

Becherer, Dirk; Bilarev, Todor

Article — Published Version

Hedging with physical or cash settlement under transient multiplicative price impact

Finance and Stochastics

Suggested Citation: Becherer, Dirk; Bilarev, Todor (2024) : Hedging with physical or cash settlement under transient multiplicative price impact, Finance and Stochastics, ISSN 1432-1122, Springer Berlin Heidelberg, Berlin/Heidelberg, Vol. 28, Iss. 2, pp. 285-328, <https://doi.org/10.1007/s00780-024-00531-7>

This Version is available at:

<https://hdl.handle.net/10419/315064>

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/>



Hedging with physical or cash settlement under transient multiplicative price impact

Dirk Becherer¹ · Todor Bilarev²

Received: 12 July 2018 / Accepted: 7 March 2023 / Published online: 15 March 2024
© The Author(s) 2024

Abstract

We solve the superhedging problem for European options in an illiquid extension of the Black–Scholes model, in which transactions have transient price impact and the costs and strategies for hedging are affected by physical or cash settlement requirements at maturity. Our analysis is based on a convenient choice of reduced effective coordinates of magnitudes at liquidation for geometric dynamic programming. The price impact is transient over time and multiplicative, ensuring nonnegativity of underlying asset prices while maintaining an arbitrage-free model. The basic (log-)linear example is a Black–Scholes model with a relative price impact proportional to the volume of shares traded, where the transience for impact on log-prices is modelled like in Obizhaeva and Wang (J. Financ. Mark. 16:1–32, 2013) for nominal prices. More generally, we allow nonlinear price impact and resilience functions. The viscosity solutions describing the minimal superhedging price are governed by the transient character of the price impact and by the physical or cash settlement specifications. The pricing equations under illiquidity extend no-arbitrage pricing à la Black–Scholes for complete markets in a non-paradoxical way (cf. Çetin et al. (Finance Stoch. 14:317–341, 2010)) even without additional frictions, and can recover it in base cases.

Keywords Transient price impact · Multiplicative impact · Hedging · Option settlement · Resilience · Viscosity solution · Geometric dynamical programming · Effective coordinates

Mathematics Subject Classification 49L20 · 49L25 · 60H30 · 91G20 · 93E20

Disclaimer: The opinions expressed in this publication are those of the authors. They do not purport to reflect views of their institutions.

✉ D. Becherer
becherer@math.hu-berlin.de

T. Bilarev
bilarev.todor@gmail.com

¹ Humboldt Universität zu Berlin, Institut für Mathematik, Unter den Linden 6, D-10099 Berlin, Germany

² FactSet Research Systems Inc., 2 Srebama Str., 1407 Sofia, Bulgaria

1 Introduction

By using methods of stochastic target problems (see Soner and Touzi [29]) and geometric dynamic programming in suitably chosen reduced effective coordinates of magnitudes at liquidation, we solve the superhedging problem for European derivatives in a market model with multiplicative transient price impact. If the market for the underlying is illiquid or if large volumes are to be traded, there is price impact and feedback effects from hedging can affect the minimal superhedging prices (see Frey [19], Schönbucher and Wilmott [28], Frey and Polte [20], Bank and Baum [3]) and the corresponding hedging strategies which almost surely superreplicate the option. Since trades at maturity can alter the price of the underlying and thereby the derivative payout, settlement specifications for the option (in cash or in physical units) become relevant and this shows in pricing and hedging equations. As our results address hedging in terms of liquidation values, i.e., “real” instead of “paper” values (see Jarrow [24]), we recover such effects, whereas Frey [19], Frey and Polte [20], Bouchard et al. [13] study hedging in terms of book (“paper”) values. The settlement constraints imposed for hedging (in Sect. 2) in combination with stability by a suitably chosen notion of value, which depends continuously on trading strategies, moreover help to avoid some known paradoxical effects in price impact modelling (see Bank and Baum [3], Çetin et al. [16], Becherer et al. [9, Remark 3.3], and cf. the comments about different notions of wealth after (2.9) and (2.10)) and overly excessive opportunities of manipulating derivative payoffs (as in Schönbucher and Wilmott [28, Sect. 4.1]).

The best-known model for transient price impact is probably the one due to Obizhaeva and Wang [26]. It states that the dynamic holdings Θ of a large trader have additive linear impact (with parameter $\lambda > 0$) on the prevailing price s of the underlying asset via

$$\begin{aligned} \text{(log-)price:} \quad & ds_t = d\bar{s}_t + \lambda dY_t, \quad \text{with} \\ \text{impact level:} \quad & dY_t = -\beta Y_t dt + d\Theta_t =: -h(Y_t)dt + d\Theta_t, \end{aligned} \quad (1.1)$$

where \bar{s} is a given unaffected (fundamental) price evolution for the underlying, while Y is a market impact level process, whose mean-reverting dynamics is driven by Θ and is linear in the asset holdings Θ of the large trader and transient over time, recovering at some resilience rate given by the parameter $\beta > 0$ in the linear resilience function h .

Assuming price impact to be additive helps for mathematical tractability (in particular if \bar{s} is a martingale) and can serve to approximate multiplicative impact on a short horizon. This is common in the literature on optimal trade execution, as explained in Busseti and Lillo [15, Sect. 6], who further describe [15, Sect. 5] how transient impact is calibrated additively to log-prices, hence multiplicatively in prices. See also the comparison in Becherer et al. [7, Example 5.5] for arguments in favour of impact to be multiplicative if combined with multiplicative price dynamics of Black–Scholes type. A large strand of literature investigates (linear–)quadratic control problems in

this realm (see e.g. Bank et al. [4], Ackermann et al. [1]) which are different from superhedging and its respective pricing problem. An undesirable property of the additive impact (1.1) in this context is that it can lead to negative prices s for the underlying asset. It is plausible that trading a quantity of stocks, that is, a fraction of company ownership, should have a relative (hence multiplicative) effect on the price. Indeed, already Bertsimas and Lo [10, Sect. 3] have argued that relative (percentage) price impact which is proportional to the traded number of stocks (i.e., additive impact with respect to log-price in first order approximation) is more plausible than absolute price impact, and they cite empirical evidence.

A simple way to obtain a multiplicative impact variant is by a log-linear interpretation of the additive Obizhaeva–Wang model (1.1), simply by taking $s = \log S$, $\bar{s} = \log \bar{S}$ to be log-prices instead of nominal prices S , \bar{S} (affected, respective fundamental). Then the impact on $S = \bar{S} \exp(\lambda Y)$ is multiplicative and log-linear, with the resilience and the (log-)price impact functions from (1.1) being linear. This is the basic log-linear example (see Example 2.1) which is covered by and motivates our transient multiplicative impact model, with unaffected price process \bar{S} for the underlying asset of Black–Scholes–Merton type. Our analysis moreover allows non-linear and nonparametric resilience and price impact functions h and f in (2.1), (2.2). The model is a multiplicative variant of the (nonlinear) additive impact model from Predoiu et al. [27], where price impact can be interpreted in terms of a limit order book shape that is static with respect to relative price perturbations with Y being their volume effect process (see Becherer et al. [7, Sect. 2.1]).

The contributions of the present paper are threefold:

- (1) We solve the superhedging problem under a transient price impact which is multiplicative, instead of additive.
- (2) Our results account for settlement specifications imposed at maturity which require analysis in *liquidation values* instead of book (paper) values, so that physical units of the underlying risky asset and cash matter at maturity (and as well at the initial time), i.e., terminal (initial) price impact cannot be treated as null. Following the terminology by Bouchard et al. [13], this means that we solve the hedging problem for non-covered instead of covered options (however, see Remark 7.1 for extensions to covered options under transient multiplicative price impact).
- (3) In this realm, the model we study is basically complete, with the transient price impact being the only digression from the frictionless Black–Scholes model assumptions (for \bar{S}), and it yields nontrivial extensions to the classical no-arbitrage pricing and hedging, while avoiding paradoxical effects from illiquidity modelling as mentioned in Çetin et al. [16], without adding further frictions (like transaction costs, or constraints on trading strategies to be “small”). In particular, the large trader neither has the ability to “manipulate” (see Jarrow [24], Bank and Baum [3]) the market to achieve unreasonable profits (see Remark 2.4 and subsequent remarks), nor can he sidestep liquidity costs entirely and trade in effect like a small trader by exploiting modelling artefacts that occur due to a lack of sensible continuity properties (cf. Becherer et al. [9, Sect. 3]).

We formulate the superhedging problem as a stochastic target problem and prove a dynamic programming principle (DPP) along reduced coordinates for *the effective price and impact processes*, which represent the price and impact levels that would

prevail if the large trader were to unwind her (long or short) position in the underlying risky asset immediately. Along the reduced coordinates, the DPP provides a way to compare at stopping times the *instantaneous liquidation wealth* and the (minimal) superhedging price. This permits characterising the superhedging price as the viscosity solution to a nonlinear pricing PDE, which is a semilinear extension of the Black–Scholes equation, with the non-linearity involving the (nonparametric) price impact and resilience functions as well. If the PDE has a sufficiently regular solution, it yields an optimal strategy which even *replicates* the option payoff in the required settlement units. This strategy incorporates the transient nature of impact in that it depends on the effective level of impact. Our analysis is also motivated by analytical tractability. It shows how effects from transience of price impact arise in a basically complete model without other additional frictions from transaction costs or constraints, with scope for results beyond those of the present paper, as outlined in Sect. 7.

While there is a large literature on optimal execution and portfolio optimisation problems under transient price impact, mostly for price impact being additive but also for multiplicative impact (see Obizhaeva and Wang [26], Alfonsi et al. [2], Busseti and Lillo [15] or Guo and Zervos [21], Becherer et al. [8, 9] and references therein), the literature on superhedging (or perfect hedging, i.e., replication) under price impact, as stated above, mostly treats permanent and purely instantaneous price impact (transaction costs, possibly nonlinear) or a combination of the two (see Frey [19], Schönbucher and Wilmott [28], Bank and Baum [3], Çetin et al. [16], Frey and Polte [20]), with the impact often taken in multiplicative form. For the implications of option settlement specifications on hedging, only few papers allow a price impact also at maturity. Clearly, it requires some relevant non-zero price impact at maturity to obtain differences between settlement specifications for options in physical or in cash units, as in Bouchard et al. [12]. However, most articles (see [19, 16, 20]) treat another hedging problem which is not posed in terms of hedgable units of assets, but instead in terms of book (that is “paper”) value, with the price impact at maturity (and possibly initiation) in the analysis effectively taken to be zero. That relates to a different hedging problem for “covered” options (see Remark 7.1). A major difference to the work by Bouchard et al. [12], which offered a fresh view to the hedging problem and inspired ours, is that the analysis in [12] is for permanent and additive impact. In contrast, our hedging results show nontrivial effects from transience of price impact, given that the impact is multiplicative. While the basic example to [12] is the Bachelier model with additive impact, our basic example is a Black–Scholes-type model with transient multiplicative price impact (see Example 2.1 and Remark 2.6) which is the log-linear variant of the model by Obizhaeva and Wang [26]. More detailed comparisons are provided throughout the paper.

The paper is organised as follows. Sections 2 and 3 introduce the model of transient multiplicative price impact and formulate the hedging problem. Effective coordinates for dynamic programming in “liquidation magnitudes” are explained in Sect. 4. Section 5 identifies hedging prices by viscosity solutions to semilinear PDEs (possibly degenerate, with delta constraints), with technical proofs deferred to the Appendix. The results are illustrated by numerical examples in Sect. 6. Finally, Sect. 7 extends the results to combined transient and permanent impact, points out further possible extensions to cross-impact with multiple assets, and comments on related results to the different hedging problem for covered options.

2 A multiplicative transient price impact model

This section describes the model for this paper. An extension with additional permanent impact is described in Sect. 7. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with countably generated \mathcal{F} , a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the usual conditions and an \mathbb{F} -Brownian motion W . We take semimartingales to have càdlàg paths, $\mathbb{R}_{++} = (0, \infty)$ and $\inf \emptyset = +\infty$.

The *unaffected* price process \bar{S} of the underlying risky asset evolves, if the large trader (she) is inactive, according to the stochastic differential equation

$$d\bar{S}_t = \bar{S}_t(\mu_t dt + \sigma dW_t), \quad \bar{S}_0 \in \mathbb{R}_{++},$$

with a constant $\sigma > 0$ and a bounded progressive process μ . The càdlàg adapted process Θ denotes the evolution of her holdings (in units of shares) in the risky asset, say a stock, which is the underlying for the derivative contingent claim in the hedging problem. The market *impact* process $Y = Y^\Theta$ is defined pathwise in the Skorohod space of càdlàg paths by

$$dY_t^\Theta = -h(Y_t^\Theta) dt + d\Theta_t, \quad Y_{0-} = y \in \mathbb{R}, \quad (2.1)$$

for a *resilience function* $h : \mathbb{R} \rightarrow \mathbb{R}$ which is a Lipschitz-continuous function with $\text{sgn}(x)h(x) \geq 0$, as in Becherer et al. [7, 9]. When the large trader trades dynamically according to a strategy Θ , the risky asset price observed on the market, which is the marginal price at which an additional infinitesimal quantity could be traded, is

$$S_t := S_t^\Theta := f(Y_t^\Theta)\bar{S}_t, \quad t \geq 0, \quad (2.2)$$

where the price *impact function* $f : \mathbb{R} \rightarrow \mathbb{R}_{++}$ is increasing and in C^1 with $f(0) = 1$. In particular, $\lambda := f'/f$ is a nonnegative and locally integrable C^0 -function satisfying

$$f(x) = \exp\left(\int_0^x \lambda(u) du\right), \quad x \in \mathbb{R}. \quad (2.3)$$

Example 2.1 The basic example is a transient proportional price impact with unaffected prices \bar{S} given by geometric Brownian motion, as in the Black–Scholes model with $\mu \in \mathbb{R}$ constant, for resilience $h(y) = \beta y$ and log-price impact $\log f(y) = \lambda y$ linear functions with constants $\beta, \lambda \in \mathbb{R}_{++}$. Then the multiplicative price impact is proportional to the number of shares $\Delta\Theta_t = \Theta_t - \Theta_{t-}$ traded at time t , that is, linear in log-prices with

$$\log S_{t+\delta t} - \log S_{t-} = \lambda \Delta\Theta_t$$

with exponential decay $\log S_{t+\delta} = \log(S_t) \exp(-\beta\delta)$ over time when there are no further trades within the time period $(t, t + \delta]$. For such a linear choice of h and $\log f$, the log-asset prices $\log S$ under multiplicative impact evolve like nominal asset prices in the seminal model by Obizhaeva and Wang [26] for additive transient price impact, as described in (1.1).

Our setting also allows the *resilience rate* β (hence h) to be zero, which makes the price impact permanent (cf. Sect. 7) and the log-price impact $\log(S_t/S_{0-})$ linear in $Y_t - Y_{0-} = \Theta_t - \Theta_{0-}$.

Next, we specify the large trader's proceeds (negative expenses) L , which are the variations of her cash account to fund the dynamic holdings Θ in the risky asset. For simplicity, we assume zero interest and a riskless asset with constant price 1 as cash, i.e., prices are discounted in units of this numeraire asset. For continuous strategies Θ of finite variation,

$$L(\Theta) = - \int_0^\cdot S^\Theta d\Theta \quad (2.4)$$

are the proceeds, and there is a unique continuous extension of the functional $\Theta \mapsto L(\Theta)$ in (2.4) to general (bounded) semimartingale strategies Θ , given by

$$L(\Theta) := \int_0^\cdot F(Y_t^\Theta) d\bar{S}_t - \int_0^\cdot \bar{S}_t(fh)(Y_t^\Theta) dt - (\bar{S}F(Y^\Theta) - \bar{S}_0F(Y_{0-}^\Theta)), \quad (2.5)$$

as shown in Becherer et al. [9, Theorem 3.8], with the function

$$F(x) := \int_0^x f(u) du, \quad x \in \mathbb{R}. \quad (2.6)$$

More precisely, every (càdlàg) semimartingale can be approximated (in probability) in the Skorokhod space $D([0, T])$ of càdlàg paths with the Skorokhod M_1 -topology (cf. [9, Sect. 3.1]) by a sequence of continuous processes of finite variation, and for semimartingales $\Theta^n \xrightarrow{\mathbb{P}} \Theta$ in $(D([0, T]), M_1)$ converging to a semimartingale Θ , we then have $L(\Theta^n) \xrightarrow{\mathbb{P}} L(\Theta)$ in $(D([0, T]), M_1)$. To define L by (2.5) is thus natural as the continuous extension of L from (2.4) to all semimartingales.

Remark 2.2 In relation to the above continuous extension, we offer two general comments with regards to 1) literature and 2) subsequent results on hedging, which may be skipped at first reading.

1) For other potential applications, it seems helpful to note that more generally, there is a unique continuous extension even beyond semimartingale strategies; see [9, Sect. 3], and also Horst and Kivman [22] and Ackermann et al. [1] for similar continuity arguments in different applications. For our hedging problem in Sect. 4, however, semimartingale strategies will suffice; see e.g. in (4.3).

2) In Sect. 4, the superhedging problem of Definition 3.2 and the superhedging price (4.4) are going to be defined with respect to a particular set of admissible strategies (see (4.3)). The form of this set (which is as in Bouchard et al. [12]) plays a technical role in proofs for the geometric dynamical programming principle (cf. Theorem 4.1). It would be natural to ask to which extent the particular choice of this set affects the superhedging price. We can offer two (partial) answers to this questions, one of which is again related to suitable continuity properties. At first, we see that in base cases, the superhedging prices w basically recover impact- and frictionless

Black–Scholes prices; see Corollary 5.12 and Remark 3.4 and likewise in [12] (with respect to the Bachelier model). This indicates *that* the superhedging price w defined later in (4.4) is robust in the sense that it does not appear to depend on particularities of the said set. To explain, secondly, *why* such a robustness holds for the almost sure superhedging problem, an almost sure uniform approximation result (in terms of physical asset and cash holdings) of more general trading strategies by a suitable set of more elementary ones would be desirable in principle. Proposition 3.12 in [9] contributes such a result for the set of continuous finite-variation strategies; but this does not quite fit with the setup for the hedging problem in Sect. 4, as the respective set (4.3) there is different.

