Finance and Stochastics (2025) 29:553–607
https://doi.org/10.1007/s00780-025-00560-w

Fast and slow optimal trading with exogenous information

Rama Cont[1] · Alessandro Micheli[2] · Eyal Neuman[2]

Received: 3 March 2023 / Accepted: 16 April 2024 / Published online: 19 March 2025

Abstract
We model the interaction between an investor executing trades at low frequency and a high-frequency trader as a multiperiod stochastic Stackelberg game. The high-frequency trader exploits price information more frequently and is subject to periodic inventory constraints. We are able to explicitly compute the equilibrium strategies, in two steps. We first derive the optimal strategy of the high-frequency trader given any strategy adopted by the investor. Then we solve the problem of the investor given the optimal strategy of the high-frequency trader, in terms of the resolvent of a Fredholm integral equation. Our results show that the high-frequency trader adopts a predatory strategy whenever the value of the trading signal is high, and follows a cooperative strategy otherwise. We also show that there is a net gain in performance for the investor from taking into account the order flow of the high-frequency trader. A U-shaped intraday pattern in trading volume is shown to arise endogenously as a result of the strategic behaviour of the agents.

Keywords Market microstructure · High-frequency trading · Optimal stochastic control · Stochastic games · Price impact · Fredholm integral equations · Trading signals · Stackelberg equilibrium

Mathematics Subject Classification 45B05 · 49N70 · 49N90 · 91A65 · 93E20 · 60H30

JEL Classification C73 · C02 · C61 · G11

AM is supported by the EPSRC Centre for Doctoral Training in Mathematics of Random Systems: Analysis, Modelling and Simulation (EP/S023925/1).

✉ E. Neuman
e.neumann@imperial.ac.uk

R. Cont
rama.cont@maths.ox.ac.uk

A. Micheli
a.micheli19@imperial.ac.uk

1 Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

2 Department of Mathematics, Imperial College London, London, SW7 1NE, UK

1 Introduction

Modern financial markets involve a range of participants who place buy and sell orders across a wide spectrum of time scales: on the one end, pension funds and mutual funds rebalance typically on a monthly time scale, while on the other end of the spectrum, electronic market makers and high-frequency trading firms submit several thousands of orders per second (see Cont [10]), while having strict inventory constraints; see U.S. Securities and Exchange Commission [26, p. 4]. Although this heterogeneity in time scales has always been present, the development of computerised trading in electronic markets has substantially widened the range of frequencies at which various market participants operate. The interaction between the flow of buy and sell orders from these different participants results in an aggregate order flow which is the superposition of components across a wide range of frequencies. The consequences of this phenomenon for market volatility, price dynamics and market stability have yet to be systematically explored.

This heterogeneity of frequencies stands in contrast with mathematical models of market microstructure which are often formulated in terms of homogeneous agents operating at a single time scale; see Casgrain and Jaimungal [8], Evangelista and Thamsten [12], Fu et al. [14], Gârleanu and Pedersen [15], Micheli et al. [19], Neuman and Schied [20], Neuman and Voß [21], Voß [28]. However, the repeated occurrence of "flash crashes" (see e.g. Kirilenko et al. [16]) demonstrates that components at different frequencies may strongly interact and possibly lead to market disruption, calling for a modelling framework which incorporates the interaction of agents operating on different time scales.

A frequent narrative is the predatory nature of high-frequency traders (HFTs) towards institutional investors. Empirical evidence is less clear: while Korajczyk and Murphy [17] and Van Kervel and Menkveld [27] find evidence pointing to predatory behaviour of HFTs with respect to institutional investor order flow, other studies, such as Chen and Garriott [9], do not find any evidence of such predatory behaviour. These studies show that the strategic interaction between HFTs and institutional investors is more complex and may depend on several factors.

To investigate these phenomena, we propose a model for the dynamics of prices and order flow in a market where an institutional investor and a high-frequency trader observe a trading signal and submit buy and sell orders at different frequencies. We model the interaction between these two agents as a multiperiod stochastic game, where the institutional investor and the high-frequency trader interact through the aggregate order flow, which generates temporary and permanent impact on the asset price. The high-frequency trader exploits the trading signal at each period and is continuously subject to inventory constraints, while the institutional investor has limited access to the signal and is only subject to a terminal inventory constraint. Given the asymmetry between the agents, it is natural to model this setting in terms of a Stackelberg game, where the HFT, the minor agent, takes advantage of the signal and the order flow generated by the institutional investor, the major agent.

Our first result describes in Theorem 3.2 the unique optimal strategy of the high-frequency trader in response to any fixed strategy adopted by the investor. The challenging part in establishing a Stackelberg equilibrium is to derive the strategy of the major agent, i.e., the one who plays first. We develop a novel approach for this class of Stackelberg games in order to derive the major agent's optimal strategy given the optimal signal-adaptive strategy of the minor agent, using tools from the theory of integral equations. Specifically, in Theorem 3.5, we describe the optimal strategy for the major agent in terms of the resolvent of a Fredholm integral equation, thus establishing the unique multiperiod Stackelberg equilibrium of the game. In Sect. 4, we describe properties of the solutions to the Stackelberg game, and in Sect. 5, we derive the additional technical steps that are needed in order to obtain such explicit results directly from Theorems 3.2 and 3.5.

We propose a novel method using tools from the theory of integral equations to solve for the equilibrium. This allows us to explicitly derive the equilibrium of this multiperiod Stackelberg game, an atypical situation in such models which are usually intractable; see e.g. Carlin et al. [4], Schöneborn and Schied [25] and Roşu [24]. These analytical and computational aspects may be of independent interest and are detailed in Sects. 5 and 10 and in Appendix A, where we describe the numerical method used to compute the optimal strategy of the major agent and provide the proof of its convergence.

From our theoretical results, we obtain explicit expressions for the agents' optimal strategies and derive interesting economic insights regarding the behaviour of high-frequency traders and best practices for institutional investors executing large trades (see Sect. 4):

(i) Our results suggest that the high-frequency trader adopts a predatory strategy whenever the value of the trading signal is high, and follows a cooperative strategy otherwise. In the language of [27], we find that HFTs "trade with the wind" when the wind is strong, i.e., when trading signals convey strong information about future price directions, but "lean against the wind" otherwise. Figure 1 below shows examples of such strategies.

(ii) Comparing the revenue of the investor's optimal order execution with the case where the agent is not taking into account the HFT's trading activity, we show that taking the HFT's impact into account leads to a net gain in performance for the investor, as shown in Fig. 5 below. In fact, the investor's optimal strategy displays on average a substantial improvement in performance. This contrasts with the belief that high-frequency traders' order flow can be regarded as a "noise" whose effect on low-frequency strategies tends to "average out".

(iii) We show that the well-known U-shaped intraday profile for trading volume arises endogenously in our model as a result of the strategic behaviour of institutional investors and HFTs: the optimal trading strategy requires agents to trade more near the open and the close, as observed in Fig. 6 below.

The emergence of predatory trading as an optimal strategy was studied by Carlin et al. [4] in a single-period setting and by Schöneborn and Schied [25] in a two-period model. Carlin et al. [4] studied a single-period model with two types of revenue-maximising agents: *sellers* who start with a positive amount of assets and *competitors* who have zero initial positions. The (Nash) equilibrium strategy is described as follows: if the seller is liquidating, then the competitor is first selling and later buying back her position due to inventory constraints. In the two-period model [25], the seller can liquidate only in the first period, while the competitor can execute her strategy over two periods. Depending on the magnitude of the price impact, there are two possible scenarios: either the competitor is buying in the first period and then selling in the second period, i.e., introducing cooperative strategies in the game [25, Fig. 8], or doing a round trip of selling first and then closing the position, all in the first period.

Our model has some critical differences with Schöneborn and Schied [25]: we assume the minor agent (HFT) is trading at a higher frequency than the major agent. This is reflected in the model via periodic inventory constraints for the HFT. The minor agent is also reacting continuously to exogenous information, while the major agent has access to the information only at the beginning of the trade. An important implication is that in contrast to [25], the minor agent's optimal portfolio is stochastic.

Another major difference with these models is the type of equilibrium. The results of [25] focus on an open-loop Nash equilibrium, which means that all traders optimise synchronously. A more realistic representation of a heterogeneous market with participants intervening at different frequencies and latencies is to consider a *Stackelberg* equilibrium where the minor agent reacts to the major agent's order flow.

As neither Carlin et al. [4] nor Schöneborn and Schied [25] take into account exogenous trading signals, their optimal strategies are found to be deterministic. One of the main conclusions of our analysis is that this aspect has a crucial effect on the behaviour of the major agent and the minor agent, which is not captured in [4] and [25].

Last but not least, despite the multiple ingredients in our model, we are able to derive explicit solutions in the multiperiod case, allowing further analytical insights into the structure of the equilibrium strategies, in contrast to [25] where only the two-period model is tractable.

Roşu [24] also studied a discrete-time model where fast traders, whose decisions depend on a market signal, trade simultaneously with slow traders, who observe a lagged version of the signal. However, in contrast to our setting, fast traders do not have different objectives or inventory constraints, which are key ingredients in our model.

Finally, in Cartea and Sánchez-Betancourt [6], a game between a broker who provides liquidity to an informed trader and to a noise trader was studied. In order to decouple the optimal execution problem of the informed trader from the broker's strategy, it was assumed that the informed trader's cost functional contains a log-likelihood ratio (also referred to as score function) between the probability measure of the unaffected asset price and a probability measure from a class in which the asset price has an additional drift term, representing a conjectured broker's strategy. In the present paper, we take a different approach in which both the institutional investor's and the high-frequency trader's problems are fully coupled (see Definition 2.1).

The emergence of U-shaped patterns in trading volume is well documented; see Wood et al. [30]. Admati and Pfleiderer [1] provide a partial explanation of intraday concentration as a result of strategic interaction between informed traders and "noise" traders. Our results sharpen the results of [1] by providing an endogenous mechanism for the U-shaped intraday pattern in trading volume.

The paper is structured as follows. Section 2 defines the ingredients of the model. Our main results regarding the explicit solution to the Stackelberg game are presented in Sect. 3. Section 4 contains the illustrations and the financial interpretation of the main results. In Sect. 5, we rigorously derive the numerical scheme that we have used in order to plot the solutions in Sect. 4. The proofs are given in Sects. 6–10 as well as in Appendices A–C.

2 Trading fast and slow: a Stackelberg game

We consider a market with two participants: an investor who is liquidating an initial amount of shares in a risky asset, and a high-frequency trader (HFT) who trades in the same asset. We model the interaction between these two agents as a Stackelberg game; in the terminology of Stackelberg games, the investor is the "major agent" and the HFT is the "minor agent".

We represent market scenarios over a horizon $T > 0$ through a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ satisfying the usual conditions of right-continuity and completeness. Let $\mathcal{H}^2$ be the class of all (special) semimartingales $P = (P_t)_{t\in[0,T]}$ whose canonical decomposition $P = M + A$ into a (local) martingale $M = (M_t)_{t\in[0,T]}$ and a predictable finite-variation process $A = (A_t)_{t\in[0,T]}$ satisfies

$$\mathbb{E}[\langle M\rangle_T] + \mathbb{E}\bigg[\bigg(\int_0^T |dA_s|\bigg)^2\bigg] < \infty. \tag{2.1}$$

We denote by $L^2([0,T])$ the space of square-integrable functions $f : [0,T] \to \mathbb{R}$ and by $\langle\cdot,\cdot\rangle_{L^2}$ the inner product on $L^2([0,T])$, that is,

$$\langle f, g\rangle_{L^2} = \int_0^T f(t)g(t)dt, \qquad f, g \in L^2([0,T]),$$

and by $\|\cdot\|_{L^2}$ the associated norm.

The major agent has an initial holding of $q_0 \in \mathbb{R}$ shares in a risky asset. Her trading rate $\nu^0 = (\nu^0_t)_{t\in[0,T]}$ is chosen from the class of fuel-constrained *deterministic* admissible strategies $\mathcal{A}^{q_0}_M$, which is defined as

$$\mathcal{A}^{q_0}_M := \bigg\{\nu : \nu \in L^2([0,T]) \text{ such that } \int_0^T \nu_t dt = q_0\bigg\}. \tag{2.2}$$

Her trading rate ν^0 affects her inventory process Q^{0,ν^0} so that

$$Q^{0,\nu^0}_t = q_0 - \int_0^t \nu^0_s ds, \qquad 0 \le t \le T. \tag{2.3}$$

The minor agent, being a proprietary high-frequency trader, is assumed to have a zero initial position in the risky asset. Her trading rate $\nu^1 = (\nu^1_t)_{t\in[0,T]}$ is chosen from a class of *adaptive* admissible strategies given by

$$\mathcal{A}_m := \bigg\{\nu : \nu \text{ progressively measurable such that } \mathbb{E}\bigg[\int_0^T \nu^2_s ds\bigg] < \infty\bigg\}. \tag{2.4}$$

Her trading rate ν^1 affects her inventory process Q^{1,ν^1} so that

$$Q_t^{1,\nu^1} = -\int_0^t \nu_s^1 ds, \qquad 0 \le t \le T. \tag{2.5}$$

Throughout, we use the notation $\nu = (\nu^0, \nu^1)$ **for the major agent's control** ν^0 **and the minor agent's control** ν^1. Once ν is fixed, the visible asset mid-price P^ν satisfies

$$P_t^\nu = P_t - Y_t^\nu, \qquad 0 \le t \le T, \tag{2.6}$$

where $P \in \mathcal{H}^2$ and Y^ν is the permanent price impact à la Almgren and Chriss [2], which is generated by both agents and given by

$$Y_t^\nu = \int_0^t (\kappa_0 \nu_s^0 + \kappa_1 \nu_s^1) ds, \tag{2.7}$$

where κ_i, $i = 1, 2$, are positive constants.

The major agent's execution price is affected instantaneously in an adverse manner through the presence of a linear temporary price impact. The major agent's execution price is taken to be

$$S_t^{0,\nu} = P_t^\nu - \lambda_0 \nu_t^0, \tag{2.8}$$

where λ_0 is a positive constant measuring the magnitude of her temporary price impact. As a result, the major agent's cash holdings satisfy

$$X_t^{0,\nu} = x_0 + \int_0^t S_s^{0,\nu} \nu_s^0 ds, \qquad 0 \le t \le T. \tag{2.9}$$

The major agent's objective is to optimally unwind her initial position q_0 by the trading horizon T to minimise her execution costs. This is equivalent to maximising the expected revenues from her liquidation; we therefore take the major agent's performance functional to be

$$H^0(\nu^0; \nu^1) := \mathbb{E}[X_T^{0,\nu}]. \tag{2.10}$$

Note that the linear temporary price impact in (2.8) translates into quadratic trading costs with respect to the trading speed ν^0 in the major agent's cost functional in (2.10).

As in the case of the major agent, the transactions of the minor agent create a temporary price impact so that the execution price of her orders is given by

$$S_t^{1,\nu} = P_t^\nu - \lambda_1 \nu_t^1, \tag{2.11}$$

where λ_1 is a positive constant. Note that the temporary price impact parameter is likely to be smaller for the minor agent as HFTs can take advantage of the order-book real-time information in order to reduce their price impact. The minor agent's cash process is given by

$$X_t^{1,\nu} = x_1 + \int_0^t S_s^{1,\nu} \nu_s^1 ds, \qquad 0 \le t \le T. \tag{2.12}$$

The minor agent wishes to maximise the amount of cash she accumulates over the interval $[0, T]$. However, as an HFT, she is inclined to avoid *overnight risk*, specifically in the form of nonzero overnight inventory. As an example, consider T to be one business week so that $[0, T]$ can be partitioned in five disjoint and contiguous intervals of equal duration τ, where each interval represents the market hours of each business day from Monday to Friday. Without loss of generality, we assume in the context of our example that the minor agent's intraday risk preferences are independent of the business day considered, and we ignore the possibility of after-hours trading. Since the minor agent wishes to close her position by the end of each day, then as often done for terminal inventory penalties in the context of single-day liquidations, we can introduce a penalisation for nonzero inventory at the end of each day. These dynamic inventory preferences can be accounted for by modelling the running inventory costs of the minor agent via a periodic function of period τ which drastically increases towards the end of each day (see e.g. (4.6) below), i.e., as t approaches $\tau, 2\tau, 3\tau, 4\tau, 5\tau$ from the left. Mathematically, in order to capture the minor agent's dynamic inventory preferences of our example and more general ones, we define the minor agent's running inventory costs in terms of a function $\phi^1 : [0, T] \to \mathbb{R}_+$ which we take to be piecewise continuous and locally bounded.

The minor agent's risk–revenue functional is therefore given by

$$H^1(\nu^1; \nu^0) := \mathbb{E}\Big[X_T^{1,\nu} + Q_T^{1,\nu^1}(P_T^{\nu} - \alpha Q_T^{1,\nu^1}) - \int_0^T \phi_t^1 (Q_t^{1,\nu^1})^2 dt \Big]. \tag{2.13}$$

The first two terms in (2.13) represent the trader's terminal wealth, that is, her final cash position, accounting for the accrued trading costs which are induced by the temporary price impact and the permanent price impact of both agents as prescribed in (2.11), as well as the mark-to-market value of her terminal risky asset position. The fourth and third terms in (2.13) implement a penalty $\phi_t^1 > 0$ and $\alpha > 0$ on her running and terminal inventory, respectively. Also observe that $H^1(\nu^1; \nu^0) < \infty$ for any pair of admissible strategies $\nu^0 \in \mathcal{A}_M^{q_0}$ and $\nu^1 \in \mathcal{A}_m$.

We formulate the competition between the major and the minor agent as a stochastic Stackelberg game in which the minor agent is reacting to the major agent's trading. Mathematically, the game unfolds in two steps:

(i) *Minor agent's problem:* For a given major agent's liquidation strategy $\nu^0 \in \mathcal{A}_M^{q_0}$, the minor agent chooses her own strategy $\nu^{1,*}(\nu^0) \in \mathcal{A}_m$ in order to maximise her objective functional H^1.

(ii) *Major agent's problem:* Given the optimal minor agent's strategy $\nu^{1,*}$ established in (i), the major agent determines the optimal liquidation strategy $\nu^{0,*} \in \mathcal{A}_M^{q_0}$ in order to maximise her objective functional H^0.

In the context of our model, we formalise the definition of a Stackelberg equilibrium as follows.

Definition 2.1 A pair $\nu^* := (\nu^{0,*}, \nu^{1,*}(\nu^{0,*}))$ where $\nu^{0,*}$ and $\nu^{1,*}(\nu^{0,*})$ solve the major and minor agent's problems, respectively, is called a *Stackelberg equilibrium*.

In order to determine the Stackelberg equilibrium of Definition 2.1, we need to perform the following two steps:

(i) Solve the minor agent's stochastic control problem for any fixed admissible major agent strategy ν^0. This will determine the minor agent's optimal strategy $\nu^{1,*}(\nu^0)$ for any given strategy of the major agent.

(ii) Solve the major agent's optimisation problem assuming that the minor agent is adopting the optimal strategy $\nu^{1,*}(\nu^0)$ obtained in step (i). This will determine the major agent's optimal strategy $\nu^{0,*}$.

Then by Definition 2.1, the pair $(\nu^{0,*}, \nu^{1,*}(\nu^{0,*}))$ determined through the steps just described is a Stackelberg equilibrium of the game.

3 Optimal strategies

Our main analytical result is an explicit characterisation of the unique Stackelberg equilibrium of the game. As stated in Sect. 2, we start by solving the minor agent's problem.

3.1 Optimal strategy of the minor agent

We denote by $L^2([0,T]^2)$ the space of measurable kernels $\mathcal{T} : [0,T]^2 \to \mathbb{R}$ such that

$$\int_0^T \int_0^T \mathcal{T}(t,s)^2 dt ds < \infty. \tag{3.1}$$

Henceforth, we make the following assumption.

Assumption 3.1 We assume that the parameters α in (2.13) and κ_1 in (2.7) are chosen such that

$$2\alpha \geq \kappa_1.$$

Let $r^1 = (r^1_t)_{t\in[0,T]}$ solve a Riccati equation with time-varying coefficient ϕ^1, given by

$$\partial_t r^1_t = \frac{1}{\lambda_1}\phi^1_t - (r^1_t)^2, \qquad r^1_T = -\frac{2\alpha - \kappa_1}{2\lambda_1}. \tag{3.2}$$

Under Assumption 3.1, the solution r^1 of (3.2) exists and is unique over $[0,T]$ (see Proposition 6.5 below). We further define

$$\xi^{\pm}_t := e^{\pm\int_0^t r^1_z dz}, \qquad 0 \leq t \leq T, \tag{3.3}$$

as well as the kernel $\mathcal{K} : [0,T]^2 \to \mathbb{R}_+$ which is given by

$$\mathcal{K}(t,s) := \xi^-_t \xi^+_s, \qquad 0 \leq t, s \leq T. \tag{3.4}$$

Note that the kernel $\mathcal{K}$ is in $L^2([0,T]^2)$ (see Lemma 6.7 below). Moreover, for any $\nu^0 \in \mathcal{A}^{q_0}_M$, we define the progressively measurable process

$$r^0_t := \frac{1}{2\lambda_1}\mathbb{E}_t\bigg[\int_t^T \mathcal{K}(t,s)(dA_s - \kappa_0 \nu^0_s ds)\bigg], \qquad 0 \leq t \leq T. \tag{3.5}$$

The solution to the minor agent's problem is given in the following result.

Theorem 3.2 *Let $\nu^0 \in \mathcal{A}_M^{q_0}$. Under Assumption 3.1, there exists a unique optimal strategy $\nu^{1,*}(\nu^0) \in \mathcal{A}_m$ that maximises (2.13). This strategy is given by*

$$\nu_t^{1,*} = -\left(r_t^0 + r_t^1 \int_0^t \mathcal{K}(s,t) r_s^0 ds\right), \qquad 0 \le t \le T. \tag{3.6}$$

The proof of Theorem 3.2 is given in Sect. 6.

Remark 3.3 In Lemma 6.2 below, we show that Assumption 3.1 is a sufficient condition to guarantee the strict concavity of the minor agent's functional (2.13), and hence the uniqueness of the solution to the minor agent's problem.

