ON CONSISTENT VALUATIONS BASED ON DISTORTED EXPECTATIONS: FROM MULTINOMIAL RANDOM WALKS TO LÉVY PROCESSES

DILIP MADAN, MARTIJN PISTORIUS, AND MITJA STADJE

Abstract. A distorted expectation is a Choquet expectation with respect to the capacity induced by a concave probability distortion. Distorted expectations are encountered in various static settings, in risk theory, mathematical finance and mathematical economics. There are a number of different ways to extend a distorted expectation to a multi-period setting, which are not all time-consistent. One time-consistent extension is to define the non-linear expectation by backward recursion, applying the distorted expectation stepwise, over single periods. In a multinomial random walk model we show that this non-linear expectation is stable when the number of intermediate periods increases to infinity: Under a suitable scaling of the probability distortions and provided that the tick-size and time step-size converge to zero in such a way that the multinomial random walks converge to a Lévy process, we show that values of random variables under the multi-period distorted expectations converge to the values under a continuous-time non-linear expectation operator, which may be identified with a certain type of Peng's g-expectation. A coupling argument is given to show that this operator reduces to a classical linear expectation when restricted to the set of pathwise increasing claims. Our results also show that a certain class of g-expectations driven by a Brownian motion and a Poisson random measure may be computed numerically by recursively defined distorted expectations.

1. Introduction

A distorted expectation is a classical example of a Choquet expectation, which is itself an instance of a non-linear expectation. While an expectation may be seen as an integral of the survival function, i.e.,

$$\mathbb{E}[\mathcal{X}] = \int_{-\infty}^{0} (S_{\mathcal{X}}(t) - 1)\mathrm{d}t + \int_{0}^{\infty} S_{\mathcal{X}}(t)\mathrm{d}t,$$

with $S_{\mathcal{X}}(t) := \mathbb{P}[\mathcal{X} > t]$, a distorted expectation is computed by integrating an *upwardly shifted* survival function. The upward shift for every survival probability is induced by a given concave *probability distortion*, say D, which is an increasing function that is a surjective mapping from the unit square onto itself—see Figure 1 for two examples of concave distortions. The distorted expectation is then the Choquet expectation defined by

$$\mathcal{C}^D[\mathcal{X}] := \int_{-\infty}^{0} (D(S_{\mathcal{X}}(t)) - 1)\mathrm{d}t + \int_{0}^{\infty} D(S_{\mathcal{X}}(t))\mathrm{d}t.$$

Shifting the survival function upwards (resulting from $D \geq \mathrm{id}$, the identity) means increasing across the board the probabilities that certain values will be exceeded, creating a safety buffer. Concavity ensures that the relative shift increases the closer one gets to the left tail. There exists a good deal of literature concerning static Choquet

Date: January 17, 2013.

2000 *Mathematics Subject Classification.* 93E20, 91B28.

Key words and phrases. g-expectation, non-linear expectation, probability distortion, option pricing, risk measurement, convergence, Lévy process, multinomial tree.

Acknowledgements: We thank Neil Walton and Bert Zwart for useful suggestions.

MRP acknowledges support by the NWO-STAR cluster, and by EPSRC Platform Grant EP/I019111/1.

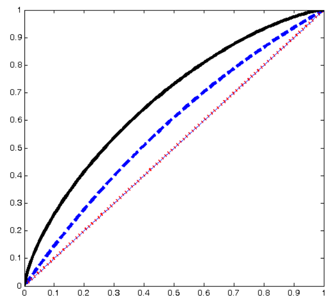

FIGURE 1. Depicted are three distortions. The diagonal line is equal to the graph of the linear distortion $\Psi_0(p) = p$, while the solid and dashed curves correspond to the graphs on the unit square of the MINMAXVAR distortion $\Psi_\gamma(p) = 1 - (1 - p^{1/(1+\gamma)})^{1+\gamma}$ for $\gamma = 0.4$ and the exponential distortion $\Psi_\alpha(p) = 1 - \frac{1-e^{-\alpha p}}{1-e^{-\alpha}}$ for $\alpha = 0.9$.

expectations and their mathematical properties—see for instance Anger [1] or Dellacherie [16]. By modifying one axiom of von Neumann-Morgenstern's expected utility theory, Yaari [46] gives an axiomatic foundation of distorted expectations in order to describe choices under uncertainty. The concavity of the distortion function is identified with uncertainty aversion. For further applications of Choquet expectations to model preferences under uncertainty see for example Sarin & Wakker [38], Schmeidler [41] and Wakker [43] and references therein.

A distorted expectation may also be interpreted in terms of model robustness in the sense of being the largest or smallest value among all members of a family of models that is induced by the distortion function— see Carlier & Dana [6]. Robust approaches of this type can be found in robust statistics (Huber [24]) and in the theory of coherent risk measures (Artzner *et al.* [2]). Kusuoka [29] showed that a distribution-invariant, comonotone additive, coherent risk measure necessarily corresponds to a distorted expectation (with Average Value at Risk being the prime example), while Dana [15] proved that any distribution-invariant risk measure that respects second-order stochastic dominance admits a representation as a supremum of Choquet expectations.

Because of their direct link to tail probabilities, distorted expectations have also been used extensively to calculate insurance premiums (see for example Wang *et al.* [44], or Wang [45]) and to model bid-ask spreads in finance (see Cherny & Madan [10] or Madan & Schoutens [31]).

While Choquet expectations play a fundamental role in static settings, this is much less the case in dynamic settings. The reason is that, contrary to what is the case for standard expectations, the collection of "conditional Choquet expectations" corresponding to the collection of "updated" probability measures, that is equal to the sequence of Choquet expectations evaluated with respect to the conditional probability measures conditioned on the sigma-algebras in a given filtration, may lead to time-*inconsistent* choices. For instance, it is possible that for two epochs s and t with $s < t$ in every future scenario the conditional Choquet expectation of X at time t will be greater than that of Y, while nevertheless at time s the conditional Choquet expectation of Y is greater than that of X. This fact suggests that a *dynamically* consistent non-linear expectation that is based on distorted expectations must apply the distortion over *single* time periods only. In a random walk setting this is equivalent to considering a collection of models which is specified by all those measures of which the one-period "transition probabilities" are dominated by the capacity induced by the distortion and the one-period transition probabilities of the random walk. The resulting multi-period valuation operator that is defined as the supremum over all values attained under conditional expectations with respect to the filtration generated by the random walk and with respect to probability measures in this collection, can indeed be shown to be time-consistent (see Proposition 2 below). This valuation operator is non-linear, which is apparent from the fact that the probability measure employed in the evaluation will be dependent on the random variable. Furthermore, this multi-period distorted expectation operator inherits the positive homogeneity and convexity from the single-period mapping, but is neither distribution-invariant nor co-monotontically additive. Valuation under dynamic risk measures in discrete-time settings was also studied by, among others, Cherny [9], Cohen & Elliot [11] Jobert & Rogers [26] and Roorda *et al.* [37].

It is an interesting question whether such a valuation method remains stable in a setting where data are observed frequently, so that also frequent updates are needed. More concretely this question may be phrased as asking if the discrete-time multi-period distorted expectations converge to a continuous-time valuation operator if the number of intermediate periods increases to infinity. It turns out that in order for this question to be answered positively it is necessary to scale the distortion function appropriately. In a Brownian setting, Stadje [42] identified the square root scaling as the only one ensuring that the discrete-time evaluations do neither explode nor converge to a simple linear expectation (which would rule out ambiguity aversion). In Madan & Schoutens [31], it was observed, in the case of multinomial random walks converging to a particular Lévy process, that the limiting g-expectation was driven only by a Brownian motion when the square root scaling was employed, suggesting that on the square root scale jumps cannot be observed in the limit. However, from a risk management perspective jumps, in particular over a short time horizon, are inherent drivers of market risk. Therefore, a key point in the development below is the identification of a suitable scaling of the series of distortions (given in Definition 9 below) that also takes into account the jump risk. We will give sufficient conditions for a suitable scaling, identify the limit and prove convergence results. The mathematical details of these proofs, especially for the jump part, are delicate. The limit turns out to belong to a certain class g-expectations driven by a Wiener process and a Poisson random measure, with driver g identified explicitly in Eqn. (4.6) below. g-expectations, which in Markovian settings are expressed in terms of the solutions of semi-linear PDEs, were originally proposed by Peng [33] in a Brownian setting. We will call the limiting g-expectation the *expectation under drift and jump-rate distortion.* It inherits a number of the properties of Choquet expectations (convexity, positive homogeneity, monotonicity), but is neither law-invariant nor comonotonically additive. In fact, the only mapping that is time-consistent, convex and law-invariant is the entropic risk measure, as shown by Kupper & Schachermayer [27]. By employing a coupling argument we also show that the limiting non-linear expectation is additive on the set of random variables that are pathwise increasing functions of the underlying Lévy process.

There is a close connection between g-expectation and the notion of time-consistency which we will also call filtration-consistency (following Coquet *et al.* [14]). We recall that in the setting of a given probability space $(\Omega, \mathcal{F}, P)$, a non-linear expectation Π was defined by Coquet *et al.* [14] to be a real-valued map on $L^2(\Omega, \mathcal{F}, P)$ that is *strictly monotone* and *preserves constants*, that is,

$$\mathcal{X} \geq \mathcal{Y} \ P\text{-a.s. implies } \Pi(\mathcal{X}) \geq \Pi(\mathcal{Y}), \text{ with equality precisely if } P(\mathcal{X} > \mathcal{Y}) = 0, \tag{1.1}$$

$$\Pi(c) = c \quad \text{for all } c \in \mathbb{R}. \tag{1.2}$$

Furthermore, given a filtration $\mathbf{F} = \{\mathcal{F}_t\}$, a collection of mappings $\{\Pi_t\}$ with $\Pi_t : L^2(\Omega, \mathcal{F}, P) \to L^2(\Omega, \mathcal{F}_t, P)$ is defined in [14] to be $\mathbf{F}$-*consistent* if it satisfies, for all t, the equation

$$\Pi(\mathcal{X}\, I_H) = \Pi(\Pi_t(\mathcal{X})\, I_H) \quad \mathcal{X} \in L^2(\Omega, \mathcal{F}, \mathbb{P}), H \in \mathcal{F}_t, \tag{1.3}$$

where I_H denotes the indicator of the set H. It is well known (see for instance Cheridito & Kupper [8]) that for monotone and constant preserving conditional valuations which are normalized in the sense that $\Pi_t(0) = 0$ time-consistency or filtration-consistency is equivalent to the condition that, for every $s \leq t$ and $\mathcal{X}, \mathcal{Y}$,

$$\Pi_t(\mathcal{X}) \geq \Pi_t(\mathcal{Y}) \quad \text{implies} \quad \Pi_s(\mathcal{X}) \geq \Pi_s(\mathcal{Y}).$$

Filtration-consistent evaluations were developed in a discrete state-space setting in Cohen & Elliot [11, 12]. Coquet *et al.* [14] (in the Brownian setting) and Royer [36] (in the case of a driving Brownian and Poisson random measure) showed that any continuous time non-linear expectation that is filtration-consistent and that satisfies a domination property must be a g-expectation for some driver g, that is, it must solve a backward stochastic differential equation with driver g.

The convergence results that are established in Section 5 also suggest an easy way to evaluate certain g-expectations and the solutions of the corresponding semi-linear PDEs numerically via Choquet expectations, if the drivers g are of the form given in Eqn. (4.6) below. As computing recursively the distorted expectations avoids calculating the Malliavin derivative (which is typically the most demanding part) and the corresponding functional of the jump part (which in a setting featuring jumps seems an even more challenging task), this method is more efficient than the computational schemes that are currently available.

Contents. The remainder of the paper is organised as follows. In Section 2 preliminary results are collected concerning Choquet integration and distortions, which will be referred to throughout the paper. Section 3 is devoted to the multi-period valuation operator defined in a multinomial random walk setting, given a concave probability distortion and a filtration. Section 4 is concerned with the non-linear expectation under drift and jump-rate uncertainty. The distortion scaling and the convergence theorem (Theorem 1) are provided in Section 5. In Section 6 the form of the non-linear expectation for pathwise increasing claims is identified, using a coupling argument (Theorem 2). By way of illustration two examples are provided in Section 7. Sections 8, 9 and 10 contain key auxiliary results, and the proofs of the upper bound and lower bound, respectively, which together form the proof of Theorem 1. Some proofs are deferred to the Appendix.

2. Preliminaries: Choquet integration and distortion

In this section key properties are collected of Choquet integrals. Dennenberg [19] and Föllmer & Schied [23, Ch. 4] provide treatments on Choquet integration (induced by distortions). Unlike the treatment in [23, Ch. 4], which is in a setting of bounded random variables on a probability space, the setting below concerns square-integrable functions and general measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel-sigma algebra over $\mathbb{R}$.

Let μ be a given (Lévy) measure on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that integrates the function $x \mapsto x^2 \wedge 1$ (with $x \wedge y = \min\{x, y\}$ for $x, y \in \mathbb{R}$), that is,

$$\int_{\mathbb{R}} [x^2 \wedge 1]\mu(\mathrm{d}x) < \infty, \tag{2.1}$$

and denote by $L^2(\mu)$ and $L^2_+(\mu)$ the collections of real-valued and non-negative measurable functions that are square-integrable with respect to μ.

Definition 1. A *(measure) distortion* is a continuous increasing function $D : \mathbb{R}_+ \to \mathbb{R}_+$ with $D(0) = 0$. In particular, a *probability-distortion* D is the restriction of a distortion to the unit interval $[0, 1]$ with $D(1) = 1$. s To a probability distortion D is associated another probability distortion $\widehat{D}$ given by

$$\widehat{D}(x) = 1 - D(1 - x), \qquad x \in [0, 1]. \tag{2.2}$$

Any distortion induces a capacity on the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Definition 2. A *measure capacity* on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a monotone set function $c : \mathcal{B}(\mathbb{R}_+) \to \mathbb{R}_+$ with $c(\emptyset) = 0$: $c(A) \le c(B)$ for all sets $A, B \in \mathcal{B}(\mathbb{R}_+)$ with $A \subset B$. A *capacity* is a measure capacity that is finite and normalised to unity $[c(\mathbb{R}_+) = 1]$.

In particular, note that any measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a measure capacity, and any probablity measure is a capacity. To any measure distortion D is associated a measure capacity $D \circ \mu$ given by

$$(D \circ \mu)(A) = D(\mu(A)), \qquad A \in \mathbb{B}(\mathbb{R}_+). \tag{2.3}$$

Furthermore, if D is a probability distortion and μ is a probability measure, then both $D \circ \mu$ and $\widehat{D} \circ \mu$ are capacities.

