CrossMark

Hedging with temporary price impact

Peter Bank[1] · H. Mete Soner[2,3] · Moritz Voß[1]

Abstract We consider the problem of hedging a European contingent claim in a Bachelier model with temporary price impact as proposed by Almgren and Chriss (J Risk 3:5–39, 2001). Following the approach of Rogers and Singh (Math Financ 20:597–615, 2010) and Naujokat and Westray (Math Financ Econ 4(4):299–335, 2011), the hedging problem can be regarded as a cost optimal tracking problem of the frictionless hedging strategy. We solve this problem explicitly for general predictable target hedging strategies. It turns out that, rather than towards the current target position, the optimal policy trades towards a weighted average of expected future target positions. This generalizes an observation of Gârleanu and Pedersen (Dynamic portfolio choice with frictions. Preprint, 2013b) from their homogenous Markovian optimal investment problem to a general hedging problem. Our findings complement a number of previous studies in the literature on optimal strategies in illiquid markets as, e.g., Gârleanu and Pedersen (Dynamic portfolio choice with frictions. Preprint, 2013b), Naujokat and Westray (Math Financ Econ 4(4):299–335, 2011), Rogers and Singh (Math Financ 20:597–615, 2010), Almgren and Li (Option hedging with smooth market impact. Preprint, 2015), Moreau et al. (Math Financ. doi:10.1111/mafi.12098, 2015), Kallsen and Muhle-Karbe (High-resilience limits of block-shaped order books. Preprint, 2014), Guasoni and Weber (Mathematical Financ. doi:10.1111/mafi.12099, 2015a; Nonlinear price impact and portfolio choice. Preprint, 2015b), where the frictionless hedging strategy is confined to

✉ Peter Bank
  bank@math.tu-berlin.de

  H. Mete Soner
  mete.soner@math.ethz.ch

  Moritz Voß
  voss@math.tu-berlin.de

[1] Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

[2] Departement für Mathematik, ETH Zürich, Rämistraße 101, 8092 Zurich, Switzerland

[3] Swiss Finance Institute, Switzerland,
  http://sfi.ch/

diffusions. The consideration of general predictable reference strategies is made possible by the use of a convex analysis approach instead of the more common dynamic programming methods.

1 Introduction

The construction of effective hedging strategies against financial risk is one of the key problems in Mathematical Finance. The seminal work of Black and Scholes [5] and Merton [19] showed how this task can be carried out in an idealized frictionless market by dynamically trading perfectly liquid assets. However, in recent years there has been a growing awareness that these idealizations can lead to misguided hedging strategies with non negligible costs, particularly when these prescribe a fast reallocation of assets in short periods of time in the presence of liquidity frictions like price impact. This has spurred the development of new financial models which take into account the impact of transactions on execution prices; see, e.g., the survey by Gökay et al. [13].

The two most widely used models go back to Almgren and Chriss [2] as well as Obizhaeva and Wang [22], respectively: Loosely speaking, the model of Almgren and Chriss is characterized by directly specifying functions describing the temporary and permanent impacts of a given order on the price. The model of Obizhaeva and Wang assumes that trading takes place in a block-shaped limit order book with persistent price impact which is vanishing at a finite resilience rate. As recently discussed in Kallsen and Muhle-Karbe [17], the former can be regarded as the high-resilience limit of the latter.

Within these two models, most of the existing literature studies the problem of optimally liquidating an exogenously given position by some fixed time horizon (cf., e.g., Alfonsi et al. [1], Almgren and Chriss [2], Almgren [4], Obizhaeva and Wang [22], Predoiu et al. [23], Schied and Schöneborn [25]). Further work is also devoted to the more involved problem of optimal portfolio choice, cf., e.g., Gârleanu and Pedersen [11,12], Guasoni and Weber [14,15], Kallsen and Muhle-Karbe [17], Moreau et al. [20]. However, only a few papers directly address the problem of hedging a contingent claim in the presence of price impact as modeled above, cf. Almgren and Li [3], Guéant and Pu [16], Rogers and Singh [24], and also Naujokat and Westray [21].

The papers most closely related to ours mathematically are Rogers and Singh [24] and Naujokat and Westray [21]. Rogers and Singh [24] analyse the problem of hedging a European contingent claim in a Black–Scholes model in the presence of purely temporary price impact as in Almgren and Chriss [2]. They relate the hedging problem to a cost optimal tracking problem of the frictionless Black–Scholes delta hedge. Naujokat and Westray [21] directly study the problem of optimally following a given target strategy in an illiquid financial market under the same type of liquidity costs; see also Cartea and Jaimungal [7] for a Markovian order flow tracking problem. By contrast to these papers, we will focus on a non-Markovian setup with general predictable target strategies.

Instead of the more common dynamic programming methods used in the papers cited above, our approach is convex analytic along the lines of Pontryagin's maximum princi-

ple. This allows us to consider general predictable target strategies and not only continuous diffusion-type processes. This is particularly important for hedging in illiquid markets when the frictionless reference hedge portfolio prescribes sizable instantaneous reallocations as, e.g., already in the case of discrete Asian options which was not covered by the literature so far. We derive first order conditions of the considered quadratic optimization problems which take the form of a linear forward backward stochastic differential equation (FBSDE). Solutions to these are explicitly available and give us the optimal frictional hedges. In fact, when considered in a Brownian setting, our approach can be viewed as a special case of the stochastic linear-quadratic control problem studied by Kohlmann and Tang [18]. Mathematically, the novelty of our contribution is the interpretation of the optimal tracking strategy. Indeed, it turns out that the optimal policy does *not* instantaneously trade from its current position towards the current target position but towards a weighted average of *expected future target positions* which does not occur in the work of Kohlmann and Tang [18]. An interesting consequence from a financial point of view is that this averaging allows us to understand how singularities in the frictionless reference strategy have to be addressed in a model with frictions: A singularity in the frictionless target hedge is smoothed out when averaging the weighted future target positions which yields sensible hedging strategies for illiquid markets. Additionally, we also study a constrained version of the problem where the terminal hedging position is restricted to a certain exogenously prescribed level. This can be viewed as a way to deal with physical delivery in derivative contracts. Our explicit solution reveals how the hedger's focus shifts systematically from tracking the frictionless target position to attaining the prescribed terminal position. Here, our convex analytic approach allows us to avoid the consideration of nonlinear Hamilton–Jacobi–Bellmann equations with singular terminal conditions and the challenges that these entail. We also give a sharp characterization of those terminal positions which can be reached with finite expected trading costs by characterizing the speed at which the size of these positions is revealed towards the end.

Conceptually, our result generalizes an observation by Gârleanu and Pedersen [12] who consider quadratic utility maximization in homogeneous Markovian models on an infinite time horizon and interpret their solution as trading towards an exponentially weighted average of future expected Markowitz portfolios. A similar interpretation is given by Naujokat and Westray [21] in their equally Markovian Example 7.1; see Cartea and Jaimungal [7] for a similarly Markovian study of tracking of order flow in high-frequency trading. These strategies as well as ours contrast with strategies targeting the present frictionless optimum directly, which are considered in many papers on asymptotically optimal portfolios with small transaction costs, including Guasoni and Weber [14,15], Moreau et al. [20], Rogers and Singh [24], and Kallsen and Muhle-Karbe [17]. In all the literature cited above, the authors confine consideration to diffusion-type target strategies which, at least asymptotically, are equivalent to our averaged targets. Our approach, by contrast, allows one to deal with general predictable frictionless target strategies and so the examples considered in this paper include strategies with jumps or even singularities where the differences between these hedges become apparent.

Almgren and Li [3] study a quite similar hedging problem but they consider a model with permanent price impact which feeds into their target strategies via the well-known functions for Black–Scholes deltas and gammas. Hence, they consider a model where the target strategy is also affected by the targeting strategy which leads to a feedback effect that we are disregarding in our problem formulation. We refer to the introduction in Rogers and Singh [24] for further discussion of this idealization.

The rest of the paper is organized as follows. In Sect. 2 we introduce the setup and motivate our problem formulation by following the approach of Rogers and Singh [24] (cf.

also Naujokat and Westray [21]). Our main result is presented in Sect. 3 and provides the explicit solution for a general hedging problem of a European style option in a Bachelier market model with temporary price impact. Section 4 contains some illustrations of optimal solutions in three examples. The technical proofs are deferred to Sect. 5.

2 Problem setup and motivation

We fix a finite deterministic time horizon $T > 0$, a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0 \le t \le T}, \mathbb{P})$ satisfying the usual conditions of right continuity and completeness and consider an agent who is trading in a financial market consisting of a risky asset, e.g., stock. The number of shares the agent holds at time $t \in [0, T]$ of the risky stock is defined as

$$X_t^u \triangleq x + \int_0^t u_s ds, \quad 0 \le t \le T, \tag{1}$$

where $x \in \mathbb{R}$ denotes her given initial holdings. The real-valued stochastic process $(u_t)_{0 \le t \le T}$ represents the agent's turnover rate, that is, the speed at which the agent trades in the risky asset. It is assumed to be chosen in the set

$$\mathscr{U} \triangleq \left\{ u : u \text{ progressively measurable s.t. } \mathbb{E} \int_0^T u_t^2 dt < \infty \right\}.$$

The square-integrability requirement ensures that the induced quadratic transaction costs which are levied on the agent's respective turnover rates due to temporary price impact as in Almgren and Chriss [2] are finite.

In such a frictional market, our agent seeks to track a target strategy which can be thought of, for instance, as a hedging strategy adopted from a frictionless setting. Mathematically, this problem can be formalized as follows: given a real-valued predictable process $(\xi_t)_{0 \le t \le T}$ in $L^2(\mathbb{P} \otimes dt)$ and a fixed constant $\kappa > 0$, the agent's objective is to minimize the performance functional

$$J(u) \triangleq \mathbb{E} \left[\frac{1}{2} \int_0^T (X_t^u - \xi_t)^2 dt + \frac{1}{2} \kappa \int_0^T u_t^2 dt \right]. \tag{2}$$

This leads to the optimal stochastic control problem

$$J(u) \to \min_{u \in \mathscr{U}}. \tag{3}$$

Since the agent's terminal position X_T^u may be important (for her future plans or physical delivery), we also consider the optimal stochastic control problem

$$J(u) \to \min_{u \in \mathscr{U}_x^\Xi} \tag{4}$$

where $\mathscr{U}_x^\Xi$ is the set of constrained policies defined as

$$\mathscr{U}_x^\Xi \triangleq \left\{ u : u \in \mathscr{U} \text{ satisfying } X_T^u \equiv x + \int_0^T u_s ds = \Xi_T \ \mathbb{P}\text{-a.s.} \right\}$$

with predetermined terminal position $\Xi_T \in L^2(\mathbb{P}, \mathscr{F}_T)$ such that

$$\int_0^T \frac{d\mathbb{E}[\Xi_t^2]}{T - t} < \infty \tag{5}$$

where $\Xi_t \triangleq \mathbb{E}[\Xi_T | \mathscr{F}_t]$ for $0 \le t \le T$.

Remark 2.1 (1) Lemma 5.4 below shows that a target Ξ_T can be reached with finite expected costs in the sense that $\mathscr{U}_x^\Xi \neq \varnothing$ if and only if (5) is satisfied. Observe that this condition implies, in particular, that $\Xi_T \in \mathscr{F}_{T-}$. In fact, (5) can be interpreted as a condition on the speed at which one learns about the ultimate target position Ξ_T as $t \uparrow T$.

(2) Concerning physical delivery at maturity T, it would be sufficient to impose the constraint $X_T^u \geq \Xi_T$. However, this would lead to an interesting, yet technically rather different optimization problem which is left for future research.

