

Horst, Ulrich; Kivman, Evgueni

**Article — Published Version**

## Optimal trade execution under small market impact and portfolio liquidation with semimartingale strategies

Finance and Stochastics

*Suggested Citation:* Horst, Ulrich; Kivman, Evgueni (2024) : Optimal trade execution under small market impact and portfolio liquidation with semimartingale strategies, Finance and Stochastics, ISSN 1432-1122, Springer Berlin Heidelberg, Berlin/Heidelberg, Vol. 28, Iss. 3, pp. 759-812, <https://doi.org/10.1007/s00780-024-00536-2>

This Version is available at:

<https://hdl.handle.net/10419/315063>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>



# Optimal trade execution under small market impact and portfolio liquidation with semimartingale strategies

Ulrich Horst<sup>1</sup> · Evgueni Kivman<sup>1</sup>

Received: 6 June 2022 / Accepted: 5 July 2023 / Published online: 14 June 2024

© The Author(s) 2024

## Abstract

We consider an optimal liquidation problem with instantaneous price impact and stochastic resilience for small instantaneous impact factors. Within our modelling framework, the optimal portfolio process converges to the solution of an optimal liquidation problem with general semimartingale controls when the instantaneous impact factor converges to zero. Our results provide a unified framework within which to embed the two most commonly used modelling frameworks in the liquidation literature and provide a foundation for the use of semimartingale liquidation strategies and the use of portfolio processes of unbounded variation. Our convergence results are based on novel convergence results for BSDEs with singular terminal conditions and novel representation results of BSDEs in terms of uniformly continuous functions of forward processes.

**Keywords** Portfolio liquidation · Singular BSDE · Stochastic liquidity · Singular control

**Mathematics Subject Classification** 93E20 · 91B70 · 60H30

**JEL Classification** G11 · G12 · G19

## 1 Introduction

The impact of limited liquidity on optimal trade execution has been extensively analysed in the mathematical finance and stochastic control literature in recent years. The majority of the optimal portfolio liquidation literature allows one of two possible price impacts. The first approach, pioneered by Bertsimas and Lo [6] and Almgren

---

✉ U. Horst  
[horst@math.hu-berlin.de](mailto:horst@math.hu-berlin.de)

E. Kivman  
[kivmanev@hu-berlin.de](mailto:kivmanev@hu-berlin.de)

<sup>1</sup> Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

and Chriss [3], divides the price impact into a purely temporary effect, which depends only on the present trading rate and does not influence future prices, and a permanent effect, which influences the price depending only on the total volume that has been traded in the past. The temporary impact is typically assumed to be linear in the trading rate, leading to a quadratic term in the cost functional. The original modelling framework has been extended in various directions, including general stochastic settings with and without model uncertainty and multi-player and mean-field-type models by many authors, including Ankirchner et al. [4], Cartea et al. [9], Fu et al. [14], Gatheral and Schied [17], Graewe et al. [19], Horst et al. [23], Kruse and Popier [24] and Neuman and Voß [29].

A second approach, initiated by Obizhaeva and Wang [30], assumes that price impact is not permanent but transient, with the impact of past trades on current prices decaying over time. When impact is transient, one often allows both absolutely continuous and singular trading strategies. When singular controls are admissible, optimal liquidation strategies usually comprise large block trades at the initial and terminal time. The work of Obizhaeva and Wang has been extended by Alfonsi et al. [2], Chen et al. [11], Fruth et al. [13], Gatheral [16], Guéant [20], Horst and Naujokat [21], Lokka and Xu [27] and Predoiu et al. [31], among many others.

Single- and multi-asset liquidation problems with instantaneous and transient market impact and stochastic resilience where trading is confined to absolutely continuous strategies have been analysed in Graewe and Horst [18] and Horst and Xia [22], respectively. This is consistent with the empirical work of Large [25] and Lo and Hall [26], which suggests that this resilience does indeed vary stochastically. Although only absolutely continuous trading strategies were admissible in [18, 22], numerical simulations reported in [18] suggest that if all model parameters are deterministic constants, then the optimal portfolio process converges to the solution in [30] with two block trades and a constant trading rate as the instantaneous impact parameter converges to zero. Cartea and Jaimungal [10] provide empirical evidence that the instantaneous price impact is indeed (much) smaller than permanent (or transient) price impact. The numerical simulations in [18] suggest that the model in [18] provides a common framework within which to embed the two most commonly used liquidation models [3, 30] as limiting cases.

This paper provides a rigorous convergence analysis within a Markovian factor model. It turns out that the stochastic setting is quite different from the deterministic one. Most importantly, we show that in the stochastic setting, the optimal portfolio processes obtained in [18] converge in the Skorohod  $\mathcal{M}_2$ -topology to a process of *infinite variation* with jumps as the instantaneous market impact parameter converges to zero. Our second main result is to prove that the limiting portfolio process is optimal in a liquidation model with semimartingale execution strategies and to explicitly compute the optimal trading cost in the semimartingale execution framework.

Showing that the limiting model solves a liquidation model with semimartingale execution strategies is more than a mere byproduct. Control problems with semimartingale strategies are usually difficult to solve because there are no canonical candidates for the value function and/or optimal strategies. We show that the solution in the limiting model is fully determined by the unique bounded solution to a one-dimensional quadratic BSDE. Our limit result provides a novel approach to solving

control problems with semimartingale strategies that complements the approaches in Ackermann et al. [1] and Lorenz and Schied [15]. They solved related models by passing to a continuous-time limit from a sequence of discrete-time models.

Within a portfolio liquidation framework, inventory processes with infinite variation were first considered by Lorenz and Schied [28] to the best of our knowledge. Later, Becherer et al. [5] considered a trading framework with generalised price impact and proved that the cost functional depends continuously on the trading strategy, considered in several topologies. Bouchard et al. [7] considered infinite-variation inventory processes in the context of hedging.

The paper closest to ours is the recent work by Ackermann et al. [1]. They considered a liquidation model with general RCLL semimartingale trading strategies. Their framework is more general than ours as they allow more general filtrations and stochastic order book depth. At the same time, their analysis is confined to risk-neutral traders. In our setting, when the model parameters are deterministic and the instantaneous price impact goes to zero, the case of risk-neutral traders – which is then a special case of the model studied in [1] – is explicitly solvable. Allowing for risk aversion renders the impact model significantly more complicated as it adds a quadratic dependence of the integrated trading rate into the HJB equation; cf. [18] for details.

Our work also complements the work of Gârleanu and Pedersen [15]. They consider an array of market impact models, including a model with purely transient costs. They write [15, beginning of Sect. 1.3] that “with purely persistent price-impact costs, the optimal portfolio policy can have jumps and infinite quadratic variation.” As in [1], they justify portfolio holdings with infinite quadratic variation by taking a limit of a sequence of discrete-time models with increasing trading frequency. They also prove that the optimal portfolio processes in the discrete-time models converge to the optimal portfolio process in the corresponding continuous-time model if either the instantaneous price impact converges to a positive constant or the instantaneous price impact factor multiplied by the (increasing) trading frequency converges to zero. However, they do not consider the general case of an instantaneous price impact factor converging to zero. Most importantly, they consider a portfolio choice problem on an infinite time horizon, which avoids the liquidation constraint at the terminal time.

Last but not least, our work complements the work of Carmona and Webster [8], who provide strong evidence that inventories of large traders often do have indeed a nontrivial quadratic-variation component. For instance, for the Blackberry stock, they analyse the inventories of “the three most active Blackberry traders” on a particular day, namely CIBC World Markets Inc., Royal Bank of Canada Capital Markets and TD Securities Inc. From their data, they “suspect that RBC (resp. TD Securities) were trading to acquire a long (resp. short) position in Blackberry” and found that the corresponding inventory processes were of infinite variation. More generally, they find that systematic tests “on different days and other actively traded stocks give systematic rejections of this null hypothesis [quadratic variation of inventory being zero], with a  $p$ -value never greater than  $10^{-5}$ ”. Our results suggest that inventories with nontrivial quadratic variation arise naturally when market depth is high and resilience and/or market risk fluctuates stochastically. This is very intuitive; in deep markets, it is comparably cheap to frequently adjust portfolios to stochastically varying market environments.

The main technical challenge in establishing our convergence results is that the solution to the limiting model cannot be obtained by taking the limit of the three-dimensional quadratic BSDE system that characterises the solution in the model with positive instantaneous impact. Instead, we prove that the limit is fully characterised by the solution to a one-dimensional quadratic BSDE. Remarkably, this BSDE is independent of the liquidation requirement. As a result, full liquidation takes place if the instantaneous impact parameter converges to zero even if full liquidation is not strictly required. The reason is a loss in book value of the remaining shares that outweighs the liquidation cost for small instantaneous impact. Our convergence result is based on a novel representation result for solutions to BSDEs driven by Itô processes in terms of *uniformly* continuous functions of the forward process and on a series of novel convergence results for sequences of singular stochastic integral equations and random ODEs, which we choose to report in an abstract setting in Appendix A and B, respectively.

The limiting portfolio process is optimal in a liquidation model with general semimartingale execution strategies. Within our modelling framework where the cost coefficients are driven by continuous factor processes, block trades optimally occur only at the beginning and the end of the trading period. This is very intuitive as one would expect large block trades to require some form of trigger such as an external shock leading to a discontinuous change of cost coefficients. The proof of optimality proceeds in three steps. We first prove that the jump part of the limiting portfolio process can be approximated by absolutely continuous processes. This allows us to approximate the trading costs in the semimartingale model by trading costs in the pre-limit models from which we finally deduce the optimality of the limiting process in the semimartingale model by using the optimality of the approximating inventory processes in the pre-limit models. As a byproduct of our approximation, we also obtain that the optimal costs are given in terms of the aforementioned one-dimensional quadratic BSDE.

The rest of the paper is organised as follows. In Sect. 2, we recall the modelling setup from [18, 22] and summarise our main results. The proofs are given in Sects. 3 and 4. A series of fairly abstract convergence results for various stochastic equations with singularities upon which our convergence results are based is postponed to two appendices.

**Notation.** Throughout, randomness is described by an  $\mathbb{R}^m$ -valued Brownian motion  $(W_t)_{t \in [0, T]}$  defined on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ , a complete probability space, where  $(\mathcal{F}_t)_{t \in [0, T]}$  denotes the filtration generated by  $W$ , augmented by the  $\mathbb{P}$ -nullsets. Unless otherwise specified, all equations and inequalities hold in the  $\mathbb{P}$ -a.s. sense. For a subset  $A \subseteq \mathbb{R}^d$ , we denote by  $\mathcal{L}_{\text{Prog}}^2(\Omega \times [0, T]; A)$  the set of all progressively measurable  $A$ -valued stochastic processes  $(X_t)_{t \in [0, T]}$  such that  $\mathbb{E}[\int_0^T |X_t|^2 dt] < \infty$ , while  $\mathcal{L}_{\text{Prog}}^\infty(\Omega \times [0, T]; A)$  denotes the subset of essentially bounded processes. By  $\mathcal{L}_{\mathcal{P}}^2(\Omega \times [0, T]; A)$  and  $\mathcal{L}_{\mathcal{P}}^\infty(\Omega \times [0, T]; A)$ , we denote the respective subsets of predictable processes. Whenever  $T -$  appears, we mean that there exists an  $\varepsilon > 0$  such that a statement holds for all  $T' \in (T - \varepsilon, T)$ .

## 2 Problem formulation and main results

In this section, we introduce two portfolio liquidation models with stochastic market impact. In the first model, analysed in Sect. 2.1, the investor is confined to absolutely continuous trading strategies. For small instantaneous market impact, we prove that the optimal liquidation strategy converges to a semimartingale with jumps. In Sect. 2.2, we therefore analyse a liquidation model with semimartingale trading strategies. We prove that the limiting process obtained in Sect. 2.1 is optimal in a model where semimartingale strategies that satisfy a suitable regularity condition are admissible.

### 2.1 Portfolio liquidation with absolutely continuous strategies

We take the liquidation model analysed in Graewe and Horst [18] and Horst and Xia [22] as our starting point and consider an investor that needs to close, within a given time interval  $[0, T]$ , a (single-asset) portfolio of  $x_0 > 0$  shares using a trading strategy  $\xi = (\xi_t)_{t \in [0, T]}$ . If  $\xi_t < 0$ , the investor is selling the asset at a rate  $\xi_t$  at time  $t \in [0, T]$ ; otherwise she is buying it. For a given strategy  $\xi$ , the corresponding portfolio process  $X^\xi = (X_t^\xi)_{t \in [0, T]}$  satisfies the ODE

$$dX_t^\xi = \xi_t dt, \quad X_0^\xi = x_0.$$

The set of *admissible strategies* is given by

$$\mathcal{A} := \mathcal{L}_{\text{Prog}}^2(\Omega \times [0, T]; \mathbb{R}).$$

For a general inventory process  $X \in \mathcal{L}_{\mathcal{P}}^2(\Omega \times [0, T]; \mathbb{R})$ , the corresponding *transient price impact* is described by the unique stochastic process  $Y^X = (Y_t^X)_{t \in [0, T]}$  that satisfies the ODE

$$dY_t^X = \gamma dX_t - \rho_t Y_t^X dt, \quad Y_0^X = 0, \quad (2.1)$$

for some constant  $\gamma > 0$  and some essentially bounded, progressively measurable,  $(0, \infty)$ -valued process  $\rho = (\rho_t)_{t \in [0, T]}$ . The process  $Y^X$  may be viewed as describing an additional shift in the unaffected benchmark price process generated by the large investor's trading activity. For  $\xi \in \mathcal{A}$ , we write  $Y^\xi := Y^{X^\xi}$ .

For any *instantaneous impact factor*  $\eta > 0$  and any *penalisation factor*

$$N \in \mathcal{N} = \mathcal{N}(\eta) := [\underline{N}(\eta), \infty],$$

where

$$\underline{N}(\eta) := \gamma + 1 + \sqrt{2\eta \max(\|\lambda\|_{L^\infty}, \gamma \|\rho\|_{L^\infty})},$$

the *cost functional* is given by

$$J^{\eta,N}(\xi) = \mathbb{E} \left[ \frac{\eta}{2} \int_0^T (\xi_t)^2 dt + \int_0^T Y_t^\xi dX_t^\xi + \frac{1}{2} \int_0^T \lambda_t (X_t^\xi)^2 dt + \frac{N}{2} (X_T^\xi)^2 - X_T^\xi Y_T^\xi \right].$$

The first term in  $J^{\eta,N}(\xi)$  captures the *instantaneous trading costs*, the second the costs from a *transient price impact*, and the third captures *market risk*, where the adapted and nonnegative process  $\lambda = (\lambda_t)_{t \in [0,T]}$  specifies the degree of risk. If  $N = \infty$ , the fourth term should formally be read as  $+\infty \mathbb{1}_{\{X_T^\xi \neq 0\}}$  with the convention  $0 \cdot \infty = 0$ . This captures the case where full liquidation is required; this case is analysed in [18]. The case  $\gamma + 1 \leq N < \infty$  is analysed in [22]. The fifth term captures an additional loss in the book value of the remaining shares. It drops out of the cost function if  $N = \infty$ ; see [18, 22] for further details on the impact costs and cost coefficients.

It has been shown in [18, 22] that the optimisation problem

$$\min_{\xi \in \mathcal{A}} J^{\eta,N}(\xi) \quad (2.2)$$

has a solution  $\hat{\xi}^{\eta,N}$  for any  $N \in \mathcal{N}$  and any  $\eta > 0$ . The solution is given in terms of a backward SDE system with possibly singular terminal condition. We index the optimal trading strategies and state processes by  $\eta$  and  $N$  as we are interested in their behaviour for small instantaneous impact factors for both finite and infinite  $N$ .

The next theorem follows directly from [18, Theorems 2.5 and 2.6] and [22, Theorem 2.1].

**Theorem 2.1** *With the above assumptions, for all  $\eta \in (0, \infty)$  and  $N \in \mathcal{N}$ , the following hold:*

i) *The BSDE system*

$$\begin{aligned} -dA_t^{\eta,N} &= \left( \lambda_t - \frac{1}{\eta} (A_t^{\eta,N} - \gamma B_t^{\eta,N})^2 \right) dt - Z_t^{\eta,N,A} dW_t, \\ -dB_t^{\eta,N} &= \left( -\rho_t B_t^{\eta,N} + \frac{1}{\eta} (\gamma C_t^{\eta,N} - B_t^{\eta,N} + 1)(A_t^{\eta,N} - \gamma B_t^{\eta,N}) \right) dt \\ &\quad - Z_t^{\eta,N,B} dW_t, \\ -dC_t^{\eta,N} &= \left( -2\rho_t C_t^{\eta,N} - \frac{1}{\eta} (\gamma C_t^{\eta,N} - B_t^{\eta,N} + 1)^2 \right) dt - Z_t^{\eta,N,C} dW_t \end{aligned}$$

*with terminal condition*

$$B_T^{\eta,N} = 1, \quad C_T^{\eta,N} = 0,$$

together with  $A_T^{\eta,N} = N$  for  $N < \infty$  and  $\lim_{t \nearrow T} A_t^{\eta,\infty} = \infty$  in  $L^\infty$  if  $N = \infty$ , has a solution

$$((A^{\eta,N}, B^{\eta,N}, C^{\eta,N}), (Z^{\eta,N,A}, Z^{\eta,N,B}, Z^{\eta,N,C}))$$

that belongs to the space  $\mathcal{L}_{\mathcal{P}}^\infty(\Omega \times [0, T]; \mathbb{R}^3) \times \mathcal{L}_{\text{Prog}}^2(\Omega \times [0, T]; \mathbb{R}^{3 \times m})$  if  $N < \infty$  and to the space  $\mathcal{L}_{\mathcal{P}}^\infty(\Omega \times [0, T-]; \mathbb{R}^3) \times \mathcal{L}_{\text{Prog}}^2(\Omega \times [0, T-]; \mathbb{R}^{3 \times m})$  if  $N = \infty$ .

ii) The liquidation problem (2.2) has a solution  $\hat{\xi}^{\eta,N}$ . The corresponding state process

$$(\hat{X}^{\eta,N}, \hat{Y}^{\eta,N}) := (X^{\hat{\xi}^{\eta,N}}, Y^{\hat{\xi}^{\eta,N}})$$

is given by the (unique) solution to the ODE system

$$\begin{aligned} \partial_t \hat{X}_t^{\eta,N} &= -\frac{1}{\sqrt{\eta}} D_t^{\eta,N} \hat{X}_t^{\eta,N} - \frac{1}{\sqrt{\eta}} E_t^{\eta,N} \hat{Y}_t^{\eta,N}, \\ \partial_t \hat{Y}_t^{\eta,N} &= -\frac{\gamma}{\sqrt{\eta}} D_t^{\eta,N} \hat{X}_t^{\eta,N} - \frac{\gamma}{\sqrt{\eta}} E_t^{\eta,N} \hat{Y}_t^{\eta,N} - \rho_t \hat{Y}_t^{\eta,N} \end{aligned} \quad (2.3)$$

with initial conditions  $\hat{X}_0^{\eta,N} = x_0$  and  $\hat{Y}_0^{\eta,N} = 0$ , where

$$D^{\eta,N} := \frac{1}{\sqrt{\eta}} (A^{\eta,N} - \gamma B^{\eta,N}), \quad E^{\eta,N} := \frac{1}{\sqrt{\eta}} (\gamma C^{\eta,N} - B^{\eta,N} + 1).$$

Let us now define the process

$$\hat{Z}^{\eta,N} := \gamma \hat{X}^{\eta,N} - \hat{Y}^{\eta,N}.$$

The benefit of defining this process is that in the ODE (2.3), the terms that are multiplied by  $\sqrt{\eta}^{-1}$  drop out so that we expect that the process  $\hat{Z}^{\eta,N}$  remains stable for small values of  $\eta$ . Next, we state a result on the optimal state process and the previously introduced process  $\hat{Z}^{\eta,N}$  that will be important for our subsequent analysis. In particular, we show that the optimal portfolio process  $\hat{X}^{\eta,N}$  never changes its sign. The proof is given in Sect. 3.1.1.

**Theorem 2.2** For all  $\eta \in (0, \infty)$ ,  $N \in \mathcal{N}$ , the process  $\hat{Z}^{\eta,N}$  is nonincreasing on  $[0, T]$ . Moreover,

$$\mathbb{P}[\hat{X}_t^{\eta,N} \in (0, x_0), \hat{Y}_t^{\eta,N} \in (-\gamma x_0, 0), \hat{Z}_t^{\eta,N} \in (0, \gamma x_0) \text{ for all } t \in (0, T)] = 1.$$

We are interested in the dynamics of the optimal portfolio processes for small instantaneous price impact. We address this problem within a factor model where the cost coefficients  $\lambda$  and  $\rho$  are driven by an Itô diffusion, which is given by the unique strong solution to the SDE

$$d\chi_t = \mu(t, \chi_t) dt + \sigma(t, \chi_t) dW_t, \quad \chi_0 = \bar{\chi}_0$$



on  $[0, T]$  with  $\bar{\chi}_0 \in \mathbb{R}^n$ . We assume **throughout** that the function

$$(\mu, \sigma): [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^{n \times m}$$

is bounded, measurable and uniformly Lipschitz-continuous in the space variable, i.e.,

$$|(\mu, \sigma)(t, x_1) - (\mu, \sigma)(t, x_2)| \leq c|x_1 - x_2|.$$

**Assumption 2.3** The processes  $\rho$  and  $\lambda$  are of the form

$$(\rho_t, \lambda_t) = f(t, \chi_t)$$

for some bounded  $C^{1,2}$  function  $f = (f^\rho, f^\lambda): [0, T] \times \mathbb{R}^n \rightarrow [0, \infty)^2$  with bounded derivatives. Moreover, the function  $f^\rho$  is bounded away from zero.

For convenience, we define the stochastic process  $\varphi = (\varphi_t)_{t \in [0, T]}$  by

$$\varphi_t := \sqrt{\lambda_t + 2\gamma\rho_t}$$

and choose constants  $\underline{\rho}, \bar{\rho}, \underline{\varphi}, \bar{\varphi}, \bar{\lambda} \in (0, \infty)$  such that

$$\underline{\rho} \leq \rho \leq \bar{\rho}, \quad 0 \leq \lambda \leq \bar{\lambda}, \quad \underline{\varphi} \leq \varphi \leq \bar{\varphi}.$$

In what follows, we heuristically argue that the processes  $\hat{X}^{\eta, N}$  converge to a limit process  $\hat{X}^0$  (independent of  $N$ ) as  $\eta \rightarrow 0$  and identify the limit  $\hat{X}^0$ . Since the ODE system (2.3) is not defined for  $\eta = 0$ , we cannot define the limiting process as the solution to this system. Instead, we first identify the limits of the coefficients of the ODE system and then derive candidate limits for the state processes in terms of the limiting coefficients.

### 2.1.1 Convergence of the coefficient processes

In this section, we state the convergence results for the coefficient processes  $D^{\eta, N}$  and  $E^{\eta, N}$  of the ODE system (2.3) as  $\eta \rightarrow 0$ . In particular, we prove that their limits  $D^0$  and  $E^0$  exist and are driven by a common factor, which is given by the solution of a quadratic BSDE.