The proceeds from a block trade of selling $\Delta\Theta_t$ shares at time t are

$$-\bar{S}_t \int_0^{\Delta\Theta_t} f(Y_{t-}^\Theta + x) dx, \quad (2.7)$$

showing that the price per share that the large trader pays (resp. obtains) for a block buy (resp. sell) order is between the price $f(Y_{t-}^\Theta)\bar{S}_t$ before the trade and the price $f(Y_t^\Theta)\bar{S}_t$ after the trade. The form of proceeds and price impact from block trades can be interpreted from the perspective of a latent limit order book, where a block trade is executed against available orders in the order book for prices between $f(Y_{t-}^\Theta)\bar{S}_t$ and $f(Y_{t-}^\Theta + \Delta\Theta_t)\bar{S}_t$, see Becherer et al. [7, Sect. 2.1], and Y can be understood as a volume effect process in the spirit of Predoiu et al. [27].

For a self-financing strategy (B, Θ) in which the dynamic holdings in cash (the riskless asset, savings account) and in the stock (the risky asset) evolve as B and Θ , the self-financing condition is

$$B = B_{0-} + L(\Theta).$$

In order to define a wealth dynamics for the large trader's strategy, it remains to specify the value of the risky asset position Θ in the portfolio in a suitable way. If the large trader is forced to liquidate her position of Θ_t stocks immediately by a hypothetical single block trade at market prices, her *liquidation wealth* $V_t^{\text{liq}} = V_t^{\text{liq}}(\Theta)$ at time $t \geq 0$ (before which the market impact is at Y_t^Θ) is

$$\begin{aligned} V_t^{\text{liq}}(\Theta) &:= B_t + \bar{S}_t \int_0^{\Theta_t} f(Y_t^\Theta - x) dx \\ &= B_{0-} + L(\Theta)_t + \bar{S}_t \int_0^{\Theta_t} f(Y_t^\Theta - x) dx. \end{aligned} \quad (2.8)$$

This wealth process is mathematically conveniently tractable, evolving continuously with

$$dV_t^{\text{liq}} = (F(Y_{t-}) - F(Y_{t-} - \Theta_{t-})) d\bar{S}_t - \bar{S}_t (f(Y_{t-}) - f(Y_{t-} - \Theta_{t-})) h(Y_t) dt \quad (2.9)$$

and $V_0^{\text{liq}} = B_{0-}$, and it inherits from the proceeds (2.5) the continuous dependence properties (on Θ) mentioned above. The notion of liquidation wealth $V^{\text{liq}}(\Theta)$ is relevant for the hedging application of Sect. 3, and is different from the so-called book

wealth process

$$V^{\text{book}}(\Theta) := B + \Theta S = B_{0-} + L(\Theta) + \Theta S, \quad (2.10)$$

in which risky assets are evaluated at the current marginal market price S . Because of price impact (monotonicity of f , positivity of f , \bar{S} , S), clearly $V_t^{\text{liq}} \leq V_t^{\text{book}}$. In the terminology of Jarrow [24, Sect. IV], V^{liq} is real wealth whereas V^{book} is paper wealth. Recently, Kolm and Webster [25] have given theoretical and practical reasons why accounting for the value (respectively the P&L, i.e., the changes in value) of a risky asset position based on current market prices S as in (2.10) can be misleading and needs to be adjusted for price impact; in their terminology, V^{liq} corresponds to fundamental wealth whereas V^{book} is accounting wealth, also referred to as mark-to-market wealth.

From (2.9), we obtain absence of arbitrage within the set of admissible strategies

$$\mathcal{A}^{\text{NA}} := \{(\Theta_t)_{t \geq 0} : \Theta \text{ is a bounded semimartingale with } \Theta_{0-} = 0 \\ \text{and } \Theta_t = 0 \text{ on } t \in [T, \infty) \text{ for some } T \in (0, \infty)\}.$$

Proposition 2.3 *The market is free of arbitrage up to any finite horizon $T \in (0, \infty)$ in the sense that there exists no $\Theta \in \mathcal{A}^{\text{NA}}$ with $\Theta_t = 0$ on $t \in [T, \infty)$ such that for the self-financing strategy (B, Θ) with $V_{0-}^{\text{liq}} := B_{0-} \leq 0$, we have $\mathbb{P}[V_T^{\text{liq}} \geq 0] = 1$ and $\mathbb{P}[V_T^{\text{liq}} > 0] > 0$. Moreover, for any such (B, Θ) , there exists a probability measure \mathbb{Q}^Θ equivalent to \mathbb{P} (on \mathcal{F}_T) such that V^{liq} is a \mathbb{Q}^Θ -martingale.*

In the terminology of [24, Sect. IV, Eq. (13)], the no-arbitrage result of Proposition 2.3 states that there exist no *market manipulation trading strategies*. Note that in contrast, there is no reason to expect a no-arbitrage result in terms of book wealth V^{book} ; there are simple counterexamples, see Example 2.5 for implications on (super-)hedging prices.

Remark 2.4 In the seminal article by Huberman and Stanzl [23], a notion of *no profitable round-trips* (stronger than no-arbitrage) is defined, which (in our notation) requires that there exists no (self-financing) strategy given by (B_{0-}, Θ) as in Proposition 2.3 with $V_0^{\text{liq}} = 0$ and $E[V_T^{\text{liq}}] > 0$. This means that there is no such strategy from zero initial holdings (with $V_0^{\text{liq}} = 0$) that achieves a terminal liquidation wealth which is positive in expectation, $E[V_T^{\text{liq}}] > 0$, within a compact time interval $[0, T]$. By definition, the liquidation wealth V^{liq} is the value of a cash-only position held after all stock holdings are liquidated.

A much cited result from [23] states that price impact needs to be linear to exclude profitable round-trips. This is not in conflict with our modelling, as the proof in [23] relies of course on some assumptions. These include permanent and additive impact. For comparison, under multiplicative permanent price impact, a linear log-price impact function $\log f$ is sufficient to conclude that V^{liq} is a martingale under \mathbb{P} (by (3.2) in Remark 3.4) if \bar{S} is a \mathbb{P} -martingale (e.g. geometric Brownian motion, like in the Black–Scholes model under the risk-neutral measure). This implies $\mathbb{E}[V_T^{\text{liq}}] = \mathbb{E}[V_0^{\text{liq}}]$, excluding profitable round-trips.

Example 2.5 To explain why absence of arbitrage in terms of liquidation values as above (with the corresponding property in terms of book values notably not available) is relevant for almost-sure hedging problems, let us compare hedging in liquidation (“real”) and book (“paper”) values in a simple example within the basic setting of Example 2.1. Consider the European option whose derivative payout (in cash, say) at maturity T is $h(S_T) := S_T(1 - \exp(-\lambda))$ which is strictly positive almost surely. This payoff cannot be superreplicated in terms of liquidation value from nonpositive initial wealth, i.e., there exists no $\Theta \in \mathcal{A}^{\text{NA}}$ with $V_{0-}^{\text{liq}} \leq 0$ and $V_T^{\text{liq}} \geq h(S_T)$. In contrast, in book values, a basic computation shows that the self-financing strategy (Θ, B) with $\mathcal{A}^{\text{NA}} \ni \Theta := 1_{[T, \infty)}$ and $B_{0-} := 0$ satisfies $V_{0-}^{\text{book}} = V_{0-}^{\text{liq}} = 0$, but $V_T^{\text{book}} = h(S_T) > 0$ (whereas $V_T^{\text{liq}} = 0$). This illustrates a major distinction between (super-)hedging problems posed in liquidation values and those posed in book values; see Remark 7.1.

Proof of Proposition 2.3 The idea of the proof is as in [9, Sect. 4], where it was additionally required for admissible strategies that V^{liq} is bounded from below. However, the latter condition can be omitted in the present setup of bounded strategies. To see this, observe that for any $\Theta \in \mathcal{A}^{\text{NA}}$, there exists an equivalent measure $\mathbb{Q}^\Theta \approx \mathbb{P}$ (on \mathcal{F}_T), constructed as in [9, proof of Theorem 4.3], under which the wealth process V^{liq} is a martingale. \square

Remark 2.6 To highlight some key differences to Bouchard et al. [12], let us explain in detail why the basic (log)-linear example for our setup is the Black–Scholes model for \bar{S} (geometric Brownian motion) with multiplicative (proportional, relative) price impact (see Example 2.1), whereas the basic example for [12] is the Bachelier model (additive Brownian motion) with additive linear price impact. Note first that [12, see Eq. (2.1)] study a general model where price impact is permanent and additive in the sense that (using our notation) the resilience h is zero, thus $Y = \Theta$ for $Y_{0-} = \Theta_{0-} := 0$, and the stock price after a small (infinitesimal) trade of size δ becomes $s(\theta + \delta) \approx s(\theta) + \delta f(s(\theta))$, where $f: \mathbb{R} \rightarrow (0, \infty)$ is a smooth function of the current stock price $s(\theta)$ which prevails if the large trader holds θ stocks just before the trade. That means more precisely that $\frac{d}{d\theta} s(\theta) = f(s(\theta))$. For comparison, it is instructive to pretend formally that one could choose a ‘multiplicative’ form $f(x) := \lambda(x)x$. With $\bar{s} := s(0)$, one then would get $s(\theta) = (\exp(\int_0^\theta \lambda(x)dx))\bar{s}$, which is reminiscent of (2.2), (2.3), and taking λ to be constant would give $s(\theta) = \exp(\lambda\theta)\bar{s}$, which is the permanent impact variant of the basic case for multiplicative (transient) impact that is studied in Sect. 5.2. However, a choice like $f(x) = \lambda x$ in linear (multiplicative) form with $\lambda > 0$ does not fit with the assumptions (H1) and (H2) in [12]: neither is $x \mapsto \lambda x$ (strictly) positive on \mathbb{R} , nor is $x \mapsto \exp(\lambda x)$ a surjective function $\mathbb{R} \rightarrow \mathbb{R}$. Observe that asset prices in [12] take values x in \mathbb{R} (instead of $(0, \infty)$). The instructive basic example for their setting is the case of fixed (constant) impact with $f(x) := \lambda > 0$, where $s(\theta) = \bar{s} + \lambda\theta$, and with the unaffected asset price \bar{s} evolving as in the Bachelier model (say), see [12, Sect. 3.4], with an additive permanent price impact. In contrast, the basic example for our setup is transient proportional impact (which is additive and linear in terms of log-prices) with respect to a Black–Scholes-type geometric Brownian motion for \bar{S} ; see Example 2.1.

3 Hedging under transient price impact

We solve in Sects. 3–6 the general problem of dynamic hedging for European options, where the issuer who wants to hedge the option receives at time $t = 0$ the option premium in cash. In an illiquid market setting with price impact, it is relevant to distinguish between *cash settlement* and *physical settlement* of an option payoff because in contrast to frictionless models with unlimited liquidity, moving funds between the bank account and the risky asset account not only induces trading costs from price impact, but also affects the price evolution of the underlying which induces feedback effects; see Frey [19], Schönbucher and Wilmott [28]. Depending on the option's settlement specifications, a terminal block trade at maturity can affect an option's payoff in different ways (see Sect. 6). We consider contingent claims of the following type.

Definition 3.1 A European option with maturity $T \geq 0$ is specified by a measurable map

$$(g_0, g_1) : \mathbb{R}_{++} \times \mathbb{R} \ni (s, y) \mapsto (g_0(s, y), g_1(s, y)) \in \mathbb{R} \times \mathbb{R}$$

representing the payoff, with cash-settlement part g_0 and physical-delivery part g_1 at maturity. It entitles its holder to receive $g_0(S_T, Y_T)$ in cash and $g_1(S_T, Y_T)$ in units of the underlying risky asset, when (S_T, Y_T) describes the risky asset price and the level of market impact at maturity.

Henceforth, T is a fixed maturity time. The optimisation task for the seller, i.e., the issuer, of the option with payoff (g_0, g_1) is to do dynamic hedging at minimal cost to avoid potential losses from her obligation to deliver the payoff at maturity. Among her admissible trading strategies in Γ (specified precisely in Sect. 4.1), she looks for the cheapest strategies to superreplicate the option's payoff in the following sense.

Definition 3.2 A *superhedging strategy* for a European option (g_0, g_1) is a self-financing strategy (B, Θ) with $\Theta \in \Gamma$, $\Theta_{0-} = 0$ and

$$B_T \geq g_0(S_T, Y_T) \quad \text{and} \quad \Theta_T = g_1(S_T, Y_T).$$

We emphasise that a hedging strategy has to deliver the physical component $g_1(S_T, Y_T)$ at maturity exactly, and that any further (long or short) position in the underlying must be unwound before options are settled at the resulting price S_T and impact level Y_T . In particular, a hedging strategy for a payoff with pure cash-delivery part is a so-called round-trip, i.e., it begins and ends with zero shares in the underlying, while the hedging strategy for a payoff with nontrivial physical-delivery part should be such that the amount of risky assets held at maturity exactly meets the physical-delivery requirement. Thus hedging strategies for European contingent claims with physical delivery can be different from those with pure cash-delivery part, and we shall see that their respective prices can also differ.

The (minimal) superhedging price of an option with payoff (g_0, g_1) is the minimal (infimum of) initial capital B_{0-} for which such a superhedging strategy (B, Θ) exists.

Note that via the impact process Y , the hedging strategy Θ clearly affects the volatility of the price process S underlying the option payout, because the price impact in (2.2) is multiplicative.

Options with pure cash settlement are described by $g_1 = 0$. Every (reasonable) option can be represented by a payoff with pure cash settlement. Indeed, if the set Γ is stable under adding additional jumps at maturity, meaning that $\Theta \in \Gamma$ implies that $\Theta + \Delta \mathbb{1}_{\{T\}} \in \Gamma$ for every \mathcal{F}_T -measurable Δ , then any European option can be represented by an option with pure cash settlement. To see this for an option with payoff (g_0, g_1) , define for $(s, y) \in \mathbb{R}_{++} \times \mathbb{R}$ the function

$$H(s, y) := \inf \left\{ g_0 \left(s \frac{f(y + \theta)}{f(y)}, y + \theta \right) + s \frac{F(y + \theta) - F(y)}{f(y)} : \right. \\ \left. \theta = g_1 \left(s \frac{f(y + \theta)}{f(y)}, y + \theta \right) \right\}. \quad (3.1)$$

The value $H(s, y)$ is the minimal amount of cash (riskless assets) needed to hedge the payoff (g_0, g_1) with a single (instant) block trade at maturity, when just before that trade (at time $T-$) the level of impact is y and there are no holdings in the risky asset whose price is s . Indeed, a block trade of size θ will result in the new price $\tilde{s} = sf(y + \theta)/f(y)$ and the impact $\tilde{y} = y + \theta$ and will incur the cost $s(F(y + \theta) - F(y))/f(y)$. Thus it will hedge the claim (g_0, g_1) if $\theta = g_1(\tilde{s}, \tilde{y})$ and if we have enough capital to pay for the block trade and to cover the cash-delivery part which after the block trade equals $g_0(\tilde{s}, \tilde{y})$; see Definition 3.2.

Example 3.3 1) A cash-settled European call option with strike K is specified by the payoff $(g_0(s, y), g_1(s, y)) = ((s - K)^+, 0)$.

2) In comparison, a European call option with strike K and physical settlement has the payoff $(-K \mathbb{1}_{\{s \geq K\}}, \mathbb{1}_{\{s \geq K\}})$. Although the payoff profile (g_0, g_1) does not directly depend on the level of impact y , the equivalent pure cash settlement profile H from (3.1) typically will depend on it if the function λ is not constant. Indeed, the effect on the relative price change $f(y + \theta)/f(y)$ from a block trade θ can depend on the level y of impact before the trade in general, unless $f(x) = \exp(\lambda x)$ for λ constant (linear log-price impact).

Remark 3.4 We now discuss an example to show how the hedging problem for the large trader can be related to hedging in a market with perfect liquidity, but with portfolio constraints, if F from (2.6) is not surjective onto \mathbb{R} . In particular, in this case, our market model will not be complete in the sense that not every contingent claim can be perfectly replicated. A prototypical example is the special case of purely permanent impact, i.e., $h \equiv 0$, with constant λ and log-linear price impact $\log f(x) = \lambda x$ (as in Example 2.1), and an option whose payoff $(H, 0)$ specifies settlement in cash only. Hence we are in the setup of Bank and Baum [3] with the smooth family of semimartingales $P(x, t) := \exp(\lambda x) \tilde{S}_t$. If $Y_{0-} = 0$ and $\lambda = 1$, (2.9) takes the form

$$dV_t^{\text{liq}} = (\exp(\Theta_t) - 1) d\tilde{S}_t. \quad (3.2)$$

By the conditions from Definition 3.2, any hedging strategy Θ satisfies $\Theta_T = 0$, and hence at maturity, we have $S_T = \tilde{S}_T$ and $Y_T = Y_{0-} = 0$. Thus the superreplication condition becomes $V_T^{\text{liq}}(\Theta) \geq H(\tilde{S}_T, 0)$. This means that after a reparametrisation $\Theta \mapsto \exp(\Theta) - 1$ of strategies, the superreplication problem in this large investor model becomes equivalent to the problem in the corresponding frictionless model (for instance, from Black–Scholes) with price process \tilde{S} for a small investor and with constraints on the delta (to be greater than -1), i.e., on the number of risky assets that a hedging strategy must hold. In particular, one expects that in such situations (where F is not invertible), the pricing equation contains gradient constraints. Note that this is different from [3] because for this particular f , the crucial Assumption 5 there is violated, and also different from Bouchard et al. [12] because their assumption (H2) does not hold in this case.

In the presence of resilience for the market impact ($h \neq 0$), the situation becomes more complex because the evolution of the price and impact processes depends on the entire history of the trading strategy, and thus a simplification as above is no longer applicable. But we shall see later in Sect. 5.2 that in the case $f(x) = \exp(\lambda x)$, a lower bound on the delta will also emerge naturally in order to make sense of the pricing equation.

4 Superhedging by geometric dynamic programming

We formulate the superhedging problem as a stochastic target problem and prove a geometric dynamic programming principle (DPP) for the control problem whose value function will be characterised. Notably, we show that a DPP (Theorem 4.1) holds with respect to suitably chosen coordinates, which correspond to modified state processes describing the evolution of *effective* price and impact levels that would result from an immediate unwinding of the risky asset holdings by the large trader. With respect to these new effective coordinates, we characterise the value function of the control problem as a viscosity solution to a partial differential equation, cf. (5.5) and (5.11) in Sect. 5, which is the pricing PDE generalising the (frictionless) Black–Scholes equation.

4.1 Stochastic target formulation

We consider strategies that take values in the constraint set $\mathcal{K} \subseteq \mathbb{R}$, for one of the two cases

$$\mathcal{K} = [-K, +\infty) \quad \text{for some } K > 0, \quad (4.1)$$

$$\mathcal{K} = \mathbb{R}. \quad (4.2)$$

The short-selling constraints (4.1) are needed when F is not surjective onto \mathbb{R} , see Remark 3.4, in which case we consider in Sect. 5.2 $f(x) = \exp(\lambda x)$ for some $\lambda > 0$, while $\mathcal{K} = \mathbb{R}$ will be in force when f is bounded away from 0 and $+\infty$, meaning that the (relative) change of the price from a block trade cannot be arbitrarily big.