Remark 3.4 The minor agent's optimal strategy in (3.6) may be written in feedback form as

$$\nu_t^{1,*} = -\left(r_t^0 + r_t^1 Q_t^{1,\nu^{1,*}}\right), \qquad 0 \le t \le T.$$

3.2 Optimal strategy of the major agent

Our next step is to derive the maximiser of the major agent's objective functional (2.10), given the minor agent's optimal strategy $\nu^{1,*}$ in (3.6). As is often the case in Stackelberg games, solving the second phase of the game is technically challenging and rarely achievable. In order to do so, we make the following simplifying assumption on the signal A in (2.1). We assume that the signal process A is given by

$$A_t = \int_0^t \mu_s ds, \qquad 0 \le t \le T,$$

where $\mu = (\mu_t)_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-progressively measurable stochastic process satisfying

$$\int_0^T \mathbb{E}[\mu_t^2] dt < \infty. \tag{3.7}$$

This is an adaptation to the present context of assumptions in previous studies by Cartea and Jaimungal [5], Lehalle and Neuman [18] in single-agent optimal execution problems. We further denote

$$\bar{\mu}_t := \mathbb{E}[\mu_t], \qquad 0 \le t \le T. \tag{3.8}$$

We now introduce some essential definitions regarding linear operators in $L^2([0,T])$. For any given linear operator $\boldsymbol{T} : L^2([0,T]) \to L^2([0,T])$, we define the operator norm

$$\|\boldsymbol{T}\| := \sup\{\|\boldsymbol{T}\psi\|_{L^2} : \psi \in L^2([0,T]), \|\psi\|_{L^2} \le 1\}. \tag{3.9}$$

Furthermore, we denote by $B(L^2([0,T]))$ the space of all bounded linear operators from $L^2([0,T])$ to $L^2([0,T])$ with respect to the operator norm (3.9).

For any kernel $\mathcal{T} \in L^2([0,T]^2)$ (see (3.1)), we say that $\boldsymbol{T}$ is the integral operator generated by the kernel $\mathcal{T}$ if for any $\psi \in L^2([0,T])$,

$$(\boldsymbol{T}\psi)(t) = \int_0^T \mathcal{T}(t,s)\psi(s)ds, \qquad 0 \le t \le T.$$

Any integral operator generated by a kernel in $L^2([0,T]^2)$ is in $B(L^2([0,T]))$ by the Cauchy–Schwarz inequality. For two operators $\boldsymbol{T}_1, \boldsymbol{T}_2 \in B(L^2([0,T]))$, we denote by $\boldsymbol{T}_2\boldsymbol{T}_1$ the operator obtained by composing $\boldsymbol{T}_2$ with $\boldsymbol{T}_1$, that is, for any $\psi \in L^2([0,T])$,

$$(\boldsymbol{T}_2\boldsymbol{T}_1\psi)(t) := \big(\boldsymbol{T}_2(\boldsymbol{T}_1\psi)\big)(t), \qquad 0 \le t \le T.$$

We now define some special operators for our setting. Recall that $\mathcal{K}$ was defined in (3.4). We introduce the kernel $\mathcal{G} : [0,T]^2 \to \mathbb{R}_+$ defined as

$$\mathcal{G}(t,s) := \int_0^{t\wedge s} \mathcal{K}(u,t)\mathcal{K}(u,s)du, \qquad 0 \le t,s \le T. \tag{3.10}$$

Note that the kernel $\mathcal{G}$ is symmetric and in $L^2([0,T]^2)$ (see Proposition 8.1 below). We define the operators $\boldsymbol{G}$ and $\boldsymbol{S}$ acting on any $\psi \in L^2([0,T])$ via

$$(\boldsymbol{G}\psi)(t) := \int_0^T \mathcal{G}(t,s)\psi(s)ds, \tag{3.11}$$

$$(\boldsymbol{S}\psi)(t) := \frac{1}{2\lambda_0}\int_0^T \mathbb{1}_{\{s\le t\}}\psi(s)ds + \frac{\kappa_1}{4\lambda_1\lambda_0}(\boldsymbol{G}\psi)(t). \tag{3.12}$$

Note that both operators $\boldsymbol{G}$ and $\boldsymbol{S}$ are in $B(L^2([0,T]))$ (see Proposition 7.4 and Lemma 7.10 below). Moreover, the operator $\boldsymbol{G}$ admits a spectral decomposition in terms of a sequence $(\zeta_n)_{n\ge1}$ of positive eigenvalues and a corresponding sequence $(\psi_n)_{n\ge1}$ of eigenfunctions in $L^2([0,T])$ (see Lemma 8.3 below). It is convenient to define the *resolvent* kernel $\mathcal{R} : [0,T]^2 \to \mathbb{R}$ as

$$\mathcal{R}(t,s) = -\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathcal{G}(t,s) + \sum_{n\ge1}\frac{1}{1+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\zeta_n}\left(\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\zeta_n\right)^2\psi_n(t)\psi_n(s) \tag{3.13}$$

for all $t,s \in [0,T]$ and where the sum converges uniformly and uniformly-absolutely over $[0,T]^2$; see Remark 3.10 below for details. Moreover, we define the *resolvent* operator $\boldsymbol{R}$ acting on any $\psi \in L^2([0,T])$ via

$$(\boldsymbol{R}\psi)(t) := \psi(t) + \int_0^T \mathcal{R}(t,s)\psi(s)ds, \qquad 0 \le t \le T. \tag{3.14}$$

The operator $\boldsymbol{R}$ is also in $B(L^2([0,T]))$; this is proved in Proposition 7.9 below. Finally, we denote by $\mathbf{1}(t)$ the constant function which equals 1 everywhere on $[0,T]$.

We are ready to state our main result regarding the solution to the major's agent problem conditional on the minor agent adopting the strategy $\nu^{1,*}$ given in (3.6). Recall that $\bar{\mu}$ was defined in (3.8) and $\boldsymbol{S}$ in (3.12).

Theorem 3.5 *Assume that $\nu^{1,*}$ is given by* (3.6) *and that Assumption* 3.1 *holds. Then there exists a unique optimal strategy $\nu^{0,*} \in \mathcal{A}_M^{q_0}$ that maximises the major agent's objective functional* (2.10). *It is given by*

$$\nu_t^{0,*} = \frac{\eta}{2\lambda_0}(\boldsymbol{R}\mathbf{1})(t) + (\boldsymbol{R}\boldsymbol{S}\bar{\mu})(t), \qquad 0 \le t \le T, \tag{3.15}$$

where

$$\eta = 2\lambda_0 \frac{q_0 - \langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}}{\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2}}. \tag{3.16}$$

Moreover, $t \mapsto \nu_t^{0,*}$ is continuous on $[0, T]$.

The proof of Theorem 3.5 is given in Sect. 7. In that proof, we also show that the constant η in (3.16) is well defined, which is an ingredient in proving the admissibility of the optimal strategy (3.15).

The following result follows immediately from Theorems 3.2 and 3.5.

Corollary 3.6 *Let $\nu^{0,*}$ and $\nu^{1,*}(\nu^{0,*})$ be as in Theorems* 3.5 *and* 3.2, *respectively. Then under Assumption* 3.1, *the pair $(\nu^{0,*}, \nu^{1,*}(\nu^{0,*})) \in \mathcal{A}_M^{q_0} \times \mathcal{A}_m$ is the unique Stackelberg equilibrium in the sense of Definition* 2.1.

The following remarks discuss the result of Corollary 3.6.

Remark 3.7 Note that $\nu^{0,*}$ in (3.15) is given in terms of the resolvent operator $\boldsymbol{R}$. In Sect. 5, we derive a numerical scheme that approximates $\nu^{0,*}$ by using finite-dimensional projections of $\boldsymbol{G}$. The problem of computing $\boldsymbol{R}$ and hence $\nu^{0,*}$ is reduced to a finite-dimensional problem of matrix inversion. We refer to Proposition 5.4 and Theorem 5.6 below for the details.

Remark 3.8 The most challenging step in obtaining a Stackelberg equilibrium is to derive the strategy of the player who acts first, namely the major agent. In our case, we needed to develop a novel approach for deriving the optimal strategy in (3.15), using tools from the theory of integral equations. In Sect. 4, we illustrate the solutions to the Stackelberg game, and in Sect. 5, we derive additional technical steps which are needed in order to plot such explicit solutions directly from Theorems 3.2 and 3.5.

Remark 3.9 Our illustrations in Sect. 4 suggest that the minor agent can adopt either a predatory or a cooperative strategy with respect to the major agent, in each period, depending on the tradeoff between the order flow of the major agent and the trading signal during the period (see Fig. 1 below). This qualitative behaviour can be compared with the deterministic model of Schöneborn and Schied [25] who showed that in the single-period case, the competitor is also selling and then buying back her position due to inventory constraints (see [25, Fig. 1]). In their two-period model, the seller trades only over the first period and she adopts a selling strategy. Depending on the price impact parameters, there are two possible scenarios: either the competitor is buying in the first period and then selling in the second period, i.e., introducing cooperative strategies in the game (see [25, Fig. 8]), or doing a round trip of selling first and then closing the position, all in the first period.

Remark 3.10 We remark that the sum appearing in (3.13) satisfies certain convergence properties. Define

$$\mathfrak{R}_N(t,s) = \sum_{n=1}^{N} \frac{1}{1+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\zeta_n}\left(\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\zeta_n\right)^2 \psi_n(t)\psi_n(s), \qquad t,s\in[0,T],$$

$$\mathfrak{R}_N^{\mathrm{abs}}(t,s) = \sum_{n=1}^{N} \left|\frac{1}{1+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\zeta_n}\left(\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\zeta_n\right)^2 \psi_n(t)\psi_n(s)\right|, \qquad t,s\in[0,T].$$

Then it follows from the proof of Porter and Stirling [23, Theorem 4.27] that $\mathfrak{R}_N$ converges uniformly to the sum in (3.13) on $(t,s)\in[0,T]^2$, and that $\mathfrak{R}_N^{\mathrm{abs}}$ is uniformly convergent. The uniform convergence of $\mathfrak{R}_N^{\mathrm{abs}}$ guarantees that the uniform convergence of $\mathfrak{R}_N$ is preserved even when the order of summation is changed. Therefore, as it is natural to expect, the solution $\nu^{0,*}$ in (3.15) is independent of how one enumerates the eigenvalues and the corresponding eigenfunctions of $\boldsymbol{G}$.

4 Examples

In this section, we further describe the equilibrium strategies derived in Theorems 3.5 and 3.2. Motivated by Lehalle and Neuman [18, Sect. 4], we consider the case where the signal μ in (3.7) follows an Ornstein–Uhlenbeck process,

$$d\mu_t = -\beta\mu_t dt + \sigma dW_t, \qquad \mu_0 = m_0, \tag{4.1}$$

where $W=(W_t)_{t\geq 0}$ is a standard Brownian motion and β and σ are positive constants. Furthermore, we assume that M in (2.1) is given by

$$M_t = M_0 + \sigma_M \widetilde{W}_t, \tag{4.2}$$

where $\widetilde{W}$ is a standard Brownian motion independent from W and M_0, σ_0 are positive constants. We fix the values of the price impact parameters λ_i, κ_i, the initial inventory q_0 of the major agent and the terminal penalty parameter α in (2.13) as

$$\kappa_0 = 2, \quad \kappa_1 = 2, \quad \lambda_0 = 1, \quad \lambda_1 = 1, \quad q_0 = 10, \quad \alpha = 10, \tag{4.3}$$

and the parameters of M in (4.2) and μ in (4.1) as

$$m_0 = -0.5, \quad \beta = 0.1, \quad \sigma = 4, \quad M_0 = 100, \quad \sigma_M = 1. \tag{4.4}$$

Our model allows different price impact parameters for the major and minor agents. While there is anecdotal evidence supporting the notion that HFTs can incur lower transaction costs than institutional investors, our numerical examples employ the same values for the price impact parameters for both major and minor agents. This decision stems from the limited research on identifying the actual values of these parameters for different market participants, as most transaction data remains anonymised. Consequently, we adopt a conservative approach and fix the price impact parameters to the same values for both major and minor agents. Note that although temporary and permanent price impact effects are at the core of our model, small variations of the price impact constants will not result in drastically different behaviours of the major and minor agents' optimal strategies. This is because other effects, such as the minor agent's periodic inventory costs ϕ^1, are also significantly contributing to the major and minor agent's optimal strategies.

The plots in this section are generated by using the numerical scheme described in detail in Sect. 5 below. We choose as a complete orthonormal basis $(a_i)_{i=1}^{\infty}$ of $L^2([0,T])$ the functions

$$a_i(t) := \begin{cases} 1/\sqrt{T}, & i = 1, \\ \sqrt{2/T}\cos(\frac{(i-1)\pi t}{T}), & i = 2, 3, \dots, \end{cases} \tag{4.5}$$

so that each corresponding degenerate kernel $\mathcal{G}_n$ defined in (5.3) represents the nth degree Fourier series approximation of the kernel $\mathcal{G}$ in (3.10). In order to strike a balance between numerical accuracy and computational efficiency, our simulations are generated by approximating the kernel $\mathcal{G}$ with the degenerate kernel $\mathcal{G}_{300}$.

The time-dependence in the minor agent's inventory costs ϕ^1 (see (2.13)) can accommodate the setting of a liquidation carried out over several days. Let $T = k\tau$ for some positive integer k and for $\tau > 0$. Moreover, we choose the function ϕ^1 to be given by the parametric form

$$\phi_t^1 = c_0\left(\frac{t}{\tau} - \left\lfloor \frac{t}{\tau} \right\rfloor\right)^{c_1}, \qquad 0 \le t \le T, \tag{4.6}$$

for two positive constants c_0 and c_1 which in the context of our simulations, we take to be $c_0 = 500$ and $c_1 = 15$. The function (4.6) is periodic of period τ and increases to its maximum value as t approaches $\tau, 2\tau, 3\tau, \dots, k\tau$ from the left, forcing the minor agent to liquidate most of her position at the end of each period. We consider a liquidation carried over a business week, from Monday to Friday, so that $T = 5$ (days) and $\tau = 1$ (day). Figure 1 illustrates three examples of a multi-day liquidation. Specifically, the top panel shows the major agent's deterministic optimal inventory (green line), deduced from (3.15), as well as three different realisations of the minor agent's optimal inventories that one obtains from (3.6) (blue, purple and red lines). The bottom panel shows the corresponding signal μ observed by the minor agent while adopting the strategies at the top panel.

From (2.6), it follows that the price impact generated by the major agent's optimal strategy $\nu^{0,*}$ is perceived as a deterministic signal. The sell-off of shares by the major agent has the effect of pushing the price downwards; it therefore generates opportunities which can be exploited by the minor agent. These considerations justify the fact that as shown in (3.6), the minor agent adopts a trading strategy which tracks the "impacted" signal $\mu_t - \kappa_0\nu_t^{0,*}$ instead of the raw market signal μ_t. Hence depending on

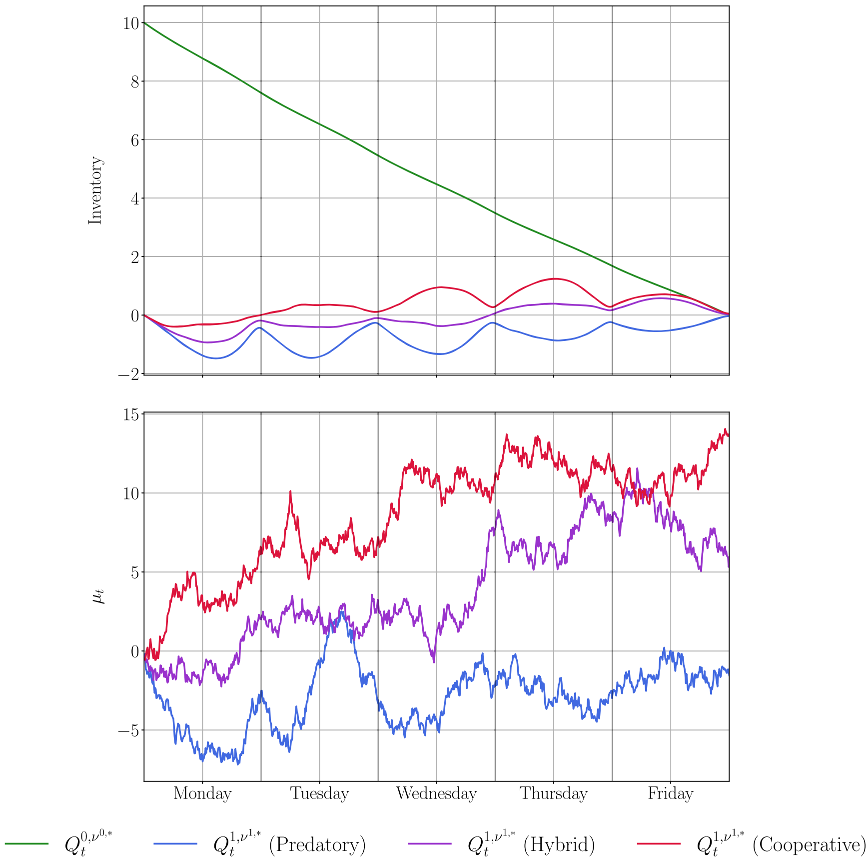

Fig. 1 In the top panel, the green line represents the major agent's optimal inventory, while the remaining solid lines represent the minor agent's optimal inventory when the minor agent is adopting a predatory trading style (blue line) or a cooperative trading style (red line) or a hybrid of both (purple line). In the bottom panel, we show the signal (μ_t) corresponding to the realisations of the minor agent's inventories in the top panel

her forecast on the impacted signal $\mu_t - \kappa_0 \nu_t^{0,*}$, during each period, the minor agent can decide whether to trade in the same direction of the major agent or not. This has the effect that over the interval $[0, T]$, the observed trading style of the minor agent can be *predatory*, i.e., front running the major agent (blue line), *cooperative* (red line) or a *hybrid* of both (purple line).

To further understand several novel features of our model in the context of the multi-day liquidation we have just analysed, it is instructive to momentarily pause our discussion and consider the simpler case of a liquidation carried out over a single day. In particular, we wish to benchmark the major agent's optimal strategy in (3.15) against the strategy $\nu^{0,\mathrm{BM}}$ the major agent would use if she were unaware of the minor agent's trading activity. The strategy $\nu^{0,\mathrm{BM}}$ can be found by solving the major agent's problem with $\kappa_1 = 0$, and it is given by

$$\nu_t^{0,\mathrm{BM}} = \frac{q_0}{T} + \frac{m_0}{2\lambda_0\beta}\,\frac{1 - \beta T e^{-\beta t} - e^{-\beta T}}{\beta T}, \qquad 0 \le t \le T. \tag{4.7}$$

Note that in the case of $m_0 = 0$ in (4.1), $\nu^{0,\mathrm{BM}}$ in (4.7) is a TWAP strategy.

We assume that the major agent wishes to liquidate his initial position over a time horizon of six hours, from 10 AM to 4 PM; hence we set $T = 6$ (hours).

In the present context, we slightly modify some of the parameters in (4.3) and (4.4) to $\sigma = 1.5$, $\alpha = 50$ and $\phi^1 \equiv 1$. The top panel of Fig. 2 shows the major agent's optimal trading rate $\nu^{0,*}$ (solid green line) and the benchmark trading rate $\nu^{0,\mathrm{BM}}$ (dashed green line). The bottom panel shows 1'000 realisations of the minor agent's

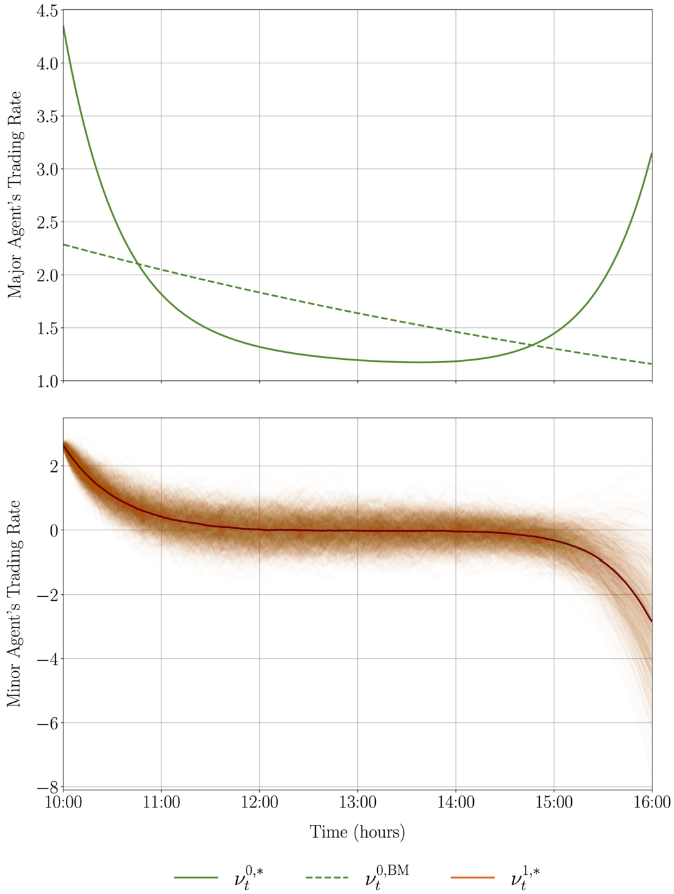

Fig. 2 The major and minor agent's optimal strategies in (3.15) and (3.6), respectively, for a single-day liquidation. In the top panel, the solid green line shows the major agent's optimal strategy, while the dashed green line shows the benchmark strategy of (4.7). In the bottom panel, the thin solid orange lines depicts different realisations of the minor agent's optimal strategy. The solid brown line is the cross-sectional mean over the realisations

optimal trading rate $\nu^{1,*}$ (thin solid orange lines) and the cross-sectional average (thick solid brown line). We observe that the major agent's optimal strategy visibly deviates from the benchmark one in order to take into account the adverse effect of the minor agent's trading activity. We remark that since the major agent adopts a deterministic strategy, her decisions are based on the cross-sectional average of the minor agent's strategy, i.e., the solid brown line in the bottom panel of Fig. 2. Initially, it is optimal to trade faster than the benchmark strategy in anticipation of the expected permanent price impact generated by the minor agent's reaction. Indeed, the early prices are more favourable to the major agent since they have not yet been affected by the extra price impact generated by the presence of the minor agent. In the middle of the trading window, the major agent keeps trading, but at a lower rate than the benchmark strategy. The explanation for this is that the major agent is aware that the minor agent could potentially trade in the same direction. Therefore, slowing down partially minimises the negative externality the minor agent's exerts on her via the aggregated permanent price impact. Finally, in the last section of the trading window, two factors determine the behaviour of the major agent's optimal strategy. First, the major agent must increase her trading rate to meet the terminal inventory constraint $Q_T^{0,\nu^{0,*}} = 0$. Second, the major agent is aware that on average, the minor agent will have to close her short position at the end of the time horizon; therefore she will have to buy shares, generating a market impact and pushing the price up again. Hence the prices at the end of the trading window are more favourable for the major agent, and therefore a substantial portion of the liquidation is postponed to the last hour. Figure 3 presents the major and minor agent's inventories corresponding to the trading rates depicted in Fig. 2.