Before proceeding to the limit result, we provide some heuristics for the convergence. Assuming for simplicity that all coefficients are deterministic, the dynamics of the coefficient processes satisfy

$$\begin{aligned} \sqrt{\eta} \dot{D}_t^{\eta, N} &= (D_t^{\eta, N})^2 - \lambda_t - \gamma \dot{B}_t^{\eta, N}, \\ \sqrt{\eta} \dot{E}_t^{\eta, N} &= 2\sqrt{\eta}\rho_t E_t^{\eta, N} + \gamma(E_t^{\eta, N})^2 - 2\rho_t(1 - B_t^{\eta, N}) - \dot{B}_t^{\eta, N}. \end{aligned}$$

Letting  $\eta \rightarrow 0$ , we expect that

$$\begin{aligned} 0 &= (D_t^0)^2 - \lambda_t - \gamma \dot{B}_t^0, \\ 0 &= \gamma(E_t^0)^2 - 2\rho_t(1 - B_t^0) - \dot{B}_t^0, \end{aligned}$$

that is, we expect that

$$D_t^0 = \sqrt{\gamma \dot{B}_t^0 + \lambda_t}, \quad E_t^0 = \sqrt{\gamma^{-1} (\dot{B}_t^0 + 2\rho_t(1 - B_t^0))}. \quad (2.4)$$

Moreover, by the choice of the coefficients  $D^{\eta,N}$  and  $E^{\eta,N}$  and by the BSDE in Theorem 2.1 i), we expect that in the limit,

$$-\dot{B}_t^0 = -\rho_t B_t^0 + D_t^0 E_t^0. \quad (2.5)$$

The three equalities combined yield

$$\begin{aligned} (D_t^0 + \gamma E_t^0)^2 &= \gamma \dot{B}_t^0 + \lambda_t + 2\gamma D_t^0 E_t^0 + \gamma (\dot{B}_t^0 + 2\rho_t(1 - B_t^0)) \\ &= (\varphi_t)^2 + 2\gamma (\dot{B}_t^0 + D_t^0 E_t^0 - \rho_t B_t^0) \\ &= (\varphi_t)^2. \end{aligned} \quad (2.6)$$

Plugging (2.5) and (2.6) back into (2.4) yields

$$D_t^0 = (\varphi_t)^{-1} (\lambda_t + \gamma \rho_t B_t^0), \quad E_t^0 = (\varphi_t)^{-1} \rho_t (2 - B_t^0).$$

Hence we expect  $D^0$  and  $E^0$  to be driven by  $B^0$  and  $B^0$  to satisfy the ODE

$$\begin{aligned} -\dot{B}_t^0 &= -\rho_t B_t^0 + D_t^0 E_t^0 \\ &= -\frac{1}{\varphi_t^2} (\gamma (\rho_t B_t^0)^2 - 2\lambda_t \rho_t (1 - B_t^0)) \end{aligned}$$

with terminal condition  $B_T^0 = 1$  (because  $B_T^{\eta,N} = 1$ ). Our heuristic also suggests that the limit processes are independent of the liquidation requirement.

**Example 2.4** If  $\lambda = C\rho$  for some constant  $C \geq 0$ , then the process  $B^0$  can be computed explicitly. If  $C = 0$ , then  $B_t^0 = 2/(2 + \int_t^T \rho_s ds)$ . If  $C > 0$ , then

$$B_t^0 = -\frac{C}{\gamma} + \frac{1}{\gamma} \sqrt{C(C + 2\gamma)} \coth \left( \operatorname{arccoth} \left( \frac{C + \gamma}{\sqrt{C(C + 2\gamma)}} \right) + \sqrt{\frac{C}{C + 2\gamma}} \int_t^T \rho_s ds \right).$$

The preceding heuristic suggests that the limiting coefficient processes are driven by a solution to the BSDE corresponding to the above ODE for  $B^0$ . The following lemma is proved in Sect. 3.1.2.

**Lemma 2.5** *There exists a unique solution  $(B^0, Z^{0,B})$  in the space*

$$\mathcal{L}_{\mathcal{P}}^\infty(\Omega \times [0, T]; (0, \infty)) \times \mathcal{L}_{\mathcal{P}}^2(\Omega \times [0, T]; \mathbb{R}^m)$$

*to the BSDE*

$$-dB_t^0 = -\frac{1}{\varphi_t^2} (\gamma (\rho_t B_t^0)^2 - 2\lambda_t \rho_t (1 - B_t^0)) dt - Z_t^{0,B} dW_t, \quad B_T^0 = 1. \quad (2.7)$$

The process  $B^0$  is bounded from above by 1 and from below by

$$\underline{B}_t^0 := \exp\left(-\underline{\varphi}^{-2}\gamma\bar{\rho}^2(T-t)\right). \quad (2.8)$$

Moreover, there exists a uniformly continuous function (a so-called “decoupling field”)  $h: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $B^0$  and  $(h(t, \chi_t))_{t \in [0, T]}$  are indistinguishable.

We prove below that the process  $B^{\eta, N}$  converges to  $B^0$  as  $\eta \rightarrow 0$  and that  $D^{\eta, N}$  and  $E^{\eta, N}$  converge to the processes

$$D_t^0 := (\varphi_t)^{-1}(\lambda_t + \gamma\rho_t B_t^0), \quad E_t^0 := (\varphi_t)^{-1}\rho_t(2 - B_t^0), \quad (2.9)$$

respectively. In view of Lemma 2.5, these processes are well defined, and so the dynamics (2.7) of the process  $B^0$  can be rewritten as

$$-dB_t^0 = -(\rho_t B_t^0 - D_t^0 E_t^0) dt - Z_t^{0, B} dW_t, \quad B_T^0 = 1.$$

Likewise, the BSDE for the process  $B^{\eta, N}$  in Theorem 2.1 can be rewritten as

$$-dB_t^{\eta, N} = -(\rho_t B_t^{\eta, N} - D_t^{\eta, N} E_t^{\eta, N}) dt - Z_t^{\eta, N, B} dW_t, \quad B_T^{\eta, N} = 1. \quad (2.10)$$

This suggests that the process  $B^{\eta, N}$  converges to  $B^0$  on the entire interval  $[0, T]$ , and this is proved below; note that all the processes  $B$  have the same terminal condition at time  $T$ . By contrast, convergence of the processes  $D^{\eta, N}$  and  $E^{\eta, N}$  can only be expected to hold on compact subintervals  $[0, T - \varepsilon]$  of  $[0, T)$  because the terminal conditions of the limiting and approximating processes are different. It turns out that we can prove a slightly stronger result for the supremum of the processes  $D$  and  $E$ , and this explains the different choices of intervals for the supremum and the infimum in the following lemma. Specifically, we have the following result. Its proof is given in Sect. 3.2.1.

**Proposition 2.6** *Let*

$$b^{\eta, N} := B^{\eta, N} - B^0, \quad d^{\eta, N} := D^{\eta, N} - D^0, \quad e^{\eta, N} := E^{\eta, N} - E^0.$$

*Then for all  $\varepsilon > 0$ , there exists an  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0]$  and for all  $N \in \mathcal{N}$ ,*

$$\mathbb{P}\left[\sup_{t \in [0, T)} \max(|b_t^{\eta, N}|, -d_t^{\eta, N}, e_t^{\eta, N}, -d_t^{\eta, N} - \gamma e_t^{\eta, N}) \leq \varepsilon, \right. \\ \left. \inf_{t \in [0, T - \varepsilon]} \min(-d_t^{\eta, N}, e_t^{\eta, N}, -d_t^{\eta, N} - \gamma e_t^{\eta, N}) \geq -\varepsilon\right] = 1.$$

## 2.1.2 Convergence of the state process

Having derived the limits of the coefficient processes, we can now heuristically derive the limits of the processes  $\hat{X}^{\eta, N}$ ,  $\hat{Y}^{\eta, N}$  and  $\hat{Z}^{\eta, N}$ , which we denote by  $\hat{X}^0$ ,  $\hat{Y}^0$  and  $\hat{Z}^0$ , respectively.

Since  $\hat{X}_T^{\eta, \infty} = 0$  for all  $\eta > 0$ , we expect that  $\hat{X}_T^0 = 0$ . We prove in Lemma 3.11 that this convergence also holds if  $N$  is finite. The proof heavily relies on the optimality of  $\hat{X}^0$  in the semimartingale portfolio liquidation model.

Assuming that the optimal trading strategy remains stable if  $\eta \rightarrow 0$ , the ODE (2.3) suggests that the term  $D^{\eta, N} \hat{X}^{\eta, N} + E^{\eta, N} \hat{Y}^{\eta, N}$  is small for small  $\eta$  and hence that

$$D_t^0 \hat{X}_t^0 = -E_t^0 \hat{Y}_t^0 \quad \text{on } [0, T).$$

We do not conjecture the above relation at the terminal time because the convergence of  $D^{\eta, N}$  and  $E^{\eta, N}$  only holds on  $[0, T)$ . Assuming that

$$\hat{Z}^0 = \gamma \hat{X}^0 - \hat{Y}^0 \quad \text{on } [0, T),$$

equation (2.6) implies

$$\hat{Z}^0 = -\frac{\varphi}{D^0} \hat{Y}^0 \quad \text{and hence} \quad \hat{Y}^0 = -\frac{D^0}{\varphi} \hat{Z}^0.$$

On the other hand, by definition,  $\partial_t \hat{Z}^{\eta, N} = \rho \hat{Y}^{\eta, N}$ . As  $\eta \rightarrow 0$ , this suggests that  $\partial_t \hat{Z}^0 = \rho \hat{Y}^0$  and hence that

$$\partial_t \hat{Z}_t^0 = -\frac{\rho_t D_t^0}{\varphi_t} \hat{Z}_t^0, \quad \hat{Z}_0^0 = \gamma x_0.$$

This motivates us to *define the process*

$$\hat{Z}_t^0 := \gamma x_0 \exp \left( - \int_0^t \frac{\rho_s}{\varphi_s} D_s^0 ds \right), \quad t \in [0, T].$$

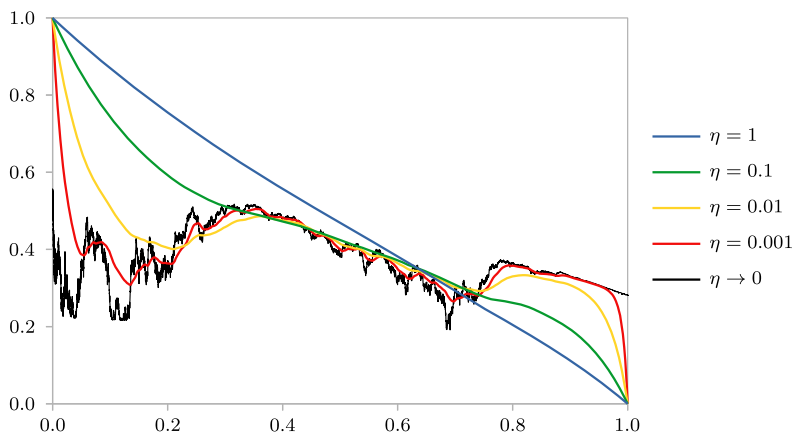
Since we expect that  $\hat{X}_T^0 = 0$  and that  $\hat{Z}^0 = \gamma \hat{X}^0 - \hat{Y}^0$ , we now introduce the *candidate limiting state processes*

$$\begin{aligned} \hat{X}_t^0 &:= \mathbb{1}_{[0, T)}(t) \frac{E_t^0}{\varphi_t} \hat{Z}_t^0, \\ \hat{Y}_t^0 &:= -\mathbb{1}_{[0, T)}(t) \frac{D_t^0}{\varphi_t} \hat{Z}_t^0 - \mathbb{1}_{\{T\}}(t) \hat{Z}_T^0, \quad t \in [0, T]. \end{aligned} \quad (2.11)$$

Since  $(\hat{X}_{0-}^0, \hat{Y}_{0-}^0) = (x_0, 0)$ , we expect that the limiting state process jumps at the initial and the terminal time. In particular, we cannot expect uniform convergence on  $[0, T]$ .

We also expect the limiting state processes to be of unbounded variation, as suggested by Fig. 1. The figure also suggests that the portfolio process is more or less monotone for large  $\eta$ , while this property is lost for small  $\eta$ . When  $\eta \rightarrow 0$ , adjustments to small changes in market environments are cheap. This is very different from round-trip strategies where an investor uses the price impact caused by him/herself to drive market prices into a favourable direction.

Figure 1 also suggests that the limiting portfolio process jumps only at the times 0 and  $T$ . This is consistent with the definition of the candidate processes (2.11) as



**Fig. 1** Optimal trading strategies  $\hat{X}^{\eta, \infty}$  for the liquidation model for different instantaneous impact factors, and their limit  $\hat{X}^0$  for  $m = n = T = x_0 = 1$ ,  $\gamma = 3$ ,  $\rho_t = 1 + 0.9 \sin(2.5W_t)$ ,  $\lambda \equiv 1$

well as the observation in Horst and Naujokat [21], according to which jumps in the optimal strategy can only be triggered by exogenous shocks like jumps in the cost coefficients, which are absent in the present model.

It remains to clarify in which sense the state processes converge. In contrast to the convergence result in Proposition 2.6, we can only expect convergence in probability because the state process follows a forward ODE while the coefficient processes follow backward SDEs; see also Appendix B.2. The following result establishes uniform convergence in probability on compact subintervals of  $(0, T)$  along with some “upper/lower convergence” at the initial and terminal time. The proof is given in Sect. 3.2.2.

**Theorem 2.7** *For all  $\varepsilon > 0$  and  $\delta > 0$ , there exists an  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0]$  and all  $N \in \mathcal{N}$ ,*

$$\mathbb{P} \left[ \sup_{t \in [\varepsilon, T)} \max(\hat{X}_t^{\eta, N} - \hat{X}_t^0, \hat{Y}_t^{\eta, N} - \hat{Y}_t^0) \leq \varepsilon, \right. \\ \left. \inf_{t \in [0, T - \varepsilon]} \min(\hat{X}_t^{\eta, N} - \hat{X}_t^0, \hat{Y}_t^{\eta, N} - \hat{Y}_t^0) \geq -\varepsilon \right] \geq 1 - \delta.$$

The preceding theorem does not provide a convergence result on the whole time interval, due to the jumps of the limit processes at the initial and terminal time. However, along with our results from Sect. 2.2, it allows us to prove the convergence of the graphs of the state processes on the entire time interval. The completed graph of an RCLL function  $X: \{0-\} \cup [0, T] \rightarrow \mathbb{R}$  with finitely many jumps is defined by

$$G_X := \{(t, x) \in [0, T] \times \mathbb{R} : x = X_t = X_{t-} \text{ or } X_{t-} < X_t, x \in [X_{t-}, X_t] \\ \text{or } X_{t-} > X_t, x \in [X_t, X_{t-}]\}.$$

The Skorohod  $\mathcal{M}_2$ -distance between  $X$  and  $Y$  is defined as the Hausdorff distance between their completed graphs, i.e.,

$$d_{\mathcal{M}_2}(X, Y) := \max \left( \sup_{p \in G_X} \min_{q \in G_Y} |p - q|_\infty, \sup_{q \in G_Y} \min_{p \in G_X} |p - q|_\infty \right) \in [0, \infty],$$

where

$$|(s, y)|_\infty := \max(|s|, |y|).$$

If strict liquidation is required, then Theorem 2.7 is sufficient to prove convergence of the state processes in the Skorohod  $\mathcal{M}_2$  sense. Even if liquidation is not required, it turns out that the terminal position converges to zero as  $\eta \rightarrow 0$ . Heuristically, this can be seen as follows.

Let  $t_0 \in (0, T)$ . Disregarding market risk costs, which we expect to be of order  $\mathcal{O}(T - t_0)$  and hence negligible if  $t_0 \rightarrow T$ , and disregarding instantaneous impact costs for the moment, the cost functional for any given admissible strategy  $\xi$  is given by

$$\begin{aligned} & \int_{t_0}^T Y_s^\xi \xi_s \, ds - X_T^\xi Y_T^\xi + \frac{N}{2} (X_T^\xi)^2 \\ &= Y_T^\xi X_T^\xi - Y_{t_0}^\xi X_{t_0}^\xi - \int_{t_0}^T \dot{Y}_s^\xi X_s^\xi \, ds - X_T^\xi Y_T^\xi + \frac{N}{2} (X_T^\xi)^2 \\ &= -Y_{t_0}^\xi X_{t_0}^\xi - \int_{t_0}^T \gamma \xi_s X_s^\xi \, ds + \int_{t_0}^T \rho Y_s^\xi X_s^\xi \, ds + \frac{N}{2} (X_T^\xi)^2 \\ &= -Y_{t_0}^\xi X_{t_0}^\xi - \frac{\gamma}{2} ((X_T^\xi)^2 - (X_{t_0}^\xi)^2) + \int_{t_0}^T \rho Y_s^\xi X_s^\xi \, ds + \frac{N}{2} (X_T^\xi)^2 \\ &= -Y_{t_0}^\xi X_{t_0}^\xi + \frac{\gamma}{2} (X_{t_0}^\xi)^2 + \frac{N - \gamma}{2} (X_T^\xi)^2 + \int_{t_0}^T \rho Y_s^\xi X_s^\xi \, ds. \end{aligned}$$

Hence we expect the *controllable* costs to satisfy

$$\mathbb{E} \left[ \frac{N - \gamma}{2} (X_T^\xi)^2 + \int_{t_0}^T \rho Y_s^\xi X_s^\xi \, ds \right] = \mathbb{E} \left[ \frac{N - \gamma}{2} (X_T^\xi)^2 \right] + \mathcal{O}(T - t_0)$$

plus instantaneous impact costs. Since  $N > \gamma$ , this suggests to make  $X_T^\xi$  small, which is cheap if  $\eta$  is small. More precisely, we have the following result; its proof is given in Sect. 3.2.3.

**Proposition 2.8** *For all  $\varepsilon > 0$  and  $\delta > 0$ , there exists an  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0]$  and all  $N \in \mathcal{N}$ ,*

$$\mathbb{P}[d_{\mathcal{M}_2}(\hat{X}^{\eta, N}, \hat{X}^0) \leq \varepsilon] \geq 1 - \delta.$$

## 2.2 Optimal liquidation with semimartingale strategies

In this section, we prove that the limit process  $\hat{X}^0$  is the optimal portfolio process in a trade execution model with semimartingale trading strategies.

In our semimartingale model, a trading strategy is given by a triple

$$\theta = (j^+, j^-, V),$$

where  $j^+$  and  $j^-$  are real-valued, nondecreasing pure jump processes and  $V$  is with respect to the Brownian filtration a real-valued continuous semimartingale starting in zero. The jump processes  $j^+$  and  $j^-$  describe the cumulative effects of buying, respectively selling, large blocks of shares, while the continuous semimartingale  $V$  describes the effect of continuously trading small amounts of the stock. The portfolio process  $X^\theta = (X_t^\theta)_{t \in [0, T]}$  associated with a strategy  $\theta$  is then given by

$$X_t^\theta = x_0 + j_t^+ - j_t^- + V_t.$$

The associated price impact process, again given by (2.1), is denoted by  $Y^\theta$ . We note that  $X^\theta$  and  $Y^\theta$  are semimartingales.

We assume that strict liquidation is required and that the cost associated with a trading strategy  $\theta$  is given by

$$J^0(\theta) := \mathbb{E} \left[ \int_{(0, T]} Y_{s-}^\theta dX_s^\theta + \int_0^T \frac{1}{2} \lambda_s (X_s^\theta)^2 ds + \frac{\gamma}{2} [X^\theta]_T \right].$$

The first term captures the transient price impact cost; the second captures market risk. The third term emerges as an additional cost term when passing from discrete to continuous time, as shown in Ackermann et al. [1]. Moreover, in the absence of this term, arbitrarily low costs can be achieved; see [1] for details.

The cost function can be conveniently rewritten as

$$\begin{aligned} J^0(\theta) = \mathbb{E} & \left[ \frac{\gamma}{2} (j_0^+ - j_0^-)^2 + \int_{(0, T]} \frac{1}{2} (Y_{s-}^\theta + Y_s^\theta) dX_s^\theta \right. \\ & \left. + \int_0^T \frac{1}{2} \lambda_s (X_s^\theta)^2 ds + \frac{\gamma}{2} [V]_T \right]. \end{aligned}$$

This representation supports our intuition that the price impact before and the price impact after the jump equally influence the total cost. The first term in this expression captures the cost of the initial block trade at time  $t = 0$ .

The cost functional is well defined under the following admissibility condition.

**Definition 2.9** A trading strategy  $\theta = (j^+, j^-, V)$  is called *admissible* if the liquidation constraint  $X_T^\theta = 0$  holds, if  $j^\pm$  are RCLL, predictable, real-valued, non-decreasing and square-integrable pure jump processes, and if  $V$  is a continuous semimartingale starting in zero with

$$\mathbb{E} \left[ \max_{t \in [0, T]} |V_t|^2 \right] < \infty. \quad (2.12)$$

The set of all admissible trading strategies is denoted by  $\mathcal{A}^0$ .

Our goal is now to solve the optimisation problem

$$\min_{\theta \in \mathcal{A}^0} J^0(\theta).$$

To this end, we verify directly that the limit process  $\hat{X}^0$  obtained in Sect. 2.1 is optimal. The results there show that the process has the representation

$$\hat{X}_t^0 = x_0 - \hat{j}_t^- + \hat{V}_t,$$

where the jump process  $\hat{j}^-$  and the continuous part are given by, respectively,

$$\begin{aligned}\hat{j}_t^- &:= x_0 - \hat{X}_0^0 + \mathbb{1}_{\{T\}}(t) \hat{X}_{t-}^0, \\ \hat{V}_t &:= (\varphi_t)^{-2} \rho_t (2 - B_t^0) \hat{Z}_t^0 - (\varphi_0)^{-2} \rho_0 (2 - B_0^0) \hat{Z}_0^0.\end{aligned}$$

In view of Assumption 2.3 and because  $B^0$  is a continuous semimartingale and  $\hat{Z}^0$  is differentiable, the process  $\hat{V}$  is a continuous semimartingale starting in zero. Hence the following holds.

**Lemma 2.10** *The strategy  $\hat{\theta} = (0, \hat{j}^-, \hat{V})$  is admissible.*

In order to prove that  $\hat{\theta}$  is optimal, we approximate the cost and the portfolio process associated with any strategy  $\theta \in \mathcal{A}^0$  by the cost and portfolio processes corresponding to absolutely continuous trading strategies. To this end, we first approximate the continuous semimartingale part  $V$  by differentiable processes. The proof of the following result is given in Sect. 4.1.