For the analysis, we need to allow jumps in the admissible trading strategies in order to obtain a DPP, following Bouchard et al. [12]. For $k \in \mathbb{N}$, let \mathcal{U}_k denote the set of random $\{0, \dots, k\}$ -valued measures ν supported on $[-k, k] \times [0, T]$ that are adapted in the following sense: for every $A \in \mathcal{B}([-k, k])$, the process $t \mapsto \nu(A, [0, t])$ is adapted to the underlying filtration. Note that the elements of \mathcal{U}_k have the representation

$$\nu(A, [0, t]) = \sum_{i=0}^k \mathbb{1}_{\{(\delta_i, \tau_i) \in A \times [0, t]\}},$$

where $0 \leq \tau_1 < \dots < \tau_k \leq T$ are stopping times and each δ_i is a $[-k, k]$ -valued \mathcal{F}_{τ_i} -measurable random variable. Consider also $\mathcal{U} := \bigcup_{k \geq 1} \mathcal{U}_k$.

The admissible trading strategies Θ that we consider are bounded, take values in \mathcal{K} and have the representation

$$\Theta_t = \Theta_{0-} + \int_0^t a_s ds + \int_0^t b_s dW_s + \int_0^t \int_{\mathbb{R}} \delta \nu(d\delta, ds), \quad t \in [0, T], \quad (4.3)$$

in which $\Theta_{0-} \in \mathcal{K}$, $\nu \in \mathcal{U}$ and $(a, b) \in \mathcal{A} := \bigcup_{k \geq 1} \mathcal{A}_k$, where for $k \geq 1$, we define

$$\mathcal{A}_k := \{(a, b) : a \text{ and } b \text{ are predictable with } |a| \vee |b| \leq k \, dt \otimes d\mathbb{P}\text{-a.e.}\}.$$

In this sense, we identify trading strategies by triplets $(a, b, \nu) \in \mathcal{A} \times \mathcal{U}$. For $k \in \mathbb{N}$, set

$$\Gamma_k := \{(a, b, \nu) \in \mathcal{A} \times \mathcal{U}_k : \Theta \text{ from (4.3) takes values in } \mathcal{K} \cap [-k, k]\}$$

and let $\Gamma := \bigcup_{k \geq 1} \Gamma_k$. To reformulate the superhedging problem in our price impact model as a stochastic target problem, consider for a strategy $\gamma \in \Gamma$ and

$$(t, z) = (t, s, y, \theta, v) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R} \times \mathcal{K} \times \mathbb{R}$$

the *state process*

$$(Z_u^{t,z,\gamma})_{u \in [t, T]} = (S_u^{t,z,\gamma}, Y_u^{t,z,\gamma}, \Theta_u^{t,z,\gamma}, V_u^{\text{liq},t,z,\gamma})_{u \in [t, T]},$$

where the processes $S^{t,z,\gamma}$, $Y^{t,z,\gamma}$, $\Theta^{t,z,\gamma}$ and $V^{\text{liq},t,z,\gamma}$ correspond to the price, impact, risky asset position and instantaneous liquidation wealth processes on $[t, T]$ for the control $\Theta^{t,z,\gamma}$ associated with the strategy γ (from a decomposition like (4.3), but on $[t, T]$), when started at time $t-$ at s, y, θ and v , respectively.

According to Sect. 3, for a European option whose payoff in cash and physical units at maturity T is described by the map $\mathbb{R}_{++} \times \mathbb{R} \ni (s, y) \mapsto (g_0(s, y), g_1(s, y))$, a strategy $\gamma \in \Gamma$ is a dynamic superhedging strategy if its state process is a.s. at maturity T within the set

$$\mathfrak{G} := \left\{ (s, y, \theta, v) \in \mathbb{R}_{++} \times \mathbb{R} \times \mathcal{K} \times \mathbb{R} : \theta = g_1(s, y), \right. \\ \left. v - s \frac{F(y) - F(y - \theta)}{f(y)} \geq g_0(s, y) \right\}$$

which we call the *target set*. The superhedging strategies are

$$\mathcal{G}(t, s, y, \theta, v) := \bigcup_{k \geq 1} \mathcal{G}_k(t, s, y, \theta, v),$$

with θ denoting the initial position in the risky asset and

$$\mathcal{G}_k(t, s, y, \theta, v) := \{\gamma \in \Gamma_k : Z_T^{t,s,y,\theta,v,\gamma} \in \mathfrak{G}\}.$$

Our aim is to derive the (minimal) superhedging price

$$\begin{aligned} w(t, s, y) &:= \inf_{k \geq 1} w_k(t, s, y), \\ w_k(t, s, y) &:= \inf\{v : \mathcal{G}_k(t, s, y, 0, v) \neq \emptyset\}, \end{aligned} \quad (4.4)$$

in the case where the hedger holds no units of the underlying asset initially. Let us note that the value function depends on the constraint set \mathcal{K} via the target set \mathfrak{G} . Note also that the set of admissible superhedging strategies, identified with $\mathcal{G}(t, s, y, 0, v)$, is a subset of \mathcal{A}^{NA} , meaning that the superhedging price is strictly positive for a nonnegative payoff H (considered as pure cash-delivery equivalent of (g_0, g_1) in (3.1)) which is positive with non-zero probability.

4.2 Effective coordinates and dynamic programming principle

For stochastic target problems, a form of the dynamic programming principle usually holds and plays a crucial role in deriving a PDE that characterises the value function (in a viscosity sense). The aim of this section is to provide a suitable DPP.

Let us first note that the above formulation for the superhedging problem looks not time-consistent because in the definition (4.4) of the superhedging price w , it is assumed that the initial position in risky assets is zero, which it typically will not be at later times. To obtain a time-consistent formulation, the first naive idea is to make the risky asset position a new variable, i.e., to work with the function \bar{w} defined on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R} \times \mathcal{K}$ by

$$\begin{aligned} \bar{w}(t, s, y, \theta) &:= \inf_{k \geq 1} \bar{w}_k(t, s, y, \theta), \\ \bar{w}_k(t, s, y, \theta) &:= \inf\{v : \mathcal{G}_k(t, s, y, \theta, v) \neq \emptyset\}. \end{aligned} \quad (4.5)$$

But the function $\bar{w}(t, \cdot, \cdot, \cdot, \cdot)$ would have to respect a functional relation along suitable orbits of the coordinates (s, y, θ) at any time t because of the equations (2.2) and (2.7), namely for $\Delta \in \mathbb{R}$ that

$$\bar{w}(t, s, y, \theta) = \bar{w}\left(t, sf(y - \Delta)/f(y), y - \Delta, \theta - \Delta\right) + (s/f(y)) \int_0^\Delta f(y - x) dx.$$

This suggests that one coordinate dimension is redundant and a ‘PDE on curves’ may be required to describe \bar{w} . Indeed, for our transient price impact problem, we show how the state space can be reduced to make the analysis more transparent by

studying the problem in suitably reduced coordinates which can be interpreted as quantities (for prices and impact y) at liquidation (of θ), and with respect to which a DPP and a viscosity characterisation are proved for the function w . While this idea is original and helps to make the analysis more transparent, in other aspects we can and do adapt techniques from Bouchard et al. [12].

To derive a dynamic programming principle for the function w , we want to compare it (evaluated at suitable coordinate processes) over time with the wealth process. Since by definition, w assumes zero initial risky asset holdings, it is natural to consider the (fictitious) state processes that would prevail if the trader were forced to liquidate her position in the risky asset immediately (with a block trade). To this end, let

$$\begin{aligned}\mathcal{S}(S_t, Y_t^\Theta, \Theta_t) &:= \bar{S}_t f(Y_t^\Theta - \Theta_t) = S_t f(Y_t^\Theta - \Theta_t) / f(Y_t^\Theta), \\ \mathcal{Y}(Y_t^\Theta, \Theta_t) &:= Y_t^\Theta - \Theta_t.\end{aligned}\quad (4.6)$$

The process $\mathcal{S}(s, y, \theta)$ is interpreted as the price of the asset that would prevail after θ assets were liquidated, when s and y are the price of the risky asset and the market impact just before the trade, while $\mathcal{Y}(y, \theta)$ would be the level of the market impact after this trade. In this sense, we refer to the processes $\mathcal{S}(S, Y^\Theta, \Theta)$ and $\mathcal{Y}(Y^\Theta, \Theta)$ as the *effective price and impact processes*, respectively, for a self-financing trading strategy Θ . Observe that both processes are continuous even though the trading strategy Θ may have jumps.

For the dynamic programming principle in Theorem 4.1, we compare the *liquidation wealth* V^{liq} defined in (2.8) with the value function w along evolutions of the *effective price and effective impact processes* $(\mathcal{S}(S, Y^\Theta, \Theta), \mathcal{Y}(Y^\Theta, \Theta))$.

Theorem 4.1 *For the geometric DPP, we fix $(t, s, y, v) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R} \times \mathbb{R}$.*

(i) *If $v > w(t, s, y)$, then there exist $\gamma \in \Gamma$ and $\theta \in \mathcal{K}$ such that*

$$V_\tau^{\text{liq}, t, z, \gamma} \geq w(\tau, \mathcal{S}(S_\tau^{t, z, \gamma}, Y_\tau^{t, z, \gamma}, \Theta_\tau^{t, z, \gamma}), Y_\tau^{t, z, \gamma} - \Theta_\tau^{t, z, \gamma})$$

for all stopping times $\tau \geq t$, where $z = (\mathcal{S}(s, y, -\theta), y + \theta, \theta, v)$.

(ii) *Let $k \geq 1$. If $v < w_{2k+2}(t, s, y)$, then for every $\gamma \in \Gamma_k$, $\theta \in \mathcal{K} \cap [-k, k]$ and stopping time $\tau \geq t$, we have with $z = (\mathcal{S}(s, y, -\theta), y + \theta, \theta, v)$ that*

$$\mathbb{P}[V_\tau^{\text{liq}, t, z, \gamma} > w_k(\tau, \mathcal{S}(S_\tau^{t, z, \gamma}, Y_\tau^{t, z, \gamma}, \Theta_\tau^{t, z, \gamma}), Y_\tau^{t, z, \gamma} - \Theta_\tau^{t, z, \gamma})] < 1.$$

Proof There are similarities and differences to Bouchard et al. [12, proof of Proposition 3.3] who treat the case for permanent additive impact; so we present the proof in full detail. As explained in Remark 3.4, the assumptions in [12] do not allow covering multiplicative price impact, and transience of impact naturally requires a further dimension in the DPP. The proof uses general ideas on dynamic programming for stochastic target problems and geometric flows; see Soner and Touzi [29]. We emphasise that for showing the DPP, our proof develops mathematical arguments in terms of effective coordinates and liquidation wealth V^{liq} , which simplifies the mathematical analysis and makes it more transparent. This also shows up in the possible extensions described in Sect. 7.

It is easy to see that for $k \geq 2$ and $(t, s, y, \theta) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R} \times (\mathcal{K} \cap [-k, k])$,

$$\bar{w}_k(t, s, y, \theta) \geq w_{k+1}(t, \mathcal{S}(s, y, \theta), \mathcal{Y}(y, \theta)), \quad (4.7)$$

$$w_{k-1}(t, \mathcal{S}(s, y, \theta), \mathcal{Y}(y, \theta)) \geq \bar{w}_k(t, s, y, \theta). \quad (4.8)$$

Now suppose that $v > w(t, s, y)$. Then by the definition of w , there exist $\theta \in \mathcal{K}$ and some $\gamma \in \mathcal{G}(t, z)$ for $z = (\mathcal{S}(s, y, -\theta), y + \theta, \theta, v)$. As in [29, proof of Theorem 3.1, Step 1], we have for all stopping times $\tau \geq t$ (the first part of) the DPP for \bar{w} , namely $V_\tau^{\text{liq}, t, z, \gamma} \geq \bar{w}(\tau, S_\tau^{t, z, \gamma}, Y_\tau^{t, z, \gamma}, \Theta_\tau^{t, z, \gamma})$. Then (i) follows from (4.7) by taking $k \rightarrow \infty$.

To prove (ii), let $v < w_{2k+2}(t, s, y)$ and suppose that there exist $\gamma \in \Gamma_k$, some $\Theta \in \mathcal{K} \cap [-k, k]$ and a stopping time $\tau \geq t$ such that

$$V_\tau^{\text{liq}, t, z, \gamma} > w_k(\tau, \mathcal{S}(S_\tau^{t, z, \gamma}, Y_\tau^{t, z, \gamma}, \Theta_\tau^{t, z, \gamma}), Y_\tau^{t, z, \gamma} - \Theta_\tau^{t, z, \gamma})$$

for $z = (\mathcal{S}(s, y, -\theta), y + \theta, \theta, v)$. Then by (4.8), we obtain

$$V_\tau^{\text{liq}, t, z, \gamma} > \bar{w}_{k+1}(S_\tau^{t, z, \gamma}, Y_\tau^{t, z, \gamma}, \Theta_\tau^{t, z, \gamma}),$$

and thus we get $v \geq \bar{w}_{2k+1}(t, \mathcal{S}(s, y, -\theta), y + \theta, \theta)$ by [29, proof of Theorem 3.1, Step 2]. In particular, we conclude from (4.7) that $v \geq w_{2k+2}(t, s, y)$, which is a contradiction. \square

Remark 4.2 Part (ii) of the theorem is stated in terms of w_k instead of w because of a measurable selection argument employed in the proof; cf. [12, Remark 3.2].

To derive the pricing PDE from the dynamic programming principle in Theorem 4.1, we need the dynamics of the continuous processes

$$t \mapsto V_t^{\text{liq}} - \varphi(t, \mathcal{S}(S_t, Y_t^\Theta, \Theta_t), \mathcal{Y}(Y_t^\Theta, \Theta_t)) \quad (4.9)$$

for sufficiently smooth functions $\varphi : [0, T] \times \mathbb{R}_{++} \times \mathbb{R}, (t, s, y) \mapsto \varphi(t, s, y)$, that will later serve as test functions when characterising value functions by viscosity solutions.

Lemma 4.3 For every $\gamma = (a, b, v) \in \Gamma$ and every $\varphi \in C^{1,2,1}([0, T] \times \mathbb{R}_{++} \times \mathbb{R})$, we have, for $\Theta = \Theta^\gamma$,

$$\begin{aligned} & d(V_t^{\text{liq}} - \varphi(t, \mathcal{S}_t, \mathcal{Y}_t)) \\ &= \mathcal{S}_t \left(\frac{F(\mathcal{Y}_t + \Theta_t) - F(\mathcal{Y}_t)}{f(\mathcal{Y}_t)} - \varphi_s \right) \left(((\mu_t - \lambda(\mathcal{Y}_t)h(\mathcal{Y}_t + \Theta_t)) dt + \sigma dW_t) \right) \\ &+ (-\varphi_t - \sigma^2 \mathcal{S}_t^2 \varphi_{ss}/2 + h(\mathcal{Y}_t + \Theta_t)\varphi_y + \mathfrak{F}(\mathcal{S}_t, \mathcal{Y}_t, \Theta_t)) dt \end{aligned}$$

with

$$\mathfrak{F}(s, y, \theta) = s h(y + \theta) \left(\lambda(y) \frac{F(y + \theta) - F(y)}{f(y)} - \frac{f(y + \theta) - f(y)}{f(y)} \right),$$

where $\mathcal{S}_t = \mathcal{S}(S_t, Y_t^\Theta, \Theta_t)$, $\mathcal{Y}_t = \mathcal{Y}(Y_t^\Theta, \Theta_t)$ and the derivatives of φ are evaluated at $(t, \mathcal{S}_t, \mathcal{Y}_t)$.

Proof Since $\mathcal{S}_t = \mathcal{S}(S_t, Y_t^\Theta, \Theta_t)$ equals $\bar{S}_t f(Y_t^\Theta - \Theta_t)$, the product rule and $f' = \lambda f$ imply

$$d\mathcal{S}_t = \mathcal{S}_t \left((\mu_t - \lambda(Y_t^\Theta - \Theta_t)h(Y_t^\Theta)) dt + \sigma dW_t \right). \quad (4.10)$$

By Itô's formula, we obtain

$$\begin{aligned} d\varphi(t, \mathcal{S}_t, Y_t^\Theta - \Theta_t) &= \varphi_t dt + \varphi_s d\mathcal{S}_t + \varphi_y d(Y_t^\Theta - \Theta_t) + \varphi_{ss}/2 d[\mathcal{S}]_t \\ &= (\varphi_t - \lambda(Y_t^\Theta - \Theta_t)h(Y_t^\Theta)\mathcal{S}_t\varphi_s - h(Y_t^\Theta)\varphi_y + \sigma^2 \mathcal{S}_t^2 \varphi_{ss}/2) dt \\ &\quad + \mu_t \mathcal{S}_t \varphi_s dt + \sigma \mathcal{S}_t \varphi_s dW_t. \end{aligned} \quad (4.11)$$

With reference to (2.9), we have

$$\begin{aligned} dV_t^{\text{liq}} &= -h(Y_t^\Theta)\mathcal{S}_t \frac{f(Y_t^\Theta) - f(Y_t^\Theta - \Theta_t)}{f(Y_t^\Theta - \Theta_t)} dt + \mu_t \mathcal{S}_t \frac{F(Y_t^\Theta) - F(Y_t^\Theta - \Theta_t)}{f(Y_t^\Theta - \Theta_t)} dt \\ &\quad + \sigma \mathcal{S}_t \frac{F(Y_t^\Theta) - F(Y_t^\Theta - \Theta_t)}{f(Y_t^\Theta - \Theta_t)} dW_t. \end{aligned} \quad (4.12)$$

Combining (4.11) and (4.12) and rearranging terms completes the proof. \square

Remark 4.4 Consider the case when λ is constant, i.e., $f(x) = \exp(\lambda x)$. Then we have $\mathfrak{F} \equiv 0$ and the dynamics of V^{liq} can be stated in a surprisingly simple form, namely

$$dV_t^{\text{liq}} = F(\Theta_t) d\mathcal{S}_t,$$

where $\mathcal{S}_t = \mathcal{S}(S_t, Y_t^\Theta, \Theta_t)$ has the dynamics (4.10). As a consequence, the superhedging price (of the large investor) of an option with maturity T and *pure cash settlement* $H(S_T)$ is at least the small investor's price of H , in the absence of the large trader, when the price process is \bar{S} instead. Indeed, for each (bounded) superhedging strategy Θ (of the large investor) with initial capital v , there exists $\mathbb{P}^\Theta \approx \mathbb{P}$ (on \mathcal{F}_T) such that $\mathcal{S} = S_0 - \mathcal{E}(\sigma \tilde{W})$ under \mathbb{P}^Θ for a \mathbb{P}^Θ -Brownian motion \tilde{W} . Hence $V^{\text{liq}}(\Theta)$ is a \mathbb{P}^Θ -martingale and thus $v \geq \mathbb{E}^{\mathbb{P}^\Theta}[H(S_T)] = \mathbb{E}^{\mathbb{P}^\Theta}[H(\mathcal{S}_T)]$ (recall that $\Theta_T = 0$, implying $S_T = \mathcal{S}_T$). On the other hand, a Feynman–Kac argument shows that $\mathbb{E}^{\mathbb{P}^\Theta}[H(\mathcal{S}_T)]$ is just the classical Black–Scholes price for a small investor in a frictionless market with risky asset process \bar{S} . As Θ was an arbitrary superhedging strategy with initial capital v , taking the infimum yields the claim.

