Finance and Stochastics (2022) 26:335–358
https://doi.org/10.1007/s00780-022-00473-y

A scaling limit for utility indifference prices in the discretised Bachelier model

Asaf Cohen[1] · **Yan Dolinsky**[2]

Received: 1 March 2021 / Accepted: 29 November 2021 / Published online: 1 March 2022

Abstract
We consider the discretised Bachelier model where hedging is done on a set of equidistant times. Exponential utility indifference prices are studied for path-dependent European options, and we compute their non-trivial scaling limit for a large number of trading times n and when risk aversion is scaled like $n\ell$ for some constant $\ell > 0$. Our analysis is purely probabilistic. We first use a duality argument to transform the problem into an optimal drift control problem with a penalty term. We further use martingale techniques and strong invariance principles and obtain that the limiting problem takes the form of a volatility control problem.

Keywords Utility indifference · Strong approximations · Path-dependent SDEs · Asymptotic analysis

Mathematics Subject Classification (2020) 91G10 · 60F15

JEL Classification G11 · C65

1 Introduction

Taking into account market frictions is an important challenge in financial modelling. In this paper, we focus on the friction that the rebalancing of the portfolio strategy is limited to occur discretely. In such a realistic situation, a general future payoff cannot

A. Cohen acknowledges the financial support of the National Science Foundation (DMS-2006305).
Y. Dolinsky is supported in part by the GIF Grant 1489-304.6/2019 and the ISF grant 230/21.

✉ Y. Dolinsky
yan.dolinsky@mail.huji.ac.il

A. Cohen
shloshim@gmail.com

1 Department of Mathematics, University of Michigan, Ann Arbor, MI, 48109, USA

2 Department of Statistics, Hebrew University, Jerusalem, Israel

be hedged perfectly even in complete market models such as the Bachelier model or the Black–Scholes model.

We consider the hedging of a path-dependent European contingent claim in the Bachelier model for the setup where the investor can hedge on a set of equidistant times. Our main result provides the asymptotic behaviour of the exponential utility indifference prices when the risk aversion goes to infinity linearly in the number of trading times. Namely, we establish a non-trivial scaling limit for indifference prices when the friction goes to zero and the risk aversion goes to infinity.

This type of scaling limits goes back to the seminal work of Barles and Soner [2] who determined the scaling limit of utility indifference prices of vanilla options for small proportional transaction costs and high risk aversion. Another work in this direction is the recent article by Bank and Dolinsky [1] which deals with scaling limits of utility indifference prices of vanilla options for hedging with vanishing delay $H \downarrow 0$ when the risk aversion is scaled like A/H for some constant A. In general, the common ground between the above two works and the present paper it that all of them start with complete markets and consider small frictions, which make the markets incomplete so that the derivative securities cannot be perfectly hedged with a reasonable initial capital. Then, instead of considering perfect hedging, these papers study utility indifference prices with exponential utilities and with large risk aversion. In contrast to the previous two papers which treated only vanilla (path-independent) options, in this paper, we are able to provide a limit theorem for path-dependent options.

Although the topic of discrete-time hedging in the Brownian setting was largely studied, the corresponding papers rather studied the optimal discretisation of given hedging strategies or stochastic integrals. Indeed, Bertsimas et al. [3], Gobet and Temam [18] and Hayashi and Mykland [19] studied the convergence rate of discrete-time delta-hedging strategies for the case where the trading is done on a set of equidistant times. In Geiss [14], the author proposed to discretise delta-hedging strategies with non-equidistant deterministic time nets and showed that this generalisation leads (for some payoffs) to better error estimates. For further research in this direction, see Geiss [15], Geiss and Toivola [16] and Gobet and Makhlouf [17]. The papers by Fukasawa [11, 12] and Cai et al. [4] study the approximation of stochastic integrals with a discretisation procedure that goes beyond deterministic nets and is performed on a set of random (stopping) times.

In the present study, instead of tracking a given hedging strategy, we follow the well-known approach of utility indifference pricing which is commonly used in the setup of incomplete markets (see Carmona [5, Chap. 2] and the references therein). In other words, this approach says that the price of a given contingent claim should be equal to the minimal amount of money that an investor has to be offered so that she becomes indifferent (in terms of utility) between the situation where she has sold the claim and the one where she has not.

We now put our contribution in the context of asymptotic analysis of risk-sensitive control problems. Such problems model situations where a decision maker aims to minimise small probability events with significant impacts. Typically, the small probability event emerges from a state process with a volatility that vanishes with the scaling parameter. The limiting behaviour of such risk-sensitive control problems is governed by deterministic differential games; see e.g. Fleming [8] and the references therein. In our case, the volatility of the state process does not vanish, but is rather of order $O(1)$, and the small probability event emerges from the discrete approximation of a stochastic integral on a grid. As a result, the structure of the limiting problem is quite different: rather than a deterministic (drift control) differential game, we obtain a stochastic (volatility) control problem. The connection between the indifference price and the difference between two values of two (non-asymptotically) risk-sensitive problems is well known; see e.g. Hernández-Hernández [20]. It stems from the fact that utility indifference pricing is a normalised version of the certainty equivalent criterion. Our study is concerned with the scaling limit of the utility indifference prices when the market friction goes to zero and the risk aversion goes to infinity.

Let us outline the key steps in establishing the asymptotic result. Our approach is purely probabilistic and allows to consider European contingent claims with path-dependent payoffs. The first step in establishing the main result goes through a dual representation of the value function (Proposition 3.1). This representation is closer in nature to the form of the limiting stochastic volatility control problem. In the dual problem, there is only one player: a maximising (adverse) player controls the drift, and the investor's role is translated into a martingale condition. The control's cost is small, hence allowing the maximiser to choose controls with high values. The second and main step is to analyse the limit behaviour of the dual representation. This is the main technical challenge of the paper and is done in Theorem 3.1. The proof is done via upper and lower bounds.

There are two main challenges in the proof of the upper bound. The first one is comparing between the penalty term of the dual problem that takes the form of a Kullback–Leibler divergence and the penalty term of the limiting problem. The second one is due to the fact that the consistent price systems appearing in the dual representation are not necessarily tight. To handle the first challenge, we work on a discrete-time grid, and for any level of penalty in the limiting problem, we are able to construct an optimal penalty term in the prelimit problem (Lemma 4.1). The structure of this best penalty term also serves us in the proof of the lower bound. We overcome the second difficulty by applying a strong invariance principle (Lemma 4.2) and not the widely used weak convergence approach which is not helpful here. Specifically, for any control in the prelimit dual problem, we construct a stochastic integral which is close in probability to the original control. This family of stochastic integrals is the set of controls in the limiting problem.

The proof of the lower bound is achieved by an explicit construction driven by Lemma 4.1. A key ingredient in the proof is a construction of a path-dependent stochastic differential equation (SDE) which translates the consistent price systems given via the drift control into the limiting volatility control problem.

The rest of the paper is organised as follows. In Sect. 2, we introduce the model and formulate the main result (Theorem 2.1). In Sect. 3, we provide the duality representation (Proposition 3.1), we formulate the main technical statement (Theorem 3.1), and use these two results to deduce Theorem 2.1. The proof of Theorem 3.1 is given in Sect. 4.

2 The model and the main results

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space carrying a one-dimensional Wiener process $(W_t)_{t\in[0,T]}$ with natural augmented filtration $(\mathcal{F}_t^W)_{t\in[0,T]}$ and time horizon $T \in (0, \infty)$. We consider a simple financial market with a riskless savings account bearing zero interest (for simplicity) and with a risky asset X with Bachelier price dynamics

$$X_t = X_0 + \sigma W_t + \mu t, \qquad t \in [0, T], \tag{2.1}$$

where $X_0 > 0$ is the initial asset price, $\sigma > 0$ is the constant volatility and $\mu \in \mathbb{R}$ is the constant drift. These parameters are fixed **throughout the paper**.

Fix $n \in \mathbb{N}$ and consider an investor who can trade the risky asset only at times from the grid $\{0, T/n, 2T/n, \dots, T\}$. For technical reasons, in addition to the risky asset $(X_t)_{t\in[0,T]}$, we assume that the financial market contains short-horizon options with payoffs in the spirit of power options. Formally, for any $k = 0, 1, \dots, n$, at time kT/n, the investor can buy but not sell European options that can be exercised at the time $(k+1)T/n$ with the payoff $|X_{(k+1)T/n} - X_{kT/n}|^3$. Denote by $h(n)$ the price of the above option. For simplicity, we assume that the price does not depend on k. Moreover, we assume the scalings

$$\lim_{n\to\infty} n^{3/2} h(n) = \infty \tag{2.2}$$

and

$$\lim_{n\to\infty} n h(n) = 0. \tag{2.3}$$

This investment opportunity can be viewed as an insurance against high values of the stock fluctuations. Roughly speaking, the term $|X_{(k+1)T/n} - X_{kT/n}|^3$ is of order $O(n^{-3/2})$. However, since the payoff of this option is quite extreme, we expect that the corresponding price will be more expensive than $O(n^{-3/2})$. The scaling given by (2.2) and (2.3) says that the option price is more expensive than $O(n^{-3/2})$, but cheaper than $O(1/n)$.

Remark 2.1 It is possible to replace the payoff $|X_{(k+1)T/n} - X_{kT/n}|^3$ with the payoff $|X_{(k+1)T/n} - X_{kT/n}|^{2+\epsilon}$ for $\epsilon > 0$ and assume that the new option price $\hat{h}(n)$ satisfies

$$\lim_{n\to\infty} n\hat{h}(n) = 0 \qquad \text{and} \qquad \lim_{n\to\infty} n^{1+\epsilon/2}\hat{h}(n) = \infty.$$

However, for simplicity we work with power 3.