Given a capacity, a corresponding integral can be defined called a Choquet integral. In particular, for any measure distortion D the Choquet integral $\mathcal{C}^D[\mathcal{X}]$ of a function $\mathcal{X}$ in $L^2_+(\mu)$ corresponding to the measure capacity $D \circ \mu$ is given by

$$\mathcal{C}^D[\mathcal{X}] = \int_{[0,\infty)} (D \circ \mu)(\mathcal{X} > x)\mathrm{d}x. \tag{2.4}$$

In the case that μ is a probability measure and D a probability distortion the Choquet integral $\mathcal{C}^D[\mathcal{X}]$ of $\mathcal{X} \in L^2(\mu)$ is given by

$$\mathcal{C}^D[\mathcal{X}] \;=\; \int_{[0,\infty)} (D \circ \mu)(\mathcal{X} > x)\mathrm{d}x - \int_{(0,\infty)} (\widehat{D} \circ \mu)(\mathcal{X} \le x)\mathrm{d}x, \tag{2.5}$$

where $\mathcal{C}^D[\mathcal{X}]$ is defined to take the value $-\infty$ if the first and second integral in eqn. (2.5) are both infinite. We will throughout we restrict our attention to the following class of distortions D:

Assumption. Assume that the (measure/probability) distortion D is *concave* or *convex* and satisfies the integrability condition

$$K^D := \begin{cases} \displaystyle\int_0^\infty D(y)\frac{\mathrm{d}y}{y\sqrt{y}} < \infty, & \text{if } D \text{ is a measure distortion,} \\ \displaystyle\int_0^1 [D(y) + \widehat{D}(y)]\frac{\mathrm{d}y}{y\sqrt{y}}\mathrm{d}y < \infty, & \text{if } D \text{ is a probability distortion.} \end{cases} \tag{2.6}$$

The integrability condition in Eqns. (2.6) guarantees that the Choquet integrals in Eqns. (2.4) and (2.5) are finite if the integrand $\mathcal{X}$ is square-integrable:

Lemma 1. *The map $\mathcal{C}^D : L^2_+(\mu) \to \mathbb{R}_+$ given by $\mathcal{X} \mapsto \mathcal{C}^D[\mathcal{X}]$ is Lipschitz continuous. In particular, for any $\mathcal{X} \in L^2_+(\mu)$ we have*

$$\mathcal{C}^D[\mathcal{X}] \leq K_D\sqrt{c} \quad \textit{with } c = \mu(\mathcal{X}^2) = \int_{\mathbb{R}_+} \mathcal{X}(z)^2 \mu(\mathrm{d}z) \tag{2.7}$$

and K_D given in Eqns. (2.6)*. If μ is a probability measure, then the estimate in Eqn.* (2.7) *holds true for any $\mathcal{X} \in L^2(\mu)$.*

Proof. For a given $\mathcal{X} \in L^2_+(\mu)$, measure μ and measure distortion D, Chebyshev's inequality and a change of variables show

$$\int_0^\infty (D \circ \mu)(\mathcal{X} > x)\mathrm{d}x \leq \int_0^\infty D(c/x^2)\mathrm{d}x = K_D\sqrt{c},$$

with c given in Eqn. (2.7). The stated Lipschitz continuity follows by combining the representation in Proposition 1 below with the estimate in Eqn. (2.7). In the case that μ is a probability measure and D a probability distortion, the integrability and Lipschitz continuity can be verified in a similar fashion. □

For later reference we collect a number of basic properties of the map $\mathcal{C}^D : L^2_+(\mu) \to \mathbb{R}_+$ given by $\mathcal{X} \mapsto \mathcal{C}^D[\mathcal{X}]$.

Lemma 2. (i) *If D is strictly increasing, $\mathcal{C}^D$ is* strictly monotone *that is, for any $\mathcal{X}, \mathcal{Y} \in L^2(\mu)$ with $\mathcal{X} \leq \mathcal{Y}$, we have $\mathcal{C}^D[\mathcal{X}] \leq \mathcal{C}^D[\mathcal{Y}]$, with equality if and only if $\mu(\mathcal{X} > \mathcal{Y}) = 0$.*

(ii) *$\mathcal{C}^D$ is* continuous from below*, that is, if $(\mathcal{X}_n), \mathcal{X} \in L^2_+(\mu)$ and we have $\mathcal{X}_n \nearrow \mathcal{X}$, then it holds $0 \leq \mathcal{C}^D[\mathcal{X}_n] \nearrow \mathcal{C}^D[\mathcal{X}]$.*

(iii) *$\mathcal{C}^D$ is* positively homogeneous *and* (positively) translation-invariant*, that is, for $\mathcal{X} \in L^2_+(\mu)$ and $c, d \in \mathbb{R}_+$ we have $\mathcal{C}^D[c\,\mathcal{X}] = c\,\mathcal{C}^D[\mathcal{X}]$ and $\mathcal{C}^D[\mathcal{X} + d] = \mathcal{C}^D[\mathcal{X}] + d$.*

If μ is a probability measure, the results in this lemma remain valid if $\mathcal{C}^D$ is defined to be a real-valued map on $L^2(\mu)$ (an observation that is a direct consequence of the definition of $\mathcal{C}^D[\mathcal{X}]$).

The Choquet integral $\mathcal{C}^D[\mathcal{X}]$ admits a representation as a supremum over a collection of measures (which was establised by Carlier & Dana [6] in the case of bounded $\mathcal{X}$). Next the L^2-version of this representation is stated, which is the result that will be deployed in the subsequent analysis. Let $\mathcal{M}^{ac}_{p,\mu}$, $p \geq 1$, denote the collection of measures m on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ that are absolutely continuous with respect to the measure μ, and have a density in $L^p(\mu)$, and denote by $\mathcal{B}^\mu(\mathbb{R})$ the subset of the sets $A \in \mathcal{B}(\mathbb{R})$ for which $\mu(A)$ is finite. In the sequel we will use the following relation between capacities which is a generalization of the notion of stochastic dominance of probability measures:

Definition 3. Given two (measure) capacities c, c' on the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ we write $c \prec c'$ and say that c' dominates c when it holds

$$c(A) \leq c'(A) \qquad \text{for all } A \in \mathcal{B}^\mu(A).$$

Proposition 1. *For any $\mathcal{X} \in L^2_+(\mu)$ the Choquet integral $\mathcal{C}^D[\mathcal{X}]$ is finite and admits the representation*

$$\mathcal{C}^D[\mathcal{X}] = \sup_{m \in \mathcal{M}^D_{2,\mu}} m[\mathcal{X}], \tag{2.8}$$

where the supremum is attained, with

$$\mathcal{M}^D_{p,\mu} = \left\{ m \in \mathcal{M}^{ac}_{p,\mu} : m \prec D \circ \mu \right\}, \quad p \geq 1. \tag{2.9}$$

Remark 1. The statement in Proposition 1 holds for any $\mathcal{X} \in L^2(\mu)$ if μ is a probability measure. Furthermore, by considering the complements of sets we deduce the equality

$$\mathcal{M}^D_{p,\mu} = \left\{ m \in \mathcal{M}^{ac}_{p,\mu} : \widehat{D} \circ \mu \prec m \prec D \circ \mu \right\}, \quad p \geq 1.$$

In the case of bounded $\mathcal{X}$ and a probability measure μ the proof of Proposition 1 can be found in Schied [40, Thm. 1.51], Föllmer & Schied [23, Thm. 4.79, 4.94]. The proof of the remaining cases is provided in the Appendix. The proof rests on a lemma concerning integrability of the Radon-Nikodym derivatives of the measures in the set $\mathcal{M}^D_1$, which we give below as it will be used in the sequel:

Lemma 3. *The collection $\mathcal{R} = \{Z = \frac{\mathrm{d}m}{\mathrm{d}\mu} : m \in \mathcal{M}^D_1\}$ is contained in a ball in $L^2(\mu)$.*

Proof. Let $Z \in \mathcal{R}$ and define, for any $M > 0$, $m_M \in \mathcal{R}$ to be the measure with Radon-Nikodym derivative given by Z_M where $Z_M(x) = (Z(x) \wedge M)I_{\{|x|>1/M\}}$. Then the estimate in Eqn. (2.7) and the fact $Z \in L^2(\mu)$ imply

$$\mu(Z_M^2) = m_M(Z_M) \leq K_D \mu(Z_M^2)^{1/2},$$

so that $\mu(Z_M^2) \leq K_D^2$, where K_D is given in Eqn. (2.6). Since we have $Z_M \nearrow Z$ as $M \nearrow \infty$, the Monotone Convergence Theorem implies $\mu(Z^2) \leq K_D^2$. □

3. Multi-period valuation by distortion of transition probabilities

Let be given a probability distortion Ψ and a filterred probability space $(\mathcal{A}, \mathcal{G}, \mathbf{G}, P)$ where $\mathbf{G} = \{\mathcal{G}_i\}_{i=0}^n$ denotes a filtration of sigma-algebras $\mathcal{G}_0, \ldots, \mathcal{G}_n$. In this section we set out to identify a non-linear expectation that is $\mathbf{G}$-consistent in the sense of Eqn. (1.3) and that, over a single period, reduces to a distorted expectation with respect to the distortion Ψ. The description of this operator that we give below is based on a number of related notions that we describe next. For any sub-sigma algebra $\mathcal{H} \subset \mathcal{G}$ denote by $L^0(\mathcal{H})$ the collection of real-valued functions on $\mathcal{A}$ that are measurable with respect to $\mathcal{H}$.

Definition 4. Let $\mathcal{H}'$ and $\mathcal{H}$ be two sigma algebras with $\mathcal{H} \subset \mathcal{H}' \subset \mathcal{G}$.

(i) A random set function $C : \mathcal{H}' \to L^0(\mathcal{H})$ is called *absolutely continuous with respect to* P if we have $C(A) = 0$ P-a.s. for all sets $A \in \mathcal{H}'$ with $P(A) = 0$.

(ii) We call a random set function $C : \mathcal{H}' \to L^0(\mathcal{H})$ an *$\mathcal{H}$-measurable capacity on* $(\mathcal{A}, \mathcal{H}', P)$ when C is absolutely continuous with respect to P and we have

(a) $C(\emptyset) = 0$ P-a.s.,
(b) [normalisation] $C(\mathcal{A}) = 1$ P-a.s., and
(b) [monotonicity] $C(A) \leq C(B)$ P-a.s., for any $A, B \in \mathcal{H}'$ with $A \subset B$.

(iii) For $\mathcal{H}$-measurable capacities C, C' on $(\mathcal{A}, \mathcal{H}', P)$ we write $C \prec' C'$ and say that C' dominates C P-a.s., if it holds

$$C(A) \leq C'(A) \qquad P\text{-a.s., for any } A \in \mathcal{H}'.$$

The distortion Ψ induces two $\mathcal{G}_i$-measurable capacities. Denoting by P_i the $\mathcal{G}_i$-conditional expectations defined on $(\mathcal{A}, \mathcal{G}_{i+1})$ by $P_i(A) = E[I_A|\mathcal{G}_i]$, for $A \in \mathcal{G}_{i+1}$ and $i = 0, \ldots, n-1$, where I_A denotes the indicator of the set A and E is the expectation under P, define the random set functions $\Psi\circ P_i$ and $\widehat{\Psi}\circ P_i : \mathcal{G}_{i+1}\to L^\infty(\mathcal{G}_i)$ by

$$(\Psi\circ P_i)(A) = \Psi(P_i(A)) \qquad (\widehat{\Psi}\circ P_i)(A) = \Psi(P_i(A)), \qquad A\in\mathcal{G}_{i+1}.$$

Furthermore, any pair of a distortion Ψ and a filtration $\mathbf{G}$ gives rise to a capacity. Define the set-function $c^{\Psi,\mathbf{G}} : \mathcal{G}\to[0,1]$ by

$$c^{\Psi,\mathbf{G}}(A) = \sup_{Q\in\mathcal{D}^\Psi(\mathbf{G})} Q(A) \qquad A\in\mathcal{G},$$

where $\mathcal{D}^\Psi(\mathbf{G})$ denotes the subset of $\mathcal{P}^{ac}_{2,P} = \{m\in\mathcal{M}^{ac}_{2,P} : m(\mathcal{A}) = 1\}$ given by

$$\mathcal{D}^\Psi(\mathbf{G}) = \left\{Q\in\mathcal{P}^{ac}_{2,P} : \widehat{\Psi}\circ P_i \prec' Q_i \prec' \Psi(P_i), \text{ for all } i = 0,1,\ldots,n-1\right\}, \tag{3.1}$$

where Q_i denotes the $\mathcal{G}_i$-conditional expectations defined on $(\mathcal{A},\mathcal{G}_{i+1})$ by $Q_i(A) = E^Q[I_A|\mathcal{G}_i]$ for $A\in\mathcal{G}_{i+1}$, where E^Q denotes the expectation under Q.

By direct verification of the relevant definitions it follows that these three maps are indeed (measurable) capacities.

Lemma 4. *(i) The random set functions $\Psi\circ P_i$ and $\widehat{\Psi}\circ P_i$ are $\mathcal{G}_i$-measurable capacities on $(\mathcal{A},\mathcal{G}_{i+1})$.*

(ii) The set-function $c^{\Psi,\mathbf{G}}$ is a capacity.

The capacity $c^{\Psi,\mathbf{G}}$ gives in turn rise to a non-linear expectation operator, that we will call the Ψ-distorted (conditional) expectation with respect to the filtration $\mathbf{G}$.

Definition 5. (i) The Ψ-distorted-expectation with respect to the filtration $\mathbf{G}$ is the map $\mathcal{C}^{\Psi,\mathbf{G}} : L^2(\mathcal{G},P)\to\mathbb{R}$ given by

$$\mathcal{C}^{\Psi,\mathbf{G}}(\mathcal{X}) = \sup_{Q\in\mathcal{D}^\Psi(\mathbf{G})} E^Q[\mathcal{X}], \qquad \mathcal{X}\in L^2(\mathcal{G},P). \tag{3.2}$$

(ii) For $i = 0,\ldots,n-1$, the Ψ-distorted $\mathcal{G}_i$-conditional expectation with respect to the filtration $\mathbf{G}$ is $\mathcal{C}^{\Psi,\mathbf{G}}(\,\cdot\,|\mathcal{G}_i) : L^2(\mathcal{G},P)\to L^2(\mathcal{G}_i,P)$ given by

$$\mathcal{C}^{\Psi,\mathbf{G}}(\mathcal{X}|\mathcal{G}_i) = \operatorname*{ess.\,sup}_{Q\in\mathcal{D}^\Psi(\mathbf{G})} E^Q(\mathcal{X}|\mathcal{G}_i), \qquad \mathcal{X}\in L^2(\mathcal{G},P). \tag{3.3}$$

The finiteness of the non-linear expectation $\mathcal{C}^{\Psi,\mathbf{G}}[\mathcal{X}]$ for $X\in L^2(\mathcal{G},P)$ is confirmed in the next result, where it is furthermore shown that the collection of maps $\mathcal{C}^\Psi(\,\cdot\,|\mathcal{G}_i)$ for $i = 0,1,\ldots,n-1$, is $\mathbf{G}$-consistent, and reduces, in a single period setting to a (conditional) Choquet integral:

Proposition 2. *(i) We have $\sup_{Q\in\mathcal{D}^\Psi(\mathbf{G})} E^Q[Z_Q^2] < \infty$, where Z_Q denotes the Radon-Nikodym derivative of Q with respect to P on $\mathcal{G}_n$.*

(ii) The collection $\Pi = \{\Pi_m, m = 0,1,\ldots,n\}$ with $\Pi_m = \mathcal{C}^\Psi(\,\cdot\,|\mathcal{G}_m)$ is $\mathbf{G}$-consistent and, for $\mathcal{X}\in L^2(\mathcal{G},P)$, $\Pi_m(\mathcal{X})$, $m = 0,\ldots,n$, satisfies the backward recursion

$$\begin{cases}\Pi_m(\mathcal{X}) = \mathcal{C}^\Psi\left(\Pi_{m+1}(\mathcal{X})|\mathcal{G}_m\right), & m<n,\\ \Pi_n(\mathcal{X}) = \mathcal{X}.\end{cases} \tag{3.4}$$

(iii) For $\mathcal{X}\in L^2(\mathcal{G}_{i+1},P)$ the random variable $\mathcal{C}^{\Psi,\mathbf{G}}(\mathcal{X}|\mathcal{G}_i)$ satisfies

$$\mathcal{C}^{\Psi,\mathbf{G}}(\mathcal{X}|\mathcal{G}_i) = \int_{[0,\infty)} \Psi(P_i(\mathcal{X}>x))\mathrm{d}x + \int_{(-\infty,0)} [\Psi(P_i(\mathcal{X}>x)) - 1]\mathrm{d}x. \tag{3.5}$$

The proof of Proposition 2(i) is a straightforward adaptation of that of Lemma 2.7, while the proof of Proposition 2(ii) is well known (e.g., a version for bounded random variables is provided in [23, Thm. 11.22]).