Having established the major and minor agent's trading patterns in the context of a single-day liquidation, we turn again to the case of the multi-day liquidation presented in Fig. 1. Analogously to Fig. 2, the top panel of Fig. 4 shows the major agent's optimal trading rate (solid green line) as well as the trading rate of the benchmark strategy (dashed green line) in the context of the multi-day liquidation initially presented in Fig. 1. Moreover, the bottom panel of Fig. 4 presents 1'000 realisations of the minor agent's trading rates (thin orange lines) as well as the cross-sectional average (solid brown line). We recover analogous trading patterns to the one observed in the single-day liquidation: the major agent's speed, when compared to the benchmark strategy, greatly increases at the beginning and at the end of each day. Moreover, on average, the minor agent acquires a short position at the beginning of each day, pushing the price down, and then, in order to meet her terminal inventory constraint at the end of each day, she pushes the price up again by buying shares. Note from Fig. 1 that the predatory, cooperative and hybrid strategies share some common features. First, at the end of each day, all the strategies have a very small inventory. This is because by introducing the periodic running inventory costs of (4.6), the minor agent is strongly discouraged to hold a nonzero position at the end of each day, independently of her forecast for the impacted signal $\mu_t - \kappa_0 \nu_t^{0,*}$. Secondly, from Fig. 1, we observe that the major agent is not liquidating at a constant speed. Indeed, over the first day, she liquidates at a speed visibly larger than for example the one employed over the last day. Such an intense liquidation in the first day generates an equally large alpha-signal through the corresponding price impact term $\kappa_0 \nu_t^{0,*}$. In the

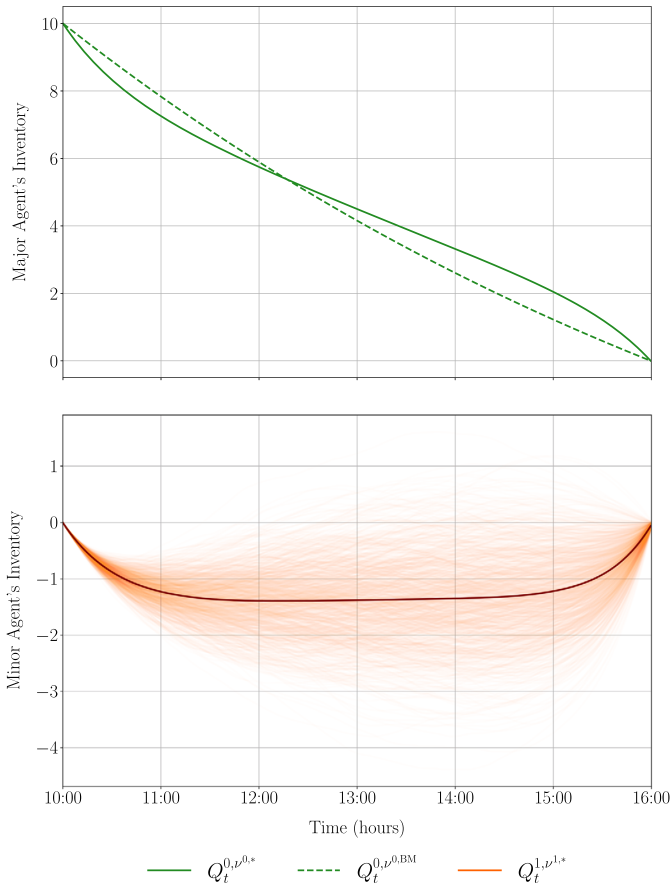

Fig. 3 The major and minor agent's optimal inventories corresponding to the strategies in (3.15) and (3.6), respectively, for a single-day liquidation. In the top panel, the solid green line shows the major agent's optimal inventory corresponding to the major agent's optimal strategy, while the dashed green line shows the inventory corresponding to the benchmark strategy in (4.7). In the bottom panel, the thin solid orange lines represent different realisations of the minor agent's optimal inventories corresponding to the strategy in (3.6), while the solid brown line is the cross-sectional mean over the realisations

first day, the market-impact-generated signal $\kappa_0 \nu_t^{0,*}$ is large enough to outweigh any realistic realisation of the exogenous signal μ_t, therefore pushing the minor agent to trade in the same direction as the major agent, independently of the trading style she will adopt later on in the remaining days.

It is of practical interest to compare the financial performance of the major agent's optimal strategy $\nu^{0,*}$ against that of the benchmark strategy $\nu^{0,\mathrm{BM}}$. In the interest of brevity, we limit ourselves to the case of the single-day liquidation presented in Fig. 3. In Fig. 5, we present a histogram of the empirical probability distribution of the performance of the major agent's optimal strategy in (3.15) relative to the benchmark strategy in (4.7), generated by using 1'000 simulations. We compare the

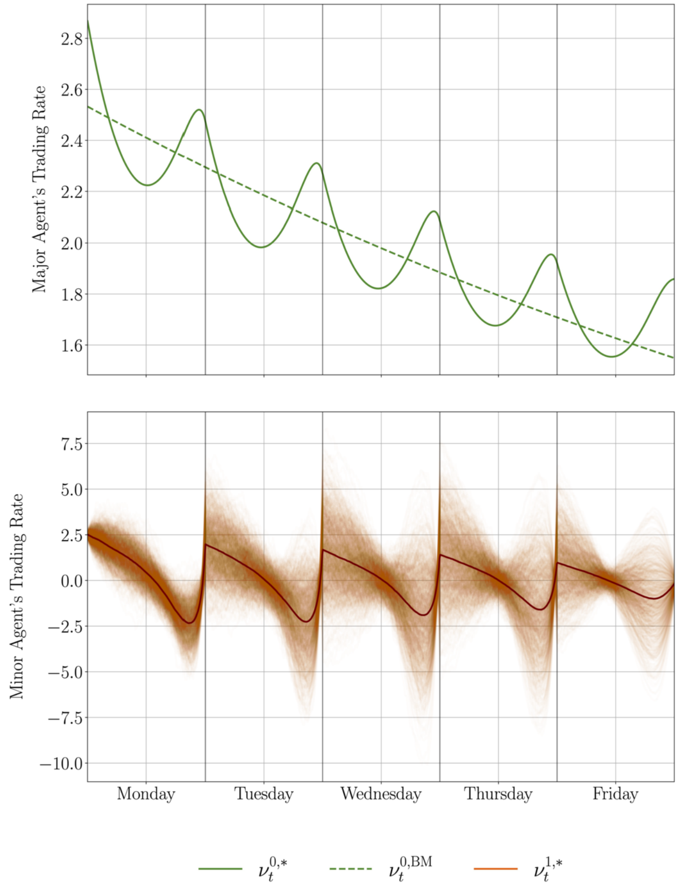

Fig. 4 We present the major and minor agent's optimal strategy in (3.15) and (3.6), respectively, for a multi-day liquidation. In the top panel, the solid green line shows the major agent's optimal strategy, while the dashed green line shows the benchmark strategy of (4.7). In the bottom panel, the thin solid orange lines depict different realisations of the minor agent's optimal strategy, while the solid brown line represents the cross-sectional mean over the realisations

profit-and-loss (PnL) of the strategies in basis points (bps) through the formula

$$\frac{X_T^{0,\nu^{0,*}} - X_T^{0,\nu^{0,\mathrm{BM}}}}{X_T^{0,\nu^{0,\mathrm{BM}}}} \times 10^4, \tag{4.8}$$

where $X_T^{0,\nu^{0,\mathrm{BM}}}$ is the terminal cash obtained from employing the benchmark strategy $\nu^{0,\mathrm{BM}}$ and $X_T^{0,\nu^{0,*}}$ is the terminal cash obtained from employing the optimal strategy $\nu^{0,*}$. The mean of the distribution in Fig. 5 is strictly positive; hence the major agent's optimal strategy on average outperforms the benchmark strategy.

Finally, we show that the strategic behaviour of the agents induces intraday patterns in trading volume. It is well documented that the intraday trading volume dis-

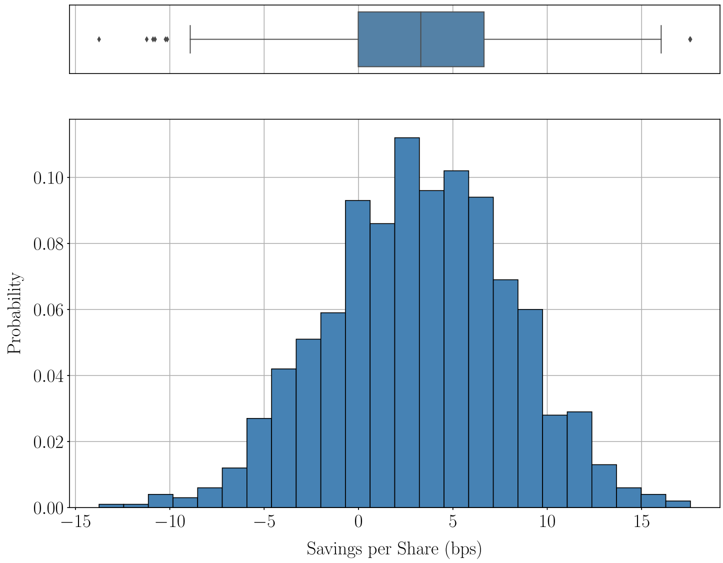

Fig. 5 The savings per share, computed using (4.8) and measured in bps, from following the major agent's optimal strategy relative to the benchmark strategy in (4.7). The top panel shows the box plot corresponding to the distribution in the bottom panel

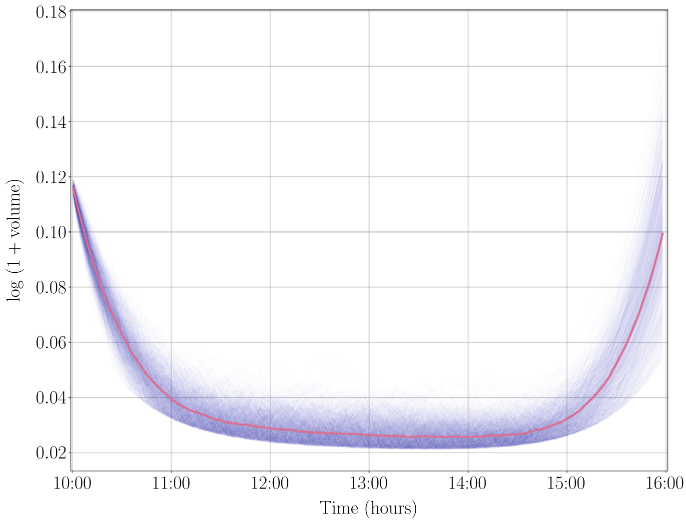

Fig. 6 Intraday volume profiles. The median is displayed in pink

plays a "U-shaped" intraday profile: volume is highest at the beginning and at the end of the day; see Wood et al. [30]. Figure 6 displays the intraday volume profile resulting from our model. We compute the total number of shares traded by the agents for each 1-minute bin, averaged over 1'000 sample paths. Each blue line in Fig. 6 represents a simulated path for the (logarithmic) volume, while the magenta line is the median value of each 1-minute bin. The volume profiles in Fig. 6 display a U-shaped pattern analogous for example to those empirically observed and reported in previous studies; see [1, 30].

5 Computing the optimal strategies

Theorem 3.5 presents the major agent's unique optimal strategy $\nu^{0,*}$ in closed form. The optimal strategy $\nu^{0,*}$ is expressed in terms of the resolvent operator $\boldsymbol{R}$, defined in (3.13), which in turn relies on the eigenvalues $(\zeta_n)_{n\geq 1}$ and eigenfunctions $(\psi_n)_{n\geq 1}$ of the operator $\boldsymbol{G}$ defined in (3.11). In several simple cases, these eigenfunctions and eigenvalues can be computed explicitly.

For example, in the case of $\phi^1 \equiv 0$, $(\zeta_n)_{n\geq 1}$ and $(\psi_n)_{n\geq 1}$ can be explicitly determined in terms of the roots of a transcendental equation (see Appendix A below). Nevertheless, a closed-form representation for the eigenvalues $(\zeta_n)_{n\geq 1}$ and the eigenfunctions $(\psi_n)_{n\geq 1}$ might be unattainable when ϕ^1 is a generic nonnegative piecewise continuous function. Therefore we dedicate this section to developing a numerical scheme to compute the major agent's optimal strategy $\nu^{0,*}$ which fully bypasses the need of determining these eigenvalues and eigenfunctions. As a byproduct, this numerical scheme also determines the Stackelberg equilibrium of Corollary 3.6.

We denote by $\boldsymbol{I}$ the identity operator on $L^2([0,T])$, that is,

$$(\boldsymbol{I}\psi)(t) = \psi(t) \qquad \text{for all } 0 \leq t \leq T,\ \psi \in L^2([0,T]).$$

As usual, the resolvent operator in (3.14) can be written as

$$\boldsymbol{R} = \left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)^{-1}.$$

This is proved rigorously in Proposition 7.9 below. It follows that the major agent's optimal strategy $\nu^{0,*}$ in (3.15) satisfies the integral operator equation

$$\left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)\nu^{0,*} = \boldsymbol{S}\bar{\mu} + \frac{\eta}{2\lambda_0}. \tag{5.1}$$

Here, the constant η is defined in (3.16) and the operators $\boldsymbol{G}$ and $\boldsymbol{S}$ are defined in (3.11) and (3.12), respectively. The idea is to replace (5.1) with a sequence of approximate equations (see (5.6) below) whose solutions converge to the desired optimal strategy $\nu^{0,*}$.

For the discussion that follows, it is convenient to recall the definitions of a finite-rank and of a compact operator; see Definition 7.2. In Proposition 7.4, we show that the operator $\boldsymbol{G}$ is compact. Therefore there exists a sequence $(\boldsymbol{G}_n)_{n\geq 1}$ of finite-rank operators in $B(L^2([0,T]))$ satisfying the approximation property

$$\lim_{n\to\infty} \|\boldsymbol{G}_n - \boldsymbol{G}\| = 0,$$

where $\|\cdot\|$ refers to the operator norm in (3.9). In order to construct such a sequence, we consider a complete orthonormal basis $(a_i)_{i=1}^{\infty}$ in $L^2([0,T])$. A possible choice of such a complete orthonormal basis in $L^2([0,T])$ is given by (4.5). Let the kernel $\mathcal{G}$ be defined in (3.10) and let the functions $(b_i)_{i=1}^{\infty}$ be defined as

$$b_i(t) := \int_0^T \mathcal{G}(t,s)a_i(s)ds. \tag{5.2}$$

We recall the definition of a degenerate kernel (see Porter and Stirling [23, Definition 3.1]) which will be useful in the following.

Definition 5.1 Let $n \geq 1$ and suppose there are finitely many functions $(a_i)_{i=1}^{n}$ and $(b_i)_{i=1}^{n}$ such that $a_i : [0,T] \to \mathbb{R}$ and $b_i : [0,T] \to \mathbb{R}$ for $i = 1,\dots,n$. Assume further that $\mathcal{T} : [0,T]^2 \to \mathbb{R}$ is a kernel such that

$$\mathcal{T}(t,s) = \sum_{i=1}^{n} a_i(t)b_i(s), \qquad t,s \in [0,T].$$

Then the kernel $\mathcal{T}$ is said to be *degenerate*.

Define the sequence $(\mathcal{G}_n)_{n\geq 1}$ of degenerate kernels as the partial sums

$$\mathcal{G}_n(t,s) := \sum_{i=1}^{n} a_i(t)b_i(s), \qquad n \geq 1. \tag{5.3}$$

Since $\mathcal{G}$ is a kernel in $L^2([0,T]^2)$ (see Proposition 8.1 below), then as shown in [23, proof of Theorem 3.4], the sequence $(\mathcal{G}_n)_{n\geq 1}$ converges to $\mathcal{G}$ in the sense that

$$\lim_{n\to\infty} \int_0^T \int_0^T \big(\mathcal{G}(t,s) - \mathcal{G}_n(t,s)\big)^2 ds dt = 0. \tag{5.4}$$

Given the degenerate kernels $(\mathcal{G}_n)_{n\geq 1}$, we can define a corresponding sequence of so-called finite-rank integral operators $(\boldsymbol{G}_n)_{n\geq 1}$ as

$$(\boldsymbol{G}_n\psi)(t) := \int_0^T \mathcal{G}_n(t,s)\psi(s)ds, \qquad \psi \in L^2([0,T]). \tag{5.5}$$

The following result, which is proved in Sect. 10, gives the convergence for the sequence $(\boldsymbol{G}_n)_{n\geq 1}$.

Proposition 5.2 *Under Assumption 3.1, let $(\boldsymbol{G}_n)_{n\geq 1}$ be defined as in (5.5) and $\boldsymbol{G}$ as in (3.11). Then the finite-rank operators $\boldsymbol{G}_n$ are in $B(L^2([0,T]))$. Moreover, we have*

$$\lim_{n\to\infty} \|\boldsymbol{G}_n - \boldsymbol{G}\| = 0.$$

Next we consider the sequence of equations approximating (5.1) given by

$$\left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n\right)\nu^{0,(n)} = \boldsymbol{S}\bar{\mu} + \frac{\eta_n}{2\lambda_0}, \qquad n \geq 1, \tag{5.6}$$

for a suitably defined sequence $(\eta_n)_{n\geq 1}$ of constants.

Remark 5.3 We remark that in (5.6), we continue to take the operator $\boldsymbol{S}$ to be defined in terms of $\boldsymbol{G}$ as in (3.12), and not in terms of the sequence $(\boldsymbol{G}_n)_{n\geq 1}$. It is not necessary to approximate the operator $\boldsymbol{S}$ since it can be explicitly expressed in terms of the kernel $\mathcal{G}$ in (3.10) via the operator $\boldsymbol{G}$; therefore it can be computed explicitly via numerical integration (see also Remark 5.8 below).

A solution to (5.6) exists if the inverse of the operator $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n$ exists, with the candidate solution $\nu^{0,(n)}$ then given by

$$\nu^{0,(n)} = \left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n\right)^{-1}\left(\boldsymbol{S}\bar{\mu} + \frac{\eta_n}{2\lambda_0}\right). \tag{5.7}$$

The next result shows that for sufficiently large n, the inverse of $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n$ exists. Moreover, we show that the problem of finding that inverse is reduced to a finite-dimensional problem of matrix inversion.

To state our results, it is convenient to introduce the sequence $(\mathbb{G}_n)_{n\geq 1}$ of matrices, where $\mathbb{G}_n \in \mathbb{R}^{n\times n}$ has entries defined as

$$(\mathbb{G}_n)_{ij} := \langle a_i, b_j\rangle_{L^2} \tag{5.8}$$

for all $1 \leq i, j \leq n$ and with a_i and b_j defined as in (5.2). Moreover, we denote by $I_n := \mathrm{diag}(1, \ldots, 1) \in \mathbb{R}^{n\times n}$ the n-dimensional identity matrix. We are now ready to state our next result, which is proved in Sect. 10.

Proposition 5.4 *Under Assumption 3.1, let $(\boldsymbol{G}_n)_{n\geq 1}$ be defined as in (5.5) and $(\mathbb{G}_n)_{n\geq 1}$ as in (5.8). Then there exists $N \geq 1$ such that for all $n \geq N$, the operator $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n$ and the matrix $I_n + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathbb{G}_n$ are both invertible. In particular, we have for all $n \geq N$ that for any $\psi \in L^2([0, T])$,*

$$\left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n\right)^{-1}\psi = \psi - \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\sum_{i,j=1}^{n}\left(I_n + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathbb{G}_n\right)^{-1}_{i,j}\langle\psi, b_j\rangle_{L^2}\, a_i. \tag{5.9}$$

Note that both operators $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ and $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n$ are invertible (see Proposition 7.9 below). However, only for the latter can the inverse operator be computed via matrix inversion by exploiting the corresponding degenerate kernel decomposition; see also Remark 5.8 below for additional discussion.

The next result shows that the candidate solutions in (5.7) converge *in mean* to the optimal strategy $\nu^{0,*}$ of Theorem 3.5. Henceforth, we take the sequence $(\eta_n)_{n\geq 1}$ of constants to be defined as

$$\eta_n := 2\lambda_0 \frac{q_0 - \langle(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n)^{-1}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}}{\langle(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n)^{-1}\mathbf{1}, \mathbf{1}\rangle_{L^2}}, \qquad n \geq 1. \tag{5.10}$$

Proposition 5.5 *Under Assumption 3.1, let $\nu^{0,*}$ and $\nu^{0,(n)}$ be defined as in (3.15) and (5.7), respectively. Then there exists $N \geq 1$ such that for all $n \geq N$, the functions $\nu^{0,(n)}$ are well defined and in $L^2([0,T])$. Moreover,*

$$\lim_{n\to\infty} \|\nu^{0,*} - \nu^{0,(n)}\|_{L^2} = 0.$$

The proof of Proposition 5.5 is postponed to Sect. 10. Lemma 10.4 below shows that for sufficiently large n, the constants η_n in (5.10) are well defined.

In order to obtain an approximating sequence which converges *uniformly* to the optimal control $\nu^{0,*}$, we introduce the sequence of candidate functions $(\hat{\nu}^{0,(n)})_{n\geq 1}$ defined as

$$\hat{\nu}_t^{0,(n)} := -\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}(\boldsymbol{G}\nu^{0,(n)})(t) + (\boldsymbol{S}\bar{\mu})(t) + \frac{\eta_n}{2\lambda_0} \tag{5.11}$$

for all $t \in [0,T]$ and $n \geq 1$. Our main result for this section is the following convergence theorem.