**Lemma 2.11** *For all  $\theta = (j^+, j^-, V) \in \mathcal{A}^0$  and all  $\beta, \delta > 0$ , there exist a constant  $v > 0$  and an adapted and continuous  $\dot{V}^{\beta, v}: \Omega \times [0, T] \rightarrow [-\frac{\beta}{v}, \frac{\beta}{v}]$  such that*

$$\max_{t \in [0, T]} \left| \int_0^t \dot{V}_s^{\beta, v} ds \right| \leq \max_{t \in [0, T]} |V_t| \quad (2.13)$$

and

$$\mathbb{P} \left[ \sup_{t \in [0, T]} \left| \int_0^t \dot{V}_s^{\beta, v} ds - V_t \right| \leq 3\beta \right] \geq 1 - \delta. \quad (2.14)$$

Next, we approximate the portfolio process  $X^\theta$  by a portfolio process associated with an absolutely continuous strategy. To this end, for all  $\theta \in \mathcal{A}^0$  and  $\beta, v, \varepsilon > 0$ , we define the integrable process

$$\xi_t^{\theta, \beta, v, \varepsilon} := \begin{cases} \frac{1}{\varepsilon} (j_t^+ - j_{t-\varepsilon}^+) - \frac{1}{\varepsilon} (j_t^- - j_{t-\varepsilon}^-) + \dot{V}_t^{\beta, v}, & t \leq T - \varepsilon, \\ -\frac{1}{\varepsilon} (x_0 + \int_0^{T-\varepsilon} \xi_u^{\theta, \beta, v, \varepsilon} du), & t > T - \varepsilon. \end{cases}$$

In view of the square-integrability of  $j^\pm$ , we see that  $\xi^{\theta, \beta, v, \varepsilon}$  belongs to  $\mathcal{A}$ . The corresponding portfolio process is denoted by  $X^{\theta, \beta, v, \varepsilon}$ . The proof of the following result is given in Sect. 4.2.



**Lemma 2.12** For all  $\theta = (j^+, j^-, V) \in \mathcal{A}^0$  and for all  $\delta > 0$ , there exist  $\beta > 0$ ,  $v > 0$  and  $\varepsilon > 0$  such that

$$\mathbb{E} \left[ \int_0^T (X_t^{\theta, \beta, v, \varepsilon} - X_t^\theta)^2 dt \right] \leq \delta.$$

For all  $\xi \in \mathcal{A}$  with  $\int_0^T \xi_s ds = -x_0$ , we can define  $V_t^\xi := \int_0^t \xi_s ds$  to obtain that  $(0, 0, V^\xi) \in \mathcal{A}^0$  and for all  $\eta > 0$ , we have

$$J^{\eta, \infty}(\xi) = J^0(0, 0, V^\xi) + \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\xi_t)^2 dt \right]. \quad (2.15)$$

The preceding lemma allows us to establish a cost estimate. The proof is given in Sect. 4.3.

**Lemma 2.13** For all  $C > 0$ , there exists a constant  $D(C) > 0$  such that for all  $\theta = (j^+, j^-, V) \in \mathcal{A}^0$  with  $\mathbb{E}[\int_0^T (X_t^\theta)^2 dt] \leq C$  and all  $\xi \in \mathcal{A}$  with  $\int_0^T \xi_s ds = -x_0$ , we have

$$\begin{aligned} & |J^0(\theta) - J^0(0, 0, V^\xi)| \\ & \leq D(C) \left( \sqrt{\mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right]} + \mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right] \right). \end{aligned}$$

As a consequence of the previous results, the optimal instantaneous price impact term converges to zero as  $\eta \rightarrow 0$ . The proof is given in Sect. 4.4.

**Lemma 2.14** We have

$$\lim_{\eta \rightarrow 0} \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\hat{\xi}_t^{\eta, \infty})^2 dt \right] = 0.$$

The cost estimate in Lemma 2.13 allows us to establish the optimality of the trading strategy  $\hat{\theta}$  by using the optimality of  $\hat{\xi}^{\eta, \infty}$  in the strict liquidation model with absolutely continuous strategies. It turns out that the minimal trading costs are fully determined by the initial value  $B_0^0$  of the process  $B^0$  along with the impact factor  $\gamma$  and the initial portfolio.

**Theorem 2.15** It holds that

$$\min_{\theta \in \mathcal{A}^0} J^0(\theta) = J^0(\hat{\theta}) = \frac{1}{2} \gamma (x_0)^2 B_0^0.$$

**Proof** Let us assume to the contrary that  $\hat{\theta}$  does not minimise the cost functional  $J^0$  over the set  $\mathcal{A}^0$ . Then there exist a strategy  $\theta = (j^+, j^-, V) \in \mathcal{A}^0$  and a constant

$\delta \in (0, \infty)$  such that  $J^0(\theta) + \delta \leq J^0(\hat{\theta})$ . Moreover, due to (2.15) and Lemma 2.13, for all  $\eta, \beta, \nu, \varepsilon \in (0, \infty)$ , we have (for convenience, let  $p(x) := x + \sqrt{x}$ )

$$\begin{aligned}
 & J^{\eta, \infty}(\xi^{\theta, \beta, \nu, \varepsilon}) - J^{\eta, \infty}(\hat{\xi}^{\eta, \infty}) \\
 & \leq J^{\eta, \infty}(\xi^{\theta, \beta, \nu, \varepsilon}) - J^0(\theta) + J^0(\hat{\theta}) - \delta - J^0(0, 0, V^{\hat{\xi}^{\eta, \infty}}) \\
 & \leq -\delta + \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\xi_t^{\theta, \beta, \nu, \varepsilon})^2 dt \right] + |J^0(\theta) - J^0(0, 0, V^{\xi^{\theta, \beta, \nu, \varepsilon}})| \\
 & \quad + |J^0(\hat{\theta}) - J^0(0, 0, V^{\hat{\xi}^{\eta, \infty}})| \\
 & \leq -\delta + \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\xi_t^{\theta, \beta, \nu, \varepsilon})^2 dt \right] \\
 & \quad + D \left( \mathbb{E} \left[ \int_0^T (X_t^\theta)^2 dt \right] \right) p \left( \mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^{\theta, \beta, \nu, \varepsilon})^2 dt \right] \right) \\
 & \quad + D \left( \mathbb{E} \left[ \int_0^T (X_t^{\hat{\theta}})^2 dt \right] \right) p \left( \mathbb{E} \left[ \int_0^T (\hat{X}_t^0 - \hat{X}_t^{\eta, \infty})^2 dt \right] \right). \quad (2.16)
 \end{aligned}$$

According to Lemma 2.12, we can first choose  $\beta, \nu, \varepsilon \in (0, \infty)$  and then, due to  $|\hat{X}_t^{\eta, \infty}| \leq x_0$  (cf. Theorem 2.2), choose  $\eta \in (0, \infty)$  small enough according to Theorem 2.7 such that the right-hand side of (2.16) is negative. This contradicts the optimality of  $\hat{\xi}^{\eta, \infty}$  for  $J^{\eta, \infty}$  and proves the optimality of  $\hat{\theta}$ .

It remains to compute  $J^0(\hat{\theta})$ . In view of (2.15), for all  $\eta \in (0, \infty)$ , we have

$$J^0(\hat{\theta}) = J^0(\hat{\theta}) - J^0(0, 0, V^{\hat{\xi}^{\eta, \infty}}) - \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\xi_t^{\eta, \infty})^2 dt \right] + J^{\eta, \infty}(\hat{\xi}^{\eta, \infty}).$$

The difference of the first two terms converges to zero as  $\eta \rightarrow 0$ , which is verified by using Lemma 2.13 and Theorem 2.7 as in the above calculation. The third term converges to zero by Lemma 2.14. Hence, using the representation of the value function given in [18],

$$\begin{aligned}
 J^0(\hat{\theta}) &= \lim_{\eta \rightarrow 0} J^{\eta, \infty}(\hat{\xi}^{\eta, \infty}) = \lim_{\eta \rightarrow 0} \frac{1}{2} A_0^{\eta, \infty} (x_0)^2 \\
 &= \frac{(x_0)^2}{2} \lim_{\eta \rightarrow 0} (\sqrt{\eta} D_0^{\eta, \infty} + \gamma B_0^{\eta, \infty}) = \frac{1}{2} \gamma (x_0)^2 B_0^0. \quad \square
 \end{aligned}$$

**Remark 2.16** If  $\lambda \equiv 0$ , our model is a special case of the model analysed in Ackermann et al. [1], which also contains cases when there is no optimal trading strategy. However, since  $\gamma$  is constant in our model, the processes “ $\mu$ ” and “ $\sigma$ ” introduced in [1] are equal to zero. This implies that the equation “ $\tilde{\beta} = Y$ ” holds in their notation. As shown in [1, Sect. 5], the process “ $M^\perp$ ” introduced therein is also equal to zero, which implies that “ $Y$ ” and hence “ $\tilde{\beta}$ ” is a semimartingale. Thus [1, Theorem 2.3 (ii)] confirms that under this property, an optimal trading strategy does indeed exist.

### 3 Proofs for Sect. 2.1

This section proves the results stated in Sect. 2.1. We start with a priori estimates and regularity properties for the coefficient processes that specify the optimal state processes.

#### 3.1 A priori estimates and regularity properties

##### 3.1.1 The case $\eta > 0$

The following estimates have been established in Graewe and Horst [18] and Horst and Xia [22], except for the upper bound on  $E^{\eta,N}$  for finite  $N$  which is stronger than the corresponding one in [22]. It can be established using the same arguments as in [18, proof of Proposition 3.2] by noting that  $D_T^{\eta,N} < \infty$  if  $N$  is finite.

**Lemma 3.1** *Let*

$$\bar{\kappa} := \sqrt{2 \max(\bar{\lambda}, \gamma \bar{\rho})}.$$

*For all  $\eta \in (0, \infty)$ ,  $N \in \mathcal{N}$  and  $s, t \in [0, T)$  with  $s \leq t$ , we have that*

$$\begin{aligned} e^{-\bar{\rho}(T-s)} &\leq B_s^{\eta,N} \leq 1, \\ 0 &\leq E_s^{\eta,N} \leq \gamma^{-1} \bar{\kappa} \tanh(\sqrt{\eta^{-1}} \bar{\kappa}(T-s)) \leq \gamma^{-1} \bar{\kappa}, \\ 0 &< D_s^{\eta,N} \leq \bar{\kappa} \coth(\sqrt{\eta^{-1}} \bar{\kappa}(T-s)) \leq \bar{\kappa} \coth(\sqrt{\eta^{-1}} \bar{\kappa}(T-t)). \end{aligned}$$

The preceding estimates allow us to prove that neither the optimal portfolio process nor the corresponding spread process change sign.

**Proof of Theorem 2.2** Put  $V(t) := \hat{X}_t^{\eta,N} \hat{Y}_t^{\eta,N}$ . Then  $V(0) = 0$  and  $V$  satisfies the ODE

$$\begin{aligned} \dot{V}(t) &= -(\sqrt{\eta^{-1}}(D_t^{\eta,N} + \gamma E_t^{\eta,N}) + \rho_t)V(t) \\ &\quad - \sqrt{\eta^{-1}}(E_t^{\eta,N}(\hat{Y}_t^{\eta,N})^2 + \gamma D_t^{\eta,N}(\hat{X}_t^{\eta,N})^2). \end{aligned}$$

As a result,

$$\begin{aligned} V(t) &= -\exp\left(-\int_0^t (\sqrt{\eta^{-1}}(D_u^{\eta,N} + \gamma E_u^{\eta,N}) + \rho_u) du\right) \\ &\quad \times \int_0^t \sqrt{\eta^{-1}}(E_s^{\eta,N}(\hat{Y}_s^{\eta,N})^2 + \gamma D_s^{\eta,N}(\hat{X}_s^{\eta,N})^2) \\ &\quad \times \exp\left(\int_0^s (\sqrt{\eta^{-1}}(D_u^{\eta,N} + \gamma E_u^{\eta,N}) + \rho_u) du\right) ds. \end{aligned}$$

In view of Lemma 3.1, this shows that  $V(t) < 0$  on  $(0, T)$ . Hence strict positivity of  $\hat{X}_0^{\eta, N}$  yields

$$\mathbb{P}[\hat{X}_t^{\eta, N} > 0, \hat{Y}_t^{\eta, N} < 0 \text{ for all } t \in (0, T)] = 1.$$

Thus the definition of  $\hat{Z}^{\eta, N}$  along with (2.3) yields  $\partial_t \hat{Z}_t^{\eta, N} = \rho_t \hat{Y}_t^{\eta, N} < 0$  on  $(0, T)$ . Moreover,

$$\hat{X}_t^{\eta, N} < \gamma^{-1} \hat{Z}_t^{\eta, N} \leq \gamma^{-1} \hat{Z}_0^{\eta, N} = x_0 \quad \text{and} \quad \hat{Y}_t^{\eta, N} > -\hat{Z}_t^{\eta, N} \geq -\gamma x_0. \quad \square$$

Next, we prove that the process  $B^{\eta, N}$  satisfies an  $L^1$  uniform continuity property. We refer to Appendix A for a discussion of general regularity properties of stochastic processes.

**Lemma 3.2** *The process  $B^{\eta, N}$  satisfies Condition A.1 stated in Appendix A on  $[0, T]$  uniformly in  $N \in \mathcal{N}$  and  $\eta \in (0, H]$ , for all  $H \in (0, \infty)$ .*

**Proof** Let  $\varepsilon, \varepsilon_1 > 0$ ,  $s > T - \varepsilon_1$  and let  $V$  and  $\tau$  be arbitrary according to the definition of Condition A.1. By Lemma 3.1, if  $\varepsilon_1$  is small enough,

$$|\mathbb{E}[V(B_t^{\eta, N} - B_s^{\eta, N})]| \leq (1 - \exp(-\bar{\rho}\varepsilon_1))\mathbb{E}[|V|] \leq \varepsilon\mathbb{E}[|V|].$$

If  $s \leq T - \varepsilon_1$ , the assertion follows from the integral representation (2.10) along with the estimates established in Lemma 3.1 by using that the stochastic integral in (2.10) is a martingale on  $[0, T - \varepsilon_1]$ . We emphasise that  $Z^{\eta, N, B}$  is possibly defined only on  $[0, T -]$ , and so the stochastic integral may be a martingale only away from the terminal time.  $\square$

### 3.1.2 The case $\eta = 0$

We are now going to establish a priori estimates on the candidate limiting coefficient processes. First, we show that Assumption 2.3 directly implies the following regularity result for the parameter processes.

**Lemma 3.3** *The processes  $\chi$ ,  $\rho$ ,  $\lambda$ ,  $\varphi$  and  $\varphi^{-1}$  satisfy Condition A.2 introduced in Appendix A.*

**Proof** The process  $\chi$  satisfies Condition A.2 due to Lemmas A.11, A.4 and A.7. The rest immediately follows by Lemma A.10.  $\square$

We are now ready to prove that the process  $B^0$  is well defined.

**Proof of Lemma 2.5** The existence result follows from a standard argument. In fact, it is well known that for any  $\bar{b} \in [1, \infty)$ , the BSDE

$$-dB_t^{\bar{b}} = \psi^{\bar{b}}(t, \chi_t, B_t^{\bar{b}}) dt - Z_t^{\bar{b}} dW_t, \quad B_T^{\bar{b}} = 1 \quad (3.1)$$

with Lipschitz-continuous driver (we recall  $f^\lambda, f^\rho$  defined in Assumption 2.3)

$$\begin{aligned} \psi^{\bar{b}}(t, \chi, b) := & -\frac{1}{f^\lambda(t, \chi) + 2\gamma f^\rho(t, \chi)} \\ & \times \left( \gamma f^\rho(t, \chi)^2 (\bar{b}^2 \wedge b^2) - 2f^\lambda(t, \chi) f^\rho(t, \chi) (1 - (b \wedge \bar{b})^+) \right) \end{aligned}$$

has a unique solution  $(B^{\bar{b}}, Z^{\bar{b}}) \in \mathcal{L}_{\mathcal{P}}^2(\Omega \times [0, T]; \mathbb{R} \times \mathbb{R}^m)$  (cf. Theorem A.13). Let us then define the functions  $\underline{\phi}, \bar{\phi}: \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$  by

$$\underline{\phi}(t, b, z) := -\varphi^{-2} \gamma \bar{\rho}^2 b, \quad \bar{\phi}(t, b, z) := 0.$$

By definition,  $(\underline{B}^0, 0)$  is the unique solution to the BSDE with driver  $\underline{\phi}$  and terminal condition 1, where the lower bound  $\underline{B}^0$  on the process  $B^0$  was defined in (2.8). Likewise,  $(1, 0)$  is the unique solution of the BSDE with driver  $\bar{\phi}$  and the same terminal condition. Since

$$\psi^{\bar{b}}(t, \chi_t, \underline{B}_t^0) > \underline{\phi}(t, \underline{B}_t^0, 0) \quad \text{and} \quad \psi^{\bar{b}}(t, \chi_t, 1) < \bar{\phi}(t, 1, 0),$$

the standard comparison principle for BSDEs with Lipschitz-continuous drivers yields

$$\underline{B}^0 \leq B^{\bar{b}} \leq 1.$$

This proves that  $(B^1, Z^1)$  is the desired unique bounded solution to the BSDE (2.7). The second assertion follows Theorem A.13 applied to the BSDE (3.1) for  $\bar{b} = 1$ .  $\square$

Having established the existence of the process  $B^0$ , the processes  $D^0$  and  $E^0$  are well defined. The following lemma establishes estimates and regularity properties for  $D^0$  and  $E^0$ .

**Lemma 3.4** *We have the a priori estimates*

$$\begin{aligned} \frac{\gamma \rho}{\bar{\varphi}} \exp \left( -\frac{\gamma \bar{\rho}^2}{\bar{\varphi}^2} T \right) & \leq D_t^0 \leq \varphi^{-1} (\bar{\lambda} + \gamma \bar{\rho}), \\ \bar{\varphi}^{-1} \underline{\rho} & \leq E_t^0 \leq 2\varphi^{-1} \bar{\rho}. \end{aligned}$$

Moreover, the processes  $D^0$  and  $E^0$  satisfy Condition A.2.

**Proof** The a priori estimates can be obtained by plugging the bounds on  $B^0$  from Lemma 2.5 into the definitions of  $D^0$  and  $E^0$  given in (2.9). Moreover, if we denote by  $h$  the function derived from Lemma 2.5, then for all  $t \in [0, T]$ ,

$$D_t^0 = \left( \frac{f^\lambda + \gamma f^\rho h}{\sqrt{f^\lambda + 2\gamma f^\rho}} \right)(t, \chi_t), \quad E_t^0 = \left( \frac{(2-h)f^\rho}{\sqrt{f^\lambda + 2\gamma f^\rho}} \right)(t, \chi_t).$$

In view of Assumption 2.3, the processes  $D^0$  and  $E^0$  can be represented as uniformly continuous functions of the factor process  $\chi$ , and hence the assertion follows from Lemmas 3.3 and A.10.  $\square$

The next lemma can be viewed as the analogue to Theorem 2.2 in the case  $\eta = 0$ .

**Lemma 3.5** *It holds*

$$\mathbb{P}[\hat{X}_t^0 \in (0, x_0), \hat{Y}_t^0 \in (-\gamma x_0, 0) \text{ for all } t \in (0, T)] = 1.$$

**Proof** Due to Lemma 3.4,  $D^0$  is positive and hence, for  $t \in (0, T)$ ,

$$0 < \hat{Z}_t^0 < x_0 \gamma.$$

Since  $D^0 + \gamma E^0 = \varphi$  on  $[0, T]$ , it thus follows from Lemma 3.4 that for all  $t \in (0, T)$ ,

$$\begin{aligned} 0 < \hat{X}_t^0 &= \frac{\varphi_t - D_t^0}{\gamma \varphi_t} \hat{Z}_t^0 < \gamma^{-1} \hat{Z}_t^0 < x_0, \\ 0 > \hat{Y}_t^0 &= -\frac{\varphi_t - \gamma E_t^0}{\varphi_t} \hat{Z}_t^0 > -\hat{Z}_t^0 > -\gamma x_0. \end{aligned} \quad \square$$

## 3.2 Proof of the convergence results

In this section, we prove our main convergence results. We start with the convergence of the coefficients of the ODE system (2.3). Subsequently, we prove that the convergence of the coefficients yields convergence of the state process.

### 3.2.1 Proof of Proposition 2.6

The proof of Proposition 2.6 is split into a series of lemmas. In the first step, we establish the convergence, as  $\eta \rightarrow 0$ , of the auxiliary processes

$$F_t^{\eta, N} := D_t^{\eta, N} + \gamma E_t^{\eta, N} \longrightarrow F^0 := \varphi$$

and

$$G^{\eta, N} := \rho B^{\eta, N} + \varphi E^{\eta, N} \longrightarrow G^0 := \rho B^0 + \varphi E^0 = 2\rho.$$

On  $[0, T]$ , the processes  $F^{\eta, N}$  and  $G^{\eta, N}$  satisfy the dynamics

$$\begin{aligned} dF_t^{\eta, N} &= \sqrt{\eta^{-1}}((F_t^{\eta, N})^2 - (\varphi_t)^2) dt + 2\gamma \rho_t E_t^{\eta, N} dt \\ &\quad + \sqrt{\eta^{-1}}(Z_t^{\eta, N, A} - 2\gamma Z_t^{\eta, N, B} + \gamma^2 Z_t^{\eta, N, C}) dW_t, \end{aligned} \quad (3.2)$$

$$\begin{aligned} -d(\varphi^{-1} G^{\eta, N})_t &= \sqrt{\eta^{-1}}(-G_t^{\eta, N} + 2\rho_t - E_t^{\eta, N}(F_t^{\eta, N} - \varphi_t)) dt - 2\rho_t E_t^{\eta, N} dt \\ &\quad - d(\varphi^{-1} \rho B^{\eta, N})_t + \sqrt{\eta^{-1}}(Z_t^{\eta, N, B} - \gamma Z_t^{\eta, N, C}) dW_t. \end{aligned} \quad (3.3)$$

A general convergence result for integral equations of the above form is established in Appendix B.1.2. It allows us to prove the following two results.

**Lemma 3.6** *Let  $f^{\eta,N} := F^{\eta,N} - \varphi$ . For all  $\varepsilon > 0$ , there exists an  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0]$  and all  $N \in \mathcal{N}$ ,*

$$\mathbb{P}\left[\sup_{s \in [0, T-\varepsilon]} f_s^{\eta,N} \leq \varepsilon, \inf_{s \in [0, T)} f_s^{\eta,N} \geq -\varepsilon\right] = 1.$$

**Proof** For every  $N \in \mathcal{N}$ , we apply Lemma B.4 to  $P^\eta := F^{\eta,N}$  with

$$\begin{aligned} \psi &:= 1, & a(x, y) &:= (x^+)^2 - y^2, & P^0 &:= \varphi, \\ q^\eta &:= 0, & L_t^\eta &:= 2\gamma \int_0^t \rho_s E_s^{\eta,N} ds, \end{aligned}$$

noticing that  $\varepsilon \mapsto \eta_0 = \eta_0(\varepsilon)$  is independent of  $N$ . Assumption B.3 i) is satisfied with  $\bar{P} := \bar{\varphi} + 3\bar{\kappa}$  due to Lemma 3.1. If  $N = \infty$ , then Assumption B.3 ii) a) is satisfied due to Theorem 2.1 and the a priori estimates derived in Lemmas 3.1 and 3.4. If  $N < \infty$  and  $\eta \leq \bar{\varphi}^{-2}$ , then  $E_T^{\eta,N} = 0$  and

$$f_T^{\eta,N} = D_T^{\eta,N} + \gamma E_T^{\eta,N} - D_T^0 - \gamma E_T^0 = \sqrt{\eta^{-1}(N - \gamma)} - \varphi_T \geq \sqrt{\eta^{-1}} - \bar{\varphi} \geq 0,$$

which shows Assumption B.3 ii) b). Assumption B.3 iii) is trivially satisfied because  $q^\eta = 0$ . Assumption B.3 iv) follows by direct computation using  $\bar{\varepsilon} := \varphi/2$  and  $\beta := \varphi/2$ . Assumption B.3 v) follows from Lemmas 3.3 and 3.1.  $\square$

It is not difficult to show that similar arguments as those used to prove the convergence of  $f^{\eta,N}$  can be applied to  $-g^{\eta,N} := -G^{\eta,N} + 2\rho$ . As a result, the intervals  $[0, T)$  and  $[0, T - \varepsilon]$  in the statement of the convergence result for  $g^{\eta,N}$  need to be swapped. We omit the proof of the corresponding next result.