In line with the above, the set $\mathcal{A}^n$ of trading strategies for the n-step model consists of pairs $(\gamma, \delta) = ((\gamma_k, \delta_k))_{0\le k\le n-1}$ such that for any k, the random variables γ_k, δ_k are $\mathcal{F}^W_{kT/n}$-measurable and in addition $\delta_k \ge 0$ (there is no short selling in the power options). The corresponding portfolio value at the grid times is given by

$$V^{\gamma,\delta}_{\frac{kT}{n}} := \sum_{i=0}^{k-1} \gamma_i (X_{(i+1)T/n} - X_{iT/n}) \\ + \sum_{i=0}^{k-1} \delta_i \big(|X_{(i+1)T/n} - X_{iT/n}|^3 - h(n)\big), \qquad k = 0, 1, \ldots, n. \tag{2.4}$$

Next, let $\mathcal{C}[0,T]$ be the space of continuous functions $z : [0,T] \to \mathbb{R}$ equipped with the uniform norm $\|z\| := \sup_{0\le t\le T} |z_t|$ and let $f : \mathcal{C}[0,T] \to [0,\infty)$ be Lipschitz-continuous with respect to the uniform norm. For the n-step model, we consider a European option with payoff of the form $f^n(X) := f(p^n(X))$, where $p^n(z)$ returns the linear interpolation of $((kT/n, z_{kT/n}) : k = 0,\ldots,n)$ for any function $z : [0,T] \to \mathbb{R}$. It is immediate from the Lipschitz-continuity of f that it has linear growth, and consequently that

$$\mathbb{E}_{\mathbb{P}}\big[\exp\big(\alpha f^n(X)\big)\big] < \infty, \qquad \forall \alpha \in \mathbb{R}. \tag{2.5}$$

The investor assesses the quality of a hedge by the resulting expected utility. Assuming exponential utility with constant absolute risk aversion $\lambda > 0$, the *utility indifference price* $\pi(n,\lambda)$ and the *certainty equivalent* $c(n,\lambda)$ of the claim $f^n(X)$ do not depend on the investor's initial wealth and, respectively, take the well-known forms

$$\pi(n,\lambda) := \frac{1}{\lambda} \log \frac{\inf_{(\gamma,\delta)\in\mathcal{A}^n} \mathbb{E}_{\mathbb{P}}[\exp(-\lambda(V_T^{\gamma,\delta} - f^n(X)))]}{\inf_{(\gamma,\delta)\in\mathcal{A}^n} \mathbb{E}_{\mathbb{P}}[\exp(-\lambda V_T^{\gamma,\delta})]}$$

and

$$c(n,\lambda) := \frac{1}{\lambda} \log\Big(\inf_{(\gamma,\delta)\in\mathcal{A}^n} \mathbb{E}_{\mathbb{P}}\Big[\exp\Big(-\lambda\big(V_T^{\gamma,\delta} - f^n(X)\big)\Big)\Big]\Big). \tag{2.6}$$

Note that

$$\pi(n,\lambda) = c(n,\lambda) - \frac{1}{\lambda}\log\Big(\inf_{(\gamma,\delta)\in\mathcal{A}^n} \mathbb{E}_{\mathbb{P}}[\exp(-\lambda V_T^{\gamma,\delta})]\Big). \tag{2.7}$$

The following scaling limits are the main results of the paper. The proof is given at the end of Sect. 3.

Theorem 2.1 *For $n \to \infty$ and re-scaled risk aversion $n\ell$ with $\ell > 0$ fixed, the certainty equivalent and the utility indifference price of $f^n(X)$ have the scaling limits*

$$\lim_{n\to\infty} c(n, n\ell) = \lim_{n\to\infty} \pi(n, n\ell) \\ = \pi(\ell) := \sup_{\nu\in\mathcal{V}} \mathbb{E}_{\mathbb{P}}\Big[f(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\Big(\frac{\nu_t}{\sigma^2}\Big)dt\Big], \tag{2.8}$$

where

$$X_t^{(\nu)} := X_0 + \int_0^t \sqrt{\nu}_u dW_u, \qquad t \in [0, T],$$

$$g(y) := y - \log y - 1, \qquad y > 0,$$

and $\mathcal{V}$ is the class of all bounded, nonnegative $(\mathcal{F}_t^W)$-predictable processes ν.

Let us finish this section with the following three remarks.

Remark 2.2 The scaling (2.3) is a technical one and needed for the proof of the upper bound. More precisely, (2.3) is used in Lemma 4.2 where we implement a strong invariance principle from Dolinsky [7]. Whether this scaling can be removed and the result still holds true is an interesting question. We expect that the proof of such a result would require additional machinery from dynamic programming and nonlinear partial differential equations. We leave this challenging question for future research.

The scaling (2.2) is used in the proof of the lower bound. Let us notice that if we allow power options (with power 3) and (2.2) does not hold, i.e., $\lim_{n\to\infty} n^{3/2}h(n) < \infty$, then there will be an additional constraint on the process ν which appears on the right-hand side of (2.8). As a result, in general, the scaling limit of the utility indifference prices will be less than or equal to that given by (2.8). Since we added power options from technical reasons, we assume the scaling (2.3), which says that these options are not too cheap.

Remark 2.3 Observe that if we take ℓ to infinity, then we have

$$\lim_{\ell\to\infty} \pi(\ell) = \sup_{\nu\in\mathcal{V}} \mathbb{E}_{\mathbb{P}}[f(X^{(\nu)})].$$

The above right-hand side can be viewed as a model-free option price; see Galichon et al. [13]. This corresponds to the case where the investor wants to superreplicate the payoff $f(X)$ without any assumptions on the volatility. For the case where $\ell \to 0$, it is straightforward to check that

$$\lim_{\ell\to 0} \pi(\ell) = \mathbb{E}_{\mathbb{P}}[f(X^{(\sigma)})],$$

where

$$X_t^{(\sigma)} = X_0 + \sigma W_t, \qquad t \in [0, T].$$

In other words, we converge to the unique price in the continuous-time complete market given by (2.1).

Remark 2.4 A natural question is whether Theorem 2.1 can be extended to the case where the risky asset $(X_t)_{t\in[0,T]}$ is given by a geometric Brownian motion, in other words, whether our scaling limit is valid for the Black–Scholes model.

The immediate conjecture is that for the Black–Scholes given by

$$\frac{d\mathrm{X}_t}{\mathrm{X}_t} = \sigma dW_t + \mu dt,$$

the scaling limit of the utility indifference prices takes the from

$$\sup_{\nu\in\mathcal{V}} \mathbb{E}_{\mathbb{P}}\left[f(\mathrm{X}^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right],$$

where

$$\mathrm{X}_t^{(\nu)} = X_0 \exp\left(\int_0^t \nu_s dW_s - \frac{1}{2}\int_0^t \nu_s^2 ds\right), \qquad t\in[0,T].$$

A proof of this conjecture is far from obvious. First, in our duality result in Proposition 3.1, we need to assume (see (3.5)) that the exponential moments of X exist. This is no longer the case for geometric Brownian motion, and so even the duality requires additional ideas. The second difficulty is to formulate and prove the “correct” analogue of Lemmas 4.1 and 4.2. At this stage, we leave the Black–Scholes setup for future research.

3 A dual representation and a scaling limit

Set $n\in\mathbb{N}$. Denote by $\mathcal{Q}^n$ the set of all probability measures $\mathbb{Q}\approx\mathbb{P}$ with finite entropy

$$\mathbb{E}_{\mathbb{Q}}\left[\log\frac{d\mathbb{Q}}{d\mathbb{P}}\right] < \infty$$

relative to $\mathbb{P}$ and such that the processes

$$(X_{kT/n})_{0\le k\le n} \quad \text{and} \quad \left(\sum_{i=0}^{k-1}|X_{(i+1)T/n} - X_{iT/n}|^3 - kh(n)\right)_{0\le k\le n}$$

are, respectively, a $\mathbb{Q}$-martingale and a $\mathbb{Q}$-supermartingale with respect to the filtration $(\mathcal{F}^W_{kT/n})_{0\le k\le n}$. Denote by $\hat{\mathbb{Q}}$ the unique probability measure such that $\hat{\mathbb{Q}}\approx\mathbb{P}$ and $(X_t)_{t\in[0,T]}$ is a $\hat{\mathbb{Q}}$-martingale. From (2.2), it follows that for sufficiently large n, we have $\hat{\mathbb{Q}}\in\mathcal{Q}^n$ and so $\mathcal{Q}^n\neq\emptyset$.

We now represent the Radon–Nikodým derivative using Girsanov kernels. For any probability measure $\mathbb{Q}\approx\mathbb{P}$, let $\psi^{\mathbb{Q}}$ be an $(\mathcal{F}^W_t)$-progressively measurable process such that $\int_0^T|\psi_s^{\mathbb{Q}}|^2ds<\infty$ $\mathbb{P}$-a.s. and

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|\mathcal{F}_t^W = \exp\left(\int_0^t \psi_s^{\mathbb{Q}}dW_s - \frac{1}{2}\int_0^t|\psi_s^{\mathbb{Q}}|^2ds\right), \qquad t\in[0,T]. \tag{3.1}$$

We refer to $\psi^{\mathbb{Q}}$ as the *Girsanov kernel* associated with $\mathbb{Q}$. The dynamics of X can be written equivalently as

$$X_t = X_0 + \sigma W_t^{\mathbb{Q}} + \sigma\int_0^t \psi_s^{\mathbb{Q}}ds + \mu t, \qquad t\in[0,T],$$

where $W_t^{\mathbb{Q}} := W_t - \int_0^t \psi_s^{\mathbb{Q}}ds$, $t\in[0,T]$, is a Wiener process under $\mathbb{Q}$.

We arrive at the dual representation (2.6) for the certainty equivalent of $f^n(X)$. Although this dual representation is quite standard (under the appropriate growth conditions), since we could not find a direct reference, we provide a self-contained proof.

Proposition 3.1 *Let $n \in \mathbb{N}$ be large enough such that $\mathcal{Q}^n \neq \emptyset$ and fix an arbitrary $\lambda > 0$. Then*

$$c(n,\lambda) = \sup_{\mathbb{Q}\in\mathcal{Q}^n} \mathbb{E}_{\mathbb{Q}}\left[f^n(X) - \frac{1}{2\lambda}\int_0^T |\psi_s^{\mathbb{Q}}|^2 ds \right]. \tag{3.2}$$

Proof Since for any $(\gamma,\delta) \in \mathcal{A}^n$ and $\lambda > 0$, we have $V_{kT/n}^{\lambda\gamma,\lambda\delta} = \lambda V_{kT/n}^{\gamma,\delta}$ for all $k = 0, 1, \ldots, n$, we take $\lambda = 1$ without loss of generality. As usual, the proof rests on the classical Legendre–Fenchel duality inequality

$$xy \le e^x + y(\log y - 1), \quad x \in \mathbb{R}, y > 0, \qquad \text{with equality iff } y = e^x. \tag{3.3}$$

We start with proving the inequality "$\geq$" in (3.2).