The above observation shows a notable difference to the model in Bank and Baum [3, Theorem 5.3], where the price for the large investor is typically smaller. This is mainly due to a different specification of superhedging strategies with less

stringent settlement constraints, according to which a large trader may be able to reduce at maturity the payoff of the option to a larger extent by exploiting her price impact on the underlying at maturity. In other words, she can vary at maturity her risky asset position in order to minimise the payoff with fewer constraints, and immediately afterwards can unwind any residual risky asset position at no additional cost (by the absence of a bid–ask spread). In contrast, our setup is more restrictive by imposing as settlement constraint on the strategies that they have to replicate the physical delivery part exactly, i.e., after settlement, the hedging strategy has to hold a nonnegative cash position without residual holdings in the risky asset.

We note that an argument as above does not apply in the general case with non-constant λ for our price impact model. In fact, the examples in Sect. 6 also reveal situations where superhedging is cheaper for the large trader; cf. Example 6.1.

5 The pricing PDEs and main results

Next, we determine the terminal value for the function w at maturity date T that will serve as a boundary condition for the pricing PDE. Recall that \mathcal{K} is the (constraint) set in which trading strategies take values and set $\mathcal{K}_n = \mathcal{K} \cap [-n, n]$ for $n \in \mathbb{N}$.

Lemma 5.1 *For the PDE boundary conditions, for $n \in \mathbb{N}$, let*

$$H_n(s, y) := \inf \left\{ g_0 \left(s \frac{f(y+\theta)}{f(y)}, y + \theta \right) + s \frac{F(y+\theta) - F(y)}{f(y)} : \right. \\ \left. \theta \in \mathcal{K}_n, \theta = g_1 \left(s \frac{f(y+\theta)}{f(y)}, y + \theta \right) \right\}.$$

Then we have $w_n(T, \cdot) = H_n(\cdot)$ and $w(T, \cdot) = H(\cdot)$, where the function H is given by

$$H := \inf_{n \geq 0} H_n. \quad (5.1)$$

Proof At the maturity time T , the hedger of the option must do a block trade of size θ in order to meet the physical-delivery part specified by g_1 , thereby moving the price of the underlying from s to $s \frac{f(y+\theta)}{f(y)}$ and the impact level from y to $y + \theta$. Such a block trade incurs costs of size $s \frac{F(y+\theta) - F(y)}{f(y)}$, and hence it superreplicates the payoff (g_0, g_1) if the hedger can cover these costs and the required cash-delivery part, which after the block trade is $g_0(s \frac{f(y+\theta)}{f(y)}, y + \theta)$. \square

Remark 5.2 Note that $H(s, y) = +\infty$ if the equation $\theta = g_1(s \frac{f(y+\theta)}{f(y)}, y + \theta)$ does not have a solution θ in \mathcal{K} .

As we do not know at this point whether the value function w is continuous, we need to work with discontinuous viscosity solutions and hence consider the relaxed semilimits

$$w_*(t, s, y) := \liminf_{(t', s', y', k) \rightarrow (t, s, y, \infty)} w_k(t', s', y'), \quad (5.2)$$

$$w^*(t, s, y) := \limsup_{(t', s', y', k) \rightarrow (t, s, y, \infty)} w_k(t', s', y'), \quad (5.3)$$

where the limits are taken over $t' < T$. Recall that w is a (discontinuous) viscosity solution (of our pricing equations, see Sects. 5.1 and 5.2) if w_* (resp. w^*) is a supersolution (resp. subsolution). For proving the viscosity property, we make the following assumption.

Assumption 5.3 The functions w_* and w^* are bounded on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$, and the payoff function H from (5.1) is regular in the sense that it is continuous, bounded, and the monotone convergence $H_n \downarrow H$ holds uniformly on compacts.

In particular, Assumption 5.3 implies that $w(T, \cdot)$ is finite. This means that the payoff is well behaved in terms of the physical-delivery part, i.e., if the trader was supposed to fulfil her obligation from selling the option immediately, she would be able to do so in any situation (in any state (s, y)) with an admissible trade, provided that she has enough capital.

5.1 Case study for a general bounded price impact function f

In this section, the following assumption is supposed to hold.

Assumption 5.4 The resilience function h is Lipschitz and bounded; the price impact function f is bounded away from 0 and ∞ , i.e., $\inf_{\mathbb{R}} f > 0$ and $\sup_{\mathbb{R}} f < \infty$; λ is bounded and continuously differentiable with bounded derivative; and $\mathcal{K} = \mathbb{R}$ (no delta constraints).

Under Assumption 5.4, the antiderivative F from (2.6) and its inverse F^{-1} are bijections $\mathbb{R} \rightarrow \mathbb{R}$ and Lipschitz-continuous with Lipschitz constants $\sup_{\mathbb{R}} f < \infty$ and $1/\inf_{\mathbb{R}} f$, respectively.

To derive the pricing PDE just formally (at first, to be justified later) in this case, let $(t, s, y) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ and formally apply part (i) of the DPP in Theorem 4.1 to $v = w(t, s, y)$ (assuming that the infimum in the definition of w is attained) and $\tau = t+$, together with Lemma 4.3 for $\varphi = w$, assuming that w is smooth enough. Thus we get the existence of θ^* such that

$$\begin{aligned} 0 \leq & s \left(\frac{F(y + \theta^*) - F(y)}{f(y)} - w_s(t, s, y) \right) \left((\mu_t - \lambda(y)h(y + \theta^*)) dt + \sigma dW_t \right) \\ & + (-w_t(t, s, y) - \sigma^2 s^2 w_{ss}(t, s, y)/2 \\ & + h(y + \theta^*)w_y(t, s, y) + \mathfrak{F}(s, y, \theta^*)) dt. \end{aligned}$$

Still arguing at a formal level, this cannot hold unless

$$F(y + \theta^*) = f(y)w_s(t, s, y) + F(y),$$

$$-w_t(t, s, y) - \sigma^2 s^2 w_{ss}(t, s, y)/2 + h(y + \theta^*)w_y(t, s, y) + \mathfrak{F}(s, y, \theta^*) \geq 0. \quad (5.4)$$

In particular, $\theta^* = \theta^*(t, y, s) = F^{-1}(f(y)w_s(t, s, y) + F(y)) - y$. The second part of the DPP in Theorem 4.1 will actually give that the drift term must be 0, i.e., we should have equality in (5.4). This formally motivates that the form of the pricing PDE for w should be

$$0 = -w_t - \frac{1}{2}\sigma^2 s^2 w_{ss} + \tilde{h}(t, s, y)(w_y + s\lambda(y)w_s) + s\tilde{h}(t, s, y)\left(1 - \frac{\tilde{f}(t, s, y)}{f(y)}\right), \quad (5.5)$$

where for $(t, s, y) \in [0, T) \times \mathbb{R}_{++} \times \mathbb{R}$, we set

$$\tilde{h}(t, s, y) := h \circ F^{-1}(f(y)w_s(t, s, y) + F(y)),$$

$$\tilde{f}(t, s, y) := f \circ F^{-1}(f(y)w_s(t, s, y) + F(y)).$$

Observe that the PDE is semilinear and degenerate (since it does not contain second-order derivatives involving the y -variable). Our main result is as follows.

Theorem 5.5 *Under Assumptions 5.3 and 5.4, the value function w of the superhedging problem is continuous and is the unique bounded viscosity solution to (5.5) with the boundary condition $w(T, \cdot) = H(\cdot)$, where H is defined in (5.1).*

Proof The viscosity property, i.e., that w_* (respectively w^*) is a viscosity supersolution (resp. subsolution), follows by the dynamic programming principle in Theorem 4.1 together with Lemma 4.3. The key arguments are presented in the Appendix in detail for the case where λ is constant, which actually leads to a slightly more involved pricing PDE (5.11) (including gradient constraints) requiring additional justifications.

The comparison result of Theorem A.5 proves uniqueness and continuity; cf. Remark A.7. \square

Let us conclude this section by commenting on some consequences from Theorem 5.5 for the superhedging price and the existence of a corresponding hedging strategy. A numerical example is presented in Sect. 6.

Remark 5.6 Like in the classical case of liquid markets (without price impact), the superhedging price does not depend on the drift in the unperturbed price process. This may be seen more directly by working under the equivalent martingale measure for \bar{S} from the beginning. On the other hand, the superhedging price depends nontrivially on the initial level of impact y and the resilience function h , and can do so even for option payoffs of the form $(g_0(s), 0)$, i.e., payoffs not depending on the level of impact. So it turns out that for pricing and hedging (cf. Remark 5.8), the deviation of the market price from the ‘unaffected’ value, determined by the impact level y , is a relevant state variable.

Remark 5.7 Observe that for only permanent impact, i.e., $h \equiv 0$, (5.5) simplifies to the classical (frictionless) Black–Scholes pricing equation. Hence the superhedging price for the large trader then equals the Black–Scholes price for the option with payoff H .

Remark 5.8 Under sufficient regularity, it turns out that a strategy can be constructed that perfectly replicates the option payout from the (minimal) superhedging price. This means that we have dynamic hedging in the sense of replication, like in the frictionless complete Black–Scholes model.

To this end, suppose that a function $w \in C_b^{1,3,1}([0, T] \times \mathbb{R}_{++} \times \mathbb{R})$ solves the pricing PDE (5.5) with the boundary condition $w(T, \cdot) = H(\cdot)$. Then for any $\varepsilon > 0$, a superhedging strategy with an initial cost of $w(0, s, y) + \varepsilon$ can be constructed as follows. Consider the self-financing strategy (B, Θ) with $B_{0-} = w(0, s, y) + \varepsilon$, $\Theta_0 = F^{-1}(f(y)w_s(0, s, y) + F(y)) - y$, meaning that a block trade of size $\Delta\Theta_0 = \Theta_0$ is performed at time 0, and

$$\Theta_t = F^{-1}\left(f(\mathcal{Y}_t^\Theta)w_s(t, \mathcal{S}(S_t, Y_t^\Theta, \Theta_t), \mathcal{Y}_t^\Theta) + F(\mathcal{Y}_t^\Theta)\right) - \mathcal{Y}_t^\Theta$$

$$\text{for } t \in [0, T), \quad (5.6)$$

$$\Theta_T = 0, \quad \text{i.e., } \Delta\Theta_T = -\Theta_{T-}, \quad (5.7)$$

where $\mathcal{Y}^\Theta = Y^\Theta - \Theta$. Then by Lemma 4.3 together with (5.6) and (5.5), we conclude that

$$\begin{aligned} \varepsilon &= V_0^{\text{liq}}(\Theta) - w(0, s, y) \\ &= V_T^{\text{liq}}(\Theta) - w(T, \mathcal{S}(S_T, Y_T^\Theta, \Theta_T), \mathcal{Y}_T^\Theta) \\ &= V_T^{\text{liq}}(\Theta) - H(\mathcal{S}(S_T, Y_T^\Theta, \Theta_T), \mathcal{Y}_T^\Theta) \\ &= V_T^{\text{liq}}(\Theta) - H(S_T, Y_T^\Theta), \\ \Theta_T &= 0, \end{aligned}$$

where the last line follows from (5.7). By the definition of H , having $H + \varepsilon$ in cash at time T is enough to superreplicate the European claim with payoff (g_0, g_1) by doing a possible additional final block trade of size Δ^ε . Note that such a block trade does not affect V_T^{liq} . Hence the strategy $\Theta + \mathbb{1}_{\{T\}}\Delta^\varepsilon$ is superreplicating for the European claim. Note that one can take $\varepsilon = 0$ if the constructed strategy is bounded and the infimum in the definition of H_n is attained (cf. Lemma 5.1), i.e., we get a replicating strategy in this case.

An application of Itô's formula gives that a strategy Θ satisfying the fixed-point problem (5.6) can be obtained, under suitable regularity, by solving the system

of SDEs

$$\begin{aligned} d\mathcal{S}_t &= \mathcal{S}_t \left((\mu_t - \lambda(\mathcal{Y}_t^\Theta) h(\mathcal{Y}_t^\Theta + \Theta_t)) dt + \sigma dW_t \right), \\ d\Theta_t &= a(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta, \Theta_t) dt + b(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta) dW_t, \\ d\mathcal{Y}_t^\Theta &= -h(\mathcal{Y}_t^\Theta + \Theta_t) dt, \end{aligned} \quad (5.8)$$

with initial conditions $\mathcal{S}_0 = s$, $\mathcal{Y}_0^\Theta = y$ and $\Theta_0 = F^{-1}(f(y)w_s(0, s, y) + F(y)) - y$, where

$$\begin{aligned} a(t, s, y, \theta) &:= h(y + \theta) \left(1 - \frac{\lambda f w_s - f - w_{sy} - \lambda s w_{ss}}{f(F^{-1}(f w_s + F))} \right) \\ &\quad + \frac{w_{ts} + s \mu_t w_{ss} + \frac{1}{2} \sigma^2 s^2 w_{sss}}{f(F^{-1}(f w_s + F))}, \\ b(t, s, y) &:= \frac{\sigma s w_{ss}}{f(F^{-1}(f w_s + F))}, \end{aligned}$$

and where we write $f = f(y)$, $\lambda = \lambda(y)$, etc., when arguments of functions have not been specified, to ease the notation. Thus an optimal (i.e., cheapest) superhedging strategy accounts for the transient nature of price impact, which shows up by the presence of the resilience function h of the impact in the formulas above.

Remark 5.9 To describe how replicating hedging strategies in our model are described by coupled forward–backward SDEs, suppose that Θ is a replicating strategy for an option with cash-equivalent payoff H and let $(\mathcal{Y}, \mathcal{S})$ be the effective impact and price processes. By a change of measure argument, we can assume without loss of generality that $\mu = 0$. Setting $Z_t := \sigma \mathcal{S}_t \frac{F(\mathcal{Y}_t + \Theta_t) - F(\mathcal{Y}_t)}{f(\mathcal{Y}_t)}$, giving

$$\Theta_t = F^{-1}(\sigma^{-1} \mathcal{S}_t^{-1} f(\mathcal{Y}_t) Z_t + F(\mathcal{Y}_t)) - \mathcal{Y}_t,$$

and using (4.12) leads to the coupled FBSDE

$$\begin{aligned} d\mathcal{Y}_t &= -(h \circ F^{-1})(\sigma^{-1} \mathcal{S}_t^{-1} f(\mathcal{Y}_t) Z_t + F(\mathcal{Y}_t)) dt, \\ d\mathcal{S}_t &= \mathcal{S}_t \left(-\lambda(\mathcal{Y}_t) (h \circ F^{-1})(\sigma^{-1} \mathcal{S}_t^{-1} f(\mathcal{Y}_t) Z_t + F(\mathcal{Y}_t)) dt + \sigma dW_t \right), \\ dV_t^{\text{liq}} &= g(\mathcal{Y}_t, \mathcal{S}_t, Z_t) dt + Z_t dW_t, \quad V_T^{\text{liq}} = H(\mathcal{S}_T, \mathcal{Y}_T), \end{aligned}$$

where the driver $g : \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R} \rightarrow \mathbb{R}$ of the FBSDE is given by

$$\begin{aligned} g(y, s, z) &= -s(h \circ F^{-1})(\sigma^{-1} s^{-1} f(y) z + F(y)) \\ &\quad \times \frac{(f \circ F^{-1})(\sigma^{-1} s^{-1} f(y) z + F(y)) - f(y)}{f(y)}. \end{aligned}$$

Example 5.10 As instructive example, consider an option with maturity $T > 0$ whose payout at maturity is the spot price of the asset, i.e., $H(s, y) = s$. In the frictionless Black–Scholes model, its arbitrage-free price is $v^{\text{BS}}(s) = s$ and a (minimal, i.e.,

cheapest) replicating strategy is to buy one share at initiation and hold it until maturity, where it is liquidated at the spot price. For the solution in our price impact model, let us consider the classical solution to (5.5) with the boundary condition H given by the function

$$w(t, s, y) = \frac{F(y + c(t, y)) - F(y)}{f(y)}s, \quad (5.9)$$

where $c : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ is a solution to the backward transport equation

$$\begin{cases} -c_t + h(y + c)c_y = 0 & \text{on } [0, T) \times \mathbb{R}, \\ c(T, y) = F^{-1}(f(y) + F(y)) - y & \text{on } \mathbb{R}. \end{cases}$$

In particular, by the dynamics of c , it holds for any strategy Θ that $c(t, \mathcal{Y}_t^\Theta) = c(0, \mathcal{Y}_0^\Theta)$ for $t \in [0, T]$, where \mathcal{Y}^Θ is the effective impact process corresponding to Θ . In particular, by (5.6), a minimal replicating strategy satisfies on $[0, T)$ the equation

$$\Theta_t^* = c(t, \mathcal{Y}_t^{\Theta^*}) = c(0, \mathcal{Y}_0^{\Theta^*}) = c(0, Y_{0-}).$$

Hence a buy-and-hold strategy is also optimal for the large trader. We can observe the following:

1) Purely permanent impact ($h \equiv 0$) yields the Black–Scholes price $w(t, s, y) = s$ and the buy-and-hold strategy with

$$c(0, y) = c(T, y) = F^{-1}(f(y) + F(y)) - y$$

shares, which does not depend on the maturity T .

2) In comparison, if the price impact is not permanent but transient ($h \neq 0$), the price (5.9) depends nontrivially on the maturity T , in addition to the price impact and resilience functions f and h , respectively.