**Lemma 3.7** *For all  $\varepsilon > 0$ , there exists an  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0]$  and all  $N \in \mathcal{N}$ ,*

$$\mathbb{P}\left[\sup_{s \in [0, T)} g_s^{\eta,N} \leq \varepsilon, \inf_{s \in [0, T-\varepsilon]} g_s^{\eta,N} \geq -\varepsilon\right] = 1.$$

We are now going to prove the almost sure convergence to zero of the process  $b^{\eta,N} := B^{\eta,N} - B^0$ . To this end, we first observe that

$$D^{\eta,N} = F^{\eta,N} - \gamma E^{\eta,N}, \quad \varphi E^{\eta,N} = G^{\eta,N} - \rho B^{\eta,N}$$

yields

$$-D^{\eta,N} E^{\eta,N} = -f^{\eta,N} E^{\eta,N} - G^{\eta,N} + \rho B^{\eta,N} + \gamma \varphi^{-2} (G^{\eta,N} - \rho B^{\eta,N})^2.$$

Plugging this into (2.10) shows that

$$dB_t^{\eta,N} = \left( -G_t^{\eta,N} - f_t^{\eta,N} E_t^{\eta,N} + 2\rho_t B_t^{\eta,N} + \gamma(\varphi_t)^{-2} (G_t^{\eta,N} - \rho_t B_t^{\eta,N})^2 \right) dt \\ + Z_t^{\eta,N,B} dW_t$$

on  $[0, T)$ . Performing an analogous computation for  $B^0$  and subtracting the two equations yields

$$db_t^{\eta,N} = \left( (b_t^{\eta,N})^2 \gamma(\varphi_t)^{-2} (\rho_t)^2 + b_t^{\eta,N} (2\rho_t - 2\gamma(\varphi_t)^{-2} \rho_t (G_t^0 - \rho_t B_t^0)) \right. \\ \left. - f_t^{\eta,N} E_t^{\eta,N} - g_t^{\eta,N} + \gamma(\varphi_t)^{-2} g_t^{\eta,N} (g_t^{\eta,N} + 2G_t^0 - \rho_t (2b_t^{\eta,N} + 2B_t^0)) \right) dt \\ + (Z_t^{\eta,N,B} - Z_t^{0,B}) dW_t$$

on  $[0, T)$ . Using that

$$2\rho_t - 2\gamma(\varphi_t)^{-2} \rho_t (G_t^0 - \rho_t B_t^0) = 2\rho_t (\varphi_t)^{-1} D_t^0, \\ g_t^{\eta,N} + 2G_t^0 - \rho_t (2b_t^{\eta,N} + 2B_t^0) = -g_t^{\eta,N} + 2\varphi_t E_t^{\eta,N}$$

shows that

$$db_t^{\eta,N} = \left( (b_t^{\eta,N})^2 \gamma(\varphi_t)^{-2} (\rho_t)^2 + b_t^{\eta,N} 2(\varphi_t)^{-1} \rho_t D_t^0 - f_t^{\eta,N} E_t^{\eta,N} \right. \\ \left. + g_t^{\eta,N} (-1 - \gamma(\varphi_t)^{-2} g_t^{\eta,N} + 2\gamma(\varphi_t)^{-1} E_t^{\eta,N}) \right) dt \\ + (Z_t^{\eta,N,B} - Z_t^{0,B}) dW_t$$

on  $[0, T)$ . This BSDE is different from (3.2) and (3.3). We apply Lemma B.2 to prove the following result.

**Lemma 3.8** *For all  $\varepsilon > 0$ , there exists an  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0]$  and all  $N \in \mathcal{N}$ ,*

$$\mathbb{P} \left[ \sup_{s \in [0, T]} |b_s^{\eta,N}| \leq \varepsilon \right] = 1.$$

**Proof** For every  $N \in \mathcal{N}$ , we apply Lemma B.2 with

$$a(t, b) := b^2 \gamma(\varphi_t)^{-2} (\rho_t)^2 + b 2(\varphi_t)^{-1} \rho_t D_t^0, \\ q_t^\eta := -f_t^{\eta,N} E_t^{\eta,N} + g_t^{\eta,N} (2\gamma(\varphi_t)^{-1} E_t^{\eta,N} - 1 - \gamma(\varphi_t)^{-2} g_t^{\eta,N}).$$

Assumption B.1 i) follows from the a priori estimates on  $B^{\eta,N}$  (Lemma 3.1),  $B^0$  (Lemma 2.5) and  $D^0$  (Lemma 3.4). Assumption B.1 ii) follows from the a priori



estimates on  $B^{\eta,N}$  and  $B^0$ , where the mapping  $\varepsilon \mapsto \delta$  is independent of  $N$ . Assumption B.1 iii) follows from the same estimates and Lemmas 3.6 and 3.7, where the choice of  $\eta_1$  is independent of  $N$ . Assumption B.1 iv) is satisfied because we can choose  $\varepsilon_1 > 0$  small enough such that

$$-\varepsilon_1 \gamma(\varphi)^{-1} \rho + 2D^0 \geq 0. \quad \square$$

### 3.2.2 Proof of Theorem 2.7

First we need to prove an auxiliary result:

**Lemma 3.9** *Let  $I \subseteq \mathbb{R}$  be a compact interval and  $Y = (Y_t)_{t \in I}$  a continuous, adapted,  $\mathbb{R}^d$ -valued stochastic process. Then the modulus of continuity*

$$\sup_{s,t \in I, |s-t| \leq v} |Y_t - Y_s|$$

*converges to 0 in probability as  $v \rightarrow 0$ .*

**Proof** For all  $\omega \in \Omega$ ,  $Y(\omega)$  is uniformly continuous; hence the modulus of continuity converges to 0 as  $v \rightarrow 0$   $\mathbb{P}$ -a.s. This implies convergence in probability.  $\square$

**Proof of Theorem 2.7** Convergence on compact subintervals of  $[0, T)$  follows from Theorem B.8 applied to the ODEs

$$\begin{aligned} \partial_t \hat{X}_t^{\eta,N} &= -\sqrt{\eta^{-1}} F_t^{\eta,N} \left( \hat{X}_t^{\eta,N} - \frac{E_t^{\eta,N}}{F_t^{\eta,N}} \hat{Z}_t^{\eta,N} \right), \\ \partial_t \hat{Z}_t^{\eta,N} &= \gamma \rho_t \hat{X}_t^{\eta,N} - \rho_t \hat{Z}_t^{\eta,N} \end{aligned} \quad (3.4)$$

and

$$\hat{X}_t^0 = \frac{E_t^0}{\varphi_t} \hat{Z}_t^0, \quad \hat{Z}_t^0 = \hat{Z}_0^0 \exp \left( \int_0^t \left( \gamma \rho_s \frac{E_s^0}{\varphi_s} - \rho_s \right) ds \right)$$

on  $[0, T)$ . In fact, Condition (B.10) follows from Proposition 2.6 and Condition (B.11) from Theorem 2.2. Finally,  $D^0 > 0$  (see Lemma 3.4), from which we deduce that  $x_0 > \frac{E_0^0}{\varphi_0} \hat{Z}_0^0$ . Thus by Theorem B.8, applied once for every  $N \in \mathcal{N}$  and every  $v > 0$ , there exists for all  $v, \delta \in (0, \infty)$  some  $\eta_1 = \eta_1(v, \delta) \in (0, \infty)$  such that for all  $\eta \in (0, \eta_1]$  and all  $N \in \mathcal{N}$ , the set

$$\begin{aligned} M_0^{v,\eta,N} := \Big\{ & \sup_{t \in [0, T-v]} \max(|\hat{Z}_t^0 - \hat{Z}_t^{\eta,N}|, |\hat{X}_t^0 - \hat{X}_t^{\eta,N}|) \leq v, \\ & \inf_{t \in [v, T-v]} (\hat{X}_t^0 - \hat{X}_t^{\eta,N}) \geq -v \Big\} \end{aligned}$$

satisfies

$$\mathbb{P}[M_0^{v,\eta,N}] > 1 - \delta/2. \quad (3.5)$$

Near the terminal time, we cannot use this theorem since the convergence  $\frac{E^{\eta,N}}{F^{\eta,N}} \rightarrow \frac{E^0}{\varphi}$  holds only on  $[0, T-]$ . Instead, we apply increment bounds on intervals of the form  $[T-\nu, T]$  to prove that for all  $\eta > 0$ ,  $N \in \mathcal{N}$  and for all sufficiently small

$$\nu \in (0, \varepsilon/2] \quad \text{and} \quad \beta \in (0, \varepsilon/3], \quad (3.6)$$

the set

$$M_1^{v,\varepsilon,\eta,N} := \left\{ \inf_{t \in [T-\nu, T)} (\hat{X}_t^0 - \hat{X}_t^{\eta,N}) \geq -\varepsilon \right\}$$

contains the set  $M_0^{v,\eta,N} \cap M_2^{\beta,v} \cap M_3^{v,\eta,N}$ , where

$$M_2^{\beta,v} := \left\{ \sup_{s,t \in [T-\nu, T)} (|\varphi_s - \varphi_t| \vee |E_s^0 - E_t^0| \vee |\hat{X}_s^0 - \hat{X}_t^0|) \leq \beta \right\},$$

$$M_3^{v,\eta,N} := \left\{ \sup_{t \in [T-\nu, T)} (e_t^{\eta,N} \vee (-f_t^{\eta,N})) \leq \nu \right\}.$$

The probability of the last two events can be made large by Lemma 3.9 and Proposition 2.6: for all  $\beta, \delta, \nu \in (0, \infty)$ , there exist  $\nu_0(\beta, \delta), \eta_2(\nu) \in (0, \infty)$  such that

$$\inf_{\nu \leq \nu_0(\beta, \delta)} \mathbb{P}[M_2^{\beta,v}] > 1 - \delta/2, \quad \inf_{\eta \leq \eta_2(\nu)} \inf_{N \in \mathcal{N}} \mathbb{P}[M_3^{v,\eta,N}] = 1.$$

In order to see that  $M_1^{v,\varepsilon,\eta,N} \supseteq M_0^{v,\eta,N} \cap M_2^{\beta,v} \cap M_3^{v,\eta,N}$ , we assume to the contrary that there exists

$$\omega \in (M_0^{v,\eta,N} \cap M_2^{\beta,v} \cap M_3^{v,\eta,N}) \setminus M_1^{v,\varepsilon,\eta,N}.$$

To obtain the desired contradiction, we show that for any such  $\omega$ , there exists a time  $s(\omega) \in [T-\nu, T)$  for which we can deduce both nonnegativity and strict negativity of  $\partial_t \hat{X}_{s(\omega)}^{\eta,N}(\omega)$  simultaneously.

We start with the choice of  $s(\omega)$ . Because we have  $\omega \notin M_1^{v,\varepsilon,\eta,N}$ , there exists some  $t \in [T-\nu, T)$  with  $\hat{X}_t^{\eta,N}(\omega) - \hat{X}_t^0(\omega) > \varepsilon$ . Since  $\omega \in M_2^{\beta,v}$  and due to (3.6), this yields that

$$\hat{X}_t^{\eta,N}(\omega) > \hat{X}_{T-\nu}^0(\omega) + \frac{2\varepsilon}{3}.$$

Since  $\omega$  belongs to  $M_0^{v,\eta,N}$  and due to (3.6), this implies that

$$\hat{X}_{T-\nu}^{\eta,N}(\omega) < \hat{X}_{T-\nu}^0(\omega) + \frac{2\varepsilon}{3} < \hat{X}_t^{\eta,N}(\omega).$$

Now we can choose  $s(\omega) \in (T-\nu, t)$  minimal with the property that

$$\hat{X}_{s(\omega)}^{\eta,N}(\omega) = \hat{X}_{T-\nu}^0(\omega) + \frac{2\varepsilon}{3}. \quad (3.7)$$

Due to the minimality of  $s(\omega)$ , we have  $\partial_t \hat{X}_{s(\omega)}^{\eta,N}(\omega) \geq 0$ .

We now show that this derivative must also be strictly negative. In fact, due to (3.4),

$$\sqrt{\eta} \partial_t \hat{X}_{s(\omega)}^{\eta, N}(\omega) = -F_{s(\omega)}^{\eta, N}(\omega) \hat{X}_{s(\omega)}^{\eta, N}(\omega) + E_{s(\omega)}^{\eta, N}(\omega) \hat{Z}_{s(\omega)}^{\eta, N}(\omega).$$

Since  $\hat{Z}^{\eta, N}$  is nonincreasing (Theorem 2.2), using (3.7) again, the right-hand side of the above equation can be bounded from above by

$$\begin{aligned} & -\varphi_{T-\nu}(\omega) \hat{X}_{T-\nu}^0(\omega) + (\varphi_{T-\nu}(\omega) - \varphi_{s(\omega)}(\omega)) \hat{X}_{T-\nu}^0(\omega) - f_{s(\omega)}^{\eta, N}(\omega) \hat{X}_{T-\nu}^0(\omega) \\ & - F_{s(\omega)}^{\eta, N}(\omega) \frac{2\varepsilon}{3} + E_{T-\nu}^0(\omega) \hat{Z}_{T-\nu}^0(\omega) + E_{T-\nu}^0(\omega) (\hat{Z}_{T-\nu}^{\eta, N}(\omega) - \hat{Z}_{T-\nu}^0(\omega)) \\ & + (E_{s(\omega)}^0(\omega) - E_{T-\nu}^0(\omega)) \hat{Z}_{T-\nu}^{\eta, N}(\omega) + e_{s(\omega)}^{\eta, N}(\omega) \hat{Z}_{T-\nu}^{\eta, N}(\omega). \end{aligned}$$

Since  $\omega \in M_3^{\nu, \eta, N}$ , we have  $F_{s(\omega)}^{\eta, N}(\omega) = \varphi_{s(\omega)}(\omega) + f_{s(\omega)}^{\eta, N}(\omega) \geq \varphi - \nu$ , which is strictly positive if  $\nu < \varphi$ . Moreover, we have  $-\varphi \hat{X}^0 + E^0 \hat{Z}^0 = 0$  on  $[0, T)$ . Because also  $\omega \in M_0^{\nu, \eta, N} \cap M_2^{\beta, \nu} \cap M_3^{\nu, \eta, N}$  and  $\hat{X}^0$ ,  $E^0$  and  $\hat{Z}^{\eta, N}$  are bounded, this shows that  $\partial_t \hat{X}_{s(\omega)}^{\eta, N}(\omega) < 0$  if  $\nu$  and  $\beta$  are chosen small enough. The convergence of  $\hat{Y}^{\eta, N} - \hat{Y}^0$  on  $[0, T-]$  follows from

$$\hat{Y}^{\eta, N} - \hat{Y}^0 = \gamma(\hat{X}^{\eta, N} - \hat{X}^0) - (\hat{Z}^{\eta, N} - \hat{Z}^0)$$

and (3.5). The convergence on  $[T-, T]$  follows from the fact that

$$\sup_{s \in [T-\nu, T]} |(\hat{Z}_s^{\eta, N} - \hat{Z}_s^0) - (\hat{Z}_{T-\nu}^{\eta, N} - \hat{Z}_{T-\nu}^0)|$$

can be made arbitrarily small by choosing  $\nu$  small since  $\partial_t \hat{Z}^{\eta, N} = \rho \hat{Y}^{\eta, N}$ ,  $\partial_t \hat{Z}^0 = \rho \hat{Y}^0$  and both  $\hat{Y}^{\eta, N}$  and  $\hat{Y}^0$  are bounded.  $\square$

### 3.2.3 Proof of Proposition 2.8

The proof of the convergence of the optimal portfolio processes in the Skorohod  $\mathcal{M}_2$ -sense follows from Theorem 2.7 if strict liquidation is required. If strict liquidation is not required, the results of Sect. 2.2 are required to establish the assertion. This is not a circular argument since the proofs of Sect. 2.2 only use the results of Sect. 2.1 concerning the liquidating case  $N = \infty$ .

Since  $\hat{\theta} = (0, \hat{j}^-, \hat{V})$  is optimal in the model introduced in Sect. 2.2 (cf. Theorem 2.15), it is easy to show that the strategy  $\hat{\theta}^q := (\hat{j}^{+, q}, \hat{j}^{-, q}, \hat{V}^q)$  is admissible for every  $q \in \mathbb{R}$ , where

$$\begin{aligned} \hat{V}_t^q &= \hat{V}_t + qt, \\ \hat{j}_t^{+, q} - \hat{j}_t^{-, q} &= -\hat{j}_t^- - \mathbb{1}_{\{T\}}(t)qT. \end{aligned}$$

The following lemma shows that we can express the cost term corresponding to the transient price impact without Itô integrals. The proof is an immediate consequence

of Itô's formula for semimartingales (see Protter [32, Theorem II.32]) and the fact that  $\gamma^2[V]$  is equal to the continuous part of  $[Y^\theta]$ . We omit it.

**Lemma 3.10** *For all  $\theta = (j^+, j^-, V) \in \mathcal{A}^0$  and all  $t \in [0, T]$ , it holds that*

$$\int_{(0,t]} \frac{1}{2} (Y_{s-}^\theta + Y_s^\theta) dX_s^\theta = \frac{1}{2\gamma} (Y_t^\theta)^2 - \frac{\gamma}{2} (j_0^+ - j_0^-)^2 - \frac{\gamma}{2} [V]_t + \frac{1}{\gamma} \int_0^t \rho_s (Y_s^\theta)^2 ds.$$

Using Lemma 3.10, it is not difficult to check that the mapping  $q \mapsto J^0(\hat{\theta}^q)$  is differentiable and that

$$\begin{aligned} 0 &= \left. \frac{\partial J^0(\hat{\theta}^q)}{\partial q} \right|_{q=0} \\ &= \mathbb{E} \left[ -\hat{Y}_T^0 e^{-\int_0^T \rho_u du} \int_0^T t \rho_t e^{\int_0^t \rho_u du} dt + \int_0^T 2t \rho_t \hat{Y}_t^0 dt \right. \\ &\quad \left. - \int_0^T 2\rho_t \hat{Y}_t^0 e^{-\int_0^t \rho_u du} \int_0^t s \rho_s e^{\int_0^s \rho_u du} ds dt + \int_0^T t \lambda_t \hat{X}_t^0 dt \right]. \quad (3.8) \end{aligned}$$

This allows us to prove that full liquidation is optimal if  $\eta \rightarrow 0$  even if it is not formally required.

**Lemma 3.11** *We have that  $\lim_{\eta \rightarrow 0} \sup_{N \in \mathcal{N}} \mathbb{E}[\hat{X}_T^{\eta, N}] = 0$ .*

**Proof** We assume to the contrary that  $\limsup_{\eta \rightarrow 0} \sup_{N \in \mathcal{N}} \mathbb{E}[\hat{X}_T^{\eta, N}] > 0$  and prove that this contradicts the optimality of  $J^{\eta, N}(\hat{\xi}^{\eta, N})$ . To this end, we consider the admissible trading strategies

$$\hat{\xi}_t^{\eta, N, q} := \hat{\xi}_t^{\eta, N} + q,$$

compute the derivative of the function  $q \mapsto J^{\eta, N}(\hat{\xi}^{\eta, N, q})$  and show that the derivative at  $q = 0$  does not vanish for small  $\eta$  if

$$\limsup_{\eta \rightarrow 0} \sup_{N \in \mathcal{N}} \mathbb{E}[\hat{X}_T^{\eta, N}] > 0.$$

Obviously,  $\hat{\xi}^{\eta, N, q} \in \mathcal{A}$  and

$$\hat{X}_t^{\eta, N, q} := X_t^{\hat{\xi}^{\eta, N, q}} = \hat{X}_t^{\eta, N} + qt.$$

Straightforward computations lead to

$$\begin{aligned} 0 &= \left. \frac{\partial J^{\eta, N}(\hat{\xi}^{\eta, N}, q)}{\partial q} \right|_{q=0} \\ &= \mathbb{E} \left[ \eta(\hat{X}_T^{\eta, N} - x_0) - \hat{Y}_T^{\eta, N} e^{-\int_0^T \rho_u du} \int_0^T t \rho_t e^{\int_0^t \rho_u du} dt + \int_0^T 2t \rho_t \hat{Y}_t^{\eta, N} dt \right. \\ &\quad - \int_0^T 2\rho_t \hat{Y}_t^{\eta, N} e^{-\int_0^t \rho_u du} \int_0^t s \rho_s e^{\int_0^s \rho_u du} ds dt + \int_0^T t \lambda_t \hat{X}_t^{\eta, N} dt \\ &\quad \left. + (N - \gamma)T \hat{X}_T^{\eta, N} + \gamma \hat{X}_T^{\eta, N} e^{-\int_0^T \rho_u du} \int_0^T t \rho_t e^{\int_0^t \rho_u du} dt \right]. \end{aligned}$$

Subtracting (3.8) yields

$$\begin{aligned} 0 &= \mathbb{E} \left[ -\eta x_0 + \hat{X}_T^{\eta, N} (\eta + (N - \gamma)T) + \int_0^T 2t \rho_t (\hat{Y}_t^{\eta, N} - \hat{Y}_t^0) dt \right. \\ &\quad + (\hat{Z}_T^{\eta, N} - \hat{Z}_T^0) e^{-\int_0^T \rho_u du} \int_0^T t \rho_t e^{\int_0^t \rho_u du} dt + \int_0^T t \lambda_t (\hat{X}_t^{\eta, N} - \hat{X}_t^0) dt \\ &\quad \left. - \int_0^T 2\rho_t (\hat{Y}_t^{\eta, N} - \hat{Y}_t^0) e^{-\int_0^t \rho_u du} \int_0^t s \rho_s e^{\int_0^s \rho_u du} ds dt \right] \\ &\geq -\eta x_0 + (\eta + (N - \gamma)T) \mathbb{E}[\hat{X}_T^{\eta, N}] - T \bar{\lambda} \mathbb{E} \left[ \int_0^T |\hat{X}_t^{\eta, N} - \hat{X}_t^0| dt \right] \\ &\quad - 2(T \bar{\rho} + \bar{\rho}^2 T^2 e^{T \bar{\rho}}) \mathbb{E} \left[ \int_0^T |\hat{Y}_t^{\eta, N} - \hat{Y}_t^0| dt \right] - T^2 \bar{\rho} e^{T \bar{\rho}} \mathbb{E}[|\hat{Z}_T^{\eta, N} - \hat{Z}_T^0|]. \end{aligned}$$

In view of Theorems 2.2 and 2.7, Lemma 3.5 and using

$$\partial_t(\hat{Z}^{\eta, N} - \hat{Z}^0) = \rho(\hat{Y}^{\eta, N} - \hat{Y}^0),$$

the sum of the three last expected values is small uniformly in  $N$  if  $\eta$  is small. Hence, if  $\limsup_{\eta \rightarrow 0} \sup_{N \in \mathcal{N}} \mathbb{E}[\hat{X}_T^{\eta, N}] > 0$ , then the sum on the right-hand side of the above inequality is strictly positive when first choosing  $\eta_0 > 0$  small enough and then choosing  $(\tilde{\eta}, \tilde{N})$  with  $\tilde{\eta} \leq \eta_0$  such that

$$\mathbb{E}[\hat{X}_T^{\tilde{\eta}, \tilde{N}}] \geq \frac{1}{2} \limsup_{\eta \rightarrow 0} \sup_{N \in \mathcal{N}} \mathbb{E}[\hat{X}_T^{\eta, N}]. \quad \square$$

We are now ready to prove the convergence of  $\hat{X}^{\eta, N}$  to  $\hat{X}^0$  in the Skorohod  $\mathcal{M}_2$ -sense. To this end, we have to bound the distance of each point of any of the graphs to the other graph. In the inner interval  $[\varepsilon, T - \varepsilon]$ , it is enough to consider  $\hat{X}^{\eta, N} - \hat{X}^0$ , which we have bounded by Theorem 2.7.