Let $(\gamma,\delta) \in \mathcal{A}^n$ satisfy

$$\mathbb{E}_{\mathbb{P}}\Big[\exp\Big(-\big(V_T^{\gamma,\delta} - f^n(X)\big)\Big)\Big] < \infty.$$

Choose an arbitrary $\mathbb{Q} \in \mathcal{Q}^n$. From the Cauchy–Schwarz inequality and (2.5), it follows that $\mathbb{E}_{\mathbb{P}}[e^{-V_T^{\gamma,\delta}/2}] < \infty$. This together with (3.3) for $x = \max(0, -V_T^{\gamma,\delta}/2)$ and $y = \frac{d\mathbb{Q}}{d\mathbb{P}}$ gives that

$$\mathbb{E}_{\mathbb{Q}}[\max(0, -V_T^{\gamma,\delta})] = 2\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \max(0, -V_T^{\gamma,\delta}/2)\right] < \infty.$$

In view of (2.4), the portfolio value process $(V_{kT/n}^{\gamma,\delta})_{0\le k\le n}$ is a local $\mathbb{Q}$-supermartingale, and therefore combining Föllmer and Schied [9, Proposition 9.6] with the inequality $\mathbb{E}_{\mathbb{Q}}[\max(0, -V_T^{\gamma,\delta})] < \infty$ yields $\mathbb{E}_{\mathbb{Q}}[V_T^{\gamma,\delta}] \le 0$.

To conclude our claim, we can use (3.3) again to show that for any $z > 0$, we have the estimates

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\Big[\exp\Big(-\big(V_T^{\gamma,\delta} - f^n(X)\big)\Big)\Big] \\
&\ge \mathbb{E}_{\mathbb{P}}\left[\left(e^{-V_T^{\gamma,\delta}} + V_T^{\gamma,\delta} z e^{-f^n(X)}\frac{d\mathbb{Q}}{d\mathbb{P}}\right)e^{f^n(X)}\right] \\
&\ge \mathbb{E}_{\mathbb{P}}\left[\left(-ze^{-f^n(X)}\frac{d\mathbb{Q}}{d\mathbb{P}}\Big(\log\Big(ze^{-f^n(X)}\frac{d\mathbb{Q}}{d\mathbb{P}}\Big) - 1\Big)\right)e^{f^n(X)}\right] \\
&= -z(\log z - 1) + z\mathbb{E}_{\mathbb{Q}}\left[f^n(X) - \log\frac{d\mathbb{Q}}{d\mathbb{P}}\right].
\end{aligned}$$

Indeed, the first inequality follows from $\mathbb{E}_{\mathbb{Q}}[V_T^{\gamma,\delta}] \le 0$ and the second by (3.3) with the choice $x = -V_T^{\gamma,\delta}$ and $y = ze^{-f^n(X)}d\mathbb{Q}/d\mathbb{P}$. Finally, the supremum over z in the last line of the last display above is equal to $\exp(\mathbb{E}_{\mathbb{Q}}[f^n(X) - \log(d\mathbb{Q}/d\mathbb{P})])$. Together with (3.1), this implies the inequality "$\geq$" in (3.2).

Next, we prove the converse inequality "$\leq$" in (3.2). Define the probability measure $\hat{\mathbb{P}}$ by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{f^n(X)}}{\mathbb{E}_{\mathbb{P}}[e^{f^n(X)}]}.$$

Without loss of generality, we can assume that the right-hand side of (3.2) is finite. Thus for any $\mathbb{Q} \in \mathcal{Q}^n$,

$$\mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\hat{\mathbb{P}}}\right] = \mathbb{E}_{\mathbb{Q}}\left[\log \frac{d\mathbb{Q}}{d\mathbb{P}} - f^n(X)\right] + \log \mathbb{E}_{\mathbb{P}}[e^{f^n(X)}].$$

Next, we show that the supremum in (3.2) is attained.

From the well-known Komlós argument, see e.g. Delbaen and Schachermayer [6, Lemma A1.1], we obtain a maximising sequence $(\mathbb{Q}_m)_{m\in\mathbb{N}} \subseteq \mathcal{Q}^n$ for which $Z_m := d\mathbb{Q}_m/d\hat{\mathbb{P}}$ converges almost surely as $m \to \infty$. Without loss of generality, $(H(\mathbb{Q}_m|\hat{\mathbb{P}}))_{m\in\mathbb{N}}$ can be assumed to be bounded, where

$$H(\mathbb{Q}_m|\hat{\mathbb{P}}) := \mathbb{E}_{\hat{\mathbb{P}}}[Z_m \log Z_m] = \mathbb{E}_{\mathbb{Q}_m}\left[\log \frac{d\mathbb{Q}_m}{d\hat{\mathbb{P}}}\right].$$

Observe that the function $y \mapsto \frac{1}{e} + y \log y$ is nonnegative. Hence,

$$\lim_{M\to\infty} \sup_{m\in\mathbb{N}} \mathbb{E}_{\hat{\mathbb{P}}}[Z_m \mathbb{1}_{\{Z_m > M\}}] \leq \lim_{M\to\infty} \frac{1}{\log M} \sup_{m\in\mathbb{N}} \left(\frac{1}{e} + H(\mathbb{Q}_m|\hat{\mathbb{P}})\right) = 0,$$

where $\mathbb{1}_A$ denotes the indicator function of the set A. Thus the sequence $(Z_m)_{m\in\mathbb{N}}$ is under $\hat{\mathbb{P}}$ uniformly integrable, and so the convergence also holds in $L^1(\hat{\mathbb{P}})$. We conclude that $Z_0 := \lim_{m\to\infty} Z_m$ yields the density (Radon–Nikodým derivative with respect to $\hat{\mathbb{P}}$) of a probability measure $\mathbb{Q}_0$. From Fatou's lemma (the function $y \mapsto y \log y$ is bounded from below), we get $H(\mathbb{Q}_0|\hat{\mathbb{P}}) \leq \liminf_{m\to\infty} H(\mathbb{Q}_m|\hat{\mathbb{P}})$ and

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_0}\left[|X_{(k+1)T/n} - X_{kT/n}|^3 \middle| \mathcal{F}_{kT/n}\right] \\
&\leq \liminf_{m\to\infty} \mathbb{E}_{\mathbb{Q}_m}\left[|X_{(k+1)T/n} - X_{kT/n}|^3 \middle| \mathcal{F}_{kT/n}\right] \leq h(n), \qquad \forall k \leq n.
\end{aligned}$$

So the attainment of the supremum in (3.2) is proved if we can argue that $\mathbb{Q}_0$ is a martingale measure for $(X_{kT/n})_{0\leq k\leq n}$. To this end, it is sufficient to argue that for any k, the measures $(\mathbb{Q}_m \circ (X_{kT/n})^{-1})_{m\in\mathbb{N}}$ are uniformly integrable in the sense that

$$\liminf_{M\to\infty} \sup_{m\in\mathbb{N}} \mathbb{E}_{\mathbb{Q}_m}\left[|X_{kT/n}| \mathbb{1}_{\{|X_{kT/n}|>M\}}\right] = 0. \tag{3.4}$$

From the symmetry of Brownian motion, the simple inequality

$$\exp\Big(\sup_{0\leq t\leq T} |z_t|^p\Big) \leq \exp\Big(\Big|\sup_{0\leq t\leq T} z_t\Big|^p\Big) + \exp\Big(\Big|\inf_{0\leq t\leq T} z_t\Big|^p\Big), \qquad \forall z \in \mathcal{C}[0,T],$$

and the fact that $\sup_{0\le t\le T} W_t \overset{(d)}{=} |W_T|$ under $\mathbb{P}$, we obtain

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\Big[\exp\Big(\sup_{0\le t\le T}|X_t|^p\Big)\Big] \\
&\le 2\mathbb{E}_{\mathbb{P}}\left[\exp\left((X_0+|\mu|T+\sigma|W_T|)^p\right)\right] \\
&= 4\int_0^\infty \exp\left((X_0+|\mu|T+\sigma x)^p\right)\frac{\exp(-x^2/2)}{\sqrt{2\pi}}dx<\infty, \qquad \forall p\in(0,2). \quad (3.5)
\end{aligned}$$

Applying Hölder's inequality, (2.5) and (3.5) yields $\mathbb{E}_{\hat{\mathbb{P}}}[\exp(|X_{kT/n}|^{3/2})]<\infty$ for any k. Hence, from (3.3), we obtain

$$\begin{aligned}
\sup_{m\in\mathbb{N}}\mathbb{E}_{\mathbb{Q}_m}[|X_{kT/n}|^{3/2}] &= \sup_{m\in\mathbb{N}}\mathbb{E}_{\hat{\mathbb{P}}}[Z_m|X_{kT/n}|^{3/2}] \\
&\le \mathbb{E}_{\hat{\mathbb{P}}}[\exp(|X_{kT/n}|^{3/2})]+\sup_{m\in\mathbb{N}}H(\mathbb{Q}_m|\hat{\mathbb{P}})<\infty,
\end{aligned}$$

and (3.4) follows.

Now we arrive at the final step of the proof. We follow the approach in Frittelli [10]. Indeed, the perturbation argument for the proof of [10, Theorem 2.3] shows that the entropy minimising measure's density is of the form

$$Z_0=\frac{e^{-\xi_0}}{\mathbb{E}_{\hat{\mathbb{P}}}[e^{-\xi_0}]}$$

for some random variable ξ_0 with $\mathbb{E}_{\mathbb{Q}_0}[\xi_0]=0$ and $\mathbb{E}_{\mathbb{Q}}[\xi_0]\le 0$ for any $\mathbb{Q}\in\mathcal{Q}^n$. The separation argument for [10, Theorem 2.4] shows that ξ_0 is contained in the $L^1(\mathbb{Q})$-closure of $\{V_T^{\gamma,\delta}:(\gamma,\delta)\in\mathcal{A}^n\}-L^\infty_+$ for any $\mathbb{Q}\in\mathcal{Q}^n$, where L^∞_+ is the set of all nonnegative random variables which are uniformly bounded. Since $\mathcal{Q}^n\ne\emptyset$, Napp [23, Lemma 3.1] yields that ξ_0 must be of the same form $\xi_0=V_T^{\gamma_0,\delta_0}-R_0$ for some $(\gamma_0,\delta_0)\in\mathcal{A}^n$ and some random variable $R_0\ge 0$. As a result, we may bound the left-hand side in (3.2) via

$$\begin{aligned}
&\log\Big(\inf_{(\gamma,\delta)\in\mathcal{A}^n}\mathbb{E}_{\mathbb{P}}\Big[\exp\Big(-\big(V_T^{\gamma,\delta}-f^n(X)\big)\Big)\Big]\Big) \\
&\le \log\mathbb{E}_{\mathbb{P}}[e^{f^n(X)}]+\log\mathbb{E}_{\hat{\mathbb{P}}}[\exp(-V_T^{\gamma_0,\delta_0})] \\
&\le \log\mathbb{E}_{\mathbb{P}}[e^{f^n(X)}]+\log\mathbb{E}_{\hat{\mathbb{P}}}[\exp(-\xi_0)] \\
&= \log\mathbb{E}_{\mathbb{P}}[e^{f^n(X)}]+\mathbb{E}_{\mathbb{Q}_0}\big[\xi_0+\log\mathbb{E}_{\hat{\mathbb{P}}}[\exp(-\xi_0)]\big] \\
&= \log\mathbb{E}_{\mathbb{P}}[e^{f^n(X)}]-\mathbb{E}_{\mathbb{Q}_0}[\log Z_0] \\
&= \sup_{\mathbb{Q}\in\mathcal{Q}^n}\mathbb{E}_{\mathbb{Q}}\Big[f^n(X)-\frac{1}{\lambda}\log\frac{d\mathbb{Q}}{d\mathbb{P}}\Big].
\end{aligned}$$

Above, the first inequality follows by the definition of $\hat{\mathbb{P}}$, the second uses $R_0\ge 0$, the first equality is derived from $\mathbb{E}_{\mathbb{Q}_0}[\xi_0]=0$, the second follows by the definition of ξ_0, and finally, the last equality is deduced by our choice of $\mathbb{Q}_0$ as a measure that attains the supremum over $\mathcal{Q}^n$.