One motivation of the objective functional in (2) and its connection to the problem of hedging a European contingent claim in the presence of temporary price impact is the following (cf. also Rogers and Singh [24] and Almgren and Li [3]): Assume the agent wants to hedge a European-type option with payoff H at maturity T in a market where, for simplicity, interest rates are zero and the price process S of the underlying risky asset follows a Brownian motion with volatility $\sigma > 0$:

$$S_t = S_0 + \sigma W_t, \quad 0 \leq t \leq T.$$

In a frictionless setting, the payoff H can be perfectly replicated by a predictable hedging strategy ξ^H. In a market with frictions where the agent faces liquidity costs as, e.g., in Almgren and Chriss [2], she may be confined to follow strategies X^u as in (1). As a consequence, starting from some initial wealth $v_0 \in \mathbb{R}$, her profits and losses from market fluctuations will incur a risk of deviating from H at maturity T that can be measured, e.g., by

$$\mathbb{E}\left[\left(H - \left(v_0 + \int_0^T X_t^u \, dS_t\right)\right)^2\right] = (\mathbb{E}[H] - v_0)^2 + \mathbb{E}\left[\int_0^T (X_t^u - \xi_t^H)^2 \sigma^2 dt\right],$$

see Föllmer and Sondermann [9]. This deviation can be made arbitrarily small if the agent is willing to incur arbitrarily high transaction costs. If, however, she puts a cap $c > 0$ on these she may want to solve the optimization problem

$$\mathbb{E}\left[\int_0^T (X_t^u - \xi_t^H)^2 \sigma^2 dt\right] \to \min_{u \in \mathscr{U}} \quad \text{subject to } \mathbb{E}\left[\int_0^T u_t^2 dt\right] \leq c, \qquad (6)$$

which in its Lagrangian formulation amounts to an objective functional of the form (2).

Remark 2.2 1. A similar hedging problem as formulated in (2) is also studied in Rogers and Singh [24] and Almgren and Li [3]. In contrast to our setting, Rogers and Singh [24] consider a Black–Scholes framework. Almgren and Li [3] also include permanent price impact.

2. Apart from hedging, the minimization problem of the objective in (2) is also related to the problem of optimally executing a VWAP order as studied using dynamic programming methods in a Markovian setup in Cartea and Jaimungal [7] and Frei and Westray [10], or, more generally, to the optimal curve following problem as discussed in Naujokat and Westray [21] as well as Cai et al. [6].

3. In a Brownian setting, our problem (3) is a special case of a stochastic linear-quadratic control problem as studied, e.g., by Kohlmann and Tang [18].

3 Main result

Our main results are the following explicit descriptions of the optimal controls for problems (3) and (4) and their corresponding minimal costs for which it is convenient to introduce

$$\tau^\kappa(t) \triangleq \frac{T-t}{\sqrt{\kappa}}, \quad 0 \leq t \leq T.$$

Theorem 3.1 *The optimal stock holdings $\hat{X}$ of problem (3) with unconstrained terminal position satisfy the linear ODE*

$$d\hat{X}_t = \frac{\tanh(\tau^\kappa(t))}{\sqrt{\kappa}} \left(\hat{\xi}_t - \hat{X}_t \right) dt, \quad \hat{X}_0 = x, \tag{7}$$

where, for $0 \leq t < T$, we let

$$\hat{\xi}_t \triangleq \mathbb{E}\left[\int_t^T \xi_u K(t,u) du \,\bigg|\, \mathscr{F}_t \right]$$

with the kernel

$$K(t,u) \triangleq \frac{\cosh(\tau^\kappa(u))}{\sqrt{\kappa}\sinh(\tau^\kappa(t))}, \quad 0 \leq t \leq u < T.$$

The minimal costs are given by

$$\inf_{u \in \mathscr{U}} J(u) = \frac{1}{2}\sqrt{\kappa}\tanh(\tau^\kappa(0)) \left(x - \hat{\xi}_0 \right)^2 + \frac{1}{2}\mathbb{E}\left[\int_0^T (\xi_t - \hat{\xi}_t)^2 dt \right]$$
$$+ \frac{1}{2}\mathbb{E}\left[\int_0^T \sqrt{\kappa}\tanh(\tau^\kappa(t)) d\langle\hat{\xi}\rangle_t \right] < \infty. \tag{8}$$

For the constrained problem we have similarly:

Theorem 3.2 *The optimal stock holdings $\hat{X}^\Xi$ of problem (4) with constrained terminal position $\Xi_T \in L^2(\mathbb{P}, \mathscr{F}_T)$ such that (5) holds satisfy the linear ODE*

$$d\hat{X}_t^\Xi = \frac{\coth(\tau^\kappa(t))}{\sqrt{\kappa}} \left(\hat{\xi}_t^\Xi - \hat{X}_t^\Xi \right) dt, \quad \hat{X}_0^\Xi = x, \tag{9}$$

where, for $0 \leq t \leq T$, we let

$$\hat{\xi}_t^\Xi \triangleq \mathbb{E}\left[\frac{1}{\cosh(\tau^\kappa(t))}\Xi_T + \left(1 - \frac{1}{\cosh(\tau^\kappa(t))} \right) \int_t^T \xi_u K^\Xi(t,u) du \,\bigg|\, \mathscr{F}_t \right],$$

with the kernel

$$K^\Xi(t,u) \triangleq \frac{\sinh(\tau^\kappa(u))}{\sqrt{\kappa}(\cosh(\tau^\kappa(t)) - 1)}, \quad 0 \leq t \leq u < T.$$

The solution $\hat{X}^\Xi$ of (9) satisfies the terminal constraint in the sense that

$$\lim_{t \uparrow T} \hat{X}_t^\Xi = \Xi_T \quad \mathbb{P}\text{-a.s.}$$

The minimal costs are given by

$$\inf_{u \in \mathscr{U}^\Xi} J(u) = \frac{1}{2}\sqrt{\kappa}\coth(\tau^\kappa(0)) \left(x - \hat{\xi}_0^\Xi \right)^2 + \frac{1}{2}\mathbb{E}\left[\int_0^T (\xi_t - \hat{\xi}_t^\Xi)^2 dt \right]$$
$$+ \frac{1}{2}\mathbb{E}\left[\int_0^T \sqrt{\kappa}\coth(\tau^\kappa(t)) d\langle\hat{\xi}^\Xi\rangle_t \right] < \infty. \tag{10}$$

The convex-analytic proofs of Theorems 3.1 and 3.2 are deferred to Sect. 5.

Note that, rather than towards the *current* target position ξ_t, the optimal frictional hedging rules in (7) and (9) prescribe to trade towards weighted averages $\hat{\xi}_t$ and $\hat{\xi}_t^{\Xi}$, respectively, of *expected future* target positions of ξ. Indeed, for each $0 \leq t \leq T$, $K(t, .)$ and $K^{\Xi}(t, .)$ specify nonnegative kernels which integrate to one over $[t, T]$, and so $\hat{\xi}$ and $\hat{\xi}^{\Xi}$ average out the expected future positions of ξ. For $\hat{\xi}^{\Xi}$ one chooses a convex combination of this average of ξ with the expected terminal position Ξ_T, where the weight shifts gradually to Ξ_T as $t \uparrow T$ since $1/\cosh(\tau^k(t)) \uparrow 1$ in that case.

According to Eqs. (7) and (9), the optimal tracking rate trades towards these targets at a speed proportional to their distance to the investor's position at any time. The coefficient of proportionality is controlled by both the cost parameter κ and the remaining time-to-maturity $T - t$. For the unconstrained solution in (7), since $\lim_{t \uparrow T} \tanh(\tau^{\kappa}(t)) = 0$, trading slows down when approaching the final time T; in other words, towards the end, the investor does not worry about tracking ξ anymore, but seeks to minimize trading costs. This becomes intuitive when comparing the effect of early interventions to later ones: with early interventions the investor ensures that she stays reasonably close to the target for the foreseeable future, but late interventions only can impact the investor's performance for very short periods and therefore do not warrant, at least asymptotically, the associated costs. For the constrained solution in (9) by contrast, we have $\lim_{t \uparrow T} \coth(\tau^{\kappa}(t)) = +\infty$ and so the optimal strategy trades with increased urgency towards $\hat{\xi}^{\Xi}$, which itself is easily seen to converge to the ultimate target position $\Xi_T = \lim_{t \uparrow T} \hat{\xi}_t^{\Xi}$ $\mathbb{P}$-a.s. (cf. Proof of Theorem 3.2 in Sect. 5).

Our tracking result generalizes an observation from Gârleanu and Pedersen [12] from their homogeneous Markovian optimal investment problem to a general hedging problem with a general predictable target strategy ξ, also allowing for a random terminal portfolio position Ξ_T. It also sheds further light on the general structure of optimal portfolio strategies in markets with frictions. Indeed, the description of (asymptotically) optimal trading strategies obtained in Kallsen and Muhle-Karbe [17], Moreau et al. [20], or Guasoni and Weber [14,15] prescribe a reversion towards the frictionless strategy ξ itself, *not* towards an average such as $\hat{\xi}$ or $\hat{\xi}^{\Xi}$. For sufficiently smooth ξ, e.g., of diffusion type, this is still optimal asymptotically for small liquidity costs as then these averages do not differ significantly from ξ. The next section, however, shows that this is no longer the case when we allow for singularities in the reference strategy.

Finally, our representations (8) and (10) for the values of the tracking problems (3) and (4), respectively, show how these depend on the initial position x and the L^2-distance between the target ξ and the respective signal processes $\hat{\xi}$ and $\hat{\xi}^{\Xi}$. It also reveals the importance of the signals' quadratic variation $\langle \hat{\xi} \rangle$, $\langle \hat{\xi}^{\Xi} \rangle$ which can be viewed as a measure for how effectively one can predict the target positions ξ and Ξ_T. To the best of our knowledge, the key role played by the signals $\hat{\xi}$, $\hat{\xi}^{\Xi}$ was not observed in the general theory of stochastic linear-quadratic control problems as discussed, e.g., by Kohlmann and Tang [18].

Remark 3.1 As mentioned in the description of our problem setup in Sect. 2, the quadratic cost term in our objective function in (2) is due to linear temporary price impact as in the model proposed by Almgren and Chriss [2]. In this regard, one might likewise extend the objective functional also in order to account for expected costs resulting from linear *permanent price impact* (cf. in Almgren and Chriss [2]). This would lead to the inclusion of the additional term

$$\mathbb{E}\left[\theta\left(\int_0^T u_t \, dt\right)^2\right] = \theta \mathbb{E}\left[(X_T^u - x)^2\right] \qquad (11)$$

for some constant $\theta > 0$. For the constrained problem in (4), this extra cost term obviously does not depend on the strategy and is thus irrelevant. For the unconstrained problem in (3), these extra costs can be regarded as a penalization term forcing the final position X_T^u to be close to the initial position x. For ease of exposition, we refrain in the present paper from inducing this additional term, since our main intention here is to outline the key role played by the optimal tracking signals $\hat{\xi}$, $\hat{\xi}^\Xi$ in the description of the optimal control as well as the corresponding minimal costs. A more general setup allowing for stochastic price impact, stochastic volatility and a penalization on the terminal position as in (11) is left for future research.

4 Illustrations

In this section we present a few case studies illustrating the structure of the optimal hedging strategies we found in Theorems 3.1 and 3.2. The first two case studies are simple deterministic toy examples which allow us to understand the effect of jumps as well as of initial and terminal positions. The final case study considers a discretely monitored Asian option where random jumps in the reference hedge occur naturally.

