cdots)\overline{\mu}$ and other terms, we have that

$$\begin{aligned}
&J(t,Z)\\
=\ &\mathbb{E}\left[\int_t^T \mathcal{X}_s^\top\left(\mathcal{Q}+2\mathcal{H}^\top A_s+\dot{A}_s\right)\mathcal{X}_s\,ds+\int_t^T \overline{\mu}_s^\top\left(-\dot{A}_s-2\mathcal{H}^\top A_s+\dot{B}_s+2\mathcal{H}^\top B_s+2\overline{\mathcal{H}}^\top B_s\right)\overline{\mu}_s\,ds\right.\\
&\left.+\int_t^T\left(2\mathcal{G}^\top B_s+\dot{D}_s^\top+D_s^\top\mathcal{H}+D_s^\top\overline{\mathcal{H}}\right)\overline{\mu}_s\,ds+\int_t^T\left(\mathcal{D}_s^\top A_s\mathcal{D}_s+D_s^\top\mathcal{G}+\dot{F}_s\right)ds\right]\\
&+\operatorname{Var}(\mu_{t-})(A_t)+\overline{\mu}_{t-}^\top B_t\overline{\mu}_{t-}+D_t^\top\overline{\mu}_{t-}+\mathbb{E}[F_t]\\
=\ &\mathbb{E}\left[\int_t^T(\mathcal{X}_s-\overline{\mu}_s)^\top\left(\mathcal{Q}+\mathcal{H}^\top A_s+A_s\mathcal{H}+\dot{A}_s\right)(\mathcal{X}_s-\overline{\mu}_s)\,ds\right.\\
&+\int_t^T\overline{\mu}_s^\top\left(\mathcal{Q}+\dot{B}_s+\mathcal{H}^\top B_s+B_s\mathcal{H}+\overline{\mathcal{H}}^\top B_s+B_s\overline{\mathcal{H}}\right)\overline{\mu}_s\,ds\\
&\left.+\int_t^T\left(2\mathcal{G}^\top B_s+\dot{D}_s^\top+D_s^\top\mathcal{H}+D_s^\top\overline{\mathcal{H}}\right)\overline{\mu}_s\,ds+\int_t^T\left(\mathcal{D}_s^\top A_s\mathcal{D}_s+D_s^\top\mathcal{G}+\dot{F}_s\right)ds\right]\\
&+\operatorname{Var}(\mu_{t-})(A_t)+\overline{\mu}_{t-}^\top B_t\overline{\mu}_{t-}+D_t^\top\overline{\mu}_{t-}+\mathbb{E}[F_t].
\end{aligned}$$

By (A.15), (A.16), (A.17) and (A.18) in the appendix it holds that

$$\begin{aligned}
\frac{(I^A)^\top I^A}{\tilde{a}} &= \mathcal{Q}+\mathcal{H}^\top A+A\mathcal{H}+\dot{A}\\
\frac{(I^B)^\top I^B}{a} &= \mathcal{Q}+\dot{B}+\mathcal{H}^\top B+B\mathcal{H}+\overline{\mathcal{H}}^\top B+B\overline{\mathcal{H}}\\
\frac{2I^D I^B}{a} &= 2\mathcal{G}^\top B_s+\dot{D}_s^\top+D_s^\top\mathcal{H}+D_s^\top\overline{\mathcal{H}}\\
\frac{(I^D)^2}{a} &= \mathcal{D}^\top A_s\mathcal{D}+D^\top\mathcal{G}+\dot{F}
\end{aligned}$$

which gives us the desired result. □

Let $\widetilde{\mathcal{X}}$ be the state process driven by the strategy $\widetilde{Z}$ given by (2.13) and (2.14). The next theorem shows that $\widetilde{\mathcal{X}}$ satisfies the equality in Proposition 4.1 thereby concluding the verification argument.

Theorem 4.2. *The state process* $\widetilde{\mathcal{X}}$ *driven by the strategy* (2.13) *and* (2.14) *satisfies*

$$I^A(\widetilde{\mathcal{X}}_s-\mathbb{E}[\widetilde{\mathcal{X}}_s])=0,\quad I^B\mathbb{E}[\widetilde{\mathcal{X}}_s]+I^D=0,\qquad s\in[t,T). \tag{4.9}$$

Proof. We first prove that (4.9) holds at $s=t$. From the definition of $\Delta\widetilde{Z}_t$ in (2.13), we know that

$$\mathbb{E}[\Delta\widetilde{Z}_t]=-\frac{I^B}{a}\mathbb{E}[\mathcal{X}]-\frac{I^D}{a},\quad \Delta\widetilde{Z}_t=-\frac{I^A}{\tilde{a}}(\mathcal{X}-\mathbb{E}[\mathcal{X}])+\mathbb{E}[\Delta\widetilde{Z}_t],$$

where we recall that $\mathcal{X}:=\mathcal{X}_{t-}$. Since $I^A\mathcal{K}=\tilde{a}$ and $I^B\mathcal{K}=a$, we have that

$$I_t^B\mathbb{E}[\widetilde{\mathcal{X}}_t]+I_t^D=I_t^B\mathbb{E}[\mathcal{X}+\mathcal{K}\Delta\widetilde{Z}_t]+I_t^D=I_t^B\mathbb{E}[\mathcal{X}]+a\mathbb{E}[\Delta\widetilde{Z}_t]+I_t^D=0$$

and

$$I_t^A(\widetilde{\mathcal{X}}_t-\mathbb{E}[\widetilde{\mathcal{X}}_t])=I_t^A(\mathcal{X}+\mathcal{K}\Delta\widetilde{Z}_t-\mathbb{E}[\mathcal{X}+\mathcal{K}\Delta\widetilde{Z}_t])=I_t^A(\mathcal{X}-\mathbb{E}[\mathcal{X}])+\tilde{a}(\Delta\widetilde{Z}_t-\mathbb{E}[\Delta\widetilde{Z}_t])=0.$$

Next, we prove that $d\Big\{I_s^A(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s])\Big\} = d\Big\{I_s^B\mathbb{E}[\widetilde{\mathcal{X}}_s] + I_s^D\Big\} = 0$ for all $s \in [t,T)$ from which (4.9) follows. In fact, the state dynamics gives us that

$$d\mathbb{E}[\widetilde{\mathcal{X}}_s] = \left((\mathcal{H}+\overline{\mathcal{H}})\mathbb{E}[\widetilde{\mathcal{X}}_s] + \mathcal{G} - \mathcal{K}\frac{\dot{I}_s^B}{a}\mathbb{E}[\widetilde{\mathcal{X}}_s] - \mathcal{K}\frac{\dot{I}_s^D}{a} - \mathcal{K}\frac{I_s^B}{a}(\mathcal{H}+\overline{\mathcal{H}})\mathbb{E}[\widetilde{\mathcal{X}}_s]) - \mathcal{K}\frac{I_s^B}{a}\mathcal{G}\right) ds,$$

and

$$d(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s]) = \left(\mathcal{H}(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s]) - \mathcal{K}\frac{\dot{I}_s^A}{\tilde{a}}(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s]) - \mathcal{K}\frac{I_s^A}{\tilde{a}}\mathcal{H}(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s])\right) ds - \mathcal{K}\frac{I_s^A}{\tilde{a}}\mathcal{D}_s\, dW_s + \mathcal{D}_s\, dW_s.$$

Hence, the desired result follows from the following equalities:

$$\begin{aligned}
d\Big\{I_s^B\mathbb{E}[\widetilde{\mathcal{X}}_s] + I_s^D\Big\} &= \dot{I}_s^B\mathbb{E}[\widetilde{\mathcal{X}}_s]\, ds + \dot{I}_s^D\, ds + I_s^B d\mathbb{E}[\widetilde{\mathcal{X}}_s] \\
&= \left(\dot{I}_s^B + I_s^B\left((\mathcal{H}+\overline{\mathcal{H}}) - \mathcal{K}\frac{\dot{I}_s^B}{a} - \mathcal{K}\frac{I_s^B}{a}(\mathcal{H}+\overline{\mathcal{H}})\right)\right)\mathbb{E}[\widetilde{\mathcal{X}}_s]\, ds \\
&\quad + \dot{I}_s^D\, ds + I_s^B\left(\mathcal{G} - \mathcal{K}\frac{I_s^B}{a}\mathcal{G} - \mathcal{K}\frac{\dot{I}_s^D}{a}\right) ds \\
&= \Big(\dot{I}_s^B + I_s^B(\mathcal{H}+\overline{\mathcal{H}}) - \dot{I}_s^B - I_s^B(\mathcal{H}+\overline{\mathcal{H}})\Big)\mathbb{E}[\widetilde{\mathcal{X}}_s]\, ds \\
&\quad + (\dot{I}_s^D + I_s^B\mathcal{G} - I_s^B\mathcal{G} - \dot{I}_s^D)\, ds \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
d\Big\{I_s^A(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s])\Big\} &= dI_s^A(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s]) + I_s^A d(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s]) \\
&= \left(\dot{I}_s^A + I_s^A\left(\mathcal{H} - \mathcal{K}\frac{\dot{I}_s^A}{\tilde{a}} - \mathcal{K}\frac{I_s^A}{\tilde{a}}\mathcal{H}\right)\right)(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s])\, ds + I_s^A\left(-\mathcal{K}\frac{I_s^A}{\tilde{a}}\mathcal{D}_s + \mathcal{D}_s\right) dW_s \\
&= \Big(\dot{I}_s^A + I_s^A\mathcal{H} - \dot{I}_s^A - I_s^A\mathcal{H}\Big)(\widetilde{\mathcal{X}}_s - \mathbb{E}[\widetilde{\mathcal{X}}_s])\, ds + \big(-I_s^A\mathcal{D}_s + I_s^A\mathcal{D}_s\big)\, dW_s \\
&= 0.
\end{aligned}$$

□

5 Numerical simulations

This section provides numerical simulations that illustrate the dependence of the optimal inventory process on various model parameters. In all cases $x = 1$, $c = y = 0$, $\sigma = 0.8$ and $T = 1$. All trajectories were generated from the same Brownian path to guarantee that the trajectories are comparable. In addition to our optimal solution we display the optimal solution in the Obizhaeva-Wang model [34], which is the canonical reference point for our model. Setting $\alpha = \beta = \sigma = 0$ our model reduces to the Obizhaeva-Wang model.