Together with (3.1), we obtain the inequality "$\leq$" in (3.2). □

The main technical challenge in this paper is showing that the scaling limit of the value function of the *drift control problem* from the right-hand side of (3.2) equals the value of the *volatility control problem* from the right-hand side of (2.8). This is summarised in the next theorem, whose proof is given in the next section.

Theorem 3.1 *We have the scaling limit*

$$\lim_{n\to\infty}\sup_{\mathbb{Q}\in\mathcal{Q}^n}\mathbb{E}_{\mathbb{Q}}\left[f^n(X)-\frac{1}{2\lambda}\int_0^T|\psi_s^{\mathbb{Q}}|^2ds\right]$$
$$=\sup_{\nu\in\mathcal{V}}\mathbb{E}_{\mathbb{P}}\left[f(X^{(\nu)})-\frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right]. \tag{3.6}$$

We end this section by proving Theorem 2.1.

Proof of Theorem 2.1 Proposition 3.1 and Theorem 3.1 imply that

$$\lim_{n\to\infty}c(n,n\ell)=\sup_{\nu\in\mathcal{V}}\mathbb{E}_{\mathbb{P}}\left[f(X^{(\nu)})-\frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right].$$

Utilising (2.7), the proof that $\lim_{n\to\infty}c(n,n\ell)=\lim_{n\to\infty}\pi(n,n\ell)$ follows once we show that

$$\lim_{n\to\infty}\frac{1}{n\ell}\log\left(\inf_{(\gamma,\delta)\in\mathcal{A}^n}\mathbb{E}_{\mathbb{P}}[\exp(-n\ell V_T^{\gamma,\delta})]\right)=0. \tag{3.7}$$

We prove this in two steps. First, notice that we can use $(\gamma,\delta)\equiv(0,0)$ to get an upper bound via

$$\inf_{(\gamma,\delta)\in\mathcal{A}^n}\mathbb{E}_{\mathbb{P}}[\exp(-n\ell V_T^{\gamma,\delta})]\leq\mathbb{E}_{\mathbb{P}}[\exp(-n\ell V_T^{0,0})]=1. \tag{3.8}$$

This establishes the relation "$\leq$" in (3.7), replacing "$=$". Next, we show that (3.7) holds with "$\geq$" instead of "$=$", which together with the last statement finishes the proof.

Fix $\ell>0$. We assume that n is sufficiently large such that $\hat{\mathbb{Q}}\in\mathcal{Q}^n$, where we recall that $\hat{\mathbb{Q}}$ is the unique martingale measure for the continuous-time Bachelier model. Let $(\gamma,\delta)\in\mathcal{A}^n$ be such that $\mathbb{E}_{\mathbb{P}}[\exp(-n\ell V_T^{\gamma,\delta})]<\infty$. This can be assumed without loss of generality due to (3.8). Using the same arguments as in the proof of Proposition 3.1, we get $\mathbb{E}_{\hat{\mathbb{Q}}}[V_T^{\gamma,\delta}]\leq 0$. Thus from Jensen's inequality for the convex function $y\mapsto e^{-n\ell y/2}$ and the Cauchy–Schwarz inequality, we obtain

$$1\leq\mathbb{E}_{\hat{\mathbb{Q}}}[\exp(-n\ell V_T^{\gamma,\delta}/2)]\leq\left(\mathbb{E}_{\mathbb{P}}[\exp(-n\ell V_T^{\gamma,\delta})]\right)^{1/2}\left(\mathbb{E}_{\mathbb{P}}\left[\left(\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right)^2\right]\right)^{1/2},$$

and so $\mathbb{E}_{\mathbb{P}}[\exp(-n\ell V_T^{\gamma,\delta})]$ is uniformly bounded from below. Thus we obtain that (3.7) holds with "≥" instead of equality. □

4 Proof of Theorem 3.1

The proof relies on two bounds, which are provided in two separate subsections.

4.1 Upper bound

This section is devoted to the proof of the inequality "≤" in (3.6). We start with the following lemma giving an intuition for the penalty term $g(y) := y - \log y - 1$, $y > 0$.

Lemma 4.1 *Consider a filtered probability space* $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_t)_{t\in[0,T]}, \bar{\mathbb{P}})$ *which supports an* $(\bar{\mathcal{F}}_t)$*-Wiener process* $\bar{W}$. *As usual, we assume that the filtration* $(\bar{\mathcal{F}}_t)$ *is right-continuous and contains the nullsets. Then for any times* $0 \le t_1 < t_2 \le T$ *and every* $(\bar{\mathcal{F}}_t)$*-progressively measurable process* ψ *satisfying* $\int_{t_1}^{t_2} \psi_u^2 du < \infty$ *and* $\bar{\mathbb{E}}[\int_{t_1}^{t_2} \psi_u du | \mathcal{F}_{t_1}] = 0$, *one has*

$$\bar{\mathbb{E}}\left[\int_{t_1}^{t_2} \psi_u^2 du \middle| \mathcal{F}_{t_1}\right] \ge g\left(\frac{1}{t_2 - t_1}\bar{\mathbb{E}}\left[\left(\bar{W}_{t_2} - \bar{W}_{t_1} + \int_{t_1}^{t_2} \psi_u du\right)^2 \middle| \mathcal{F}_{t_1}\right]\right). \tag{4.1}$$

Moreover, define the parametrised (by β*) processes* $(\theta_t^\beta)_{t\in[t_1,t_2]}$ *and* $(\vartheta_t^\beta)_{t\in[t_1,t_2]}$ *by*

$$\theta_t^\beta := \int_{t_1}^{t} \big(\beta(t_2 - t_1) - (t_2 - s)\big)^{-1} d\bar{W}_s, \qquad \beta > 1, \tag{4.2}$$

$$\vartheta_t^\beta := -\int_{t_1}^{t} \big(\beta(t_2 - t_1) + (t_2 - s)\big)^{-1} d\bar{W}_s, \qquad \beta > 0. \tag{4.3}$$

Then (4.1) holds with equality for $(\psi_t)_{t\in[t_1,t_2]}$, *given by*

$$\psi_t = \theta_t^{\bar{\beta}} \mathbb{1}_{\{\mathbb{E}[(\bar{W}_{t_2} - \bar{W}_{t_1} + \int_{t_1}^{t_2} \psi_u du)^2 | \mathcal{F}_{t_1}] > t_2 - t_1\}} + \vartheta_t^{\bar{\beta}} \mathbb{1}_{\{\mathbb{E}[(\bar{W}_{t_2} - \bar{W}_{t_1} + \int_{t_1}^{t_2} \psi_u du)^2 | \mathcal{F}_{t_1}] < t_2 - t_1\}}, \tag{4.4}$$

where

$$\bar{\beta} = \frac{\mathbb{E}[(\bar{W}_{t_2} - \bar{W}_{t_1} + \int_{t_1}^{t_2} \psi_u du)^2 | \mathcal{F}_{t_1}]}{|\mathbb{E}[(\bar{W}_{t_2} - \bar{W}_{t_1} + \int_{t_1}^{t_2} \psi_u du)^2 | \mathcal{F}_{t_1}] - (t_2 - t_1)|}. \tag{4.5}$$

Proof By the independent increments and the scaling property of the Wiener process, we assume without loss of generality that $t_1 = 0$, $t_2 = 1$ and $\mathcal{F}_0$ is the trivial σ-algebra.

Obviously, (4.1) holds trivially if $\bar{\mathbb{E}}[\int_0^1 \psi_t^2 dt] = \infty$. Hence for the rest of the proof, we also assume that $\bar{\mathbb{E}}[\int_0^1 \psi_t^2 dt] < \infty$. This in turn implies that $\bar{\mathbb{E}}[\psi_t^2] < \infty$ for almost every $t \in [0, 1]$ with respect to Lebesgue measure. Thus without loss of generality, we may assume that $\bar{\mathbb{E}}[\psi_t] = 0$ for any $t \in [0, 1]$. Indeed, set $\bar{\psi}_t := \psi_t - \bar{\mathbb{E}}[\psi_t]$, $t \in [0, 1]$. Then if we prove (4.1) for $\bar{\psi}$, then

$$\begin{aligned}\bar{\mathbb{E}}\left[\int_0^1 \psi_t^2 dt\right] &\geq \bar{\mathbb{E}}\left[\int_0^1 \bar{\psi}_u^2 dt\right]\\ &\geq g\left(\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \bar{\psi}_t dt\right)^2\right]\right)\\ &= g\left(\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \psi_t dt\right)^2\right]\right),\end{aligned}$$

where the second inequality follows from (4.1) applied to $\bar{\psi}$ and the equality follows since $\int_0^1 \bar{\mathbb{E}}[\psi_t] dt = 0$.

Next, by applying standard density arguments in $L^2(dt \otimes \bar{\mathbb{P}})$, we can assume that ψ is a simple process (see Karatzas and Shreve [22, Sect. 3.2]) in the sense that ψ is bounded and there exists a deterministic partition $0 = t_0 < t_1 < \cdots < t_m = 1$ such that ψ is a (random) constant on each interval $(t_i, t_{i+1}]$. Hence for the rest of the proof, we fix a simple process ψ which satisfies $\bar{\mathbb{E}}[\psi_t] = 0$ for all $t \in [0, 1]$.