In the first two cases we assume the initial position to be $x = 0$ and consider a time horizon of $T = 1$ when, in the constrained case, the position has to be liquidated, i.e., $\Xi_T = 0$. We depict ξ along with its averages $\hat{\xi}$ and $\hat{\xi}^\Xi$, respectively, as well as the corresponding optimal frictional hedges $\hat{X}$ and $\hat{X}^\Xi$. We also include a "myopic" benchmark strategy $\tilde{X}$ which directly targets ξ (without final constraint) given by

$$d\tilde{X}_t = \frac{1}{\sqrt{\kappa}}(\xi_t - \tilde{X}_t)dt, \quad 0 \leq t \leq T,$$

in order to compare with analogous strategies considered in Guasoni and Weber [14,15], Moreau et al. [20], Rogers and Singh [24], and Kallsen and Muhle-Karbe [17].
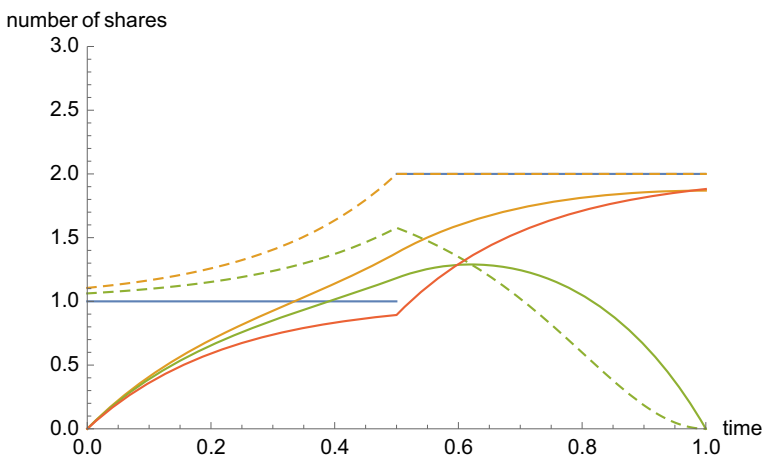

Fig. 1 Frictionless hedge ξ with a jump at $t = T/2$ (*blue*), corresponding unconstrained (*orange*, *dashed*) and constrained (*green*, *dashed*) targets $\hat{\xi}$ and $\hat{\xi}^\Xi$, respectively, as well as the corresponding frictional hedges $\hat{X}$ (*orange line*) and $\hat{X}^\Xi$ (*green line*). The myopic benchmark hedge $\tilde{X}$ is plotted in *red*
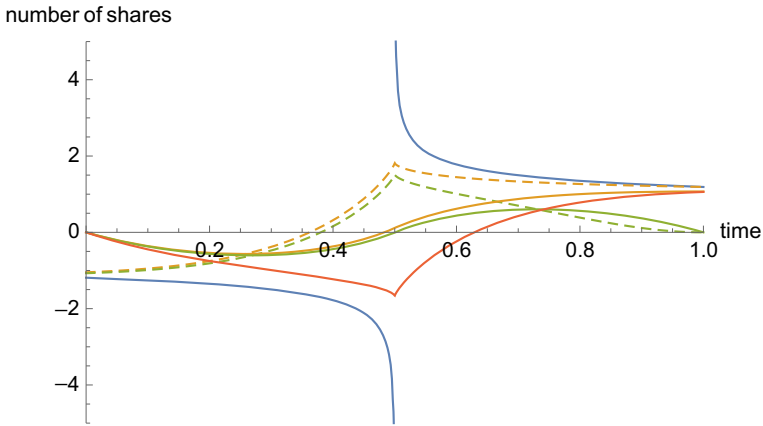
number of shares

Fig. 2 Frictionless hedge ξ with a singularity at $t = T/2$ (*blue*), corresponding unconstrained (*orange, dashed*) and constrained (*green, dashed*) targets $\hat{\xi}$ and $\hat{\xi}^\Xi$, respectively, as well as the corresponding frictional hedges $\hat{X}$ (*orange line*) and $\hat{X}^\Xi$ (*green line*). The myopic benchmark hedge $\tilde{X}$ is plotted in *red*

4.1 Frictionless deterministic hedge with a jump

In our first case study we consider a deterministic target strategy ξ (solid blue line in Fig. 1) which can be viewed as a stock-buying schedule that prescribes to hold one stock until time $T/2$ when the position is doubled by a jump. One can observe that the effective target strategies $\hat{\xi}$ and $\hat{\xi}^\Xi$ of the optimal controls $\hat{u}$ and $\hat{u}^\Xi$, respectively, are smoothing out the jump of ξ. The target $\hat{\xi}^\Xi$ additionally takes into account the liquidation constraint $\Xi_T = 0$ of the agent's position until maturity T. As expected, the optimal frictional hedges $\hat{X}$ and $\hat{X}^\Xi$ are indeed anticipating the upward jump of the target strategy ξ at $t = T/2$ by building up their positions beyond the actual current position of ξ even before the occurrence of the jump. This is not the case for the myopic benchmark strategy $\tilde{X}$ which increases its position much more slowly and exhibits a kink when the jump occurs after which trading speed picks up significantly. Finally, the optimal holdings $\hat{X}^\Xi$ in the constrained setting, where the position has to be unwound ultimately, are decreasing when time approaches maturity and end up in the final desired position $\hat{X}_T^\Xi = 0$.

4.2 Frictionless deterministic hedge with a singularity

The second target strategy ξ (solid blue line in Fig. 2) is again deterministic and also exhibits a singularity midway at $t = T/2$, this time, however, it is a jump from $-\infty$ to $+\infty$. Once more, one can observe that the effective target strategies $\hat{\xi}$ and $\hat{\xi}^\Xi$ of the optimal controls $\hat{u}$ and $\hat{u}^\Xi$, respectively, are smoothing out the singularity of ξ. Again, the target $\hat{\xi}^\Xi$ additionally takes into account the liquidation constraint $\Xi_T = 0$ of the agent's position until maturity T. In contrast to the benchmark strategy $\tilde{X}$, the optimal frictional hedges $\hat{X}$ and $\hat{X}^\Xi$ are anticipating the singularity of the target strategy ξ at $t = T/2$ by gradually building up their positions before the singularity occurs. Actually, they are trading *away* from the current target positions of ξ for some time prior to $T/2$. This is in stark contrast with the myopic benchmark strategy which keeps selling short more and more intensely even milliseconds before the reference strategy jumps to $+\infty$.

4.3 Discrete Asian option

In this final example we investigate a situation where the target strategy ξ is stochastic and exhibits a random jump. Specifically, we consider hedging a discrete Asian call with maturity $T > 0$ in the Bachelier model where the underlying risky asset S is modeled by a Brownian motion with volatility $\sigma > 0$:

$$S_t = S_0 + \sigma W_t, \quad 0 \le t \le T.$$

For simplicity, we assume that the average is discretely monitored over two fixing dates $T/2$ and T. That is, the payoff at maturity T is given by

$$H \triangleq \left(\frac{1}{2}(S_{T/2} + S_T) - K \right)^+$$

for some strike $K \in \mathbb{R}$. The Bachelier price of the discrete Asian option at time $t \in [0, T)$ can be computed as

$$\pi_t \triangleq \begin{cases} \sigma\sqrt{5T/8 - t}\, \varphi\left(\frac{S_t - K}{\sigma\sqrt{5T/8 - t}}\right) + S_t \Phi\left(\frac{S_t - K}{\sigma\sqrt{5T/8 - t}}\right), & 0 \le t < T/2 \\ \frac{1}{2}\sigma\sqrt{T-t}\, \varphi\left(\frac{S_{T/2} + S_t - 2K}{\sigma\sqrt{T-t}}\right) + \left(\frac{1}{2}(S_{T/2} + S_t) - K\right)\Phi\left(\frac{S_{T/2} + S_t - 2K}{\sigma\sqrt{T-t}}\right), & T/2 \le t < T \end{cases}$$

where φ and Φ denote the density and the cumulative distribution function of the standard normal distribution, respectively. Accordingly, the frictionless delta-hedging strategy is

$$\xi_t = \begin{cases} \Phi\left(\frac{S_t - K}{\sigma\sqrt{5T/8 - t}}\right), & 0 \le t \le T/2 \\ \frac{1}{2}\Phi\left(\frac{S_{T/2} + S_t - 2K}{\sigma\sqrt{T-t}}\right), & T/2 < t < T. \end{cases}$$

Note that the delta-hedge exhibits a negative random jump at time $T/2$ since

$$\xi_{\frac{T}{2}+} - \xi_{\frac{T}{2}-} \triangleq \lim_{t \downarrow \frac{T}{2}} \xi_t - \lim_{t \uparrow \frac{T}{2}} \xi_t = -\frac{1}{2}\Phi\left(\frac{S_{T/2} - K}{\sigma\sqrt{T/8}}\right).$$

We assume that the initial position x coincides with the initial frictionless delta, i.e., e.g., $x = 1/2$ in the case of an at-the-money option with $K = S_0$. This allows us to focus on the hedging performance itself and avoids distortions from the initial built up of a sensible hedging position. As before, the terminal position will be allowed to be either unconstrained or mandating liquidation, i.e., $\Xi_T = 0$.

The effective targets $\hat{\xi}$ and $\hat{\xi}^\Xi$ of the optimal frictional hedging strategy in (7) and (9), respectively, can be explicitly computed:

$$\hat{\xi}_t = \begin{cases} \Phi\left(\frac{2(S_t - K)}{\sigma\sqrt{5T/2 - 4t}}\right)\left(1 - \frac{1}{2}\frac{\sinh(\tau^\kappa(T/2))}{\sinh(\tau^\kappa(t))}\right), & 0 \le t < T/2, \\ \xi_t, & T/2 \le t < T, \end{cases}$$

and

$$\hat{\xi}_t^\Xi = \begin{cases} \Phi\left(\frac{2(S_t - K)}{\sigma\sqrt{5T/2 - 4t}}\right)\left(1 - \frac{1}{2}\frac{\cosh(\tau^\kappa(T/2)) + 1}{\cosh(\tau^\kappa(t))}\right), & 0 \le t < T/2 \\ \left(1 - \frac{1}{\cosh(\tau^\kappa(t))}\right)\xi_t, & T/2 \le t < T. \end{cases}$$

Observe that the Bachelier delta-hedge ξ is a martingale on $[T/2, T]$ and thus the signal $\hat{\xi}$ coincides with it in this period. However, the optimal target $\hat{\xi}$ differs from the frictionless hedge ξ on $[0, T/2]$ since it is anticipating and systematically smoothing out the random jump at $T/2$ whose size is determined by the option's moneyness at this point. The constrained
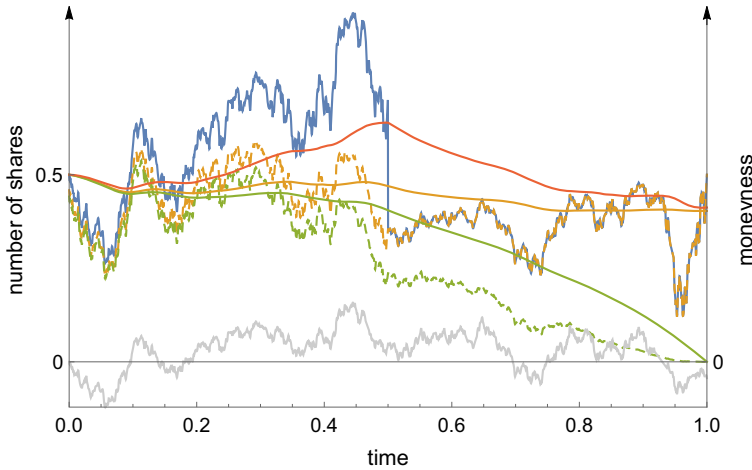
Fig. 3 Frictionless hedge ξ with a jump at $t = T/2$ (*blue*), corresponding unconstrained (*orange*, *dashed*) and constrained (*green*, *dashed*) targets $\hat{\xi}$ and $\hat{\xi}^{\Xi}$, respectively, as well as the corresponding frictional hedges $\hat{X}$ (orange line) and $\hat{X}^{\Xi}$ (*green line*). The myopic benchmark hedge $\tilde{X}$ is plotted in *red*. The moneyness is indicated by the *light gray line*

target $\hat{\xi}^{\Xi}$ anticipates the liquidation requirement at maturity which plays a more and more dominating role after time $T/2$.