3) The large trader's price $w(t, s, y)$ dominates the Black–Scholes price $v^{\text{BS}}(t, s)$ (which is equal to s in this example) if and only if $c(t, y) > c(T, y)$. Moreover, there are situations where this condition holds and situations where it is violated. The intuitive reason is that there are two counterbalancing effects: at initiation, where the large trader buys shares to set up the initial delta hedge, thereby moving prices in an unfavourable direction, and at maturity, when she liquidates the delta and moves prices in a direction favourable to her. Which of these two effects dominates overall depends in a nontrivial way on the level of liquidity at initiation and at maturity, and on the settlement specifications of the option; see the discussion in Example 6.1.

Let us comment here on Assumption 5.4 which implies bijectivity of F on \mathbb{R} . Observe that the inverse F^{-1} is used to describe the optimal control θ^* . Similar conditions are also crucial for the results in Bank and Baum [3] and Bouchard et al. [12]; see the surjectivity assumption (A5) in [3] and the invertibility assumption (H2) in [12]. The next section shows how departing from this assumption leads naturally to singularities in the pricing PDE with respect to the gradient. Indeed, the lack of invertibility of F requires conditions on w_s so that θ^* can be derived. Therefore, the analysis there will involve constraints on the ‘delta’, i.e., on the holdings in the risky asset, which in PDE terms translates to constraints on the spatial gradient w_s .

5.2 Case study for price impact of exponential form

We extend the analysis to a natural case where the antiderivative of the price impact function is not assumed to be surjective. To this end, the price impact function is taken to be of exponential form $f(x) = \exp(\lambda x)$ with λ a constant (i.e., $\log f$ is linear), meaning that the relative marginal price impact function $\lambda = f'/f > 0$ is constant. A distinctive feature of this case is that at any time t , knowing the (marginal) stock price S_t is sufficient to determine the impact from an instant block trade, since after a block trade of size Δ , the price is $\bar{S}_t f(Y_t + \Delta) = S_t \exp(\lambda \Delta)$. Hence the relative displacement $f(Y^\Theta)$ of S from the fundamental price \bar{S} is immaterial to determine the price impact from a block trade, in contrast to the situation of Sect. 5.1. Motivated by Remark 3.4, we impose short-selling constraints by requiring trading strategies to evolve in $\mathcal{K} = [-K, \infty)$ for some $K > 0$.

To derive (only heuristically at first, we justify it rigorously later) the pricing PDE, let us apply formally Theorem 4.1 for $v = w(t, s, y)$ at $t, s, y, \tau = t+$, provided that w is smooth enough, to get the existence of $\theta^* \in \mathcal{K}$ such that using Lemma 4.3, we have

$$\mathcal{L}^{\theta^*} w(t, s, y) dt - s(w_s(t, s, y) - e^{\lambda \theta^*} / \lambda + 1 / \lambda)(\sigma dW_t + \eta_t dt) \geq 0, \quad (5.10)$$

where $\eta_t = \mu_t - \lambda h(y + \theta^*)$ and

$$\mathcal{L}^{\theta^*} w(t, s, y) := -w_t(t, s, y) + h(y + \theta^*)w_y(t, s, y) - \frac{1}{2}\sigma^2 s^2 w_{ss}(t, s, y).$$

As in Sect. 5.1, the diffusion part in (5.10) should vanish, giving the optimal control

$$\theta^* = \frac{1}{\lambda} \log(\lambda w_s(t, s, y) + 1),$$

and from the drift part, we identify the pricing PDE $\mathcal{L}^{\theta^*} w(t, s, y) = 0$. The constraint $\theta^* \in \mathcal{K}$ is now equivalent to $\mathcal{H}_{\mathcal{K}} w(t, s, y) \geq 0$, where for a smooth function φ , we set

$$\mathcal{H}_{\mathcal{K}} \varphi(t, s, y) := \lambda \varphi_s(t, s, y) + 1 - e^{-\lambda K}.$$

Thus we conclude formally that w should be a solution to the variational inequality

$$\mathcal{F}_{\mathcal{K}}[w] := \min\{\mathcal{L}^{\theta[w]} w, \mathcal{H}_{\mathcal{K}} w\} = 0 \quad \text{on } [0, T) \times \mathbb{R}_{++} \times \mathbb{R}, \quad (5.11)$$

where

$$\theta[w](t, s, y) := \frac{1}{\lambda} \log(\lambda w_s(t, s, y) + 1). \quad (5.12)$$

As usual, the gradient constraints propagate to the boundary, meaning that the boundary condition for (5.11) should be

$$\min\{w(T, \cdot) - H, \mathcal{H}_{\mathcal{K}} w\} = 0. \quad (5.13)$$

After this motivation, we state the main result for the exponential price impact function $f(x) = \exp(\lambda x)$.

Theorem 5.11 *Suppose that the resilience function h is Lipschitz-continuous and Assumption 5.3 holds. Then the value function w of the superhedging problem is continuous and is the unique bounded viscosity solution to the variational inequality (5.11) with boundary condition (5.13).*

Proof The technical proofs are deferred to the Appendix. The viscosity super-/sub-solution properties are proved in Theorems A.2 and A.3, respectively, while uniqueness and continuity follow from the comparison result of Theorem A.6; cf. Remark A.7. \square

Corollary 5.12 *In the setup from Theorem 5.11, suppose moreover that the payoff function (g_0, g_1) does not depend on the level of impact y , but only on the price s of the underlying. Then the superhedging price is a function in (t, s) only, and the pricing PDE (5.11) simplifies to a Black–Scholes PDE with gradient constraints. In this case, if the face-lifted payoff*

$$F_K[H](s) := \sup_{x \leq 0} \left(H(s+x) + \frac{1 - e^{-\lambda K}}{\lambda} x \right), \quad s \in \mathbb{R}_{++},$$

is continuously differentiable in s with bounded derivative, with the convention that $H = H(0)$ on $(-\infty, 0]$, then the superhedging price (for the large trader) coincides with the frictionless Black–Scholes price for the face-lifted payoff $F_K[H]$.

Proof If (g_0, g_1) is a function of the price s of the underlying only (but not of y), then it is easy to see that H is such as well and that the dimension of the state process can be reduced by omitting the impact process Y . In this case, the stochastic target problem in Sect. 4 can be formulated for the new state process and thus the value function becomes a function of (t, s) only. The same analysis can be carried over to derive the pricing PDE and to prove the viscosity solution property of the value function. The pricing PDE in this case becomes the Black–Scholes PDE with gradient constraints since the term $h(Y)\varphi_y$ in Lemma 4.3 is not present. Hence the superhedging price in our large investor model coincides with the superhedging price under delta constraints in the small investor model for the payoff H (because it solves the same PDE). In this one-dimensional setup, this price coincides with the Black–Scholes price for the face-lifted payoff $F_K[H]$; cf. Chassagneux et al. [17, Proposition 3.1]. \square

6 Numerical examples

We discuss numerical calculations of the superhedging price w characterised by (5.5) to illustrate our results. For the computations, we consider the impact function

$$f(x) = 1 + \arctan(x)/10, \quad x \in \mathbb{R}, \quad (6.1)$$

satisfying Assumption 5.4. Note that $\lambda(x) = 1/(10(1+x^2)f(x))$ varies most within the range of about $(-4, 4)$ and here the change in impact is significant; see Fig. 1(a). Apart from satisfying our assumptions and having the antiderivative

$$F(x) = x + (x \arctan x - \log(1+x^2))/2)/10$$

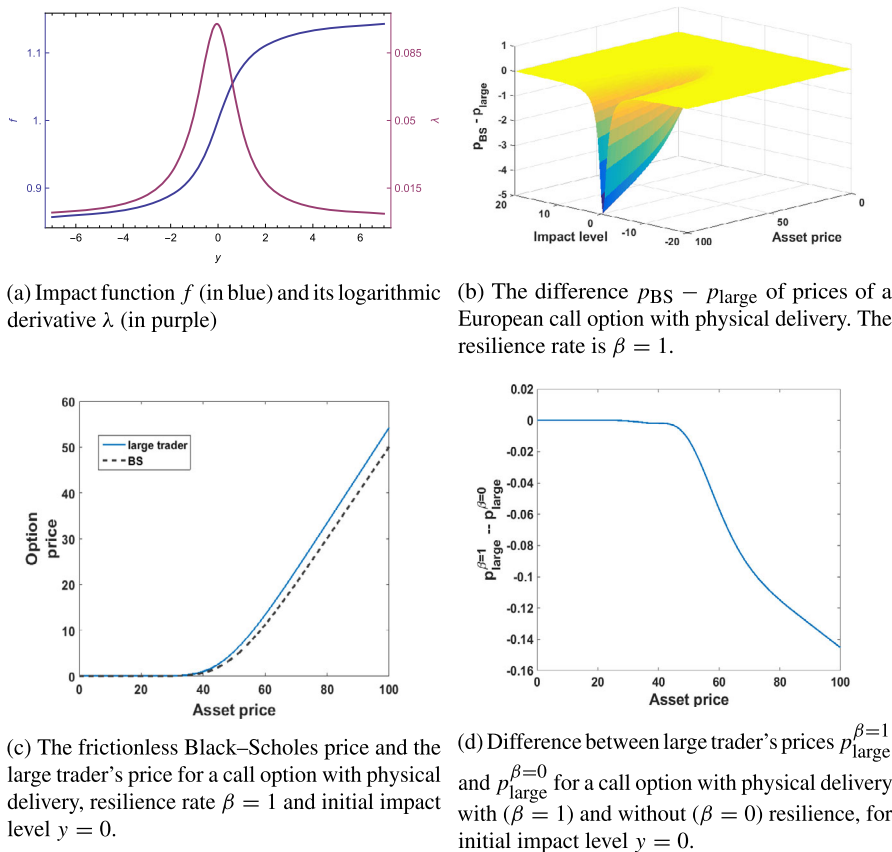


Fig. 1 Numerical computations with the impact function f from (6.1). Other parameters are $\sigma = 0.3$, $T = 0.5$, strike $K = 50$, resilience function $h(y) = \beta y$.

in explicit form, which is useful for the implementation, a similar shape of impact has been observed in the calibration to real data of a related propagator model; see Busseti and Lillo [15, Appendix].

For $h(y) = \beta y$ with $\beta = 1$, we compare the large trader’s price (denoted by p_{large}) for a European call option with physical delivery at maturity $T = 0.5$ and strike $K = 50$, and the option’s frictionless price, i.e., the classical Black–Scholes price (denoted by p_{BS}) of a European call option for the same model parameters. Recall that the case $f = 1$ in our price impact model coincides with the Black–Scholes model. The volatility σ is set to 0.3. The payoff for the large trader is $H(s, y) = (s \frac{F(y+1) - F(y)}{f(y)} - K) \mathbb{1}_{\{s \geq K\}}$ that we “smooth out” by approximating the indicator function by linearly interpolating 0 and 1 between $K - 0.5$ and K .

To approximate both prices, we solve the corresponding PDEs via a (semi-implicit) finite-difference scheme in the bounded region $(y, s) \in [-20, 20] \times [0, 200]$.

For our numerical approximation, we set for $t < T$ the boundary conditions

$$\begin{aligned}\frac{\partial w}{\partial s} &= (F(y+1) - F(y))/f(y) && \text{on } [-20, 20] \times \{200\}, \\ \frac{\partial w}{\partial y} &= 0 && \text{on } (\{-20, 20\} \times [0, 200]) \cup ([-20, 20] \times \{0\}).\end{aligned}$$

Indeed, for initial impact y close to -20 or $+20$, the impact function is approximately constant, and until maturity T , resilience is unlikely to bring back the level of impact to the region where the changes in f are significant; see Fig. 1(a). Thus we expect that the price does not depend much on the level of impact. On the other hand, for larger values of s , one expects the price to depend approximately linearly on s like the payoff profile. The difference between the Black–Scholes price and the large trader’s price (as a function of the risky asset price s and the level of impact y) is shown in Fig. 1(b). Let us point out that the Black–Scholes price does not depend on the level of impact y .

The numerical results of Fig. 1(c) illustrate that the superreplication price for the large trader dominates the frictionless Black–Scholes price for the call option with physical delivery. But we note that this property need not hold in general. For instance, it does not appear to do so for a European call with pure cash delivery where numerical computations show that for the large investor, the price can also be smaller, typically if the impact level at inception is away from zero; see also Example 6.1 below. The intuition for this more complex behaviour is that for pure cash delivery, the net turnover until maturity of the traded assets for a (super-)hedging strategy must be zero (as $\Theta_{0-} = \Theta_T = 0$ then), while non-zero resilience ($h \neq 0$) induces an additional drift which turns out to be less costly to the large trader if it moves the underlying price paths into regions with lower (or zero) option payout.

On the other hand, superhedging becomes more expensive for the large trader when she has to deliver the underlying asset physically at maturity since if the call option settles in the money, she needs to do a final block trade to buy what is lacking in the pre-terminal delta position for the one physical unit required. But this last price impact at maturity is costly in that it further increases the issuer’s call option payout for physical delivery in comparison to cash settlement, where selling the long delta position decreases the payout.

In addition, observe that the presence of resilience renders the level of impact (or the displacement from the fundamental price) a relevant state variable for the problem. For the setup of our numerical example for instance, the price of a European call option with physical delivery, when hedging is initiated at neutral impact level $y = 0$, is cheaper in the presence of resilience than in the case of no resilience, i.e., of only permanent impact; see Fig. 1(d). This is, however, not always the case, for example if the impact at initiation is negative ($y < 0$). To conclude, the dependence on y of the option price is complex. Apart from the drift that the level of impact induces on the prices, it also determines the price impact from intermediate trading and the final trade (enforced by settlement rules). Moreover, we have mentioned examples where superhedging is less or more expensive for the large investor in the presence or absence of resilience.

Example 6.1 The price of a European option in the Black–Scholes model (for a small investor) can indeed be greater than the superhedging price for the large trader of this option with pure cash delivery. To see this, consider for maturity $T > 0$ the solution v^{BS} of the Black–Scholes PDE with a bounded and smooth terminal condition H that has bounded derivatives, where we moreover assume that $\partial_S H \geq 0$; for instance, think of a smooth approximation of a bull call spread option. Note that in particular $\partial_S v^{\text{BS}} \geq 0$ and the derivatives of v^{BS} are bounded. We compare now $v^{\text{BS}}(0, \cdot)$ with $v(0, \cdot, y)$ for large values of y , where $v = w$ with w from Theorem 5.5 with terminal condition H . Note that when $y = Y_{0-} > 0$, the affected price process includes an additional drift in a favourable direction for the large trader.

Let Θ with $\Theta_{0-} = 0$ be such that $\Theta_T = 0$ (corresponding to pure cash delivery at maturity) and for $t \in [0, T-]$, set

$$\Theta_t = F^{-1}(\partial_S v^{\text{BS}}(t, \mathcal{S}_t) f(\mathcal{Y}_t^\Theta) + F(\mathcal{Y}_t^\Theta)) - \mathcal{Y}_t^\Theta, \quad (6.2)$$

where $\mathcal{Y}^\Theta = Y^\Theta - \Theta$ and $\mathcal{S} = f(\mathcal{Y}^\Theta) \bar{S}$. Since v^{BS} is smooth, the arguments in Remark 5.8 ensure the existence of such a Θ , while nonnegativity of $\partial_S v^{\text{BS}}$ implies that $\Theta \geq 0$ on $[0, T]$. Now for the self-financing portfolio (B, Θ) with initial cash holdings $B_{0-} = v^{\text{BS}}(0, S_{0-})$, we have by (4.10), (4.12) and (6.2) (recall that $S_{0-} = \mathcal{S}_0$) that

$$\begin{aligned} V_T^{\text{liq}} &= v^{\text{BS}}(0, \mathcal{S}_0) + \int_0^T \partial_S v^{\text{BS}}(t, \mathcal{S}_t) d\mathcal{S}_t \\ &\quad - \int_0^T \mathcal{S}_t h(Y_t^\Theta) \left(\frac{f(Y_t^\Theta) - f(Y_t^\Theta - \Theta_t)}{F(Y_t^\Theta) - F(Y_t^\Theta - \Theta_t)} - \lambda(Y_t^\Theta - \Theta_t) \right) dt \\ &= H(\mathcal{S}_T) - \int_0^T \mathcal{S}_t h(Y_t^\Theta) \left(\frac{f(Y_t^\Theta) - f(Y_t^\Theta - \Theta_t)}{F(Y_t^\Theta) - F(Y_t^\Theta - \Theta_t)} - \lambda(Y_t^\Theta - \Theta_t) \right) dt. \end{aligned} \quad (6.3)$$

In particular, if the last integrand in (6.3) is negative on $[0, T]$, then (B, Θ) is a superhedging strategy for the large trader with initial capital $B_{0-} = v^{\text{BS}}(0, S_{0-})$ and hence

$$v(0, S_{0-}, Y_{0-}) \leq v^{\text{BS}}(0, S_{0-}). \quad (6.4)$$

One can show that the integrand is negative for instance when $Y^\Theta \geq 0$ on $[0, T]$ and λ is strictly decreasing (at least on a compact set containing the range of Y^Θ and $Y^\Theta - \Theta$); such a situation can arise if for example Y_{0-} is large enough. Alternatively, a negative integrand can also occur if Y^Θ is negative on $[0, T]$, for instance if Y_{0-} is small enough, and λ is strictly increasing. Let us mention that equality in (6.4) cannot hold in general for all values of S_{0-}, Y_{0-} , as this would imply that v does not depend on the initial level of impact Y_{0-} , which is not the case for general payoff functions H ; see e.g. Fig. 1(b).

7 Extensions: permanent price impact, covered options, and cross-impact among multiple illiquid assets

This section explains possible extensions and variations of the previous results on hedging under multiplicative transient price impact. We first show how the results generalise to *combined* transient and permanent price impact, and explain how working in suitable effective coordinates further enables extensions to *multiple illiquid assets*. We also comment on and give references for the solution to the different but related hedging problem for *covered* options.