Figure 1 displays the optimal position for two extreme choice of the market risk parameter. When the investor is highly risk averse the optimal holding in our models is relatively close to the one in the Obizhaeva-Wang model with added twist that in our model the investor may take short positions generating additional sell child order flow and buy the stock back while benefitting from the additional sell order flow. Overselling own positions should not be viewed as a fraudulent attempt to manipulate prices. Instead, the large investor rationally anticipates his/her impact on future order flow when making

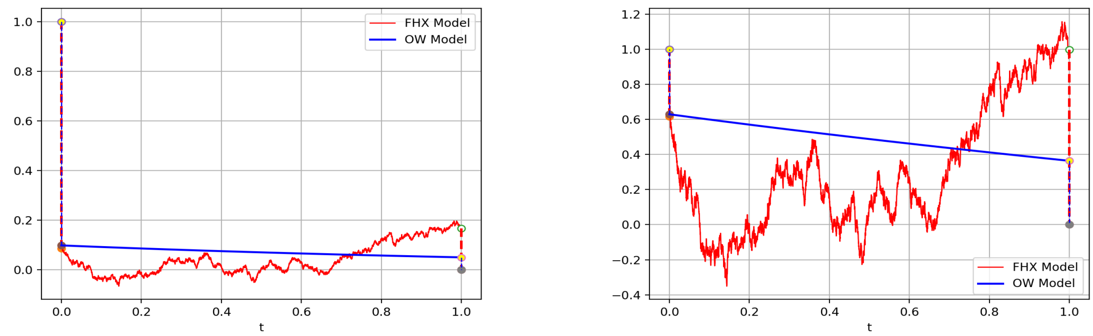

Figure 1: Dependence of the optimal position on the risk parameter λ for $\lambda = 1.5$ (left) and $\lambda = 0$(right). Other parameters are chosen as $\rho = 0.7$, $\gamma_1 = 0.1$, $\gamma_2 = 0.5$, $\alpha = 0.5$, $\beta = 1.1$.

trading decisions and uses it to his/her advantage. Similar effects have previously been observed in the literature; see [23] and references therein for a more detailed discussion of different manipulation strategies in portfolio liquidation models.

When the investor is risk-neutral, then the variations in the optimal inventory process are much larger; portfolio holdings range from about -0.4 to 1.2. This, too is very intuitive. Large portfolio holdings are much "cheaper" for a risk-neutral than a risk-averse trader.

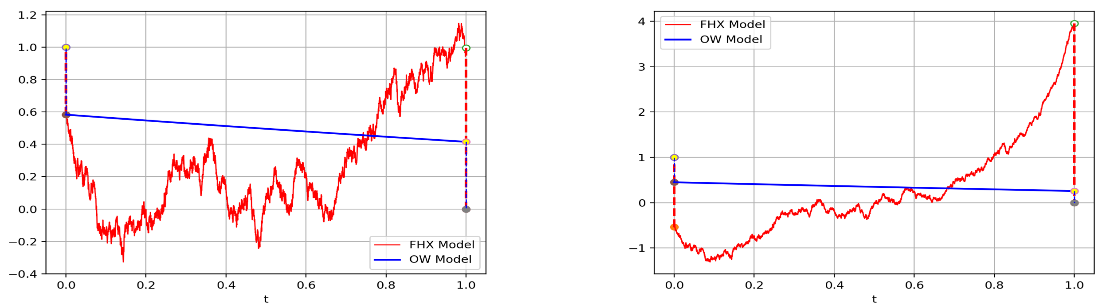

Figure 2: Dependence of the optimal position on the impact parameter α for $\alpha = 0$ (left) and $\alpha = 1.8$ (right). Other parameters are chosen as $\rho = 0.4$, $\lambda = 0$, $\gamma_1 = 0.1$, $\gamma_2 = 0.5$, $\beta = 3$.

Figure 2 displays the optimal position for varying degrees of child order flow. Even in the absence of any feedback effect ($\alpha = 0$) we see that it may be optimal to take short positions, that is, to drive the benchmark price down and then to close the position submitting a large order at a favorable price at the end of the trading period. This effect is much stronger in the presence of child order flow where the price decrease due to own selling may be very strong and may well outweigh the cost of block trade at the end of the trading period. Optimal positions for different transient market impact parameters are shown in Figure 3.

6 Conclusion

We considered a novel mean-field control problem with semimartingale strategies. We obtained a candidate value function by passing to the limit from a sequence of discrete time models. The value function can be described in terms of the solution to a fully coupled system of Riccati equations. A

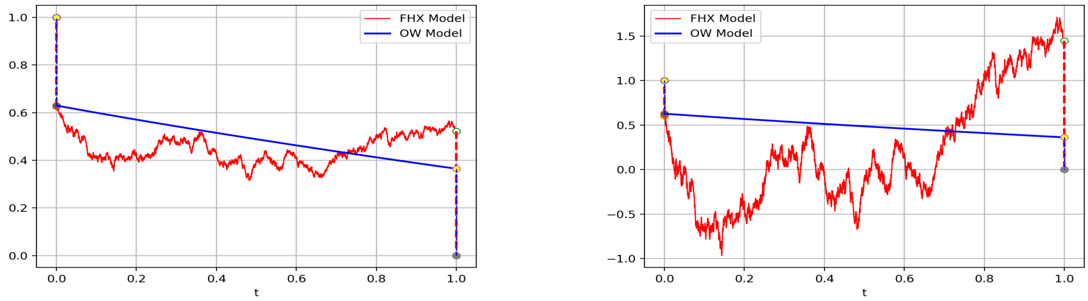

Figure 3: Dependence of the optimal position on the impact parameter γ_2 for $\gamma_2 = 2$ (left) and $\gamma_2 = 0.3$ (right). Other parameters are chosen as $\rho = 0.7$, $\lambda = 0$, $\gamma_1 = 0.1$, $\alpha = 0.5$, $\beta = 1.1$.

sophisticated transformation shows that the system has a unique solution and that the candidate optimal strategy is indeed optimal.

Several avenues are open for future research. Let us just mention two. First, except the volatility of the spread all cost coefficients in our model are deterministic constants. Although there are many liquidation models where similar assumptions are made, these assumptions seem restrictive from a mathematical perspective. However, as far as we can tell there is no obvious way to extend the heuristics outlined in the appendix to the case of random coefficients. Second, we only considered a single-player model. While there is a substantial literature on N-player games and, more so on mean-field games (MFGs) with singular controls (see e.g. [9, 10, 11, 18, 20, 22, 26]), MFGs with semimartingale strategies have not yet been considered in the literature to the best of our knowledge.

A Heuristic derivation of the optimal solution

In this appendix we consider a discrete time model from which we heuristically derive the Riccati equations in terms of which we can represent both the value function and the optimal strategy in our continuous time model. The idea is to derive a recursive dynamics in a discrete-time setting and then to take formal limits as the time between two consecutive trading times tends to zero.

A.1 The discrete time model

Let us assume that there are $N+1$ trading times $0, \Delta, 2\Delta, \cdots, N\Delta$. Let the volume traded at time $i\Delta$ be denoted by ξ_i. The state immediately before this control is implemented is denoted by

$$\mathcal{X}_{i\Delta-} := (X_{i\Delta-}, Y_{i\Delta-}, C_{i\Delta-}).$$

Let (ϵ_n) be a sequence of i.i.d. $N(0,\Delta)$-distributed random variables. In terms of the quantities

$$\mathcal{L} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^\top, \quad \mathcal{R} = \frac{\gamma_2}{2}, \quad \mathcal{Q}_\Delta = \begin{pmatrix} \Delta\lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

the discrete-time cost functional is given by

$$J(\xi) = \mathbb{E}\left[\sum_{i=0}^{N} \mathcal{L}^\top \mathcal{X}_{i\Delta-}\xi_i + \mathcal{R}\xi_i^2 + \mathcal{X}_{i\Delta-}^\top \mathcal{Q}_\Delta \mathcal{X}_{i\Delta-}\right],$$

and the discrete time state dynamics in matrix form reads

$$\mathcal{X}_{(n+1)\Delta-} = \mathcal{A}\mathcal{X}_{n\Delta-} + \overline{\mathcal{A}}\mathbb{E}[\mathcal{X}_{n\Delta-}] + \mathcal{B}\xi_n + \overline{\mathcal{B}}\mathbb{E}[\xi_n] + \mathcal{C} + \mathcal{D}\epsilon_{n+1}, \tag{A.1}$$

where

$$\begin{aligned}
&\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-\Delta\rho & -\Delta\gamma_1(\beta-\alpha) \\ 0 & 0 & 1-(\beta-\alpha)\Delta \end{pmatrix}, \quad && \overline{\mathcal{A}} = \begin{pmatrix} 0 & 0 & 0 \\ -\alpha\Delta\gamma_1 & 0 & 0 \\ -\alpha\Delta & 0 & 0 \end{pmatrix}, \\
&\mathcal{B} = \begin{pmatrix} -1 & (1-\Delta\rho)\gamma_2 & 0 \end{pmatrix}^\top, && \overline{\mathcal{B}} = \begin{pmatrix} 0 & \alpha\Delta\gamma_1 & \alpha\Delta \end{pmatrix}^\top, \\
&\mathcal{C} = \begin{pmatrix} 0 & \alpha\Delta\gamma_1\mathbb{E}[x_0] & \alpha\Delta\mathbb{E}[x_0] \end{pmatrix}^\top, && \mathcal{D} = \begin{pmatrix} 0 & \sigma & 0 \end{pmatrix}^\top.
\end{aligned} \tag{A.2}$$