Again, the myopic benchmark strategy

$$d\tilde{X}_t = \frac{\sigma}{\sqrt{\kappa}}(\xi_t - \tilde{X}_t)dt, \quad 0 \le t < T$$

is not taking into account the random jump at time $T/2$ and keeps on tracking the frictionless delta-hedge even milliseconds before $T/2$ (see Fig. 3).

5 Proofs

In order to prove our main Theorems 3.1 and 3.2 we use tools from convex analysis. Note that the performance functional $u \mapsto J(u)$ in (2) is strictly convex. Given a control $u \in \mathscr{U}$ recall the definition of the Gâteaux derivative of J at u in the direction of $w \in L^2(\mathbb{P} \otimes dt)$:

$$\langle J'(u), w \rangle \triangleq \lim_{\rho \to 0} \frac{J(u + \rho w) - J(u)}{\rho}.$$

The following lemma provides an explicit expression for the Gâteaux derivative of our performance functional J:

Lemma 5.1 *For* $u \in \mathscr{U}$ *we have*

$$\langle J'(u), w \rangle = \mathbb{E}\left[\int_0^T w_s \left(\kappa u_s + \int_s^T (X_t^u - \xi_t)dt\right)ds\right]$$

for any $w \in L^2(\mathbb{P} \otimes dt)$.

Proof Let $\rho > 0$, $u \in \mathscr{U}$ and $w \in L^2(\mathbb{P} \otimes dt)$. Note that $X_t^{u+\rho w} = X_t^u + \rho \int_0^t w_s ds$. Then, we have

$$J(u + \rho w) - J(u) = \rho \mathbb{E}\left[\int_0^T \kappa u_t w_t + \left(\int_0^t w_s ds\right)(X_t^u - \xi_t)dt\right]$$
$$+ \rho^2 \mathbb{E}\left[\frac{\kappa}{2}\int_0^T w_t^2 dt + \frac{1}{2}\int_0^T \left(\int_0^t w_s ds\right)^2 dt\right].$$

Hence,

$$\langle J'(u), w\rangle = \mathbb{E}\left[\int_0^T \kappa u_t w_t + \left(\int_0^t w_s ds\right)(X_t^u - \xi_t)dt\right].$$

Note that due to Fubini's Theorem we can write the second part of the above integral as

$$\int_0^T \left(\int_0^t w_s ds\right)(X_t^u - \xi_t)dt = \int_0^T \left(\int_s^T (X_t^u - \xi_t)dt\right) w_s ds$$

which finally yields the assertion.   □

Let us next derive necessary and sufficient first order conditions for problems (3) and (4).

Lemma 5.2 (First order conditions)

1. *In the unconstrained problem (3), a control $\hat{u} \in \mathscr{U}$ with $X \triangleq X^{\hat{u}}$ minimizes the functional J if and only if X satisfies*

$$X_0 = x, \quad d\dot{X}_t = \frac{1}{\kappa}(X_t - \xi_t)dt + dM_t \text{ for } 0 \leq t \leq T, \quad \dot{X}_T = 0, \qquad (12)$$

   for a suitable square integrable martingale $(M_t)_{0 \leq t \leq T}$.
2. *In the constrained problem (4), a control $\hat{u} \in \mathscr{U}_x^{\Xi}$ with $X \triangleq X^{\hat{u}}$ minimizes the functional J if and only if X satisfies*

$$X_0 = x, \quad d\dot{X}_t = \frac{1}{\kappa}(X_t - \xi_t)dt + dM_t \text{ for } 0 \leq t < T, \quad X_T = \Xi_T, \qquad (13)$$

   for a suitable square integrable martingale $(M_t)_{0 \leq t < T}$.

In other words, the first order conditions in (12) and (13) are taking the form of a coupled linear forward backward stochastic differential equation (FBSDE) for the pair (X, u):

$$dX_t = u_t dt,$$
$$du_t = \frac{1}{\kappa}(X_t - \xi_t)dt + dM_t,$$

with some square integrable martingale M subject to

$$X_0 = x \text{ and } \begin{cases} u_T = 0 & \text{unconstrained case,} \\ X_T = \Xi_T & \text{constrained case.} \end{cases}$$

Proof (1) We start with the unconstrained problem (3). Since we are minimizing the strictly convex functional $u \mapsto J(u)$ over $\mathscr{U}$, a necessary and sufficient condition for the optimality of $\hat{u} \in \mathscr{U}$ with corresponding $X^{\hat{u}} = x + \int_0^{\cdot} \hat{u}_s ds$ is given by

$$\langle J'(\hat{u}), w\rangle = 0 \text{ for all } w \in \mathscr{U}$$

(cf., e.g., Ekeland and Témam [8]). In view of Lemma 5.1 this means that $\hat{u} \in \mathscr{U}$ is optimal if and only if

$$\mathbb{E}\left[\int_0^T w_s \left(\kappa \hat{u}_s + \int_s^T (X_t^{\hat{u}} - \xi_t)dt\right) ds\right] = 0 \qquad (14)$$

for all $w \in \mathscr{U}$. We will now show that the first order condition in (14) is satisfied (i.e., $\hat{u} \in \mathscr{U}$ is optimal) if and only if $X^{\hat{u}}$ satisfies the dynamics in (12).

Necessity Assume that $\hat{u} \in \mathscr{U}$ with $X^{\hat{u}} = x + \int_0^\cdot \hat{u}_s ds$ minimizes J, i.e., condition (14) is satisfied by $\hat{u}$. Then, by Fubini's Theorem and optional projection, we also get that

$$\mathbb{E}\left[\int_0^T w_s \left(\kappa \hat{u}_s + \mathbb{E}\left[\int_s^T (X_t^{\hat{u}} - \xi_t)dt \,\middle|\, \mathscr{F}_s\right]\right) ds\right] = 0$$

for all $w \in \mathscr{U}$. However, this is only possible if

$$\hat{u}_s = -\frac{1}{\kappa}\mathbb{E}\left[\int_s^T (X_t^{\hat{u}} - \xi_t)dt \,\middle|\, \mathscr{F}_s\right] \quad d\mathbb{P} \otimes ds\text{-a.e. on } \Omega \times [0, T]. \qquad (15)$$

Hence, by defining the square integrable martingale

$$M_s \triangleq \mathbb{E}\left[\int_0^T (X_t^{\hat{u}} - \xi_t)dt \,\middle|\, \mathscr{F}_s\right], \quad 0 \le s \le T, \qquad (16)$$

we obtain the representation

$$\hat{u}_s = -\frac{1}{\kappa}\left(M_s - \int_0^s (X_t^{\hat{u}} - \xi_t)dt\right) \quad d\mathbb{P} \otimes ds\text{-a.e. on } \Omega \times [0, T], \qquad (17)$$

in other words, $X^{\hat{u}}$ satisfies the dynamics in (12). In particular, $X_0^{\hat{u}} = x$ and $\dot{X}_T^{\hat{u}} = \hat{u}_T = 0$ $\mathbb{P}$-a.s.

Sufficiency Assume now that $\hat{u} \in \mathscr{U}$ with corresponding $X^{\hat{u}}$ satisfies the dynamics in (12) with $X_0^{\hat{u}} = x$ and $\dot{X}_T^{\hat{u}} = 0$ $\mathbb{P}$-a.s. Note that the unique strong solution to this linear FBSDE in (12) is indeed given by (15) or, equivalently, by (17). However, using this representation of $\hat{u}$ and applying Fubini's Theorem yields

$$\mathbb{E}\left[\int_0^T w_s \left(\kappa \hat{u}_s + \int_s^T (X_t^{\hat{u}} - \xi_t)dt\right) ds\right] = \mathbb{E}\left[\int_0^T w_s (M_T - M_s) ds\right]$$

$$= \mathbb{E}\left[\int_0^T w_s \mathbb{E}[M_T - M_s | \mathscr{F}_s] ds\right] = \int_0^T \mathbb{E}\left[w_s \left(\mathbb{E}[M_T | \mathscr{F}_s] - M_s\right)\right] ds = 0$$

for all $w \in \mathscr{U}$, since M is a martingale. Consequently, the first order condition in (14) is satisfied and $\hat{u} \in \mathscr{U}$ is optimal.

(2) Similar as above, a necessary and sufficient condition for the optimality of $\hat{u}^\Xi \in \mathscr{U}_x^\Xi$ with corresponding $X^{\hat{u}^\Xi} = x + \int_0^\cdot \hat{u}_s^\Xi ds$ satisfying $X_T^{\hat{u}^\Xi} = \Xi_T$ $\mathbb{P}$-a.s. for the constrained problem (4) is given by

$$\langle J'(\hat{u}^\Xi), w \rangle = 0 \text{ for all } w \in \mathscr{U}_0^0.$$

In contrast to the unconstrained case, observe now that we have an additional constraint on w. Again, in view of Lemma 5.1, we get that $\hat{u}^\Xi \in \mathscr{U}_x^\Xi$ is optimal if and only if

$$\mathbb{E}\left[\int_0^T w_s \left(\kappa \hat{u}_s^\Xi + \int_s^T (X_t^{\hat{u}^\Xi} - \xi_t)dt\right) ds\right] = 0 \text{ for all } w \in \mathscr{U}_0^0. \qquad (18)$$

We will now show that the first order condition in (18) is fulfilled (i.e., $\hat{u}^\Xi \in \mathcal{U}_x^\Xi$ is optimal) if and only if $X^{\hat{u}^\Xi}$ satisfies the dynamics in (13).

Sufficiency Assume that $\hat{u}^\Xi \in \mathcal{U}_x^\Xi$ with corresponding $X^{\hat{u}^\Xi}$ satisfies the dynamics in (13) with $X_0^{\hat{u}^\Xi} = x$ and $X_T^{\hat{u}^\Xi} = \Xi_T$ $\mathbb{P}$-a.s. That is, we have the representation

$$\hat{u}_t^\Xi = \hat{u}_0^\Xi + \frac{1}{\kappa} \int_0^t (X_s^{\hat{u}^\Xi} - \xi_s) ds + M_t \quad d\mathbb{P} \otimes dt\text{-a.e. on } \Omega \times [0, T)$$

for some square integrable martingale $(M_t)_{0 \le t < T}$. From $\hat{u}^\Xi, \xi \in L^2(\mathbb{P} \otimes dt)$, it follows that $\mathbb{E}[\int_0^T M_s^2 ds] < \infty$. Defining the square integrable martingale

$$N_s^\Xi \triangleq \mathbb{E}\left[\int_0^T (X_t^{\hat{u}^\Xi} - \xi_t) dt \,\middle|\, \mathscr{F}_s\right], \quad 0 \le s \le T,$$

the above representation of $\hat{u}^\Xi$ yields

$$\mathbb{E}\left[\int_0^T w_s \left(\kappa \hat{u}_s^\Xi + \int_s^T (X_t^{\hat{u}^\Xi} - \xi_t) dt\right) ds\right]$$
$$= \mathbb{E}\left[\int_0^T w_s \left(\kappa \hat{u}_0^\Xi + N_T^\Xi + \kappa M_s\right) ds\right]$$
$$= \mathbb{E}\left[(\kappa \hat{u}_0^\Xi + N_T^\Xi) \int_0^T w_s ds\right] + \kappa \mathbb{E}\left[\int_0^T w_s M_s ds\right]$$
$$= 0 \text{ for all } w \in \mathcal{U}_0^0$$

by virtue of Lemma 5.3 below. Consequently, the first order condition in (18) is satisfied and $\hat{u}^\Xi \in \mathcal{U}_x^\Xi$ is optimal.

Necessity As shown in the proof of Theorem 3.2 below (which does *not* use the necessity assertion of the present lemma), the optimal control $\hat{u}^\Xi$ in (9) satisfies the dynamics in (13). Moreover, by strict concavity of the objective functional in (2), the solution to problem (4) is unique. Therefore, the assertion is indeed necessary. □

The following technical lemma is needed in the proof of Lemma 5.2 for the constrained problem (3).