We split the proof into two cases, namely (I) $\bar{\mathbb{E}}[(\bar{W}_1 + \int_0^1 \psi_t dt)^2] > 1$ and (II) $\bar{\mathbb{E}}[(\bar{W}_1 + \int_0^1 \psi_t dt)^2] < 1$. When the expected value equals 1, (4.1) follows immediately since $g(1) = 0$.

Case I: $\bar{\mathbb{E}}[(\bar{W}_1 + \int_0^1 \psi_t dt)^2] > 1$. Let $\bar{\beta}$ be given by (4.5). Observe that

$$\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \psi_t dt\right)^2\right] = \frac{\bar{\beta}}{\bar{\beta} - 1}.$$

In order to prove the inequality (4.1), it is sufficient to show that

$$\begin{aligned}\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \psi_t dt\right)^2 - \bar{\beta}\int_0^1 \psi_t^2 dt\right] &\leq \frac{\bar{\beta}}{\bar{\beta} - 1} - \bar{\beta} g\left(\frac{\bar{\beta}}{\bar{\beta} - 1}\right)\\ &= \bar{\beta} \log \frac{\bar{\beta}}{\bar{\beta} - 1}.\end{aligned} \tag{4.6}$$

Denote by $(\mathcal{F}_t^{\bar{W}})_{t \in [0,1]}$ the augmented filtration generated by the Brownian motion $(\bar{W}_t)_{t \in [0,1]}$ and let $(u_t)_{t \in [0,1]}$ be the optional projection of ψ on $(\mathcal{F}_t^{\bar{W}})_{t \in [0,1]}$ (this exists since ψ is bounded). Set $v_t := \psi_t - u_t$, $t \in [0, T]$. Clearly, $\bar{W}_{[t,1]} - \bar{W}_t$ is independent of ψ_t and $\bar{W}_{[0,t]}$, where $\bar{W}_{[a,b]}$ is the restriction of $\bar{W}$ to the interval $[a, b]$. This together with the fact that $\mathcal{F}_1^{\bar{W}}$ is generated by $\mathcal{F}_t^{\bar{W}}$ and $\bar{W}_{[t,1]} - \bar{W}_t$ yields $u_t = \bar{\mathbb{E}}[\psi_t | \mathcal{F}_t^{\bar{W}}] = \bar{\mathbb{E}}[\psi_t | \mathcal{F}_1^{\bar{W}}]$ for all t. Thus

$$\begin{aligned}&\bar{\mathbb{E}}\left[\left(\bar{W}_1+\int_0^1\psi_t dt\right)^2-\bar{\beta}\int_0^1\psi_t^2dt\right]\\&=\bar{\mathbb{E}}\left[\left(\bar{W}_1+\int_0^1 u_t dt\right)^2-\bar{\beta}\int_0^1 u_t^2dt\right]+\bar{\mathbb{E}}\left[\left(\int_0^1 v_t dt\right)^2-\bar{\beta}\int_0^1 v_t^2dt\right]\\&\le\bar{\mathbb{E}}\left[\left(\bar{W}_1+\int_0^1 u_t dt\right)^2-\bar{\beta}\int_0^1 u_t^2dt\right],\end{aligned}\tag{4.7}$$

where the inequality follows from Jensen's inequality and the fact that $\bar{\beta}>1$.

From the martingale representation theorem and the fact that ψ is simple, it follows that there exists a (jointly) measurable map $\kappa:[0,1]^2\times\Omega\to\mathbb{R}$ such that $\kappa_{t,s}$ is $\mathcal{F}^{\bar{W}}_{t\wedge s}$-measurable for all $t,s\in[0,1]$ and

$$u_t=\int_0^t\kappa_{t,s}d\bar{W}_s\qquad(dt\otimes\bar{\mathbb{P}})\text{-a.e.}$$

Here, for each $t\in[0,T]$, we are applying the martingale representation theorem for the random variable u_t. Utilising the fact that ψ is piecewise constant, we obtain that κ is jointly measurable in s and t. This is essential in the sequel when we apply Fubini's theorem.

Define the processes

$$\zeta_s:=\int_s^1\kappa_{t,s}dt,\quad \eta_s:=\int_s^1\kappa_{t,s}^2dt,\qquad s\in[0,1].$$

We get

$$\begin{aligned}\bar{\mathbb{E}}\left[\left(\bar{W}_1+\int_0^1 u_t dt\right)^2-\bar{\beta}\int_0^1 u_t^2dt\right]&=\bar{\mathbb{E}}\left[\int_0^1\left((1+\zeta_s)^2-\bar{\beta}\eta_s\right)ds\right]\\&\le\bar{\mathbb{E}}\left[\int_0^1\left((1+\zeta_s)^2-\frac{\bar{\beta}\zeta_s^2}{1-s}\right)ds\right]\\&\le\int_0^1\frac{\bar{\beta}}{\bar{\beta}-(1-s)}ds=\bar{\beta}\log\frac{\bar{\beta}}{\bar{\beta}-1}.\end{aligned}$$

Indeed, the first equality follows from the stochastic Fubini theorem (see Revuz and Yor [24, Sect. IV.5]) and the Itô isometry, the first inequality follows from the Cauchy–Schwarz inequality, the second follows from maximising the quadratic function (for a given s) $z\mapsto(1+z)^2-\frac{\bar{\beta}z^2}{1-s}$ and the last equality is a simple computation. This together with (4.7) completes the proof of (4.6).

Next recall the process θ^{β} given by (4.2). Observe that for $\kappa_{t,s}:=\theta_s^{\bar{\beta}}$ with $t,s\in[0,1]$, the above two inequalities are in fact equalities. Moreover, it easy to check that

$$\bar{\mathbb{E}}\left[\left(\bar{W}_1+\int_0^1 u_t\right)^2\right]=\int_0^1\left(1+(1-s)\theta_s^{\bar{\beta}}\right)^2ds=\frac{\bar{\beta}}{\bar{\beta}-1},$$

and so for $\psi=\theta^{\bar{\beta}}$, we have equality in (4.1).

Case II: $\bar{\mathbb{E}}[(\bar{W}_1 + \int_0^1 \psi_t dt)^2] < 1$. Let $\bar{\beta}$ be given by (4.5). Observe that

$$\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \psi_t dt\right)^2\right] = \frac{\bar{\beta}}{\bar{\beta}+1}.$$

In order to prove the inequality (4.1), it sufficient to show that

$$\begin{aligned}\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \psi_t dt\right)^2 + \bar{\beta}\int_0^1 \psi_t^2 dt\right] &\geq \frac{\bar{\beta}}{\bar{\beta}+1} + \bar{\beta} g\left(\frac{\bar{\beta}}{\bar{\beta}+1}\right) \\ &= \bar{\beta}\log\frac{\bar{\beta}+1}{\bar{\beta}}. \end{aligned} \tag{4.8}$$

Let u, v, ζ, η be defined as in Case I. Recall the process ϑ^{β} given by (4.3). Then by using similar arguments as in Case I, we obtain

$$\begin{aligned}\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 \psi_t dt\right)^2 + \bar{\beta}\int_0^1 \psi_t^2 dt\right] &\geq \bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 u_t dt\right)^2 + \bar{\beta}\int_0^1 u_t^2 dt\right] \\ &= \bar{\mathbb{E}}\left[\int_0^1 \big((1+\zeta_s)^2 + \bar{\beta}\eta_s\big) ds\right] \\ &\geq \bar{\mathbb{E}}\left[\int_0^1 \big((1+\zeta_s)^2 + \bar{\beta}\zeta_s^2/(1-s)\big) ds\right] \\ &\geq \int_0^1 \frac{\bar{\beta}}{\bar{\beta}+(1-s)} ds = \bar{\beta}\log\frac{\bar{\beta}+1}{\bar{\beta}},\end{aligned}$$

and (4.8) follows. Finally, we notice that for $\kappa_{t,s} := \vartheta_s^{\bar{\beta}}$ with $t, s \in [0, 1]$, the above inequalities are in fact equalities. In addition, it is easy to check that

$$\bar{\mathbb{E}}\left[\left(\bar{W}_1 + \int_0^1 u_t dt\right)^2\right] = \int_0^1 \big(1 + (1-s)\vartheta_s^{\bar{\beta}}\big)^2 ds = \frac{\bar{\beta}}{\bar{\beta}+1},$$

and so for $\psi = \vartheta^{\bar{\beta}}$, we have equality in (4.1). □

Next, fix $n \in \mathbb{N}$. The next lemma provides a bound for an expected payoff calculated with respect to a given discrete-time martingale M, which later on will stand for $(X_{kT/n})_{0\leq k\leq n}$. The idea is to construct a continuous-time martingale that is close in distribution to the process M on the discrete set of times and whose volatility is piecewise constant between two consecutive points on the discrete-time set.

Lemma 4.2 *Let $(M_k)_{0\leq k\leq n}$ be a martingale, defined on some probability space, with $M_0 = X_0$ and satisfying for any $k = 0, \dots, n-1$ that*

$$\hat{\mathbb{E}}\big[|M_{k+1} - M_k|^3 \big| M_0, \dots, M_k\big] \leq h(n), \tag{4.9}$$

where $\hat{\mathbb{E}}$ *is the expectation with respect to the given probability space. Assume that for some* $K > 0$,

$$\hat{\mathbb{E}}\left[f^n(M) - \frac{1}{2n\ell}\sum_{k=0}^{n-1} g\left(\frac{n}{\sigma^2 T}\hat{\mathbb{E}}\left[|M_{k+1} - M_k|^2 \big| M_0, \ldots, M_k\right]\right)\right] \geq -K, \quad (4.10)$$

where $f^n(M) = f(p^n(M))$ *and (with abuse of notation)* $p^n(M)$ *is the linear interpolation of* $((kT/n, M_k) : k = 0, \ldots, n)$ *so that* $p^n(M)$ *is a random element in* $\mathcal{C}[0, T]$. *Then there exists a constant* $C > 0$ *(that depends only on* ℓ, K *and* f, *through the Lipschitz property and the linear growth), which is independent of* n, *such that*

$$\hat{\mathbb{E}}\left[f^n(M) - \frac{1}{2n\ell}\sum_{k=0}^{n-1} g\left(\frac{n}{\sigma^2 T}\hat{\mathbb{E}}\left[|M_{k+1} - M_k|^2 \big| M_0, \ldots, M_k\right]\right)\right]$$
$$\leq C\big(h(n)n\big)^{1/8} + \sup_{\nu\in\mathcal{V}} \mathbb{E}_{\mathbb{P}}\left[f^n(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right]. \quad (4.11)$$

Proof The Lipschitz-continuity of f implies that it has linear growth. This together with the Doob inequality for the martingale M, the simple bound $g(y) \geq y/2 - 1$ and (4.10) gives that there exists a constant $\hat{C} > 0$ (that depends only on ℓ, K and f, through the Lipschitz property and the linear growth), which is independent of n, such that

$$\hat{\mathbb{E}}\left[\max_{0\leq k\leq n} M_k^2\right] \leq \hat{C}. \quad (4.12)$$

From Dolinsky [7, Lemma 3.2] and (4.9), it follows that we can construct the martingale $(M_k)_{0\leq k\leq n}$ on a new probability space (meaning that the joint distribution of $(M_0, \ldots, M_n)$ is the same as before) which supports a sequence of identically distributed random variables $(Y_k)_{1\leq k\leq n}$ having the standard normal distribution and such that the following hold:

(I) For each $k = 0, 1, \ldots, n-1$, Y_{k+1} is independent of $(M_i)_{0\leq i\leq k}$ and $(Y_i)_{1\leq i\leq k}$.