**Proof of Proposition 2.8** To prove that the probability of

$$d_{\mathcal{M}_2}(\hat{X}^{\eta, N}, \hat{X}^0) \leq \varepsilon$$

is large for small  $\eta > 0$ , we need to prove that the distance of any point  $(t, x)$  from either  $G_{\hat{X}^{\eta, N}(\omega)}$  or  $G_{\hat{X}^0(\omega)}$  to the respective other graph is small on a set of large probability. To this end, we fix a small enough  $\nu \in (0, \varepsilon)$ . If  $t \in [\nu, T - \nu]$ , the claim follows directly from Theorem 2.7. For  $t \in [0, \nu) \cup (T - \nu, T]$ , we use Theorem 2.7 along with the facts that (i) the completed graph of a discontinuous function contains the line segments joining the values of the function at the points of discontinuity, (ii) the increments of  $\hat{X}^0$  are small in the sense of Lemma 3.9, and (iii)  $\lim_{\eta \rightarrow 0} \sup_{N \in \mathcal{N}} \mathbb{E}[\hat{X}_T^{\eta, N}] = 0$  by Lemma 3.11. For instance, consider an  $\omega \in \Omega$  with

$$\sup_{u \in [0, T - \nu]} (\hat{X}_u^0(\omega) - \hat{X}_u^{\eta, N}(\omega)) \leq \nu$$

and assume that  $(t, x) \in G_{\hat{X}^{\eta, N}(\omega)}$  with  $t < \nu$  and  $x < \hat{X}_0^0(\omega) - \nu$ . Since  $x = \hat{X}_t^{\eta, N}(\omega)$ , the mean value theorem yields an  $s \in [0, t]$  such that  $\hat{X}_s^0(\omega) = x + \nu$ , which proves that  $(s, x + \nu) \in G_{\hat{X}^0(\omega)}$  and

$$d_\infty((t, x), (s, x + \nu)) = (t - s) + \nu \leq 2\nu. \quad \square$$

## 4 Proofs for Sect. 2.2

### 4.1 Proof of Lemma 2.11

In order to prove Lemma 2.11, we first define, for all  $\beta, \nu \in (0, \infty)$  and  $x \in \mathbb{R}$ ,

$$f^{\beta, \nu}(x) := \begin{cases} -\beta/\nu, & x \leq -\beta, \\ x/\nu, & -\beta \leq x \leq \beta, \\ \beta/\nu, & x \geq \beta. \end{cases}$$

For any admissible strategy  $(j^+, j^-, V) \in \mathcal{A}^0$ , there exists a unique pathwise differentiable, adapted stochastic process  $(\tilde{V}_t^{\beta, \nu})_{t \in [0, T]}$  such that  $\tilde{V}_0^{\beta, \nu} = 0$  and its time derivative satisfies

$$\dot{V}_t^{\beta, \nu}(\omega) = f^{\beta, \nu}(V_t(\omega) - \tilde{V}_t^{\beta, \nu}(\omega)).$$

Now (2.13) can easily be verified by the comparison principle: For all  $\varepsilon > 0$  and all  $t \in [0, T]$ , we have

$$f^{\beta, \nu}\left(V_t - \left(\max_{s \in [0, T]} |V_s| + \varepsilon\right)\right) < 0$$

and hence, for all  $\varepsilon > 0$  and  $t \in [0, T]$ ,

$$\tilde{V}_t^{\beta, \nu} \leq \max_{s \in [0, T]} |V_s| + \varepsilon.$$

Analogously, we can prove that  $\tilde{V}_t^{\beta, \nu} \geq -\max_{s \in [0, T]} |V_s| - \varepsilon$ .

Now we are going to prove (2.14). For  $\beta, \nu \in (0, \infty)$ , let

$$M := \left\{ \sup_{s, t \in [0, T], |s-t| \leq \nu} |V_s - V_t| \leq \beta \right\}.$$

In view of Lemma 3.9, it is enough to prove that

$$|V_t(\omega) - \tilde{V}_t^{\beta, \nu}(\omega)| < 3\beta$$

for all  $\omega \in M$  and  $t \in [0, T]$ . In order to see this, let us assume to the contrary that the statement is wrong. Then by continuity, since  $V_0 - \tilde{V}_0^{\beta, \nu} = 0$ , there exist some  $\omega \in M$  and  $t_1 \in [0, T]$  such that

$$|V_{t_1}(\omega) - \tilde{V}_{t_1}^{\beta, \nu}(\omega)| = 3\beta.$$

We choose the smallest such  $t_1 \in [0, T]$ . Then for each  $t \in [(t_1 - \nu)^+, t_1]$ ,

$$\begin{aligned} |V_t(\omega) - \tilde{V}_t^{\beta, \nu}(\omega)| &\geq |V_{t_1}(\omega) - \tilde{V}_{t_1}^{\beta, \nu}(\omega)| - |V_{t_1}(\omega) - V_t(\omega)| \\ &\quad - |\tilde{V}_{t_1}^{\beta, \nu}(\omega) - \tilde{V}_t^{\beta, \nu}(\omega)| \\ &\geq 3\beta - \beta - (t_1 - t) \frac{\beta}{\nu} \geq \beta. \end{aligned}$$

Since  $V_0 - \tilde{V}_0^{\beta, \nu} = 0$ , we have  $t_1 > \nu$ . Since  $V_t(\omega) - \tilde{V}_t^{\beta, \nu}(\omega)$  is continuous and has no roots in  $[t_1 - \nu, t_1]$ , it does not change sign on this interval. We may hence without loss of generality assume that  $V_{t_1}(\omega) - \tilde{V}_{t_1}^{\beta, \nu}(\omega) = 3\beta$ . This implies that

$$V_t(\omega) - \tilde{V}_t^{\beta, \nu}(\omega) \geq \beta$$

for all  $t \in [t_1 - \nu, t_1]$ . By definition, for all those  $t$ , we have  $\dot{V}_t^{\beta, \nu}(\omega) = \beta/\nu$ . This, however, contradicts the minimality of  $t_1$  as  $\dot{V}_s^{\beta, \nu} = \beta/\nu$  and

$$\begin{aligned} V_{t_1-\nu}(\omega) - \tilde{V}_{t_1-\nu}^{\beta, \nu}(\omega) &= V_{t_1-\nu}(\omega) - V_{t_1}(\omega) + V_{t_1}(\omega) - \tilde{V}_{t_1}^{\beta, \nu}(\omega) \\ &\quad + \tilde{V}_{t_1}^{\beta, \nu}(\omega) - \tilde{V}_{t_1-\nu}^{\beta, \nu}(\omega) \\ &\geq -\beta + 3\beta + \int_{t_1-\nu}^{t_1} \dot{V}_s^{\beta, \nu} \, ds = 3\beta. \end{aligned}$$

This finishes the proof of Lemma 2.11.  $\square$

## 4.2 Proof of Lemma 2.12

The proof of Lemma 2.12 requires the following result on the jump processes.

**Lemma 4.1** *Let  $(j_t)_{t \in [0, T]}$  be a  $[0, \infty)$ -valued, nondecreasing and progressively measurable stochastic process with  $\mathbb{E}[(j_T)^2] < \infty$ . Then (putting  $j_t = 0$  for  $t < 0$ )*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \int_0^T (j_t - j_{t-\varepsilon})^2 \, dt \right] = 0.$$

**Proof** First we want to bound  $\int_0^T (j_t - j_{t-\varepsilon})^2 dt$  through a term that depends on  $\int_0^T (j_t - j_{t-2\varepsilon})^2 dt$  and apply this bound inductively. For every  $\varepsilon > 0$ , we have

$$\begin{aligned} \int_0^T (j_t - j_{t-2\varepsilon})^2 dt &= \int_0^T (j_t - j_{t-\varepsilon} + j_{t-\varepsilon} - j_{t-2\varepsilon})^2 dt \\ &\geq \int_0^T (j_t - j_{t-\varepsilon})^2 dt + \int_0^{T-\varepsilon} (j_t - j_{t-\varepsilon})^2 dt \\ &= 2 \int_0^T (j_t - j_{t-\varepsilon})^2 dt - \int_{T-\varepsilon}^T (j_t - j_{t-\varepsilon})^2 dt. \end{aligned}$$

Moreover,  $j_t \leq j_T$  and  $j_{t-\varepsilon} \geq j_{T-\varepsilon-\varepsilon}$  for  $t \in [T - \varepsilon, T]$  give

$$\int_{T-\varepsilon}^T (j_t - j_{t-\varepsilon})^2 dt \leq \int_{T-\varepsilon}^T (j_T - j_{T-2\varepsilon})^2 dt \leq \varepsilon (j_T - j_{T-2\varepsilon})^2.$$

As a result,

$$\frac{1}{\varepsilon} \int_0^T (j_t - j_{t-\varepsilon})^2 dt \leq \frac{1}{2\varepsilon} \int_0^T (j_t - j_{t-2\varepsilon})^2 dt + \frac{1}{2} (j_T - j_{T-2\varepsilon})^2.$$

Hence, inductively, for all  $\varepsilon \in (0, T)$  and  $k \in \{1, 2, \dots\}$ , we have

$$\begin{aligned} \frac{1}{\varepsilon} \int_0^T (j_t - j_{t-\varepsilon})^2 dt &\leq \frac{1}{2\varepsilon} \int_0^T (j_t - j_{t-2\varepsilon})^2 dt + \frac{1}{2} (j_T - j_{T-2\varepsilon})^2 \\ &\leq \frac{1}{4\varepsilon} \int_0^T (j_t - j_{t-4\varepsilon})^2 dt + \frac{1}{4} (j_T - j_{T-4\varepsilon})^2 \\ &\quad + \frac{1}{2} (j_T - j_{T-2\varepsilon})^2 \\ &\leq \dots \\ &\leq \frac{1}{2^k \varepsilon} \int_0^T (j_t - j_{t-2^k \varepsilon})^2 dt + \sum_{\ell=1}^k \frac{1}{2^\ell} (j_T)^2 \\ &\leq \frac{1}{2^k \varepsilon} \int_0^T (j_t - j_{t-2^k \varepsilon})^2 dt + (j_T)^2. \end{aligned}$$

Hence for all  $k \in \mathbb{N}$ ,

$$\mathbb{E} \left[ \int_0^T (j_t - j_{t-\varepsilon})^2 dt \right] \leq \left( \frac{T}{2^k} + \varepsilon \right) \mathbb{E}[(j_T)^2].$$

This shows the desired result.  $\square$

We are now ready to prove the approximation of arbitrary portfolio processes by absolutely continuous ones.



**Proof** of Lemma 2.12 By the triangle inequality, for all  $\beta, \nu, \varepsilon > 0$ ,

$$\begin{aligned}
 & \sqrt{\mathbb{E} \left[ \int_0^T (X_t^{\theta, \beta, \nu, \varepsilon} - X_t^\theta)^2 dt \right]} \\
 & \leq \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_{(T-\varepsilon) \wedge t}^t \dot{V}_s^{\beta, \nu} ds \right)^2 dt \right]} \\
 & \quad + \sum_{* \in \{+, -\}} \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_0^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) ds - j_t^* \right)^2 dt \right]} \\
 & \quad + \sqrt{\mathbb{E} \left[ \int_0^T \left( \frac{1}{\varepsilon} (t - (T - \varepsilon))^+ X_{T-\varepsilon}^{\theta, \beta, \nu, \varepsilon} \right)^2 dt \right]} \\
 & \quad + \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_0^t \dot{V}_s^{\beta, \nu} ds - V_t \right)^2 dt \right]}. \tag{4.1}
 \end{aligned}$$

We analyse the four terms separately. The first can be bounded by

$$\begin{aligned}
 \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_{(T-\varepsilon) \wedge t}^t \dot{V}_s^{\beta, \nu} ds \right)^2 dt \right]} & \leq \sqrt{\mathbb{E} \left[ \int_{T-\varepsilon}^T \left( \int_{T-\varepsilon}^t \dot{V}_s^{\beta, \nu} ds \right)^2 dt \right]} \\
 & \leq \frac{\varepsilon^{\frac{3}{2}} \beta}{\nu} \longrightarrow 0 \quad \text{for } \varepsilon \rightarrow 0.
 \end{aligned}$$

Regarding the second term, for all  $\varepsilon < T/2$ ,

$$\int_0^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) ds - j_t^* = - \int_{((T-\varepsilon) \wedge t) - \varepsilon}^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_t^* - j_s^*) ds.$$

Using the monotonicity of the jump processes, it follows from Lemma 4.1 that

$$\begin{aligned}
 & \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_0^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) \, ds - j_t^* \right)^2 \, dt \right]} \\
 &= \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_{((T-\varepsilon) \wedge t) - \varepsilon}^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_t^* - j_s^*) \, ds \right)^2 \, dt \right]} \\
 &\leq \frac{1}{\varepsilon} \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_{((T-\varepsilon) \wedge t) - \varepsilon}^{(T-\varepsilon) \wedge t} (j_t^* - j_s^*) \, ds \right)^2 \, dt \right]} \\
 &= \sqrt{\mathbb{E} \left[ \int_0^{T-\varepsilon} (j_t^* - j_{t-\varepsilon}^*)^2 \, dt \right]} + \mathbb{E} \left[ \int_{T-\varepsilon}^T (j_t^* - j_{T-2\varepsilon}^*)^2 \, dt \right] \\
 &\leq \sqrt{\mathbb{E} \left[ \int_0^T (j_t^* - j_{t-\varepsilon}^*)^2 \, dt \right]} + \varepsilon \mathbb{E}[(j_T^*)^2] \longrightarrow 0 \quad \text{for } \varepsilon \rightarrow 0. \quad (4.2)
 \end{aligned}$$

For the third term, we conclude from the Itô isometry and the definition of  $\xi^{\theta, \beta, v, \varepsilon}$  that

$$\begin{aligned}
 & \sqrt{\mathbb{E} \left[ \int_0^T \left( \frac{1}{\varepsilon} (t - (T - \varepsilon))^+ X_{T-\varepsilon}^{\theta, \beta, v, \varepsilon} \right)^2 \, dt \right]} \\
 &= \sqrt{\frac{\varepsilon}{3} \mathbb{E} \left[ \left( x_0 + \int_0^{T-\varepsilon} \xi_s^{\theta, \beta, v, \varepsilon} \, ds \right)^2 \right]} \\
 &\leq x_0 \sqrt{\frac{\varepsilon}{3}} + \sum_{* \in \{+, -\}} \sqrt{\frac{\varepsilon}{3} \mathbb{E} \left[ \left( \int_0^{T-\varepsilon} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) \, ds \right)^2 \right]} \\
 &\quad + \sqrt{\frac{\varepsilon}{3} \mathbb{E} \left[ \left( \int_0^{T-\varepsilon} \dot{V}_s^{\beta, v} \, ds \right)^2 \right]}.
 \end{aligned}$$

Now we can bound

$$\sqrt{\frac{\varepsilon}{3} \mathbb{E} \left[ \left( \int_0^{T-\varepsilon} \dot{V}_s^{\beta, v} \, ds \right)^2 \right]} \leq T \frac{\beta}{v} \sqrt{\frac{\varepsilon}{3}} \longrightarrow 0 \quad \text{for } \varepsilon \rightarrow 0.$$

Moreover, in view of (4.2),

$$\begin{aligned}
 & \sqrt{\frac{\varepsilon}{3} \mathbb{E} \left[ \left( \int_0^{T-\varepsilon} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) \, ds \right)^2 \right]} \\
 &= \sqrt{\frac{1}{3} \mathbb{E} \left[ \int_{T-\varepsilon}^T \left( \int_0^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) \, ds \right)^2 dt \right]} \\
 &\leq \sqrt{\frac{1}{3} \mathbb{E} \left[ \int_0^T \left( \int_0^{(T-\varepsilon) \wedge t} \frac{1}{\varepsilon} (j_s^* - j_{s-\varepsilon}^*) \, ds - j_t^* \right)^2 dt \right]} + \sqrt{\frac{1}{3} \mathbb{E} \left[ \int_{T-\varepsilon}^T (j_t^*)^2 dt \right]} \\
 &\leq \frac{1}{\sqrt{3}} \sqrt{\mathbb{E} \left[ \int_0^T (j_t^* - j_{t-\varepsilon}^*)^2 dt \right]} + \varepsilon \mathbb{E}[(j_T^*)^2] + \sqrt{\frac{\varepsilon}{3} \mathbb{E}[(j_T^*)^2]} \longrightarrow 0 \quad \text{for } \varepsilon \rightarrow 0.
 \end{aligned}$$

It remains to consider the fourth term in (4.1). To this end, let

$$M^{\beta, \nu} := \left\{ \sup_{t \in [0, T]} \left| \int_0^t \dot{V}_s^{\beta, \nu} \, ds - V_t \right| \leq 3\beta \right\}.$$

Then using (2.13) in the last step,

$$\begin{aligned}
 & \sqrt{\mathbb{E} \left[ \int_0^T \left( \int_0^t \dot{V}_s^{\beta, \nu} \, ds - V_t \right)^2 dt \right]} \\
 &\leq \sqrt{\mathbb{E} \left[ \int_0^T \mathbb{1}_{M^{\beta, \nu}} \left( \int_0^t \dot{V}_s^{\beta, \nu} \, ds - V_t \right)^2 dt \right]} + \sqrt{\mathbb{E} \left[ \mathbb{1}_{\Omega \setminus M^{\beta, \nu}} \int_0^T (V_t)^2 dt \right]} \\
 &\quad + \sqrt{\mathbb{E} \left[ \mathbb{1}_{\Omega \setminus M^{\beta, \nu}} \int_0^T \left( \int_0^t \dot{V}_s^{\beta, \nu} \, ds \right)^2 dt \right]} \\
 &\leq 3\beta\sqrt{T} + \sqrt{T \mathbb{E} \left[ \mathbb{1}_{\Omega \setminus M^{\beta, \nu}} \max_{t \in [0, T]} |V_t|^2 \right]} \\
 &\quad + \sqrt{T \mathbb{E} \left[ \mathbb{1}_{\Omega \setminus M^{\beta, \nu}} \max_{t \in [0, T]} \left( \int_0^t \dot{V}_s^{\beta, \nu} \, ds \right)^2 \right]} \\
 &\leq 3\beta\sqrt{T} + 2\sqrt{T \mathbb{E} \left[ \mathbb{1}_{\Omega \setminus M^{\beta, \nu}} \max_{t \in [0, T]} |V_t|^2 \right]}.
 \end{aligned}$$

By (2.12), the single random variable  $\max_{t \in [0, T]} |V_t|^2$  is in  $L^1$  and hence uniformly integrable. So we can first choose  $\beta > 0$  small enough, then, according to Lemma 2.11, choose  $\nu > 0$  small enough such that  $\mathbb{P}[\Omega \setminus M^{\beta, \nu}]$  is sufficiently small, and finally choose  $\varepsilon > 0$  small enough in order to obtain the desired result.  $\square$

### 4.3 Proof of Lemma 2.13

We start with a technical lemma.

**Lemma 4.2** *Let  $(M, \mathcal{A}, \mu)$  be a measure space and  $u, v \in \mathcal{L}^2(M)$ . Then*

$$\int_M |u^2 - v^2| d\mu \leq \int_M (u - v)^2 d\mu + 2\sqrt{\int_M (u - v)^2 d\mu} \sqrt{\int_M v^2 d\mu}.$$

**Proof** Due to the Hölder and triangle inequalities,

$$\begin{aligned} \int_M |u^2 - v^2| d\mu &= \sqrt{\int_M (u - v)^2 d\mu} \sqrt{\int_M (u + v)^2 d\mu} \\ &\leq \sqrt{\int_M (u - v)^2 d\mu} \left( \sqrt{\int_M (u - v)^2 d\mu} + \sqrt{\int_M (2v)^2 d\mu} \right). \quad \square \end{aligned}$$

The following technical lemma provides useful estimates for the impact process.