Lemma 5.3 *Let M be an adapted càdlàg process on $[0, T)$ with $\mathbb{E}[\int_0^T M_s^2 ds] < \infty$. Then,*

$$\mathbb{E}\left[\int_0^T w_s M_s ds\right] = 0 \text{ for all } w \in \mathcal{U}_0^0 \tag{19}$$

if and only if M is a square integrable martingale on $[0, T)$.

Proof First, assume that M is a square integrable martingale on $[0, T)$ with $\mathbb{E}[\int_0^T M_s^2 ds] < \infty$. Consider a $w \in \mathcal{U}_0^0$ such that $w = 0$ on $\Omega \times [T - \varepsilon, T]$ for some $\varepsilon > 0$. Then, by applying Fubini's Theorem we have

$$\mathbb{E}\left[\int_0^T w_s M_s ds\right] = \mathbb{E}\left[\int_0^{T-\varepsilon} w_s \mathbb{E}[M_{T-\varepsilon} | \mathscr{F}_s] ds\right] = \mathbb{E}\left[M_{T-\varepsilon} \int_0^T w_s ds\right] = 0.$$

Now, let $w \in \mathcal{U}_0^0$ be arbitrary and consider an approximating sequence $(w^{(n)})_{n \ge 1} \subset \mathcal{U}_0^0$ with $w^{(n)} = 0$ on $\Omega \times [T - \varepsilon_n, T]$ for some $\varepsilon_n \downarrow 0$ such that $w^{(n)} \to w$ in $L^2(\Omega \times [0, T], \mathbb{P} \otimes dt)$

for $n \to \infty$. Then, by the Cauchy-Schwarz inequality we obtain

$$\lim_{n \to \infty} \mathbb{E} \left[\int_0^T |(w_s^{(n)} - w_s) M_s| ds \right] = 0.$$

Consequently,

$$\mathbb{E} \left[\int_0^T w_s M_s ds \right] = \lim_{n \to \infty} \mathbb{E} \left[\int_0^T w_s^{(n)} M_s ds \right] = 0,$$

where the last identity follows from our initial consideration for ws with support in $[T - \varepsilon, T]$, $\varepsilon > 0$. Hence, the condition in (19) is satisfied.

Conversely, assume now that the condition in (19) is satisfied. We have to show that M is a square integrable martingale on $[0, T)$. Let $0 \le t < u < T$, $A \in \mathscr{F}_t$, be arbitrary. For any $\varepsilon > 0$ such that $t + \varepsilon, u + \varepsilon < T$ we define

$$w_s^\varepsilon(\omega) \triangleq 1_A(\omega) \frac{1}{\varepsilon} \left(1_{[t,t+\varepsilon]}(s) - 1_{[u,u+\varepsilon]}(s) \right) \text{ on } \Omega \times [0, T].$$

Obviously, w is progressively measurable, in $L^2(\mathbb{P} \otimes ds)$ and satisfies $\int_0^T w_s ds = 0$ $\mathbb{P}$-a.s. Hence, by assumption (19) we have

$$0 = \mathbb{E} \left[\int_0^T w_s^\varepsilon M_s ds \right] = \mathbb{E} \left[1_A \frac{1}{\varepsilon} \int_t^{t+\varepsilon} M_s ds \right] - \mathbb{E} \left[1_A \frac{1}{\varepsilon} \int_u^{u+\varepsilon} M_s ds \right].$$

Passing to the limit $\varepsilon \downarrow 0$, we obtain by right-continuity of M,

$$0 = \mathbb{E} \left[1_A (M_t - M_u) \right] \text{ for all } 0 \le t < u < T.$$

Consequently, M is a martingale on $[0, T)$. By assumption, we have that $\mathbb{E}[\int_0^T M_s^2 ds] < \infty$ which implies that M is square integrable on $[0, T)$. $\square$

Now, we are ready to prove our main result by simple verification. We start with Theorem 3.1 for the unconstrained problem (3).

Proof of Theorem 3.1 We divide the proof in two parts. First, we prove optimality of the solution given in (7). Then, we compute the corresponding minimal costs given in (8).

Optimality of (7): In order to show that our candidate in (7) is the optimal solution for problem (3), we need to check the first order condition in Lemma 5.2(1). For this, define the processes

$$Y_t \triangleq \int_0^t \xi_s \cosh(\tau^\kappa(s)) ds \text{ and } \tilde{M}_t \triangleq \mathbb{E}[Y_T | \mathscr{F}_t], \quad 0 \le t \le T.$$

Since $Y_T \in L^2(\mathbb{P})$, we have that $(\tilde{M}_t)_{0 \le t \le T}$ is a square integrable martingale. Moreover, note that $Y, \tilde{M} \in L^2(\mathbb{P} \otimes dt)$. Hence, the process $\hat{\xi}$ in Theorem 3.1 can be written as

$$\hat{\xi}_t = \frac{1}{\sqrt{\kappa} \sinh(\tau^\kappa(t))} \left(\tilde{M}_t - Y_t \right) \quad d\mathbb{P} \otimes dt\text{-a.e. on } \Omega \times [0, T) \tag{20}$$

with corresponding dynamics

$$d\hat{\xi}_t = -\frac{\coth(\tau^\kappa(t))}{\sqrt{\kappa}} (\xi_t - \hat{\xi}_t) dt + \frac{1}{\sqrt{\kappa} \sinh(\tau^\kappa(t))} d\tilde{M}_t \quad \text{on } [0, T). \tag{21}$$

Due to Lemma 5.5(b), we know that $\hat{\xi} \in L^2(\mathbb{P} \otimes dt)$. Now, the density of the solution from (7) satisfies

$$
\begin{aligned}
d\hat{u}_t &= -\frac{1}{\kappa}(1 - \tanh(\tau^\kappa(t))^2)\left(\hat{\xi}_t - \hat{X}_t\right)dt + \frac{1}{\sqrt{\kappa}}\tanh(\tau^\kappa(t))\left(d\hat{\xi}_t - d\hat{X}_t\right) \\
&= \frac{1}{\kappa}\left(\left(\hat{X}_t - \xi_t\right)dt + \frac{1}{\cosh(\tau^\kappa(t))}d\tilde{M}_t\right) \quad d\mathbb{P} \otimes dt\text{-a.e. on } \Omega \times [0, T],
\end{aligned}
$$

that is, $\hat{u}$ satisfies the BSDE-dynamics in (12). Obviously, it holds that $\hat{X}_0 = x$. Solving equation (7) for $\hat{X}$ yields upon differentiation

$$
\begin{aligned}
\hat{u}_t = &-\frac{1}{\sqrt{\kappa}}\frac{\sinh(\tau^\kappa(t))}{\cosh(\tau^\kappa(0))}x \\
&-\frac{1}{\kappa}\sinh(\tau^\kappa(t))\int_0^t \hat{\xi}_s\frac{\sinh(\tau^\kappa(s))}{\cosh(\tau^\kappa(s))^2}ds + \frac{1}{\kappa}\frac{\tilde{M}_t - Y_t}{\cosh(\tau^\kappa(t))}
\end{aligned} \tag{22}
$$

and we observe that $\lim_{t \uparrow T} \hat{u}_t = 0$ $\mathbb{P}$-a.s., i.e., the terminal condition in (12) is indeed satisfied. It remains to show that $\hat{u} \in L^2(\mathbb{P} \otimes dt)$. Since $\tilde{M}, Y \in L^2(\mathbb{P} \otimes dt)$, it suffices to observe that $\sinh(\tau^\kappa(s))/\cosh(\tau^\kappa(s))^2$ is bounded and therefore

$$
\begin{aligned}
\mathbb{E}\left[\int_0^T \left(\int_0^t \hat{\xi}_s\frac{\sinh(\tau^\kappa(s))}{\cosh(\tau^\kappa(s))^2}ds\right)^2 dt\right] &\leq \text{const}\,\mathbb{E}\left[\int_0^T \left(\int_0^t |\hat{\xi}_s|ds\right)^2 dt\right] \\
&\leq \text{const}\,\frac{T^2}{2}\|\hat{\xi}\|^2_{L^2(\mathbb{P} \otimes dt)} < \infty.
\end{aligned}
$$

Computation of minimal costs To compute the minimal costs associated to the optimal control $\hat{u}$ given in (8), note first that $\hat{u} \in L^2(\mathbb{P} \otimes dt)$ implies $\hat{X} \in L^2(\mathbb{P} \otimes dt)$ and thus $J(\hat{u}) < \infty$. For ease of presentation, we define

$$
c(t) \triangleq \sqrt{\kappa}\tanh(\tau^\kappa(t)), \quad 0 \leq t \leq T,
$$

so that $\hat{u}_t = c(t)(\hat{\xi}_t - \hat{X}_t)/\kappa$. Hence, the minimal costs can be written as

$$
\begin{aligned}
\infty > J(\hat{u}) &= \mathbb{E}\left[\frac{1}{2}\int_0^T (\hat{X}_s - \xi_s)^2 ds + \frac{1}{2}\kappa\int_0^T \hat{u}_s^2 ds\right] \\
&= \lim_{t \uparrow T}\left\{\frac{1}{2}\mathbb{E}\left[\int_0^t \hat{X}_s^2 ds\right] - \mathbb{E}\left[\int_0^t \hat{X}_s\xi_s ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t \xi_s^2 ds\right]\right. \\
&\quad + \frac{1}{2\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{\xi}_s^2 ds\right] - \frac{1}{\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{X}_s\hat{\xi}_s ds\right] \\
&\quad \left. + \frac{1}{2\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{X}_s^2 ds\right]\right\},
\end{aligned} \tag{23}
$$

due to monotone convergence. Observe that, using integration by parts and the dynamics of $\hat{\xi}$ from (21), we have, for all $t < T$,

$$
\begin{aligned}
\mathbb{E}[c(t)\hat{X}_t^2] = &\,c(0)x^2 + \frac{2}{\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{X}_s\hat{\xi}_s ds\right] \\
&-\frac{1}{\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{X}_s^2 ds\right] - \mathbb{E}\left[\int_0^t \hat{X}_s^2 ds\right]
\end{aligned}
$$

as well as

$$\mathbb{E}[c(t)\hat{X}_t\hat{\xi}_t] = c(0)\hat{\xi}_0 x + \frac{1}{\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{\xi}_s^2 ds\right] - \mathbb{E}\left[\int_0^t \hat{X}_s\xi_s ds\right]$$

and

$$\mathbb{E}[c(t)\hat{\xi}_t^2] = c(0)\hat{\xi}_0^2 + \frac{1}{\kappa}\mathbb{E}\left[\int_0^t c(s)^2\hat{\xi}_s^2 ds\right] - 2\mathbb{E}\left[\int_0^t \hat{\xi}_s\xi_s ds\right]$$
$$+ \mathbb{E}\left[\int_0^t \hat{\xi}_s^2 ds\right] + \mathbb{E}\left[\int_0^t c(s)d\langle\hat{\xi}\rangle_s\right].$$

Using these identities, we can write (23) as

$$\infty > J(\hat{u}) = \lim_{t\uparrow T}\left\{\frac{1}{2}c(0)(x - \hat{\xi}_0)^2 + \frac{1}{2}\mathbb{E}\left[\int_0^t (\hat{\xi}_s - \xi_s)^2 ds\right]\right.$$
$$\left. + \frac{1}{2}\mathbb{E}\left[\int_0^t c(s)d\langle\hat{\xi}\rangle_s\right] - \frac{1}{2}c(t)\mathbb{E}[(\hat{X}_t - \hat{\xi}_t)^2]\right\}. \qquad (24)$$

To conclude our assertion for the minimal costs in (8), observe that

$$\mathbb{E}[(\hat{X}_t - \hat{\xi}_t)^2] \le 2\left(\mathbb{E}[\hat{X}_t^2] + \mathbb{E}[\hat{\xi}_t^2]\right),$$

and let us argue why

$$\lim_{t\uparrow T} c(t)\mathbb{E}[\hat{X}_t^2] = 0 \quad \text{and} \quad \lim_{t\uparrow T} c(t)\mathbb{E}[\hat{\xi}_t^2] = 0. \qquad (25)$$