Theorem 5.6 *Under Assumption 3.1, let $\nu^{0,*}$, $\hat{\nu}^{0,(n)}$ and $\nu^{1,*}$ be defined as in (3.15), (5.11) and (3.6), respectively. Then there exists $N \geq 1$ such that for all $n \geq N$, the functions $\hat{\nu}^{0,(n)}$ are in $L^2([0,T])$ and the controls $\nu^{1,*}(\hat{\nu}^{0,(n)})$ are in $\mathcal{A}_m$. Furthermore, we have that*

(i) $\lim_{n\to\infty} \sup_{t\in[0,T]} |\nu_t^{0,*} - \hat{\nu}_t^{0,(n)}| = 0;$

(ii) $\lim_{n\to\infty} \sup_{t\in[0,T]} |\nu_t^{1,*}(\nu^{0,*}) - \nu_t^{1,*}(\hat{\nu}^{0,(n)})| = 0 \qquad \mathbb{P}$*-a.s.*

The proof of Theorem 5.6 is postponed to Sect. 10.

Proposition 5.4 and Theorem 5.6 show that the infinite-dimensional problem of determining the solution to (5.1) can be reduced to a finite-dimensional problem of matrix inversion.

Remark 5.7 The proofs of Proposition 5.5 and Theorem 5.6 do not rely on the existence of the orthonormal expansion in (5.3) and the corresponding convergence (5.4). Indeed, our result can be extended to any generic sequence $(\boldsymbol{G}_n)_{n\geq 1}$ of operators in $B(L^2([0,T]))$ which satisfy the approximation property of Proposition 5.2, but do not necessarily enjoy an integral representation of the form (5.5).

Remark 5.8 The matrix entries in (5.8) must be computed numerically. The use of a numerical evaluation in (5.8) will lead to numerical errors in the entries of the matrix $I_n + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathbb{G}_n$. As shown in Atkinson [3, Chap. 2.3.4], for a sufficiently accurate computation of the entries of the matrix in (5.8), this numerical error is negligible. A similar discussion applies among others to the numerical evaluation of the inverse of the matrix $I_n + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathbb{G}_n$ and of the integral $\boldsymbol{S}\bar{\mu}$. These are all elementary and well-understood convergence problems in numerical analysis, and the corresponding convergence rates could be easily incorporated in the convergence results of this section. Hence our discussion assumes that the aforementioned quantities are taken to be exact and that the corresponding numerical errors are negligible.

Remark 5.9 The numerical scheme we have presented has an advantage from an implementation standpoint, too. Specifically, if one were to determine the major agent's optimal strategy by using the result of Theorem 3.5, she would need to mathematically determine the eigenvalues $(\zeta_n)_{n\geq 1}$ and eigenfunctions $(\psi_n)_{n\geq 1}$ each time she wishes to change the function ϕ^1, as shown for example in Appendix A. On the other hand, with the numerical scheme of Theorem 5.6, to achieve the same goal, it is sufficient to change the expression of ϕ^1 in the numerical solver of the Riccati ODE (3.2), which usually amounts to changing only a few lines of code.

6 Proof of Theorem 3.2

We show how the Stackelberg equilibrium can be found by *backward induction*, that is, by first solving the minor agent's problem and then the major agent's problem. We determine the minor agent's optimal strategy via a calculus of variations argument, as similarly done in Neuman and Voß [22]. The following results also borrow ideas from Casgrain and Jaimungal [7].

Henceforth, we assume that $\nu^0 \in \mathcal{A}_M^{q_0}$ is a fixed liquidation strategy for the major agent, and with a slight abuse of notation, we write $H^1(\nu)$ for $H^1(\nu;\nu^0)$. We start by determining an alternative representation for the minor agent's objective.

Lemma 6.1 *The minor agent's objective H^1 in (2.13) can be alternatively represented, for any $\nu^1 \in \mathcal{A}_m$, as*

$$\begin{aligned} H^1(\nu^1) = x_1 - \mathbb{E}\Big[\lambda_1 \int_0^T (\nu_t^1)^2 dt + \alpha (Q_T^{1,\nu^1})^2 + \int_0^T \phi_t^1 (Q_t^{1,\nu^1})^2 dt \\ + \int_0^T Q_t^{1,\nu^1}(\kappa_0 \nu_t^0 dt + \kappa_1 \nu_t^1 dt - dA_t)\Big]. \end{aligned} \tag{6.1}$$

Proof We use (2.12), (2.11) and Itô's product rule on $Q_T^{1,\nu^1} P_T^{\nu}$ to get

$$\begin{aligned} &\mathbb{E}[X_T^{1,\nu^1} + Q_T^{1,\nu^1} P_T^{\nu}] \\ &= x_1 + \mathbb{E}\Big[\int_0^T (P_t^{\nu} - \lambda_1 \nu_t^1)\nu_t^1 dt + \int_0^T Q_t^{1,\nu^1} dP_t^{\nu} + \int_0^T P_t^{\nu} dQ_t^{1,\nu^1}\Big], \end{aligned} \tag{6.2}$$

where we also used $Q_0^{1,\nu^1} = 0$ by (2.5). Recall that $P = M + A$. We apply (2.5)–(2.7) to (6.2) in order to obtain

$$\begin{aligned} \mathbb{E}[X_T^{1,\nu^1} + Q_T^{1,\nu^1} P_T^{\nu}] &= x_1 + \mathbb{E}\Big[-\lambda_1 \int_0^T (\nu_t^1)^2 dt + \int_0^T Q_t^{1,\nu^1} dP_t^{\nu}\Big] \\ &= x_1 - \mathbb{E}\Big[\lambda_1 \int_0^T (\nu_t^1)^2 dt - \int_0^T Q_t^{1,\nu^1} dM_t \\ &\quad + \int_0^T Q_t^{1,\nu^1}(\kappa_0 \nu_t^0 dt + \kappa_1 \nu_t^1 dt - dA_t)\Big]. \end{aligned} \tag{6.3}$$

Since $\nu^1 \in \mathcal{A}_m$ (see (2.4)), we can drop the martingale term in (6.3) and obtain

$$\mathbb{E}[X_T^{1,\nu^1} + Q_T^{1,\nu^1} P_T^{\nu}] = x_1 - \mathbb{E}\Big[\lambda_1 \int_0^T (\nu_t^1)^2 dt + \int_0^T Q_t^{1,\nu^1}(\kappa_0 \nu_t^0 dt + \kappa_1 \nu_t^1 dt - dA_t)\Big]. \quad (6.4)$$

Substituting (6.4) in (2.13) returns (6.1). □

In the next result, we use the representation (6.1) to show that the minor agent's objective is strictly concave.

Lemma 6.2 *Under Assumption 3.1, the functional H^1 defined in (2.13) is strictly concave for $\nu^1 \in \mathcal{A}_m$.*

Proof To prove that H^1 is strictly concave, we must show that for any $0 < \rho < 1$ and $\nu, \omega \in \mathcal{A}_m$ which are $(d\mathbb{P} \otimes dt)$-distinguishable, it holds that

$$\mathcal{I}^1(\rho, \nu, \omega) := H^1\big(\rho\nu + (1-\rho)\omega\big) - \rho H^1(\nu) - (1-\rho)H^1(\omega) > 0. \quad (6.5)$$

It is convenient to introduce the constant $\theta := \frac{2\alpha - \kappa_1}{2}$ and the function Γ^1 via

$$\Gamma_t^1 = \begin{pmatrix} \lambda_1 & -\theta \\ -\theta & \phi_t^1 \end{pmatrix}, \qquad 0 \le t \le T. \quad (6.6)$$

Note that under Assumption 3.1, we have $\theta \ge 0$. From (2.5) and integration by parts, we get

$$(Q_T^{1,\nu^1})^2 = -2\int_0^T Q_t^{1,\nu^1} \nu_t^1 dt. \quad (6.7)$$

Using (6.7), we rewrite the minor agent's objective in (6.1) in terms of Γ^1 as

$$H^1(\nu^1) = x_1 - \mathbb{E}\Big[\int_0^T \begin{pmatrix} \nu_t^1 \\ Q_t^{1,\nu^1} \end{pmatrix}^{\top} \Gamma_t^1 \begin{pmatrix} \nu_t^1 \\ Q_t^{1,\nu^1} \end{pmatrix} dt + \int_0^T Q_t^{1,\nu^1}(\kappa_0 \nu_t^0 dt - dA_t)\Big]. \quad (6.8)$$

Note that given the representation in (6.8), the strict concavity of $H^1(\nu^1)$ follows in the case of $\theta > 0$ and $\phi_t^1 > 0$ for all $t \in [0, T]$ from (2.5) and the fact that Γ_t^1 is a positive definite matrix for all $t \in [0, T]$. In what follows, we use (6.8) to show that $H^1(\nu^1)$ is strictly concave also if we only have $\phi_t^1 \ge 0$ and $\theta \ge 0$.

We observe that $Q^{1,\nu}$ is linear with respect to ν, that is,

$$Q_t^{1,\rho\nu+(1-\rho)\omega} = \rho Q_t^{1,\nu} + (1-\rho)Q_t^{1,\omega} \qquad \text{for all } \rho \in [0,1],\ \nu, \omega \in \mathcal{A}_m.$$

We substitute (6.8) in (6.5) and use the linearity of $Q^{1,\cdot}$ to cancel out the terms $Q_t^{1,\cdot}(\kappa_0 \nu_t^0 dt - dA_t)$. This yields

$$\begin{aligned}
&\mathcal{I}^1(\rho, \nu, \omega) \\
&= \mathbb{E}\Big[\int_0^T \rho \begin{pmatrix} \nu_t \\ Q_t^{1,\nu} \end{pmatrix}^\top \Gamma_t^1 \begin{pmatrix} \nu_t \\ Q_t^{1,\nu} \end{pmatrix} dt + (1-\rho) \begin{pmatrix} \omega_t \\ Q_t^{1,\omega} \end{pmatrix}^\top \Gamma_t^1 \begin{pmatrix} \omega_t \\ Q_t^{1,\omega} \end{pmatrix} dt \\
&\qquad - \Big(\rho\begin{pmatrix} \nu_t \\ Q_t^{1,\nu} \end{pmatrix} + (1-\rho)\begin{pmatrix} \omega_t \\ Q_t^{1,\omega} \end{pmatrix}\Big)^\top \Gamma_t^1 \Big(\rho\begin{pmatrix} \nu_t \\ Q_t^{1,\nu} \end{pmatrix} + (1-\rho)\begin{pmatrix} \omega_t \\ Q_t^{1,\omega} \end{pmatrix}\Big) dt\Big],
\end{aligned}$$

and after multiplying out all the terms, we get

$$\begin{aligned}
&\mathcal{I}^1(\rho, \nu, \omega) \\
&= \mathbb{E}\Big[\int_0^T \rho(1-\rho)\Big(\begin{pmatrix} \nu_t \\ Q_t^{1,\nu} \end{pmatrix} - \begin{pmatrix} \omega_t \\ Q_t^{1,\omega} \end{pmatrix}\Big)^\top \Gamma_t^1 \Big(\begin{pmatrix} \nu_t \\ Q_t^{1,\nu} \end{pmatrix} - \begin{pmatrix} \omega_t \\ Q_t^{1,\omega} \end{pmatrix}\Big) dt\Big].
\end{aligned} \tag{6.9}$$

It is convenient to introduce the function $\delta_t := \nu_t - \omega_t$ for $t \in [0, T]$. From (2.5), it follows that $Q_t^{1,\delta} = Q_t^{1,\nu} - Q_t^{1,\omega}$ for $t \in [0, T]$. We can rewrite (6.9) in terms of δ and $Q^{1,\delta}$ as

$$\begin{aligned}
\mathcal{I}^1(\rho, \nu, \omega) = \rho(1-\rho)\Big(&\mathbb{E}\Big[\int_0^T \lambda_1 \delta_t^2 dt\Big] + \mathbb{E}\Big[\int_0^T \phi_t^1 (Q_t^{1,\delta})^2 dt\Big] \\
&- \mathbb{E}\Big[\int_0^T 2\theta \delta_t Q_t^{1,\delta} dt\Big]\Big),
\end{aligned}$$

where we have used (6.6). Since $\phi_t^1 \geq 0$ for $t \in [0, T]$, we have

$$\mathbb{E}\Big[\int_0^T \phi_t^1 (Q_t^{1,\delta})^2 dt\Big] \geq 0.$$

From (2.5), it holds that $Q_t^{1,\delta} = -\int_0^t \delta_t dt$; therefore (6.7) yields

$$-\mathbb{E}\Big[\int_0^T 2\delta_t Q_t^{1,\delta} dt\Big] = \mathbb{E}[(Q_T^{1,\delta})^2] \geq 0.$$

Finally, notice that since ν and ω are $(d\mathbb{P} \otimes dt)$-distinguishable, we have

$$\mathbb{E}\Big[\int_0^T \delta_t^2 dt\Big] > 0.$$

This shows that $\mathcal{I}^1(\rho, \nu, \omega) > 0$ for any $\theta \geq 0$, $0 < \rho < 1$ and $\nu, \omega \in \mathcal{A}_m$ which are $(d\mathbb{P} \otimes dt)$-distinguishable. □

As similarly shown in Neuman and Voß [22], a probabilistic and convex analytic calculus of variations approach can be readily applied to derive a system of coupled linear FBSDEs which characterises the unique solution to the minor agent's problem.

Since the map $\nu^1 \to H^1(\nu^1)$ in (6.1) is strictly concave under Assumption 3.1, it admits a unique maximiser characterised by the critical point at which the Gâteaux derivative

$$\langle \mathcal{D}H^1(\nu^1), \omega\rangle := \lim_{\epsilon\to 0} \frac{H^1(\nu^1+\epsilon\omega) - H^1(\nu^1)}{\epsilon} \tag{6.10}$$

vanishes. In the following result, we obtain an explicit expression for the Gâteaux derivative of H^1.

Lemma 6.3 *The Gâteaux derivative of H^1 in (6.1) in the direction $\omega \in \mathcal{A}_m$ is for any $\nu^1 \in \mathcal{A}_m$ given by*

$$\begin{aligned}\langle \mathcal{D}H^1(\nu^1), \omega\rangle = \mathbb{E}\bigg[\int_0^T \omega_t \bigg(& -2\lambda_1 \nu_t^1 + 2\alpha Q_T^{1,\nu^1} - \kappa_1 Q_t^{1,\nu^1} + A_t - A_T \\ & + \int_t^T (2\phi_s^1 Q_s^{1,\nu^1} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^1) ds\bigg) dt\bigg].\end{aligned} \tag{6.11}$$

The proof of Lemma 6.3 is given in Appendix B.

From the explicit expression of the Gâteaux derivative in (6.11), we can derive a first-order optimality condition. It takes the form of a coupled system of linear forward–backward stochastic differential equations (FBSDEs), as described in the following result which is proved in Appendix B.

Lemma 6.4 *Under Assumption 3.1, the control $\nu^{1,*} \in \mathcal{A}_m$ is the unique maximiser to the minor agent's objective functional H^1 in (6.1) if and only if $(Q^{1,\nu^{1,*}}, \nu^{1,*})$ satisfies the coupled linear FBSDE system*

$$\begin{cases} dQ_t^{1,\nu^{1,*}} = -\nu_t^{1,*} dt, \qquad Q_0^{1,\nu^{1,*}} = 0, \\ d\nu_t^{1,*} = \frac{1}{2\lambda_1} d\mathcal{N}_t + \frac{1}{2\lambda_1} d\mathcal{M}_t - \frac{\phi_t^1}{\lambda_1} Q_t^{1,\nu^{1,*}} dt - \frac{\kappa_0}{2\lambda_1}\nu_t^0 dt + \frac{1}{2\lambda_1} dA_t, \\ \nu_T^{1,*} = \frac{2\alpha - \kappa_1}{2\lambda_1} Q_T^{1,\nu^{1,*}} \end{cases} \tag{6.12}$$

$(d\mathbb{P}\otimes dt)$-a.e. on $\Omega\times[0,T]$, where $\mathcal{M} = (\mathcal{M}_t)_{t\in[0,T]}$ and $\mathcal{N} = (\mathcal{N}_t)_{t\in[0,T]}$ are two suitable square-integrable martingales.

For the remainder of this section, we focus on the derivation of the explicit solution to (6.12). We begin by describing some heuristics of the proof.

The solution to the FBSDE system (6.12) determines the solution to the minor agent's problem. The main obstacle in solving the system (6.12) is that it presents a general time-dependent coefficient ϕ_t^1. In order to solve this equation, we formulate an ansatz for the minor agent's optimal strategy $\nu^{1,*}$. Then we demonstrate that the ansatz solution for $\nu^{1,*}$ is the unique solution to (6.12) and therefore the solution to the minor agent's problem. Due to the linear structure of the system (6.12), we make the ansatz that there are two progressively measurable processes $r^0 = (r_t^0)_{t\in[0,T]}$ and $r^1 = (r_t^1)_{t\in[0,T]}$ such that $\nu^{1,*}$ can be expressed as

$$\nu_t^{1,*} = -(r_t^0 + r_t^1 Q_t^{1,\nu^{1,*}}), \qquad 0 \le t \le T. \tag{6.13}$$

We differentiate (6.13) via Itô's lemma and use $dQ_t^{1,\nu^{1,*}} = -\nu_t^{1,*}dt$ to get

$$d\nu_t^{1,*} = -dr_t^0 - Q_t^{1,\nu^{1,*}} dr_t^1 + \nu_t^{1,*} r_t^1 dt. \tag{6.14}$$

We plug (6.14) into (6.12) and arrive at

$$\begin{aligned} 0 = &\big(2\lambda_1 dr_t^1 - 2\phi_t^1 dt + 2\lambda_1 (r_t^1)^2 dt\big) Q_t^{1,\nu^{1,*}} \\ &+ (2\lambda_1 dr_t^0 + 2\lambda_1 r_t^1 r_t^0 dt - \kappa_0 \nu_t^0 dt + dA_t + d\mathcal{M}_t + d\mathcal{N}_t). \end{aligned} \tag{6.15}$$

Equation (6.15) must hold $(d\mathbb{P} \otimes dt)$-almost everywhere for all values $Q_t^{1,\nu^{1,*}}$. We conjecture that the terms within each bracket must vanish independently. The terms from (6.15) lead to two coupled differential equations for r^0 and r^1 independent of the process $Q^{1,\nu^{1,*}}$, where we determine the terminal conditions from (6.12). Specifically, the process r^1 must satisfy dt-a.e. the non-autonomous Riccati ODE

$$\partial_t r_t^1 = \frac{1}{\lambda_1}\phi_t^1 - (r_t^1)^2, \qquad r_T^1 = -\frac{2\alpha - \kappa_1}{2\lambda_1}, \tag{6.16}$$

while r^0 must satisfy the BSDE

$$-dr_t^0 = r_t^1 r_t^0 dt - \frac{\kappa_0}{2\lambda_1}\nu_t^0 dt + \frac{1}{2\lambda_1} dA_t + \frac{1}{2\lambda_1} d\mathcal{M}_t + \frac{1}{2\lambda_1} d\mathcal{N}_t, \qquad r_T^0 = 0. \tag{6.17}$$

An explicit formula for the solution of (6.16) does not exist when $t \mapsto \phi_t^1$ is a general piecewise continuous function as in the case at hand. Nevertheless, we prove that the solution to (6.16) exists and is unique. Once a solution to (6.16) is found, we can plug it into (6.17) and derive r^0.

In the following result, which is proved in Appendix C, we derive the existence and uniqueness of the solutions to (6.16) and (6.17).

Proposition 6.5 *Under Assumption 3.1, there exists a unique continuous function r^1 that satisfies the non-autonomous Riccati ODE (6.16) dt-a.e. on $[0, T]$. Furthermore, the BSDE (6.17) admits a closed-form solution r^0 given by (3.5). Moreover,*

$$\mathbb{E}\left[\int_0^T (r_t^0)^2 dt\right] < \infty. \tag{6.18}$$

Remark 6.6 As stated in Proposition 6.5, the function r^1 satisfies the Riccati ODE (6.16) only dt-almost everywhere. This is to be expected since ϕ^1 is assumed to be piecewise continuous and the derivatives of r^1 need not exist at points of discontinuity. Nevertheless, as we show in the proof of Theorem 3.2, this is sufficient for our needs as we wish to solve the FBSDE system (6.12) only $(d\mathbb{P} \otimes dt)$-almost everywhere.

In order to prove Theorem 3.2, we need the following result which is also proved in Appendix C.

Lemma 6.7 *Let r^1 be the unique solution of* (6.16) *and recall the kernel $\mathcal{K}$ from* (3.4). *Then $\mathcal{K}$ is jointly continuous over* $[0,T]^2$. *In particular, $\mathcal{K}$ is bounded over* $[0,T]^2$ *and is in* $L^2([0,T]^2)$.

We are now ready to prove Theorem 3.2. In order to simplify the notation, we often denote the process $\nu^{1,*}(\nu^0)$ by $\nu^{1,*}$.

Proof of Theorem 3.2 Let $\nu^{1,*}$ as in (3.6) with r^1 as in (6.16) and r^0 as in (3.5).

As a first step, we determine an explicit expression for $Q^{1,\nu^{1,*}}$. We argue that

$$Q_t^{1,\nu^{1,*}} = \int_0^t \mathcal{K}(s,t) r_s^0 ds. \tag{6.19}$$

It follows from (3.4) that verifying the expression in (6.19) is equivalent to verifying that

$$Q_t^{1,\nu^{1,*}} = \xi_t^+ \int_0^t \xi_s^- r_s^0 ds. \tag{6.20}$$

From (3.3), we note that ξ^+ satisfies the ODE

$$\frac{d\xi_t^+}{dt} = r_t^1 \xi_t^+, \qquad 0 \le t \le T. \tag{6.21}$$

Taking the derivative in (6.20) and using (3.3), (3.4) and (6.21), we arrive at

$$dQ_t^{1,\nu^{1,*}} = r_t^1\left(\xi_t^+ \int_0^t \xi_s^- r_s^0 ds\right)dt + r_t^0 dt = \left(r_t^0 + r_t^1 \int_0^t \mathcal{K}(s,t) r_s^0 ds\right)dt. \tag{6.22}$$

From (6.22) and (2.5), we get (6.19).