**Lemma 4.3** *Let  $X \in \mathcal{L}^2_{\mathcal{P}}(\Omega \times [0, T]; \mathbb{R})$ . Then the transient price impact process  $Y^X$  given by (2.1) satisfies  $Y^X \in \mathcal{L}^2_{\mathcal{P}}(\Omega \times [0, T]; \mathbb{R})$  and*

$$\sqrt{\mathbb{E}[(Y_T^X)^2]} \leq \gamma \left( X_0 + \sqrt{\mathbb{E}[(X_T)^2]} + \bar{\rho} \exp(T\bar{\rho}) \sqrt{T \mathbb{E} \left[ \int_0^T (X_s)^2 ds \right]} \right), \quad (4.3)$$

$$\sqrt{\mathbb{E} \left[ \int_0^T (Y_t^X)^2 dt \right]} \leq \gamma X_0 \sqrt{T} + \gamma (1 + T\bar{\rho} \exp(T\bar{\rho})) \sqrt{\mathbb{E} \left[ \int_0^T (X_t)^2 dt \right]}. \quad (4.4)$$

If  $X_0 = 0$ , then we additionally have for all  $s, t \in [0, T]$  with  $s < t$  that

$$\begin{aligned} \sqrt{\mathbb{E} \left[ \int_s^t (Y_u^X)^2 du \right]} &\leq \gamma \bar{\rho} \exp(T\bar{\rho}) \sqrt{(t - s) T \mathbb{E} \left[ \int_0^s (X_u)^2 du \right]} \\ &\quad + \gamma (1 + \bar{\rho} \exp(T\bar{\rho}) \sqrt{(t - s) T}) \sqrt{\mathbb{E} \left[ \int_s^t (X_u)^2 du \right]}. \quad (4.5) \end{aligned}$$

**Proof** Inequality (4.3) follows from the explicit formula

$$Y_t^X = \gamma X_t - \gamma \exp \left( - \int_0^t \rho_u du \right) \left( X_0 + \int_0^t \rho_s X_s \exp \left( \int_0^s \rho_u du \right) ds \right) \quad (4.6)$$

and the triangle inequality. Moreover,

$$\begin{aligned} & \sqrt{\mathbb{E}\left[\int_0^T (Y_t^X)^2 dt\right]} \\ & \leq \gamma \sqrt{\mathbb{E}\left[\int_0^T (X_t)^2 dt\right]} \\ & \quad + \gamma \sqrt{\mathbb{E}\left[\int_0^T \exp\left(-2 \int_0^t \rho_u du\right) \left(X_0 + \int_0^t \rho_s X_s \exp\left(\int_0^s \rho_u du\right) ds\right)^2 dt\right]}. \end{aligned} \quad (4.7)$$

Substituting the inequality

$$\begin{aligned} & \sqrt{\mathbb{E}\left[\int_0^T \exp\left(-2 \int_0^t \rho_u du\right) \left(X_0 + \int_0^t \rho_s X_s \exp\left(\int_0^s \rho_u du\right) ds\right)^2 dt\right]} \\ & \leq \sqrt{\mathbb{E}[T(X_0)^2]} + \bar{\rho} \exp(T\bar{\rho}) \sqrt{T \mathbb{E}\left[\left(\int_0^T |X_s| ds\right)^2\right]} \\ & \leq X_0 \sqrt{T} + T\bar{\rho} \exp(T\bar{\rho}) \sqrt{\mathbb{E}\left[\int_0^T (X_s)^2 ds\right]} \end{aligned}$$

into (4.7) yields (4.4). To prove (4.5), let  $X_0 = 0$ . Then due to (4.6),

$$|Y_u^X - \gamma X_u| \leq \gamma \bar{\rho} \exp(T\bar{\rho}) \int_0^u |X_r| dr$$

for all  $u \in [0, T]$  and so

$$\mathbb{E}\left[\int_s^t (Y_u^X - \gamma X_u)^2 du\right] \leq (t-s)T\gamma^2\bar{\rho}^2 \exp(2T\bar{\rho}) \mathbb{E}\left[\int_0^t (X_u)^2 du\right].$$

Using the subadditivity of the square root, we now obtain (4.5) from

$$\begin{aligned} & \sqrt{\mathbb{E}\left[\int_s^t (Y_u^X)^2 du\right]} \\ & \leq \sqrt{\mathbb{E}\left[\int_s^t (Y_u^X - \gamma X_u)^2 du\right]} + \sqrt{\mathbb{E}\left[\int_s^t (\gamma X_u)^2 du\right]} \\ & \leq \gamma \bar{\rho} \exp(T\bar{\rho}) \sqrt{(t-s)T \mathbb{E}\left[\int_0^s (X_u)^2 du\right]} \\ & \quad + \gamma \bar{\rho} \exp(T\bar{\rho}) \sqrt{(t-s)T \mathbb{E}\left[\int_s^t (X_u)^2 du\right]} + \gamma \sqrt{\mathbb{E}\left[\int_s^t (X_u)^2 du\right]}. \quad \square \end{aligned}$$

We are now ready to prove our approximation result for the cost functional.

**Proof** of Lemma 2.13 For  $\theta = (j^+, j^-, V) \in \mathcal{A}^0$  and  $\xi \in \mathcal{A}$ ,

$$\begin{aligned} & |J^0(\theta) - J^0(0, 0, V^\xi)| \\ & \leq \left| \mathbb{E} \left[ \frac{\gamma}{2} (j_0^+ - j_0^-)^2 + \int_{(0,T]} \frac{1}{2} (Y_{t-}^\theta + Y_t^\theta) dX_t^\theta - \int_0^T Y_t^\xi dX_t^\xi + \frac{\gamma}{2} [V]_T \right] \right| \\ & \quad + \frac{\bar{\lambda}}{2} \mathbb{E} \left[ \int_0^T |(X_t^\theta)^2 - (X_t^\xi)^2| dt \right]. \end{aligned}$$

Due to Lemma 4.2, the last term can be estimated as

$$\begin{aligned} & \mathbb{E} \left[ \int_0^T |(X_t^\theta)^2 - (X_t^\xi)^2| dt \right] \\ & \leq \mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right] + 2\sqrt{\mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right] \mathbb{E} \left[ \int_0^T (X_t^\theta)^2 dt \right]}. \end{aligned}$$

Moreover, using first Lemma 3.10 and then Lemmas 4.3 and 4.2, we obtain

$$\begin{aligned} & \left| \mathbb{E} \left[ \frac{\gamma}{2} (j_0^+ - j_0^-)^2 + \int_{(0,T]} \frac{1}{2} (Y_{t-}^\theta + Y_t^\theta) dX_t^\theta - \int_0^T Y_t^\xi dX_t^\xi + \frac{\gamma}{2} [V]_T \right] \right| \\ & = \left| \mathbb{E} \left[ \frac{\gamma}{2} (j_0^+ - j_0^-)^2 + \frac{1}{2\gamma} (Y_T^\theta)^2 - \frac{\gamma}{2} (j_0^+ - j_0^-)^2 - \frac{\gamma}{2} [V]_T + \frac{1}{\gamma} \int_0^T \rho_t (Y_t^\theta)^2 dt \right. \right. \\ & \quad \left. \left. - \frac{1}{2\gamma} (Y_T^\xi)^2 - \frac{1}{\gamma} \int_0^T \rho_t (Y_t^\xi)^2 dt + \frac{\gamma}{2} [V]_T \right] \right| \\ & \leq \frac{1}{2\gamma} \mathbb{E} [| (Y_T^\theta)^2 - (Y_T^\xi)^2 |] + \frac{\bar{\rho}}{\gamma} \mathbb{E} \left[ \int_0^T |(Y_t^\theta)^2 - (Y_t^\xi)^2| dt \right] \\ & \leq \left( \frac{1}{2} T \gamma \bar{\rho}^2 e^{2T\bar{\rho}} + \bar{\rho} \gamma (1 + T \bar{\rho} e^{T\bar{\rho}})^2 \right) \mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right] \\ & \quad + \sqrt{T \gamma \bar{\rho} e^{T\bar{\rho}}} \sqrt{\mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right]} \left( x_0 + \bar{\rho} e^{T\bar{\rho}} \sqrt{T \mathbb{E} \left[ \int_0^T (X_t^\theta)^2 dt \right]} \right) \\ & \quad + 2\bar{\rho} (1 + T \bar{\rho} e^{T\bar{\rho}}) \sqrt{\mathbb{E} \left[ \int_0^T (X_t^\theta - X_t^\xi)^2 dt \right]} \\ & \quad \times \left( \gamma x_0 \sqrt{T} + \gamma (1 + T \bar{\rho} e^{T\bar{\rho}}) \sqrt{\mathbb{E} \left[ \int_0^T (X_t^\theta)^2 dt \right]} \right). \quad \square \end{aligned}$$

#### 4.4 Proof of Lemma 2.14

We assume the contrary, i.e., that there exists a constant  $c > 0$  such that for all  $H > 0$ , there exists some  $\eta \in (0, H)$  with

$$\frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\hat{\xi}_t^{\eta, \infty})^2 dt \right] > c. \quad (4.8)$$

The optimality of  $\hat{\xi}^{\eta, \infty}$  and (2.15) imply that for all  $\eta, \nu, \beta, \varepsilon > 0$ ,

$$\begin{aligned} 0 &\geq J^{\eta, \infty}(\hat{\xi}^{\eta, \infty}) - J^{\eta, \infty}(\xi^{\hat{\theta}, \beta, \nu, \varepsilon}) \\ &= \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\hat{\xi}_t^{\eta, \infty})^2 dt \right] - \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\xi_t^{\hat{\theta}, \beta, \nu, \varepsilon})^2 dt \right] \\ &\quad + J^0(0, 0, V^{\hat{\xi}^{\eta, \infty}}) - J^0(\hat{\theta}) + J^0(\hat{\theta}) - J^0(0, 0, V^{\xi^{\hat{\theta}, \beta, \nu, \varepsilon}}). \end{aligned} \quad (4.9)$$

We now prove that (4.8) contradicts (4.9). By Theorem 2.13 and since  $|\hat{X}_t^0| \leq x_0$ , we obtain (for convenience, let  $p(x) := x + \sqrt{x}$ )

$$|J^0(0, 0, V^{\hat{\xi}^{\eta, \infty}}) - J^0(\hat{\theta})| \leq D(T(x_0)^2) p \left( \mathbb{E} \left[ \int_0^T (\hat{X}_t^{\eta, \infty} - \hat{X}_t^0)^2 dt \right] \right)$$

and

$$|J^0(\hat{\theta}) - J^0(0, 0, V^{\xi^{\hat{\theta}, \beta, \nu, \varepsilon}})| \leq D(T(x_0)^2) p \left( \mathbb{E} \left[ \int_0^T (X_t^{\hat{\theta}} - X_t^{\hat{\theta}, \beta, \nu, \varepsilon})^2 dt \right] \right).$$

Plugging the results into (4.9) yields that for all  $\eta > 0$  that satisfy (4.8), we have

$$\begin{aligned} 0 &> c - \frac{\eta}{2} \mathbb{E} \left[ \int_0^T (\xi_t^{\hat{\theta}, \beta, \nu, \varepsilon})^2 dt \right] \\ &\quad - D(T(x_0)^2) \left( p \left( \mathbb{E} \left[ \int_0^T (\hat{X}_t^{\eta, \infty} - \hat{X}_t^0)^2 dt \right] \right) \right. \\ &\quad \left. + p \left( \mathbb{E} \left[ \int_0^T (X_t^{\hat{\theta}} - X_t^{\hat{\theta}, \beta, \nu, \varepsilon})^2 dt \right] \right) \right). \end{aligned} \quad (4.10)$$

Due to Lemma 2.12, we can first choose  $\beta, \nu, \varepsilon > 0$  sufficiently small to get

$$D(T(x_0)^2) p \left( \mathbb{E} \left[ \int_0^T (X_t^{\hat{\theta}} - X_t^{\hat{\theta}, \beta, \nu, \varepsilon})^2 dt \right] \right) < \frac{c}{2}.$$

Since  $|\hat{X}_t^{\eta, \infty}| \leq x_0$  (cf. Theorem 2.2) and in view of Theorem 2.7, we can then choose  $\eta > 0$  sufficiently small and satisfying (4.8) such that the right-hand side of (4.10) is larger than zero, which is a contradiction. This finishes the proof of Lemma 2.14.

## Appendix A: Regularity properties of Itô processes and BSDEs

In this appendix, we introduce some regularity properties of stochastic processes, which we use to prove various convergence results for stochastic processes.

We consider a continuous, adapted,  $\mathbb{R}^d$ -valued stochastic process  $Y = (Y_t)_{t \in I}$  on some interval  $I$  (i.e., a convex subset of  $\mathbb{R}$ ) and introduce the following continuity conditions.

**Condition A.1** For all  $\varepsilon \in (0, \infty)$ , there exists  $\delta \in (0, \infty)$  such that for all  $s \in I$ ,  $\mathcal{F}_s$ -measurable and integrable  $V: \Omega \rightarrow \mathbb{R}$  and stopping times  $\tau: \Omega \rightarrow [s, s + \delta] \cap I$ ,

$$|\mathbb{E}[V(Y_\tau - Y_s)]| \leq \varepsilon \mathbb{E}[|V|].$$

**Condition A.2** For all  $\varepsilon \in (0, \infty)$ , there exists  $\delta \in (0, \infty)$  such that for all  $s \in I$ ,

$$\mathbb{E} \left[ \sup_{t \in [s, s + \delta] \cap I} |Y_t - Y_s| \middle| \mathcal{F}_s \right] \leq \varepsilon.$$

**Definition A.3** A family of stochastic processes is said to *uniformly* satisfy Condition A.1 or A.2 on  $I$  if all processes satisfy the respective property and  $\delta$  can be chosen uniformly for all processes.

In what follows, we list some auxiliary results.

**Lemma A.4** An  $\mathbb{R}^d$ -valued stochastic process  $Y = ((Y_t^1, \dots, Y_t^d))_{t \in I}$  satisfies Condition A.1 (resp. Condition A.2) if and only if all components  $Y^i$  satisfy Condition A.1 (resp. Condition A.2).

**Lemma A.5** Condition A.2 implies Condition A.1.

**Proof** Due to Lemma A.4, it is enough to consider a real-valued  $Y$ . To show the assertion, we then calculate

$$|\mathbb{E}[V(Y_\tau - Y_s)]| = |\mathbb{E}[V \mathbb{E}[Y_\tau - Y_s | \mathcal{F}_s]]| \leq \mathbb{E}[|V| \mathbb{E}[|Y_\tau - Y_s| | \mathcal{F}_s]] \leq \mathbb{E}[|V|] \varepsilon. \quad \square$$

**Remark A.6** If we want to bound a term of the form  $|\mathbb{E}[\mathbb{1}_N(Y_\tau - Y_s)]|$  for some  $N \subseteq M \in \mathcal{F}_s$  with  $N \notin \mathcal{F}_s$ , then Condition A.1 is not enough; in this case, we need the stronger Condition A.2.

**Lemma A.7** Let  $X = (X_t)_{t \in I}$  and  $Y = (Y_t)_{t \in I}$  be continuous, adapted, real-valued processes which satisfy Condition A.1 (resp. Condition A.2). Then  $X + Y$  satisfies Condition A.1 (resp. Condition A.2).

**Theorem A.8** Let  $X = (X_t)_{t \in I}$  and  $Y = (Y_t)_{t \in I}$  be continuous, adapted, essentially bounded, real-valued processes. Then:

- If  $X$  satisfies Condition A.2 and  $Y$  satisfies Condition A.1, then  $XY$  satisfies Condition A.1.
- If  $X$  and  $Y$  both satisfy Condition A.2, then  $XY$  also satisfies Condition A.2.



**Proof** The second part of the statement can be proved directly via

$$|X_t Y_t - X_s Y_s| \leq |X_t| |Y_t - Y_s| + |Y_s| |X_t - X_s|.$$

The first part of the statement follows from

$$|\mathbb{E}[V(X_\tau Y_\tau - X_s Y_s)]| \leq |\mathbb{E}[V Y_\tau (X_\tau - X_s)]| + |\mathbb{E}[V X_s (Y_\tau - Y_s)]|. \quad \square$$

Next, we prove some properties of the concave envelope of the modulus of continuity of a uniformly continuous function. We use these results to show that a uniformly continuous function of an Itô process with bounded coefficients satisfies Condition A.2.

**Lemma A.9** *Let  $(X, \|\cdot\|_X)$  be a normed space with a nonempty and convex subset  $D$ , let  $(Y, d_Y)$  be a metric space, let  $f: D \rightarrow Y$  and for every  $t \in [0, \infty)$ , let*

$$\tilde{\omega}_f(t) := \sup_{x_1, x_2 \in D, \|x_1 - x_2\|_X \leq t} d_Y(f(x_1), f(x_2)),$$

$$\omega_f(t) := \inf\{\psi(t) : \psi: [0, \infty) \rightarrow [0, \infty) \text{ is concave and } \psi \geq \tilde{\omega}_f\}.$$

*Then  $\tilde{\omega}_f: [0, \infty) \rightarrow [0, \infty]$  is nondecreasing and satisfies  $\tilde{\omega}_f(0) = 0$ , and for all  $s, t \in [0, \infty)$  with  $s > 0$ , we have*

$$\tilde{\omega}_f(t) \leq \frac{t+s}{s} \tilde{\omega}_f(s). \quad (\text{A.1})$$

*If  $\tilde{\omega}_f$  is finite-valued, then  $\omega_f$  takes values in  $[0, \infty)$ , is concave, nondecreasing and for all  $x_1, x_2 \in D$ ,*

$$d_Y(f(x_1), f(x_2)) \leq \omega_f(\|x_1 - x_2\|_X). \quad (\text{A.2})$$

*If  $f$  is uniformly continuous, then  $\tilde{\omega}_f$  is finite-valued,  $\tilde{\omega}_f$  and  $\omega_f$  are continuous in 0 and  $\omega_f(0) = 0$ .*

**Proof** By definition,  $\tilde{\omega}_f$  is nondecreasing and satisfies  $\tilde{\omega}_f(0) = 0$ . To prove (A.1), let  $s, t \in [0, \infty)$  with  $s > 0$  and  $x_1, x_2 \in D$  with  $\|x_1 - x_2\|_X \leq t$ . Let  $N := \lceil t/s \rceil$  and

$$x^k := \frac{k}{N} x_1 + \frac{N-k}{N} x_2 \quad \text{for } k = 0, \dots, N.$$

Then  $x^k \in D$  because  $D$  is convex and  $\|x^k - x^{k+1}\|_X = \frac{1}{N} \|x_1 - x_2\|_X \leq s$ . Hence

$$\begin{aligned} d_Y(f(x_1), f(x_2)) &\leq \sum_{k=0}^{N-1} d_Y(f(x^k), f(x^{k+1})) \\ &\leq N \max_k \tilde{\omega}_f(\|x^k - x^{k+1}\|_X) \leq (t+s) \frac{\tilde{\omega}_f(s)}{s}. \end{aligned}$$

Having established (A.1), we know that if  $\tilde{\omega}_f$  is finite-valued, then it is bounded above by an affine (hence concave) function. In particular, the concave envelope is well defined and satisfies (A.2). Concavity and monotonicity of  $\omega_f$  are clear. Finally, uniform continuity of  $f$  implies that  $\tilde{\omega}_f$  is finite-valued and continuous in 0. In order to see  $\omega_f(0) = 0$  and continuity of  $\omega_f$  in 0, let  $\varepsilon > 0$  be arbitrary and let  $\delta > 0$  be such that  $\tilde{\omega}_f(\delta) \leq \varepsilon/2$ . Due to (A.1),  $\tilde{\omega}_f$  and hence  $\omega_f$  is bounded above by the affine and hence concave function  $t \mapsto ((t + \delta)/\delta)(\varepsilon/2)$ . In particular, for all  $t \leq \delta$ , we obtain  $\omega_f(t) \leq \varepsilon$ .  $\square$

**Lemma A.10** *Let  $X = (X_t)_{t \in I}$  be a continuous, adapted,  $\mathbb{R}^d$ -valued process which satisfies Condition A.2 and let  $f: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  be uniformly continuous. Then  $(f(t, X_t))_{t \in [0, T]}$  also satisfies Condition A.2.*

**Proof** There exists a function  $\omega_f$  with the properties listed by Lemma A.9 (with the max-norm  $|\cdot|_\infty$  on  $[0, T] \times \mathbb{R}^d$ ). Now let  $\delta > 0$  be arbitrary. Using (A.2) and the concavity and monotonicity of  $\omega_f$ , we obtain that

$$\begin{aligned} & \mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} |f(t, X_t) - f(s, X_s)| \middle| \mathcal{F}_s \right] \\ & \leq \mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} \omega_f(|t - s| \vee |X_t - X_s|_\infty) \middle| \mathcal{F}_s \right] \\ & \leq \mathbb{E} \left[ \omega_f \left( \delta \vee \max_{s \leq t \leq T \wedge (s+\delta)} |X_t - X_s|_\infty \right) \middle| \mathcal{F}_s \right] \\ & \leq \omega_f \left( \delta + \mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} |X_t - X_s|_\infty \middle| \mathcal{F}_s \right] \right). \end{aligned}$$

Since  $\omega_f$  is continuous in 0 and  $X$  satisfies Condition A.2, we can choose  $\delta > 0$  small enough such that this term is not greater than  $\varepsilon$ .  $\square$

**Lemma A.11** *Let  $(\tilde{\sigma}_t)_{t \in [0, T]} \in \mathcal{L}^\infty_{\mathcal{P}}(\Omega \times [0, T]; \mathbb{R}^m)$ . Then  $(\int_0^t \tilde{\sigma}_u dW_u)_{t \in [0, T]}$  satisfies Condition A.2.*

**Proof** Let  $\bar{\sigma} > 0$  be a componentwise bound on  $\tilde{\sigma}$ . By Jensen's inequality,

$$\mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} \left| \int_s^t \tilde{\sigma}_u dW_u \right| \middle| \mathcal{F}_s \right] \leq \sqrt{\mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} \left| \int_s^t \tilde{\sigma}_u dW_u \right|^2 \middle| \mathcal{F}_s \right]}.$$

In order to prove that the conditional expectation on the right-hand side is bounded, we apply the classical Doob maximal inequality for the conditional measures with respect to all sets  $A \in \mathcal{F}_s$  with positive probability and obtain

$$\sqrt{\mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} \left| \int_s^t \tilde{\sigma}_u dW_u \right|^2 \middle| \mathcal{F}_s \right]} \leq 2 \sqrt{\mathbb{E} \left[ \left| \int_s^{T \wedge (s+\delta)} \tilde{\sigma}_u dW_u \right|^2 \middle| \mathcal{F}_s \right]}.$$

Then by combining these inequalities and using the Itô isometry for conditional expectations, we finally obtain

$$\mathbb{E} \left[ \max_{s \leq t \leq T \wedge (s+\delta)} \left| \int_s^t \tilde{\sigma}_u dW_u \right| \middle| \mathcal{F}_s \right] \leq 2\sqrt{m \mathbb{E} \left[ \int_s^{T \wedge (s+\delta)} m \bar{\sigma}^2 du \middle| \mathcal{F}_s \right]} \leq 2m\bar{\sigma}\sqrt{\delta}.$$

□

**Remark A.12** Obviously, we have the same result for the drift part of an Itô process, i.e., if  $\tilde{\mu} \in \mathcal{L}_{\text{Prog}}^\infty(\Omega \times [0, T]; \mathbb{R}^m)$ , then  $(\int_0^t \tilde{\mu}_u du)_{t \in [0, T]}$  satisfies Condition A.2. However, we cannot expect the weaker Condition A.1 under weaker integrability assumptions, as the following example shows. Take  $T = 2$ ,  $\tilde{\mu}_u := \mathbb{1}_{[1, 2]}(u)|W_1|$ ,  $\varepsilon := 1$ ,  $\delta \in (0, 1)$  arbitrary and  $s := 1$ ,  $\tau := 1 + \delta$ ,  $V := \mathbb{1}_{\{\delta|W_1| > 1\}}$ . Then

$$\frac{|\mathbb{E}[V \int_1^{1+\delta} \tilde{\mu}_u du]|}{\mathbb{E}[|V|]} = \frac{\mathbb{E}[\mathbb{1}_{\{\delta|W_1| > 1\}} \delta |W_1|]}{\mathbb{E}[\mathbb{1}_{\{\delta|W_1| > 1\}}]} > 1.$$

Finally, we prove that the strong Condition A.2 holds for a certain class of BSDEs driven by forward SDEs. Specifically, we prove that the solution to the BSDE can be expressed as a *uniformly continuous* function of the forward process and then we apply Lemma A.10. The representation of the solution in terms of a *continuous* function has already been proved by El Karoui [12].