By Jensen's inequality, we have

$$\mathbb{E}[\hat{X}_t^2] \le t\mathbb{E}\left[\int_0^t \hat{u}_s^2 ds\right] \le T\,\mathbb{E}\left[\int_0^T \hat{u}_s^2 ds\right] < \infty.$$

Hence, due to $\lim_{t\uparrow T} c(t) = 0$, the first convergence in (25) holds true. Concerning the second convergence in (25), we use the representation in (20) for $\hat{\xi}$ to obtain, again with Jensen's inequality as well as the Cauchy-Schwarz inequality,

$$0 \le c(t)\mathbb{E}[\hat{\xi}_t^2] = \frac{c(t)}{\kappa\sinh(\tau^\kappa(t))^2}\mathbb{E}[(\tilde{M}_t - Y_t)^2]$$
$$\le \frac{c(t)}{\kappa\sinh(\tau^\kappa(t))^2}\mathbb{E}[(Y_T - Y_t)^2]$$
$$= \frac{c(t)}{\kappa\sinh(\tau^\kappa(t))^2}\mathbb{E}\left[\left(\int_t^T \xi_s\cosh(\tau^\kappa(s))ds\right)^2\right]$$
$$\le \frac{\cosh(\tau^\kappa(0))^2}{\sqrt{\kappa}\cosh(\tau^\kappa(t))}\frac{1}{\sinh(\tau^\kappa(t))}(T-t)\mathbb{E}\left[\int_t^T \xi_s^2 ds\right]$$
$$\le \frac{\cosh(\tau^\kappa(0))^2}{\cosh(\tau^\kappa(t))}\mathbb{E}\left[\int_t^T \xi_s^2 ds\right]\xrightarrow[t\uparrow T]{} 0,$$

where for the last inequality we used that $\sinh(\tau) \ge \tau$ for all $\tau \ge 0$. In other words, also the second convergence in (25) holds true. This finishes our proof of the representation of the minimal costs in (8). □

Next, we come to the proof of Theorem 3.2 concerning the constrained problem (4).

Proof of Theorem 3.2 Again, we will proceed in two steps. First, we prove optimality of the solution given in (9). Then, we compute the corresponding minimal costs given in (10).

Optimality of (9): The verification of the optimality of $\hat{X}^\Xi = x + \int_0^\cdot \hat{u}_t^\Xi dt$ in Theorem 3.2 for the constrained problem (4) follows along the same lines as in the unconstrained case. Again, we have to check the first order condition in Lemma 5.2(2). For this, we define the processes

$$Y_t \triangleq \frac{1}{\sqrt{\kappa}} \int_0^t \xi_s \sinh(\tau^\kappa(s)) ds \quad \text{and} \quad \tilde{M}_t^\Xi \triangleq \mathbb{E}[Y_T + \Xi_T | \mathscr{F}_t]$$

for all $0 \leq t \leq T$. Since $Y_T, \Xi_T \in L^2(\mathbb{P})$, we have that $(\tilde{M}_t^\Xi)_{0 \leq t \leq T}$ is a square integrable martingale. Moreover, note that $Y, \tilde{M}^\Xi \in L^2(\mathbb{P} \otimes dt)$. Hence, the process $\hat{\xi}^\Xi$ in Theorem 3.2 can be written as

$$\hat{\xi}_t^\Xi = \frac{1}{\cosh(\tau^\kappa(t))} \left(\tilde{M}_t^\Xi - Y_t \right) \quad d\mathbb{P} \otimes dt\text{-a.e. on } \Omega \times [0, T] \tag{26}$$

with corresponding dynamics

$$d\hat{\xi}_t^\Xi = -\frac{\tanh(\tau^\kappa(t))}{\sqrt{\kappa}} (\xi_t - \hat{\xi}_t^\Xi) dt + \frac{1}{\cosh(\tau^\kappa(t))} d\tilde{M}_t^\Xi \quad \text{on } [0, T]. \tag{27}$$

In particular, we observe that $\hat{\xi}^\Xi \in L^2(\mathbb{P} \otimes dt)$. Similar to the unconstrained case above, one easily checks that

$$d\hat{u}_t^\Xi = \frac{1}{\kappa} (\hat{X}_t^\Xi - \xi_t) dt + \frac{1}{\sqrt{\kappa}} \frac{1}{\sinh(\tau^\kappa(t))} d\tilde{M}_t^\Xi \quad d\mathbb{P} \otimes dt\text{-a.e. on } \Omega \times [0, T),$$

that is, $\hat{u}^\Xi$ satisfies the dynamics in (13). Obviously, it holds that $\hat{X}_0^\Xi = x$.

Next, we have to check the terminal condition in (13), that is, $\lim_{t \uparrow T} \hat{X}_t^\Xi = \Xi_T$ $\mathbb{P}$-a.s. In order to show this, first note that we can consider a càdlàg version of $(\hat{\xi}_t^\Xi)_{0 \leq t \leq T}$ due to its representation in (26). Hence, since Ξ_T is $\mathscr{F}_{T-}$-measurable by assumption (5) we obtain the $\mathbb{P}$-a.s. limit

$$\lim_{t \uparrow T} \hat{\xi}_t^\Xi = \mathbb{E}[\Xi_T | \mathscr{F}_{T-}] = \Xi_T$$

in (26). In other words, for every $\varepsilon > 0$ there exists a random time $\Upsilon_\varepsilon \in [0, T)$ such that $\mathbb{P}$-a.s.

$$\Xi_T - \varepsilon \leq \hat{\xi}_t^\Xi \leq \Xi_T + \varepsilon \quad \text{for all } t \in [\Upsilon_\varepsilon, T].$$

For $\lim_{t \uparrow T} \hat{X}_t^\Xi = \Xi_T$ $\mathbb{P}$-a.s., it clearly suffices to show that for any $\varepsilon > 0$ it holds that

$$\limsup_{t \uparrow T} \hat{X}_t^\Xi \leq \Xi_T + \varepsilon \quad \text{and} \quad \liminf_{t \uparrow T} \hat{X}_t^\Xi \geq \Xi_T - \varepsilon \quad \mathbb{P}\text{-a.s.}$$

Define $X_t^\varepsilon \triangleq \Xi_T + \varepsilon - \hat{X}_t^\Xi$ so that $\hat{\xi}_t^\Xi - \hat{X}_t^\Xi \leq X_t^\varepsilon$ $\mathbb{P}$-a.s. for $t \in [\Upsilon_\varepsilon, T)$. This yields

$$dX_t^\varepsilon = -d\hat{X}_t^\Xi = -\frac{1}{\sqrt{\kappa}} \coth(\tau^\kappa(t))(\hat{\xi}_t^\Xi - \hat{X}_t^\Xi) dt$$

$$\geq -\frac{1}{\sqrt{\kappa}} \coth(\tau^\kappa(t)) X_t^\varepsilon dt.$$

Moreover, note that for all $\omega \in \Omega$ the linear ODE on $[\Upsilon_\varepsilon(\omega), T)$ given by

$$dZ_t = -\frac{1}{\sqrt{\kappa}} \coth(\tau^\kappa(t)) Z_t dt, \quad Z_{\Upsilon_\varepsilon(\omega)} = X_{\Upsilon_\varepsilon(\omega)}^\varepsilon(\omega),$$

admits the solution

$$Z_t = X^\varepsilon_{\Upsilon_\varepsilon} \exp\left(-\frac{1}{\sqrt{\kappa}} \int_{\Upsilon_\varepsilon}^t \coth(\tau^\kappa(s))ds\right) = X^\varepsilon_{\Upsilon_\varepsilon} \frac{\sinh(\tau^\kappa(t))}{\sinh(\tau^\kappa(\Upsilon_\varepsilon))}, \quad t < T,$$

with $\lim_{t\uparrow T} Z_t = 0$. By the comparison principle for ODEs, we get $\mathbb{P}$-a.s. $X^\varepsilon_t \geq Z_t$ for all $t \in [\Upsilon_\varepsilon, T)$. Hence,

$$\liminf_{t\uparrow T} X^\varepsilon_t \geq \lim_{t\uparrow T} Z_t = 0 \quad \mathbb{P}\text{-a.s.},$$

that is, $\limsup_{t\uparrow T} \hat{X}^\Xi_t \leq \Xi_T + \varepsilon$ $\mathbb{P}$-a.s. Similarly, define $\tilde{X}^\varepsilon_t \triangleq \Xi_T - \varepsilon - \hat{X}^\Xi_t$ and observe as above that $\mathbb{P}$-a.s. on $[\Upsilon_\varepsilon, T)$ we have

$$d\tilde{X}^\varepsilon_t \leq -\frac{1}{\sqrt{\kappa}} \coth(\tau^\kappa(t))\tilde{X}^\varepsilon_t dt.$$

Again, as above by comparison principle we obtain

$$\limsup_{t\uparrow T} \tilde{X}^\varepsilon_t \leq 0 \quad \mathbb{P}\text{-a.s.},$$

i.e., $\liminf_{t\uparrow T} \hat{X}^\Xi_t \geq \Xi_T - \varepsilon$ $\mathbb{P}$-a.s. as remained to be shown for (13).

Finally, we have to argue that $\hat{u}^\Xi \in L^2(\mathbb{P} \otimes dt)$. For this, we may assume without loss of generality that $x = 0$. Moreover, let us denote $\hat{u}^{\Xi,\xi} \triangleq \hat{u}^\Xi$, $\hat{X}^{\Xi,\xi} \triangleq \hat{X}^\Xi$ and $\hat{\xi}^{\Xi,\xi} \triangleq \hat{\xi}^\Xi$ to emphasize also the dependence on the given target process ξ. With this notation it holds that

$$\hat{u}^{\Xi,\xi} = \hat{u}^{\Xi,0} + \hat{u}^{0,\xi}.$$

Hence, we have to show that $\hat{u}^{\Xi,0} \in L^2(\mathbb{P} \otimes dt)$ and $\hat{u}^{0,\xi} \in L^2(\mathbb{P} \otimes dt)$.

Concerning $\hat{u}^{\Xi,0}$, observe that, using $\hat{\xi}^{\Xi,0}_t = \Xi_t/\cosh(\tau^\kappa(t))$ with $\Xi_t \triangleq \mathbb{E}[\Xi_T|\mathscr{F}_t]$, $0 \leq t \leq T$, as well as the explicit solution $\hat{X}^{\Xi,0}_t$ for the ODE in (9), we obtain

$$\begin{aligned}
\hat{u}^{\Xi,0}_t &= \frac{\coth(\tau^\kappa(t))}{\sqrt{\kappa}} \left(\hat{\xi}^{\Xi,0}_t - \hat{X}^{\Xi,0}_t\right) \\
&= \frac{\coth(\tau^\kappa(t))}{\sqrt{\kappa}} \left(e^{-\int_0^t \frac{\coth(\tau^\kappa(u))}{\sqrt{\kappa}}du} \hat{\xi}^{\Xi,0}_0 \right. \\
&\qquad \left. + e^{-\int_0^t \frac{\coth(\tau^\kappa(u))}{\sqrt{\kappa}}du} \int_0^t e^{\int_0^s \frac{\coth(\tau^\kappa(u))}{\sqrt{\kappa}}du} d\hat{\xi}^{\Xi,0}_s\right) \\
&= \frac{\cosh(\tau^\kappa(t))}{\sqrt{\kappa}\sinh(\tau^\kappa(0))} \hat{\xi}^{\Xi,0}_0 + \frac{\cosh(\tau^\kappa(t))}{\kappa} \int_0^t \frac{\Xi_s}{\cosh(\tau^\kappa(s))^2} ds \\
&\quad + \frac{\cosh(\tau^\kappa(t))}{\sqrt{\kappa}} \int_0^t \frac{2}{\sinh(2\tau^\kappa(s))} d\Xi_s,
\end{aligned} \tag{28}$$

where we used integration by parts in the second line. Obviously, the first two terms in (28) belong to $L^2(\mathbb{P} \otimes dt)$. The third term is in $L^2(\mathbb{P} \otimes dt)$ since, using Fubini's Theorem as well

as $\sinh(\tau) \geq \tau$ for all $\tau \geq 0$, we get

$$
\mathbb{E}\left[\int_0^T \left(\int_0^t \frac{2 d\Xi_s}{\sinh(2\tau^\kappa(s))}\right)^2 dt\right] = \mathbb{E}\left[\int_0^T \int_0^t \left(\frac{2}{\sinh(2\tau^\kappa(s))}\right)^2 d\langle\Xi\rangle_s dt\right]
$$
$$
= \mathbb{E}\left[\int_0^T (T-s)\left(\frac{2}{\sinh(2\tau^\kappa(s))}\right)^2 d\langle\Xi\rangle_s\right] \leq \mathbb{E}\left[\int_0^T \frac{\kappa}{T-s} d\langle\Xi\rangle_s\right]
$$
$$
= \kappa \int_0^T \frac{d\mathbb{E}[\Xi_s^2]}{T-s} < \infty
$$

by assumption (5).