Secondly, we show that $\nu^{1,*}$ solves (6.12). Note that from (2.5) and (3.6), we can rewrite $\nu^{1,*}$ as

$$\nu_t^{1,*} = -(r_t^0 + r_t^1 Q_t^{1,\nu^{1,*}}), \qquad 0 \le t \le T. \tag{6.23}$$

By plugging (6.23) into (6.12), we conclude that it is enough to prove that (6.15) holds. Since r^0 satisfies (6.17) and r^1 satisfies (6.16) dt-a.e. on $[0,T]$, (6.15) holds $(d\mathbb{P}\otimes dt)$-almost everywhere. Next, using the terminal conditions of r^0 and r^1 from (6.17) and (6.16), we deduce from (6.23) that the terminal condition in (6.12) is satisfied. Notice that it follows from (2.5) that the differential equation for $Q^{1,\nu^{1,*}}$ in (6.12) is satisfied. Therefore $(Q^{1,\nu^{1,*}}, \nu^{1,*})$ solve the system (6.12) $(d\mathbb{P}\otimes dt)$-almost everywhere.

Finally, we show that $\nu^{1,*} \in \mathcal{A}_m$. From (2.4), it follows that we need to verify that

$$\mathbb{E}\left[\int_0^T (\nu_t^{1,*})^2 dt\right] < \infty.$$

From (6.19), Proposition 6.5, Lemma 6.7 and the Cauchy–Schwarz inequality, we get

$$\mathbb{E}\left[\int_0^T (Q_t^{1,\nu^{1,*}})^2 dt\right] \le \mathbb{E}\left[\int_0^T \left(\int_0^t \mathcal{K}(s,t)^2 ds\right)\left(\int_0^t (r_s^0)^2 ds\right) dt\right] < \infty.$$

By Proposition 6.5, r^1 is continuous, hence bounded, on $[0,T]$ and r^0 is square-integrable. Using this, (6.23) and the Cauchy–Schwarz inequality gives

$$\mathbb{E}\left[\int_0^T (\nu_t^{1,*})^2 dt\right] \le 2\mathbb{E}\left[\int_0^T (r_t^0)^2 dt\right] + 2\mathbb{E}\left[\int_0^T (r_t^1)^2 (Q_t^{1,\nu^{1,*}})^2 dt\right] < \infty.$$

Therefore $\nu^{1,*}$ is admissible and solves (6.12). Hence by Lemma 6.4, it is the unique maximiser to the minor agent's objective functional H^1 in (6.1). □

7 Proof of Theorem 3.5

In this section, we derive the major agent's optimal strategy via a calculus of variations argument. We start by defining operators which are essential to our proofs. Then we derive an equivalent representation of the major agent's objective H^0 which is more convenient for our method of proof. **Throughout this section**, we assume that Assumption 3.1 holds and that the minor agent is adopting the strategy $\nu^{1,*}$ from Theorem 3.2. **Henceforth**, with a slight abuse of notation, we write $H^0(\nu^0)$ for $H^0(\nu^0, \nu^{1,*}(\nu^0))$.

We now introduce some essential definitions regarding operators on $L^2([0,T])$. We denote by $\mathcal{T}^*$ the adjoint kernel of $\mathcal{T}$ for $\langle \cdot, \cdot \rangle_{L^2}$, that is,

$$\mathcal{T}^*(t,s) := \mathcal{T}(s,t), \qquad s,t \in [0,T], \tag{7.1}$$

and by $\boldsymbol{T}^*$ the corresponding adjoint integral operator. We define the kernel

$$\mathcal{K}_1 : [0,T]^2 \to \mathbb{R}_+$$

by

$$\mathcal{K}_1(t,s) := \mathcal{K}(s,t)\mathbb{1}_{\{s\le t\}}, \qquad s,t \in [0,T], \tag{7.2}$$

where $\mathcal{K}$ is given in (3.4). We let $\boldsymbol{K}_1$ be the integral operator generated by the kernel $\mathcal{K}_1$, that is,

$$(\boldsymbol{K}_1\psi)(t) := \int_0^T \mathcal{K}_1(t,s)\psi(s)ds, \qquad t \in [0,T], \psi \in L^2([0,T]). \tag{7.3}$$

The following result, which is proved in Sect. 8, outlines some useful properties of $\boldsymbol{K}_1$. Recall that the class $B(L^2([0,T]))$ of operators was defined after (3.9).

Lemma 7.1 *The operator $\boldsymbol{K}_1$ is in $B(L^2([0,T]))$. Moreover, $\boldsymbol{K}_1^* \in B(L^2([0,T]))$ is given by*

$$(\boldsymbol{K}_1^*\psi)(t) = \int_0^T \mathcal{K}(t,s)\psi(s)\mathbb{1}_{\{t\le s\}}ds, \qquad t \in [0,T], \psi \in L^2([0,T]). \tag{7.4}$$

We recall the definition of a compact operator from Porter and Stirling [23, Definition 3.2].

Definition 7.2 An operator $\boldsymbol{T} : L^2([0,T]) \to L^2([0,T])$ is said to have *finite rank* if its image $\{\boldsymbol{T}\psi : \psi \in L^2([0,T])\}$ has finite dimension. Furthermore, an operator $\boldsymbol{L} \in B(L^2([0,T]))$ is said to be *compact* if there is a sequence $(\boldsymbol{L}_n)_{n\geq 1}$ of finite-rank operators in $B(L^2([0,T]))$ such that $\|\boldsymbol{L}_n - \boldsymbol{L}\| \to 0$ as $n \to \infty$.

A particularly important result that we use states that any operator generated by a kernel in $L^2([0,T]^2)$ is compact (see [23, Theorem 3.4]).

Next we define nonnegative and positive operators in $B(L^2([0,T]))$ as in [23, Definition 6.1].

Definition 7.3 A self-adjoint operator $\boldsymbol{T} \in B(L^2([0,T]))$ is said to be *nonnegative* if $\langle \boldsymbol{T}\psi, \psi\rangle_{L^2} \geq 0$ for all $\psi \in L^2([0,T])$. It is said to be *positive* if $\langle \boldsymbol{T}\psi, \psi\rangle_{L^2} > 0$ for all $\psi \neq 0$ in $L^2([0,T])$. If there is a positive constant m for which $\langle \boldsymbol{T}\psi, \psi\rangle_{L^2} \geq m\|\psi\|^2_{L^2}$ for all $\psi \in L^2([0,T])$, then $\boldsymbol{T}$ is called *positive and bounded below*.

Our next result outlines several important properties of the operator $\boldsymbol{G}$ in (3.11). Recall that $\boldsymbol{K}_1$ was defined in (7.3).

Proposition 7.4 *Let $\boldsymbol{G}$ be defined as in* (3.11). *Then $\boldsymbol{G} \in B(L^2([0,T]))$ is a positive, compact and self-adjoint operator. Moreover, it satisfies*

$$(\boldsymbol{G}\psi)(t) = (\boldsymbol{K}_1\boldsymbol{K}_1^*\psi)(t), \qquad t \in [0,T], \psi \in L^2([0,T]). \tag{7.5}$$

The proof of Proposition 7.4 is given in Sect. 8.

In the following result, we determine an alternative representation for the major agent's objective functional.

Lemma 7.5 *Let H^0 be the major agent's objective functional in* (2.10). *Then for any $\nu^0 \in \mathcal{A}_M^{q_0}$, it holds that*

$$\begin{aligned} H^0(\nu^0) &= x_0 + M_0 q_0 - \kappa_0 \frac{q_0^2}{2} - \frac{\kappa_1\kappa_0}{2\lambda_1}\int_0^T \nu_t^0 (\boldsymbol{G}\nu^0)(t)dt \\ &\quad + \int_0^T \left(\frac{\kappa_1}{2\lambda_1}\nu_t^0(\boldsymbol{G}\bar{\mu})(t) + Q_t^{0,\nu^0}\bar{\mu}_t\right)dt - \lambda_0\int_0^T (\nu_t^0)^2 dt. \end{aligned} \tag{7.6}$$

The proof of Lemma 7.5 is postponed to Sect. 9.

The following result establishes the uniqueness of the maximiser of H^0.

Proposition 7.6 *There exists at most one admissible maximiser to the major agent's objective functional H^0 in* (2.10).

Proof To show the result, it is sufficient to show that H^0 is strictly concave over $\mathcal{A}_M^{q_0}$.

Notice that H^0 is finite-valued since $\nu^{1,*} \in \mathcal{A}_m$, $\nu^0 \in \mathcal{A}_M^{q_0}$ and μ satisfies (3.7). Therefore, in order to show that H^0 is strictly concave over $\mathcal{A}_M^{q_0}$, we must verify that

$$\mathcal{I}^0(\rho,\nu,\omega) := H^0\big(\rho\nu + (1-\rho)\omega\big) - \rho H^0(\nu) - (1-\rho)H^0(\omega) > 0 \tag{7.7}$$

for any $\rho \in (0,1)$ and for any dt-distinguishable $\nu, \omega \in \mathcal{A}_M^{q_0}$. Fix such ρ and ν, ω. From (2.3) and (3.11), it follows that $Q^{0,\nu}$ and $\boldsymbol{G}\nu$ are linear in ν. Then from (7.6) and (7.7), we get

$$\mathcal{I}^0(\rho,\nu,\omega) = \rho(1-\rho)\lambda_0 \int_0^T (\nu_t - \omega_t)^2 dt + \rho(1-\rho)\frac{\kappa_1\kappa_0}{2\lambda_1}\int_0^T (\nu_t - \omega_t)\big(\boldsymbol{G}(\nu-\omega)\big)(t)dt. \tag{7.8}$$

Note that we are only left to show that the second term on the right-hand side of (7.8) is nonnegative. Since $\boldsymbol{G}$ is a positive operator by Proposition 7.4 and ν and ω are dt-distinguishable, the result follows. □

As similarly done in Sect. 6, we can apply a convex analytic calculus of variations approach to derive a first-order optimality condition which characterises the unique solution to the major agent's problem.

Since the map $\nu^0 \to H^0(\nu^0)$ in (2.10) is strictly concave, H^0 admits a unique maximiser characterised by the critical point at which the Gâteaux derivative

$$\langle \mathcal{D}H^0(\nu^0), \omega\rangle = \lim_{\epsilon\to 0}\frac{H^0(\nu^0+\epsilon\omega) - H^0(\nu^0)}{\epsilon}$$

vanishes. We remind the reader that the minor agent's optimal strategy $\nu^{1,*}$ was fixed to be the one in (3.6), making the major agent's objective functional H^0 only a function of the major agent's control ν^0. In the next result, we explicitly compute the first-order Gâteaux derivative of H^0. Recall that $\mathcal{A}_M^{q_0}$ was defined in (2.2).

Lemma 7.7 *The Gâteaux derivative of H^0 in (7.6) in the direction $\omega \in \mathcal{A}_M^0$ is given for any $\nu^0 \in \mathcal{A}_M^{q_0}$ by*

$$\langle \mathcal{D}H^0(\nu^0), \omega\rangle = \int_0^T \omega_t\Big(-2\lambda_0\nu_t^0 - \frac{\kappa_1\kappa_0}{\lambda_1}(\boldsymbol{G}\nu^0)(t) + \frac{\kappa_1}{2\lambda_1}(\boldsymbol{G}\bar{\mu})(t) - \int_t^T \bar{\mu}_s ds\Big)dt. \tag{7.9}$$

Proof Let $\epsilon > 0$, $\nu^0 \in \mathcal{A}_M^{q_0}$ and $\omega \in \mathcal{A}_M^0$. From (7.6), we get

$$\begin{aligned} &H^0(\nu^0+\epsilon\omega) - H^0(\nu^0) \\ &= \epsilon\Big(\int_0^T \omega_t\Big(-2\lambda_0\nu_t^0 + \frac{\kappa_1}{2\lambda_1}(\boldsymbol{G}\bar{\mu})(t)\Big)dt - \frac{\kappa_1\kappa_0}{2\lambda_1}\int_0^T \omega_t(\boldsymbol{G}\nu^0)(t)dt \\ &\quad - \frac{\kappa_1\kappa_0}{2\lambda_1}\int_0^T \nu_t^0(\boldsymbol{G}\omega)(t)dt - \int_0^T \bar{\mu}_t\Big(\int_0^t \omega_s ds\Big)dt\Big) \\ &\quad + \epsilon^2\Big(-\lambda_0\int_0^T \omega_s^2 ds - \frac{\kappa_1\kappa_0}{2\lambda_1}\int_0^T \omega_t(\boldsymbol{G}\omega)(t)dt\Big). \end{aligned} \tag{7.10}$$

Notice that all the terms in (7.10) are finite as a consequence of $\omega \in \mathcal{A}_M^0$, $\nu^0 \in \mathcal{A}_M^{q_0}$, μ satisfying (3.7) and $\boldsymbol{G} \in B(L^2([0,T]))$ as shown in Proposition 7.4. It follows that

$$\langle \mathcal{D}H^0(\nu^0), \omega \rangle = \int_0^T \omega_t \Big(-2\lambda_0 \nu_t^0 - \frac{\kappa_1 \kappa_0}{2\lambda_1} (\boldsymbol{G}\nu^0)(t) + \frac{\kappa_1}{2\lambda_1} (\boldsymbol{G}\bar{\mu})(t) \Big) dt$$
$$- \frac{\kappa_1 \kappa_0}{2\lambda_1} \int_0^T \nu_t^0 (\boldsymbol{G}\omega)(t) dt - \int_0^T \bar{\mu}_t \Big(\int_0^t \omega_s ds \Big) dt. \tag{7.11}$$

Proposition 7.4 shows that $\boldsymbol{G} \in B(L^2([0,T]))$. Since μ satisfies (3.7), $\bar{\mu}$ is in $L^2([0,T])$ by Jensen's inequality. Since $\omega \in \mathcal{A}_M^0$ and $\bar{\mu} \in L^2([0,T])$, we can apply Fubini's theorem to obtain

$$\int_0^T \bar{\mu}_t \Big(\int_0^t \omega_s ds \Big) dt = \int_0^T \omega_s \int_s^T \bar{\mu}_t dt ds. \tag{7.12}$$

As shown in Proposition 7.4, $\boldsymbol{G}$ is self-adjoint; therefore it holds that

$$\int_0^T \nu_t^0 (\boldsymbol{G}\omega)(t) dt = \langle \nu^0, (\boldsymbol{G}\omega) \rangle_{L^2} = \langle \omega, (\boldsymbol{G}\nu^0) \rangle_{L^2} = \int_0^T \omega_t (\boldsymbol{G}\nu^0)(t) dt. \tag{7.13}$$

Finally, we use (7.12) and (7.13) in (7.11) to get (7.9). □

Recall the definition of the operator $\boldsymbol{S}$ in (3.12). In the following result, we derive an optimality condition that takes the form of an integral equation.

Proposition 7.8 *A strategy $\nu^{0,*} \in \mathcal{A}_M^{q_0}$ maximises the major agent's performance functional (2.10) if there exists a constant $\eta \in \mathbb{R}$ such that*

$$\nu_t^{0,*} + \frac{\kappa_1 \kappa_0}{2\lambda_0 \lambda_1} (\boldsymbol{G}\nu^{0,*})(t) = (\boldsymbol{S}\bar{\mu})(t) + \frac{\eta}{2\lambda_0}, \qquad 0 \le t \le T. \tag{7.14}$$

Proof In the proof of Proposition 7.6, we have shown that H^0 is strictly concave over $\mathcal{A}_M^{q_0}$. Therefore, by Ekeland and Témam [11, Proposition II.2.1], if an admissible strategy $\nu^{0,*}$ satisfies

$$\langle \mathcal{D}H^0(\nu^{0,*}), \omega \rangle = 0 \qquad \text{for all } \omega \in \mathcal{A}_M^0, \tag{7.15}$$

then it is the maximiser of H^0. Recalling the result of Lemma 7.7, we plug (7.14) into (7.9) to get

$$\langle \mathcal{D}H^0(\nu^{0,*}), \omega \rangle = -\int_0^T \omega_t \Big(2\lambda_0 (\boldsymbol{S}\bar{\mu})(t) + \eta - \frac{\kappa_1}{2\lambda_1} (\boldsymbol{G}\bar{\mu})(t) + \int_t^T \bar{\mu}_s ds \Big) dt$$
$$= -\Big(\int_0^T \bar{\mu}_s ds + \eta \Big) \int_0^T \omega_t dt,$$

where we used (3.12) in the second equality. Recall that $\omega \in \mathcal{A}_M^0$; then from (2.2), we get (7.15). □

The following result, which is proved in Sect. 8, derives some properties of the operator $\boldsymbol{R}$ in (3.14) that are crucial to the proof of Theorem 3.5.

Proposition 7.9 *For $\boldsymbol{R}$ defined as in (3.14) and $\boldsymbol{G}$ as in (3.11), the following hold:*
(i) *The inverse operator of $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ exists and satisfies*

$$\boldsymbol{R} = \left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)^{-1}. \tag{7.16}$$

(ii) *The operator $\boldsymbol{R}$ is positive, bounded from below and in $B(L^2([0,T]))$.*
(iii) *If $f \in C([0,T])$, then $\boldsymbol{R}f \in C([0,T])$.*

The following result derives some essential properties of the operator $\boldsymbol{S}$ in (3.12).

Lemma 7.10 *Suppose that $\boldsymbol{S}$ is as in (3.12). Then $\boldsymbol{S}$ is in $B(L^2([0,T]))$, and for any $\psi \in L^2([0,T])$, $t \mapsto (\boldsymbol{S}\psi)(t)$ is continuous on $[0,T]$.*

Proof Proposition 7.4 proves that $\boldsymbol{G}$ is in $B(L^2([0,T]))$; hence it follows from (3.12) that $\boldsymbol{S}$ is also in $B(L^2([0,T]))$. Next, let $\psi \in L^2([0,T])$. From Proposition 9.4 below, it follows that $\boldsymbol{G}\psi$ is continuously differentiable on $[0,T]$. Hence by (3.12), the continuity of $(\boldsymbol{S}\psi)(\cdot)$ follows. □

We are now ready to prove Theorem 3.5.

Proof of Theorem 3.5 Recall that $\nu^{0,*}$ and η were defined in (3.15) and (3.16), respectively. We split the proof into the following steps.

1) We show that $(\nu^{0,*}, \eta)$ is the unique solution to (7.14). Recall that $\boldsymbol{R}$ was defined in (3.14). From (3.7) and Lemma 7.10, it follows that $\boldsymbol{S}\bar{\mu}(\cdot)$ is in $L^2([0,T])$. Then from Proposition 7.9, we get that the unique solution to (7.14) is given by

$$\begin{aligned}\nu_t^{0,*} &= \left(\left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)^{-1}\left(\frac{\eta}{2\lambda_0} + \boldsymbol{S}\bar{\mu}\right)\right)(t) \\ &= \frac{\eta}{2\lambda_0}(\boldsymbol{R}\mathbf{1})(t) + (\boldsymbol{R}\boldsymbol{S}\bar{\mu})(t), \qquad 0 \le t \le T. \end{aligned} \tag{7.17}$$

2) We verify that $\nu^{0,*}$ satisfies the fuel constraint in (2.2). Note that we need to show that $\langle \nu^{0,*}, \mathbf{1}\rangle_{L^2} = q_0$. From (3.15) and (3.16), we get

$$\begin{aligned}\langle \nu^{0,*}, \mathbf{1}\rangle_{L^2} &= \frac{\eta}{2\lambda_0}\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2} + \langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2} \\ &= \frac{q_0 - \langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}}{\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2}}\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2} + \langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2} \\ &= q_0. \end{aligned}$$

Note that by Proposition 7.9(ii), the operator $\boldsymbol{R}$ is positive and bounded from below. Therefore the denominator in (3.16) is strictly positive and the constant η is well defined.

3) We verify that $\nu^{0,*}$ is in $L^2([0,T])$. From Lemma 7.10, it follows that $t \mapsto (\boldsymbol{S}\bar{\mu})(t)$ is continuous on $[0,T]$. Proposition 7.9(iii) and (7.17) then imply that $t \mapsto \nu_t^{0,*}$ is continuous and therefore bounded on $[0,T]$. This concludes the proof. □

8 Proofs of Lemma 7.1 and Propositions 7.4 and 7.9

Proof of Lemma 7.1 Recall that the kernels $\mathcal{K}$ and $\mathcal{K}_1$ were defined in (3.4) and (7.2), respectively. Let ψ be in $L^2([0,T])$. From (7.1), we get

$$\mathcal{K}_1^*(t,s) = \mathcal{K}(t,s)\mathbb{1}_{\{t\le s\}}, \qquad 0 \le t,s \le T, \tag{8.1}$$

which verifies (7.4). By Lemma 6.7, $\mathcal{K}$ is in $L^2([0,T]^2)$. Then by (7.2) and (8.1), also $\mathcal{K}_1^*$ and $\mathcal{K}_1$ are in $L^2([0,T]^2)$. Thus $\boldsymbol{K}_1$ and $\boldsymbol{K}_1^*$ are in $B(L^2([0,T]))$. □

The following result describes several properties of the kernel $\mathcal{G}$ in (3.10) which are essential for the proof of Proposition 7.9.

Lemma 8.1 *Let $\mathcal{G}$ be as in (3.10). Then $\mathcal{G}$ is symmetric and jointly continuous on $[0,T]^2$. Moreover, $\mathcal{G}$ is in $L^2([0,T]^2)$.*

Proof From (3.10), it follows that $\mathcal{G}$ is symmetric. Recall that ξ^+ and ξ^- were defined in (3.3). Note that ξ^+ and ξ^- are continuous on $[0,T]$ and that ξ^- belongs to $L^2([0,T])$. It follows that $t \mapsto \int_0^t (\xi_u^-)^2 du$ is continuous on $[0,T]$. From (3.4) and (3.10), it follows that the kernel $\mathcal{G}$ can be rewritten as

$$\mathcal{G}(t,s) = \xi_t^+ \xi_s^+ \int_0^{t\wedge s} (\xi_u^-)^2 du, \qquad 0 \le t,s \le T.$$

Therefore $\mathcal{G}(t,s)$ is jointly continuous on $[0,T]^2$, and hence is in $L^2([0,T]^2)$. □

The following result is needed for the proof of Proposition 7.4.