(II) There exists a universal constant $\bar{C} > 0$, which is independent of the parameters in the model, for which

$$\hat{\mathbb{P}}\left[\max_{k=0,\ldots,n} |M_k - \hat{X}_k| > \big(h(n)n\big)^{1/4}\right] < \bar{C}\big(h(n)n\big)^{1/4}, \quad (4.13)$$

with

$$\hat{X}_k := X_0 + \sum_{i=0}^{k-1} Y_{i+1}\sqrt{\hat{\mathbb{E}}[(M_{i+1} - M_i)^2 | M_0, \ldots, M_i]}, \qquad k = 0, \ldots, n.$$

We abuse notation here and use $\hat{\mathbb{P}}$ both for the probability measure on the original space and the new space and keep using the notation M for the martingale on the new probability space.

Let us remark that in the formulation of [7, Lemma 3.2], we have that $(Y_k)_{1\leq k\leq n}$ are independent and identically distributed random variables with the standard normal distribution such that for any k, Y_{k+1} is independent of $(M_i)_{0\le i\le k}$. However, the construction in the proof of [7, Lemma 3.2] provides a stronger property which is property (I) above.

Next, we use this representation to embed the law of $(\hat{X}_k)_{0\le k\le n}$ into the original Brownian probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^W)_{t\in[0,T]}, \mathbb{P})$ from Sect. 2. By a classical result of Skorohod [25, Theorem 1], we obtain that there exist measurable functions $\chi_k : \mathbb{R}^{2k+1} \to \mathbb{R}$, $k = 1, \dots, n$, such that for any $k = 0, 1, \dots, n-1$,

$$\begin{aligned}&\mathcal{L}(Y_1, \dots, Y_{k+1}, M_0, \dots, M_k, M_{k+1})\\&= \mathcal{L}\big(Y_1, \dots, Y_{k+1}, M_0, \dots, M_k, \chi_{k+1}(Y_1, \dots, Y_{k+1}, M_0, \dots, M_k, \xi)\big),\end{aligned}$$

where ξ has the standard normal distribution and is independent of $(Y_i)_{1\le i\le k+1}$ and $(M_i)_{0\le i\le k}$.

Define on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^W)_{t\in[0,T]}, \mathbb{P})$ the processes $(\bar{Y}_i)_{1\le i\le n}$, $(\bar{\xi}_i)_{1\le i\le n}$ and $(\bar{M}_i)_{0\le i\le n}$ as follows. First, for any $k = 1, \dots, n$, set

$$\begin{aligned}\bar{Y}_k &:= \sqrt{n/T}\big(W_{kT/n} - W_{(k-1)T/n}\big),\\ \bar{\xi}_k &:= \frac{3W_{(k-1)T/n+T/(3n)} - 2W_{(k-1)T/n+T/(2n)} - W_{(k-1)T/n}}{\sqrt{\mathrm{Var}[3W_{(k-1)T/n+T/(3n)} - 2W_{(k-1)T/n+T/(2n)} - W_{(k-1)T/n}]}}.\end{aligned}$$

Next, define by recursion $\bar{M}_0 := X_0$ and for any $k = 0, 1, \dots, n-1$,

$$\bar{M}_{k+1} := \chi_{k+1}(\bar{Y}_1, \dots, \bar{Y}_{k+1}, \bar{M}_0, \dots, \bar{M}_k, \bar{\xi}_{k+1}).$$

Clearly, $(\bar{Y}_i)_{1\le i\le n}$ and $(\bar{\xi}_i)_{1\le i\le n}$ have the standard normal distribution. Observe that for any k, $\bar{Y}_{k+1}$ and $\bar{\xi}_{k+1}$ are independent of $W_{[0,kT/n]}$, and so they are independent of $(\bar{Y}_i)_{1\le i\le k}$ and $(\bar{M}_i)_{0\le i\le k}$. Moreover, we notice that for any k, $\bar{\xi}_k$ is independent of $\bar{Y}_k$ (they are bivariate normal and uncorrelated). Thus for any k, $\bar{\xi}_{k+1}$ is independent of $(\bar{Y}_i)_{1\le i\le k+1}$ and $(\bar{M}_i)_{0\le i\le k}$. From the definition of the functions χ_k, $k = 1, \dots, n$, we conclude (by induction) that

$$\mathcal{L}\big((\bar{M}_i)_{0\le i\le n}, (\bar{Y}_i)_{1\le i\le n}\big) = \mathcal{L}\big((M_i)_{0\le i\le n}, (Y_i)_{1\le i\le n}\big). \tag{4.14}$$

Next, let $\varphi_k : \mathbb{R}^{k+1} \to \mathbb{R}_+$, $k = 0, 1, \dots, n-1$, be measurable functions such that

$$\sqrt{\mathbb{E}[(M_{k+1} - M_k)^2 | M_0, \dots, M_k]} = \varphi_k(M_0, \dots, M_k), \qquad k = 0, 1, \dots, n-1.$$

Introduce the process $\nu \in \mathcal{V}$ by

$$\nu_t := \frac{n}{T}\sum_{k=0}^{n-1} \varphi_k^2(\bar{M}_0, \dots, \bar{M}_k)\mathbb{1}_{\{kT/n\le t<(k+1)T/n\}}, \qquad t \in [0, T].$$

From (4.14), it follows that the law of $(\nu_{kT/n} : k = 0, \dots, n-1)$ equals the law of $(\frac{n}{T}\mathbb{E}[(M_{k+1} - M_k)^2 | M_0, \dots, M_k] : k = 0, \dots, n-1)$. Since ν is constant on each of the intervals $[kT/n, (k+1)T/n]$, $k=0,\dots,n-1$, we conclude that

$$\hat{\mathbb{E}}\left[\frac{1}{n}\sum_{k=0}^{n-1} g\left(\frac{n}{\sigma^2 T}\hat{\mathbb{E}}[(M_{k+1}-M_k)^2|M_0,\dots,M_k]\right)\right] = \mathbb{E}_{\mathbb{P}}\left[\frac{1}{T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right]. \tag{4.15}$$

Finally, consider the process $X^{(\nu)}$. Observe that

$$X^{(\nu)}_{kT/n} = X_0 + \sum_{i=0}^{k-1} \bar{Y}_{i+1}\varphi_i(\bar{M}_0,\dots,\bar{M}_i), \qquad k=0,1,\dots,n.$$

This together with (4.14) yields that $(X^{(\nu)}_{kT/n})_{0\le k\le n}$ and $(\hat{X}_k)_{0\le k\le n}$ have the same distribution. Therefore, because $f\ge 0$ is Lipschitz-continuous, we obtain that there exists a constant c_1, which does not depend on n, such that

$$\begin{aligned}\hat{\mathbb{E}}[f^n(M)] &\le \mathbb{E}_{\mathbb{P}}[f^n(X^{(\nu)})] + c_1\big(h(n)n\big)^{1/4} \\ &\quad + \hat{\mathbb{E}}\big[f^n(M)\mathbb{1}_{\{\max_{k=0,\dots,n}|M_k-\hat{X}_k|>(h(n)n)^{1/4}\}}\big] \\ &\le \mathbb{E}_{\mathbb{P}}[f^n(X^{(\nu)})] + C\big(h(n)n\big)^{1/8}\end{aligned} \tag{4.16}$$

for some constant C which does not depend on n. The last inequality follows from the Cauchy–Schwarz inequality, the linear growth of f, the scaling assumption (2.3) and (4.12), (4.13).

By combining (4.15) and (4.16), we complete the proof of (4.11). □

We are now ready to prove the upper bound.

Proof of the inequality "≤" in (3.6) Fix $\ell>0$. By passing to a subsequence (which is still denoted by n), we assume without loss of generality that

$$\lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}^n}\left[f^n(X) - \frac{1}{2n\ell}\int_0^T |\psi_t^{\mathbb{Q}^n}|^2 dt\right] > -\infty \tag{4.17}$$

(otherwise the statement is obvious). Fix n and introduce the $\mathbb{Q}^n$-martingale M via $M_k := X_{kT/n}$, $k=0,1,\dots,n$, where we recall that

$$X_t = X_0 + \sigma W_t^{\mathbb{Q}^n} + \sigma\int_0^t \psi_s^{\mathbb{Q}^n} ds + \mu t, \qquad t\in[0,T],$$

and $W^{\mathbb{Q}^n}$ is a Wiener process under $\mathbb{Q}^n$. Observe that for any $0\le k\le n-1$, we have

$$\mathbb{E}_{\mathbb{Q}^n}\left[\int_{kT/n}^{(k+1)T/n}(\psi_t^{\mathbb{Q}^n}+\mu/\sigma)dt\,\middle|\,\mathcal{F}_{kT/n}\right] = 0.$$

This together with Lemma 4.1 and the scaling property of Brownian motion gives

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^n}\Big[\int_0^T |\psi_t^{\mathbb{Q}^n}|^2 dt\Big] \\
&= \mathbb{E}_{\mathbb{Q}^n}\Big[\sum_{k=0}^{n-1} \mathbb{E}_{\mathbb{Q}^n}\Big[\int_{kT/n}^{(k+1)T/n} |\psi_t^{\mathbb{Q}^n}|^2 dt \Big| \mathcal{F}_{kT/n}\Big]\Big] \\
&= \mathbb{E}_{\mathbb{Q}^n}\Big[\sum_{k=0}^{n-1} \mathbb{E}_{\mathbb{Q}^n}\Big[\int_{kT/n}^{(k+1)T/n} (\psi_t^{\mathbb{Q}^n} + \mu/\sigma)^2 dt \Big| \mathcal{F}_{kT/n}\Big]\Big] - \mu^2 T/\sigma^2 \\
&\geq \mathbb{E}_{\mathbb{Q}^n}\Big[\sum_{k=0}^{n-1} g\Big(\frac{n}{\sigma^2 T}\mathbb{E}_{\mathbb{Q}^n}[(M_{k+1}-M_k)^2|\mathcal{F}_{kT/n}]\Big)\Big] - \mu^2 T/\sigma^2 \\
&\geq \mathbb{E}_{\mathbb{Q}^n}\Big[\sum_{k=0}^{n-1} g\Big(\frac{n}{\sigma^2 T}\mathbb{E}_{\mathbb{Q}^n}[(M_{k+1}-M_k)^2|M_0,\ldots,M_k]\Big)\Big] - \mu^2 T/\sigma^2,
\end{aligned}$$

where the last inequality follows from Jensen's inequality for the convex function g.