A.1.1 The value function

To derive a representation of the value function in discrete time we denote by μ the law of the random variable $\mathcal{X} := \mathcal{X}_{n\Delta-}$ and set

$$V_n(\mu) := \inf_{(\xi_i)_{i=n}^N} \mathbb{E}\left[\sum_{i=n}^N \mathcal{L}^\top \mathcal{X}_{i\Delta-}\xi_i + \mathcal{R}\xi_i^2 + \mathcal{X}_{i\Delta-}^\top \mathcal{Q}_\Delta \mathcal{X}_{i\Delta-}\right].$$

By the dynamic programming principle given in [16, Corollary 4.1] we have that

$$V_n(\mu) = \inf_\xi \left\{\mathbb{E}\left[\mathcal{L}^\top\mathcal{X}\xi + \mathcal{R}\xi^2 + \mathcal{X}^\top\mathcal{Q}_\Delta\mathcal{X}\right] + V_{n+1}\left(\mathbb{P}\circ\mathcal{X}_{(n+1)\Delta-}^{(\xi),-1}\right)\right\}, \tag{A.3}$$

where $\mathcal{X}^{(\xi)}$ denotes the state corresponding to the control ξ. Let

$$\xi = \widehat{\xi} + \delta, \quad \widehat{\mathcal{X}} = (\widehat{X}_{n\Delta-}, \widehat{Y}_{n\Delta-}, \widehat{C}_{n\Delta-})^\top = (X_{n\Delta-} - \delta, Y_{n\Delta-} + \gamma_2\delta, C_{n\Delta-})^\top = \mathcal{X} + (-\delta, \gamma_2\delta, 0)^\top.$$

A straightforward calculation shows that

$$\begin{aligned}
V_n(\mu) = & \inf_{\widehat{\xi}} \left\{\mathbb{E}\left[\mathcal{L}^\top\widehat{\mathcal{X}}\widehat{\xi} + \mathcal{R}\widehat{\xi}^2 + \widehat{\mathcal{X}}^\top\mathcal{Q}_\Delta\widehat{\mathcal{X}} + V_{n+1}(\mathbb{P}\circ\widehat{\mathcal{X}}_{(n+1)\Delta-}^{(\widehat{\xi}),-1})\right]\right\} \\
& + \mathbb{E}\left[\delta Y_{n\Delta-} + \frac{\gamma_2}{2}\delta^2 + \lambda\Delta\delta^2 + 2\lambda\Delta\delta\widehat{X}_{n\Delta-}\right] \\
= & V_n(\widehat{\mu}) + \mathbb{E}\left[\delta Y_{n\Delta-} + \frac{\gamma_2}{2}\delta^2 + \lambda\Delta\delta^2 + 2\lambda\Delta\delta\widehat{X}_{n\Delta-}\right] \\
= & V_n(\widehat{\mu}) + \mathbb{E}\left[(2\lambda\Delta\delta, \delta, 0)\mathcal{X}\right] + \frac{\gamma_2}{2}\delta^2 - \lambda\Delta\delta^2 \\
= & V_n(\widehat{\mu}) + (2\lambda\Delta\delta, \delta, 0)\overline{\mu} + \frac{\gamma_2}{2}\delta^2 - \lambda\Delta\delta^2,
\end{aligned} \tag{A.4}$$

where $\widehat{\mathcal{X}}$ follows the distribution $\widehat{\mu}$. Let us now make the ansatz

$$V_n(\mu) = \mathrm{Var}(\mu)(A_n) + \overline{\mu}^\top B_n\overline{\mu} + D_n^\top\overline{\mu} + F_n, \tag{A.5}$$

where for each n, $A_n, B_n \in \mathbb{S}^3$, $D_n \in \mathbb{R}^3$ and $F_n \in \mathbb{R}$ are to be determined. Along with equation (A.4) the ansatz yields that

$$\begin{aligned}
& \mathbb{E}\left[\mathcal{X}^\top A_n\mathcal{X}\right] - \mathbb{E}[\mathcal{X}^\top]A_n\mathbb{E}[\mathcal{X}] + \mathbb{E}[\mathcal{X}^\top]B_n\mathbb{E}[\mathcal{X}] + D_n^\top\mathbb{E}[\mathcal{X}] + F_n \\
= & \mathbb{E}\left[\widehat{\mathcal{X}}^\top A_n\widehat{\mathcal{X}}\right] - \mathbb{E}[\widehat{\mathcal{X}}^\top]A_n\mathbb{E}[\widehat{\mathcal{X}}] + \mathbb{E}[\widehat{\mathcal{X}}^\top]B_n\mathbb{E}[\widehat{\mathcal{X}}] + D_n^\top\mathbb{E}[\widehat{\mathcal{X}}] + F_n \\
& + (2\lambda\Delta\delta, \delta, 0)\overline{\mu} + \frac{\gamma_2}{2}\delta^2 - \lambda\Delta\delta^2.
\end{aligned}$$

Dividing by δ on both sides and letting $\delta \to 0$, we get that

$$(-1, \gamma_2, 0)B_n\overline{\mu} + \frac{1}{2}D_n^\top(-1, \gamma_2, 0)^\top + (\lambda\Delta, \frac{1}{2}, 0)\overline{\mu} = 0.$$

The fact that this equation needs to hold for all $\bar{\mu}$ suggests that the coefficients multiplying the entries of the vector $\overline{\mu}$ are all equal to zero and so are the entries of the second summand. This yields that

$$\begin{cases} -B_{n,11} + \gamma_2 B_{n,21} + \lambda\Delta = 0 \\ -B_{n,12} + \gamma_2 B_{n,22} + \dfrac{1}{2} = 0 \\ -B_{n,13} + \gamma_2 B_{n,23} = 0 \\ -D_{n,1} + \gamma_2 D_{n,2} = 0. \end{cases} \tag{A.6}$$

We conjecture that the entries of the matrix A_n satisfy the same equations as those of B_n, i.e.

$$\begin{cases} -A_{n,11} + \gamma_2 A_{n,21} + \lambda\Delta = 0 \\ -A_{n,12} + \gamma_2 A_{n,22} + \dfrac{1}{2} = 0 \\ -A_{n,13} + \gamma_2 A_{n,23} = 0. \end{cases} \tag{A.7}$$

From the equations (A.3) and (A.5), we will derive recursive equations for the coefficients A_n, B_n, D_n and F_n from which we will then derive the candidate continuous time dynamics.

A.1.2 The optimal strategy for the discrete time model

To derive the desired recursion, we first need to determine the candidate optimal strategy for the discrete time model. To this end, we first use (A.1) and (A.2) to conclude that

$$\mathcal{X}_{(n+1)\Delta-} - \mathbb{E}\left[\mathcal{X}_{(n+1)\Delta-}\right] = \mathcal{A}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) + \mathcal{B}\left(\xi_n - \mathbb{E}[\xi_n]\right) + \mathcal{D}\epsilon_{n+1}. \tag{A.8}$$

Hence,

$$\begin{aligned}
&\mathrm{Var}\Big(\mathbb{P} \circ \mathcal{X}^{(\xi),-1}_{(n+1)\Delta-}\Big)(A_{n+1}) \\
&= \mathbb{E}\left[\left(\mathcal{X}^{(\xi)}_{(n+1)\Delta-} - \mathbb{E}\left[\mathcal{X}^{(\xi)}_{(n+1)\Delta-}\right]\right)^\top A_{n+1}\left(\mathcal{X}_{(n+1)\Delta-} - \mathbb{E}\left[\mathcal{X}^{(\xi)}_{(n+1)\Delta-}\right]\right)\right] \\
&= \mathbb{E}\Big[\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) + \mathcal{B}\left(\xi_n - \mathbb{E}[\xi_n]\right) + \mathcal{D}\epsilon_{n+1}\Big\}^\top A_{n+1} \\
&\quad \cdot \Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) + \mathcal{B}\left(\xi_n - \mathbb{E}[\xi_n]\right) + \mathcal{D}\epsilon_{n+1}\Big\}\Big] \\
&= \mathbb{E}\Big[\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) + \mathcal{B}\left(\xi_n - \mathbb{E}[\xi_n]\right)\Big\}^\top A_{n+1}\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) + \mathcal{B}\left(\xi_n - \mathbb{E}[\xi_n]\right)\Big\}\Big] \\
&\quad + \Delta\mathcal{D}^\top A_{n+1}\mathcal{D}.
\end{aligned}$$