Lemma 8.2 *Let $\boldsymbol{K}_1$ be defined as in (7.3) and let $\psi \in L^2([0,T])$. If $(\boldsymbol{K}_1\psi)(t) = 0$ for $0 \le t \le T$, then $\psi(t) = 0$ a.e. on $[0,T]$.*

Proof Let $\psi \in L^2([0,T])$ and assume that $(\boldsymbol{K}_1\psi)(t) = 0$ for $0 \le t \le T$. From (3.4), (7.2) and (7.3), we get

$$(\boldsymbol{K}_1\psi)(t) = \xi_t^+ \left(\int_0^t \xi_s^- \psi(s) ds \right) = 0, \qquad 0 \le t \le T. \tag{8.2}$$

From (3.3), it follows that $\xi_t^\pm > 0$ for $0 \le t \le T$; therefore

$$\int_0^t \xi_s^- \psi(s) ds = 0, \qquad 0 \le t \le T. \tag{8.3}$$

Since $t \mapsto \xi_t^{\pm}$ are also continuous on $[0, T]$ and $\psi \in L^2([0, T])$, $\xi^- \psi$ is integrable. Then from the Lebesgue differentiation theorem and (8.3), we conclude that

$$\xi_t^- \psi(t) = 0 \qquad dt\text{-a.e. on } [0, T].$$

But again since $\xi_t^- > 0$ for all $t \in [0, T]$, we get that

$$\psi(t) = 0 \qquad dt\text{-a.e. on } [0, T],$$

which proves the result. □

We are now ready to prove Proposition 7.4.

Proof of Proposition 7.4 Let $\psi \in L^2([0, T])$. From (3.10), we get

$$\begin{aligned}(\boldsymbol{G}\psi)(t) &= \int_0^T \mathcal{G}(t, u)\psi(u)du \\ &= \int_0^T \psi(u) \int_0^{t\wedge u} \mathcal{K}(s, t)\mathcal{K}(s, u)dsdu \\ &= \int_0^T \int_0^T \mathcal{K}(s, t)\mathcal{K}(s, u)\psi(u)\mathbb{1}_{\{s\le t\}}\mathbb{1}_{\{s\le u\}}duds.\end{aligned}$$

Together with (7.2) and (8.1), it follows that

$$\begin{aligned}(\boldsymbol{G}\psi)(t) &= \int_0^T \mathcal{K}_1(t, s) \int_0^T \mathcal{K}_1^*(s, u)\psi(u)duds \\ &= (\boldsymbol{K}_1\boldsymbol{K}_1^*\psi)(t), \qquad 0 \le t \le T,\end{aligned} \tag{8.4}$$

which proves (7.5). Since $\boldsymbol{K}_1$, $\boldsymbol{K}_1^*$ are in $B(L^2([0, T]))$ by Lemma 7.1, an application of the Cauchy–Schwarz inequality and (8.4) allow us to show that $\boldsymbol{G} \in B(L^2([0, T]))$. Note that (8.4) also implies that $\boldsymbol{G}$ is self-adjoint, that is,

$$\boldsymbol{G}^* = (\boldsymbol{K}_1\boldsymbol{K}_1^*)^* = (\boldsymbol{K}_1^*)^*\boldsymbol{K}_1^* = \boldsymbol{G}.$$

Next, using (3.11), we prove that $\boldsymbol{G}$ is compact (see Definition 7.2). From Lemma 8.1, it follows that $\mathcal{G}$ is in $L^2([0, T]^2)$. Then the result follows from [23, Theorem 3.4] which shows that any integral operator generated by a kernel in $L^2([0, T]^2)$ is compact. Finally, we prove that $\boldsymbol{G}$ is a positive operator in the sense of Definition 7.3. We have shown that $\boldsymbol{G} = \boldsymbol{K}_1\boldsymbol{K}_1^*$. Since we have for any $\psi \in L^2([0, T])$ that

$$\langle \boldsymbol{G}\psi, \psi\rangle_{L^2} = \langle \boldsymbol{K}_1\boldsymbol{K}_1^*\psi, \psi\rangle_{L^2} = \|\boldsymbol{K}_1\psi\|^2 \ge 0,$$

it follows that $\boldsymbol{G}$ is nonnegative. Moreover, by Lemma 8.2, we have $(\boldsymbol{K}_1\psi)(t) = 0$ for all $t \in [0, T]$ only for $\psi = 0$ a.e. on $[0, T]$. Therefore $\boldsymbol{G}$ is positive. □

In the following, we present a sequence of results which are essential to the proof of Proposition 7.9. Using the results of Proposition 7.4, we are in a position to fully characterise the spectral properties of the integral operator $\boldsymbol{G}$.

Lemma 8.3 *Let $\boldsymbol{G}$ be defined as in (3.11). Then $\boldsymbol{G}$ has a sequence $(\zeta_n)_{n\geq 1}$ of positive eigenvalues and a corresponding orthonormal sequence $(\psi_n)_{n\geq 1}$ of eigenfunctions in $L^2([0,T])$ such that for each $\varphi \in L^2([0,T])$, we have that*

$$\boldsymbol{G}\varphi = \sum_{n\geq 1} \zeta_n \langle \varphi, \psi_n \rangle_{L^2} \psi_n.$$

Moreover, for all $N \geq 1$, define $\boldsymbol{G}_N, \boldsymbol{G}_N^{\mathrm{abs}} \in B(L^2([0,T]))$ by

$$\boldsymbol{G}_N\varphi = \sum_{n=1}^{N} \zeta_n \langle \varphi, \psi_n \rangle_{L^2} \psi_n,$$

$$\boldsymbol{G}_N^{\mathrm{abs}}\varphi = \sum_{n=1}^{N} |\zeta_n \langle \varphi, \psi_n \rangle_{L^2} \psi_n|.$$

Then $\boldsymbol{G}_N\varphi$ converges uniformly to $\boldsymbol{G}\varphi$, i.e.,

$$\lim_{N\to\infty} \sup_{t\in[0,T]} |(\boldsymbol{G}_N\varphi)(t) - (\boldsymbol{G}\varphi)(t)| = 0, \tag{8.5}$$

and $\boldsymbol{G}_N^{\mathrm{abs}}\varphi$ is uniformly convergent, i.e., there exists a function $\Phi \in L^2([0,T])$ with

$$\lim_{N\to\infty} \sup_{t\in[0,T]} |(\boldsymbol{G}_N^{\mathrm{abs}}\varphi)(t) - \Phi(t)| = 0. \tag{8.6}$$

Proof By Proposition 7.4, $\boldsymbol{G}$ is a self-adjoint compact operator in $B(L^2([0,T]))$. Thus by Porter and Stirling [23, Theorem 4.15], there are a sequence $(\zeta_n)_{n\geq 1}$ of nonzero eigenvalues of $\boldsymbol{G}$ and a corresponding orthonormal sequence $(\psi_n)_{n\geq 1}$ of eigenfunctions in $L^2([0,T])$ with $\boldsymbol{G}\phi = \sum_{n\geq 1} \zeta_n \langle \phi, \psi_n \rangle_{L^2} \phi_n$ for any ϕ in $L^2([0,T])$. Moreover, the operators $\boldsymbol{G}_N$ converge to $\boldsymbol{G}$ in mean, i.e., $\|\boldsymbol{G} - \boldsymbol{G}_N\| \to 0$ as $N \to \infty$. By Proposition 7.4, $\boldsymbol{G}$ is positive and self-adjoint; hence [23, Lemma 6.1] gives that all its eigenvalues $(\zeta_n)_{n\geq 1}$ are positive. Since $\mathcal{G}$ is continuous and symmetric by Proposition 8.1, [23, Theorem 4.22] implies that $\boldsymbol{G}_N\phi$ and $\boldsymbol{G}_N^{\mathrm{abs}}\phi$ satisfy the convergences in (8.5) and (8.6). □

Remark 8.4 In Appendix A, we provide an example for the spectral decomposition of $\boldsymbol{G}$ in Lemma 8.3.

Lemma 8.5 *Let $\boldsymbol{G}$ be defined as in (3.11). Then the operator $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ is positive and bounded from below in the sense of Definition 7.3.*

Proof Let $\psi \in L^2([0,T])$. Since $\kappa_0, \kappa_1, \lambda_0, \lambda_1 > 0$ and $\boldsymbol{G}$ is positive by Proposition 7.4, we obtain

$$\left\langle \left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)\psi, \psi \right\rangle_{L^2} \geq \|\psi\|_{L^2}^2.$$

Therefore $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ is positive and bounded from below. □

Recall that we assume that the constants $\lambda_0, \lambda_1, \kappa_1, \kappa_0$ are strictly positive and that we proved in Lemma 8.3 that the eigenvalues of $\boldsymbol{G}$ are all positive. The following result is therefore an easy corollary.

Lemma 8.6 *Let* $\zeta^* = -\frac{2\lambda_0\lambda_1}{\kappa_1\kappa_0}$ *and let* $\boldsymbol{G}$ *be defined as in* (3.11). *Then* ζ^* *is not an eigenvalue of the integral operator* $\boldsymbol{G}$.

We recall [23, Theorem 4.27] which will be useful in the proof of Proposition 7.9.

Proposition 8.7 *Let* $f \in L^2([0,T])$. *Suppose* $\mathcal{T} : [0,T]^2 \to \mathbb{R}$ *is a continuous symmetric kernel and* $\boldsymbol{T}$ *is the integral operator generated by* $\mathcal{T}$. *Let* $(\mu_n)_{n\geq 1}$ *and* $(\varphi_n)_{n\geq 1}$ *be the sequences of eigenvalues and eigenfunctions of the operator* $\boldsymbol{T}$. *Moreover, suppose* $\frac{1}{\lambda}$ *is not an eigenvalue of* $\boldsymbol{T}$. *Then the unique solution to the integral equation*

$$\psi(t) - \lambda \int_0^T \mathcal{T}(t,s)\psi(s)dt = f(t), \qquad t \in [0,T],$$

is given by

$$\psi(t) = f(t) + \int_0^T \widetilde{\mathcal{R}}(t,s) f(s)ds, \qquad t \in [0,T],$$

where

$$\widetilde{\mathcal{R}}(t,s) = \lambda\mathcal{T}(t,s) + \sum_{n\geq 1} \frac{\lambda\mu_n}{1-\lambda\mu_n}\varphi_n(t)\varphi_n(s)$$

is jointly continuous on $[0,T]^2$.

We are now ready to prove Proposition 7.9.

Proof of Proposition 7.9 (i) Note that the operator $\boldsymbol{G}$ and the corresponding kernel $\mathcal{G}$ satisfy the assumptions of Proposition 8.7. Specifically, it follows from Lemma 8.6 that $-\frac{2\lambda_0\lambda_1}{\kappa_1\kappa_0}$ is not an eigenvalue of $\boldsymbol{G}$. Moreover, we have shown in Proposition 8.1 that $\mathcal{G}$ is continuous and symmetric on $[0,T]^2$. Therefore we can apply Proposition 8.7 to the integral equation

$$\psi(t) + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\int_0^T \mathcal{G}(t,s)\psi(s)ds = f(t), \qquad t \in [0,T], \tag{8.7}$$

and deduce that the unique solution to (8.7) is given by

$$\psi(t) = f(t) + \int_0^T \mathcal{R}(t,s) f(s)ds \tag{8.8}$$

with $\mathcal{R}$ as in (3.13). Moreover, it follows from Proposition 8.7 that the kernel $\mathcal{R}$ is jointly continuous on $[0,T]^2$.

Next, we show that the inverse of $\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ is given by $\boldsymbol{R}$. Since by Lemma 8.6, $-\frac{2\lambda_0\lambda_1}{\kappa_1\kappa_0}$ is not an eigenvalue of $\boldsymbol{G}$, the operator $(\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G})^{-1}$ exists. Let ψ be the solution of (8.7). Since $\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ is invertible, it follows from (8.7) that ψ can be written as

$$\psi=\left(\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)^{-1}f. \tag{8.9}$$

On the other hand, we have from (3.14) and (8.8) that

$$\psi=\boldsymbol{R}f. \tag{8.10}$$

Therefore by comparing (8.9) and (8.10), we find that

$$\left(\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}\right)^{-1}f=\boldsymbol{R}f. \tag{8.11}$$

Since (8.11) holds for any $f\in L^2([0,T])$, (i) follows.

(ii) By Proposition 7.4, $\boldsymbol{G}$ is a compact operator. Since also the inverse of $\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ exists by (i), we get from the remark below the proof of [23, Theorem 3.3] that the inverse of $\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ is in $B(L^2([0,T]))$. From (7.16), it follows that $\boldsymbol{R}$ is also in $B(L^2([0,T]))$. Recall that Lemma 8.5 shows that $\boldsymbol{I}+\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}$ is positive and bounded from below. Hence it follows from [23, Lemma 6.2] that its inverse is positive and bounded from below. From (7.16), we conclude that $\boldsymbol{R}$ is positive and bounded from below in the sense of Definition 7.3.

(iii) If $f\in C([0,T])$, then (iii) follows from (8.10) and since $\mathcal{R}$ is jointly continuous on $[0,T]^2$ by (i). □

9 Proof of Lemma 7.5

Throughout this section, we impose Assumption 3.1 so that the minor agent's optimal control $\nu^{1,*}$ is well defined. Before proving Lemma 7.5, we prove several intermediate results.

Lemma 9.1 *Let r^0 be defined as in (3.5) and $\boldsymbol{K}_1$ as in (7.3). Then for any $\nu^0\in\mathcal{A}_M^{q_0}$,*

$$\mathbb{E}[r_t^0(\nu^0)]=\frac{1}{2\lambda_1}(\boldsymbol{K}_1^*\bar{\mu})(t)-\frac{\kappa_0}{2\lambda_1}(\boldsymbol{K}_1^*\nu^0)(t),\qquad 0\le t\le T, \tag{9.1}$$

Moreover, $(\mathbb{E}[r_t^0(\nu^0)])_{t\in[0,T]}$ is in $L^2([0,T])$.

Proof Fix $\nu^0\in\mathcal{A}_M^{q_0}$ and recall that $\mathcal{K}$ was defined in (3.4). From Lemma 6.7, (3.7) and (3.8), it follows that the conditions of Fubini's theorem are satisfied and we get

$$\mathbb{E}\left[\int_t^T\mathcal{K}(t,s)\mu_s ds\right]=\int_t^T\mathcal{K}(t,s)\bar{\mu}_s ds,\qquad 0\le t\le T. \tag{9.2}$$

Using (3.5), (9.2) and the tower property gives

$$\mathbb{E}[r_t^0(\nu^0)] = \frac{1}{2\lambda_1}\mathbb{E}\Big[\mathbb{E}_t\Big[\int_t^T \mathcal{K}(t,s)(\mu_s - \kappa_0 \nu_s^0)ds\Big]\Big]$$
$$= \frac{1}{2\lambda_1}\int_0^T \mathcal{K}(t,s)(\bar{\mu}_s - \kappa_0 \nu_s^0)\mathbb{1}_{\{t\le s\}}ds. \tag{9.3}$$

Using the expression for $\boldsymbol{K}_1^*$ from (7.4) in (9.3), we arrive at (9.1).

By Lemma 7.1, the operators $\boldsymbol{K}_1$ and $\boldsymbol{K}_1^*$ are in $B(L^2([0,T]))$. By assumption, $\nu^0, \bar{\mu} \in L^2([0,T])$, and so it follows from (9.1) that $(\mathbb{E}[r_t^0(\nu^0)])_{t\in[0,T]}$ is in $L^2([0,T])$. □

Lemma 9.2 *Let $\nu^{1,*}$ be defined as in (3.6). Then for any $\nu^0 \in \mathcal{A}_M^{q_0}$, we have for $0 \le t \le T$ that*

$$\mathbb{E}[\nu_t^{1,*}(\nu^0)] = \frac{\kappa_0}{2\lambda_1}\big((\boldsymbol{K}_1^*\nu^0)(t) + r_t^1(\boldsymbol{G}\nu^0)(t)\big) - \frac{1}{2\lambda_1}\big((\boldsymbol{K}_1^*\bar{\mu})(t) + r_t^1(\boldsymbol{G}\bar{\mu})(t)\big).$$

Proof Lemmas 9.1 and 6.7 prove that $\mathbb{E}[r_\cdot^0(\nu^0)] \in L^2([0,T])$ and $\mathcal{K} \in L^2([0,T]^2)$, respectively. Then from (3.6) and Fubini's theorem, it follows that

$$\mathbb{E}[\nu_t^{1,*}(\nu^0)] = -\mathbb{E}[r_t^0(\nu^0)] - r_t^1\int_0^t \mathcal{K}(s,t)\mathbb{E}[r_s^0(\nu^0)]ds.$$

Together with (9.1), we get

$$\mathbb{E}[\nu_t^{1,*}(\nu^0)] = -\frac{1}{2\lambda_1}(\boldsymbol{K}_1^*\bar{\mu})(t) + \frac{\kappa_0}{2\lambda_1}(\boldsymbol{K}_1^*\nu^0)(t)$$
$$- r_t^1\int_0^t \mathcal{K}(s,t)\Big(\frac{1}{2\lambda_1}(\boldsymbol{K}_1^*\bar{\mu})(s) - \frac{\kappa_0}{2\lambda_1}(\boldsymbol{K}_1^*\nu^0)(s)\Big)ds. \tag{9.4}$$

By using (7.2), (7.3) and (7.5) in (9.4), we get the result. □

The following result simply follows from (2.2) and integration by parts; hence we omit the proof.

Lemma 9.3 *Let M be a square-integrable martingale over* $[0,T]$. *Then for any $\nu^0 \in \mathcal{A}_M^{q_0}$, we have*

$$\mathbb{E}\Big[\int_0^T M_t \nu_t^0 dt\Big] = M_0 q_0.$$

In the following result, we derive an operator differential equation which is satisfied by the operator $\boldsymbol{G}$ in (3.11).

Proposition 9.4 *For any* $\psi \in L^2([0,T])$*, the operator* $\boldsymbol{G}$ *satisfies the differential equation*

$$\frac{d}{dt}(\boldsymbol{G}\psi)(t) = r^1_t(\boldsymbol{G}\psi)(t) + (\boldsymbol{K}^*_1\psi)(t), \quad 0 \le t \le T, \qquad (\boldsymbol{G}\psi)(0) = 0. \quad (9.5)$$

Moreover, $t \mapsto (\boldsymbol{G}\psi)(t)$ *is continuously differentiable on* $[0,T]$.

Proof Fix $\psi \in L^2([0,T])$. By Proposition 8.1, $\mathcal{G}$ is jointly continuous on $[0,T]^2$; hence by (3.11), we get that $t \mapsto (\boldsymbol{G}\psi)(t)$ is continuous on $[0,T]$.

Note that by (3.3), we have

$$\frac{d\xi^{\pm}}{dt} = \pm r^1_t \xi^{\pm}_t, \qquad 0 \le t \le T. \quad (9.6)$$

Since r^1 is continuous on $[0,T]$ by Proposition 6.5, it follows from (9.6) that $\xi^{\pm}$ are continuously differentiable on $[0,T]$. From (8.2), (8.4) and (9.6), we get that

$$\begin{aligned}\frac{d}{dt}(\boldsymbol{G}\psi)(t) &= \frac{d}{dt}\bigg(\xi^+_t \int_0^t \xi^-_s (\boldsymbol{K}^*_1\psi)(s)ds\bigg) \\ &= r^1_t\bigg(\xi^+_t \int_0^t \xi^-_s (\boldsymbol{K}^*_1\psi)(s)ds\bigg) + (\boldsymbol{K}^*_1\psi)(t) \\ &= r^1_t(\boldsymbol{G}\psi)(t) + (\boldsymbol{K}^*_1\psi)(t).\end{aligned}$$

Since the operator $\boldsymbol{G}$ can by (7.5) be represented in terms of $\boldsymbol{K}_1$ and $\boldsymbol{K}^*_1$ and Lemma 6.7 shows that $\mathcal{K}$ is jointly continuous on $[0,T]^2$, it follows from (7.3) and (7.4) that $t \mapsto (\boldsymbol{G}\psi)(t)$ and $t \mapsto (\boldsymbol{K}^*_1\psi)(t)$ are continuous on $[0,T]$. As we have shown that r^1 is also continuous, it follows that $\frac{d}{dt}(\boldsymbol{G}\psi)$ is continuous on $[0,T]$. Finally, note that

$$(\boldsymbol{K}_1\psi)(0) = \int_0^T \mathcal{K}(s,0)\mathbb{1}_{\{s\le 0\}}\psi(s)ds = 0.$$

From (7.5), we have $(\boldsymbol{G}\psi)(t) = (\boldsymbol{K}_1\boldsymbol{K}^*_1\psi)(t)$. This proves that $(\boldsymbol{G}\psi)(0) = 0$ and completes the proof. □

Now we are ready to prove Lemma 7.5.

Proof of Lemma 7.5 Let $\nu^0 \in \mathcal{A}^{q_0}_M$. Recall that the minor agent's strategy is assumed to be $\nu^{1,*}$ in (3.6). We define

$$Z^\nu_t = Y^\nu_t - \int_0^t \mu_s ds, \qquad 0 \le t \le T. \quad (9.7)$$

Note that it follows from (2.7) that $Z^\nu_0 = 0$.