For $\eta \geq 0$, the marginal price of the risky asset (for an extra infinitesimal quantity) is

$$S_t := f(\eta\Theta_t + Y_t^\Theta)\bar{S}_t, \quad (7.1)$$

in a generalisation of (2.2), with Y^Θ given by (2.1). Following the arguments in Becherer et al. [9, Sect. 5.4], the proceeds from a general semimartingale strategy Θ are (recall that Θ (and Y) may jump at $t = 0$ with Θ_{0-} denoting the initial value)

$$\begin{aligned} \tilde{L}(\Theta) := & \frac{1}{1+\eta} \left(\int_0^\cdot F(\eta\Theta_t + Y_t^\Theta) d\bar{S}_t - \int_0^\cdot \bar{S}_t f(\eta\Theta_t + Y_t^\Theta) h(Y_t^\Theta) dt \right. \\ & \left. - \bar{S}F(\eta\Theta + Y^\Theta) \Big|_{0-} \right). \end{aligned}$$

In particular, a block trade $\Delta\Theta_t$ at time t yields the proceeds

$$-\bar{S}_t \frac{1}{1+\eta} \int_0^{(1+\eta)\Delta\Theta_t} f(\eta\Theta_{t-} + Y_{t-}^\Theta + x) dx.$$

Thus following the discussion in Sect. 2, the volume effect process (in the spirit of Predoiu et al. [27]) in this case is $\eta\Theta + Y^\Theta$ and thereby has a permanent and a transient component. The dynamics of the instantaneous liquidation value process \tilde{V}^{liq} now satisfies

$$\begin{aligned} (1+\eta) d\tilde{V}_t^{\text{liq}} = & (F(\eta\Theta_t + Y_t^\Theta) - F(Y_t^\Theta - \Theta_t)) d\bar{S}_t \\ & - h(Y_t^\Theta)(f(\eta\Theta_t + Y_t^\Theta) - f(Y_t^\Theta - \Theta_t)) dt. \end{aligned}$$

It is worth noting that the generalisation by an additional permanent impact effect does not change the price and impact processes $\mathcal{S}(\mathcal{S}, Y^\Theta, \Theta)$ and $\mathcal{V}(Y^\Theta, \Theta)$ in effective coordinates, because the permanent component vanishes for asset holdings with zero shares in the risky asset. Thanks to this, the previous analysis carries over to additional permanent impact quite seamlessly, with only minor adjustments as follows:

- The boundary condition in Lemma 5.1 needs to be modified by adding the prefactor $1 + \eta$ to θ , when θ appears as an argument of a function.

- In Lemma 4.3, $F(Y^\Theta)$ must be substituted by $F(\eta\Theta + Y^\Theta)$, all the fractions must be divided by $1 + \eta$, and \mathfrak{F} now becomes \mathfrak{F}^η with

$$\mathfrak{F}^\eta(s, y, \theta) := sh(y + \theta) \left(\lambda(y) \frac{F(y + (1 + \eta)\theta) - F(y)}{(1 + \eta)f(y)} - \frac{f(y + (1 + \eta)\theta) - f(y)}{(1 + \eta)f(y)} \right).$$

Let us first discuss the setup of Sect. 5.1 which essentially required F to be invertible. In this case, the pricing PDE has the same structure as (5.5) with modifications $\tilde{h} = \tilde{h}^\eta$ and $\tilde{f} = \tilde{f}^\eta$ replacing the former \tilde{h} and \tilde{f} , namely

$$\begin{aligned} \tilde{h}^\eta(t, s, y) &= h \left(\frac{1}{1 + \eta} F^{-1}((1 + \eta)f(y)\varphi_s(t, s, y) + F(y)) + \frac{\eta}{1 + \eta} y \right), \\ \tilde{f}^\eta(t, s, y) &= f \circ F^{-1}((1 + \eta)f(y)\varphi_s(t, s, y) + F(y)). \end{aligned}$$

An optimal (i.e., cheapest) superhedging strategy Θ^* , if it exists, satisfies (as in Remark 5.8 for $\eta = 0$) the equation

$$(1 + \eta)\Theta_t^* = F^{-1}((1 + \eta)f(\mathcal{Y}_t^*)\varphi_s(t, \mathcal{S}_t^*, \mathcal{Y}_t^*) + F(\mathcal{Y}_t^*)) - \mathcal{Y}_t^*,$$

where $\mathcal{S}^* = \mathcal{S}(S, Y^{\Theta^*}, \Theta^*)$ and $\mathcal{Y}^* = \mathcal{Y}(Y^{\Theta^*}, \Theta^*)$. Hence by this equation, the large trader's optimal strategy also depends on the permanent component of the price impact (which shows by its dependence on η) in addition to the displacement from the fundamental price process tracked by Y^{Θ^*} .

In the setup from Sect. 5.2, we again consider portfolio constraints $\theta \in \mathcal{K}$ for $\mathcal{K} = [-K, \infty)$ in order to derive the pricing PDE. Thanks to $\mathfrak{F}^\eta = 0$, the pricing PDE here simplifies to

$$\begin{aligned} 0 &= \min \left\{ -w_t - \frac{\sigma^2}{2} s^2 w_{ss} + h(y + \theta^*) w_y, \lambda(1 + \eta)\varphi_s + 1 - e^{-\lambda(1 + \eta)K} \right\} \\ &\text{for } (t, s, y) \in [0, T) \times \mathbb{R}_{++} \times \mathbb{R}, \end{aligned}$$

where $\theta^* = \frac{1}{\lambda(1 + \eta)} \log(\lambda(1 + \eta)w_s + 1)$, with boundary condition

$$\min\{w(T, \cdot) - H, \lambda(1 + \eta)\varphi_s + 1 - e^{-\lambda(1 + \eta)K}\} = 0$$

for H being the modified boundary condition from Lemma 5.1, with the modifications for \tilde{h} and \tilde{f} as explained above. In particular, the pricing PDE with permanent impact coincides with the pricing PDE with purely transient impact but with a suitably modified λ , which in this case becomes $\lambda(1 + \eta)$.

Remark 7.1 We now explain how to obtain further results and note key differences in the related hedging problem for so-called *covered options*, as in Bouchard et al. [13] but in contrast for multiplicative transient price impact. Our analysis with effective coordinates carries over similarly by adopting arguments of [13] from the case

of additive permanent price impact, as shown in detail in Becherer and Bilarev [6, Sect. 8].

In contrast to the problem studied in the main body of the present paper and in Bouchard et al. [12] for *non-covered* options, the stochastic target problem for *covered* options is very different in that there is no price impact at inception and at maturity in the hedging problem for *covered* options. The reason (see [13]) is that the buyer of a covered option has to provide (upon request and at the discretion of the hedger) the required initial (delta) hedging position as a part of the option premium, and accepts any mix of cash and stocks (at a suitable book value if evaluated at current marginal market prices S) as an option settlement. In this way, the hedger is not exposed to initial and terminal impact for meeting settlement specifications when forming and unwinding the hedging position for covered options. We mention that similar assumptions are made in the literature by Frey [19], Frey and Polte [20], Çetin et al. [16], where the analysis is in terms of book value instead of liquidation value; see also Bank and Baum [3] and Bouchard et al. [12].

In the previous sections, the superhedging price for (non-covered) options under transient multiplicative price impact was characterised by a degenerate semilinear PDE, whose non-linearity involves the resilience function h and the price impact function f . It can involve gradient constraints (i.e., delta constraints), reducing to the Black–Scholes equation with gradient constraints in the situation of Corollary 5.12.

In contrast, for *covered options*, the corresponding pricing equation turns out to be fully nonlinear and singular in the second-order term. This induces *gamma* constraints, whereas for non-covered options, a singularity arises in the first-order derivative and induces *delta* constraints; see Sect. 5.2. For covered options, it can be shown (see Becherer and Bilarev [6, Sect. 8]) that the resilience of the price impact is immaterial for the hedging price, irrespectively of a particular form for the resilience function, which has been observed likewise in [13, Sect. 4] for additive impact. We emphasise that this is very different to Sect. 5.1 where the resilience function enters the pricing equation in a nontrivial way. It turns out that the current deviation of the asset price from the unaffected price becomes a relevant state variable for describing the solution. Moreover, one can show (see Becherer and Bilarev [6, Remark 8.2, 2])) that the superhedging price is decreasing in the impact function λ in the sense that if $\lambda \geq \tilde{\lambda}$, then the price with respect to λ dominates the one with respect to $\tilde{\lambda}$. For a dual formulation for the hedging of covered options, we refer to Bouchard and Tan [14].

Remark 7.2 As explained in Sect. 4, working in effective coordinates further permits extending results about transient price impact, in additive or multiplicative form, to multiple risky assets with cross-impact from transactions across different assets (described in Bilarev [11, Chap. 5, see Example 5.1.6]). To this end, a key idea is that the impact function needs to be the gradient field of a suitable potential in order to avoid a form of instantaneously profitable round-trips (see [11, Theorem 5.1.4]). Thereby, results like from previous sections (or Bouchard et al. [12] for permanent impact) can be extended to multiple assets in an additive transient cross-impact model. One obtains a geometric DPP and a viscosity PDE to characterise superhedging prices, which involves the resilience function h of the transient impact (see [11, Sect. 5.3.2]). Moreover, under certain conditions, one recovers as instructive reference case again

results as in a multidimensional Bachelier model with its natural pricing formula ([11, Remark 5.3.8]) that does not involve the price impact. This extends to multiple dimensions the instructive one-dimensional linear permanent impact example from [12, Sect. 2.4], which also yields the familiar Bachelier pricing formula. Notice that the hedging strategy is affected by the price impact, though closely related to the usual Bachelier delta-hedging strategy formula, by being computed at liquidation magnitudes for the stock price (i.e., in effective coordinates, analogously to those in (4.6) with \mathcal{S} instead of S). This is entirely analogous to Black–Scholes formula related quantities (for pricing and hedging) occurring under (permanent) multiplicative impact in the basic log-linear example of our model (see Example 2.1 and Remark 2.6).

Appendix: Proofs

This section provides the proofs relegated from Sect. 5, in particular the proof of Theorem 5.11. Recall that in this case, $f(x) = \exp(\lambda x)$ for $\lambda > 0$, and thus the effective price simplifies to $\mathcal{S}(s, y, \theta) = se^{-\lambda\theta} =: \mathcal{S}(s, \theta)$, i.e., the level of impact is not needed in order to determine the price change of a block trade, given the price before the trade. We consider strategies taking values in $\mathcal{K} = [-K, \infty)$ for $K > 0$. This yields a gradient constraint for the PDE that is needed because of a singularity in the PDE, for the expression (5.12) for the form of the optimal strategy to be finitely defined.

First, we verify in Appendix A.1 that if the pricing PDE (5.11) admits a sufficiently smooth classical solution, a replicating strategy in feedback form can be constructed. Such a construction is also needed for the contradiction argument in the proof of the subsolution property in Sect. A.2 where, using smooth test functions, one constructs locally strategies which, roughly speaking, behave like replicating strategies. The viscosity property proofs are collected in Appendix A.2, and in Appendix A.3, we prove comparison results that imply uniqueness of the viscosity solutions of the pricing PDEs and continuity of the value function for the superhedging problem.

A.1 Verification argument for exponential impact function

Suppose that the function $w \in C^{1,2,1}([0, T] \times \mathbb{R}_{++} \times \mathbb{R})$ has the property that for any $(t, s, y) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$, we have

- 1) $\theta[w](t, s, y) \in \mathcal{K}$, recalling the definition in (5.12);
- 2) $\mathcal{L}^{\theta[w](t,s,y)} w(t, s, y) = 0$ when $t < T$;
- 3) $w(T, s, y) = H(s, y)$.

Suppose further that w is sufficiently regular (see Remark A.1 below) so that there exists an admissible strategy $\Theta \in \Gamma$ of the form

$$\begin{aligned} \Theta_t &= \frac{1}{\lambda} \log \left(\lambda w_s(t, \mathcal{S}(S_t, \Theta_t), Y_t - \Theta_t) + 1 \right) \quad \text{for } t \in [0, T), \\ \Theta_T &= 0, \quad \text{i.e., } \Delta \Theta_T = -\Theta_{T-}. \end{aligned} \quad (\text{A.1})$$

In particular, $\Theta_0 = \log(\lambda w_s(0, s, y) + 1)/\lambda$ and $\Delta\Theta_T \in \mathcal{K}$. Consider the self-financing portfolio (β, Θ) with $\beta_{0-} = w(0, s, y)$. Then as in Remark 5.8, we get

$$V_T^{\text{liq}}(\Theta) = H(S_T, Y_T^\Theta), \quad \Theta_T = 0.$$

By the definition of H , this shows that $V_T^{\text{liq}}(\Theta)$ at maturity T is enough capital to (super-)replicate the European claim with payoff (g_0, g_1) with a possible additional block trade (provided that the infima in the definition of H , see Lemma 5.1, are attained). Hence (β, Θ) will be a (super-)replicating strategy for the European claim (g_0, g_1) with initial capital $w(0, s, y)$, which is equal to the superreplication price $w(0, s, y)$, meaning that a replicating hedging strategy as described exists and is an optimal (i.e., cheapest) superreplication strategy under the given assumptions.

Remark A.1 To construct a replicating strategy as in (A.1), we suppose moreover that $w \in C^{1,3,1}([0, T] \times \mathbb{R}_{++} \times \mathbb{R})$ and apply Itô's formula, similarly as in Remark 5.8, to get for $t < T$ the equation

$$\begin{aligned} d\Theta_t &= \frac{1}{\lambda} \left(\frac{1}{\lambda w_s + 1} d(\lambda w_s + 1) - \frac{1}{2(\lambda w_s + 1)^2} d[\lambda w_s + 1]_t \right) \\ &= a(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta, \Theta_t) dt + b(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta) dW_t, \end{aligned}$$

where for $\mathcal{S}_t := \mathcal{S}(S_t, \Theta_t)$ and $\mathcal{Y}_t^\Theta = Y_t^\Theta - \Theta_t$, we set

$$\begin{aligned} a(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta, \Theta_t) &:= \frac{1}{\lambda w_s + 1} \left(w_{ts} + w_{ss} \mathcal{S}_t (\mu_t - \lambda h(Y_t^\Theta)) - w_{sy} h(Y_t^\Theta) \right. \\ &\quad \left. + \frac{1}{2} w_{sss} \sigma^2 \mathcal{S}_t^2 - \frac{\lambda^2 \sigma^2 \mathcal{S}_t^2 w_{ss}}{2(\lambda w_s + 1)} \right), \\ b(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta) &:= \frac{\sigma \mathcal{S}_t w_{ss}}{\lambda w_s + 1}, \end{aligned}$$

with all derivatives of w evaluated at $(t, \mathcal{S}(S_t, \Theta_t), Y_t - \Theta_t)$. Thus a replicating strategy, which is superhedging the payout at a minimal cost (see the arguments preceding the current remark), can be constructed as the $(\Theta_t)_{t \in [0, T]}$ -part (plus a terminal block trade) from a solution, if it exists, to the SDE system, for $t \in [0, T]$,

$$\begin{aligned} d\mathcal{S}_t &= \mathcal{S}_t \left((\mu_t - \lambda h(\mathcal{Y}_t^\Theta + \Theta_t)) dt + \sigma dW_t \right), \\ d\Theta_t &= a(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta, \Theta_t) dt + b(t, \mathcal{S}_t, \mathcal{Y}_t^\Theta) dW_t, \\ d\mathcal{Y}_t^\Theta &= -h(\mathcal{Y}_t^\Theta + \Theta_t) dt, \end{aligned} \tag{A.2}$$

with initial condition $\mathcal{S}_0 = s$, $\mathcal{Y}_0^\Theta = y$ and $\Theta_0 = \log(\lambda w_s(0, s, y) + 1)/\lambda$.

A.2 Viscosity solution property of w for exponential impact function

For the results from Sect. 5.2, we now prove the viscosity property.

Theorem A.2 *The function w_* from (5.2) is a viscosity supersolution of (5.11) on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ with the boundary condition (5.13) on $\{T\} \times \mathbb{R}_{++} \times \mathbb{R}$.*

Proof First, let $(t_0, s_0, y_0) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ and $\varphi \in C_b^\infty([0, T] \times \mathbb{R}_{++} \times \mathbb{R})$ be a smooth function such that we have a strict (meaning uniquely attained) minimum,

$$(\text{strict}) \min_{[0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (w_* - \varphi) = (w_* - \varphi)(t_0, s_0, y_0) = 0.$$

Case 1: Suppose that $\mathcal{H}_K \varphi(t_0, s_0, y_0) < 0$. By the continuity of the operator \mathcal{H}_K , there exists an open neighbourhood \mathcal{O} of (t_0, s_0, y_0) whose closure is contained in $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ such that $\mathcal{H}_K \varphi(t, s, y) < -\varepsilon$ in \mathcal{O} for some $\varepsilon > 0$. Therefore, after possibly shrinking the neighbourhood \mathcal{O} , there exists a constant $k_\varepsilon > 0$ such that

$$s|\varphi_S(t, s, y) + 1/\lambda - e^{\lambda\theta}/\lambda| \geq k_\varepsilon \quad \text{for all } \theta \in \mathcal{K}, (t, s, y) \in \mathcal{O}. \quad (\text{A.3})$$

Let $(t_n, s_n, y_n)_{n \in \mathbb{N}} \subseteq \mathcal{O}$ be a sequence converging to (t_0, s_0, y_0) with

$$w(t_n, s_n, y_n) \longrightarrow w_*(t_0, s_0, y_0),$$

where w_* is the lower-semicontinuous envelope of w . Set $v_n := w(t_n, s_n, y_n) + 1/n$. Since $v_n > w(t_n, s_n, y_n)$, Theorem 4.1 implies the existence of $\theta_n \in \mathcal{K}$ and strategies $\gamma_n \in \Gamma$ such that for stopping times $\tau_n \geq t_n$ (to be suitably chosen later), we have \mathbb{P} -a.s. for $t \in [t_n, T]$ that

$$V_{t \wedge \tau_n}^{\text{liq}, t_n, z_n, \gamma_n} \geq w(\cdot, \mathcal{S}(S^{t_n, z_n, \gamma_n}, \Theta^{t_n, z_n, \gamma_n}), Y^{t_n, z_n, \gamma_n} - \Theta^{t_n, z_n, \gamma_n})_{t \wedge \tau_n}, \quad (\text{A.4})$$

where $z_n = (s_n e^{\lambda\theta_n}, y_n + \theta_n, \theta_n, v_n)$. To abbreviate notation, we write in the sequel n as superscript instead of the full argument (t_n, z_n, γ_n) , so that

$$\mathcal{S}^n := \mathcal{S}(S^{t_n, z_n, \gamma_n}, \Theta^{t_n, z_n, \gamma_n}), \quad \mathcal{Y}^n := Y^{t_n, z_n, \gamma_n} - \Theta^{t_n, z_n, \gamma_n}.$$

Take $\tau_n = \inf\{t \geq t_n : (t, \mathcal{S}_t^n, \mathcal{Y}_t^n) \notin \mathcal{O}\}$, which is the first entrance time of the parabolic boundary of the open region \mathcal{O} . In particular, $\tau_n < T$. Since $w \geq w_* \geq \varphi$ and $w_* - \varphi$ has a strict local minimum at (t_0, s_0, y_0) , there exists $\iota > 0$ such that

$$(w - \varphi)(\tau_n, \mathcal{S}_{\tau_n}^n, \mathcal{Y}_{\tau_n}^n) \geq \iota.$$