Moreover,

$$\overline{\mathbb{P} \circ \mathcal{X}^{(\xi),-1}_{(n+1)\Delta-}} = \mathbb{E}\left[\mathcal{X}^{(\xi)}_{(n+1)\Delta-}\right] = (\mathcal{A} + \overline{\mathcal{A}})\overline{\mu} + (\mathcal{B} + \overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}.$$

From the ansatz (A.5) we get that

$$\begin{aligned}
&V_{n+1}\Big(\mathbb{P} \circ \mathcal{X}^{(\xi),-1}_{(n+1)\Delta-}\Big) \\
&= \mathbb{E}\Big[\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \overline{\mu}\right) + \mathcal{B}\left(\xi - \mathbb{E}[\xi]\right)\Big\}^\top A_{n+1}\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \overline{\mu}\right) + \mathcal{B}\left(\xi - \mathbb{E}[\xi]\right)\Big\}\Big] \\
&\quad + \Delta\mathcal{D}^\top A_{n+1}\mathcal{D} + \Big\{(\mathcal{A} + \overline{\mathcal{A}})\overline{\mu} + (\mathcal{B} + \overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}\Big\}^\top B_{n+1}\Big\{(\mathcal{A} + \overline{\mathcal{A}})\overline{\mu} + (\mathcal{B} + \overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}\Big\} \\
&\quad + D_{n+1}^\top\Big\{(\mathcal{A} + \overline{\mathcal{A}})\overline{\mu} + (\mathcal{B} + \overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}\Big\} + F_{n+1}.
\end{aligned}$$

Thus, the DPP (A.3) implies that

$$
\begin{aligned}
V_n(\mu) = & \inf_{\xi} \Big\{ \mathbb{E}\left[\mathcal{L}^\top \mathcal{X}_{n\Delta-}\xi + \mathcal{R}\xi^2 + \mathcal{X}_{n\Delta-}^\top \mathcal{Q}_\Delta \mathcal{X}_{n\Delta-}\right] \\
& + \mathbb{E}\Big[\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \overline{\mu}\right) + \mathcal{B}\left(\xi - \mathbb{E}[\xi]\right)\Big\}^\top A_{n+1}\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \overline{\mu}\right) + \mathcal{B}\left(\xi - \mathbb{E}[\xi]\right)\Big\}\Big] + \Delta\mathcal{D}^\top A_{n+1}\mathcal{D} \\
& + \Big\{(\mathcal{A}+\overline{\mathcal{A}})\overline{\mu} + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}\Big\}^\top B_{n+1}\Big\{(\mathcal{A}+\overline{\mathcal{A}})\overline{\mu} + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}\Big\} \\
& + D_{n+1}^\top\Big\{(\mathcal{A}+\overline{\mathcal{A}})\overline{\mu} + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi] + \mathcal{C}\Big\} + F_{n+1}\Big\} \\
:= & \inf_{\xi} \mathcal{J}(\xi).
\end{aligned}
$$

Let ξ^* be the candidate optimal strategy and let $\widetilde{\xi}$ be an arbitrary strategy. Then

$$
\begin{aligned}
\mathcal{J}(\xi^* + \varepsilon\widetilde{\xi}) - \mathcal{J}(\xi^*) = & \ \varepsilon\mathbb{E}\Big\{\mathcal{L}^\top \mathcal{X}_{n\Delta-}\widetilde{\xi} + 2\mathcal{R}\xi^*\widetilde{\xi} \\
& + 2\Big(\mathcal{A}(\mathcal{X}_{n\Delta-} - \overline{\mu}) + \mathcal{B}(\xi^* - \mathbb{E}[\xi^*])\Big)^\top A_{n+1}\mathcal{B}(\widetilde{\xi} - \mathbb{E}[\widetilde{\xi}]) \\
& + 2\Big((\mathcal{A}+\overline{\mathcal{A}})\overline{\mu} + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi^*] + \mathcal{C}\Big)^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\widetilde{\xi}] \\
& + D_{n+1}^\top(\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\widetilde{\xi}]\Big\} + O(\varepsilon^2) \\
= & \ \varepsilon\mathbb{E}\Big\{\Big[\mathcal{L}^\top \mathcal{X}_{n\Delta-} + 2\mathcal{R}\xi^* + 2\Big(\mathcal{A}(\mathcal{X}_{n\Delta-} - \overline{\mu}) + \mathcal{B}(\xi^* - \mathbb{E}[\xi^*])\Big)^\top A_{n+1}\mathcal{B} \\
& + 2\Big((\mathcal{A}+\overline{\mathcal{A}})\overline{\mu} + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi^*] + \mathcal{C}\Big)^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}}) + D_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\Big]\widetilde{\xi}\Big\} \\
& + O(\varepsilon^2).
\end{aligned}
$$

Since $\liminf_{\varepsilon\to 0} \frac{\mathcal{J}(\xi^*+\varepsilon\widetilde{\xi}) - \mathcal{J}(\xi^*)}{\varepsilon} \geq 0$ for any $\widetilde{\xi}$, we conclude that

$$
\begin{aligned}
& \mathcal{L}^\top \mathcal{X}_{n\Delta-} + 2\mathcal{R}\xi^* + 2\Big(\mathcal{A}(\mathcal{X}_{n\Delta-} - \overline{\mu}) + \mathcal{B}(\xi^* - \mathbb{E}[\xi^*])\Big)^\top A_{n+1}\mathcal{B} \\
& + 2\Big((\mathcal{A}+\overline{\mathcal{A}})\overline{\mu} + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi^*] + \mathcal{C}\Big)^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}}) \\
& + D_{n+1}^\top(\mathcal{B}+\overline{\mathcal{B}}) \\
= & \ 0.
\end{aligned} \tag{A.9}
$$

Taking expectations on both sides of (A.9), we get that

$$
\begin{aligned}
\mathbb{E}[\xi^*] = & -\Big\{2\mathcal{R} + 2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\Big\}^{-1}\Big\{\mathcal{L}^\top + 2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{A}+\overline{\mathcal{A}})\Big\}\overline{\mu} \\
& -\Big\{2\mathcal{R} + 2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\Big\}^{-1}\Big\{2\mathcal{C}^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}}) + D_{n+1}^\top(\mathcal{B}+\overline{\mathcal{B}})\Big\}.
\end{aligned} \tag{A.10}
$$

Combining (A.9) and (A.10) yields that

$$
\xi^* - \mathbb{E}[\xi^*] = -\Big\{2\mathcal{R} + 2\mathcal{B}^\top A_{n+1}\mathcal{B}\Big\}^{-1}\left(\mathcal{L}^\top + 2\mathcal{B}^\top A_{n+1}\mathcal{A}\right)(\mathcal{X}_{n\Delta-} - \overline{\mu}), \tag{A.11}
$$

and from (A.10) and (A.11) we have that

$$
\begin{aligned}
\xi^* = &-\Big\{2\mathcal{R}+2\mathcal{B}^\top A_{n+1}\mathcal{B}\Big\}^{-1}\left(\mathcal{L}^\top+2\mathcal{B}^\top A_{n+1}\mathcal{A}\right)(\mathcal{X}_{n\Delta-}-\overline{\mu})\\
&-\Big\{2\mathcal{R}+2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\Big\}^{-1}\Big\{\mathcal{L}^\top+2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{A}+\overline{\mathcal{A}})\Big\}\overline{\mu}\\
&-\Big\{2\mathcal{R}+2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\Big\}^{-1}\Big\{2\mathcal{C}^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})+D_{n+1}^\top(\mathcal{B}+\overline{\mathcal{B}})\Big\}.
\end{aligned}
$$

In terms of the notation

$$
\begin{aligned}
\tilde{a}_n &= \mathcal{R}+\mathcal{B}^\top A_{n+1}\mathcal{B}, & a_n &= \mathcal{R}+(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}}),\\
I_n^A &= \frac{1}{2}\mathcal{L}^\top+\mathcal{B}^\top A_{n+1}\mathcal{A}, & I_n^B &= \frac{1}{2}\mathcal{L}^\top+(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{A}+\overline{\mathcal{A}}),\\
I_n^D &= \mathcal{C}^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})+\frac{1}{2}D_{n+1}^\top(\mathcal{B}+\overline{\mathcal{B}}),
\end{aligned}
\tag{A.12}
$$

the optimal strategy at time $n\Delta$ as a function of the initial state $\mathcal{X}_{n\Delta-}$ can be written as

$$
\xi_n^* = \xi_n^*(\mathcal{X}_{n\Delta-}) = -\frac{I_n^A}{\tilde{a}_n}(\mathcal{X}_{n\Delta-}-\overline{\mu})-\frac{I_n^B}{a_n}\overline{\mu}-\frac{I_n^D}{a_n}. \tag{A.13}
$$

A.2 Heuristic derivation of (A_t, B_t, D_t, F_t)

To obtain the continuous time dynamics of the coefficient processes (A, B, D, F) we are now going to derive a recursive dynamics for the processes A_n, B_n, D_n and F_n and then formal derivatives.