Using (2.9) and (2.8), we get

$$\mathbb{E}[X^{0,\nu^0}_T] = x_0 + \mathbb{E}\bigg[\int_0^T (P^\nu_t - \lambda_0\nu^0_t)\nu^0_t dt\bigg].$$

Together with (2.6), (2.3) and (9.7), we arrive at

$$\mathbb{E}[X_T^{0,\nu^0}] = x_0 + \mathbb{E}\bigg[\int_0^T M_t \nu_t^0 dt\bigg] + \mathbb{E}\bigg[\int_0^T Z_t^\nu dQ_t^{0,\nu^0}\bigg] - \int_0^T \lambda_0 (\nu_t^0)^2 dt. \quad (9.8)$$

Recall that $Z_0^\nu = 0$ and $Q_T^{0,\nu^0} = 0$. Using integration by parts, (2.7), (9.7) and Fubini's theorem, we obtain

$$\begin{aligned}\mathbb{E}\bigg[\int_0^T Z_t^\nu dQ_t^{0,\nu^0}\bigg] &= -\mathbb{E}\bigg[\int_0^T Q_t^{0,\nu^0} dZ_t^\nu\bigg] \\ &= -\int_0^T Q_t^{0,\nu}\Big(\kappa_0 \nu_t^0 + \kappa_1 \mathbb{E}[\nu_t^{1,*}(\nu^0)] - \bar{\mu}_t\Big) dt. \quad (9.9)\end{aligned}$$

Moreover, it follows from (2.3) that

$$\int_0^T Q_t^{0,\nu^0} \nu_t^0 dt = \frac{q_0^2}{2}. \quad (9.10)$$

By substituting (9.3), (9.9) and (9.10) into (9.8), we get

$$\begin{aligned}\mathbb{E}[X_T^{0,\nu^0}] = x_0 + M_0 q_0 - \kappa_0 \frac{q_0^2}{2} - \int_0^T Q_t^{0,\nu^0}\big(\kappa_1 \mathbb{E}[\nu_t^{1,*}(\nu^0)] - \bar{\mu}_t\big) dt \\ - \int_0^T \lambda_0 (\nu_t^0)^2 dt. \quad (9.11)\end{aligned}$$

Notice that from Proposition 9.4 and Lemma 9.2, we have

$$\mathbb{E}[\nu_t^{1,*}(\nu^0)] = \frac{\kappa_0}{2\lambda_1}\frac{d}{dt}(\boldsymbol{G}\nu^0)(t) - \frac{1}{2\lambda_1}\frac{d}{dt}(\boldsymbol{G}\bar{\mu})(t), \qquad 0 \le t \le T. \quad (9.12)$$

Plugging (9.12) into (9.11) gives

$$\begin{aligned}\mathbb{E}[X_T^{0,\nu^0}] = x_0 + M_0 q_0 - \kappa_0 \frac{q_0^2}{2} - \frac{\kappa_1 \kappa_0}{2\lambda_1}\int_0^T Q_t^{0,\nu^0}\frac{d}{dt}(\boldsymbol{G}\nu^0)(t) dt \\ + \int_0^T Q_t^{0,\nu^0}\bigg(\frac{\kappa_1}{2\lambda_1}\frac{d}{dt}(\boldsymbol{G}\bar{\mu})(t) + \bar{\mu}_t\bigg) dt - \lambda_0 \int_0^T (\nu_t^0)^2 dt. \quad (9.13)\end{aligned}$$

Since $\bar{\mu} \in L^2([0, T])$ by (3.7) and (3.8) and also $\nu^0 \in L^2([0, T])$ by (2.2), we get from Proposition 9.4 that $(\boldsymbol{G}\bar{\mu})(0) = (\boldsymbol{G}\nu^0)(0) = 0$. Then one more integration by parts and recalling that $Q_T^{0,\nu^0} = 0$ gives

$$\int_0^T Q_t^{0,\nu^0}\frac{d}{dt}(\boldsymbol{G}\nu^0)(t) dt = \int_0^T \nu_t^0 (\boldsymbol{G}\nu^0)(t) dt \quad (9.14)$$

as well as

$$\int_0^T Q_t^{0,\nu^0}\frac{d}{dt}(\boldsymbol{G}\bar{\mu})(t) dt = \int_0^T \nu_t^0 (\boldsymbol{G}\bar{\mu})(t) dt. \quad (9.15)$$

Hence by plugging (9.14) and (9.15) into (9.13), we obtain

$$\mathbb{E}[X_T^{0,\nu^0}] = x_0 + M_0 q_0 - \kappa_0 \frac{q_0^2}{2} - \frac{\kappa_1 \kappa_0}{2\lambda_1} \int_0^T \nu_t^0 (\boldsymbol{G}\nu^0)(t) dt$$
$$+ \int_0^T \left(\frac{\kappa_1}{2\lambda_1} \nu_t^0 (\boldsymbol{G}\bar{\mu})(t) + Q_t^{0,\nu^0} \bar{\mu}_t \right) dt - \lambda_0 \int_0^T (\nu_t^0)^2 dt,$$

which together with (2.10) proves the result. □

10 Convergence of the numerical scheme

We now discuss the convergence of the numerical scheme proposed in Sect. 5 under Assumption 3.1. Our first goal is to prove Proposition 5.2, but before getting to the proof, we introduce an auxiliary result.

Lemma 10.1 *Let $(\mathcal{G}_n)_{n\geq 1}$ be defined as in (5.3). Then $\mathcal{G}_n$ is in $L^2([0,T]^2)$ for any $n \geq 1$.*

Proof Recall that $\boldsymbol{G}$ was defined in (3.11). By Proposition 7.4, $\boldsymbol{G}$ is an operator in $B(L^2([0,T]))$. Recall that $(a_i)_{i\geq 1}$ is a complete orthonormal basis in $L^2([0,T])$; hence (5.2) implies that the $(b_i)_{i\geq 1}$ are in $L^2([0,T])$. We therefore get from (5.3) that

$$\int_0^T \int_0^T \mathcal{G}_n(t,s)^2 ds dt \leq n \sum_{i=1}^n \left(\int_0^T a_i^2(t) dt \right) \left(\int_0^T b_i^2(s) ds \right) < \infty,$$

and the result follows. □

We provide a proof of Proposition 5.2.

Proof of Proposition 5.2 The result follows directly from Lemma 10.1 and (5.4). □

Before we prove Proposition 5.4, we need to introduce the following result presented in Atkinson [3, Theorem 2.1.1].

Theorem 10.2 *Let $\boldsymbol{G}$ be in $B(L^2([0,T]))$ and $\lambda \in \mathbb{R}$. Assume that $\boldsymbol{I} - \lambda \boldsymbol{G}$ is invertible on $L^2([0,T])$. Furthermore, assume that $(\boldsymbol{G}_n)_{n\geq 1}$ is a sequence of operators in $B(L^2([0,T]))$ with*

$$\lim_{n\to\infty} \|\boldsymbol{G} - \boldsymbol{G}_n\| = 0.$$

Then the following hold:

(i) *There exists $N \geq 1$ such that for all $n \geq N$, the operator $(\boldsymbol{I} - \lambda \boldsymbol{G}_n)^{-1}$ exists and is in $B(L^2([0,T]))$.*

(ii) $(\boldsymbol{I} - \lambda \boldsymbol{G}_n)^{-1}$ *converges to $(\boldsymbol{I} - \lambda \boldsymbol{G})^{-1}$ in $B(L^2([0,T]))$, that is,*

$$\lim_{n\to\infty} \|(\boldsymbol{I} - \lambda \boldsymbol{G}_n)^{-1} - (\boldsymbol{I} - \lambda \boldsymbol{G})^{-1}\| = 0.$$

(iii) $\|(\boldsymbol{I} - \lambda \boldsymbol{G}_n)^{-1}\|$ *converges to* $\|(\boldsymbol{I} - \lambda \boldsymbol{G})^{-1}\|$, *that is*,

$$\lim_{n\to\infty} \|(\boldsymbol{I} - \lambda \boldsymbol{G}_n)^{-1}\| = \|(\boldsymbol{I} - \lambda \boldsymbol{G})^{-1}\|.$$

We define

$$\boldsymbol{R}_n := \left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n\right)^{-1}, \qquad n \geq 1. \tag{10.1}$$

We are now ready to prove Proposition 5.4.

Proof of Proposition 5.4 Recall that $(\boldsymbol{G}_n)_{n\geq 1}$ was defined in (5.5) and $\boldsymbol{G}$ in (3.11). From Propositions 5.2, 7.4 and 7.9, it follows that the assumptions of Theorem 10.2 hold. Hence there exists $N \geq 1$ such that for all $n \geq N$, the operator $\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n$ is invertible. Since the corresponding kernels $\mathcal{G}_n$ in (5.3) are degenerate for any $n \geq N$, it follows from [3, Theorem 2.1.2] that the matrices $I_n + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathbb{G}_n$ are invertible (recall (5.8) for the definition of $\mathbb{G}_n$).

Let $g, \psi \in L^2([0, T])$ and define for any $n \geq N$

$$\gamma_i = -\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\sum_{j=1}^{n}\left(I_n + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\mathbb{G}_n\right)^{-1}_{ij}\langle \psi, b_j\rangle_{L^2}, \qquad i = 1, \ldots, n.$$

As shown in Porter and Stirling [23, Eqs. (3.5)–(3.7)], the unique solution to

$$\left(\boldsymbol{I} + \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\boldsymbol{G}_n\right)g = \psi$$

is given by

$$g(t) = \psi(t) + \sum_{i=1}^{n}\gamma_i a_{i,n}(t), \qquad 0 \leq t \leq T, n \geq N,$$

and (5.9) follows. □

Before proving Proposition 5.5, we need to present two intermediate results.

Lemma 10.3 *Let* $\boldsymbol{R}$ *be defined as in* (3.14) *and* $(\boldsymbol{R}_n)_{n\geq 1}$ *as in* (10.1). *Then the following hold*:

(i) $\liminf_{n\to\infty}\langle \boldsymbol{R}_n\mathbf{1}, \mathbf{1}\rangle_{L^2} > 0$.

(ii) $\lim_{n\to\infty}\dfrac{1}{\langle \boldsymbol{R}_n\mathbf{1}, \mathbf{1}\rangle_{L^2}} = \dfrac{1}{\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2}}$.

Proof (i) We have shown in the proof of Proposition 5.4 that the assumptions of Theorem 10.2 hold; hence there exists $N \geq 1$ such that the operators $\boldsymbol{R}_n$ exist for all $n \geq N$. From the Cauchy–Schwarz inequality and since $\|\mathbf{1}\|_{L^2} = T$, we get

$$|\langle \boldsymbol{R}_n\mathbf{1}, \mathbf{1}\rangle_{L^2} - \langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2}| \leq \|\boldsymbol{R}_n - \boldsymbol{R}\|\|\mathbf{1}\|^2_{L^2} \leq \|\boldsymbol{R}_n - \boldsymbol{R}\|T^2. \tag{10.2}$$

By Proposition 7.9(ii), the operator $\boldsymbol{R}$ is bounded from below. Therefore by Definition 7.3, there exists $\varepsilon > 0$ such that

$$\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2} > \varepsilon. \tag{10.3}$$

By Theorem 10.2(iii), there exists $N_1 \geq N$ such that for all $n \geq N_1$, we have

$$\|\boldsymbol{R}_n - \boldsymbol{R}\| < \varepsilon T^{-2}/2. \tag{10.4}$$

Now (10.2)–(10.4) yield (i).

(ii) This follows directly from (10.2), (10.3) and (i). □

Next we prove the convergence of the sequence $(\eta_n)_{n\geq 1}$ of constants from (5.10).

Lemma 10.4 *Let η and η_n be defined as in (3.16) and (5.10), respectively. Then there exists $N \geq 1$ such that for all $n \geq N$, the constant η_n is well defined. Moreover,*

$$\lim_{n\to\infty} \eta_n = \eta. \tag{10.5}$$

Proof From Lemma 10.3(i), (3.16) and (10.1), it follows that η_n is well defined for all n sufficiently large. Furthermore, we remind the reader that the constant η is also well defined as shown in the proof of Theorem 3.5, step 2). From the Cauchy–Schwarz inequality, we get

$$|\langle \boldsymbol{R}_n \boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2} - \langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}| \leq \|\boldsymbol{R} - \boldsymbol{R}_n\| \|\boldsymbol{S}\bar{\mu}\|_{L^2} \|\mathbf{1}\|_{L^2},$$

and together with Theorem 10.2(iii) and Lemma 7.10, it follows that

$$\lim_{n\to\infty} |\langle \boldsymbol{R}_n \boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2} - \langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}| = 0. \tag{10.6}$$

From (3.16) and (5.10), we have

$$|\eta_n - \eta| = \left| \frac{\langle \boldsymbol{R}_n \boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}}{\langle \boldsymbol{R}_n \mathbf{1}, \mathbf{1}\rangle_{L^2}} - \frac{\langle \boldsymbol{R}\boldsymbol{S}\bar{\mu}, \mathbf{1}\rangle_{L^2}}{\langle \boldsymbol{R}\mathbf{1}, \mathbf{1}\rangle_{L^2}} \right|,$$

and so (10.5) follows from Lemma 10.3 and (10.6). □

We are now ready to prove Proposition 5.5.

Proof of Proposition 5.5 From (3.15), (5.7) and (10.1), we have

$$\nu^{0,*} - \nu^{0,(n)} = (\boldsymbol{R} - \boldsymbol{R}_n)(\boldsymbol{S}\bar{\mu}) + \frac{1}{2\lambda_0}(\eta \boldsymbol{R}\mathbf{1} - \eta_n \boldsymbol{R}_n \mathbf{1}).$$

It follows that

$$\begin{aligned} \|\nu^{0,*} - \nu^{0,(n)}\|_{L^2} \leq \|(\boldsymbol{R} - \boldsymbol{R}_n)(\boldsymbol{S}\bar{\mu})\|_{L^2} + \frac{\eta}{2\lambda_0}\|(\boldsymbol{R} - \boldsymbol{R}_n)\mathbf{1}\|_{L^2} \\ + \frac{1}{2\lambda_0}|\eta_n - \eta| \|\boldsymbol{R}_n \mathbf{1}\|_{L^2}. \end{aligned}$$

Hence by Lemma 10.4, Theorem 10.2(iii) and following similar lines as in the proof of (10.6), we get

$$\lim_{n\to\infty} \|\nu^{0,*} - \nu^{0,(n)}\|_{L^2} = 0. \qquad \square$$

Finally, we prove Theorem 5.6.

Proof of Theorem 5.6 Throughout the proof, we consider n large enough such that the results of Lemma 10.4 and Proposition 5.5 hold, even if this is not stated explicitly.

(i) From (5.11) and (7.14), we get

$$\begin{aligned} \nu_t^{0,*} - \hat{\nu}_t^{0,(n)} &= -\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\big(\boldsymbol{G}(\nu^{0,*} - \nu^{0,(n)})\big)(t) + \frac{\eta - \eta_n}{2\lambda_0} \\ &= -\frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\int_0^T \mathcal{G}(t,s)(\nu_s^{0,*} - \nu_s^{0,(n)})ds + \frac{\eta-\eta_n}{2\lambda_0}, \end{aligned}$$

where we used (3.11) in the second equality. From the Cauchy–Schwarz inequality, we get for $0 \le t \le T$ and n sufficiently large that

$$|\nu_t^{0,*} - \hat{\nu}_t^{0,(n)}| \le \frac{\kappa_1\kappa_0}{2\lambda_0\lambda_1}\left(\int_0^T |\mathcal{G}(t,s)|^2 ds\right)^{1/2} \|\nu^{0,*} - \nu^{0,(n)}\|_{L^2} + \frac{|\eta-\eta_n|}{2\lambda_0}.$$

Hence Proposition 8.1, Lemma 10.4 and Proposition 5.5 yield (i).

(ii) By Proposition 6.5, r^1 is bounded over $[0,T]$ and by Lemma 6.7, the kernel $\mathcal{K}$ is bounded over $[0,T]^2$. Together with (3.6), we get that there exists a constant $C > 0$ such that for $0 \le t \le T$ and n sufficiently large, we have

$$\begin{aligned} |\nu_t^{1,*}(\nu^{0,*}) - \nu_t^{1,*}(\hat{\nu}^{0,(n)})| &\le |r_t^0(\hat{\nu}^{0,(n)}) - r_t^0(\nu^{0,*})| \\ &\quad + C\left|\int_0^t \big(r_s^0(\hat{\nu}^{0,(n)}) - r_s^0(\nu^{0,*})\big)ds\right|. \end{aligned} \tag{10.7}$$

By plugging (3.5) into (10.7), we observe that the net contribution of the terms depending on the finite-variation process A is zero. Hence we conclude that

$$\begin{aligned} |\nu_t^{1,*}(\nu^{0,*}) - \nu_t^{1,*}(\hat{\nu}^{0,(n)})| &\le C_1\int_t^T |\hat{\nu}_s^{0,(n)} - \nu_s^{0,*}|ds \\ &\quad + C_2\int_0^t\left(\int_s^T |\hat{\nu}_r^{0,(n)} - \nu_r^{0,*}|dr\right)ds \end{aligned} \tag{10.8}$$

for some constants $C_1, C_2 > 0$ independent from n and t. Then (ii) follows from (10.8) and (i). $\square$

Appendix A: An example of the spectral decomposition of $\boldsymbol{G}$

In this section, we give an example of the spectral decomposition of $\boldsymbol{G}$ in Lemma 8.3 for the case where $\phi^1 = 0$. We continue to impose Assumption 3.1.

Lemma A.1 *Let $\psi \in C([0,T])$ and recall $\boldsymbol{K}_1$ defined in (7.3). Then $\boldsymbol{K}_1$ satisfies for $0 \le t \le T$ that*

$$\frac{d}{dt}(\boldsymbol{K}_1^*\psi)(t) = -r_t^1(\boldsymbol{K}_1^*\psi)(t) - \psi(t), \qquad (\boldsymbol{K}_1^*\psi)(T) = 0.$$

In particular, $t \mapsto (\boldsymbol{K}_1^*\psi)(t)$ is continuously differentiable on $[0,T]$.

Proof This goes as the proof of Proposition 9.4; hence we just give an outline. We take the derivative of $\boldsymbol{K}_1^*\psi$ with respect to time using (7.4), (3.3) and (3.4) to get for $0 \le t \le T$ that

$$\frac{d}{dt}(\boldsymbol{K}_1^*\psi)(t) = \frac{d}{dt}\left(\xi_t^- \int_t^T \xi_s^+ \psi(s)ds\right) = -r_t^1(\boldsymbol{K}_1^*\psi)(t) - \psi(t).$$

Note that (7.4) implies that

$$(\boldsymbol{K}_1^*\psi)(T) = 0. \qquad \square$$

Proposition A.2 *Let $\boldsymbol{G}$ be defined as in (3.11) and assume $\phi^1 \equiv 0$. Let $(z_n)_{n\ge1}$ be the increasing sequence of real positive roots of the equation*

$$\cot(z) = -\frac{2\alpha - \kappa_1}{\lambda_1}\frac{T}{z}. \tag{A.1}$$

Then the eigenvalues $(\zeta_n)_{n\ge1}$ and eigenfunctions $(\psi_n)_{n\ge1}$ of $\boldsymbol{G}$ are given by

$$\psi_n(t) = \frac{2}{\sqrt{\zeta_n}} \frac{\sin(\frac{t}{\sqrt{\zeta_n}})}{\sqrt{\frac{2T}{\sqrt{\zeta_n}} - \sin(\frac{2T}{\sqrt{\zeta_n}})}}, \qquad \zeta_n = \frac{T^2}{z_n^2}. \tag{A.2}$$

Proof We first show that the eigenvalues $(\zeta_n)_{n\ge1}$ and eigenfunctions $(\psi_n)_{n\ge1}$ arise from solutions to an ODE. Then we show that the solutions of the ODE can be determined in terms of the roots to (A.1).

Let ζ be an eigenvalue of $\boldsymbol{G}$ and ψ the corresponding eigenfunction, i.e., ζ and ψ satisfy

$$(\boldsymbol{G}\psi)(t) = \zeta\psi(t), \qquad 0 \le t \le T. \tag{A.3}$$

From Lemma 8.3, it follows that $\zeta > 0$ and $\psi \in L^2([0,T])$. Proposition 9.4 shows that $t \mapsto (\boldsymbol{G}\psi)(t)$ is continuously differentiable on $[0,T]$; therefore it follows from (A.3) that $t \mapsto \psi(t)$ is continuously differentiable on $[0,T]$. We take the derivative on both sides of (A.3) to obtain that (ζ,ψ) must satisfy

$$\frac{d}{dt}(\boldsymbol{G}\psi)(t) = \zeta\psi'(t), \qquad 0 \le t \le T. \tag{A.4}$$

Proposition 9.4 shows that $\boldsymbol{G}\psi$ is the solution to (9.5); therefore we can substitute (9.5) in (A.4) to obtain that ψ must satisfy

$$r_t^1(\boldsymbol{G}\psi)(t) + (\boldsymbol{K}_1^*\psi)(t) = \zeta\psi'(t), \qquad 0 \le t \le T. \tag{A.5}$$

By Proposition 6.5, r^1 is the solution to (6.16). When $\phi^1 \equiv 0$, it can be computed explicitly as

$$r_t^1 = \frac{2\alpha - \kappa_1}{(t-T)(2\alpha - \kappa_1) - 2\lambda_1}, \qquad 0 \le t \le T.$$

Note that under Assumption 3.1, r^1 is continuously differentiable on $[0, T]$.

Since we have proved that $t \mapsto \psi(t)$ is continuous, it follows from Proposition A.1 that $t \mapsto (\boldsymbol{K}_1^* \psi)(t)$ is continuously differentiable on $[0, T]$. We take the derivative on both sides of (A.5) to get

$$\frac{dr_t^1}{dt}(\boldsymbol{G}\psi)(t) + r_t^1 \frac{d}{dt}(\boldsymbol{G}\psi)(t) + \frac{d}{dt}(\boldsymbol{K}_1^*\psi)(t) = \zeta \psi''(t)$$

and then use (9.5) to show that ψ satisfies for $0 \le t \le T$ that

$$\frac{dr_t^1}{dt}(\boldsymbol{G}\psi)(t) + (r_t^1)^2(\boldsymbol{G}\psi)(t) + r_t^1(\boldsymbol{K}_1^*\psi)(t) + \frac{d}{dt}(\boldsymbol{K}_1^*\psi)(t) = \zeta \psi''(t). \tag{A.6}$$

Applying (6.16), (A.5) and (A.1) to (A.6) yields that ψ must satisfy

$$-\psi(t) = \zeta \psi''(t), \qquad 0 \le t \le T.$$

Recall that $\zeta > 0$; hence it follows from (A.3) and Proposition 9.4 that ψ satisfies the initial condition $\psi(0) = 0$. The terminal condition $\psi'(T) = -(\frac{2\alpha - \kappa_1}{\lambda_1})\psi(T)$ follows by combining (A.5) with (6.16), (A.3) and $(\boldsymbol{K}_1^*\psi)(T) = 0$ (see Proposition A.1). It follows that (ζ, ψ) satisfy

$$\begin{aligned}
\psi''(t) &= -\frac{1}{\zeta}\psi(t), \qquad 0 < t < T, \\
\psi(0) &= 0, \\
\psi'(T) &= -\frac{2\alpha - \kappa_1}{\lambda_1}\psi(T).
\end{aligned} \tag{A.7}$$

We show that (A.1) has an infinite number of positive roots. To see this, note that since $2\alpha - \kappa_1 \ge 0$ by Assumption 3.1, we have for any $n \ge 1$ that

$$\begin{aligned}
\lim_{z \searrow (n-1)\pi} \cot(z) + \frac{2\alpha - \kappa_1}{\lambda_1}\frac{T}{z} &= +\infty, \\
\lim_{z \nearrow (n-1)\pi} \cot(z) + \frac{2\alpha - \kappa_1}{\lambda_1}\frac{T}{z} &= -\infty.
\end{aligned}$$

Since $z \mapsto \cot(z) + \frac{2\alpha - \kappa_1}{\lambda_1}\frac{T}{z}$ is continuous over the interval $((n-1)\pi, n\pi)$ for any $n \ge 1$, it follows by the intermediate value theorem that (A.1) must have a root in the interval $((n-1)\pi, n\pi)$ for any $n \ge 1$.