Finally, from the assumption (4.17), it follows that we can apply Lemma 4.2 (i.e., (4.10) holds true for some constant K). We conclude that

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^n}\Big[f^n(X) - \frac{1}{2n\ell}\int_0^T |\psi_t^{\mathbb{Q}^n}|^2 dt\Big] \\
&\leq C\big(h(n)n\big)^{1/8} + \mu^2 T/(2\ell\sigma^2 n) + \sup_{\nu\in\mathcal{V}} \mathbb{E}_{\mathbb{P}}\Big[f^n(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\Big(\frac{\nu_t}{\sigma^2}\Big)dt\Big].
\end{aligned}$$

The proof is completed by using (2.3) and taking $n \to \infty$. □

4.2 Lower bound

This section is devoted to the proof of the inequality "≥" in (3.6).

Proof of the inequality "≥" in (3.6) Fix $\ell > 0$. The proof is done in three steps. In the first step, we construct a sequence of controls on the Brownian probability space which asymptotically achieve the supremum on the right-hand side of (3.6) and have a simple structure. In the second step, we use the processes θ^β, ϑ^β from Lemma 4.1 in order to construct a sequence of probability measures $\mathbb{Q}^n \in \mathcal{Q}^n$ together with their Girsanov kernels $\psi^{\mathbb{Q}^n}$. The construction of the measures $\mathbb{Q}^n$ is done in a semi-explicit manner. Namely, the processes θ^β and ϑ^β are constructed via the process $\bar{W}$, which in our case translates to $W^{\mathbb{Q}^n}$. Note that the measure $\mathbb{Q}^n$ is determined by the Girsanov kernel $\psi^{\mathbb{Q}^n}$. To achieve this construction, we use integration by parts and introduce a path-dependent SDE. As a by-product, our process $\psi^{\mathbb{Q}^n}$ is measurable with respect to the original filtration $(\mathcal{F}_t^W)$. Finally, we show convergence of the payoff components.

Step 1: For any $K > 0$ and $n \in \mathbb{N}$, let $\mathcal{V}_K^n \subseteq \mathcal{V}$ be the set of all volatility processes of the form

$$\nu_t = \sum_{k=0}^{n-1} \phi_k(W_0, W_{T/n}, \ldots, W_{kT/n}) \mathbb{1}_{t \in [kT/n,(k+1)T/n)}, \tag{4.18}$$

where $\phi_k : \mathbb{R}^{k+1} \to [1/K, K]$, $k = 0, 1, \ldots, n-1$, are continuous functions.

Set $\epsilon > 0$. In this step, we argue that there exist $K = K(\epsilon)$ and $N = N(\epsilon)$ such that for any $n > N$,

$$\begin{aligned}&\sup_{\nu \in \mathcal{V}} \mathbb{E}_{\mathbb{P}}\left[f(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right] \\ &< \epsilon + \sup_{\nu \in \mathcal{V}_K^n} \mathbb{E}_{\mathbb{P}}\left[f^n(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right].\end{aligned} \tag{4.19}$$

To this end, observe first that by standard density arguments, we get the same supremum on the right-hand side of (3.6) if instead of letting ν vary over all of $\mathcal{V}$ there, we confine it to be of the form

$$\nu_t := \sum_{j=0}^{J-1} \phi_j(W_{t_0}, \ldots, W_{t_j}) \mathbb{1}_{t \in [t_j, t_{j+1})}, \qquad t \in [0, T],$$

where $0 = t_0 < t_1 < \cdots < t_J = T$ is a finite deterministic partition of $[0, T]$ and each $\phi_j : \mathbb{R}^{j+1} \to \mathbb{R}_+$, $j = 0, \ldots, J-1$, is continuous, bounded and bounded away from zero. Let ν be of that form. There exists K such that ν has values in $[1/K, K]$ a.s. For any $n \in \mathbb{N}$, set

$$t_j^n := \min\left\{t \in \{0, T/n, 2T/n, \ldots, T\} : t \geq t_j\right\}, \qquad j = 0, 1, \ldots, J,$$

and define $\nu^n \in \mathcal{V}_K^n$ by

$$\nu_t^n := \sum_{j=0}^{J-1} \phi_j(W_{t_0^n}, \ldots, W_{t_j^n}) \mathbb{1}_{t \in [t_j^n, t_{j+1}^n)}, \qquad t \in [0, T].$$

Observe that $\nu_n \to \nu$ in $L^2(dt \otimes \mathbb{P})$. This together with the Lipschitz-continuity of f and the Lipschitz-continuity of g on the interval $[1/(K\sigma^2), K/\sigma^2]$ gives that

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}}\left[f(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right] \\ &= \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}\left[f(X^{(\nu^n)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t^n}{\sigma^2}\right)dt\right].\end{aligned}$$

Thus in order to establish (4.19), it remains to show that

$$\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}[f^n(X^{(\nu^n)}) - f(X^{(\nu^n)})] = 0.$$

Indeed, from the Lipschitz-continuity of f, the Burkholder–Davis–Gundy inequality and the fact that $\nu^n \le K$ for all n, it follows that there exist constants $\tilde{C}_1$, $\tilde{C}_2$ (independent of n) such that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\big[\big(f^n(X^{(\nu^n)}) - (X^{(\nu^n)})\big)^4\big] &\le \tilde{C}_1 \mathbb{E}_{\mathbb{P}}\Big[\max_{0\le k\le n-1} \sup_{kT/n\le t\le (k+1)T/n} \big(X_t^{(\nu^n)} - X_{kT/n}^{(\nu^n)}\big)^4\Big] \\ &\le \tilde{C}_1 \sum_{k=0}^{n-1} \mathbb{E}_{\mathbb{P}}\Big[\sup_{kT/n\le t\le (k+1)T/n} \big(X_t^{(\nu^n)} - X_{kT/n}^{(\nu^n)}\big)^4\Big] \\ &\le \tilde{C}_1 n \tilde{C}_2 n^{-2} \\ &= \tilde{C}_1 \tilde{C}_2 / n.\end{aligned}$$

This completes the proof of (4.19).

Step 2: Fix $K, A > 0$ and choose $n \in \mathbb{N}$. Following (4.4) and (4.5), we define the functions $\beta : [1/K, K] \to \mathbb{R}_+$ and $\theta : [1/K, K] \times [0, T/n] \to \mathbb{R}$ by

$$\begin{aligned}\beta(u) &:= \frac{u/\sigma^2}{|u/\sigma^2 - 1|} \mathbb{1}_{\{u\neq\sigma^2\}}, \\ \theta(u,t) &:= \big(\beta(u)T/n - (T/n - t)\big)^{-1} \mathbb{1}_{\{u>\sigma^2\}} \\ &\quad - \big(\beta(u)T/n + (T/n - t)\big)^{-1} \mathbb{1}_{\{u<\sigma^2\}}.\end{aligned} \tag{4.20}$$

Introduce the map $\Phi^A : [1/K, K] \times \mathcal{C}[0, T/n] \to \mathcal{C}[0, T/n]$ via

$$\Phi_t^A(u, z) := (-A) \vee \left(\theta(u,t)z_t - \theta(u,0)z_0 - \int_0^t z_s \frac{\partial\theta(u,s)}{\partial s} ds\right) \wedge A \tag{4.21}$$

for $t \in [0, T/n]$. Observe that $\Phi_t^A(u, z)$ depends only on $z_{[0,t]}$. For given $u \in [1/K, K]$ and $y \in \mathbb{R}$, consider the path-dependent SDE

$$dY_t = dW_t - \big(\Phi_t^A(u, Y) - \mu/\sigma\big)dt, \qquad t \in [0, T/n], Y_0 = y. \tag{4.22}$$

Let us notice that for a given u, $\Phi^A(u, \cdot) : \mathcal{C}[0, T/n] \to \mathcal{C}[0, T/n]$ is a Lipschitz-continuous function with respect to the sup-norm. Thus Revuz and Yor [24, Theorem IX.2.1] yield a unique strong solution for the above SDE. Obviously, (4.22) can be reformulated as

$$d(U_t, Y_t) = (0, dW_t) - (0, \Phi_t^A(U_t, Y) - \mu/\sigma)dt, \quad t \in [0, T/n], U_0 = u, Y_0 = y.$$

The last formulation allows us to apply Kallenberg [21, Theorem 1] and obtain the existence of a *jointly* measurable function

$$\Psi^A : [1/K, K] \times \mathbb{R} \times \mathcal{C}[0, T/n] \to \mathcal{C}[0, T/n]$$

such that for any $u > 0$ and $y \in \mathbb{R}$, $Y_{[0,T/n]} := \Psi^A(u, y, W_{[0,T/n]})$ is the unique strong solution to (4.22).

Next, let $\nu \in \mathcal{V}_K^n$ be given by (4.18). Define inductively the random variables $u_k^{A,n,\nu} = u_k$ and $Y_{[kT/n,(k+1)T/n]}^{A,n,\nu} = Y_{[kT/n,(k+1)T/n]}^A$, $k = 0, 1, \ldots, n-1$, as follows. Set $u_0 := \nu_0$, $Y_{[0,T/n]}^A := \Psi^A(u_0, 0, W_{[0,T/n]})$, and for $k = 1, \ldots, n-1$,

$$u_k := \phi_k(Y_0^A, \ldots, Y_{kT/n}^A),$$

$$Y_{[kT/n,(k+1)T/n]}^A := S^k\Big(\Psi^A\big(u_k, Y_{kT/n}^A, (W_{t+kT/n} - W_{kT/n})_{t\in[0,T/n]}\big)\Big),$$

where $S_k : \mathcal{C}[0,T/n] \to \mathcal{C}[kT/n,(k+1)T/n]$ is the shift operator (bijection) given by $(S_k(z))_t := z_{t-kT/n}$.