Proof of Proposition 2(iii). The proof consists of two steps. The first step is to observe that the left-hand side of Eqn. (3.5) is equal to

$$\widetilde{\mathcal{C}}^{\Psi,\mathbf{G}}(\mathcal{X}|\mathcal{G}_i) := \underset{Q\in\mathcal{M}_i}{\text{ess. sup}}\, E_Q[\mathcal{X}|\mathcal{G}_i]$$

where $\mathcal{M}_i$ is equal to the set of probability measures given by $\mathcal{M}_i = \{Q \in \mathcal{P}^{ac}_{2,P} : \widehat{\Psi}\circ P_i \prec' Q_i \prec' \Psi\circ P_i\}$. This observation follows by noting that the measure that attains the supremum in the definition of $\mathcal{C}^{\Psi,\mathbf{G}}[\mathcal{X}]$ in fact also attains the supremum in the definition of $\widetilde{\mathcal{C}}^{\Psi,\mathbf{G}}(\mathcal{X}|\mathcal{G}_i)$. The second step is to note that the right-hand side of Eqn. (3.5) is equal to $\widetilde{\mathcal{C}}^{\Psi,\mathbf{G}}(\mathcal{X}|\mathcal{G}_i)$, which is a fact that follows by a line of reasoning that is analogous to the one that was used in the proof of Proposition 1. □

4. A continuous-time non-linear expectation

4.1. **Continuous-time setting.** Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process, that is, a stochastic process with independent and stationary increments that has right-continuous paths with left-limits and starts at zero, $X_0 = 0$. Let X be defined as the coordinate process on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$, i.e. $X_t(\omega) = \omega(t)$ for $t \in \mathbb{R}_+$ and $\omega \in \Omega$. Here $\Omega = D([0, T], \mathbb{R})$ is the Skorokhod space endowed with the Skorokhod metric, $\mathcal{F}$ denotes the Borel sigma algebra, and $\{\mathcal{F}_t\}_{t\in[0,T]}$ the right-continuous filtration generated by X. We denote by $\mathbb{E}$ the expectation under the measure $\mathbb{P}$.

The law of the Lévy process X is determined by its characteristic exponent $\psi : \mathbb{R} \to \mathbb{C}$ that is for any $t \in \mathbb{R}_+$ given by $E[\mathrm{e}^{\mathbf{i}\theta X_t}] = \exp(-t\psi(\theta))$ and that is by the Lévy-Khintchine formula of the form

$$\psi(\theta) = \frac{\sigma^2}{2}\theta^2 - \mathbf{i}\mathtt{d}\theta + \int_{\mathbb{R}} (1 - \exp\{\mathbf{i}\theta x\} + \mathbf{i}\theta x)\Lambda(\mathrm{d}x), \qquad \theta \in \mathbb{R},$$

where $\mathtt{d} = \mathbb{E}[X_t]/t$, $\mathtt{d} \in \mathbb{R}$, is the drift of X, $\sigma^2 = \mathrm{Var}[X^c_t]/t$, $\sigma^2 \geq 0$ is the instantaneous variance of the continuous martingale part X^c of X, and Λ is the Lévy measure of X, which is in turn characterised by the property that for any Borel set $A \subset \mathbb{R}_0 = \mathbb{R}\backslash\{0\}$, $\Lambda(A)t$ is equal to the expected number of jumps ΔX_s, $s \in [0, t]$ of X with $\Delta X_s \in A$. The measure Λ is defined on the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with $\Lambda(\{0\}) = 0$ and satisfies the integrability condition in Eqn. (2.1).

In order to ensure a finite second moment in an exponential Lévy model, we restrict ourselves to Lévy processes X that admit some finite exponential moment, by assuming that there exist a $q \geq 1$ such that we have

$$\mathbb{E}[\exp(2qX_t)] < \infty, \qquad \forall t \in [0, T]. \tag{4.1}$$

These moment conditions are equivalent to the condition (see e.g. [39] for a proof)

$$\int_{\mathbb{R}\backslash[-1,1]} [\exp(2qx)]\Lambda(\mathrm{d}x) < \infty. \tag{4.2}$$

4.2. **Notation.** Before proceeding we fix some notation that will be used in the sequel. We denote by $\mathcal{P}$ the predictable sigma-algebra, write $\widetilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+)$, and use the following notation:

$\mathcal{P}^{ac}_{2,\mathbb{P}}$: the set of probability measures on $(\Omega, \mathcal{F})$ that are absolutely continuous w.r.t. $\mathbb{P}$ with Radon-Nikodym derivative in $L^2(\Omega, \mathcal{F}, \mathbb{P})$,

$\mathcal{A}^2$: the set of predictable processes A satisfying $\mathbb{E}\left[\left(\int_0^T A_s \mathrm{d}s\right)^2\right] < \infty$,

$\mathcal{L}^2$: the set of predictable processes H satisfying $\mathbb{E}\left[\int_0^T H_s^2 \mathrm{d}s\right] < \infty$,

$\widetilde{\mathcal{L}}^p$: the set of $\widetilde{\mathcal{P}}$-measurable processes U satisfying $\mathbb{E}\left[\int_0^T \int_{\mathbb{R}} |U(t,x)|^p \Lambda(\mathrm{d}x)\mathrm{d}t\right] < \infty$, $p > 0$.

4.3. **Lévy-Itô processes.** Under an arbitrary probability measure Q that is absolutely continuous with respect to $\mathbb{P}$ the process X is in general not a Lévy process, but will be a Lévy-Itô process. A square-integrable Lévy-Itô process Y on $(\Omega, \mathcal{F}, \mathbb{P})$ is a stochastic process $Y = \{Y(t), t \in [0,T]\}$ of the form

$$Y(t) = \int_0^t A(s)\mathrm{d}s + \int_0^t H(s)\mathrm{d}X_s^c + \int_{[0,t]\times\mathbb{R}} U(s,x)\widetilde{\mu}^X(\mathrm{d}s \times \mathrm{d}x), \tag{4.3}$$

where $A \in \mathcal{A}^2$, $H \in \mathcal{L}^2$ and $U \in \widetilde{\mathcal{L}}^2$, and $\widetilde{\mu}^X = \mu^X - \nu^X$ is the compensated Poisson random measure with compensator $\nu^X(\mathrm{d}x \times \mathrm{d}t) = \Lambda(\mathrm{d}x)\mathrm{d}t$. A Lévy-Itô process Y is called a pure-jump Lévy-Itô subordinator in the case that $U \in \widetilde{\mathcal{L}}^1 \cap \widetilde{\mathcal{L}}^2$ is non-negative and the representation in Eqn. (4.3) can be simplified to

$$Y(t) = \int_{[0,t]\times\mathbb{R}_+} U(s,y)\mu^X(\mathrm{d}s \times \mathrm{d}x).$$

The compensator of the jumps of a Lévy-Itô process Y is equal to the random measure Λ^Y on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ given by $\Lambda_t^Y(\mathrm{d}y) = (\Lambda \circ U_t^{-1})(\mathrm{d}y)$ where $U_t^{-1}(A) = \{z : U(t,z) \in A\}$ for any set $A \in \mathcal{B}(\mathbb{R})$, and will be referred to as the Lévy-Itô measure of Y. The triplet $(A, \sigma^2 H^2, \Lambda^Y)$ will be called the characteristic triplet of Y. Refer to Jacod & Shiryaev [25] for further background concerning Lévy-Itô processes.

By virtue of Kunita & Watanabe [28]'s martingale representation theorem any square-integrable $\mathbf{F}$-martingale is a Lévy-Itô process, that is, any square-integrable $\mathbf{F}$-martingale M on $(\Omega, \mathcal{F}, \mathbb{P})$ admits the representation

$$M_t = \int_0^t H(s)\mathrm{d}X_s^c + \int_{[0,t]\times\mathbb{R}} U(s,x)\widetilde{\mu}(\mathrm{d}s \times \mathrm{d}x), \qquad t \in [0,T], \tag{4.4}$$

with $H \in \mathcal{L}^2$ and $U \in \widetilde{\mathcal{L}}^2$. We will refer to (H,U) as the *representing pair* of the martingale M.

Furthermore, the process X under absolutely continuous probability measures Q has the same law as a Lévy-Itô process. In particular, if the stochastic logarithm M^Q of the Radon-Nikodym derivative of Q with respect to $\mathbb{P}$ is a square-integrable $\mathbf{F}$-martingale M^Q with representing pair $(H^Q, U^Q - 1)$ Girsanov's theorem (e.g. [25, Thms. III.3.24, III.5.19]) states that (X,Q) (X underthe measure Q) has the same law as a Lévy-Itô process with representing triplet $(c^Q, (\sigma^Q)^2, \Lambda^Q)$ under $\mathbb{P}$ given by

$$\begin{cases} c^Q(t) = \mathtt{d}_Q(t) + \displaystyle\int_{\mathbb{R}} (U^Q(t,x) - 1)\, x\, \Lambda(\mathrm{d}x) \qquad \text{with} \qquad \mathtt{d}_Q(t) = \mathtt{d} + H^Q(t)\sigma^2, \\ (\sigma^Q)^2(t) \equiv \sigma^2, \\ \Lambda_t^Q(\mathrm{d}x) = U^Q(t,x)\Lambda(\mathrm{d}x). \end{cases} \tag{4.5}$$

Note that $\mathtt{d}_Q(t)$ is equal to the instantaneous drift at $t \in [0,T]$ of the linear Brownian motion with drift $X_t^c + \mathtt{d}t$ under the measure Q.

4.4. **Expectation under drift and jump-rate distortion.** We next define a real-valued operator on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ that in the next section is shown to arise as the limit of multi-period distorted expectation operators that were defined in Section 3.

Definition 6. A *drift-shift* is a pair of constants $\Delta = (\Delta_+, \Delta_-) \in \mathbb{R}^2_+$ and a *jump-rate distortion* is a pair $\Gamma = (\Gamma_+, \Gamma_-)$ of increasing maps $\Gamma_+, \Gamma_- : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\Gamma_+ - \mathrm{id}$ and $\mathrm{id} - \Gamma_-$ are (measure) distortions, where id denotes the identity map given by $\mathrm{id} : \mathbb{R}_+ \to \mathbb{R}_+$, $\mathrm{id}(x) = x$.

In the sequel we also use the notion of measurable measure capacities, which are defined by extending Definition 4 to measure capacities.

Definition 7. Let $\mathcal{H}, \mathcal{H}'$ be two sigma algebras with $\mathcal{H} \subset \mathcal{H}' \subset \mathcal{B}(\mathbb{R})$.

(i) We call a random set function $\gamma : \mathcal{H}' \to L^0(\mathcal{H})$ an *$\mathcal{H}$-measurable measure capacity* on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \Lambda)$ if γ is absolutely continuous with respect to the measure Λ (cf. Definition 4(i)), and we have $\gamma(\emptyset) = 0$ Λ-a.s., and

$$\gamma(A) \leq \gamma(B) \quad \Lambda\text{-a.s. for any } A, B \in \mathcal{H}' \text{ with } A \subset B \text{ and } \gamma(A) < \infty.$$

(ii) For $\mathcal{H}$-measurable measure capacities γ, γ' on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \Lambda)$ we write $\gamma \prec' \gamma'$ and say that γ' dominates γ Λ-a.s., if we have

$$\gamma(A) \leq \gamma'(A) \;\; \Lambda\text{-a.s., for any } A \in Ḩ' \text{ with } \gamma(A) < \infty\ .$$

Given a drift-shift Δ and a jump-distortion Γ we will consider the collection of probability measures Q under which the drift $\mathtt{d}_Q$ and jump-measure Λ_Q are shifted away from of the drift $\mathtt{d}$ and the Lévy measure Λ of X under P by a certain amount that is expressed in terms of Δ and Γ as follows:

$$\mathcal{D}_{\Delta,\Gamma} = \left\{ Q \in \mathcal{P}^{ac}_{2,\mathbb{P}} : \begin{array}{c} (X,Q) \text{ satisfies for a.e. } (t,\omega) \in [0,T] \times \Omega, \\ \mathtt{d} - \sigma^2 \Delta_- \leq \mathtt{d}_Q(t) \leq \mathtt{d} + \sigma^2 \Delta_+ \\ \Gamma_- \circ \Lambda \prec' \Lambda^Q_t \prec' \Gamma_+ \circ \Lambda \end{array} \right\}.$$

The set $\mathcal{D}_{\Delta,\Gamma}$ is contained in a ball in L^2:

Lemma 5. *We have* $\sup_{t\in[0,T]} \sup_{Q\in\mathcal{D}_{\Delta,\Gamma}} \mathbb{E}[(Z^Q_t)^2] < \infty$.

The proof of Lemma 5 is given at the end of this section. In terms of the set $\mathcal{D}_{\Delta,\Gamma}$ we define a non-linear expectation as follows:

Definition 8. Let Δ be a drift-shift and Γ a jump-rate distortion.

(i) The *expectation under drift and jump-rate distortion* $\mathcal{E}_{\Delta,\Gamma} : L^2(\Omega, \mathcal{F}, P) \to \mathbb{R}$ is given by

$$\mathcal{E}_{\Delta,\Gamma}(\mathcal{X}) = \sup \left\{ \mathbb{E}^Q[\mathcal{X}] : Q \in \mathcal{D}_{\Delta,\Gamma} \right\}.$$

(ii) The *$\mathcal{F}_t$–conditional expectation under drift and jump-rate distortion* $\mathcal{E}_{\Delta,\Gamma}(\,\cdot\,|\mathcal{F}_t) : L^2(\Omega, \mathcal{F}, P) \to L^2(\Omega, \mathcal{F}_t, P)$ is given by

$$\mathcal{E}_{\Delta,\Gamma}(\mathcal{X}|\mathcal{F}_t) = \text{ess.sup} \left\{ \mathbb{E}^Q[\mathcal{X}|\mathcal{F}_t] : Q \in \mathcal{D}_{\Delta,\Gamma} \right\}, \qquad t \in [0,T].$$

Remark 2. It can be verified directly from the definition that the collection $\{\mathcal{E}_{\Delta,\Gamma}(\mathcal{X}|\mathcal{F}_t), t \in [0,T]\}$ is $\mathbf{F}$-consistent, which is a property that also follows from the fact (as shown in Appendix D) that $\mathcal{E}_{\Delta,\Gamma}$ is equal to a g-expectation with driver function $g_{D,G} : \mathbb{R} \times L^2(\Lambda) \to \mathbb{R}$ given by

$$g_{\Delta,\Gamma}(h,u) = h^+\Delta_+\sigma^2 + h^-\Delta_-\sigma^2 + \mathcal{C}^{\Gamma_+-\mathrm{id}}(u^+) + \mathcal{C}^{\mathrm{id}-\Gamma_-}(u^-), \tag{4.6}$$

where, for any $x \in \mathbb{R}$, $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$.