Next we identify ζ_n as in (A.2). Let z_n be the nth positive root of (A.1) and let ζ_n, ψ_n be defined as in (A.2). First note that since $z_n > 0$, we have

$$2z_n - \sin(2z_n) > 0. \tag{A.8}$$

From (A.2) and (A.8), it follows that

$$\frac{2T}{\sqrt{\zeta_n}} - \sin\left(\frac{2T}{\sqrt{\zeta_n}}\right) > 0,$$

and so the function $t \mapsto \psi_n(t)$ is well defined on $[0, T]$ and $\|\psi_n\|_{L^2} = 1$. Using the identity – which arises from (A.7) –

$$\cos(z_n) = -\frac{2\alpha - \kappa_1}{\lambda_1}\frac{T}{z_n}\sin(z_n),$$

it is easy to verify that ψ_n in (A.2) solves (A.7) with $\zeta = \zeta_n$ for any $n \geq 1$. This completes the proof. □

Appendix B: Proofs of Lemmas 6.3 and 6.4

Proof of Lemma 6.3 Let $\epsilon > 0$ and $\nu^1, \omega \in \mathcal{A}_m$. We note that (2.5) implies that

$$Q_t^{1,\nu^1+\epsilon\omega} = Q_t^{1,\nu^1} - \epsilon \int_0^t \omega_s ds, \qquad 0 \leq t \leq T. \tag{B.1}$$

We use the alternative representation of H^1 in (6.1) and (B.1) to get

$$\begin{aligned}
&H^1(\nu^1 + \epsilon\omega) - H^1(\nu^1) \\
&= \epsilon\mathbb{E}\bigg[\int_0^T \omega_t(-2\lambda_1\nu_t^1 + 2\alpha Q_T^{1,\nu^1} - \kappa_1 Q_t^{1,\nu^1})dt \\
&\qquad + \int_0^T \left(\int_0^t \omega_s ds\right)(2\phi_t^1 Q_t^{1,\nu^1} dt + \kappa_0\nu_t^0 dt + \kappa_1\nu_t^1 dt - dA_t)\bigg] \\
&\quad + \epsilon^2\mathbb{E}\bigg[-\lambda_1\int_0^T \omega_s^2 ds - \alpha\left(\int_0^T \omega_s ds\right)^2 \\
&\qquad - \int_0^T \phi_t^1\left(\int_0^t \omega_s ds\right)^2 dt + \kappa_1\int_0^T \omega_t\left(\int_0^t \omega_s ds\right)dt\bigg].
\end{aligned} \tag{B.2}$$

From (6.10) and (B.2), we get

$$\begin{aligned}
&\langle \mathcal{D}H^1(\nu^1), \omega\rangle \\
&= \mathbb{E}\bigg[\int_0^T \omega_t(-2\lambda_1\nu_t^1 + 2\alpha Q_T^{1,\nu^1} - \kappa_1 Q_t^{1,\nu^1})dt \\
&\qquad + \int_0^T \left(\int_0^t \omega_s ds\right)(2\phi_t^1 Q_t^{1,\nu^1} dt + \kappa_0\nu_t^0 dt + \kappa_1\nu_t^1 dt - dA_t)\bigg].
\end{aligned} \tag{B.3}$$

Since $\nu^1, \omega \in \mathcal{A}_m$, $\nu^0 \in \mathcal{A}_M^{q_0}$ and $\mathbb{E}[(\int_0^T |dA_t|)^2] < \infty$, we can use Fubini's theorem in (B.3) to get

$$\langle \mathcal{D}H^1(\nu^1), \omega \rangle = \mathbb{E}\bigg[\int_0^T \omega_t \bigg(-2\lambda_1 \nu_t^1 + 2\alpha Q_T^{1,\nu^1} - \kappa_1 Q_t^{1,\nu^1} + A_t - A_T + \int_t^T (2\phi_s^1 Q_s^{1,\nu^1} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^1) ds \bigg) dt \bigg],$$

which concludes the proof. □

Proof of Lemma 6.4 In Lemma 6.2, we have shown that under Assumption 3.1, H^1 is strictly concave on $\mathcal{A}_m$. Therefore we may apply Ekeland and Témam [11, Proposition II.2.1] to obtain that

$$\langle \mathcal{D}H^1(\nu^{1,*}), \omega \rangle = 0 \text{ for all } \omega \in \mathcal{A}_m \quad \Longleftrightarrow \quad \nu^{1,*} = \arg\sup_{\nu \in \mathcal{A}_m} H^1(\nu). \tag{B.4}$$

The strict concavity of H^1 guarantees that the optimiser $\nu^{1,*}$ is unique.

We first prove necessity, that is, if $\nu^{1,*}$ is the maximiser of H^1, then it satisfies (6.12). We assume that

$$\nu^{1,*} = \arg\sup_{\nu \in \mathcal{A}_m} H^1(\nu).$$

Then (B.4) and (6.11) imply that for all $\omega \in \mathcal{A}_m$, we have

$$\langle \mathcal{D}H^1(\nu^{1,*}), \omega \rangle = \mathbb{E}\bigg[\int_0^T \omega_t \bigg(-2\lambda_1 \nu_t^{1,*} + 2\alpha Q_T^{1,\nu^{1,*}} - \kappa_1 Q_t^{1,\nu^{1,*}} + A_t - A_T + \int_t^T (2\phi_s^1 Q_s^{1,\nu^{1,*}} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^{1,*}) ds \bigg) dt \bigg] = 0.$$

By applying the optional projection theorem, we get

$$\mathbb{E}\bigg[\int_0^T \omega_t \bigg(-2\lambda_1 \nu_t^{1,*} + \mathbb{E}_t[2\alpha Q_T^{1,\nu^{1,*}} - A_T] - \kappa_1 Q_t^{1,\nu^{1,*}} + A_t + \mathbb{E}_t\bigg[\int_t^T (2\phi_s^1 Q_s^{1,\nu^{1,*}} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^{1,*}) ds\bigg] \bigg) dt \bigg] = 0. \tag{B.5}$$

As (B.5) holds for all $\omega \in \mathcal{A}_m$ we deduce the first-order condition

$$0 = -2\lambda_1 \nu_t^{1,*} + \mathbb{E}_t[2\alpha Q_T^{1,\nu^{1,*}} - A_T] - \kappa_1 Q_t^{1,\nu^{1,*}} + A_t + \mathbb{E}_t\bigg[\int_t^T (2\phi_s^1 Q_s^{1,\nu^{1,*}} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^{1,*}) ds\bigg] \tag{B.6}$$

$(d\mathbb{P} \otimes dt)$-a.e. on $\Omega \times [0, T]$. We define the martingales

$$\mathcal{M}_t := \mathbb{E}_t\bigg[\int_0^T (2\phi_s^1 Q_s^{1,\nu^{1,*}} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^{1,*}) ds\bigg],$$
$$\mathcal{N}_t := \mathbb{E}_t[2\alpha Q_T^{1,\nu^{1,*}} - A_T].$$

Note that $\mathcal{M}$ and $\mathcal{N}$ are square-integrable since $\mathbb{E}[(\int_0^T |dA_t|)^2] < \infty$, $\nu^{1,*}, \omega \in \mathcal{A}_m$ and $\nu^0 \in \mathcal{A}_M^{q_0}$. We plug $\mathcal{M}$ and $\mathcal{N}$ into (B.6) to get

$$\begin{aligned} 0 = -2\lambda_1 \nu_t^{1,*} + \mathcal{N}_t - \kappa_1 Q_t^{1,\nu^{1,*}} + A_t + \mathcal{M}_t \\ - \int_0^t (2\phi_s^1 Q_s^{1,\nu^{1,*}} + \kappa_0 \nu_s^0 + \kappa_1 \nu_s^{1,*}) ds. \end{aligned} \tag{B.7}$$

From (2.5) and (B.7), it follows that $\nu^{1,*}$ solves the BSDE

$$d\nu_t^{1,*} = \frac{1}{2\lambda_1} d\mathcal{N}_t + \frac{1}{2\lambda_1} d\mathcal{M}_t - \frac{1}{\lambda_1}\phi_t^1 Q_t^{1,\nu^{1,*}} dt - \frac{\kappa_0}{2\lambda_1}\nu_t^0 dt + \frac{1}{2\lambda_1} dA_t$$

with $\nu_T^{1,*} = \frac{2\alpha-\kappa_1}{2\lambda_1} Q_T^{1,\nu^{1,*}}$. This gives (6.12).

Next we prove sufficiency, that is, if $\nu^{1,*}$ satisfies (6.12), then it is a maximiser of H^1. Assume that $(Q^{1,\nu^{1,*}}, \nu^{1,*})$ solves (6.12) $(d\mathbb{P} \otimes dt)$-a.e. and that $\nu^{1,*} \in \mathcal{A}_m$. We show that $\langle \mathcal{D}H^1(\nu^{1,*}), \omega\rangle$ vanishes for all $\omega \in \mathcal{A}_m$, which combined with (B.4) implies that $\nu^{1,*}$ is the solution to the minor agent's problem. Since $(Q^{1,\nu^{1,*}}, \nu^{1,*})$ solves (6.12), we get

$$\begin{aligned} 2\lambda_1 \nu_t^{1,*} &= \mathbb{E}_t[(2\alpha - \kappa_1) Q_T^{1,\nu^{1,*}}] - \mathbb{E}_t\bigg[\int_t^T dA_s\bigg] + \mathbb{E}_t\bigg[\int_t^T (\kappa_0 \nu_s^0 + 2\phi_t^1 Q_s^{1,\nu^{1,*}}) ds\bigg] \\ &= \mathbb{E}_t[2\alpha Q_T^{1,\nu^{1,*}} - A_T] - \kappa_1 Q_t^{1,\nu^{1,*}} + A_t \\ &\quad + \mathbb{E}_t\bigg[\int_t^T (\kappa_0 \nu_s^0 + 2\phi_t^1 Q_s^{1,\nu^{1,*}} + \kappa_1 \nu_s^{1,*}) ds\bigg] \qquad (d\mathbb{P} \otimes dt)\text{-a.e.}, \end{aligned}$$

where we used (2.5) in the second equality. Hence $\nu^{1,*}$ satisfies (B.6), and so the left-hand side of (B.4) holds. □

Appendix C: Proofs of Proposition 6.5 and Lemma 6.7

Before we prove Lemma 6.7, we introduce the following result.

Lemma C.1 *Under Assumption 3.1, the Riccati equation (6.16) has a unique continuous solution.*

Proof Let $\hat{r}^1$ be the solution to the equation, for $0 \leq t \leq T$,

$$\partial_t \hat{r}_t^1 = -\frac{1}{\lambda_1}\phi_t^1 + (\hat{r}_t^1)^2, \qquad \hat{r}_T^1 = \frac{2\alpha - \kappa_1}{2\lambda_1}. \tag{C.1}$$

Since $\hat{r}_T^1 \geq 0$ and ϕ^1 is a piecewise continuous, locally bounded nonnegative function on $[0, T]$, there exists by Wonham [29, Theorem 2.1] a unique solution $\hat{r}^1$, which is absolutely continuous on $[0, T]$ (see also Freiling [13, Theorem 3.5] for a more recent reference). As stated by [29], the function $\hat{r}^1$ satisfies (C.1) only dt-almost everywhere. Note that by taking $r_t^1 = -\hat{r}_t^1$, it follows that r^1 is an absolutely continuous solution to (6.16). Uniqueness of the solution to (6.16) then follows by the uniqueness for (C.1). □

We are now ready to prove Lemma 6.7.

Proof of Lemma 6.7 Lemma C.1 proves that r^1 is continuous on $[0, T]$. Then it follows from (3.3) that the functions $t \mapsto \xi_t^{\pm}$ are continuous. Therefore we get from (3.4) that $\mathcal{K}$ is jointly continuous on $[0, T]^2$; hence it is bounded on $[0, T]^2$ and so $\mathcal{K} \in L^2([0, T]^2)$. □

Now we are ready to prove Proposition 6.5.

Proof of Proposition 6.5 In Lemma C.1, we have established that (6.16) has a unique continuous solution r^1. We prove the rest of the claims in the following two steps.

1) We show that r^0 given by (3.5) solves the BSDE (6.17). Note that since r^1 is continuous, the function ξ^+ in (3.3) is the unique solution of the ODE

$$\frac{d\xi_t^+}{dt} = r_t^1 \xi_t^+, \qquad \xi_0^+ = 1. \tag{C.2}$$

Since r^1 is continuous on $[0, T]$, it holds that

$$\int_0^T (\xi_t^+)^2 dt < \infty. \tag{C.3}$$

Since ξ^+ satisfies (C.3), the process

$$\xi_t^+ r_t^0 := \frac{1}{2\lambda_1}\mathbb{E}_t\left[\int_t^T \xi_s^+ (dA_s - \kappa_0 \nu_s^0 ds)\right], \qquad 0 \leq t \leq T, \tag{C.4}$$

is the unique strong solution to the linear BSDE

$$d(\xi_t^+ r_t^0) = \frac{\xi_t^+}{2\lambda_1}(\kappa_0 \nu_t^0 dt - dA_t) - \frac{1}{2\lambda_1}\xi_t^+ d\mathcal{M}_t - \frac{1}{2\lambda_1}\xi_t^+ d\mathcal{N}_t, \qquad \xi_T^+ r_T^0 = 0. \tag{C.5}$$

We multiply both sides in (C.4) by ξ_t^- from (3.3) and use the identity $\xi_t^- \xi_t^+ = 1$. By doing so, we obtain for r_t^0 the expression

$$r_t^0 = \frac{1}{2\lambda_1}\mathbb{E}_t\left[\int_t^T \xi_t^- \xi_s^+ (dA_s - \kappa_0 \nu_s^0 ds)\right]. \tag{C.6}$$

We now show that r^0 from (C.6) is the solution to (6.17). From (C.5) and Itô's product rule, we get

$$r_t^0 d\xi_t^+ + \xi_t^+ dr_t^0 = \frac{\xi_t^+}{2\lambda_1}(\kappa_0 \nu_t^0 dt - dA_t) - \frac{1}{2\lambda_1}\xi_t^+ d\mathcal{M}_t - \frac{1}{2\lambda_1}\xi_t^+ d\mathcal{N}_t. \tag{C.7}$$

Next we use (C.2) and (C.7) to get

$$\xi_t^+ \left(r_t^1 r_t^0 dt + dr_t^0 - \frac{1}{2\lambda_1}(\kappa_0 \nu_t^0 dt - dA_t) + \frac{1}{2\lambda_1} d\mathcal{M}_t + \frac{1}{2\lambda_1} d\mathcal{N}_t \right) = 0.$$

Since $\xi_t^+ > 0$ for $t \in [0, T]$, we have

$$r_t^1 r_t^0 dt + dr_t^0 - \frac{1}{2\lambda_1}(\kappa_0 \nu_t^0 dt - dA_t) + \frac{1}{2\lambda_1} d\mathcal{M}_t + \frac{1}{2\lambda_1} d\mathcal{N}_t = 0$$

with terminal condition $r_T^0 = 0$ which follows from (C.6). By comparing this with (6.17), it follows that r^0 is the solution to the BSDE (6.17). Recall the definition of $\mathcal{K}$ in (3.4). We substitute the expression for $\mathcal{K}(t, s)$ from (3.4) into (C.6), from which it follows that r^0 given by (3.5) solves (6.17).

2) We show (6.18). From (3.4), it follows that $\mathcal{K}(t, s) > 0$ for $t, s \in [0, T]^2$. Moreover, $\mathcal{K}$ is bounded on $[0, T]^2$ by Lemma 6.7. Since A is of finite variation and $\mathcal{K}$ is bounded, the conditional Jensen inequality and the tower property give

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left(\mathbb{E}_t\left[\int_t^T \mathcal{K}(t, s) dA_s \right] \right)^2 \right] \le C\mathbb{E}\left[\left(\int_0^T |dA_s| \right)^2 \right] < \infty, \tag{C.8}$$

where we used (2.1) in the last inequality. Similarly, we can obtain that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left(\mathbb{E}_t\left[\int_t^T \mathcal{K}(t, s) \nu_s^0 ds \right] \right)^2 \right] \le C\mathbb{E}\left[\int_0^T (\nu_s^0)^2 ds \right] < \infty, \tag{C.9}$$

where we used the fact that $\nu^0 \in \mathcal{A}_M^{q_0}$ and (2.2). From (3.5), (C.8) and (C.9), we get

$$\sup_{t\in[0,T]} \mathbb{E}[(r_t^0)^2] < \infty$$

and (6.18) follows. □

Acknowledgements We are very grateful to the Associate Editor and the anonymous referees for careful reading of the manuscript and for a number of useful comments and suggestions that significantly improved this paper.

Declarations

Competing interests The authors declare no competing interests.

References

1. Admati, A.R., Pfleiderer, P.: A theory of intraday patterns: volume and price variability. Rev. Financ. Stud. **1**, 3–40 (2015)
2. Almgren, R., Chriss, N.: Optimal execution of portfolio transactions. J. Risk **3**(2), 5–39 (2000)
3. Atkinson, K.E.: The Numerical Solution of Integral Equations of the Second Kind. Cambridge University Press, Cambridge (1997)
4. Carlin, B.I., Lobo, M.S., Viswanathan, S.: Episodic liquidity crises: cooperative and predatory trading. J. Finance **65**, 2235–2274 (2007)
5. Cartea, Á., Jaimungal, S.: Incorporating order-flow into optimal execution. Math. Financ. Econ. **10**, 339–364 (2016)
6. Cartea, Á., Sánchez-Betancourt, L.: Brokers and informed traders: Dealing with toxic flow and extracting trading signals. Preprint (2022). Available online at https://ssrn.com/abstract=4265814
7. Casgrain, P., Jaimungal, S.: Mean field games with partial information for algorithmic trading. Preprint (2019). Available online at https://arxiv.org/abs/1803.04094
8. Casgrain, P., Jaimungal, S.: Mean-field games with differing beliefs for algorithmic trading. Math. Finance **30**, 995–1034 (2020)
9. Chen, M., Garriott, C.: High-frequency trading and institutional trading costs. J. Empir. Finance **56**, 74–93 (2020)
10. Cont, R.: Statistical modeling of high-frequency financial data. IEEE Signal Process. **28**(5), 16–25 (2011)
11. Ekeland, I., Témam, R.: Convex Analysis and Variational Problems. SIAM, Philadelphia (1999)
12. Evangelista, D., Thamsten, Y.: On finite population games of optimal trading. Preprint (2020). Available online at https://arxiv.org/abs/2004.00790
13. Freiling, G.: A survey of nonsymmetric Riccati equations. Linear Algebra Appl. **351**(352), 243–270 (2002)
14. Fu, G., Graewe, P., Horst, U., Popier, A.: A mean field game of optimal portfolio liquidation. Math. Oper. Res. **46**, 1235–1657 (2021)
15. Gârleanu, N., Pedersen, L.H.: Dynamic portfolio choice with frictions. J. Econ. Theory **165**, 487–516 (2016)
16. Kirilenko, A., Kyle, A.S., Samadi, M., Tuzun, T.: The flash crash: the impact of high frequency trading on an electronic market. J. Finance **72**, 967–998 (2017)
17. Korajczyk, R.A., Murphy, D.: High-frequency market making to large institutional trades. Rev. Financ. Stud. **32**, 1034 (2018)
18. Lehalle, C.A., Neuman, E.: Incorporating signals into optimal trading. Finance Stoch. **23**, 275–311 (2019)
19. Micheli, A., Muhle-Karbe, J., Neuman, E.: Closed-loop Nash competition for liquidity. Math. Finance **33**, 1082–1118 (2023)
20. Neuman, E., Schied, A.: Protecting pegged currency markets from speculative investors. Math. Finance **32**, 405–420 (2022)
21. Neuman, E., Voß, M.: Trading with the crowd. Math. Finance **33**, 548–617 (2023)
22. Neuman, E., Voß, M.: Optimal signal-adaptive trading with temporary and transient price impact. SIAM J. Financ. Math. **13**, 551–575 (2022)
23. Porter, D., Stirling, D.S.G.: Integral Equations: A Practical Treatment, from Spectral Theory to Applications. Cambridge University Press, Cambridge (1990)
24. Roşu, I.: Fast and slow informed trading. J. Financ. Mark. **43**, 1–30 (2019)

25. Schöneborn, T., Schied, A.: Liquidation in the face of adversity: Stealth vs. sunshine trading. EFA 2008 Athens Meetings Paper (2009). Available online at https://ssrn.com/abstract=1007014
26. Securities, U.S., Commission, E.: Equity market structure literature review. Part II: High frequency trading (2014). Available online at https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf
27. van Kervel, V., Menkveld, A.J.: High-frequency trading around large institutional orders. J. Finance **74**, 1091–1137 (2019)
28. Voß, M.: A two-player portfolio tracking game. Math. Financ. Econ. **16**, 779–809 (2022)
29. Wonham, W.M.: On a matrix Riccati equation of stochastic control. SIAM J. Control **6**, 681–697 (1968)
30. Wood, R., McInish, T., Ord, J.: An investigation of transactions data for NYSE stocks. J. Finance **40**, 723–739 (1985)