Hence \mathbb{P} -a.s., we have $V_{\tau_n}^{\text{liq}, n} - \varphi(\tau_n, \mathcal{S}_{\tau_n}^n, \mathcal{Y}_{\tau_n}^n) \geq \iota$. Now Lemma 4.3 together with the fact that $\mathcal{S}_{t_n}^n = s_n$, $\mathcal{Y}_{t_n}^n = y_n$ gives that \mathbb{P} -a.s.,

$$\begin{aligned} \iota &\leq v_n - \varphi(t_n, s_n, y_n) \\ &\quad - \int_{t_n}^{\tau_n} \mathcal{S}_u^n \left(\varphi_S(u, \mathcal{S}_u^n, \mathcal{Y}_u^n) + 1/\lambda - e^{\lambda\Theta_u^n}/\lambda \right) (\sigma \, dW_u + \zeta_u^n \, du), \end{aligned} \quad (\text{A.5})$$

where

$$\zeta_t^n := \eta_t^n - \frac{\mathcal{L}^{\Theta_t^n} \varphi}{\mathcal{S}_t^n (\varphi_S(u, \mathcal{S}_t^n, \mathcal{Y}_t^n) + 1/\lambda - e^{\lambda\Theta_t^n}/\lambda)} \quad \text{for } t \in [t_n, \tau_n]$$

with $\eta_t^n := \mu_t - \lambda h(Y_t^n)$. Note that ζ_t^n is well defined on $[t_n, \tau_n]$ and uniformly bounded, noting (A.3) and the fact that Y^n is bounded since Θ^n is. Hence by Girsanov's theorem, there exists a measure \mathbb{P}^n equivalent to \mathbb{P} such that

$$\int_{t_n}^{t \wedge \tau_n} \mathcal{S}_u(\varphi_S(u, \mathcal{S}_u^n, \mathcal{Y}_u^n) + 1/\lambda - e^{\lambda \Theta_u}/\lambda)(\sigma dW_u + \zeta_u^n du), \quad t \geq t_n,$$

is a square-integrable martingale under \mathbb{P}^n , as the integrand of the stochastic integral is uniformly bounded because of the definition of τ_n , the continuity of φ_S and the boundedness of the range of Θ , noting $\tau_n \leq T$. Taking expectations under \mathbb{P}^n of the right-hand side of (A.5) leads to $v_n - \varphi(t_n, s_n, y_n) \geq \iota > 0$, which yields a contradiction as by our choice of v_n and of the sequence $(t_n, s_n, y_n)_{n \in \mathbb{N}}$, we have

$$v_n - \varphi(t_n, s_n, y_n) \longrightarrow w_*(t_0, s_0, y_0) - \varphi(t_0, s_0, y_0) = 0.$$

Case 2: From Case 1, we know that $\mathcal{H}_K \varphi(t_0, s_0, y_0) \geq 0$. Hence

$$\theta[\varphi](t_0, s_0, y_0) = \frac{1}{\lambda} \log(\lambda \varphi_S(t_0, s_0, y_0) + 1)$$

is well defined (also in a neighbourhood of (t_0, s_0, y_0)).

Let us suppose that $\mathcal{L}^{\theta[\varphi]} \varphi(t_0, s_0, y_0) < 0$. By the continuity of the operator \mathcal{L} , there exist then an open neighbourhood $\mathcal{O} \subseteq [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ of (t_0, s_0, y_0) and some $r > 0$ and $\varepsilon > 0$ such that

$$\mathcal{L}^\theta \varphi(t, s, y) < -\varepsilon \quad \text{for } (t, s, y) \in \mathcal{O}, \theta \in (\theta[\varphi](t, s, y) - r, \theta[\varphi](t, s, y) + r).$$

In particular, by the continuity of the involved functions, we have (after possibly shrinking the open set \mathcal{O}) that for every $(t, s, y) \in \mathcal{O}$ and for some $r' > 0$,

$$\mathcal{L}^\theta \varphi(t, s, y) < -\varepsilon \quad \text{whenever } |\varphi_S(t, s, y) + 1/\lambda - e^{\lambda \theta}/\lambda| \leq r'.$$

As in Case 1, consider a sequence (t_n, s_n, y_n) in \mathcal{O} which converges to (t_0, s_0, y_0) and such that $w(t_n, s_n, y_n) \rightarrow w_*(t_0, s_0, y_0)$. Set $v_n := w(t_n, s_n, y_n) + 1/n$ and let $\theta_n \in \mathcal{K}$ and strategies $\gamma_n \in \Gamma$ be such that the dynamic programming principle (A.4) holds for the stopping times τ_n that are the first exit times of $(\cdot, \mathcal{S}^n, \mathcal{Y}^n)$ from the set \mathcal{O} . Now a contradiction follows similarly as in Case 1 with the following adjustment: We have

$$\begin{aligned} & V_{t \wedge \tau_n}^{\text{liq}, n} - \varphi(\cdot, \mathcal{S}^n, \mathcal{Y}^n)_{t \wedge \tau_n} \\ &= v_n - \varphi(t_n, s_n, y_n) \\ &\quad - \int_{t_n}^{t \wedge \tau_n} \mathcal{S}_u^n(\varphi_S + 1/\lambda - e^{\lambda \Theta_u^n}/\lambda)(\sigma dW_u + \zeta_u^n du) \\ &\quad + \int_{t_n}^{t \wedge \tau_n} \mathcal{L}^{\Theta_u^n} \varphi(u, \mathcal{S}_u^n, \mathcal{Y}_u^n) \mathbb{1}_{\{|\varphi_S + 1/\lambda - e^{\lambda \Theta_u^n}/\lambda| \leq r'\}} du \\ &\leq v_n - \varphi(t_n, s_n, y_n) - \int_{t_n}^{t \wedge \tau_n} \mathcal{S}_u^n(\varphi_S + 1/\lambda - e^{\lambda \Theta_u^n}/\lambda)(\sigma dW_u + \zeta_u^n du), \end{aligned}$$

where we set

$$\zeta_t^n := \eta_t^n - \frac{\mathcal{L}^{\Theta_t^n} \varphi}{\mathcal{S}_t^n(\varphi_S + 1/\lambda - e^{\lambda \Theta_t^n}/\lambda)} \mathbb{1}_{\{|\varphi_S + 1/\lambda - e^{\lambda \Theta_t^n}/\lambda| \geq r'\}} \quad \text{for } t \in [t_n, \tau_n],$$

with the functions φ and φ_S above evaluated at $(\cdot, \mathcal{S}^n, \mathcal{Y}^n)$. The contradiction now follows by taking expectations under $\mathbb{P}^n \approx \mathbb{P}$ and letting $n \rightarrow \infty$.

Boundary condition. Let $(s_0, y_0) \in \mathbb{R}_{++} \times \mathbb{R}$ and φ be a smooth function with

$$(\text{strict}) \quad \min_{[0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (w_* - \varphi) = (w_* - \varphi)(T, s_0, y_0) = 0.$$

Suppose that

$$\min\{w_*(T, s_0, y_0) - H(s_0, y_0), \mathcal{H}_{\mathcal{K}}\varphi(T, s_0, y_0)\} < 0.$$

The case $\mathcal{H}_{\mathcal{K}}\varphi(T, s_0, y_0) < 0$ leads to a contradiction by the same arguments as in Case 1 above, using that $\mathcal{H}_{\mathcal{K}}\varphi < 0$ in a small neighbourhood of (T, s_0, y_0) . Hence we obtain $\mathcal{H}_{\mathcal{K}}\varphi(T, s_0, y_0) \geq 0$.

Now if $w_*(T, s_0, y_0) < H(s_0, y_0)$, then also $\varphi(T, s_0, y_0) - H(s_0, y_0) < 0$. After possibly modifying the test function φ by $(t, s, y) \mapsto \varphi(t, s, y) - \sqrt{T - t}$, we can assume that $\partial_t \varphi(t, s, y) \rightarrow \infty$ when $t \rightarrow T$, uniformly on compacts. Hence in an ε -neighbourhood $[T - \varepsilon, T] \times B_\varepsilon(s_0, y_0)$ around (T, s_0, y_0) , we have $\mathcal{L}^{\theta[\varphi]}\varphi < 0$. Moreover, after possibly decreasing ε , we have $\varphi(T, \cdot) \leq H(\cdot) - \iota_1$ on $B_\varepsilon(s_0, y_0)$ for some $\iota_1 > 0$. We can argue as in Cases 1 and 2 above, by starting from (t_n, s_n, y_n) in $[T - \varepsilon, T] \times B_\varepsilon(s_0, y_0)$, with $(t_n, s_n, y_n) \rightarrow (T, s_0, y_0)$ and $w(t_n, s_n, y_n) \rightarrow w_*(T, s_0, y_0)$, stopping at the (parabolic) boundary at time τ_n and using $w(T, \cdot) = H(\cdot)$, to get

$$V_{\tau_n}^{\text{liq}, n} - \varphi(\cdot, \mathcal{S}(S^n, \Theta^n), Y^n - \Theta^n)_{\tau_n} \geq \iota_1 \wedge \iota_2,$$

where $\iota_2 := \inf_{[T - \varepsilon, T] \times \partial B_\varepsilon(s_0, y_0)} (w_* - \varphi) > 0$. A contradiction follows as in Case 2 above. \square

Now we prove the subsolution property.

Theorem A.3 *The function w^* from (5.3) is a viscosity subsolution of (5.11) on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ with the boundary condition (5.13) on $\{T\} \times \mathbb{R}_{++} \times \mathbb{R}$.*

Proof The proof is similar to and inspired by the one for the subsolution property in [12, Theorem 3.7]. The reason is that in this case, the gradient constraints ensure that a test function φ that would possibly contradict the subsolution property must satisfy $\mathcal{H}_{\mathcal{K}}\varphi > 0$ locally and hence is sufficiently “nice” to define (locally) control processes (employing the verification argument in Remark A.1) that lead to a contradiction like in [12]. For completeness, we outline differences in the line of proof and sketch the main steps.

Let $\varphi \in C_b^\infty([0, T], \mathbb{R}_{++} \times \mathbb{R})$ be a test function with the property that the point $(t_0, s_0, y_0) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ is a strict (local) maximum of $w^* - \varphi$, i.e.,

$$(\text{strict}) \quad \max_{[0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (w^* - \varphi) = (w^* - \varphi)(t_0, s_0, y_0) = 0.$$

First assume that $t_0 < T$. To ease notation, we use the variable x to denote the pair (s, y) . Because of the special form of the second part of the DPP in Theorem 4.1(ii), we need to employ w_k (instead of w as in the proof of the supersolution property). By Barles [5, Lemma 6.1], we can take a sequence $(k_n, t_n, x_n)_{n \in \mathbb{N}}$ such that $k_n \rightarrow \infty$, any (t_n, x_n) is a local maximum of $w_{k_n}^* - \varphi$ and $(t_n, x_n, w_{k_n}(t_n, x_n)) \rightarrow (t_0, x_0, w^*(t_0, x_0))$.

Assume that $\mathcal{F}_{\mathcal{K}}[\varphi](t_0, x_0) > 0$ and let

$$\varphi_n(t, x) = \varphi(t, x) + |t - t_n|^2 + |y - y_n|^2 + |s - s_n|^4.$$

Then $\mathcal{F}_{\mathcal{K}}[\varphi_n] > 0$ holds in a neighbourhood B of (t_0, x_0) that contains (t_n, x_n) for all n large enough. Since we work on the local neighbourhood B where also $\mathcal{H}_{\mathcal{K}}\varphi_n > 0$, we can modify (in a smooth way) the functions h and φ_n outside of B to be supported on a slightly bigger compact set where $\mathcal{H}_{\mathcal{K}}\varphi_n > 0$ holds. Thus after possibly passing to a suitable subsequence, there exist $\gamma_n \in \Gamma_{k_n}$ such that

$$\Theta_t^{t_n, z_n, \gamma_n} = \frac{1}{\lambda} \log \left(\lambda \frac{\partial \varphi_n}{\partial s}(t, \mathcal{S}_t^{t_n, z_n, \gamma_n}, \mathcal{Y}_t^{t_n, z_n, \gamma_n}) + 1 \right), \quad t \geq t_n,$$

where we set $\mathcal{S}_t^{t_n, z_n, \gamma_n} = \mathcal{S}(S_t^{t_n, z_n, \gamma_n}, \Theta_t^{t_n, z_n, \gamma_n})$ and $\mathcal{Y}_t^{t_n, z_n, \gamma_n} = (Y - \Theta)_t^{t_n, z_n, \gamma_n}$ for $z_n = (s_n, y_n, 0, w_{k_n}(t_n, x_n) - n^{-1})$; see Remark A.1. Let τ_n be the first time after t_n at which the process $(\mathcal{S}_t^{t_n, z_n, \gamma_n}, \mathcal{Y}_t^{t_n, z_n, \gamma_n})_{t \geq t_n}$ leaves B . Like in [12, proof of Theorem 3.7], we conclude by applying Itô's formula, using Lemma 4.3 and $\mathcal{F}_{\mathcal{K}}[\varphi_n] > 0$ on B that \mathbb{P} -a.s.,

$$V_{\tau_n}^{\text{liq}, t_n, z_n, \gamma_n} \geq \varphi_n(\tau_n, \mathcal{S}_{\tau_n}^{t_n, z_n, \gamma_n}, (Y - \Theta)_{\tau_n}^{t_n, z_n, \gamma_n}) + v_n - \varphi_n(t_n, x_n).$$

Now a contradiction follows as in [12, proof of Theorem 3.7, subsolution property, (a)].

For the boundary condition, i.e., the case $t_0 = T$, the arguments are exactly the same as in [12, proof of Theorem 3.7, subsolution property, (b)]. \square

A.3 Comparison results for viscosity solutions

First we provide a comparison result for the pricing PDE (5.5), needed for the proof of Theorem 5.5. Note that (5.5) has the structure

$$\begin{aligned} 0 = & -\varphi_t - \frac{\sigma^2 s^2}{2} \varphi_{ss} \\ & - B_1(y, f(y)\varphi_s)\varphi_y - s B_2(y, f(y)\varphi_s)\varphi_s - s B_3(y, f(y)\varphi_s), \end{aligned} \quad (\text{A.6})$$

where $B_i : \mathbb{R}^2 \rightarrow \mathbb{R}$, $i = 1, 2, 3$, are bounded and Lipschitz-continuous functions. By a change of coordinates, one can transform the PDE as follows.

Lemma A.4 *Let u be a viscosity subsolution (resp. supersolution) of the PDE (A.6). Fix $\kappa > 0$. Then the function \tilde{u} defined by*

$$\tilde{u}(t, s, y) = e^{\kappa t} u(t, sf(y), y) \quad \text{for } (t, s, y) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$$

is a viscosity subsolution (resp. supersolution) of the PDE

$$0 = \kappa\varphi - \varphi_t - \frac{\sigma^2 s^2}{2} \varphi_{ss} - B_1(y, e^{-\kappa t} \varphi_s) \varphi_y + \lambda(y) B_1(y, e^{-\kappa t} \varphi_s) \varphi_s \\ - s B_2(y, e^{-\kappa t} \varphi_s) \varphi_s - e^{\kappa t} s f(y) B_3(y, e^{-\kappa t} \varphi_s). \quad (\text{A.7})$$

Proof To prove the super-(resp. sub-)solution property, take any point (t_0, s_0, y_0) in $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ and a test function $\tilde{\varphi} \in C_b^\infty([0, T] \times \mathbb{R}_{++} \times \mathbb{R})$ for \tilde{u} at (t_0, s_0, y_0) , i.e.,

$$\min_{[0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (\text{resp. max}) (\tilde{u} - \tilde{\varphi}) = \tilde{u}(t_0, s_0, y_0) - \tilde{\varphi}(t_0, s_0, y_0) = 0. \quad (\text{A.8})$$

Consider $\varphi(t, s, y) := e^{-\kappa t} \tilde{\varphi}(t, s/f(y), y)$ for $(t, s, y) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$. We have by definition that $e^{\kappa t} \varphi(t, s f(y), y) = \tilde{\varphi}(t, s, y)$ for $(t, s, y) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}$. In particular, φ is a test function for u at $(t_0, s_0 f(y_0), y_0)$ since by (A.8), we get

$$\min_{[0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (\text{resp. max}) (u - \varphi) = u(t_0, s_0 f(y_0), y_0) - \tilde{\varphi}(t_0, s_0 f(y_0), y_0) \\ = 0. \quad (\text{A.9})$$

We also have

$$\begin{aligned} \tilde{\varphi}_s(t, s, y) &= e^{\kappa t} f(y) \varphi_s(t, s f(y), y), \\ \tilde{\varphi}_{ss}(t, s, y) &= e^{\kappa t} f^2(y) \varphi_{ss}(t, s f(y), y), \\ \tilde{\varphi}_y(t, s, y) &= e^{\kappa t} \lambda(y) f(y) \varphi_s(t, s f(y), y) + e^{\kappa t} \varphi_y(t, s f(y), y) \\ &= \lambda(y) \tilde{\varphi}_s(t, s, y) + e^{\kappa t} \varphi_y(t, s f(y), y), \\ \tilde{\varphi}_t(t, s, y) &= e^{\kappa t} \varphi_t(t, s f(y), y) + \kappa e^{\kappa t} \varphi(t, s f(y), y). \end{aligned}$$

By direct application of these identities, we derive from the right-hand side of (A.7) for $\tilde{\varphi}$ evaluated at (t_0, s_0, y_0) exactly the right-hand side of (A.6) for φ at $(t_0, s_0 f(y_0), y_0)$. By the viscosity property of u and (A.9), we thus conclude that (A.7) holds for $\tilde{\varphi}$ at (t_0, s_0, y_0) with “ \geq ” (resp. “ \leq ”). This proves the claim. \square

By Lemma A.4, it now suffices to prove comparison for (A.7) since this implies a comparison result for (A.6). This is done in the following result.

Theorem A.5 *Let u (respectively v) be a bounded upper-semicontinuous subsolution (resp. lower-semicontinuous supersolution) on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$ of (A.7). Suppose that $u \leq v$ on $\{T\} \times \mathbb{R}_{++} \times \mathbb{R}$. Then $u \leq v$ on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$.*

Proof To prove the claim by contradiction, let us suppose that

$$\sup_{(t, s, y) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (u - v)(t, s, y) > 0.$$

Then we can find $R > 1$ such that with $\mathcal{O}_R := (1/R, R)$, we have

$$\sup_{(t,s,y) \in [0,T] \times \mathcal{O}_R \times [-R,R]} (u - v)(t, s, y) > 0.$$

In particular, there exist $\delta > 0$ and $(t_0, s_0, y_0) \in \overline{\mathcal{O}_R} \times [-R, R]$ with the property that $(u - v)(t_0, s_0, y_0) = \delta > 0$.