Proof of Lemma 5. Let $Q \in \mathcal{D}_{\Delta,\Gamma}$ and denote by M^Q the square-integrable martingale that is such that its stochastic exponential, denoted by Z^Q, is equal to the Radon-Nikodym derivative of Q w.r.t. $\mathbb{P}$. Denoting by $(H^Q, U^Q - 1)$ the representing pair of M^Q, it follows in view of Girsanov's theorem and the definition of $\mathcal{D}_{\Delta,\Gamma}$ that we have $H_Q \in [\sigma\Delta_-, \sigma\Delta_+]$, $P \times \mathrm{d}t$-a.e. Furthermore, by an analogous reasoning as used in the proof of Lemma 3 it follows

$$\int_{\mathbb{R}} (U_s^Q(x) - 1)^2 I_{\{U_s^Q(x)>1\}} \Lambda(\mathrm{d}x) \leq C_+^2, \qquad \int_{\mathbb{R}} (U_s^Q(x) - 1)^2 I_{\{U_s^Q(x)<1\}} \Lambda(\mathrm{d}x) \leq C_-^2,$$

for $\mathbb{P} \times \mathrm{d}t$-a.e. $(\omega, s) \in \Omega \times [0,T]$, where $C_+ = (K_{\Gamma_+-\mathrm{id}} + \tilde{K}_{\Gamma_+-\mathrm{id}})$ and $C_- = (K_{\mathrm{id}-\Gamma_-} + \tilde{K}_{\mathrm{id}-\Gamma_-})$ (the definitions of these constants were given in Eqns. (2.6)). As a consequence, the Radon-Nikodym derivative with respect to the Lebesgue measure $\mathrm{d}t$ of the angle-bracket process of M^Q is bounded: for $\mathbb{P} \times \mathrm{d}t$-a.e. $(\omega, t) \in \Omega \times [0,T]$, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle M^Q \rangle_t = \left(H_t^Q\right)^2 + \int_{\mathbb{R}} (U_t^Q(x) - 1)^2 \Lambda(\mathrm{d}x) \leq \sigma^2\Delta_-^2 + \sigma^2\Delta_+^2 + C_+^2 + C_-^2.$$

This estimate implies that also $\langle Z^Q \rangle_T$ is bounded: we have for $\mathbb{P} \times \mathrm{d}t$-a.e. $(\omega, t) \in \Omega \times [0,T]$

$$\langle Z^Q \rangle_t = \int_0^t (Z_s^Q)^2 \mathrm{d}\langle M^Q \rangle_s \leq ((\sigma\Delta)^2 + C_+^2 + C_-^2) \int_0^t (Z_s^Q)^2 \mathrm{d}s.$$

By taking expectations and applying Gronwall's lemma, it follows $\mathbb{E}[(Z_t^Q)^2] \leq \exp(t((\sigma\Delta)^2 + C_-^2 + C_+^2))$. Thus, the collection of random variables $\{Z_t^Q, t \in [0,T], Q \in \mathcal{D}_{\Delta,\Gamma}\}$, is contained in a ball in L^2. □

5. Scaling and limit

In this section it is shown that the multi-period distorted expectations converge to the instantaneously distorted expectations if the number of time-steps increases to infinity and the state-space $\mathbb{Z}$ and the transition probabilities are chosen appropriately and the distortion Ψ are scaled in a suitable manner.

5.1. **Random walk setting.** Let $R = (R_n)_{n\in\mathbb{N}}$ be a random walk with $R_0 = 0$ and $R_n = \sum_{i=1}^n Z_i$, with IID increments Z_i that take values in the integers $\mathbb{Z}$. The process R is a time- and space-homogeneous Markov process, with transition probabilities determined by the probability mass function of the increment $Z_1 = R_1 - R_0$

$$P(R_n = y | R_{n-1} = x) = P(Z_1 = x - y) = p_{x-y}, \qquad x, y \in \mathbb{Z}, n \in \mathbb{N},$$

where $(p_n)_{n\in\mathbb{Z}}$ is a probability distribution, that is, $\sum_{n\in\mathbb{Z}} p_n = 1$, $p_n \geq 0$. Denote by $(\mathcal{A}, \mathcal{G}, \mathbf{G}, P)$ a filtered probability space that carries the random walk R where $\mathcal{A} = \mathbb{Z}^{\mathbb{N}} = \{(\omega_i)_{i\in\mathbb{N}} : \omega_i \in \mathbb{Z}\}$ denotes the sample space of paths, $\mathbf{G} = \{\mathcal{G}_n\}_{n\in\mathbb{N}\cup\{0\}}$ the filtration generated by the random walk R, with $\mathcal{G}_0 = \{\emptyset, \Omega\}$ the trivial sigma algebra, and $\mathcal{G}$ is the sigma-algebra given by the power-set 2^{Ω}. The random walk R is defined by the coordinate process $R_n(\alpha) = \alpha_n$ for $n \in \mathbb{N}$ with $R_0 = 0$.

5.2. **Scaled price processes.** For a given a time-step $\delta > 0$ and tick-size $h \in (0,1)$ consider the time-space grid given by $\mathbb{G}^{\delta,h}_{\mathtt{d}} = \{(t_i, x_k) : t_i = i\delta, x_k = \mathtt{d}t_i + hk, i \in \mathbb{N}_*, k \in \mathbb{Z}\}$, where $\mathtt{d} = \mathbb{E}[X_t]/t$ denote the mean of the Lévy process X.

For a non-zero value of $\mathtt{d}$ the grid $\mathbb{G}^{\delta,h}_{\mathtt{d}}$ is obtained from the normalised grid $\mathbb{G} := \mathbb{G}^{1,1}_0 = \mathbb{N}_* \times \mathbb{Z}$ by changing the slope of any horizontal line to $\mathtt{d}$ and changing the horizontal and vertical stepsizes to δ and h respectively. The problem to find a random walk $R^{\delta,h}$ approximating X on the grid $\mathbb{G}^{\delta,h}_{\mathtt{d}}$ is equivalent to the problem to construct a random walk R approximating the martingale $h^{-1}(X_{t\delta} - \mathtt{d}t\delta)$ on the standard grid $\mathbb{G}$. In fact, $R^{\delta,h}$ can be obtained from R via the transformation $R^{\delta,h}_t = hR_{t/\delta} + \mathtt{d}t$, $t \in \delta\mathbb{N}_*$. When the choice of h is clear from the context, we will supress h and write R^δ for $R^{\delta,h}$.

We next turn to the definition of the approximating random walk $R = \{R_n, n \in \mathbb{N}_*\}$, by specifying how the probability distribution of an increment of R is obtained from the characteristic triplet of X. The probability of a "large" jump of the random walk of size $J \in \mathbb{Z}$ is set to be equal to the Lévy measure Λ integrated over a neightbourhood of $J \cdot h$ while its transition probabilities of "small" jumps is chosen such that the mean and variance of an increment of the random walk R matches those of an increment of $h\,X$ over a time interval of length δ. Here a jump sizes is called large if its absolute value is larger than a cut-off value $a > 0$. The common probability distribution $(p_k)_{k\in\mathbb{Z}}$ of the step-sizes Z_i of R under the probability measure P is given in terms of the characteristics of X by

$$p_k = P(Z_i = k) = \begin{cases} \delta\,\Lambda([kh,(k+1)h)) & \text{for } k \in \mathbb{Z}, k \geq a, \\ \delta\,\Lambda(((k-1)h), kh]) & \text{for } k \in \mathbb{Z}, k \leq -a, \end{cases} \tag{5.1}$$

for some integer $a \in \mathbb{N}$ with $a \geq 2$, where the probabilities p_k, $k = -a+1, \ldots, a-1$, are specified so as to match the mean and variance of X_δ, i.e.

$$h \cdot E[Z_i] = E[X_\delta] - \delta\,\mathtt{d} = 0, \tag{5.2}$$

$$h^2 \cdot \mathrm{Var}[Z_i] = \mathrm{Var}[X_\delta] = \delta\,(\sigma^2 + \Sigma^2(\mathbb{R})), \tag{5.3}$$

where $E[Z_i]$ and $\mathrm{Var}[Z_i]$ denote the expectation and the variance of the random variable Z_i under the probability measure P, and where we denote $\Sigma^2(I) = \int_I x^2\Lambda(\mathrm{d}x)$ for any interval I. Under suitable conditions on the step-size h and time-step δ and the integer a existence can be established of a probability distribution satisfying Eqns. (5.1)–(5.2)–(5.3).

Lemma 6. *(i) For any triplet (h, δ, a) of positive real numbers satisfying the conditions*

$$h^2 = 3\,\delta\,\widetilde{\Sigma}^2(a) \qquad \widetilde{\Sigma}^2(a) := \begin{cases} \sigma^2 + a^{-1}\,\Sigma^2(\mathbb{R}) & \textit{in the case } \sigma > 0, \\ \Sigma^2(\mathbb{R}) & \textit{in the case } \sigma = 0, \end{cases} \tag{5.4}$$

$$a^{-1}\Sigma^2(\mathbb{R}) - \sigma^2 \leq \Sigma^2((-ah, ah)) \leq 2\sigma^2 + \Sigma^2(\mathbb{R})(I_{\{\sigma^2=0\}}3(1-a^{-1}) + a^{-1} - 2a^{-2}), \tag{5.5}$$

there exists a probability distribution $(p_k)_{k\in\mathbb{Z}}$ that solves the system in Eqns. (5.1), (5.2) *and* (5.3) *and satisfies*

$$p_k = 0 \textit{ for all } k \in \mathbb{N} \textit{ with } |k| \geq 2 \textit{ and } |k| \leq a-1. \tag{5.6}$$

(ii) Let $(p_k(h))_{k\in\mathbb{Z}}$ satisfy Eqns. (5.1), (5.2), (5.3) *and* (5.6). *Let (h,δ,a) be a collection of triplets satisfying Eqns.* (5.4)–(5.5) *such that $a\cdot h\leq 1$ and, when $h\searrow 0$, $a\to\infty$ and $a\cdot h\to 0$. When $h\searrow 0$, we have*

$$\begin{cases} p_{-1}(h)\to p_{-1}^{BM}:=\frac{1}{6}, & p_0(h)\to p_0^{BM}:=\frac{2}{3}, \quad p_1(h)\to p_1^{BM}:=\frac{1}{6}, \quad \textit{if } \sigma^2>0,\\ p_{-1}(h)\to 0, & p_0(h)\to 1, \quad p_1(h)\to 0, \quad \textit{if } \sigma^2=0. \end{cases} \tag{5.7}$$

(iii) If, in addition to the assumptions in (ii), the triplets (h,δ,a) satisfy $a^2\cdot h\to\infty$, then it holds

$$p_{-1}(h)\wedge p_1(h)\geq I_{\{\sigma^2>0\}}\left[\frac{1}{6}-\frac{1}{2}c^*_{a,h}\sqrt{h}\right]\frac{\sigma^2}{\widetilde{\Sigma}^2(a)}, \tag{5.8}$$

$$p_{-1}(h)+p_0(h)+p_1(h)\geq 1-c^{**}_{a,h}h, \tag{5.9}$$

when $h\searrow 0$, where $c^*_{a,h}=(a^2h)^{-1/2}\cdot c_\sigma$ and $c^{**}_{a,h}=(a^2h)^{-1}\cdot c_\sigma$ with

$$c_\sigma = I_{\{\sigma^2=0\}}\cdot\frac{1}{3}+I_{\{\sigma^2>0\}}\cdot\frac{\Sigma^2(\mathbb{R})}{3\sigma^2}.$$

It follows from Lemma 6 that, for each given $\delta>0$, there exists a random walk on some grid that matches the first two moments of increments of X over a time-step. From now on it is assumed that for each given $\delta>0$ the transition probabilities of a random walk have been fixed as in Lemma 6. We will also consider two related stochastic processes, namely the skip-free random walk $R^{(c)}=\{R^{(c)}_n\}_{n\in\mathbb{N}}$ that starts at zero, and has increments $Z_i\mathbf{1}_{\{Z_i=\pm1\}}$, and the compensated counting process $\overline{N}^A$ which is for any Borel set $A\subset\mathbb{R}\backslash$ given by

$$\overline{N}^A_0=0, \qquad \overline{N}^A_n=\sum_{k=1}^n\sum_{i:i\in A}[I^k_i-p_i], \qquad n\in\mathbb{N}, \tag{5.10}$$

where $I^k_i=I_{\{Z_k=i\}}$. For each $\delta>0$ continuous time stochastic processes $Y^\delta=\{Y^\delta_t,t\in[0,T]\}$, $Y^{\delta(c)}=\{Y^{\delta(c)}_t,t\in[0,T]\}$ and $Z^{\delta,A}=\{Z^{\delta,A}_t,t\in[0,T]\}$ on Ω can be constructed from the discrete time processes by piecewise constant extension of the paths, defining $Y^\delta_t=hR_{\lfloor t/\delta\rfloor}+\mathtt{d}t$, $Y^{\delta(c)}_t=hR^{(c)}_{\lfloor t/\delta\rfloor}$ and $Z^{\delta,A}_t=\overline{N}^A_{\lfloor t/\delta\rfloor}$ for any $t\geq 0$, where $\lfloor s\rfloor$ is the largest integer that is smaller than $s\in\mathbb{R}$. Standard convergence arguments, which are presented in Section 8, yield weak convergence as the time-step δ tends to zero.

Lemma 7. *For any positive sequence $(\delta_n)_n$ with $\delta_n\to 0$ and any set $A\subset\mathbb{R}\backslash\{0\}$ with boundary $\partial A=\overline{A}\backslash A^o$ satisfying $\Lambda(\partial A)=0$ we have*

$$Y^{\delta_n}\Rightarrow X, \qquad Y^{\delta_n(c)}\Rightarrow X^{(c)}, \qquad Z^{\delta,A}\Rightarrow\tilde{\mu}(A\times\cdot),$$

where $\Rightarrow$ denotes weak convergence in the Skorokhod J_1-topology on Ω.

5.3. **Scaled distortions.** To ensure that the discrete time valuations converge to a non-degenerate limit the distortions need to be scaled suitably:

Definition 9. For given maps $\xi:[0,1]\to\mathbb{R}_+$ and Γ_+, and $\Gamma_-:\mathbb{R}_+\to\mathbb{R}$ such that (Γ_+,Γ_-) is a jump-rate distortion, the (ξ,Γ_+,Γ_-)-*scaling family of distortions* $\{\Psi(\cdot,\delta),\delta>0\}$ is defined as follows:

(i) For any $\delta>0$, the map $\Psi(\cdot,\delta)$ is a continuous concave probability distortion.

(ii) With $\sigma^*=\sigma/(2\sqrt{3})$, we have the uniform limit

$$\lim_{\delta\searrow 0}\sup_{p\in[0,1]}\left(\frac{\Psi(p,\delta)-p}{\sqrt{\delta}}-\xi(p)\sigma^*\right)=0, \tag{5.11}$$

(iii) With $\widehat{\Psi}(p,\delta) = 1 - \Psi(1-p,\delta)$, it holds

$$\lim_{\delta\searrow 0} \sup_{0<\lambda\leq 1/\delta} \left\{ \frac{\frac{\Psi(\delta\lambda,\delta)}{\delta} - \Gamma_+(\lambda)}{\Gamma_+(\lambda)-\lambda} \right\} = 0, \qquad \lim_{\delta\searrow 0} \sup_{0<\lambda\leq 1/\delta} \left\{ \frac{\frac{\widehat{\Psi}(\delta\lambda,\delta)}{\delta} - \Gamma_-(\lambda)}{\lambda-\Gamma_-(\lambda)} \right\} = 0, \tag{5.12}$$

where the denominators are taken equal to 1 in the cases $\Gamma_+ = \mathrm{id}$ or $\Gamma_- = \mathrm{id}$.

Remark 3. Since the concavity and monotonicity of the distortions $p \mapsto \Psi(p,\delta)$ are preserved under the point-wise limit of δ tending to zero, it follows that ξ and Γ_+ are concave, while Γ_- is convex, and Γ_+ and Γ_- are increasing. Moreover, the fact $\Psi \geq \mathrm{id}$ yields $\Gamma_+ \geq \mathrm{id} \geq \Gamma_-$.