Recall the first paragraph of Sect. 4.2. We now use the process Y^A in order to generate *at the same time* a measure $\mathbb{Q}^A$ and its Girsanov kernel $\psi^{\mathbb{Q}^A}$. Observe that $(Y_t^A)_{t\in[0,T]}$ satisfies the equation

$$Y_t^A = W_t + \frac{\mu t}{\sigma} - \sum_{k=0}^{n-1} \int_{t\wedge(kT/n)}^{t\wedge((k+1)T/n)} \Phi_t^A\big(u_k, S_k^{-1}(Y_{[kT/n,(k+1)T/n]}^A)\big)dt. \tag{4.23}$$

Since Φ^A is a bounded function, we obtain from Girsanov's theorem that there exists a probability measure $\mathbb{Q}^{A,n,\nu} = \mathbb{Q}^A \approx \mathbb{P}$ with finite entropy $\mathbb{E}_{\mathbb{Q}^A}[\log(d\mathbb{Q}^A/d\mathbb{P})] < \infty$ such that $W^{\mathbb{Q}^A} := Y^A$ is a Wiener process under $\mathbb{Q}^A$. From (4.21), (4.23) and the integration by parts formula, it follows that

$$\psi_t^{\mathbb{Q}^A} = (-A) \vee \left(\int_{kT/n}^t \theta\big(\phi_k(W_0^{\mathbb{Q}^A}, \ldots, W_{kT/n}^{\mathbb{Q}^A}), s - kT/n\big)dW_s^{\mathbb{Q}^A}\right) \wedge A - \frac{\mu}{\sigma}$$
$$\text{for } t \in [kT/n,(k+1)T/n), \qquad k = 0, 1, \ldots, n-1. \tag{4.24}$$

We end this step by arguing that there exists $N = N(K)$ such that for any $n > N(K)$, we have $\mathbb{Q}^A \in \mathcal{Q}^n$. First, we establish the martingale property. Indeed, from (4.24), it follows that for any $0 \le k \le n-1$ and $t \in [kT/n,(k+1)T/n)$, the conditional distribution (under $\mathbb{Q}^A$) of $\psi_t^{\mathbb{Q}^A}$ given $\mathcal{F}_{kT/n}^W$ is symmetric around $-\mu/\sigma$, and so $(X_{kT/n})_{0\le k\le n}$ is a $\mathbb{Q}^A$-martingale. Finally, we establish the supermartingale property. Clearly, there exists a constant $\bar{c} > 0$ which depends on K such that $|\theta(u,t)| \le \bar{c}$ for all $u \in [1/K, K]$ and $t \in [0, T/n]$. This together with (2.2) and (4.24) implies that there exists a parameter $N = N(K)$ such that for any $n > N(K)$,

$$\mathbb{E}_{\mathbb{Q}^A}\big[|X_{(k+1)T/n} - X_{kT/n}|^3 \big| \mathcal{F}_{kT/n}^W\big] \le h(n), \qquad k = 0, 1, \ldots, n-1,$$

as required.

Step 3: In this step, we fix arbitrary $n > N(K)$ and $\nu \in \mathcal{V}_n^K$. Then in view of (4.19), in order to complete the proof, it remains to show that

$$\sup_{\mathbb{Q}\in\mathcal{Q}^n} \mathbb{E}_{\mathbb{Q}}\left[f^n(X) - \frac{1}{2\lambda}\int_0^T |\psi_s^{\mathbb{Q}}|^2 ds\right]$$
$$\ge \mathbb{E}_{\mathbb{P}}\left[f^n(X^{(\nu)}) - \frac{1}{2\ell T}\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right)dt\right] - \frac{\mu^2 T}{2n\ell\sigma^2}. \tag{4.25}$$

By definition,

$$\sup_{\mathbb{Q}\in\mathcal{Q}^n} \mathbb{E}_{\mathbb{Q}}\left[f^n(X) - \frac{1}{2\lambda}\int_0^T |\psi_s^{\mathbb{Q}}|^2 ds \right] \geq \liminf_{A\to\infty} \mathbb{E}_{\mathbb{Q}^A}\left[f^n(X) - \frac{1}{2n\ell}\int_0^T |\psi_t^{\mathbb{Q}^A}|^2 dt \right]. \tag{4.26}$$

From (4.20), (4.24) and Lemma 4.1, it follows that for any $A > 0$ and $0 \leq k \leq n-1$,

$$\mathbb{E}_{\mathbb{Q}^A}\left[\int_{kT/n}^{(k+1)T/n} |\psi_t^{\mathbb{Q}^A}|^2 dt \middle| \mathcal{F}^W_{kT/n} \right] \leq \frac{\mu^2 T}{\sigma^2 n} + g\left(\frac{\phi_k(W_0^{\mathbb{Q}^A}, \ldots, W_{kT/n}^{\mathbb{Q}^A})}{\sigma^2} \right).$$

Thus

$$\mathbb{E}_{\mathbb{Q}^A}\left[\int_0^T |\psi_t^{\mathbb{Q}^A}|^2 dt \right] \leq \mu^2 T/\sigma^2 + \frac{n}{T}\mathbb{E}_{\mathbb{P}}\left[\int_0^T g\left(\frac{\nu_t}{\sigma^2}\right) dt \right]. \tag{4.27}$$

Finally, from (2.1) and (4.24), we obtain

$$\mathbb{Q}^A \circ (X_0, X_{T/n}, \ldots, X_T)^{-1} \Longrightarrow \mathbb{P} \circ (Y_0, Y_1, \ldots, Y_n)^{-1} \qquad \text{as } A \to \infty, \tag{4.28}$$

where $Y_0 := X_0$ and for any $k = 0, 1, \ldots, n-1$,

$$\begin{aligned} &Y_{k+1} - Y_k \\ &:= \sigma(W_{(k+1)T/n} - W_{kT/n}) \\ &\quad + \sigma \int_{kT/n}^{(k+1)T/n} \left(\int_{kT/n}^t \theta\big(\phi_k(W_0, \ldots, W_{kT/n}), s - kT/n\big) dW_s \right) dt \\ &= \sigma \int_{kT/n}^{(k+1)T/n} \Big(1 + \big((k+1)T/n - s\big)\theta\big(\phi_k(W_0, \ldots, W_{kT/n}), s - kT/n\big)\Big) dW_s, \end{aligned}$$

where the last equality follows from Fubini's theorem. From the Itô isometry and (4.20), it follows that for any k,

$$\mathbb{E}_{\mathbb{P}}[(Y_{k+1} - Y_k)^2 | Y_0, \ldots, Y_k] = \frac{T}{n}\phi_k(W_0, \ldots, W_{kT/n}).$$

We conclude that

$$\mathbb{P} \circ (Y_0, Y_1, \ldots, Y_n)^{-1} = \mathbb{P} \circ (X_0^\nu, X_{\frac{T}{n}}^\nu, \ldots, X_T^\nu)^{-1},$$

and so from (4.28) and Fatou's lemma,

$$\liminf_{A\to\infty} \mathbb{E}_{\mathbb{Q}^A}[f^n(X)] \geq \mathbb{E}_{\mathbb{P}}[f^n(X^{(\nu)})].$$

This together with (4.26) and (4.27) completes the proof of (4.25). □

Acknowledgements The authors thank the anonymous AE and reviewer for their valuable reports and comments which helped to improve the quality of this paper.

References

1. Bank, P., Dolinsky, Y.: A note on utility indifference pricing with delayed information. SIAM J. Financ. Math. **12**, SC31–SC43 (2021)
2. Barles, G., Soner, H.M.: Option pricing with transaction costs and a nonlinear Black–Scholes equation. Finance Stoch. **2**, 369–397 (1998)
3. Bertsimas, D., Kogan, L., Lo, A.W.: When is time continuous? J. Financ. Econ. **55**, 173–204 (2000)
4. Cai, J., Fukasawa, M., Rosenbaum, M., Tankov, P.: Optimal discretization of hedging strategies with directional views. SIAM J. Financ. Math. **7**, 34–69 (2016)
5. Carmona, R.: Indifference Pricing: Theory and Applications. Princeton University Press, Princeton (2012)
6. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**, 463–520 (1994)
7. Dolinsky, Y.: Numerical schemes for G-expectations. Electron. J. Probab. **17**, 1–15 (2012)
8. Fleming, W.H.: Risk sensitive stochastic control and differential games. Commun. Inf. Syst. **6**, 161–177 (2006)
9. Föllmer, H., Schied, A.: Stochastic Finance. An Introduction in Discrete Time, 4th edn. de Gruyter, Berlin (2016)
10. Frittelli, M.: The minimal entropy martingale measure and the valuation problem in incomplete markets. Math. Finance **10**, 39–52 (2000)
11. Fukasawa, M.: Discretization error of stochastic integrals. Ann. Appl. Probab. **21**, 1436–1465 (2011)
12. Fukasawa, M.: Efficient discretization of stochastic integrals. Finance Stoch. **18**, 175–208 (2014)
13. Galichon, A., Henry-Labordère, P., Touzi, N.: A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. Ann. Appl. Probab. **24**, 312–336 (2014)
14. Geiss, S.: Quantitative approximation of certain stochastic integrals. Stoch. Stoch. Rep. **73**, 241–270 (2002)
15. Geiss, S.: Weighted BMO and discrete time hedging within the Black–Scholes model. Probab. Theory Relat. Fields **132**, 13–38 (2005)
16. Geiss, S., Toivola, A.: Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. Bernoulli **15**, 925–954 (2009)
17. Gobet, E., Makhlouf, A.: The tracking error rate of the delta-gamma hedging strategy. Math. Finance **22**, 277–309 (2012)
18. Gobet, E., Temam, E.: Discrete time hedging errors for options with irregular payoffs. Finance Stoch. **5**, 357–367 (2001)
19. Hayashi, T., Mykland, P.A.: Evaluating hedging errors: an asymptotic approach. Math. Finance **15**, 309–343 (2005)
20. Hernández-Hernández, D.: On the relation between risk sensitive control and indifference pricing. In: 2008 American Control Conference, pp. 1013–1016. IEEE Press, New York (2008)
21. Kallenberg, O.: On the existence of universal functional solutions to classical SDEs. Ann. Probab. **24**, 196–205 (1996)
22. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus, 2nd edn. Springer, New York (1991)
23. Napp, C.: The Dalang–Morton–Willinger theorem under cone constraints. J. Math. Econ. **39**, 111–126 (2003)
24. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin (1999)
25. Skorohod, A.V.: On a representation of random variables. Teor. Veroâtn. Primen. **21**, 645–648 (1976)