For all  $(t, x) \in [0, T] \times \mathbb{R}^n$ , we consider on  $[t, T]$  the SDE

$$dX_s^{t,x} = \tilde{\mu}(s, X_s^{t,x}) ds + \tilde{\sigma}(s, X_s^{t,x}) dW_s, \quad X_t^{t,x} = x,$$

where  $\tilde{\mu}, \tilde{\sigma} : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^{n \times m}$  are measurable and bounded and satisfy the Lipschitz condition

$$|(\tilde{\mu}, \tilde{\sigma})(t, x_1) - (\tilde{\mu}, \tilde{\sigma})(t, x_2)| \leq L|x_1 - x_2|.$$

By the previous results, we obtain that  $X^{t,x}$  satisfies Condition A.2. Standard computations show that there exists a constant  $C \in (0, \infty)$  such that for all  $t \in [0, T]$  and  $x_1, x_2 \in \mathbb{R}^n$ ,

$$\mathbb{E} \left[ \sup_{s \in [t, T]} |X_s^{t,x_1} - X_s^{t,x_2}|^2 \right] \leq C|x_1 - x_2|^2. \quad (\text{A.3})$$

**Theorem A.13** Let  $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$  be bounded,  $\psi : [0, T] \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  continuous and bounded and let both functions satisfy the Lipschitz condition

$$|\Psi(x_1) - \Psi(x_2)| + |\psi(t, x_1, y_1) - \psi(t, x_2, y_2)| \leq L(|x_1 - x_2| + |y_1 - y_2|).$$

For  $(t, x) \in [0, T] \times \mathbb{R}^n$ , let  $(Y^{t,x}, Z^{t,x}) \in \mathcal{L}_{\mathcal{P}}^2(\Omega \times [0, T]; \mathbb{R} \times \mathbb{R}^m)$  be the unique solution to the BSDE on  $[t, T]$  given by

$$-dY_s^{t,x} = \psi(s, X_s^{t,x}, Y_s^{t,x}) ds - Z_s^{t,x} dW_s, \quad Y_T^{t,x} = \Psi(X_T^{t,x})$$

driven by the forward process  $X^{t,x}$  ( $Y^{t,x}$  exists due to [12, Theorem 2.1]). Then there exists a uniformly continuous function  $h: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ , which does not depend on  $(t, x)$ , such that

$$\mathbb{P}[h(s, X_s^{t,x}) = Y_s^{t,x} \text{ for all } s \in [t, T]] = 1. \quad (\text{A.4})$$

In particular, the process  $Y^{t,x}$  satisfies Condition A.2 on  $[t, T]$ .

**Proof** By [12, Theorem 3.4] and the remark following it, there exists a decoupling field  $h: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  that is  $1/2$ -Hölder continuous in  $t$  and locally Lipschitz-continuous in  $x$  such that (A.4) holds. It therefore remains to prove the uniform continuity of  $h$ . Since

$$h(t, x) = \mathbb{E} \left[ \Psi(X_T^{t,x}) + \int_t^T \psi(u, X_u^{t,x}, h(u, X_u^{t,x})) du \right],$$

we see that  $h$  is bounded, say  $|h| \leq \bar{h}$ . Then

$$\tilde{\omega}_h(s, \delta) := \sup_{|x_1 - x_2| \leq \delta} |h(s, x_1) - h(s, x_2)|$$

is finite-valued, and Lemma A.9 therefore allows us to define its concave envelope  $\omega_h(s, \cdot): [0, \infty) \rightarrow [0, \infty)$ .

Using the Lipschitz-continuity of  $\psi$ , we obtain for all  $x_1, x_2 \in \mathbb{R}^n$  and  $t \in [0, T]$  that

$$\begin{aligned} & |h(t, x_1) - h(t, x_2)| \\ & \leq L \mathbb{E}[|X_T^{t,x_1} - X_T^{t,x_2}|] \\ & \quad + L \int_t^T (\mathbb{E}[|X_u^{t,x_1} - X_u^{t,x_2}|] + \mathbb{E}[\tilde{\omega}_h(u, |X_u^{t,x_1} - X_u^{t,x_2}|)]) du. \end{aligned}$$

Using (A.3) along with Lemma A.9, if  $|x_1 - x_2| \leq \delta$ , then

$$\mathbb{E}[|X_T^{t,x_1} - X_T^{t,x_2}|] \leq C\delta$$

and

$$\mathbb{E}[\tilde{\omega}_h(u, |X_u^{t,x_1} - X_u^{t,x_2}|)] \leq \omega_h(u, \mathbb{E}[|X_u^{t,x_1} - X_u^{t,x_2}|]) \leq \omega_h(u, C\delta)$$

for some constant  $C \in (0, \infty)$  that depends only on  $\tilde{\mu}, \tilde{\sigma}$  and  $L$ . In view of (A.1),

$$\omega_h(u, C\delta) \leq (C + 1)\tilde{\omega}_h(u, \delta).$$

This shows that

$$\begin{aligned} \tilde{\omega}_h(t, \delta) & \leq LC\delta + L \int_t^T (C\delta + (C + 1)\tilde{\omega}_h(u, \delta)) du \\ & = LC\delta(1 + T - t) + L(C + 1) \int_t^T \tilde{\omega}_h(u, \delta) du. \end{aligned}$$

Hence Gronwall's inequality implies that  $h$  is uniformly Lipschitz-continuous in the second variable. Since  $h$  is uniformly  $1/2$ -Hölder-continuous in the first variable, this proves the assertion.  $\square$

## Appendix B: Convergence results for SDEs and random ODEs

In this appendix, we establish three convergence results for stochastic integral equations that are useful to establish the convergence of our coefficient and state processes. In Sect. B.1, we consider sequences of SDEs parametrised by  $\eta \rightarrow 0$  with “positive feedback”, i.e., sequences of equations that are driven away from their limits. In Sect. B.2, we consider sequences of random ODE systems parametrised by  $\eta \rightarrow 0$  with “negative feedback”, which are driven towards their limit.

For equations with positive feedback, we require a priori information about the terminal value and the increments of the coefficient processes to be bounded independently of their past (see Conditions A.1 and A.2). These bounds prevent the stochastic process from reaching its terminal value if a large difference between the limit and the pre-limit occurs at some time  $s \in (0, T)$ . This will imply a.s. uniform convergence. An almost sure statement under negative feedback cannot be expected. Instead, we follow a pathwise approach and establish uniform convergence in probability.

### B.1 SDEs with positive feedback

#### B.1.1 An equation without scaling

We first consider a family of real-valued, adapted, continuous stochastic processes  $b^\eta = (b_t^\eta)_{t \in (0, T)}$  that satisfy the integral equation

$$db_t^\eta = (a(t, b_t^\eta) + q_t^\eta) dt + \overline{Z}_t^\eta dW_t \quad \text{on } (0, T), \quad (\text{B.1})$$

where  $a: \Omega \times (0, T) \times \mathbb{R} \rightarrow \mathbb{R}$  is adapted and continuous and  $(q_t^\eta)_{t \in (0, T)}$  is an adapted, real-valued, continuous stochastic process and

$$(\overline{Z}_t^\eta)_{t \in (0, T)} \in \mathcal{L}_{\text{Prog}}^2(\Omega \times (0, T-]; \mathbb{R}^m).$$

Our goal is to prove that the processes  $(b_t^\eta)_{t \in (0, T)}$  with terminal conditions  $b_T^\eta = 0$  converge to 0 as  $\eta \rightarrow 0$  uniformly on  $(0, T)$  if the process  $q^\eta$  does and if the mapping  $a$  is such that  $b^\eta$  is driven away from 0. Intuitively, the last condition makes it impossible for  $b^\eta$  to return to 0 at the terminal time once it has left a neighbourhood of 0. Specifically, we assume that the following assumption is satisfied.

**Assumption B.1** i) There exists a real constant  $\overline{a} > 0$  such that

$$\sup_{t \in (0, T)} \sup_{\eta > 0} |a(t, b_t^\eta)| \leq \overline{a}.$$

ii) For all  $\varepsilon > 0$ , there exists a  $\delta > 0$  such that for all  $\eta > 0$ ,

$$\mathbb{P} \left[ \sup_{t \in [T-\delta, T)} |b_t^\eta| \leq \varepsilon \right] = 1.$$

iii) There exists a function  $\eta_1: (0, \infty) \rightarrow (0, \infty)$  such that for all  $\varepsilon > 0$  and  $\eta \leq \eta_1(\varepsilon)$ ,

$$\mathbb{P}\left[\sup_{t \in (0, T-\varepsilon]} |q_t^\eta| \leq \varepsilon\right] = 1.$$

iv) There exists an  $\varepsilon_1 > 0$  such that

$$\inf_{t \in (0, T)} \inf_{b \in [-\varepsilon_1, \varepsilon_1]} ba(t, b) \geq 0.$$

**Lemma B.2** *Under Assumption B.1, there exists a function  $\eta_0: (0, \infty) \rightarrow (0, \infty)$ , which depends only on  $\eta_1$  and the mapping  $\varepsilon \mapsto \delta$  introduced in Assumption B.1 ii), such that for all  $\varepsilon > 0$  and all  $\eta \in (0, \eta_0(\varepsilon)]$ ,*

$$\mathbb{P}\left[\sup_{s \in (0, T)} |b_s^\eta| \leq \varepsilon\right] = 1.$$

**Proof** Let  $\varepsilon \in (0, \varepsilon_1]$ . In the first step, we show that there exists an  $\eta_0(\varepsilon) > 0$  such that

$$\sigma^{\varepsilon, \eta} := \inf\{t \in (0, T) : \mathbb{P}[b_s^\eta < -\varepsilon] = 0 \text{ for all } s \in [t, T)\}$$

satisfies  $\sigma^{\varepsilon, \eta} = 0$  for all  $\eta \in (0, \eta_0(\varepsilon)]$ . Since  $b^\eta$  is continuous, this will imply that for all  $\eta \in (0, \eta_0(\varepsilon)]$ ,

$$\mathbb{P}\left[\inf_{s \in (0, T)} b_s^\eta \geq -\varepsilon\right] = 1.$$

The other part of the statement can be proved analogously.

We now assume to the contrary that  $0 < \sigma^{\varepsilon, \eta}$  and prove that this leads to a contradiction if  $\eta$  is small enough. For all  $s \in (0, T)$  and  $\eta \in (0, \infty)$ , we can define the stopping time

$$\tau^s := \tau^{s, \varepsilon, \eta} := \inf\{u \in [s, T) : b_u^\eta = -\varepsilon/2\}.$$

By Assumption B.1 ii), there exists a small enough  $\delta > 0$  such that for all  $s \in (0, T)$  and  $\eta \in (0, \infty)$ ,

$$\mathbb{P}[b_s^\eta < -\varepsilon, \tau^s > T - \delta] = 0. \quad (\text{B.2})$$

Since  $\tau^s$  is a stopping time, the square-integrability of  $\bar{Z}^\eta$ , the optional sampling theorem and Assumption B.1 i), iv), (B.1) and (B.2) yield that for all  $s \in (0, \sigma^{\varepsilon, \eta})$  and  $\eta \in (0, \infty)$ ,

$$\begin{aligned} \frac{\varepsilon}{2} \mathbb{P}[b_s^\eta < -\varepsilon] &\leq \mathbb{E}[\mathbb{1}_{\{b_s^\eta < -\varepsilon\}}(b_{\tau^s}^\eta - b_s^\eta)] \\ &= \mathbb{E}\left[\mathbb{1}_{\{b_s^\eta < -\varepsilon\}}\left(\int_s^{\tau^s \wedge \sigma^{\varepsilon, \eta}} a(t, b_t^\eta) dt \right. \right. \\ &\quad \left. \left. + \int_{\tau^s \wedge \sigma^{\varepsilon, \eta}}^{\tau^s} a(t, b_t^\eta) dt + \int_s^{\tau^s} q_t^\eta dt\right)\right]. \end{aligned}$$

Inside the second integral, if  $\sigma^{\varepsilon, \eta} < \tau^s$ , we have  $b_t^\eta \in [-\varepsilon, -\varepsilon/2]$ . Hence

$$\frac{\varepsilon}{2} \mathbb{P}[b_s^\eta < -\varepsilon] \leq \mathbb{P}[b_s^\eta < -\varepsilon] \left( (\sigma^{\varepsilon, \eta} - s) \bar{a} + T \sup_{t \in [0, T-\delta]} |q_t^\eta| \right). \quad (\text{B.3})$$

The assumption  $0 < \sigma^{\varepsilon, \eta}$  implies that for any  $\alpha > 0$ , we can find

$$\sigma^{\varepsilon, \eta} - \alpha \leq s \leq \sigma^{\varepsilon, \eta}$$

with  $\mathbb{P}[b_s^\eta < -\varepsilon] > 0$ . However, due to Assumption B.1 ii) and iii), we can choose  $\alpha \leq \varepsilon/(5\bar{a})$ ,  $\delta \leq \varepsilon/(5T)$  and  $\eta_0(\varepsilon) < \eta_1(\delta)$  such that (B.3) leads to a contradiction because  $\mathbb{P}[b_s^\eta < -\varepsilon] > 0$ . This implies  $0 = \sigma^{\varepsilon, \eta}$ .  $\square$

### B.1.2 An equation with scaling

In this section, we establish an abstract convergence result for stochastic processes  $(P_t^\eta)_{t \in (0, T)}$  indexed by some parameter  $\eta > 0$  that satisfy the integral equation

$$d(\psi P^\eta)_t = \sqrt{\eta^{-1}} (a(P_t^\eta, P_t^0) + q_t^\eta) dt + dL_t^\eta + \bar{Z}_t^\eta dW_t \quad \text{on } (0, T), \quad (\text{B.4})$$

where  $a : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a measurable mapping,  $(\psi_t)_{t \in (0, T)}$ ,  $(P_t^\eta)_{t \in (0, T)}$ ,  $(P_t^0)_{t \in (0, T)}$ ,  $(q_t^\eta)_{t \in (0, T)}$  and  $(L_t^\eta)_{t \in (0, T)}$  are adapted, real-valued, continuous stochastic processes and

$$(\bar{Z}_t^\eta)_{t \in (0, T)} \in \mathcal{L}_{\text{Prog}}^2(\Omega \times (0, T-]; \mathbb{R}^m).$$

Our goal is to prove that the processes  $P^\eta$  converge to  $P^0$  as  $\eta \rightarrow 0$  uniformly on compact subintervals of  $(0, T)$  if the mapping  $a$  is such that  $P^\eta$  is driven away from  $P^0$  and if the boundary condition  $\lim_{t \rightarrow T} (P_t^\eta - P_t^0) \geq 0$  holds. If the integral equation (B.4) holds on the whole interval  $(0, T]$ , then it is enough to assume that  $P_T^\eta \geq P_T^0$ . When applying the abstract convergence result to the BSDEs (3.2) and (3.3), the former condition holds if  $N = \infty$ , while the latter holds if  $N$  is finite.

**Assumption B.3** i) There exist positive constants  $\underline{\psi}$ ,  $\bar{\psi}$ ,  $\bar{P}$  such that  $\underline{\psi} \leq \psi \leq \bar{\psi}$ ,  $|P^0| \leq \bar{P}$  and  $P^\eta \geq -\bar{P}$ . Moreover, there exists a function  $\eta_1 : (0, \infty) \rightarrow (0, \infty)$  such that

$$P^\eta|_{\Omega \times (0, T-\varepsilon]} \leq \bar{P} \quad \text{for all } \varepsilon > 0 \text{ and } \eta \leq \eta_1(\varepsilon).$$

ii) One of the following two “boundary conditions” holds:

a) For all  $\eta > 0$ , there exists a  $T_\eta < T$  such that

$$\mathbb{P} \left[ \inf_{T_\eta \leq t < T} (P_t^\eta - P_t^0) \geq 0 \right] = 1.$$

b) The integral equation (B.4) holds on the whole interval  $(0, T]$  and  $P_T^\eta \geq P_T^0$ . Thus in particular, the processes  $\psi$ ,  $P^\eta$ ,  $P^0$ ,  $q^\eta$  and  $L^\eta$  are continuous on  $(0, T]$  and  $\bar{Z}^\eta \in \mathcal{L}_{\text{Prog}}^2(\Omega \times (0, T]; \mathbb{R}^m)$ .

iii) There exists a function  $\eta_2: (0, \infty) \rightarrow (0, \infty)$  such that for all  $\varepsilon > 0$  and all  $\eta \in (0, \eta_2(\varepsilon)]$ ,

$$\mathbb{P}\left[\inf_{t \in (0, T-\varepsilon]} q_t^\eta \geq -\varepsilon, \sup_{t \in (0, T)} q_t^\eta \leq \varepsilon\right] = 1.$$

iv) The function  $a: \mathbb{R}^2 \rightarrow \mathbb{R}$  is nondecreasing and differentiable in the first variable and  $a(P_t^0(\omega), P_t^0(\omega)) \equiv 0$ . Moreover, there exist  $\bar{\varepsilon}, \beta > 0$  such that

$$\frac{\partial}{\partial x} a(x, P_t^0(\omega)) \geq \beta \quad \text{for all } |x - P_t^0(\omega)| \leq \bar{\varepsilon}.$$

v) The processes  $\psi$  and  $P^0$  satisfy Condition A.2, and the processes  $L^\eta$  ( $\eta > 0$ ) satisfy Condition A.1 uniformly in  $\eta > 0$ .

**Lemma B.4** *Under Assumption B.3, there exists a mapping  $\eta_0: (0, \infty) \rightarrow (0, \infty)$  that depends only on  $\psi, \bar{P}, \eta_1, \eta_2, P^0, \bar{\varepsilon}, \beta$  and the mapping  $\varepsilon \mapsto \delta$  corresponding to the uniform satisfaction of Condition A.1 by  $(L^\eta)_{\eta>0}$  such that for all  $\varepsilon > 0$  and  $\eta \in (0, \eta_0(\varepsilon)]$ ,*

$$\mathbb{P}\left[\sup_{s \in (0, T-\varepsilon]} (P_s^\eta - P_s^0) \leq \varepsilon, \inf_{s \in (0, T)} (P_s^\eta - P_s^0) \geq -\varepsilon\right] = 1.$$

**Proof** Let  $p^\eta := P^\eta - P^0$  and fix  $\varepsilon \in (0, \bar{\varepsilon}]$ . We first show that there exists  $\eta_0(\varepsilon) > 0$  such that

$$\sup_{\eta \leq \eta_0(\varepsilon)} \sup_{s \in (0, T-\varepsilon]} \mathbb{P}[p_s^\eta > \varepsilon] = 0. \quad (\text{B.5})$$

Since  $p^\eta$  is continuous, this proves that, for all  $\eta \leq \eta_0(\varepsilon)$ ,

$$\mathbb{P}\left[\sup_{s \in (0, T-\varepsilon]} p_s^\eta \leq \varepsilon\right] = 1.$$

To show (B.5), we fix  $\delta \leq \varepsilon/2$  and  $s \in (0, T - \varepsilon]$ . For all  $\eta > 0$ , we define the stopping time

$$\tau^\eta := (s + \delta) \wedge \inf\{u \in [s, T) : |p_u^\eta| = \varepsilon/2\} \leq T - \frac{\varepsilon}{2}.$$

The  $L^2$ -integrability of  $\bar{Z}^\eta$  on  $[s, T - \varepsilon/2]$  and the optional sampling theorem give

$$\begin{aligned} 0 = \mathbb{E}\left[\mathbb{1}_{\{p_s^\eta > \varepsilon\}} \left( \sqrt{\eta^{-1}} \int_s^{\tau^\eta} (a(P_u^\eta, P_u^0) + q_u^\eta) du + L_{\tau^\eta}^\eta - L_s^\eta \right. \right. \\ \left. \left. + \psi_s P_s^0 - \psi_{\tau^\eta} P_{\tau^\eta}^0 + \psi_s p_s^\eta - \psi_{\tau^\eta} p_{\tau^\eta}^\eta \right) \right]. \end{aligned} \quad (\text{B.6})$$

We now analyse the various terms in (B.6) separately to deduce that  $\mathbb{P}[p_s^\eta > \varepsilon] = 0$  if

$$\eta \leq \eta_0(\varepsilon) := \min\left(\eta_1(\varepsilon/2), \eta_2(\min(\delta, \beta\varepsilon/4))\right).$$



We first consider the term

$$\mathbb{E} \left[ \mathbb{1}_{\{p_s^\eta > \varepsilon\}} \sqrt{\eta^{-1}} \int_s^{\tau^\eta} (a(P_u^\eta, P_u^0) + q_u^\eta) du \right].$$

On  $\{p_s^\eta > \varepsilon\}$ , we have  $p_u^\eta \geq \varepsilon/2$  for all  $u \in [s, \tau^\eta]$  and hence

$$a(P_u^\eta, P_u^0) = a(P_u^0 + p_u^\eta, P_u^0) \geq a(P_u^0 + \varepsilon/2, P_u^0) \geq \frac{\beta\varepsilon}{2}. \quad (\text{B.7})$$

Together with our assumption on the process  $q^\eta$  and our choice of  $\eta$ , this implies by using (B.7) and  $q_u^\eta \geq -\beta\varepsilon/4$  that

$$\begin{aligned} & \mathbb{E} \left[ \mathbb{1}_{\{p_s^\eta > \varepsilon\}} \sqrt{\eta^{-1}} \int_s^{\tau^\eta} (a(P_u^\eta, P_u^0) + q_u^\eta) du \right] \\ & \geq \mathbb{E} \left[ \mathbb{1}_{\{p_s^\eta > \varepsilon\}} \sqrt{\eta^{-1}} (\tau^\eta - s) \frac{\beta\varepsilon}{4} \right] \\ & \geq \delta \sqrt{\eta^{-1}} \frac{\beta\varepsilon}{4} \mathbb{P}[p_s^\eta > \varepsilon \text{ and } \tau^\eta = s + \delta]. \end{aligned}$$

By Assumption B.3 v), Lemma A.5 and Theorem A.8, we can choose for any  $\varepsilon_0 > 0$  some  $\delta > 0$ , which does not depend on  $\eta$ , such that

$$|\mathbb{E}[\mathbb{1}_{\{p_s^\eta > \varepsilon\}}(Y_{\tau^\eta} - Y_s)]| \leq \varepsilon_0 \mathbb{P}[p_s^\eta > \varepsilon] \quad \text{for } Y = \psi P^0, L^\eta \quad (\text{B.8})$$

and

$$\mathbb{E}[|\psi_{\tau^\eta} - \psi_s| | \mathcal{F}_s] \leq \varepsilon_0.$$

Using (B.8), the fact that  $P^\eta|_{\Omega \times (0, T-\varepsilon]} \leq \bar{P}$  by Assumption B.3 i) and  $p_{\tau^\eta}^\eta \leq 2\bar{P}$  yields