Taking the equation (A.13) back into the cost function $\mathcal{J}(\cdot)$ we get that

$$
\begin{aligned}
\mathcal{J}(\xi_n^*) = &\ \mathbb{E}\Big\{\mathcal{L}^\top(\mathcal{X}_{n\Delta-}-\overline{\mu})\xi_n^*+\mathcal{L}^\top\overline{\mu}\xi_n^*+\mathcal{R}(\xi_n^*)^2+(\mathcal{X}_{n\Delta-}-\overline{\mu})^\top\mathcal{Q}_\Delta(\mathcal{X}_{n\Delta-}-\overline{\mu})+\overline{\mu}^\top\mathcal{Q}_\Delta\overline{\mu}\Big\}\\
&+\mathbb{E}\Big\{(\mathcal{X}_{n\Delta-}-\overline{\mu})^\top\mathcal{A}^\top A_{n+1}\mathcal{A}(\mathcal{X}_{n\Delta-}-\overline{\mu})+2(\mathcal{X}_{n\Delta-}-\overline{\mu})^\top\mathcal{A}^\top A_{n+1}\mathcal{B}(\xi_n^*-\mathbb{E}[\xi_n^*])\\
&+\mathcal{B}^\top A_{n+1}\mathcal{B}(\xi_n^*-\mathbb{E}[\xi_n^*])^2+\Delta\mathcal{D}^\top A_{n+1}\mathcal{D}\Big\}\\
&+\overline{\mu}^\top(\mathcal{A}+\overline{\mathcal{A}})^\top B_{n+1}(\mathcal{A}+\overline{\mathcal{A}})\overline{\mu}+2\overline{\mu}^\top(\mathcal{A}+\overline{\mathcal{A}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi_n^*]\\
&+2(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}\mathcal{C}\mathbb{E}[\xi_n^*]+(\mathcal{B}+\overline{\mathcal{B}})^\top B_{n+1}(\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi_n^*]^2\\
&+2\overline{\mu}^\top(\mathcal{A}+\overline{\mathcal{A}})^\top B_{n+1}\mathcal{C}+\mathcal{C}^\top B_{n+1}\mathcal{C}\\
&+D_{n+1}^\top(\mathcal{A}+\overline{\mathcal{A}})\overline{\mu}+D_{n+1}^\top(\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi_n^*]+D_{n+1}^\top\mathcal{C}+F_{n+1}\\
= &\ \mathrm{Var}(\mu)\Big(\mathcal{Q}_\Delta+\mathcal{A}^\top A_{n+1}\mathcal{A}-(I_n^A)^\top\tilde{a}_n^{-1}I_n^A\Big)+\overline{\mu}^\top\Big(\mathcal{Q}_\Delta+(\mathcal{A}+\overline{\mathcal{A}})^\top B_{n+1}(\mathcal{A}+\overline{\mathcal{A}})-(I_n^B)^\top a_n^{-1}I_n^B\Big)\overline{\mu}\\
&+\Big(-2a_n^{-1}I_n^D I_n^B+(2\mathcal{C}^\top B_{n+1}+D_{n+1}^\top)(\mathcal{A}+\overline{\mathcal{A}})\Big)\overline{\mu}\\
&-\frac{(I_n^D)^2}{4a_n}+\Delta\mathcal{D}^\top A_{n+1}\mathcal{D}+\mathcal{C}^\top B_{n+1}\mathcal{C}+D_{n+1}^\top\mathcal{C}+F_{n+1}.
\end{aligned}
$$

Comparing the coefficients of $\mathrm{Var}(\mu)(\cdots)$, $\overline{\mu}^\top(\cdots)\overline{\mu}$, $(\cdots)\overline{\mu}$ and the remaining terms respectively, we see that A_n, B_n, D_n and F_n satisfy the following recursive equations:

$$
\begin{cases}
A_n = \mathcal{Q}_\Delta+\mathcal{A}^\top A_{n+1}\mathcal{A}-(I_n^A)^\top\tilde{a}_n^{-1}I_n^A,\\
B_n = \mathcal{Q}_\Delta+(\mathcal{A}+\overline{\mathcal{A}})^\top B_{n+1}(\mathcal{A}+\overline{\mathcal{A}})-(I_n^B)^\top a_n^{-1}I_n^B,\\
D_n^\top = -2a_n^{-1}I_n^D I_n^B+(2\mathcal{C}^\top B_{n+1}+D_{n+1}^\top)(\mathcal{A}+\overline{\mathcal{A}}),\\
F_n = -\dfrac{(I_n^D)^2}{4a_n}+\Delta\mathcal{D}^\top A_{n+1}\mathcal{D}+\mathcal{C}^\top B_{n+1}\mathcal{C}+D_{n+1}^\top\mathcal{C}+F_{n+1}.
\end{cases}
\tag{A.14}
$$

Let (A, B, D, F) be the formal limit of (A_n, B_n, D_n, F_n) as $\Delta \to 0$. Using (A.6) and (A.7), the driver of (A, B, D, F) is formally obtained by taking the limit

$$\lim_{\Delta \to 0} \frac{\Phi_{n+1} - \Phi_n}{\Delta}, \qquad \Phi = A, B, D, F.$$

In terms of the notation

$$\begin{aligned}
&J^A = \gamma_1 A_{11} + \gamma_2 A_{13}, \quad \widetilde{J}^A = \gamma_1 A_{13} + \gamma_2 A_{33}, \quad \widehat{J}^A = -\rho A_{13} + \frac{\gamma_1(\beta - \alpha)}{2}, \\
&J^B = \gamma_1 B_{11} + \gamma_2 B_{13}, \quad \widetilde{J}^B = \gamma_1 B_{13} + \gamma_2 B_{33}, \quad \widehat{J}^B = -\rho B_{13} + \frac{\gamma_1(\beta - \alpha)}{2}, \quad \breve{J}^B = -\rho B_{11} + \frac{\alpha\gamma_1}{2}, \\
&J^D = \gamma_1 D_1 + \gamma_2 D_3, \quad \breve{J}^D = -\rho D_1 - \alpha\gamma_1 \mathbb{E}[x_0],
\end{aligned}$$

we obtain

- the following matrix-valued ODE for A:

$$\begin{cases}
-\dfrac{dA_t}{dt} = \begin{pmatrix} 0 & -\rho\frac{A_{11,t}}{\gamma_2} & -\frac{\beta-\alpha}{\gamma_2} J_t^A \\ -\rho\frac{A_{11,t}}{\gamma_2} & -\rho\frac{2A_{11,t}-\gamma_2}{\gamma_2^2} & -\frac{\beta-\alpha}{\gamma_2^2} J_t^A + \frac{\widehat{J}_t^A}{\gamma_2} \\ -\frac{\beta-\alpha}{\gamma_2} J_t^A & -\frac{\beta-\alpha}{\gamma_2^2} J_t^A + \frac{\widehat{J}_t^A}{\gamma_2} & -\frac{2(\beta-\alpha)}{\gamma_2} \widetilde{J}_t^A \end{pmatrix} \\
\qquad - \dfrac{1}{\lambda + \gamma_2\rho} \begin{pmatrix} -\rho A_{11,t} - \lambda \\ -\rho\frac{A_{11,t}}{\gamma_2} + \rho \\ \frac{\gamma_1(\beta-\alpha)}{2} - \rho A_{13,t} \end{pmatrix} \begin{pmatrix} -\rho A_{11,t} - \lambda \\ -\rho\frac{A_{11,t}}{\gamma_2} + \rho \\ \frac{\gamma_1(\beta-\alpha)}{2} - \rho A_{13,t} \end{pmatrix}^\top \\
\qquad + \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
A_T = \begin{pmatrix} \frac{\gamma_2}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},
\end{cases} \tag{A.15}$$

- the following matrix-valued Riccati equation for B:

$$\begin{cases}
-\dfrac{dB_t}{dt} = \begin{pmatrix} -\frac{2\alpha}{\gamma_2} J_t^B & -\frac{\alpha}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \breve{J}_t^B & -\frac{\beta-\alpha}{\gamma_2} J_t^B - \frac{\alpha}{\gamma_2} \widetilde{J}_t^B \\ -\frac{\alpha}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \breve{J}_t^B & -\rho\frac{2B_{11,t}-\gamma_2}{\gamma_2^2} & -\frac{(\beta-\alpha)}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \widehat{J}_t^B \\ -\frac{\beta-\alpha}{\gamma_2} J_t^B - \frac{\alpha}{\gamma_2} \widetilde{J}_t^B & -\frac{(\beta-\alpha)}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \widehat{J}_t^B & -\frac{2(\beta-\alpha)}{\gamma_2} \widetilde{J}_t^B \end{pmatrix} \\
\qquad - \dfrac{1}{\lambda + \gamma_2\rho - \alpha\gamma_1} \begin{pmatrix} \frac{\alpha}{\gamma_2} J_t^B - \breve{J}_t^B - \lambda \\ \frac{\alpha}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \breve{J}_t^B + \frac{\rho\gamma_2 - \alpha\gamma_1}{\gamma_2} \\ \widehat{J}_t^B + \frac{\alpha}{\gamma_2} \widetilde{J}_t^B \end{pmatrix} \begin{pmatrix} \frac{\alpha}{\gamma_2} J_t^B - \breve{J}_t^B - \lambda \\ \frac{\alpha}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \breve{J}_t^B + \frac{\rho\gamma_2 - \alpha\gamma_1}{\gamma_2} \\ \widehat{J}_t^B + \frac{\alpha}{\gamma_2} \widetilde{J}_t^B \end{pmatrix}^\top \\
\qquad + \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
B_T = \begin{pmatrix} \frac{\gamma_2}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},
\end{cases} \tag{A.16}$$