Example 1. Let Ψ_1, Ψ_2, Ψ_3 be arbitrary concave probability distortions that are such that the right-derivatives $\Psi_1'(0+)$ and $\Psi_3'(0+)$ at zero are finite, while we have $\Psi_3'(0+) < \Psi_2'(0+) \in (1,\infty]$. let Ψ_0 denote the linear distortion, $\Psi_0(p) = p$ for $p \in [0,1]$. For some $\delta_0 > 0$ to be specified shortly, consider the collection $\{\Psi(\cdot, \delta \wedge \delta_0), \delta > 0\}$ given by

$$\Psi(p,\delta) = C_0(\delta)\Psi_0(p) + \sqrt{\delta}\{\Psi_1(p) - p\}\sigma + \delta\Psi_2\left(1 - \mathrm{e}^{-p/\delta}\right) + \delta\left\{\Psi_3\left(1-\mathrm{e}^{-(1-p)/\delta}\right) - \Psi_3\left(1-\mathrm{e}^{-1/\delta}\right)\right\},$$

for $\delta \in (0,\delta_0)$, where $C_0(\delta) = 1 - \delta\Psi_2\left(1-\mathrm{e}^{-1/\delta}\right) + \delta\Psi_3\left(1-\mathrm{e}^{-1/\delta}\right)$ and where $\delta_0 \in (0,1)$ is chosen sufficiently small to guarantee that $C_0(\delta)$ is positive for all $\delta \in (0,\delta_0)$.

It is straightforward to verify that, for any $\delta \in (0,\delta_0)$, the function $p \mapsto \Psi(p,\delta)$ is a concave probability distortion, and that, moreover, the family $\{\Psi(\cdot,\delta\wedge\delta_0), \delta > 0\}$ is a $(\xi, \Gamma_+, \Gamma_-)$-distortion scaling family with the functions Γ_+ and Γ_- given by

$$\xi(p) = 2\sqrt{3}(\Psi_1(p) - p), \quad \Gamma_+(\lambda) = \lambda + \Psi_2\left(1-\mathrm{e}^{-\lambda}\right), \quad \Gamma_-(\lambda) = \lambda - \Psi_3\left(1-\mathrm{e}^{-\lambda}\right). \tag{5.13}$$

In the absence of jumps (that is, in the case $\Lambda = 0$), the forms of Γ_+ and Γ_- will not affect the limit. Thus, in that case, one arrives at the same scaling limit by simply taking $\Gamma_+ = \Gamma_- = \mathrm{id}$ and replacing the family $\{\Psi(\cdot,\delta\wedge\delta_0), \delta > 0\}$ by the family $\{\widetilde{\Psi}(\cdot,\delta), \delta > 0\}$ given by $\widetilde{\Psi}(p,\delta) = p + \sigma\sqrt{\delta}(\Psi_1(p) - p)$. This choice is in agreement with the square-root scaling identified in Stadje [42] in the study of convergence of approximations of BSDE by BSΔEs in a Brownian setting.

5.4. **Convergence theorem.** In the setting described earlier in this section we will show the convergence of the multi-period distorted expectations of a class of path-functionals of the random walk $R^{\delta,h}$, as the time-step δ and spatial mesh size h converge to zero, to the instantaneously distorted expectation of the corresponding path-functional of X.

The class in question is given by the collection of functionals $F : \Omega \to \mathbb{R}$ that are continous in the Skorokhod J_1-topology and satisfy the bound

$$|F(x)| \leq C\exp(q\|x\|_\infty) \qquad \text{for all } x \in \Omega \tag{5.14}$$

with $\|x\|_\infty = \sup_{0\leq s\leq T}|x_s|$ and some $C > 0$, where $q > 0$ was given in Eqn. (4.1). The corresponding path-functionals F^δ of the random walk R^δ are obtained by embedding paths from the space $(h\mathbb{Z})^n$ into Ω by piecewise constant extension, as follows:

$$F_\delta : (\omega_1, \ldots, \omega_n) \mapsto F(\overline{\omega}) \qquad \overline{\omega}_t = \omega_{\lfloor t/\delta \rfloor}, \qquad \delta = T/n, \qquad t \in [0,T].$$

The bound in Eqn. (5.14) suffices to guarantee square-integrability of the random variables $F(X)$ and $F_\delta(R^\delta)$:

Lemma 8. *Assume $F : \Omega \to \mathbb{R}$ is continuous, that the conditions in Eqn.* (5.14) *are satisfied and that X satisfies* (4.1). *Then it holds*

$$\sup_{\delta>0} \mathbb{E}[\exp(2q\|Y^\delta\|_\infty)] < \infty.$$

Proof of Lemma 8. Denote by X' the Lévy process $X'_t = X_t + \mathtt{d}^- t$ where $\mathtt{d}^- = \max\{-\mathtt{d}, 0\}$ and note that X' is a sub-martingale. In view of the fact that $\exp(qX')$ are sub-martingales, Doob's maximal inequality implies that there exists a constant C such that we have

$$\mathbb{E}[\exp(2q\|X'\|_\infty)] \leq C\mathbb{E}[\exp(2q|X'_T|)],$$

which is finite by the assumption in Eqn. (4.1). Since we have $|\|X\| - \|X'\|| \leq \mathtt{d}^- T$, the assertion follows.

The second bound follows by an analogous argument applied to the embedded random walk Y^δ. □

Denoting by $\mathbf{G}_\delta$ the filtration generated by the scaled random walk R^δ, the announced result is phrased as follows:

Theorem 1. *Let $\{\Psi(p,\delta), \delta > 0\}$ be a $(\xi, \Gamma_+, \Gamma_-)$-scaling family of distortions and let $F : \Omega \to \mathbb{R}$ be a continuous map satisfying the bound in Eqn.* (5.14). *For any sequence $\delta_n \searrow 0$, we have*

$$\mathcal{C}^{\Psi_{\delta_n}, \mathbf{G}_{\delta_n}}\left(F_{\delta_n}\left(R^{\delta_n}\right)\right) \longrightarrow \mathcal{E}_{\Delta,\Gamma}(F(X)), \qquad \text{as } n \to \infty,$$

with $\Psi_{\delta_n} = \Psi(\cdot, \delta_n)$ and $\Gamma = (\Gamma_+, \Gamma_-)$, and with $\Delta = (\Delta_+, \Delta_-)$ given by $\Delta_+ = \xi(\frac{1}{6})$ and $\Delta_- = \xi(\frac{5}{6})$.

The proof of Theorem 1 is given in Sections 8–10

6. Valuation of pathwise increasing claims

While the map $\mathcal{E}_{\Delta,\Gamma} : L^2(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is not additive, it will be shown below that, when restricted to the set of random variables that are a pathwise increasing function of the Lévy process X, the map $\mathcal{E}_{\Delta,\Gamma}$ is additive. Similarly, it holds that the restriction of the map $\mathcal{E}_{\Delta,\Gamma}$ to the set of random variables that are pathwise decreasing is additive. As the derivations in the two cases are similar we will restrict ourselves below to the case of pathwise increasing functionals.

Definition 10. The random variable $\mathcal{X} \in L^0(\Omega, \mathcal{F})$ is *pathwise increasing* if $\mathcal{X}(\omega) \geq \mathcal{X}(\omega')$ whenever $\omega(t) \geq \omega'(t)$ for all $t \in [0,T]$.

Assume a drift-shift $\Delta = (\Delta_+, \Delta_-)$ and a jump-distortion $\Gamma = (\Gamma_+, \Gamma_-)$ have been fixed. Denote by $Q^\#$ the probability measure on $(\Omega, \mathcal{F})$ under which X is a Lévy process with characteristic triplet given by $(\gamma_\#, \sigma^2_\#, \Lambda_\#)$, where

$$\gamma_\# = \mathtt{d} + \Delta_+\sigma + \int_{\mathbb{R}} (Z^\#(x) - 1)x\Lambda(\mathrm{d}x), \qquad \sigma^2_\# = \sigma^2, \tag{6.1}$$

with $Z^\#$ denoting the Radon-Nikodym derivative of the Lévy measure $\Lambda_\#$ with respect to Λ. The tail-functions $\overline{\Lambda}_\#$ and $\underline{\Lambda}_\#$ of $\Lambda^\#$, given by $\overline{\Lambda}^\#(x) = \Lambda^\#((x,\infty))$ and $\underline{\Lambda}^\#(-x) = \Lambda^\#((-\infty,-x))$ for $x > 0$, are expressed in terms of the tail functions $\overline{\Lambda}$ and $\underline{\Lambda}$ of Λ by

$$\overline{\Lambda}_\#(x) = \Gamma_+\left(\overline{\Lambda}(x)\right), \qquad \underline{\Lambda}_\#(-x) = \Gamma_-\left(\underline{\Lambda}(-x)\right), \qquad x > 0. \tag{6.2}$$

Lemma 9. *The measure $Q_\#$ is element of the set $\mathcal{D}_{D,G}$, and we have*

$$\max_{Q\in\mathcal{D}_{\Delta,\Gamma}} \overline{\Lambda}^Q(x) = \Gamma_+\left(\overline{\Lambda}(x)\right), \qquad \Gamma_-\left(\underline{\Lambda}(-x)\right) = \min_{Q\in\mathcal{D}_{\Delta,\Gamma}} \underline{\Lambda}^Q(-x), \quad x>0. \tag{6.3}$$

Proof. We show in three steps that the measure $Q_\#$ satisfies the conditions stated in the definition of the set $\mathcal{D}_{\Delta,\Gamma}$. Firstly, we note that the Radon-Nikodym derivative $Z_\#$ is square integrable, in view of Lemma 3. Secondly, we observe that Girsanov's theorem implies $\mathtt{d}_{Q_\#}(t) = \mathtt{d} + \Delta_+\sigma$. Finally, reasoning as in the proof of Lemma 18(ii) it follows

$$\Gamma_-(\Lambda(A)) \leq \int_A Z_\#(y)\Lambda(\mathrm{d}y) \leq \Gamma_+(\Lambda(A)), \qquad A\in\mathcal{B}^\Lambda(\mathbb{R}).$$

Hence, we deduce $Q_\#\in\mathcal{D}_{D,G}$, and it follows that the two identities in Eqn. (6.3) both hold with inequality ($\geq$) instead of equality. That we have in fact equalities in Eqn. (6.3) follows from the explicit forms of $\overline{\Lambda}_\#(x)$ and $\underline{\Lambda}_\#(-x)$ in Eqn. (6.2). ☐

Theorem 2. *If $\Upsilon:\Omega\to\mathbb{R}$ is $\mathcal{F}$-measurable and pathwise increasing and $\Upsilon(X)\in L^2(\Omega,\mathcal{F},\mathbb{P})$, then it holds*

$$\mathcal{E}_{\Delta,\Gamma}(\Upsilon(X)|\mathcal{F}_t) = \mathbb{E}^{Q^\#}[\Upsilon(X)|\mathcal{F}_t], \qquad t\in[0,T]. \tag{6.4}$$

Theorem 2 implies that for any increasing $\mathcal{B}(\mathbb{R})$-measurable map $H:\mathbb{R}\to\mathbb{R}$ satisfying $H(X_T)\in L^2(\Omega,\mathcal{F},\mathbb{P})$ we have

$$\mathcal{E}_{\Delta,\Gamma}(H(X_T)|\mathcal{F}_t) = \mathbb{E}^{Q^\#}[H(X_T)|\mathcal{F}_t], \qquad t\in[0,T].$$

6.1. **Proof by coupling.** The proof is based on the following auxiliary coupling result:

Lemma 10. *For any $Q\in\mathcal{D}_{D,G}$ there exists a probability space $(\Omega^{(Q)},\mathcal{F}^{(Q)},\mathbb{P}^{(Q)})$ supporting a stochastic process $(Y^Q,Y^\#)=\{(Y^Q_t,Y^\#_t),t\in[0,T]\}$ that satisfies*

$$(Y^Q,\mathbb{P}^{(Q)}) \stackrel{\mathcal{L}}{=} (X,Q), \qquad (Y^\#,\mathbb{P}^{(Q)}) \stackrel{\mathcal{L}}{=} (X,Q^\#), \qquad \text{and}$$
$$Y^Q(t,\omega)\leq Y^\#(t,\omega) \qquad \text{for } \mathbb{P}^{(Q)}\text{-a.e. } \omega\in\Omega^{(Q)} \text{ and all } t\in[0,T],$$

where $\stackrel{\mathcal{L}}{=}$ denotes equality in law.

The proof of Lemma 10 in turn relies on a coupling of Lévy-Itô subordinators.

Lemma 11. *Let $Y^1=(Y^1_t)_{t\in[0,T]}$ and $Y^2=(Y^2_t)_{t\in[0,T]}$ be two pure jump Lévy-Itô subordinators with Lévy-Itô measures ν^1 and ν^2 satisfying the domination condition*

$$\overline{\nu}^1_{t,\omega}(x)\geq\overline{\nu}^2_{t,\omega}(x) \qquad \text{for all } x>0, \text{ and all } t\in[0,T], \text{ and } \mathbb{P}\text{-a.e. } \omega\in\Omega. \tag{6.5}$$

Then there exists a probability space $(\Omega^*,\mathcal{F}^*,\mathbb{P}^*)$ that supports a stochastic process $(Z^1,Z^2)=\{(Z^1_t,Z^2_t),t\in[0,T]\}$ satisfying

$$Z^1\stackrel{\mathcal{L}}{=}Y^1, \qquad Z^2\stackrel{\mathcal{L}}{=}Y^2, \qquad Z^1(t,\omega)\geq Z^2(t,\omega) \qquad \text{for } \mathbb{P}^*\text{-a.e. } \omega\in\Omega^* \text{ and all } t\times[0,T].$$

Proof of Theorem 2. In view of Girsanov's theorem and the definition of the non-linear expectation $\mathcal{E}_{D,G}$ it follows

$$\mathcal{E}_{D,G}(\Upsilon(X)|\mathcal{F}_t) = \operatorname*{ess.sup}_{Q\in\mathcal{D}_{D,G}} \mathbb{E}^Q[\Upsilon(X)|\mathcal{F}_t], \qquad t\in[0,T]. \tag{6.6}$$

Lemma 10 and the fact that Υ is pathwise increasing imply $\mathbb{E}^Q[\Upsilon(X)|\mathcal{F}_t] \le \mathbb{E}^{Q_\#}[\Upsilon(X)|\mathcal{F}_t]$ for any $t \in [0,T]$. The statement hence follows from the fact that $Q_\#$ is element of $\mathcal{D}_{D,G}$. □

We next turn to the proofs of the Lemmas 10 and 11.

Proof of Lemma 11. We present an explicit contruction of processes Z^1 and Z^2 with the required properties first assuming that ν^1 and ν^2 have finite mass. Let $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ denote a probability space that supports the processes Y^1 and Y^2, a counting process N and a collection of IID $U(0,1)$ random variables $\mathbb{U} = \{U_i, i \in \mathbb{N}\}$. Denoting by $\mathcal{H}$ the filtration generated by the processes Y^1 and Y^2, assume that, conditional on the curve $C = \{C_t, t \in [0,T]\}$ with $C_t = \nu_t^1(\mathbb{R}_+) \vee \nu_t^2(\mathbb{R}_+)$, the process N is a time-inhomogeneous Poisson process with time-dependent rates C_t that is independent of $\mathcal{H}$. Also assume that $\mathbb{U}$ is independent of N and of $\mathcal{H}$. Define the processes $Z^i = \{Z_t^i, t \in [0,T]\}$ for $i = 1, 2$ by

$$Z^i(t) = \sum_{j=1}^{N_t} \left(F_t^{(i)}\right)^{-1}(U_j), \qquad t \in [0,T], \tag{6.7}$$

where $\left(F_t^{(i)}\right)^{-1}$, $i = 1, 2$, denote the right-inverses of the maps $F_t^{(i)} : \mathbb{R}_+ \to [0,1]$ given by

$$F_t^{(i)}(x) = \begin{cases} c_t^i / C_t, & x = 0, \\ [\nu_t^i((0,x]) + c_t^i]/C_t, & x > 0, \end{cases} \tag{6.8}$$

where $c_t^i = C_t - \nu_t^i(\mathbb{R}_+)$. Since any jump size $(F_t^{(i)})^{-1}(U_j)$ is non-negative and, conditional on $\mathcal{H}$, follows the distribution $F_t^{(i)}$, it follows that Z^1 and Z^2 are Lévy-Itô subordinators with Lévy-Itô measures given by $I_{(0,\infty)}(x) C_t F_t^{(i)}(\mathrm{d}x) = \nu_t^i(\mathrm{d}x)$, $i = 1, 2$. Furthermore, in view of the implications

$$\left\{\forall x > 0\ \overline{\nu}_t^1(x) \ge \overline{\nu}_t^2(x)\right\} \Leftrightarrow \left\{\forall x > 0\ F_t^{(1)}(x) \le F_t^{(2)}(x)\right\} \Leftrightarrow \left\{\forall x > 0\ \left(F_t^{(1)}\right)^{-1}(x) \ge \left(F_t^{(2)}\right)^{-1}(x)\right\}$$

for any $t \in [0,T]$, it follows from Eqn. (6.7) that we have $Z^1 \ge Z^2$.