Now consider for $n \in \mathbb{N}$ the bounded upper-semicontinuous function

$$\Phi_n(t, s_1, s_2, y_1, y_2) := u(t, s_1, y_1) - v(t, s_2, y_2) - \frac{n}{2}(s_1 - s_2)^2 - \frac{n}{2}(y_1 - y_2)^2.$$

It attains its maximum at some $(t^n, s_1^n, s_2^n, y_1^n, y_2^n) \in [0, T] \times \overline{\mathcal{O}_R}^2 \times [-R, R]^2$ by compactness of that set, and we clearly have

$$\Phi_n(t^n, s_1^n, s_2^n, y_1^n, y_2^n) \geq \delta \quad \text{for all } n \in \mathbb{N}. \quad (\text{A.10})$$

By arguments as in [12, proof of Lemma 3.11], one obtains (after possibly passing to a subsequence) that

$$n(s_1^n - s_2^n)^2 + n(y_1^n - y_2^n)^2 \longrightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (\text{A.11})$$

Note that (A.11) also implies $n(s_1^n - s_2^n)(y_1^n - y_2^n) \rightarrow 0$ as $n \rightarrow \infty$.

Now by Ishii's lemma as stated in Crandall et al. [18, Theorem 8.3], there exist (b^n, X^n, Y^n) in $\mathbb{R} \times S_2 \times S_2$ such that with $p^n = n(s_1^n - s_2^n)$ and $q^n = n(y_1^n - y_2^n)$, we have

$$\begin{aligned} (b^n, (p^n, q^n), X^n) &\in \bar{\mathcal{P}}_{\mathcal{O}_a}^{2,+} u(t^n, s_1^n, y_1^n), \\ (b^n, (p^n, q^n), Y^n) &\in \bar{\mathcal{P}}_{\mathcal{O}_a}^{2,-} v(t^n, s_2^n, y_2^n), \end{aligned}$$

where X^n and Y^n satisfy

$$\begin{pmatrix} X^n & 0 \\ 0 & -Y^n \end{pmatrix} \leq 3n \begin{pmatrix} I_2 & -I_2 \\ -I_2 & I_2 \end{pmatrix}. \quad (\text{A.12})$$

Here, S_2 denotes the set of 2×2 symmetric nonnegative matrices and $I_2 \in S_2$ is the identity matrix. Using the viscosity property of u and v at (t^n, s_1^n, y_1^n) and (t^n, s_2^n, y_2^n) , respectively, we have

$$\begin{aligned} \kappa u(t^n, s_1^n, y_1^n) - b_n - \frac{1}{2}\sigma^2(s_1^n)^2 X_{11}^n + L(s_1^n, y_1^n, p^n, q^n) &\leq 0, \\ \kappa v(t^n, s_2^n, y_2^n) - b_n - \frac{1}{2}\sigma^2(s_2^n)^2 Y_{11}^n + L(s_2^n, y_2^n, p^n, q^n) &\geq 0, \end{aligned}$$

where

$$\begin{aligned} L(t, s, y, p, q) &:= -B_1(y, e^{-\kappa t} p)q + \lambda(y)B_1(y, e^{-\kappa t} p)p \\ &\quad - sB_2(y, e^{-\kappa t} p)p - e^{\kappa t} sf(y)B_3(y, e^{-\kappa t} p). \end{aligned}$$

As a consequence,

$$\begin{aligned} 0 &< \kappa \delta < \kappa (u(t^n, s_1^n, y_1^n) - v(t^n, s_2^n, y_2^n)) \\ &\leq -\frac{1}{2} \sigma^2 (s_2^n)^2 Y_{11}^n + \frac{1}{2} \sigma^2 (s_1^n)^2 X_{11}^n \\ &\quad + L(t^n, s_2^n, y_2^n, p^n, q^n) - L(t^n, s_1^n, y_1^n, p^n, q^n). \end{aligned} \quad (\text{A.13})$$

On the other hand, by (A.12), we get that

$$\frac{1}{2} \sigma^2 (s_1^n)^2 X_{11}^n - \frac{1}{2} \sigma^2 (s_2^n)^2 Y_{11}^n \leq \frac{3}{2} \sigma^2 n (s_1^n - s_2^n)^2,$$

which converges to 0 for $n \rightarrow \infty$ due to (A.11). Let us now analyse the difference $L(t^n, s_2^n, y_2^n, p^n, q^n) - L(t^n, s_1^n, y_1^n, p^n, q^n)$. With C (resp. C_R) denoting a Lipschitz constant (depending on R) that may change from line to line, we get estimates for the corresponding terms via

$$\begin{aligned} |B_1(y_1^n, e^{-\kappa t^n} p^n) q^n - B_1(y_2^n, e^{-\kappa t^n} p^n) q^n| &\leq C |y_1^n - y_2^n| |q^n|, \\ |\lambda(y_1^n) B_1(y_1^n, e^{-\kappa t^n} p^n) p^n \\ &\quad - \lambda(y_2^n) B_1(y_2^n, e^{-\kappa t^n} p^n) p^n| \leq C |y_1^n - y_2^n| |p^n|, \\ |s_1^n B_2(y_1^n, e^{-\kappa t^n} p^n) p^n - s_2^n B_2(y_2^n, e^{-\kappa t^n} p^n) p^n| &\leq C |(s_1^n - s_2^n) p^n| \\ &\quad + C_R |y_1^n - y_2^n| p^n|, \\ |e^{\kappa t^n} s_1^n f(y_1^n) B_3(y_1^n, e^{-\kappa t^n} p^n) \\ &\quad - e^{\kappa t^n} s_2^n f(y_2^n) B_3(y_2^n, e^{-\kappa t^n} p^n)| \leq C_R (|s_1^n - s_2^n| + |y_1^n - y_2^n|). \end{aligned}$$

As all estimates from above vanish for $n \rightarrow \infty$, the right-hand side in (A.13) is bounded by something that converges to 0 as $n \rightarrow \infty$. But this yields a contradiction for large n . \square

Because of lack of a precise reference, we provide a comparison result also in the case of delta constraints leading to the variational inequality (5.11).

Theorem A.6 *Suppose that the resilience function h is Lipschitz-continuous and Assumption 5.3 holds. Let u (resp. v) be a bounded upper-(resp. lower-)semicontinuous viscosity subsolution (resp. supersolution) of the variational inequality (5.11) with the terminal condition (5.13). Then $u \leq v$ on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$.*

Proof We argue by contradiction. For any $a > 0$, set $\mathcal{O}_a := [a, \infty) \times [-1/a, 1/a]$. If $\sup_{[0, T] \times \mathbb{R}_{++} \times \mathbb{R}} (u - v) > 0$, there exists $a > 0$ with $\sup_{[0, T] \times \mathcal{O}_a} (u - v) > 0$. For $\kappa > 0$, consider $\tilde{u} := e^{\kappa t} u$ and $\tilde{v} := e^{\kappa t} v$. Then \tilde{u} (resp. \tilde{v}) is a viscosity sub-(resp. super-)solution of

$$\min\{\kappa \varphi + \tilde{\mathcal{L}}[\varphi], \mathcal{H}_{\mathcal{K}, t} \varphi\} = 0$$

with the boundary condition $\min\{\varphi(T, \cdot) - H(\cdot), \mathcal{H}_{\mathcal{K}, T}\varphi\} = 0$, where

$$\tilde{\mathcal{L}}[\varphi](t, s, y) = -\partial_t \varphi + h(y + \log(\lambda e^{-\kappa t} \partial_s \varphi + 1)/\lambda) \partial_y \varphi - \sigma^2 s^2 \partial_{ss} \varphi / 2$$

and $\mathcal{H}_{\mathcal{K}, t}\varphi = \lambda e^{-\kappa t} \partial_s \varphi + 1 - e^{-\lambda K}$ for $t \in [0, T]$.

Consider

$$\Theta_n := \sup_{(t, x_1, x_2) \in [0, T] \times \mathcal{O}_a^2} \left(\tilde{u}(t, x_1) - \tilde{v}(t, x_2) - \frac{n}{2} |x_1 - x_2|^2 \right).$$

We have $\Theta_n > \iota$ for some $\iota > 0$. Since $\tilde{u} - \tilde{v}$ is upper-semicontinuous, the supremum is attained for any n as a maximum at some (t_n, x_1^n, x_2^n) in the compact set $[0, T] \times \mathcal{O}_a^2$. By arguments as in [12, proof of Lemma 3.11], after possibly passing to a subsequence, we obtain

$$\lim_{n \rightarrow \infty} \Theta_n = \sup_{[0, T] \times \mathcal{O}_a} (\tilde{v} - \tilde{u}) \geq \iota > 0, \quad (\text{A.14})$$

$$n |x_1^n - x_2^n|^2 \longrightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (\text{A.15})$$

Note also that

$$\lim_{n \rightarrow \infty} (\tilde{u}(t_n, x_1^n) - \tilde{v}(t_n, x_2^n)) \geq \iota. \quad (\text{A.16})$$

Case 1: Suppose, after passing to a subsequence, that $t_n = T$ for all n . Then Ishii's lemma together with the viscosity property of \tilde{u} and \tilde{v} gives

$$\min\{\tilde{u}(T, x_1^n) - H(x_1^n), \lambda e^{-\kappa T} p_n + 1 - e^{-\lambda K}\} \leq 0,$$

$$\min\{\tilde{v}(T, x_2^n) - H(x_2^n), \lambda e^{-\kappa T} p_n + 1 - e^{-\lambda K}\} \geq 0,$$

where $p_n = n(s_1^n - s_2^n)$. Hence we conclude that $\tilde{u}(T, x_1^n) \leq H(x_1^n)$ for all n . However, in this case, since $\tilde{v}(T, x_2^n) \geq H(x_2^n)$ for all n , we have

$$\tilde{v}(T, x_2^n) \geq H(x_2^n) \geq H(x_2^n) - H(x_1^n) + \tilde{u}(T, x_1^n),$$

which contradicts (A.16) for large n by continuity of H .

Case 2: We can now assume (after passing to a subsequence) that $t_n < T$ for all n . Set

$$p_n := n(s_1^n - s_2^n), \quad q_n := n(y_1^n - y_2^n).$$

By Ishii's lemma, see [18, Theorem 8.3], using the viscosity property of \tilde{u} and \tilde{v} , there exist $a_n \in \mathbb{R}$ and symmetric 2×2 matrices A_n, B_n (that satisfy a bound like in (A.12)) with

$$(a_n, (p_n, q_n), A_n) \in \bar{\mathcal{P}}_{\mathcal{O}_a}^{2,+} \tilde{u}(t_n, x_1^n), \quad (a_n, (p_n, q_n), B_n) \in \bar{\mathcal{P}}_{\mathcal{O}_a}^{2,-} \tilde{v}(t_n, x_2^n),$$

and such that

$$\begin{aligned} \min \left\{ -a_n + L(t_n, x_1^n, \tilde{u}(t_n, x_1^n), p_n, q_n, A_n), \lambda e^{-\kappa t_n} p_n + 1 - e^{-\lambda K} \right\} &\leq 0, \\ \min \left\{ -a_n + L(t_n, x_2^n, \tilde{v}(t_n, x_2^n), p_n, q_n, B_n), \lambda e^{-\kappa t_n} p_n + 1 - e^{-\lambda K} \right\} &\geq 0, \end{aligned}$$

where for $t \in [0, T]$, $x = (x_1, y_1) \in \mathbb{R}^2$, $\ell, p, q \in \mathbb{R}$ and a 2×2 matrix A , we define

$$L(t, x, \ell, p, q, A) := \kappa \ell + h(y_1 + \log(\lambda e^{-\kappa t} p + 1)/\lambda)q - \sigma^2 x_1^2 A_{11}/2.$$

Therefore we have $-a_n + L(t_n, x_1^n, \tilde{u}(t_n, x_1^n), p_n, q_n, A_n) \leq 0$.

On the set $\{(t, y, p) \in [0, T] \times \mathbb{R} \times \mathbb{R} : \lambda e^{-\kappa t} p + 1 - e^{-\lambda K} \geq 0\}$, the function

$$(t, y, p) \mapsto h(y + \log(\lambda e^{-\kappa t} p + 1)/\lambda)$$

is Lipschitz-continuous. Thus we are exactly in the setup of the proof of Theorem A.5, and a contradiction argument yields the claim like there: One gets the estimate

$$\kappa(\tilde{u}(t_n, x_1^n) - \tilde{v}(t_n, x_2^n)) \leq C(n|x_1^n - x_2^n|^2 + 1/n)$$

for a constant $C > 0$ not depending on n , which contradicts (A.14) for n large. \square

Remark A.7 By Theorems A.2 and A.3, w_* (resp. w^*) is a supersolution (subsolution) of (5.11) with boundary condition (5.13). By Theorem A.6, we have $w_* \geq w^*$ on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$. It is clear by definition that $w_* \leq w^*$, and so $w_* = w^*$ coincide on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$. On the other hand, we have $w_* \leq w \leq w^*$ on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$. To show equality also at $t = T$, note that the super-(sub-)solution property of w_* (respectively w^*) implies $w_*(T, \cdot) \geq H(\cdot)$ and $w_*(T, \cdot) \leq H(\cdot)$. So w_* equals H at T . Since also $H(\cdot) = w(T, \cdot)$, the equality $w_* = w^* = w$ holds on $\{T\} \times \mathbb{R}_{++} \times \mathbb{R}$. Hence $w_* = w^* = w$ on $[0, T] \times \mathbb{R}_{++} \times \mathbb{R}$, which implies continuity.

The same conclusion holds for (5.5) with the boundary condition (5.1).

Funding Open Access funding enabled and organized by Projekt DEAL. Support by the German Science Foundation DFG is gratefully acknowledged.

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

- Ackermann, J., Kruse, T., Urusov, M.: Reducing Obizhaeva–Wang type trade execution problems to LQ stochastic control problems (2022). To appear in *Finance Stoch.* Available online at [arXiv:2206.03772](https://arxiv.org/abs/2206.03772)
- Alfonsi, A., Schied, A., Slynko, A.: Order book resilience, price manipulation, and the positive portfolio problem. *SIAM J. Financ. Math.* **3**, 511–533 (2012)
- Bank, P., Baum, D.: Hedging and portfolio optimization in financial markets with a large trader. *Math. Finance* **14**, 1–18 (2004)
- Bank, P., Soner, H.M., Voß, M.: Hedging with temporary price impact. *Math. Financ. Econ.* **11**, 215–239 (2017)
- Barles, G.: *An Introduction to the Theory of Viscosity Solutions for First-Order Hamilton–Jacobi Equations and Applications*. Springer, Berlin (2013)
- Becherer, D., Bilarev, T.: Hedging with transient price impact for non-covered and covered options (2018). Preprint, Available online at [arXiv:1807.05917v1](https://arxiv.org/abs/1807.05917v1)
- Becherer, D., Bilarev, T., Frentrup, P.: Optimal asset liquidation with multiplicative transient price impact. *Appl. Math. Optim.* **78**, 643–676 (2018)
- Becherer, D., Bilarev, T., Frentrup, P.: Optimal liquidation under stochastic liquidity. *Finance Stoch.* **22**, 39–68 (2018)
- Becherer, D., Bilarev, T., Frentrup, P.: Stability for gains from large investors’ strategies in M1/J1 topologies. *Bernoulli* **25**, 1105–1140 (2019)
- Bertsimas, D., Lo, A.W.: Optimal control of execution costs. *J. Financ. Mark.* **1**, 1–50 (1998)
- Bilarev, T.: Feedback effects in stochastic control problems with liquidity frictions. PhD thesis, Humboldt-Universität zu Berlin (2018). Available online at <https://doi.org/10.18452/19592>
- Bouchard, B., Loeper, G., Zou, Y.: Almost-sure hedging with permanent price impact. *Finance Stoch.* **20**, 741–771 (2016)
- Bouchard, B., Loeper, G., Zou, Y.: Hedging of covered options with linear market impact and gamma constraint. *SIAM J. Control Optim.* **55**, 3319–3348 (2017)
- Bouchard, B., Tan, X.: Understanding the dual formulation for the hedging of path-dependent options with price impact. *Ann. Appl. Probab.* **32**, 1705–1733 (2022)
- Busseti, E., Lillo, F.: Calibration of optimal execution of financial transactions in the presence of transient market impact. *J. Stat. Mech. Theory Exp.* **2012**, P09010 (2012)
- Çetin, U., Soner, H.M., Touzi, N.: Option hedging for small investors under liquidity costs. *Finance Stoch.* **14**, 317–341 (2010)
- Chassagneux, J.-F., Elie, R., Kharroubi, I.: When terminal facelift enforces delta constraints. *Finance Stoch.* **19**, 329–362 (2015)
- Crandall, M.G., Ishii, H., Lions, P.-L.: User’s guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc. (N.S.)* **27**, 1–67 (1992)
- Frey, R.: Perfect option hedging for a large trader. *Finance Stoch.* **2**, 115–141 (1998)
- Frey, R., Polte, U.: Nonlinear Black–Scholes equations in finance: associated control problems and properties of solutions. *SIAM J. Control Optim.* **49**, 185–204 (2011)
- Guo, X., Zervos, M.: Optimal execution with multiplicative price impact. *SIAM J. Financ. Math.* **6**, 281–306 (2015)
- Horst, U., Kivman, E.: Optimal trade execution under small market impact and portfolio liquidation with semimartingale strategies. To appear in *Finance Stoch.* (2024). Available online at [arXiv:2103.05957](https://arxiv.org/abs/2103.05957)
- Huberman, G., Stanzl, W.: Price manipulation and quasi-arbitrage. *Econometrica* **72**, 1247–1275 (2004)
- Jarrow, R.A.: Market manipulation, bubbles, corners, and short squeezes. *J. Financ. Quant. Anal.* **27**, 311–336 (1992)
- Kolm, P.N., Webster, K.: Do you really know your P&L? The importance of impact-adjusting the P&L (2023). Preprint, Available online at <https://ssrn.com/abstract=4331027>
- Obizhaeva, A., Wang, J.: Optimal trading strategy and supply/demand dynamics. *J. Financ. Mark.* **16**, 1–32 (2013)
- Predoiu, S., Shaikhet, G., Shreve, S.: Optimal execution in a general one-sided limit-order book. *SIAM J. Financ. Math.* **2**, 183–212 (2011)
- Schönbucher, P.J., Wilmott, P.: The feedback effect of hedging in illiquid markets. *SIAM J. Appl. Math.* **61**, 232–272 (2000)
- Soner, H.M., Touzi, N.: Dynamic programming for stochastic target problems and geometric flows. *J. Eur. Math. Soc.* **4**, 201–236 (2002)

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