- the following vector-valued ODE for D:

$$
\begin{cases}
-\dfrac{dD_t}{dt} = \begin{pmatrix} \frac{2\alpha\mathbb{E}[x_0]}{\gamma_2} J_t^B - \frac{\alpha}{\gamma_2} J_t^D \\ \frac{2\alpha\mathbb{E}[x_0]}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \check{J}_t^D \\ \frac{2\alpha\mathbb{E}[x_0]}{\gamma_2} \widetilde{J}_t^B - \frac{\beta-\alpha}{\gamma_2} J_t^D \end{pmatrix} \\
\qquad\qquad - \dfrac{1}{\lambda + \gamma_2\rho - \alpha\gamma_1} \left(\dfrac{\alpha}{\gamma_2} J_t^D + \check{J}_t^D \right) \begin{pmatrix} \frac{\alpha}{\gamma_2} J_t^B + \check{J}_t^B - \lambda \\ \frac{\alpha}{\gamma_2^2} J_t^B + \frac{1}{\gamma_2} \check{J}_t^B + \rho \\ \frac{\alpha}{\gamma_2} \widetilde{J}_t^B + \check{J}_t^B \end{pmatrix} \\
D_T = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},
\end{cases}
\tag{A.17}
$$

- and the following BSDE for F:

$$
\begin{cases}
-dF_t = \Big\{ \sigma_t^2 \dfrac{2A_{11,t} - \gamma_2}{2\gamma_2^2} + \alpha\gamma_1 \mathbb{E}[x_0] \dfrac{D_{1,t}}{\gamma_2} + \alpha\mathbb{E}[x_0] D_{3,t} \\
\qquad - \dfrac{1}{4(\lambda + \gamma_2\rho - \alpha\gamma_1)} \left(-\alpha\gamma_1 \mathbb{E}[x_0] + (\gamma_1\alpha - \gamma_2\rho) \dfrac{D_{1,t}}{\gamma_2} + \alpha D_{3,t} \right)^2 \Big\} dt - Z_t^F \, dW_t \\
F_T = 0.
\end{cases}
\tag{A.18}
$$

In view of (A.6) and (A.7) the above systems reduce to (2.6)-(2.9).

A.3 Heuristic derivation of the optimal strategy in continuous time

In this section, we construct a continuous-time candidate optimal strategy by taking limits of the discrete time model. Intuitively, and in view of the results established in [28, 29] we expect the optimal strategy to jump only at the initial and the terminal time, and to follow an SDE on the open interval $(0, T)$.

A.3.1 The jumps

The final jump size is X_{T-} in order to close the open position at T. To determine the initial jump we first deduce from (A.12) that

$$
\begin{aligned}
\tilde{a}_n &= \Delta(\lambda + \gamma_2\rho) + O(\Delta^2), \\
a_n &= \Delta(\lambda + \rho\gamma_2 - \alpha\gamma_1) + O(\Delta^2), \\
I_n^A &= \Delta \left(-(\rho A_{11,n\Delta} + \lambda) + O(\Delta), (\rho - \frac{A_{11,n\Delta}}{\gamma_2}) + O(\Delta), (\frac{\gamma_1(\beta-\alpha)}{2} - \rho A_{13,n\Delta}) + O(\Delta) \right), \\
I_n^B &= \Delta \begin{pmatrix} (\frac{\alpha\gamma_1}{\gamma_2} - \rho) B_{11,n\Delta} + \alpha B_{13,n\Delta} - \lambda + \frac{\alpha\gamma_1}{2} + O(\Delta) \\ \frac{\alpha\gamma_1 - \gamma_2\rho}{\gamma_2^2} B_{11,n\Delta} + \alpha \frac{B_{13,n\Delta}}{\gamma_2} + \rho - \frac{\alpha\gamma_1}{2\gamma_2} + O(\Delta) \\ \frac{\gamma_1(\beta-\alpha)}{2} + (\gamma_1\alpha - \gamma_2\rho) \frac{B_{13,n\Delta}}{\gamma_2} + \alpha B_{33,n\Delta} + O(\Delta) \end{pmatrix}^\top, \\
I_n^D &= -\Delta \left(\frac{\alpha\gamma_1}{2} \mathbb{E}[x_0] + \frac{(\alpha\gamma_1 - \rho\gamma_2)}{2\gamma_2} D_{1,n\Delta} + \frac{\alpha\Delta D_{3,n\Delta}}{2} + O(\Delta) \right).
\end{aligned}
$$

Thus, letting $n\Delta = t$ and $n \to \infty$ in (A.13), we see that

$$
\begin{aligned}
\xi_n^* &= -\frac{I_n^A}{\tilde{a}_n} (\mathcal{X}_{n\Delta-} - \overline{\mu}) - \frac{I_n^B}{a_n} \overline{\mu} - \frac{I_n^D}{a_n} \\
&\to -\frac{I_t^A}{\tilde{a}} (\mathcal{X}_{t-} - \overline{\mu}) - \frac{I_t^B}{a} \overline{\mu} - \frac{I_t^D}{a} := \Delta Z_t,
\end{aligned}
\tag{A.19}
$$

where

$$\tilde{a} = \gamma_2\rho + \lambda, \qquad a = \gamma_2\rho - \gamma_1\alpha + \lambda,$$

$$I^A = \begin{pmatrix} -\rho A_{11} - \lambda \\ -\rho\frac{A_{11}}{\gamma_2} + \rho \\ \frac{\gamma_1(\beta-\alpha)}{2} - \rho A_{13} \end{pmatrix}^\top, \quad I^B = \begin{pmatrix} \frac{\alpha\gamma_1-\gamma_2\rho}{\gamma_2}B_{11} + \alpha B_{13} - \lambda + \frac{\alpha\gamma_1}{2} \\ \frac{\alpha\gamma_1-\gamma_2\rho}{\gamma_2^2}B_{11} + \alpha\frac{B_{13}}{\gamma_2} + \rho - \frac{\alpha\gamma_1}{2\gamma_2} \\ \frac{\gamma_1(\beta-\alpha)}{2} + (\gamma_1\alpha - \gamma_2\rho)\frac{B_{13}}{\gamma_2} + \alpha B_{33} \end{pmatrix}^\top, \tag{A.20}$$

$$I^D = -\frac{\alpha\gamma_1}{2}\mathbb{E}[x_0] + (\gamma_1\alpha - \gamma_2\rho)\frac{D_1}{2\gamma_2} + \frac{\alpha}{2}D_3.$$

The limit ΔZ_t in (A.19) is the candidate initial jump when starting with a position $\mathcal{X}_{t-}$ at time $t \in [0,T)$.

A.3.2 The candidate strategy on (t,T).

Next, we derive recursive dynamics for the discrete-time optimal strategy from which we deduce a candidate optimal continuous-time strategy. The strategy at time $(n+1)\Delta$ satisfies

$$\begin{aligned}
&\xi^*_{n+1}\Big(\mathcal{X}_{(n+1)\Delta-}\Big)\\
=\ &-\frac{I^A_{n+1}}{\tilde{a}_{n+1}}\left(\mathcal{X}_{(n+1)\Delta-} - \mathbb{E}[\mathcal{X}_{(n+1)\Delta-}]\right) - \frac{I^B_{n+1}}{a_{n+1}}\mathbb{E}[\mathcal{X}_{(n+1)\Delta-}] - \frac{I^D_{n+1}}{a_{n+1}}\\
=\ &-\frac{I^A_{n+1}}{\tilde{a}_{n+1}}\Big\{\mathcal{A}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) + \mathcal{B}\left(\xi^*_n - \mathbb{E}[\xi^*_n]\right) + \mathcal{D}\epsilon_{n+1}\Big\}\\
&-\frac{I^B_{n+1}}{a_{n+1}}\Big\{(\mathcal{A}+\overline{\mathcal{A}})\mathbb{E}[\mathcal{X}_{n\Delta-}] + (\mathcal{B}+\overline{\mathcal{B}})\mathbb{E}[\xi^*_n] + \mathcal{C}\Big\} - \frac{I^D_{n+1}}{a_{n+1}}\\
=\ &-\frac{I^A_{n+1}}{\tilde{a}_{n+1}}\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) - \frac{I^B_{n+1}}{a_{n+1}}\mathbb{E}[\mathcal{X}_{n\Delta-}] - \frac{I^D_{n+1}}{a_{n+1}}\\
&-\frac{I^A_{n+1}}{\tilde{a}_{n+1}}\mathcal{K}\left(\xi^*_n - \mathbb{E}[\xi^*_n]\right) - \frac{I^B_{n+1}}{a_{n+1}}\mathcal{K}\mathbb{E}[\xi^*_n] - \frac{I^A_{n+1}}{\tilde{a}_{n+1}}\mathcal{D}\epsilon_{n+1} - \frac{I^B_{n+1}}{a_{n+1}}\mathcal{C}\\
&-\frac{I^A_{n+1}}{\tilde{a}_{n+1}}(\mathcal{A} - I_{3\times3})\left(\mathcal{X}_{n\Delta-} - \mathbb{E}[\mathcal{X}_{n\Delta-}]\right) - \frac{I^B_{n+1}}{a_{n+1}}(\mathcal{A}+\overline{\mathcal{A}} - I_{3\times3})\mathbb{E}[\mathcal{X}_{n\Delta-}]\\
&-\frac{I^A_{n+1}}{\tilde{a}_{n+1}}(\mathcal{B} - \mathcal{K})\left(\xi^*_n - \mathbb{E}[\xi^*_n]\right) - \frac{I^B_{n+1}}{a_{n+1}}(\mathcal{B}+\overline{\mathcal{B}} - \mathcal{K})\mathbb{E}[\xi^*_n],
\end{aligned}$$