Next we remove the assumption of boundedness. Let ν^1, ν^2 be as stated and fix $\epsilon > 0$ arbitrary. Applying the first part of the proof to the truncated measures $\nu_\epsilon^i(\mathrm{d}x) = I_{(\epsilon,\infty)}(x)\nu^i(\mathrm{d}x)$ for $i = 1, 2$, shows

$$Z^{1,\epsilon}(t,\omega) \ge Z^{2,\epsilon}(t,\omega) \text{ for any } \epsilon > 0,\ t \in [0,T],\ \omega \in \Omega^*. \tag{6.9}$$

The sequence $(Z^{1,\epsilon}, Z^{2,\epsilon})_\epsilon$ of two-dimensional Lévy-Ito processes converges weakly in the Skorokhod topology as $\epsilon \searrow 0$ (see [25]). The limit (Z_*^1, Z_*^2) is a Lévy-Itô process of which the components Z_*^i are Lévy-Itô subordinators with Lévy-Itô measures ν^i, $i = 1, 2$. In particular, passing to a suitable subsequence (ϵ'), it follows $Z^{i,\epsilon}(t) \to Z_*^i(t)$ $\mathbb{P}^*$-a.s., for any $t \in [0,T]$. Taking in Eqn. (6.9) the limit of $\epsilon' \searrow 0$ along the sequence (ϵ') shows $Z_*^1(t,\omega) \ge Z_*^2(t,\omega)$ for $\mathbb{P}^*$-a.e. $\omega \in \Omega^*$ and any $t \in [0,T]$. □

Proof of Lemma 10. According to Girsanov's theorem, the process X under the measure $Q \in \mathcal{D}_{D,G}$ has the same law as a Lévy-Itô process X^Q with characteristics given in Eqn. (4.5). Denote by $(\Omega^{(Q)}, \mathcal{F}^{(Q)}, \mathbb{P}^{(Q)})$ a probability space that supports a Wiener process W and a random measure μ^Q with compensator given by $\nu_t^Q(\mathrm{d}x) = (U^Q(t,x) \vee 1)\nu_t(\mathrm{d}x)$. Let $J^Q = \{J_t^Q, t \in [0,T]\}$ be the compensated jump-process given by $J_t^Q = \int_{[0,t]\times\mathbb{R}} x(\mu^Q(\mathrm{d}s,\mathrm{d}x) - \nu_s^Q(\mathrm{d}x)\mathrm{d}s)$. From the random measure μ^Q can be constructed by thinning the compensated jump processes A^Q and B^Q with compensators $\nu_t(\mathrm{d}x)\mathrm{d}t$ and $U^Q(t,x)\nu_t(\mathrm{d}x)\mathrm{d}t$ respectively, and the Lévy-Itô subordinators $Y^{Q,+}, Z^{Q,+}$ and negatives of Lévy-Itô subordinators $Y^{Q,-}, Z^{Q,-}$ with compensators $\overline{Y}^{Q,\pm}, \overline{Z}^{Q,\pm}$ given by

$$\overline{Y}_t^{Q,\pm} = \int_{[0,t]\times\mathbb{R}_\pm} (U^Q(s,y)-1)^+\Lambda(\mathrm{d}y)\mathrm{d}s \quad \text{and} \quad \overline{Z}_t^{Q,\pm} = \int_{[0,t]\times\mathbb{R}_\pm} (U^Q(s,z)-1)^-\Lambda(\mathrm{d}z)\mathrm{d}s.$$

Observe that J^Q admits the decompositions

$$J^Q = B^Q + Z^{Q,+} + Z^{Q,-} - \overline{Z}^{Q,+} - \overline{Z}^{Q,-} \tag{6.10}$$

$$= A^Q + Y^{Q,+} + Y^{Q,-} - \overline{Y}^{Q,+} - \overline{Y}^{Q,-}. \tag{6.11}$$

Hence, by comparing the characteristics of X to those of X^Q given in Eqn. (4.5) and by combining this with Eqns. (6.10)—(6.11), it follows that the process X^Q can be constructed on $(\Omega^{(Q)}, \mathcal{F}^{(Q)}, \mathbb{P}^{(Q)})$ by

$$X^Q = X' + D^Q + Y^{Q,+} + Y^{Q,-} - (Z^{Q,+} + Z^{Q,-}), \tag{6.12}$$

where X' has the same law as X and $D^Q = \{D_t^Q, t \in [0,T]\}$ is given by $D_t^Q = \int_0^t H_s^Q \sigma^2 \mathrm{d}s$. The definition of the set $\mathcal{D}_{D,G}$ implies $D_t^Q \leq \Delta_+\sigma^2 t$. Moreover, in combination with Lemma 11 this definition also yields the inequalities

$$Y^{Q,+} \leq Y^{Q_\#,+}, \qquad Y^{Q,-} \leq 0 = Y^{Q_\#,-}, \qquad Z^{Q,+} \geq 0 = Z^{Q_\#,+} \qquad \text{and} \qquad Z^{Q,-} \geq Z^{Q_\#,-}.$$

Thus, from Eqn. (6.12) it follows $X^Q \leq X^{Q_\#}$ for each $Q \in \mathcal{D}_{D,G}$. Since $(X^{Q_\#}, \mathbb{P}^{(Q_\#)})$ has the same law as $(X, Q_\#)$, the proof is complete. □

7. Examples

7.1. **Geometric Brownian motion.** Under Samuelson's model the price process of a risky stock $S = \{S_t, t \in [0,T]\}$ is modelled by a geometric Brownian motion given by

$$S_t = S_0 \exp(X_t), \qquad t \in [0,T], \tag{7.1}$$

where the log-price process X is a linear Brownian motion given by $X_t = ct + \sigma B_t$ where $c = \mu - \frac{\sigma^2}{2}$, B is a standard Brownian motion and $\mu, \sigma \in \mathbb{R}$ denote the drift and volatility of S.

We consider the distortion valuation with respect to the family of distortions $\{\Psi(\cdot, \delta\wedge 1), \delta > 0\}$ given by

$$\Psi(p,\delta) = p + \sqrt{\delta}\left(\Psi_\alpha(p) - p\right), \qquad \Psi_\alpha(p) = \frac{1-\mathrm{e}^{-\alpha p}}{1-\mathrm{e}^{-\alpha}}, \qquad p \in [0,1], \alpha, \delta > 0.$$

Note that the extreme cases of $\delta = 0$ and $\delta = 1$ correspond to a linear distortion $\Psi_0(p) = p$ and the exponential distortion Ψ_α, respectively. It is straightforward that this collection is a $(\xi_\alpha, \Gamma_-, \Gamma_+)$-scaling family of distortions with $\Gamma_- = \Gamma_+ = 0$ and ξ given by $\xi_\alpha(p) = \Psi_\alpha(p) - p$. We refer to Cherny & Madan [10] and Wang [45] for different examples of distortions which can be deployed analogously to construct scaling families of distortions.

Let $R = (R_n, n \in \mathbb{N}\cup\{0\})$ denote the trinomial random walk given by $R_0 = 0$ and $R_n = \sum_{k=1}^n U_k$ with $U_k \in \{\pm 1, 0\}$ and transition probabilities given by

$$p_1 = p_{-1} = \frac{1}{6}, \qquad p_0 = \frac{2}{3}. \tag{7.2}$$

Then the sequence of scaled processes $Y^\delta = \{Y_t^\delta, t \in [0,T]\}$ defined by $Y_t^\delta = ct + hR_{\lceil t/\delta\rceil}$, where δ and h are linked via the classical relation $3\delta = \sigma^2 h^2$, converges weakly to X in the Skorokhod topology as $\delta \to 0$.

The multi-period distortion expectation values of a given claim $\mathcal{X} \in L^2(\mathcal{F}_T)$ satisfy the recursion

$$\Pi_m^\delta(\mathcal{X}) = \alpha_1 P_1^{m+1} + \alpha_0 P_0^{m+1} + \alpha_{-1} P_{-1}^{m+1}, \qquad m = 0, \ldots, n-1,$$

with $\Pi_n(\mathcal{X}) = \mathcal{X}$ and $\delta = T/n$, where we denoted $P_i^{m+1} = \Pi_{m+1}(\mathcal{X}|\mathcal{G}_m) I_{\{\Delta R_m = ih\}}$, $i \in \{-1, 0, 1\}$. If $\mathcal{X}$ is a claim of the form as stated in Theorem 1 then according to Theorem 1 we have $\Pi_0^\delta(\mathcal{X}) \to \mathcal{E}_{\Delta,\Gamma}(\mathcal{X})$ as $\delta \searrow 0$.

In the case that $\mathcal{X}$ is pathwise increasing, Theorem 2 states that $\mathcal{E}_{\Delta,\Gamma}(\mathcal{X})$ is equal to the (classical linear) expectation $\mathbb{E}^{Q_\#}[\mathcal{X}]$ under the measure $Q_\#$ under which the process X is given by

$$X_t = \sigma B_t^\# + \left(c_\# - \frac{\sigma^2}{2}\right) t, \qquad c_\# = \Delta_+ \sigma^2 + \mu, \qquad \Delta_+ = \xi_\gamma\left(\frac{5}{6}\right),$$

where $B^\#$ is a standard Brownian motion under the measure $Q_\#$. In particular, in the cases of a call-option pay-off, $\mathcal{X} = (S_T - K)^+$, and a up-and-in digital pay-off $\mathcal{Y} = I_{\{\sup_{s\in[0,T]} S_s \geq H\}}$, we find

$$\mathcal{E}_{\Delta,\Gamma}(\mathcal{X}) = S_0 \mathrm{e}^{c_\# T}\Phi(d_+) - K\Phi(d_-), \qquad \mathcal{E}_{\Delta,\Gamma}(\mathcal{Y}) = \left(\frac{H}{S_0}\right)^{2a} \overline{\Phi}(e_+) + \overline{\Phi}(e_-), \tag{7.3}$$

$$d_+ = \frac{\log(S_0/K) + aT}{\sigma\sqrt{T}}, \qquad d_- = d_+ - \sigma\sqrt{T}, \qquad a = c_\# + \frac{\sigma^2}{2}, \tag{7.4}$$

$$e_+ = \frac{\log(H/S_0) + aT}{\sigma\sqrt{T}}, \qquad e_- = e_+ - 2a\sqrt{T}/\sigma, \tag{7.5}$$

where Φ and $\overline{\Phi}$ denote the the standard normal distribution function and complementary distribution function.

7.2. **Tail-CGMY model.** Assume that the price process of a stock $S = \{S_t, t \in [0,T]\}$ follows an exponential Lévy model given by Eqn. (7.1) with

$$X_t = (\mu - \kappa)t + Y_t, \qquad \kappa = \int_{\mathbb{R}} (\mathrm{e}^x - 1 - x)\nu(\mathrm{d}x), \tag{7.6}$$

where Y is a pure-jump martingale Lévy process with Lévy measure $\Lambda(\mathrm{d}x) = k(x)\mathrm{d}x$ that has density

$$k(x) = C\left(\frac{Y + Mx}{x^{1+Y}}\mathrm{e}^{-Mx} I_{(0,\infty)}(x) + \frac{Y + G|x|}{|x|^{1+Y}}\mathrm{e}^{-G|x|} I_{(-\infty,0)}(x)\right), \qquad x \in \mathbb{R}.$$

with $M > 1, C > 0$, $G \geq 0$ and $Y \in (0,1]$. The Lévy density $k(x)$ decays exponentially fast when $|x|$ tends to infinity and has a power-type singularity at the origin, and is a mixture of classical CGMY/KoBoL Lévy densities (see [13] for the definition of a CGMY/KoBoL process). In fact, the process Y falls into the class of regular Lévy processes of exponential type (RLPE) introduced in Boyarchenko and Levendorskii [5, Def. 3.3]. By integrating the expression in the previous display it follows that the corresponding right- and left-tails functions $\overline{\Lambda}(x)$ and $\underline{\Lambda}(-x)$ of the measure Λ are given by

$$\overline{\Lambda}(x) = C\frac{\mathrm{e}^{-Mx}}{x^Y}, \qquad \underline{\Lambda}(-x) = C\frac{\mathrm{e}^{-G|x|}}{|x|^Y}, \qquad x > 0.$$

We will consider the distorted valuations that are based a distortion family constructed from the MINMAXVAR distortion Ψ_γ, with $\gamma \in \mathbb{R}_+$, that was introduced in Cherny & Madan [10] and is given by

$$\Psi_\gamma(p) = 1 - (1 - p^{1/(1+\gamma)})^{1+\gamma}.$$

A corresponding scaling family of distortions $\{\Psi(\,\cdot\,, \delta \wedge 1), \delta > 0\}$ is constructed by taking convex combinations,

$$\Psi(p, \delta) = \Psi_0(p)(1 - C_\gamma(\delta)) + \Psi_\gamma(p)C_\gamma(\delta), \qquad \delta \in (0,1)$$

where $C_\gamma(\delta)=\frac{\gamma}{1+\gamma}\delta^{\frac{\gamma}{1+\gamma}}$. Note that for any $\delta\in(0,1)$, $p\mapsto\Psi(p,\delta)$ is concave and increasing on $[0,1]$. It is straightforward to check that $\{\Psi(\,\cdot\,,\delta\wedge 1),\delta>0\}$ is a (ξ,Γ_+,Γ_-)-scaling family. In particular, we note that $\frac{\partial\Psi(0+,\delta)}{\partial p}=+\infty$, and $\frac{\partial\Psi(1-,\delta)}{\partial p}=1-C_\gamma(\delta)$. Furthermore, we see that, as $\delta\searrow 0$,

$$\frac{\Psi(p,\delta)-\Psi(p,0)}{\sqrt{\delta}}\to 0=\xi(p),\qquad \frac{\Psi(\lambda\delta,\delta)}{\delta}\to\lambda+\gamma\lambda^{1/(1+\gamma)}:=\Gamma_+(\lambda),\quad \frac{\widehat{\Psi}(\lambda\delta,\delta)}{\delta}\to\lambda:=\Gamma_-(\lambda).$$

To construct a sequence of multinomial processes approximating X, we consider the multinomial random walk R with step-probabillities

$$\begin{cases} p_k=\delta\overline{\Lambda}(kh),\quad p_{-k}=\delta\underline{\Lambda}(-kh) & k\in\mathbb{Z},|k|\geq a,\\ p_k=0 & 2\leq|k|<a-1,\end{cases}\tag{7.7}$$

where p_1,p_{-1} and p_0 are the solution of the system in Eqns. (5.2)—(5.3) which arises from matching moments. The spatial mesh size h and the step-size δ are related via $h^2=3\delta\Sigma^2(\mathbb{R})$ with

$$\Sigma^2(\mathbb{R})=2C\Gamma(2-Y)[M^{Y-2}+G^{Y-2}],$$

where Γ denotes the Gamma function, and the parameter a is specified by

$$a=a(h)=\left\{\left(\frac{\Sigma^2(2-Y)}{2CY}\right)^{1/(3-Y)}h^{(Y-2)/(3-Y)}\right\}\vee h^{-1/2}|\log h|.\tag{7.8}$$

For this choice of a the conditions stated in Lemma 6 are satisfied, as we verify next. Note that the form of a implies that we have $ah\to 0$ and $a^2h\to\infty$ when h tends to zero. Furthermore, for any $a\geq 2$, it holds $\Sigma^2(-ah,,ah)$ is bounded above by $\Sigma^2(\mathbb{R})(3-2a^{-1}-2a^{-2})$ (as this factor is larger than 1 for all $a\geq 2$) and, for a as specified in Eqn. (7.8), we have

$$\Sigma^2(-ah,,ah)\geq 2CY\int_0^{ah}\frac{x^2}{x^{1+Y}}\mathrm{d}y=\frac{2CY}{2-Y}(ah)^{2-Y}\geq\frac{\Sigma^2(\mathbb{R})}{a}.\tag{7.9}$$

The multi-period distortion expectation values of a given claim $\mathcal{X}\in L^2(\mathcal{F}_T)$ satisfy a recursion given by

$$\Pi_m^\delta(\mathcal{X})=\sum_{k\in\mathbb{Z}}\alpha_k P_k^{m+1},\qquad m=0,\ldots,n-1,$$

with $\Pi_n(\mathcal{X})=\mathcal{X}$ and $\delta=T/n$, where we denoted $P_i^{m+1}=\Pi_{m+1}(\mathcal{X}|\mathcal{G}_m)I_{\{\Delta R_m=kh\}}$, $k\in\mathbb{Z}$. If $\mathcal{X}$ is a claim of the form as stated in Thm 1 then according to Theorem 1 we have $\Pi_0^\delta(\mathcal{X})\to\mathcal{E}_{D,G}(\mathcal{X})$ as $\delta\searrow 0$.