Concerning $\hat{u}_t^{0,\xi}$, we use the explicit expressions for $\hat{\xi}_t^{0,\xi}$ and $\hat{X}_t^{0,\xi}$ to obtain in (9) that

$$
\hat{u}_t^{0,\xi} = \frac{\coth(\tau^\kappa(t))}{\sqrt{\kappa}}\left(\hat{\xi}_t^{0,\xi} - \hat{X}_t^{0,\xi}\right)
$$
$$
= \frac{\cosh(\tau^\kappa(t))-1}{\sqrt{\kappa}\sinh(\tau^\kappa(t))}\mathbb{E}\left[\int_t^T \xi_u K^\Xi(t,u)du \,\Big|\, \mathscr{F}_t\right]
$$
$$
- \frac{\cosh(\tau^\kappa(t))}{\kappa}\int_0^t \frac{\cosh(\tau^\kappa(s))-1}{\sinh(\tau^\kappa(s))^2}\mathbb{E}\left[\int_s^T \xi_u K^\Xi(s,u)du \,\Big|\, \mathscr{F}_s\right]ds. \qquad (29)
$$

Note that all the ratios in (29) involving the functions $\cosh(\cdot)$ and $\sinh(\cdot)$ are actually bounded on $[0, T]$. Moreover, we have by Lemma 5.5(c) below that

$$
\mathbb{E}\left[\int_t^T \xi_u K^\Xi(t,u)du \,\Big|\, \mathscr{F}_t\right] \in L^2(\mathbb{P}\otimes dt),
$$

as well as, using Jensen's inequality,

$$
\mathbb{E}\left[\int_0^T \left(\int_0^t \mathbb{E}\left[\int_s^T \xi_u K^\Xi(s,u)du \,\Big|\, \mathscr{F}_s\right]ds\right)^2 dt\right]
$$
$$
\leq \frac{T^2}{2}\mathbb{E}\left[\int_0^T \left(\mathbb{E}\left[\int_s^T \xi_u K^\Xi(s,u)du \,\Big|\, \mathscr{F}_s\right]\right)^2 ds\right] < \infty.
$$

Together, this shows $\hat{u}^\Xi \in L^2(\mathbb{P}\otimes dt)$ as desired.

Computation of minimal costs Now, we compute the minimal costs associated to the optimal control $\hat{u}^\Xi$ given in (10). We will follow along the same lines as in the unconstrained case above. First of all, note that $\hat{u}^\Xi \in L^2(\mathbb{P}\otimes dt)$ implies $\hat{X}^\Xi \in L^2(\mathbb{P}\otimes dt)$ and hence $J(\hat{u}) < \infty$. For ease of presentation, we define

$$
c(t) \triangleq \sqrt{\kappa}\coth(\tau^\kappa(t)), \quad 0 \leq t < T,
$$

i.e., $\hat{u}_t^\Xi = c(t)(\hat{\xi}_t^\Xi - \hat{X}_t^\Xi)/\kappa$. Analogously to the unconstrained case above, we can write $J(\hat{u}^\Xi)$ as

$$
\infty > J(\hat{u}^\Xi) = \lim_{t\uparrow T}\left\{\frac{1}{2}c(0)(x - \hat{\xi}_0^\Xi)^2 + \frac{1}{2}\mathbb{E}\left[\int_0^t (\hat{\xi}_s^\Xi - \xi_s)^2 ds\right]\right.
$$
$$
\left. + \frac{1}{2}\mathbb{E}\left[\int_0^t c(s)d\langle\hat{\xi}^\Xi\rangle_s\right] - \frac{1}{2}c(t)\mathbb{E}[(\hat{X}_t^\Xi - \hat{\xi}_t^\Xi)^2]\right\}. \qquad (30)
$$

To conclude our assertion for the minimal costs in (10), observe that

$$
\mathbb{E}[(\hat{X}_t^\Xi - \hat{\xi}_t^\Xi)^2] \leq 2\left(\mathbb{E}[(\hat{X}_t^\Xi - \Xi_t)^2] + \mathbb{E}[(\Xi_t - \hat{\xi}_t^\Xi)^2]\right),
$$

where $\Xi_t \triangleq \mathbb{E}[\Xi_T|\mathscr{F}_t]$, $0 \le t \le T$, and let us argue why

$$\lim_{t\uparrow T} c(t)\mathbb{E}[(\hat{X}_t^\Xi - \Xi_t)^2] = 0 \quad \text{and} \quad \lim_{t\uparrow T} c(t)\mathbb{E}[(\Xi_t - \hat{\xi}_t^\Xi)^2] = 0. \tag{31}$$

Concerning the first convergence in (31), Jensen's inequality, monotonicity of the function $\cosh(\cdot)$ as well as the estimate $\sinh(\tau) \ge \tau$ for all $\tau \ge 0$ yield

$$\begin{aligned}
c(t)\mathbb{E}[(\hat{X}_t^\Xi - \Xi_t)^2] &\le c(t)\mathbb{E}[(\hat{X}_t^\Xi - \hat{X}_T^\Xi)^2] \\
&\le \frac{\kappa \cosh(\tau^\kappa(0))}{T-t} \mathbb{E}\left[\left(\int_t^T \hat{u}_s^\Xi ds\right)^2\right] \\
&\le \kappa \cosh(\tau^\kappa(0))\mathbb{E}\left[\int_t^T (\hat{u}_s^\Xi)^2 ds\right] \xrightarrow[t\uparrow T]{} 0, \tag{32}
\end{aligned}$$

since $\Xi_T = \hat{X}_T^\Xi$ and $\hat{u}^\Xi \in L^2(\mathbb{P} \otimes dt)$.

Concerning the second convergence in (31), we insert the definition for $\hat{\xi}^\Xi$ to obtain that

$$\begin{aligned}
&c(t)\mathbb{E}[(\Xi_t - \hat{\xi}_t^\Xi)^2] \\
&= c(t)\mathbb{E}\left[\left(\frac{\cosh(\tau^\kappa(t))-1}{\cosh(\tau^\kappa(t))}\Xi_t - \frac{\cosh(\tau^\kappa(t))-1}{\cosh(\tau^\kappa(t))}\mathbb{E}\left[\int_t^T \xi_u K^\Xi(t,u)du\Big|\mathscr{F}_t\right]\right)^2\right] \\
&\le 2c(t)\left(\frac{\cosh(\tau^\kappa(t))-1}{\cosh(\tau^\kappa(t))}\right)^2 \mathbb{E}[\Xi_T^2] + 2c(t)\left(\frac{\cosh(\tau^\kappa(t))-1}{\cosh(\tau^\kappa(t))}\right)^2 \mathbb{E}\left[\int_t^T \xi_u^2 K^\Xi(t,u)du\right] \\
&\le \frac{2\sqrt{\kappa}}{\cosh(\tau^\kappa(t))} \frac{(\cosh(\tau^\kappa(t))-1)^2}{\sinh(\tau^\kappa(t))}\mathbb{E}[\Xi_T^2] \\
&\quad + \frac{2\sinh(\tau^\kappa(0))}{\cosh(\tau^\kappa(t))}\frac{\cosh(\tau^\kappa(t))-1}{\sinh(\tau^\kappa(t))}\mathbb{E}\left[\int_t^T \xi_u^2 du\right] \xrightarrow[t\uparrow T]{} 0,
\end{aligned}$$

since $\Xi_T \in L^2(\mathbb{P})$, $\xi \in L^2(\mathbb{P} \otimes dt)$ and $\lim_{t\uparrow T}(\cosh(\tau^\kappa(t))-1)/\sinh(\tau^\kappa(t)) = 0$. Consequently, also the second convergence in (31) holds true. This finishes our proof of the representation of the minimal costs in (10). □

The next lemma shows that the set $\mathscr{U}_x^\Xi$ is not empty under the assumption (5).

Lemma 5.4 *For $\Xi_T \in L^2(\mathbb{P}, \mathscr{F}_T)$ we have that $\mathscr{U}_x^\Xi \ne \varnothing$ if and only if condition (5) holds, i.e., if and only if $\int_0^T \frac{d\mathbb{E}[\Xi_t^2]}{T-t} < \infty$ with $\Xi_t \triangleq \mathbb{E}[\Xi_T|\mathscr{F}_t]$ for all $0 \le t \le T$.*

Proof Let $\Xi_T \in L^2(\mathbb{P}, \mathscr{F}_T)$. We first prove necessity. Assume there exists $u \in \mathscr{U}_x^\Xi$, i.e., $u \in L^2(\mathbb{P} \otimes dt)$ such that

$$X_T^u = x + \int_0^T u_s ds = \Xi_T.$$

Then, applying Fubini's Theorem, we obtain

$$\int_0^T \frac{d\mathbb{E}[\Xi_t^2]}{T-t} = \frac{1}{T}(\mathbb{E}[\Xi_T^2] - \mathbb{E}[\Xi_0^2]) + \int_0^T \mathbb{E}[\Xi_T^2 - \Xi_s^2]d\left(\frac{1}{T-s}\right).$$

Moreover, $\mathbb{E}[\Xi_T^2 - \Xi_s^2] = \mathbb{E}[(\Xi_T - \Xi_s)^2] \leq \mathbb{E}[(X_T^u - X_s^u)^2]$ due to the L^2-projection property of conditional expectations. Hence, we get

$$\int_0^T \frac{d\mathbb{E}[\Xi_t^2]}{T-t} \leq \frac{1}{T}(\mathbb{E}[\Xi_T^2] - \mathbb{E}[\Xi_0^2]) + \int_0^T \mathbb{E}\left[\left(\int_s^T u_r dr\right)^2\right] d\left(\frac{1}{T-s}\right)$$

$$= \frac{1}{T}(\mathbb{E}[\Xi_T^2] - \mathbb{E}[\Xi_0^2]) + \mathbb{E}\left[\int_0^T \left(\frac{1}{T-s}\int_s^T u_r dr\right)^2 ds\right] < \infty$$

by $\Xi_T \in L^2(\mathbb{P})$ and Lemma 5.5(a).

For sufficiency, simply consider the optimizer $\hat{u}^\Xi$ from Theorem 3.2 which we proved to be in $\mathscr{U}_x^\Xi$ under the condition (5). □

The final lemma collects estimates concerning the $L^2(\mathbb{P} \otimes dt)$-norm which are needed several times in the proofs above.