where $I_{3\times3}$ is the identity matrix of order 3. In view of (A.13) the first line in the third equality equals $\xi^*_{n+1}(\mathcal{X}_{n\Delta-})$. Thus,

$$\begin{aligned}
&\xi^*_{n+1}(\mathcal{X}_{(n+1)\Delta-})\\
&= \xi^*_{n+1}(\mathcal{X}_{n\Delta-}) - \frac{I^A_{n+1}}{\tilde{a}_{n+1}}\mathcal{K}\left(\xi^*_n - \mathbb{E}[\xi^*_n]\right) - \frac{I^B_{n+1}}{a_{n+1}}\mathcal{K}\mathbb{E}[\xi^*_n] - \frac{I^A_{n+1}}{\tilde{a}_{n+1}}\mathcal{D}\epsilon_{n+1} - \frac{I^B_{n+1}}{a_{n+1}}\mathcal{C}\\
&\quad - \frac{I^A_{n+1}}{\tilde{a}_{n+1}}(\mathcal{A} - I_{3\times3})\left(\mathcal{X}_{n\Delta-} + \mathcal{K}\xi^*_n - \mathbb{E}[\mathcal{X}_{n\Delta-} + \mathcal{K}\xi^*_n]\right) - \frac{I^B_{n+1}}{a_{n+1}}(\mathcal{A}+\overline{\mathcal{A}} - I_{3\times3})\mathbb{E}[\mathcal{X}_{n\Delta-} + \mathcal{K}\xi^*_n]\\
&\quad - \frac{I^A_{n+1}}{\tilde{a}_{n+1}}(\mathcal{B} - \mathcal{K} - (\mathcal{A} - I_{3\times3})\mathcal{K})\left(\xi^*_n - \mathbb{E}[\xi^*_n]\right)\\
&\quad - \frac{I^B_{n+1}}{a_{n+1}}(\mathcal{B}+\overline{\mathcal{B}} - \mathcal{K} - (\mathcal{A}+\overline{\mathcal{A}} - I_{3\times3})\mathcal{K})\mathbb{E}[\xi^*_n].
\end{aligned}$$

Direct computations show that

$$\begin{cases}
I^A_{n+1}\mathcal{K} = \tilde{a}_{n+1}, \quad I^B_{n+1}\mathcal{K} = a_{n+1},\\
I^A_{n+1}\mathcal{K} = \tilde{a}_{n+1}, \quad I^B_{n+1}\mathcal{K} = a_{n+1},\\
\mathcal{B} - \mathcal{K} - (\mathcal{A} - I_{3\times3})\mathcal{K} = 0_{3\times1},\\
\mathcal{B} + \overline{\mathcal{B}} - \mathcal{K} - (\mathcal{A}+\overline{\mathcal{A}} - I_{3\times3})\mathcal{K} = 0_{3\times1},
\end{cases} \tag{A.21}$$

where $0_{3\times1}$ is 3×1 vector with zero entries. Hence we have that

$$\begin{aligned}
&\xi^*_{n+1}\Big(\mathcal{X}_{(n+1)\Delta-}\Big)\\
&=\xi^*_{n+1}(\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n)-\xi^*_n(\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n)+O(\Delta^2)\\
&\quad-\frac{I^A_{n+1}}{\tilde a_{n+1}}(\mathcal{A}-I_{3\times3})\left(\mathcal{X}_{n\Delta+}-\mathbb{E}[\mathcal{X}_{n\Delta+}]\right)-\frac{I^B_{n+1}}{a_{n+1}}(\mathcal{A}+\overline{\mathcal{A}}-I_{3\times3})\mathbb{E}[\mathcal{X}_{n\Delta+}]-\frac{I^B_{n+1}}{a_{n+1}}\mathcal{C}\\
&\quad-\frac{I^A_{n+1}}{\tilde a_{n+1}}\mathcal{D}\epsilon_{n+1}\\
&:=\mathbb{I}+\mathbb{III}+\mathbb{IIII},
\end{aligned}\tag{A.22}$$

where we use $\mathcal{X}_{n\Delta+}$ to denote the state at $n\Delta$ after ξ^*_n is implemented. Let $\widetilde{Z}$ be the candidate optimal strategy, and let $\delta\widetilde{Z}_{(n+1)\Delta}:=\xi^*_{n+1}(\mathcal{X}_{(n+1)\Delta-})$. (A.22) can be written as

$$\delta\widetilde{Z}_{(n+1)\Delta}=\Big(\frac{\mathbb{I}+\mathbb{III}}{\Delta}\Big)\Delta-\frac{I^A_{n+1}}{\tilde a_{n+1}}\mathcal{D}\epsilon_{n+1}.$$

This suggests to study the limit of $\frac{\mathbb{I}}{\Delta}$ and limit of $\frac{\mathbb{III}}{\Delta}$ as $\Delta\to0$. By (A.13), we have that

$$\begin{aligned}
\frac{\mathbb{I}}{\Delta}=&\frac{\xi^*_{n+1}(\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n)-\xi^*_n(\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n)}{\Delta}+\frac{O(\Delta^2)}{\Delta}\\
=&-\frac{I^A_{n+1}-I^A_n}{\Delta\tilde a_{n+1}}\left(\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n-\mathbb{E}[\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n]\right)\\
&-\frac{I^A_n}{\Delta}\left(\frac{1}{\tilde a_{n+1}}-\frac{1}{\tilde a_n}\right)\left(\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n-\mathbb{E}[\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n]\right)\\
&-\frac{I^B_{n+1}-I^B_n}{\Delta a_{n+1}}\mathbb{E}[\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n]-\frac{I^B_n}{\Delta}\left(\frac{1}{a_{n+1}}-\frac{1}{a_n}\right)\mathbb{E}[\mathcal{X}_{n\Delta+}-\mathcal{K}\xi^*_n]\\
&-\frac{I^D_{n+1}-I^D_n}{\Delta a_{n+1}}-\frac{I^D_n}{\Delta}\left(\frac{1}{a_{n+1}}-\frac{1}{a_n}\right)+\frac{O(\Delta^2)}{\Delta}.
\end{aligned}\tag{A.23}$$

Since $\tilde a_{n+1},\tilde a_n,a_{n+1},a_n=O(\Delta)$, we get that

$$\begin{aligned}
\frac{\tilde a_{n+1}-\tilde a_n}{\Delta^3}&=\rho^2\frac{A_{11,n+1}-A_{11,n}}{\Delta}\to\rho^2\frac{\dot A_{11,t}}{dt}\\
\frac{a_{n+1}-a_n}{\Delta^3}&=\frac{(\gamma_1\alpha-\gamma_2\rho)^2}{\gamma_2^2}\frac{B_{11,n+1}-B_{11,n}}{\Delta}+\frac{2\alpha(\gamma_1\alpha-\gamma_2\rho)}{\gamma_2}\frac{B_{13,n+1}-B_{13,n}}{\Delta}+\alpha^2\frac{B_{33,n+1}-B_{33,n}}{\Delta}\\
&\to\frac{(\gamma_1\alpha-\gamma_2\rho)^2}{\gamma_2^2}\dot B_{11,t}+\frac{2\alpha(\gamma_1\alpha-\gamma_2\rho)}{\gamma_2}\dot B_{13,t}+\alpha^2\dot B_{33,t},
\end{aligned}$$

from which we deduce that

$$\begin{aligned}
\frac{1}{\tilde a_{n+1}}-\frac{1}{\tilde a_n}&=-\frac{\tilde a_{n+1}-\tilde a_n}{\tilde a_{n+1}\tilde a_n}=-\frac{\Delta^2}{\tilde a_{n+1}\tilde a_n}\cdot\frac{\tilde a_{n+1}-\tilde a_n}{\Delta^3}\Delta\to0,\\
\frac{1}{a_{n+1}}-\frac{1}{a_n}&=-\frac{a_{n+1}-a_n}{a_{n+1}a_n}=-\frac{\Delta^2}{a_{n+1}a_n}\cdot\frac{a_{n+1}-a_n}{\Delta^3}\Delta\to0.
\end{aligned}$$