Theorem 2 implies that the distorted expectation value of any pathwise increasing claim $\mathcal{X}\in L^2(\mathcal{F}_T)$ is given by $\mathcal{E}_{\Delta,\Gamma}(\mathcal{X})=\mathbb{E}^{Q_\#}[\mathcal{X}]$ where under the measure $Q_\#$ the process Y is a Lévy process with characteristic triplet $(c_\#,0,\Lambda_\#)$ given by

$$\overline{\Lambda}_\#(x)\;=\;\overline{\Lambda}(x)+\gamma C^{1/(1+\gamma)}\frac{\mathrm{e}^{-Mx/(1+\gamma)}}{x^{Y/(1+\gamma)}},\quad\underline{\Lambda}_\#(-x)=\underline{\Lambda}(-x),\qquad x>0,\tag{7.10}$$

$$c_\#\;=\;\gamma\int_0^\infty\mathrm{e}^{-Mx/(1+\gamma)}x^{-Y/(1+\gamma)}\mathrm{d}x=\Gamma\left(u_\gamma\right)\left(\frac{M}{1+\gamma}\right)^{-u_\gamma},\qquad u_\gamma=\frac{1+\gamma-Y}{1+\gamma}.\tag{7.11}$$

Thus, the asymmetric nature of the MINMAXVAR distortion Ψ_γ carries over in the limit, with the Lévy measure $\Lambda_\#$ having larger mass than Λ in the right-tail and the same left-tail as Λ. For the claims $\mathcal{X}$ and $\mathcal{Y}$ given in the previous example, we note that the Fourier transform of $\mathcal{E}_{\Delta,\Gamma}(\mathcal{X})$ in the log-strike $k=\log K$ and a Fourier-Laplace transform of $\mathcal{E}_{D,G}(\mathcal{Y})$ in the variables $h=\log H$ and T can be explicitly expressed in terms of the characteristic exponent of X (see e.g. [7] and [5, Sect. 7.2], respectively).

8. Proof of convergence

The proof of Theorem 1 is based on the representation of the multi-period distorted expectation as supremum of expectations of $F_\delta\left(R^\delta\right)$ under probability measures contained in the set

$$\mathcal{C}^{\Psi_\delta,\mathbf{G}_\delta}\left(F_\delta\left(R^\delta\right)\right) = \sup_{Q\in\mathcal{D}^\delta} E^Q[F_\delta\left(R^\delta\right)], \qquad \text{with} \tag{8.1}$$

$$\mathcal{D}^\delta = \mathcal{D}^{\Psi^\delta}(\mathbf{G}^\delta) = \left\{Q\in\mathcal{M}^\delta : \widehat{\Psi}^\delta\circ P_i^\delta \prec' Q_i \prec' \Psi^\delta\circ P_i^\delta \text{ for all } i=0,\ldots,n-1\right\}. \tag{8.2}$$

Throughout Sections 8–10 we fix a (ξ,Γ_+,Γ_-)-scaling family of distortions $\{\Psi(\cdot,\delta),\delta>0\}$ and we will write $\Psi(\cdot)=\Psi(\cdot,\delta)$ and $\widehat{\Psi}(\cdot)=\widehat{\Psi}(\cdot,\delta)$ to simplify the notation. We denote the corresponding jump-distortion and drift-shift by $\Gamma=(\Gamma_+,\Gamma_-)$, and $\Delta=(\Delta_+,\Delta_-)$ with $\Delta_+=\xi(\frac{1}{6})$ and $\Delta_-=\xi(\frac{5}{6})$. We denote by $F:\Omega\to\mathbb{R}$ a continuous map satisfying the bound in Eqn. (5.14). Theorem 1 follows from the following convergence result:

Proposition 3. *For any sequence $\delta_n\searrow 0$ we have*

$$\limsup_{n\to\infty}\sup_{Q\in\mathcal{D}^{\delta_n}} E^Q[F_{\delta_n}\left(R^{\delta_n}\right)] \quad\leq\quad \sup_{Q\in\mathcal{D}_{\Delta,\Gamma}}\mathbb{E}^Q[F(X)] \tag{8.3}$$

$$\liminf_{n\to\infty}\sup_{Q\in\mathcal{D}^{\delta_n}} E^Q[F_{\delta_n}\left(R^{\delta_n}\right)] \quad\geq\quad \sup_{Q\in\mathcal{D}_{\Delta,\Gamma}}\mathbb{E}^Q[F(X)]. \tag{8.4}$$

The proofs of the limits in Eqns. (8.3) and (8.4) in Proposition 3 is given in Sections 9 and 10, respectively. These proofs rest on two properties of the set of Radon-Nikodym derivatives with respect to the probability measure P of the probability measures in the sets $\mathcal{D}^{\delta_n}$: it is shown below that these sets are contained in a ball in L^2 (Lemma 12) and are sandwhiched by members of a certain family $\left\{\mathcal{D}_*^{\delta_n,\epsilon},\epsilon\in(-1,1)\right\}$ of sets of probability measures that is defined below (Lemma 13).

8.1. **Preliminaries: Radon-Nikodym derivarives.** Before proceeding we first introduce some extra notation. Denote by P^δ the probability measure on $(\Omega,\mathcal{G})$ corresponding to the random walk R, and by $\mathcal{M}^\delta$ the collection of all probability measures on $(\Omega,\mathcal{G})$ that are absolutely continuous with respect to P^δ. For $i\in\mathbb{N}\cup\{0\}$ and any probability measure $Q\in\mathcal{M}^\delta$ we denote by Q_i and P_i the conditional probabilities given by

$$Q_i(A)=Q(A|\mathcal{G}_i), \qquad P_i(A)=P_i^\delta(A|\mathcal{G}_i) \qquad \text{for any set } A\in\mathcal{G}.$$

For any $k\in\mathbb{N}$, let $I_i^k = I_{\left\{R_k^\delta-R_{k-1}^\delta=\mathtt{d}\delta+ih\right\}}$ be a Bernoulli random variable and denote the conditional probability that I_i^k is equal to one by q_i^k,

$$q_i^k = Q(I_i^k=1|\mathcal{G}_{k-1}) = Q\left(R_k^\delta-R_{k-1}^\delta=i\middle|\mathcal{G}_{k-1}\right), \qquad i\in\mathbb{N}.$$

From the sequence of Bernoulli random variables $(I_i^k,i\in\mathbb{Z},k\in\mathbb{N})$ a number of martingales will be constructed. Observe that the increments of the process $\overline{N}^A$ that was defined in (5.10) are expressed as

$$\Delta\overline{N}_k^A = \sum_{i:i\in A}[I_i^k-p_i], \qquad k=1,\ldots,N,$$

where $\Delta\overline{N}_k^A=\overline{N}_k^A-\overline{N}_{k-1}^A$ denotes the increment of Z^A at time k. Since the increments $\Delta\overline{N}_k^A$ have zero conditional expectation, $E[\Delta\overline{N}_k^A|\mathcal{G}_{k-1}]=0$, where E denotes the expectation under the measure $P^{\delta,h}$, it follows that $\overline{N}^A$ is a $(P^\delta,\mathbf{G}^\delta)$-martingale. Given a measure $Q\in\mathcal{M}^{\delta,h}$ two further martingales, M^Q and Z^Q, can be defined as follows. Denote by $Z_\infty^Q=\frac{\mathrm{d}Q}{\mathrm{d}P^\delta}$ the Radon-Nikodym derivative of the measure Q with respect to the measure P^δ. The process $Z^Q = (Z^Q_k)_{k\in\mathbb{N}\cup\{0\}}$, defined by $Z^Q_k = E[Z^Q_\infty|\mathcal{G}_k]$ for $k\in\mathbb{N}$, is a **G**-martingale, and takes the form

$$Z^Q_0 = 1, \qquad \Delta Z^Q_k = Z^Q_{k-1}\Delta M^Q_k, \qquad k\in\mathbb{N}, \tag{8.5}$$

where $\Delta M^Q_k = M^Q_k - M^Q_{k-1}$ is an increment of the martingale $M^Q = (M^Q_k, k\in\mathbb{N}^*)$ given by

$$M^Q_0 = 0, \qquad \Delta M^Q_k = \prod_{i\in\mathbb{Z}}\left(\frac{q^k_i}{p_i}\right)^{I^k_i} - 1 = \sum_{i\in\mathbb{Z}}\left(\frac{q^k_i}{p_i} - 1\right)[I^k_i - p_i], \qquad k\in\mathbb{N}. \tag{8.6}$$

It follows from Eqns. (8.5) and (8.6) that Z^Q is equal to the stochastic exponential of M^Q, and that for any $n\in\mathbb{N}$, Z^Q_n is equal to

$$Z^Q_n = \prod_{k\le n, k\in\mathbb{N}}\prod_{i\in\mathbb{Z}}\left(\frac{q^k_i}{p_i}\right)^{I^k_i}.$$

8.2. **Martingales in a ball in L^2.** The collection of processes Z^Q for $Q\in\mathcal{D}^\delta$ is contained in a ball in L^2:

Lemma 12. *There exists a $\widetilde{c}>0$ such that, for any $i=1,\ldots,n$ and any measure $Q\in\mathcal{D}^\delta$, we have*

$$E\left[\left(Z^Q_i - Z^Q_{i-1}\right)^2\middle|\mathcal{G}_{i-1}\right] \le \widetilde{c}\delta\left(Z^Q_{i-1}\right)^2. \tag{8.7}$$

with $Z^Q = \frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}$. In particular, $E\left[\left(Z^Q_i\right)^2\right] \le 1 + T\widetilde{d}$, $i=1,\ldots,n$, with $\widetilde{d} = (1-(\widetilde{c}\delta)^{n+1})/(1-\widetilde{c}\delta)$.

Proof of Lemma 12. Note that the ratio $\bar{Z}^Q_i = Z^Q_i/Z^Q_{i-1}$ can be decomposed as

$$\bar{Z}^Q_i - 1 = \Delta M^Q_i = \Delta M^{Q_c}_i + \Delta M^{Q_d}_i, \qquad i=1,\ldots,n, \tag{8.8}$$

where M^{Q_c} and $M^{Q_d}_i$ are martingales with $M^{Q_c}_0 = M^{Q_d}_0 = 0$ and with orthogonal increments given by

$$\Delta M^{Q_c}_i = \sum_{j\in\{\pm1,0\}}\left(\frac{q^i_j}{p_j} - 1\right)[I^i_j - p_j], \qquad \Delta M^{Q_d}_i = \sum_{j\notin\{\pm1,0\}}\left(\frac{q^i_j}{p_j} - 1\right)[I^i_j - p_j].$$

In view of the definition of $\mathcal{D}^\delta$, the independence of the I^i_j and the definition of the scaling family of distortions, we find

$$\begin{aligned}
E\left[\left(\Delta M^{Q_c}_i\right)^2\middle|\mathcal{G}_{i-1}\right] &= \sum_{j\in\{\pm1,0\}}\left(\frac{q^i_j}{p_j} - 1\right)^2 p_j[1-p_j] \\
&\le \sum_{j\in\{\pm1,0\}}\left(\frac{[\Psi(p_j)-p_j]\vee[p_j - \widehat{\Psi}(p_j)]}{p_j}\right)^2 p_j[1-p_j] \\
&\le \sum_{j\in\{\pm1,0\}}\left((\xi(p_j)\vee\xi(1-p_j))\sigma^*\sqrt{\delta} + o(\sqrt{\delta})\right)^2\frac{[1-p_j]}{p_j} \\
&= \delta\,(\sigma^*)^2\overline{C} + o(\delta),\ \text{where } \overline{C} = \left[10\left(\xi\left(\frac{1}{6}\right)\vee\xi\left(\frac{5}{6}\right)\right)^2 + \frac{1}{2}\left(\xi\left(\frac{1}{3}\right)\vee\xi\left(\frac{2}{3}\right)\right)^2\right],
\end{aligned} \tag{8.9}$$

where we used that $p_1, p_{-1}\to\frac{1}{6}$ and $p_0\to\frac{2}{3}$, by Lemma 6. We note that the $o(\delta)$ term holds uniformly over $i=1,\ldots,n$ and $Q\in\mathcal{D}^\delta$, in view of Eqn. (5.11) in the definition of scaling family.