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{\{p_s^\eta > \varepsilon\}}(\psi_s p_s^\eta - \psi_{\tau^\eta} p_{\tau^\eta}^\eta)] & \geq \mathbb{E}[\mathbb{1}_{\{p_s^\eta > \varepsilon \text{ and } \tau^\eta < s + \delta\}}(\psi_s \varepsilon/2 + (\psi_s - \psi_{\tau^\eta})\varepsilon/2)] \\ & \quad + \mathbb{E}[\mathbb{1}_{\{p_s^\eta > \varepsilon \text{ and } \tau^\eta = s + \delta\}}(\psi_s \varepsilon - \psi_{\tau^\eta} p_{\tau^\eta}^\eta)] \\ & \geq \frac{\varepsilon}{2} \underline{\psi} \mathbb{P}[p_s^\eta > \varepsilon \text{ and } \tau^\eta < s + \delta] - \frac{\varepsilon \varepsilon_0}{2} \mathbb{P}[p_s^\eta > \varepsilon] \\ & \quad + (\varepsilon \underline{\psi} - 2\bar{\psi} \bar{P}) \mathbb{P}[p_s^\eta > \varepsilon \text{ and } \tau^\eta = s + \delta]. \end{aligned}$$

Altogether this yields

$$\begin{aligned} 0 & \geq \mathbb{P}[p_s^\eta > \varepsilon \text{ and } \tau^\eta = s + \delta] \left( \delta \sqrt{\eta_0(\varepsilon)^{-1}} \frac{\beta\varepsilon}{4} - 2\varepsilon_0 - \frac{\varepsilon \varepsilon_0}{2} + \varepsilon \underline{\psi} - 2\bar{\psi} \bar{P} \right) \\ & \quad + \mathbb{P}[p_s^\eta > \varepsilon \text{ and } \tau^\eta < s + \delta] \left( -2\varepsilon_0 + \frac{\varepsilon}{2} \underline{\psi} - \frac{\varepsilon \varepsilon_0}{2} \right). \quad (\text{B.9}) \end{aligned}$$

Now if we choose first  $\varepsilon_0$  and then  $\eta_0(\varepsilon)$  small enough, the coefficients that multiply the probabilities in (B.9) become positive. Hence both probabilities must be equal to

zero, and so we have  $\mathbb{P}[p_s^\eta > \varepsilon] = 0$  if  $\eta \leq \eta_0(\varepsilon)$ . Analogously, we can prove that  $\mathbb{P}[p_s^\eta < -\varepsilon] = 0$  for all  $s \in (0, T)$ . The main difference is that  $\tau^\eta < T - \delta$  does not hold on the set  $\{p_s^\eta < -\varepsilon\}$ . Instead, we only obtain  $\tau^\eta < T$  (if Assumption B.3 ii) b) holds) or  $\tau^\eta < T_\eta$  (if Assumption B.3 ii) a) holds).  $\square$

## B.2 ODE systems with negative feedback

In this section, we establish a convergence result for ODEs with random coefficients and a “nearly singular” driver. Specifically, we consider pairs of continuously differentiable stochastic processes  $(X^\eta, Z^\eta)$  which satisfy the ODE system

$$\begin{aligned}\dot{X}_t^\eta &= -A_t^\eta(X_t^\eta + B_t^\eta Z_t^\eta), & X_0^\eta &= x_0, \\ \dot{Z}_t^\eta &= C_t X_t^\eta + D_t Z_t^\eta, & Z_0^\eta &= z_0\end{aligned}$$

on some time interval  $[0, S]$  for all  $\eta > 0$ , where  $A^\eta, B^\eta, C$  and  $D$  are continuous, adapted, real-valued stochastic processes. We assume that the process  $A^\eta$  converges in probability to infinity as  $\eta \rightarrow 0$ , that the process  $B^\eta$  converges to a process  $B^0$  in probability and that the processes  $(X^\eta, Z^\eta)$  are uniformly bounded in probability.

**Assumption B.5** There exists a continuous adapted process  $B^0$  such that for all  $\varepsilon, \delta \in (0, \infty)$ , there exists some  $\eta_0 = \eta_0(\varepsilon, \delta) > 0$  such that for all  $\eta \in (0, \eta_0]$ ,

$$\mathbb{P}\left[\inf_{t \in [0, S]} A_t^\eta \geq \frac{1}{\varepsilon}, \sup_{t \in [0, S]} |B_t^\eta - B_t^0| \leq \varepsilon\right] \geq 1 - \delta. \quad (\text{B.10})$$

Moreover, for all  $\delta \in (0, 1)$ , there exists an  $L > 0$  such that for all  $\eta > 0$ ,

$$\mathbb{P}\left[\sup_{t \in [0, S]} (|X_t^\eta| + |Z_t^\eta|) \leq L\right] \geq 1 - \delta. \quad (\text{B.11})$$

In terms of the process  $B^0$ , we define

$$\begin{aligned}X_t^0 &:= -B_t^0 Z_t^0, \\ Z_t^0 &:= z_0 \exp\left(\int_0^t (-C_s B_s^0 + D_s) ds\right).\end{aligned}$$

Our goal is to prove that  $(X^\eta, Z^\eta) \rightarrow (X^0, Z^0)$  in a suitable sense as  $\eta \rightarrow 0$ . Since the initial condition of the processes  $X^\eta$  does not vary with  $\eta$  and since  $A_0^\eta \uparrow \infty$ , we cannot expect convergence in  $t = 0$ . Instead, we first prove that  $\frac{\dot{X}^\eta}{A^\eta}$  converges to 0 on any compact subinterval of  $(0, S]$ , with lower, respectively upper convergence near the origin. More precisely, we have the following result.

**Lemma B.6** *If Assumption B.5 holds and if  $x_0 + B_0^0 z_0 > 0$ , then for all  $\varepsilon, \delta \in (0, \infty)$ , there exists  $\eta_1 = \eta_1(\varepsilon, \delta) > 0$  such that for all  $\eta \in (0, \eta_1]$ ,*

$$\mathbb{P}\left[\sup_{t \in [0, S]} \frac{\dot{X}_t^\eta}{A_t^\eta} \leq \varepsilon, \inf_{t \in [\varepsilon, S]} \frac{\dot{X}_t^\eta}{A_t^\eta} \geq -\varepsilon\right] \geq 1 - \delta.$$

If  $x_0 + B_0^0 z_0 < 0$ , then the intervals  $[0, S]$  and  $[\varepsilon, S]$  in the above statement are to be swapped. The mapping  $\eta_1$  depends only on  $B^0, C, D, x_0, z_0$  and the mappings  $(\varepsilon, \delta) \mapsto \eta_0(\varepsilon, \delta)$  and  $\delta \mapsto L$  introduced in Assumption B.5.

**Proof** Let  $y_0 := x_0 + B_0^0 z_0$ . We assume without loss of generality that  $y_0 > 0$ . Then there exists  $\alpha_0 > 0$  such that

$$y_0 - \alpha |z_0| > 0 \quad \text{for all } \alpha \leq \alpha_0.$$

Let us now put

$$\begin{aligned} M^{L,\eta} &:= \left\{ \sup_{t \in [0, S]} (|B_t^0| + |C_t| + |D_t| + |X_t^\eta| + |Z_t^\eta|) \leq L \right\}, \\ M^{\alpha,\eta} &:= \left\{ \inf_{t \in [0, S]} A_t^\eta \geq \frac{1}{\alpha}, \sup_{t \in [0, S]} |B_t^\eta - B_t^0| \leq \alpha \right\}, \\ M^{\beta,v} &:= \left\{ \sup_{s, t \in [0, S], |s-t| \leq v} |B_s^0 - B_t^0| \leq \beta \right\}. \end{aligned}$$

Due to Assumption B.5 and Lemma 3.9, for all  $\delta > 0$ ,  $\beta > 0$  and  $\alpha \leq \alpha_0$ , there exist constants  $L_0(\delta)$ ,  $\eta_0(\alpha, \delta)$  and  $v_0(\beta, \delta)$  such that

$$\inf_{L \geq L_0} \inf_{\eta} \mathbb{P}[M^{L,\eta}] \geq 1 - \frac{\delta}{6}, \quad \inf_{\eta \leq \eta_0} \mathbb{P}[M^{\alpha,\eta}] \geq 1 - \frac{\delta}{6}, \quad \inf_{v \leq v_0} \mathbb{P}[M^{\beta,v}] \geq 1 - \frac{\delta}{6}.$$

Moreover, for all  $\alpha \leq \alpha_0$  and  $\mathbb{P}$ -a.a.  $\omega \in M^{\alpha,\eta}$ ,

$$\dot{X}_0^\eta(\omega) \leq -A_0^\eta(\omega)(y_0 - |B_0^\eta(\omega) - B_0^0(\omega)||z_0|) \leq -\frac{1}{\alpha}(y_0 - \alpha|z_0|) < 0. \quad (\text{B.12})$$

We are now ready to prove that for all  $\varepsilon, \delta > 0$ , there exists  $\eta_1 > 0$  such that for all  $\eta \leq \eta_1$ ,

$$\mathbb{P} \left[ \sup_{t \in [0, S]} \frac{\dot{X}_t^\eta}{A_t^\eta} \leq \varepsilon \right] \geq 1 - \frac{\delta}{2}. \quad (\text{B.13})$$

To this end, we take an  $\omega \in M^{L,\eta} \cap M^{\alpha,\eta} \cap M^{\beta,v}$  where (B.12) holds and assume that there exists some  $t \in [0, S]$  such that  $\frac{\dot{X}_t^\eta(\omega)}{A_t^\eta(\omega)} \geq \varepsilon$ . It is enough to show that such a  $t$  cannot exist for  $L$  large and  $\alpha, \beta, v, \eta$  small enough. Since  $\frac{\dot{X}^\eta(\omega)}{A^\eta(\omega)}$  is continuous and  $\frac{\dot{X}_0^\eta(\omega)}{A_0^\eta(\omega)} < 0$  if  $\alpha \leq \alpha_0$ , we can choose a minimal  $t \in [0, S]$  with the property that  $\frac{\dot{X}_t^\eta(\omega)}{A_t^\eta(\omega)} = \varepsilon$ . Let

$$s := \sup \left\{ u \in (0, t) : \frac{\dot{X}_u^\eta(\omega)}{A_u^\eta(\omega)} = \varepsilon/2 \right\}.$$

Then  $\frac{\dot{X}_s^\eta(\omega)}{A_s^\eta(\omega)} = \frac{\varepsilon}{2}$  and  $\frac{\dot{X}_u^\eta(\omega)}{A_u^\eta(\omega)} > \frac{\varepsilon}{2}$  for all  $u \in (s, t]$ . We now distinguish two cases. If  $t - s \geq \nu$ , then the fact that  $\omega \in M^{L,\eta} \cap M^{\alpha,\eta}$  yields

$$2L \geq X_t^\eta(\omega) - X_s^\eta(\omega) = \int_s^t \frac{\dot{X}_u^\eta(\omega)}{A_u^\eta(\omega)} A_u^\eta(\omega) du \geq \frac{\nu\varepsilon}{2\alpha}. \quad (\text{B.14})$$

If  $t - s < \nu$ , we have

$$\begin{aligned} \frac{\varepsilon}{2} &= \frac{\dot{X}_t^\eta(\omega)}{A_t^\eta(\omega)} - \frac{\dot{X}_s^\eta(\omega)}{A_s^\eta(\omega)} \\ &= -X_t^\eta(\omega) - B_t^\eta(\omega)Z_t^\eta(\omega) + X_s^\eta(\omega) + B_s^\eta(\omega)Z_s^\eta(\omega). \end{aligned}$$

Using that  $\omega \in M^{\alpha,\eta}$  yields  $-X_t^\eta(\omega) + X_s^\eta(\omega) = -\int_s^t \frac{\dot{X}_u^\eta(\omega)}{A_u^\eta(\omega)} A_u^\eta(\omega) du < 0$ . Using that  $\omega \in M^{L,\eta} \cap M^{\beta,\nu}$  yields

$$\begin{aligned} \frac{\varepsilon}{2} &\leq -B_t^\eta(\omega) \int_s^t \dot{Z}_u^\eta(\omega) du - Z_s^\eta(\omega)(B_t^\eta(\omega) - B_s^\eta(\omega)) \\ &\leq 2L^2\nu(|B_t^0(\omega)| + |B_t^\eta(\omega) - B_t^0(\omega)|) \\ &\quad + L(|B_t^\eta(\omega) - B_t^0(\omega)| + |B_t^0(\omega) - B_s^0(\omega)| + |B_s^0(\omega) - B_s^\eta(\omega)|) \\ &\leq 2L^2\nu(L + \alpha) + L(2\alpha + \beta). \end{aligned} \quad (\text{B.15})$$

Now we first choose  $L \geq L_0(\delta)$ , then  $\beta > 0$  such that  $3L\beta < \frac{\varepsilon}{4}$ , then  $\nu \leq \nu_0(\beta, \delta)$  such that  $2L^2\nu(L + \alpha_0) < \frac{\varepsilon}{4}$ , then  $\alpha \leq \alpha_0$  such that  $\frac{\nu\varepsilon}{2\alpha} > 2L$  and  $\alpha \leq \beta$ , and finally  $\eta_1(\varepsilon, \delta) \leq \eta_0(\alpha, \delta)$ . Then both (B.14) and (B.15) are violated. As a result,

$$M^{L,\eta} \cap M^{\alpha,\eta} \cap M^{\beta,\nu} \subseteq \left\{ \sup_{t \in [0, S]} \frac{\dot{X}_t^\eta}{A_t^\eta} \leq \varepsilon \right\},$$

and hence (B.13) holds.

It remains to prove that for suitably chosen parameters,

$$M^{L,\eta} \cap M^{\alpha,\eta} \cap M^{\beta,\nu} \subseteq \left\{ \inf_{t \in [\varepsilon, S]} \frac{\dot{X}_t^\eta}{A_t^\eta} \geq -\varepsilon \right\}.$$

To this end, we fix  $\omega \in M^{L,\eta} \cap M^{\alpha,\eta} \cap M^{\beta,\nu}$  and assume that there exists a minimal  $t \in [\varepsilon, S]$  such that  $\frac{\dot{X}_t^\eta(\omega)}{A_t^\eta(\omega)} \leq -\varepsilon$ . Using that  $\omega \in M^{L,\eta} \cap M^{\alpha,\eta}$ , it is straightforward to show that for all sufficiently small  $\alpha$ , there exists an  $r \in [0, \varepsilon]$  such that  $\frac{\dot{X}_r^\eta(\omega)}{A_r^\eta(\omega)} \geq -\varepsilon/2$ . Hence we can define

$$s := \sup \left\{ u \in [0, t) : \frac{\dot{X}_u^\eta(\omega)}{A_u^\eta(\omega)} = -\varepsilon/2 \right\}.$$

Then  $0 \leq s < t$  and  $\frac{\dot{X}_u^\eta(\omega)}{A_u^\eta(\omega)} \in [-\varepsilon, -\varepsilon/2]$  for all  $u \in [s, t]$ . We can now use the same arguments as in the first step to conclude that

$$\mathbb{P} \left[ \inf_{t \in [\varepsilon, S]} \frac{\dot{X}_t^\eta}{A_t^\eta} \geq -\varepsilon \right] \geq 1 - \frac{\delta}{2}. \quad (\text{B.16})$$

Combining (B.13) and (B.16) yields the desired result.  $\square$

**Lemma B.7** *Let Assumption B.5 be satisfied and let  $x_0 + B_0^0 z_0 \neq 0$ . Then for all  $\varepsilon, \delta \in (0, \infty)$ , there exists  $\eta_2 = \eta_2(\varepsilon, \delta) > 0$  such that for all  $\eta \in (0, \eta_2]$ ,*

$$\mathbb{P} \left[ \sup_{t \in [0, S]} |Z_t^\eta - Z_t^0| \leq \varepsilon \right] \geq 1 - \delta.$$

*The mapping  $\eta_2$  depends only on  $B^0, C, D, x_0, z_0$  and the mappings  $(\varepsilon, \delta) \mapsto \eta_0(\varepsilon, \delta)$  and  $\delta \mapsto L$  introduced in Assumption B.5.*

**Proof** Since

$$\begin{aligned} Z_t^\eta - Z_t^0 &= \exp \left( \int_0^t (D_u - B_u^\eta C_u) du \right) \\ &\quad \times \int_0^t \exp \left( - \int_0^s (D_u - B_u^\eta C_u) du \right) \left( (B_s^0 - B_s^\eta) C_s Z_s^0 - C_s \frac{\dot{X}_s^\eta}{A_s^\eta} \right) ds, \end{aligned}$$

the assertion follows from Assumption B.5 and Lemma B.6.  $\square$

The next theorem follows by Assumption B.5 and Lemmas B.6 and B.7.

**Theorem B.8** *If Assumption B.5 holds and  $x_0 + B_0^0 z_0 > 0$ , then for all  $\varepsilon, \delta \in (0, \infty)$ , there exists  $\eta_3 = \eta_3(\varepsilon, \delta) > 0$  such that for all  $\eta \in (0, \eta_3]$ ,*

$$\mathbb{P} \left[ \sup_{t \in [0, S]} \max(|Z_t^0 - Z_t^\eta|, X_t^0 - X_t^\eta) \leq \varepsilon, \inf_{t \in [\varepsilon, S]} (X_t^0 - X_t^\eta) \geq -\varepsilon \right] \geq 1 - \delta.$$

*If  $x_0 + B_0^0 z_0 < 0$ , then the intervals  $[0, S]$  and  $[\varepsilon, S]$  in the statement are to be swapped. The mapping  $\eta_3$  depends only on  $B^0, C, D, x_0, z_0$  and the mappings  $(\varepsilon, \delta) \mapsto \eta_0(\varepsilon, \delta)$  and  $\delta \mapsto L$  introduced in Assumption B.5.*

**Acknowledgements** Financial support through the Stiftung der Deutschen Wirtschaft (sdw) is gratefully acknowledged. We thank Peter Bank, Mikhail Urusov and seminar participants at various institutions for valuable comments and remarks. We thank two anonymous referees for valuable comments and suggestions that greatly helped to improve the presentation of the results.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Ackermann, J., Kruse, T., Urusov, M.: Càdlàg semimartingale strategies for optimal trade execution in stochastic order book models. *Finance Stoch.* **25**, 757–810 (2021)
2. Alfonsi, A., Fruth, A., Schied, A.: Optimal execution strategies in limit order books with general shape functions. *Quant. Finance* **10**, 143–157 (2010)
3. Almgren, R., Chriss, N.: Optimal execution of portfolio transactions. *J. Risk* **3**, 5–40 (2001)
4. Ankirchner, S., Jeanblanc, M., Kruse, T.: BSDEs with singular terminal condition and a control problem with constraints. *SIAM J. Control Optim.* **52**, 893–913 (2014)
5. Becherer, D., Bilarev, T., Frentrop, P.: Stability for gains from large investors' strategies in  $M_1/J_1$  topologies. *Bernoulli* **25**, 1105–1140 (2019)
6. Bertsimas, D., Lo, A.W.: Optimal control of execution costs. *J. Financ. Mark.* **1**, 1–50 (1998)
7. Bouchard, B., Loeper, G., Zou, Y.: Almost-sure hedging with permanent price impact. *Finance Stoch.* **20**, 741–771 (2016)
8. Carmona, R., Webster, K.: The self-financing equation in limit order book markets. *Finance Stoch.* **23**, 729–759 (2019)
9. Cartea, Á., Donnelly, R., Jaimungal, S.: Portfolio liquidation and ambiguity aversion. In: Dempster, M.A.H., et al. (eds.) *High-Performance Computing in Finance*, pp. 77–114. Chapman & Hall, London (2018)
10. Cartea, Á., Jaimungal, S.: Incorporating order-flow into optimal execution. *Math. Financ. Econ.* **10**, 339–364 (2016)
11. Chen, N., Kou, S., Wang, C.: A partitioning algorithm for Markov decision processes with applications to market microstructure. *Manag. Sci.* **64**, 784–803 (2018)
12. El Karoui, N.: Backward stochastic differential equations. A general introduction. In: El Karoui, N., Mazliak, L. (eds.) *Pitman Research Notes in Mathematics Series*, pp. 7–26 (1997)
13. Fruth, A., Schöneborn, T., Urusov, M.: Optimal trade execution in order books with stochastic liquidity. *Math. Finance* **29**, 507–541 (2019)
14. Fu, G., Graewe, P., Horst, U., Popier, A.: A mean field game of optimal portfolio liquidation. *Math. Oper. Res.* **46**, 1250–1281 (2021)
15. Gârleanu, N., Pedersen, L.H.: Dynamic portfolio choice with frictions. *J. Econ. Theory* **165**, 487–516 (2016)
16. Gatheral, J.: No-dynamic-arbitrage and market impact. *Quant. Finance* **10**, 749–759 (2010)
17. Gatheral, J., Schied, A.: Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework. *Int. J. Theor. Appl. Finance* **14**, 353–368 (2011)
18. Graewe, P., Horst, U.: Optimal trade execution with instantaneous price impact and stochastic resilience. *SIAM J. Control Optim.* **55**, 3707–3725 (2017)
19. Graewe, P., Horst, U., Qiu, J.: A non-Markovian liquidation problem and backward SPDEs with singular terminal conditions. *SIAM J. Control Optim.* **53**, 690–711 (2015)
20. Guéant, O.: Optimal execution and block trade pricing: a general framework. *Appl. Math. Finance* **22**, 336–365 (2015)
21. Horst, U., Naujokat, F.: When to cross the spread? Trading in two-sided limit order books. *SIAM J. Financ. Math.* **5**, 278–315 (2014)
22. Horst, U., Xia, X.: Multi-dimensional optimal trade execution under stochastic resilience. *Finance Stoch.* **23**, 889–923 (2019)
23. Horst, U., Xia, X., Zhou, C.: Optimal portfolio liquidation under factor uncertainty. *Ann. Appl. Probab.* **32**, 80–123 (2022)

24. Kruse, T., Popier, A.: Minimal supersolutions for BSDEs with singular terminal condition and application to optimal position targeting. *Stoch. Process. Appl.* **126**, 2554–2592 (2016)
25. Large, J.: Measuring the resiliency of an electronic limit order book. *J. Financ. Mark.* **10**, 1–25 (2007)
26. Lo, D.K., Hall, A.D.: Resiliency of the limit order book. *J. Econ. Dyn. Control* **61**, 222–244 (2015)
27. Lokka, A., Xu, J.: Optimal liquidation for a risk averse investor in a one-sided limit order book driven by a Lévy process (2020). Preprint, Available online at <https://arxiv.org/abs/2002.03379>
28. Lorenz, C., Schied, A.: Drift dependence of optimal trade execution strategies under transient price impact. *Finance Stoch.* **17**, 743–770 (2013)
29. Neuman, E., Voß, M.: Optimal signal-adaptive trading with temporary and transient price impact. *SIAM J. Financ. Math.* **13**, 551–575 (2022)
30. Obizhaeva, A.A., Wang, J.: Optimal trading strategy and supply/demand dynamics. *J. Financ. Mark.* **16**, 1–32 (2013)
31. Predoiu, S., Shaikhet, G., Shreve, S.: Optimal execution in a general one-sided limit-order book. *SIAM J. Financ. Math.* **2**, 183–212 (2011)
32. Protter, P.E.: *Stochastic Integration and Differential Equations*, 2nd edn. Springer, Berlin (2004)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.