Furthermore, we have that

$$\begin{aligned}
\frac{(I^A_{n+1}-I^A_n)\mathcal{K}}{\Delta^3}&=\rho^2\frac{A_{11,n+1}-A_{11,n}}{\Delta}\to\rho^2\dot A_{11,t}\\
\frac{(I^B_{n+1}-I^B_n)\mathcal{K}}{\Delta^3}&=\frac{(\gamma_1\alpha-\gamma_2\rho)^2}{\gamma_2^2}\frac{B_{11,n+1}-B_{11,n}}{\Delta}+\frac{2\alpha(\gamma_1\alpha-\gamma_2\rho)}{\gamma_2}\frac{B_{13,n+1}-B_{13,n}}{\Delta}+\alpha^2\frac{B_{33,n+1}-B_{33,n}}{\Delta}\\
&\to\frac{(\gamma_1\alpha-\gamma_2\rho)^2}{\gamma_2^2}\dot B_{11,t}+\frac{2\alpha(\gamma_1\alpha-\gamma_2\rho)}{\gamma_2}\dot B_{13,t}+\alpha^2\dot B_{33,t}.
\end{aligned}$$

Altogether, we conclude that

$$\frac{\mathbb{I}}{\Delta} \to -\frac{\dot{I}_t^A}{\tilde{a}}(\mathcal{X}_t - \mathbb{E}[\mathcal{X}_t]) - \frac{\dot{I}_t^B}{a}\mathbb{E}[\mathcal{X}_t] - \frac{\dot{I}_t^D}{a}.$$

The limit of $\frac{\mathbb{III}}{\Delta}$ as $\Delta \to 0$ reads

$$\begin{aligned}\frac{\mathbb{III}}{\Delta} = & \frac{1}{\Delta}\Big(-\frac{I_{n+1}^A}{\tilde{a}_{n+1}}(\mathcal{A} - I_{3\times 3})\,(\mathcal{X}_{n\Delta+} - \mathbb{E}[\mathcal{X}_{n\Delta+}]) \\ & - \frac{I_{n+1}^B}{a_{n+1}}(\mathcal{A} + \overline{\mathcal{A}} - I_{3\times 3})\mathbb{E}[\mathcal{X}_{n\Delta+}] - \frac{I_{n+1}^B}{a_{n+1}}\mathcal{C}\Big) \\ \to & -\frac{I_t^A}{\tilde{a}}\mathcal{H}(\mathcal{X}_t - \mathbb{E}[\mathcal{X}_t]) - \frac{I_t^B}{a}\Big((\mathcal{H} + \overline{\mathcal{H}})\mathbb{E}[\mathcal{X}_t] + \mathcal{G}\Big).\end{aligned}$$

Moreover, heuristically,

$$\mathbb{IIII} = -\frac{I_{n+1}^A}{\tilde{a}_{n+1}}\mathcal{D}\epsilon_{n+1} \to -\frac{I_t^A}{\tilde{a}}\mathcal{D}_t\, dW_t.$$

Combining the above limits we obtain the candidate trading strategy on (t, T):

$$\begin{aligned}-dX_s = d\widetilde{Z}_s = & \Bigg(-\frac{\dot{I}_s^A}{\tilde{a}}(\mathcal{X}_s - \mathbb{E}[\mathcal{X}_s]) - \frac{\dot{I}_s^B}{a}\mathbb{E}[\mathcal{X}_s] - \frac{\dot{I}_s^D}{a} - \frac{I_s^A}{\tilde{a}}\mathcal{H}(\mathcal{X}_s - \mathbb{E}[\mathcal{X}_s]) \\ & -\frac{I_s^B}{a}((\mathcal{H} + \overline{\mathcal{H}})\mathbb{E}[\mathcal{X}_s] + \mathcal{G})\Bigg)\, ds - \frac{I_s^A}{\tilde{a}}\mathcal{D}_s\, dW_s.\end{aligned}$$

References

[1] J. Ackermann, T. Kruse, and M. Urusov. Càdlàg semimartingale strategies for optimal trade execution in stochastic order book models. *Finance and Stochastics*, 25(4):757–810, 2021.

[2] A. Alfonsi and P. Blanc. Dynamic optimal execution in a mixed-market-impact Hawkes price model. *Finance and Stochastics*, 20(1):183–218, 2016.

[3] S. Ankirchner, M. Jeanblanc, and T. Kruse. BSDEs with singular terminal condition and a control problem with constraints. *SIAM Journal on Control and Optimization*, 52(2):893–913, 2014.

[4] P. Bank. Optimal control under dynamic fuel constraint. *SIAM Journal on Control and Optimization*, 44(4):1529–1541, 2005.

[5] P. Bank and N. El Karoui. A stochastic representation theorem with applications to optimization and obstacle problems. *Annals of Probability*, 32(1B):1030–1067, 2004.

[6] P. Bank and F. Riedel. Optimal consumption choice with intertemporal substitution. *Annals of Applied Probability*, 11(3):750–788, 2001.

[7] D. Becherer, T. Bilarev, and P. Frentrup. Stability for gains from large investors' strategies in M1/J1 topologies. *Bernoulli*, 25(2):1105–1140, 2019.

[8] M.K. Brunnermeier and L.H. Pedersen. Predatory trading. *Journal of Finance*, 60(4):1825–1863, 2005.

[9] L. Campi, T. De Angelis, M. Ghio, and G. Livieri. Mean-field games of finite-fuel capacity expansion with singular controls. *to appear in Annals of Applied Probability*, 2020.

[10] H. Cao, J. Dianetti, and G. Ferrari. Stationary discounted and ergodic mean field games of singular control. *arXiv:2105.07213*, 2021.

[11] H. Cao and X. Guo. MFGs for partially reversible investment. *to appear in Stochastic Processes and their Applications*, 2020.

[12] B. Carlin, M. Lobo, and S. Viswanathan. Episodic liquidity crises: Cooperative and predatory trading. *Journal of Finance*, 62(5):2235–2274, 2007.

[13] R. Carmona and K. Webster. The self-financing equation in limit order book markets. *Finance and Stochastics*, 23(3):729–759, 2019.

[14] Á. Cartea, S. Jaimungal, and J. Ricci. Algorithmic trading, stochastic control, and mutually exciting processes. *SIAM Review*, 60(3):673–703, 2018.

[15] Y. Chen, U. Horst, and H. Tran. Optimal trade execution under endogenous order flow. *https://ssrn.com/abstract=3933364*, 2021.

[16] A. Cosso and H. Pham. Zero-sum stochastic differential games of generalized McKean–Vlasov type. *Journal de Mathématiques Pures et Appliquées*, 129:180–212, 2019.

[17] J. Dianetti and G. Ferrari. Multidimensional singular control and related Skorokhod problem: sufficient conditions for the characterization of optimal controls. *arXiv:2103.08487*, 2021.

[18] J. Dianetti, G. Ferrari, M. Fischer, and M. Nendel. A unifying framework for submodular mean field games. *arXiv:2201.07850*, 2022.

[19] R. Elie, L. Moreau, and D. Possamaï. On a class of path-dependent singular stochastic control problems. *SIAM Journal on Control and Optimization*, 56(5):3260–3295, 2018.

[20] G. Fu. Extended mean field games with singular controls. *arXiv preprint arXiv:1909.04154*, 2019.

[21] G. Fu, P. Graewe, U. Horst, and A. Popier. A mean field game of optimal portfolio liquidation. *Mathematics of Operations Research*, 46(4):1235–1657, 2021.

[22] G. Fu and U. Horst. Mean field games with singular controls. *SIAM Journal on Control and Optimization*, 55(6):3833–3868, 2017.

[23] G. Fu, U. Horst, and X. Xia. Portfolio liquidation games with self-exciting order flow. *to appear in Mathematical Finance*, 2022.

[24] N. Gârleanu and L.H. Pedersen. Dynamic portfolio choice with frictions. *Journal of Economic Theory*, 165:487–516, 2016.

[25] X. Guo, H. Pham, and X. Wei. Itô's formula for flow of measures on semimartingales. *arXiv:2010.05288*, 2020.

[26] X. Guo and R. Xu. Stochastic games for fuel followers problem: N vs MFG. *SIAM Journal on Control and Optimization*, 57(1):659–692, 2020.

[27] M. Hafayed. A mean-field necessary and sufficient conditions for optimal singular stochastic control. *Communications in Mathematics and Statistics*, 1:417–435, 2013.

[28] U. Horst and E. Kivman. Optimal trade execution under small market impact and portfolio liquidation with semimartingale strategies. *arXiv:2103.05957*, 2022.

[29] U. Horst and F. Naujokat. When to cross the spread? Trading in two-sided limit order books. *SIAM Journal on Financial Mathematics*, 5(1):278–315, 2014.

[30] U. Horst, X. Xia, and C. Zhou. Portfolio liquidation under factor uncertainty. *Annals of Applied Probability*, 32(1):80–123, 2022.

[31] Y. Hu, B. Øksendal, and A. Sulem. Singular mean-field control games. *Stochastic Analysis and Applications*, 35(5):823–851, 2017.

[32] M. Kohlmann and S. Tang. Multidimensional backward stochastic Riccati equations and applications. *SIAM Journal on Control and Optimization*, 41(6):1696–1721, 2003.

[33] C. Lorenz and A. Schied. Drift dependence of optimal trade execution strategies under transient price impact. *Finance and Stochastics*, 17(4):743–770, 2013.

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

[35] A. Popier and C. Zhou. Second order BSDE under monotonicity condition and liquidation problem under uncertainty. *Annals of Applied Probability*, 29(3), 2019.

[36] P. E. Protter. *Stochastic Integration and Differential Equations.* Springer, 2005.

[37] A. Schied and T. Schöneborn. Liquidation in the face of adversity: Stealth vs. sunshine trading. *ssrn.1007014*, 2009.