Lemma 5.5 *Let $(\zeta_t)_{0 \leq t \leq T} \in L^2(\mathbb{P} \otimes dt)$ be progressively measurable. Moreover, let $K(t, u)$, $K^\Xi(t, u)$, $0 \leq t \leq u < T$, denote the kernels from Theorems 3.1 and 3.2, respectively.*

a) For $\bar{\zeta}_t \triangleq \frac{1}{T-t}\int_t^T \zeta_s ds$, $t < T$, we have

$$\|\bar{\zeta}\|_{L^2(\mathbb{P} \otimes dt)} \leq 2\|\zeta\|_{L^2(\mathbb{P} \otimes dt)}.$$

b) For $\zeta_t^K \triangleq \mathbb{E}[\int_t^T \zeta_u K(t, u) du | \mathscr{F}_t]$, $t < T$, we have

$$\|\zeta^K\|_{L^2(\mathbb{P} \otimes dt)} \leq c\|\zeta\|_{L^2(\mathbb{P} \otimes dt)}$$

for some constant $c > 0$.

c) For $\zeta_t^{K^\Xi} \triangleq \mathbb{E}[\int_t^T \zeta_u K^\Xi(t, u) du | \mathscr{F}_t]$, $t < T$, we have

$$\|\zeta^{K^\Xi}\|_{L^2(\mathbb{P} \otimes dt)} \leq c\|\zeta\|_{L^2(\mathbb{P} \otimes dt)}$$

for some constant $c > 0$.

Proof a) By Fubini's Theorem and the Cauchy-Schwarz inequality, we have

$$\|\bar{\zeta}\|_{L^2(\mathbb{P} \otimes dt)}^2 = \mathbb{E}\left[\int_0^T \int_0^T \zeta_r \zeta_s \int_0^{r \wedge s} \left(\frac{1}{T-t}\right)^2 dt dr ds\right]$$

$$= \mathbb{E}\left[\int_0^T \int_0^T \zeta_r \zeta_s \frac{1}{T - r \wedge s} dr ds\right] - \frac{1}{T}\mathbb{E}\left[\left(\int_0^T \zeta_s ds\right)^2\right]$$

$$\leq \mathbb{E}\left[2\int_0^T \zeta_r \int_0^r \zeta_s \frac{1}{T-s} ds dr\right]$$

$$= 2\mathbb{E}\left[\int_0^T \zeta_s \left(\frac{1}{T-s}\int_s^T \zeta_r dr\right) ds\right]$$

$$\leq 2\|\zeta\|_{L^2(\mathbb{P} \otimes dt)}\|\bar{\zeta}\|_{L^2(\mathbb{P} \otimes dt)}$$

and hence the assertion.

b) First, assume that $(\zeta_t)_{0 \le t \le T}$ is deterministic, and so $\zeta_t^K = \int_t^T \zeta_u K(t,u)du$. By similar computations as in (a) we obtain

$$
\begin{aligned}
\|\zeta^K\|_{L^2(dt)}^2 &= \int_0^T \int_0^T \zeta_r \zeta_s \int_0^{r \wedge s} K(t,r)K(t,s)dt\,dr\,ds \\
&\le \int_0^T \int_0^T \zeta_r \zeta_s \frac{1}{\sqrt{\kappa}} \cosh(\tau^\kappa(r)) \cosh(\tau^\kappa(s)) \coth(\tau^\kappa(r \wedge s))dr\,ds \\
&= 2 \int_0^T \zeta_r \frac{\cosh(\tau^\kappa(r))}{\sqrt{\kappa}} \int_0^r \zeta_s \cosh(\tau^\kappa(s)) \coth(\tau^\kappa(s))ds\,dr \\
&= 2 \int_0^T \zeta_s \cosh(\tau^\kappa(s))^2 \zeta_s^K ds \\
&\le 2\cosh(\tau^\kappa(0))^2 \|\zeta\|_{L^2(dt)} \|\zeta^K\|_{L^2(dt)},
\end{aligned}
$$

i.e., $\|\zeta^K\|_{L^2(dt)} \le c\|\zeta\|_{L^2(dt)}$ for some constant $c > 0$. Now, for general $(\zeta_t)_{0 \le t \le T} \in L^2(\mathbb{P} \otimes dt)$ progressively measurable, we get with Fubini's Theorem

$$
\mathbb{E}\left[\int_0^T (\zeta_t^K)^2 dt\right] = \int_0^T \int_t^T \int_t^T \mathbb{E}\left[\mathbb{E}[\zeta_r|\mathscr{F}_t]\mathbb{E}[\zeta_s|\mathscr{F}_t]\right] K(t,r)K(t,s)dr\,ds\,dt.
$$

Again, application of Cauchy-Schwarz's and Jensen's inequalities yields

$$
\mathbb{E}\left[\mathbb{E}[\zeta_r|\mathscr{F}_t]\mathbb{E}[\zeta_s|\mathscr{F}_t]\right] \le \|\zeta_r\|_{L^2(\mathbb{P})} \|\zeta_s\|_{L^2(\mathbb{P})}, \quad t \le r,s \le T.
$$

Consequently,

$$
\begin{aligned}
\|\zeta^K\|_{L^2(\mathbb{P} \otimes dt)}^2 &\le \int_0^T \int_t^T \int_t^T \|\zeta_r\|_{L^2(\mathbb{P})} \|\zeta_s\|_{L^2(\mathbb{P})} K(t,r)K(t,s)dr\,ds\,dt \\
&= \int_0^T \left(\int_t^T \|\zeta_r\|_{L^2(\mathbb{P})} K(t,r)dr\right)^2 dt.
\end{aligned}
$$

Now, put $\tilde{\zeta}_t \triangleq \|\zeta_t\|_{L^2(\mathbb{P})}$ and apply the estimate already proved for deterministic functions to conclude

$$
\begin{aligned}
\|\zeta^K\|_{L^2(\mathbb{P} \otimes dt)}^2 &= \int_0^T \left(\int_t^T \tilde{\zeta}_r K(t,r)dr\right)^2 dt \\
&\le c \int_0^T |\tilde{\zeta}_t|^2 dt = c \int_0^T \mathbb{E}[\zeta_t^2]dt = c\|\zeta\|_{L^2(\mathbb{P} \otimes dt)}^2.
\end{aligned}
$$

c) Jensen's inequality and Fubini's Theorem give

$$
\begin{aligned}
\|\zeta^{K^\Xi}\|_{L^2(\mathbb{P} \otimes dt)}^2 &= \mathbb{E}\left[\int_0^T (\zeta_t^{K^\Xi})^2 dt\right] \le \int_0^T \int_t^T \mathbb{E}[\zeta_u^2]K^\Xi(t,u)du\,dt \\
&= \int_0^T \mathbb{E}[\zeta_u^2] \int_0^u K^\Xi(t,u)dt\,du.
\end{aligned}
$$

Now, using $\cosh(\tau) - 1 \ge \tau^2/2$ for all $\tau \ge 0$, we get

$$
\begin{aligned}
0 \le \int_0^u K^\Xi(t,u)dt &= \int_0^u \frac{\sinh(\tau^\kappa(u))}{\sqrt{\kappa}(\cosh(\tau^\kappa(t)) - 1)}dt \\
&\le \frac{\sinh(\tau^\kappa(u))}{\sqrt{\kappa}} \int_0^u \frac{2\kappa}{(T-t)^2}dt \le 2\sqrt{\kappa}\frac{\sinh(\tau^\kappa(u))}{T-u} \xrightarrow[u \uparrow T]{} 1.
\end{aligned}
$$

Thus, the above integral over K^Ξ is bounded uniformly in $0 \le u \le T$ by some constant $c > 0$, and so

$$\|\zeta^{K^\Xi}\|^2_{L^2(\mathbb{P}\otimes dt)} \le c \int_0^T \mathbb{E}[\zeta_u^2]du = c \|\zeta\|^2_{L^2(\mathbb{P}\otimes dt)}$$

yielding the assertion in (c).   □

References

1. Alfonsi, A., Fruth, A., Schied, A.: Optimal execution strategies in limit order books with general shape functions. Quant. Financ. **10**(2), 143–157 (2010). doi:10.1080/14697680802595700
2. Almgren, R., Chriss, N.: Optimal execution of portfolio transactions. J. Risk **3**, 5–39 (2001)
3. Almgren, R., Li, T.M.: Option hedging with smooth market impact. Preprint, June 2015
4. Almgren, R.F.: Optimal execution with nonlinear impact functions and trading-enhanced risk. Appl. Math. Financ. **10**(1), 1–18 (2003). doi:10.1080/135048602100056
5. Black, F., Scholes, M.: The pricing of options and corporate liabilities. J. Polit. Econ. **81**, 637–654 (1973)
6. Cai, J., Rosenbaum, M., Tankov, P.: Asymptotic lower bounds for optimal tracking: a linear programming approach. Preprint, October 2015
7. Cartea, Á., Jaimungal, S.: A closed-form execution strategy to target VWAP. Preprint, January 2015
8. Ekeland, I., Témam, R.: Convex Analysis and Variational Problems. Society for Industrial and Applied Mathematics, Philadelphia (1999). doi:10.1137/1.9781611971088
9. Föllmer, H., Sondermann, D.: Hedging of non-redundant contingent claims. In: Contributions to Mathematical Economics, pp. 205–223. North-Holland, Amsterdam (1986)
10. Frei, C., Westray, N.: Optimal execution of a VWAP order: a stochastic control approach. Math. Financ. (2013). doi:10.1111/mafi.12048
11. Gârleanu, N., Pedersen, L.H.: Dynamic trading with predictable returns and transaction costs. J. Financ. **68**(6), 2309–2340 (2013a). ISSN 1540-6261. doi:10.1111/jofi.12080
12. Gârleanu, N., Pedersen, L.H.: Dynamic portfolio choice with frictions. Preprint, May 2013b
13. Gökay, S., Roch, A.F., Soner, H.M.: Liquidity models in continuous and discrete time. In: Di Nunno, G., Øksendal, B. (eds.) Advanced Mathematical Methods for Finance, pp. 333–365. Springer, Berlin (2011). ISBN 978-3-642-18412-3. doi:10.1007/978-3-642-18412-3_13
14. Guasoni, P., Weber, M.: Dynamic trading volume. Math. Financ. (2015a). ISSN 1467-9965. doi:10.1111/mafi.12099
15. Guasoni, P., Weber, M.: Nonlinear price impact and portfolio choice. Preprint, June 2015b
16. Guéant, O., Pu, J.: Option pricing and hedging with execution costs and market impact. Preprint, April 2015. arxiv:1311.4342
17. Kallsen, J., Muhle-Karbe, J.: High-resilience limits of block-shaped order books. Preprint, September 2014
18. Kohlmann, M., Tang, S.: Global adapted solution of one-dimensional backward stochastic Riccati equations, with application to the mean-variance hedging. Stoch. Process. Appl. **97**(2), 255–288 (2002). ISSN 0304-4149. doi:10.1016/S0304-4149(01)00133-8
19. Merton, R.C: Theory of rational option pricing. Bell J. Econ. Manag. Sci. **4**, 141–183 (1973). ISSN 0741-6261
20. Moreau, L., Muhle-Karbe, J., Soner, H.M.: Trading with small price impact. Math. Financ. (2015). ISSN 1467-9965. doi:10.1111/mafi.12098
21. Naujokat, F., Westray, N.: Curve following in illiquid markets. Math. Financ. Econ. **4**(4), 299–335 (2011). ISSN 1862-9679. doi:10.1007/s11579-011-0042-5
22. Obizhaeva, A.A., Wang, J.: Optimal trading strategy and supply/demand dynamics. J. Financ. Mark. **16**(1), 1–32 (2013). ISSN 1386-4181. doi:10.1016/j.finmar.2012.09.001. http://www.sciencedirect.com/science/article/pii/S1386418112000328
23. Predoiu, S., Shaikhet, G., Shreve, S.: Optimal execution in a general one-sided limit-order book. SIAM J. Financ. Math. **2**(1), 183–212 (2011). doi:10.1137/10078534X

24. Rogers, L.C.G., Singh, S.: The cost of illiquidity and its effects on hedging. Math. Financ. **20**, 597–615 (2010)
25. Schied, A., Schöneborn, T.: Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. Financ. Stoch. **13**(2), 181–204 (2009). ISSN 0949-2984