Denoting by Q_d the measure with Radon-Nikodym derivatrive Z^{Q_d} with respect to P^δ given by the stochastic exponential of the martingale M^{Q_d}, we find

$$\begin{aligned} E\left[\left(\Delta M_i^{Q_d}\right)^2\middle|\mathcal{G}_{i-1}\right] &= E\left[(\widetilde{Z}^{Q_d}-1)\Delta M_i^{Q_d}\middle|\mathcal{G}_{i-1}\right] \\ &= E^{Q_d}\left[\Delta M_i^{Q_d}\middle|\mathcal{G}_{i-1}\right] - E\left[\Delta M_i^{Q_\delta}\middle|\mathcal{G}_{i-1}\right]. \end{aligned}$$

Denoting $Y_i^+ = (\Delta M_i^{Q_d})^+$ and $Y_i^- = (\Delta M_i^{Q_d})^-$, where, for any $x \in \mathbb{R}$, $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$, we find, in view of the definition of $\mathcal{D}^\delta$,

$$\begin{aligned} E^{Q_d}[Y_i^+|\mathcal{G}_{i-1}] - E[Y_i^+|\mathcal{G}_{i-1}] &= \int_{\mathbb{R}_+}[Q_d(Y_i^+ > x|\mathcal{G}_{i-1}) - P(Y_i^+ > x|\mathcal{G}_{i-1})]\mathrm{d}x \\ &\leq \int_{\mathbb{R}_+}[\Psi(P(Y_i^+ > x|\mathcal{G}_{i-1})) - P(Y_i^+ > x|\mathcal{G}_{i-1})]\mathrm{d}x \\ &= \delta\int_{\mathbb{R}_+}[\delta^{-1}\Psi(\delta\Lambda_{Y,i}(x,\infty)) - \Lambda_{Y,i}(x,\infty)]\mathrm{d}x \\ &= \delta\int_{\mathbb{R}_+}A_\delta(x)[\Gamma(\Lambda_{Y,i}(x,\infty)) - \Lambda_{Y,i}(x,\infty)]\mathrm{d}x := I^+_{\delta,i}, \end{aligned} \tag{8.10}$$

where $\Lambda_{Y,i}$ is the measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ given by $\Lambda_{Y,i}(\mathrm{d}x) = P(Y_i^+ \in \mathrm{d}x|\mathcal{G}_{i-1})/\delta$ and

$$A_\delta(x) = \frac{\delta^{-1}\Psi(\delta\Lambda_{Y,i}(x,\infty)) - \Lambda_{Y,i}(x,\infty)}{\Gamma_+(\Lambda_{Y,i}(x,\infty)) - \Lambda_{Y,i}(x,\infty)}.$$

Since the measure $\Lambda_{Y,i}$ has a finite second moment, $m_{2,i} = E[(Y_i^+)^2|\mathcal{G}_{i-1}]/\delta$, and $\Gamma_+ - \mathrm{id}$ is increasing it follows by Chebyshev's inequality

$$\Gamma_+(\Lambda_{Y,i}(x,\infty)) - \Lambda_{Y,i}(x,\infty) \leq (\Gamma_+ - \mathrm{id})(m_{2,i}/x^2), \qquad x > 0.$$

Moreover, for δ sufficiently small, $\sup_{x>0} A_\delta(x)$ is bounded by some finite constant $A_+ > 0$, in view of Eqn. (5.12) in the definition of distortion scaling family. A change of variables yields that the integral $I^+_{\delta,i}$ is bounded in terms of $m_{2,i}$: $I^+_{\delta,i} \leq \delta\sqrt{m_{2,i}}A_+C_+$, where

$$C_+ = \int_{\mathbb{R}_+}[(\Gamma_+ - \mathrm{id})(y)]\frac{\mathrm{d}y}{2y\sqrt{y}}.$$

Thus, we find

$$E^{Q_d}[Y_i^+|\mathcal{G}_{i-1}] - E[Y^+|\mathcal{G}_{i-1}] \leq A_+C_+\sqrt{\delta}\cdot\sqrt{E[(Y_i^+)^2|\mathcal{G}_{i-1}]},$$

for all δ sufficiently small. By an analogous line of reasoning, we find

$$E^{Q_d}[Y_i^-|\mathcal{G}_{i-1}] - E[Y_i^-|\mathcal{G}_{i-1}] \geq A_-C_-\sqrt{\delta}\cdot\sqrt{E[(Y_i^-)^2|\mathcal{G}_{i-1}]},$$

for δ sufficiently small, with $C_- = \int_{\mathbb{R}_+}[(\mathrm{id} - \Gamma_-)(y)]\frac{\mathrm{d}y}{2y\sqrt{y}}$ and some finite constant $A_- < 0$.

Since we have $E[(Y_i^+)^2|\mathcal{G}_{i-1}] \vee E[(Y_i^-)^2|\mathcal{G}_{i-1}] \leq E[(\Delta M_i^Q)^2|\mathcal{G}_{i-1}]$ we find the estimate

$$E[(\Delta M_i^Q)^2|\mathcal{G}_{i-1}] \leq \delta\overline{C} + C_0\sqrt{\delta}\sqrt{E[(\Delta M_i^Q)^2|\mathcal{G}_{i-1}]} + o(\delta),$$

as $\delta \to 0$, where $C_0 = A_+C_+ + |A_-|C_-$ and $\overline{C}$ is defined in Eqn. (8.9). Thus, by solving this inequality and taking expectations we find

$$E[(\Delta M_i^Q)^2] \leq C\delta + o(\delta), \text{ where } C = \frac{1}{4}\left(C_0 + \sqrt{(C_0)^2 + 4\overline{C}}\right)^2, \tag{8.11}$$

as $\delta \to 0$, where we recall that the $o(\delta)$ term in Eqn. (8.11) is valid uniformly across $i = 1, \ldots, n$ and $Q \in \mathcal{D}^\delta$. In view of the relation between M^Q and Z^Q (in Eqn. (8.8)), we see that the bound on Z^Q in Eqn. (8.7) is valid with $\widetilde{c}$ taken equal to C given in Eqn. (8.11). Furthermore, orthogonality of martingale increments implies

$$E[(Z_i^Q)^2] = E[(Z_0^Q)^2] + \sum_{j=1}^{i} E[(Z_j^Q - Z_{j-1}^Q)^2] \leq 1 + \widetilde{c}\delta \sum_{j=1}^{i} E[(Z_{j-1}^Q)^2] \qquad i = 1, \ldots, n,$$

with $\delta = T/n$. Solving this recursion yields the stated bound. □

8.3. **Related sets of probability measures.** The collection $\mathcal{D}^\delta$ is contained and contains sets of probability measures that are defined in terms of the covariance between the stochastic logarithm of the martingale density process and the martingales $R^{(+1)}$, $R^{(-1)}$ and $R^{(0)}$ that start at the origin, $R_0^{(+1)} = R_0^{(-1)} = R^{(0)} = 0$, and have increments

$$\Delta R_n^{(+1)} = h(I_1^n - p_1), \quad \Delta R_n^{(-1)} = h(I_{-1}^n - p_{-1}) \ \text{ and } \ \Delta R_n^{(0)} = h(I_1^n + I_{-1}^n - p_1 - p_{-1}), \qquad n \in \mathbb{N}.$$

For $\epsilon \in (-1, 1)$ and $\delta > 0$ the collections in question are defined as follows:

$$\mathcal{D}_*^{\delta,\epsilon} = \left\{ Q \in \mathcal{M}^\delta : \begin{array}{ll} M^Q \text{ is such that the following hold for all } k \in \mathbb{N}: & \\ E\left[\Delta M_k^Q \Delta R_k^{(i)} \middle| \mathcal{G}_{k-1}\right] \in \left[-\delta \underline{\Delta}_\epsilon^{(i)}, \delta \overline{\Delta}_\epsilon^{(i)}\right] \ \text{ for } i \in \{-1, 0, 1\} & \\ E\left[\Delta M_k^Q \Delta \overline{N}_k^{A_h} \middle| \mathcal{G}_{k-1}\right] \leq \delta \overline{\gamma}_\epsilon(A) & \\ E\left[\Delta M_k^Q \Delta \overline{N}_k^{A_h} \middle| \mathcal{G}_{k-1}\right] \geq \delta \underline{\gamma}_\epsilon(A) & \text{for all closed sets } A \subset \mathbb{R}\backslash\{0\} \end{array} \right\}.$$

where we denote $A_h = A \cap (h\mathbb{Z})$ and $A_h^c = h\mathbb{Z}\backslash A_h$, and the constants $\overline{\Delta}_\epsilon^{(i)}$ and $\underline{\Delta}_\epsilon^{(i)}$, $i = \pm 1$ and $\overline{\gamma}_\epsilon(A)$ and $\underline{\gamma}_\epsilon(A)$ are given by

$$\begin{aligned} &\overline{\Delta}_\epsilon^{(i)} = (1+\epsilon)\frac{5}{12}\sigma^2 \overline{\xi}^{(i)} + \epsilon^+ \mathbf{1}_{\{\overline{\xi}^{(i)} \cdot \sigma^2 = 0\}}, \qquad \underline{\Delta}_\epsilon^{(i)} = (1+\epsilon)\frac{5}{12}\sigma^2 \underline{\xi}^{(i)} + \epsilon^+ \mathbf{1}_{\{\underline{\xi}^{(i)} \cdot \sigma^2 = 0\}}, \qquad i \in \{\pm 1, 0\}, \\ &\overline{\gamma}_\epsilon(A) = (1+\epsilon)\left[\Gamma_+(\Lambda(A)) - \Lambda(A)\right] + \epsilon^+ \mathbf{1}_{\{\Gamma_+(\Lambda(A)) = \Lambda(A)\}}, \\ &\underline{\gamma}_\epsilon(A) = (1+\epsilon)\left[\Gamma_-(\Lambda(A)) - \Lambda(A)\right] - \epsilon^+ \mathbf{1}_{\{\Gamma_-(\Lambda(A)) = \Lambda(A)\}}, \end{aligned}$$

for closed sets $A \subset \mathbb{R}\backslash\{0\}$, with $\epsilon^+ = \max\{\epsilon, 0\}$ and

$$\overline{\xi}^{(1)} = \overline{\xi}^{(-1)} = \Delta_+, \ \underline{\xi}^{(1)} = \underline{\xi}^{(-1)} = \Delta_-, \ \overline{\xi}^0 = \xi\left(\frac{1}{3}\right) \ \text{ and } \underline{\xi}^0 = \xi\left(\frac{2}{3}\right).$$

Lemma 13. *Fix $\epsilon \in (0, 1)$. Then for all δ sufficiently small, we have $\mathcal{D}^\delta \subset \mathcal{D}_*^{\delta,\epsilon}$ and $D_*^{\delta,-\epsilon} \subset \mathcal{D}^\delta$.*

Before turning to the proof we collect some general observations. We shall simplify the notation by writing, for any set $A \subset h\mathbb{Z}$, $p(A) = P_i^\delta(\Delta R_{i+1} \in A|\mathcal{G}_i)$ and $q(A) = Q_i(\Delta R_{i+1} \in A|\mathcal{G}_i)$. In view of the definition of M^Q in (8.6) we find the following identity for any subset $A \in 2^\mathbb{Z}$

$$E\left[\Delta M_k^Q \Delta \overline{N}_k^A \middle| \mathcal{G}_{k-1}\right] = \sum_{i: ih \in A} \left(\frac{q_i^k}{p_i} - 1\right) E\left[(I_i^k - p_i)^2\right] = \sum_{i: ih \in A} (q_i^k - p_i)(1 - p_i), \tag{8.12}$$

where the last sum can be estimated by

$$\sum_{i:ih\in A}(q_i^k-p_i)(1-p_i)-[q(A)-p(A)] \;=\; -\sum_{i:ih\in A}(q_i^k-p_i)p_i\le p(A)^2 \tag{8.13}$$

$$\sum_{i:ih\in A}(q_i^k-p_i)(1-p_i)-[q(A)-p(A)] \;\ge\; -[q(A^+)-p(A^+)]\max_{i:ih\in A^+}p_i, \tag{8.14}$$

with $A^+=\{i: ih\in A, q_i^k>p_i\}$.

Furthermore, in view of the definitions of the martingales M^Q, $R^{(1)}$, $R^{(-1)}$ we have

$$E[\Delta M_k^Q\Delta R_k^{(\pm1)}]=E\left[\sum_{i\in\mathbb{Z}}\left(\frac{q_i^k}{p_i}-1\right)[I_i^k-p_i]\,(\pm h(I_{\pm1}-p_{\pm1}))\right]=\pm h(q_{\pm1}-p_{\pm1})(1-p_{\pm1}) \tag{8.15}$$

$$E[\Delta M_k^Q\Delta R_k^{(0)}]=h[(q_1-p_1)(1-p_1)+(q_{-1}-p_{-1})(1-p_{-1})]. \tag{8.16}$$

Proof of Lemma 13. In the proof both inclusions will be considered separately.

(i) (Proof of the inclusion $\mathcal{D}^\delta\subset\mathcal{D}_*^{\delta,\epsilon}$): Let $Q\in\mathcal{D}^\delta$. The proof of this inclusion rests on the following three observations:

(a) In the case $\sigma^2>0$ we have in view of Eqn. (8.15), Lemma 6(i,ii) and Definition 9

$$\begin{aligned}E[\Delta M_k^Q\Delta R_k^{(1)}] &\le h\left(\frac{\xi(p_1)\sigma}{2\sqrt{3}}\sqrt{\delta}+o(\sqrt{\delta})\right)(1-p_1)=\frac{1}{2}\cdot\frac{5}{6}\xi\left(\frac{1}{6}\right)\delta\sigma^2+o(\delta)\\ &\ge -h\left(\frac{\xi(1-p_1)\sigma}{2\sqrt{3}}\sqrt{\delta}+o(\sqrt{\delta})\right)(1-p_1)=-\frac{5}{12}\xi\left(\frac{5}{6}\right)\delta\sigma^2+o(\delta),\end{aligned}$$

when δ tends to zero. By a similar line of reasoning it follows that $E[\Delta M_k^Q\Delta R_k^{(-1)}]$ satisfies the same bounds as in above display. Furthermore, in the case $\sigma^2=0$ Lemma 6 and Definition 9 imply that we have $E\left[\Delta M_k^Q\Delta R_k^{(j)}\right]\in\delta[-\epsilon,\epsilon]$ for all δ sufficiently small and $j=-1,1$.

Hence, for all δ sufficiently small, we have

$$-\underline{\Delta}_\epsilon^{(i)}\le E\left[\Delta M_k^Q\Delta R_k^{(i)}\right]\le\overline{\Delta}_\epsilon^{(i)}\qquad\text{for } i=\pm1. \tag{8.17}$$

(b) Let $A\subset\mathbb{R}\backslash\{0\}$ be a closed set and recall the notation $A_h=A\cap h\mathbb{Z}$. Recall from the definition of the transition probabilities p_k in Eqn. (5.1) that we have $p(A_h)=\delta\Lambda(A)$ for all h sufficiently small. By combining Eqns. (8.12) and (8.14) with the definition of $\mathcal{D}^\delta$ and Definition 9 we find

$$E\left[\Delta M_k^Q\Delta\overline{N}_k^{A_h}\middle|\mathcal{G}_{k-1}\right]\le\Psi(p(A_h))-p(A_h)+p(A_h)^2\le\delta(1+\epsilon/2)[\Gamma_+(\Lambda(A))-\Lambda(A)]+\delta^2\Lambda(A)^2]$$

for all δ sufficiently small, which is bounded above by $\delta(1+\epsilon)[\Gamma_+(\Lambda(A))-\Lambda(A)]$ for all δ sufficiently small in the case $\Gamma_+(\Lambda(A))>\Lambda(A)$, and by $\delta\epsilon$ otherwise.

(c) Let $A\subset\mathbb{R}\backslash\{0\}$ be a closed set. Then we have

$$E\left[\Delta M_k^Q\Delta\overline{N}_k^{A_h}\middle|\mathcal{G}_{k-1}\right]\ge-[q(A^+)-p(A^+)]\max_{i:ih\in A^+}p_i\ge-(\Psi(p(A^+))-p(A^+))p(A^+)\ge-\Psi(p(A))p(A),$$

where we used that Ψ is increasing. Then, similarly as above ot follows from Eqns. (8.12) and (8.13)

$$\begin{aligned}E\left[\Delta M_k^Q\Delta\overline{N}_k^{A_h}\middle|\mathcal{G}_{k-1}\right] &\ge \widehat{\Psi}(p(A_h))-p(A_h)(1+\Psi(p(A_h)))\\ &\ge \delta(1-\epsilon/2)[\Gamma_-(\Lambda(A))-\Lambda(A)]-\delta^2\Lambda(A)(1+\Gamma_+(\Lambda))\end{aligned}$$

for all δ sufficiently small, which is bounded below by $\delta(1-\epsilon)[\Gamma_-(\Lambda(A)) - \Lambda(A)]$ for all δ sufficiently small in the case $\Gamma_-(\Lambda(A)) < \Lambda(A)$ and by $-\delta\epsilon$ otherwise.

In conclusion, for all δ sufficiently small and for arbitrary $Q \in \mathcal{D}^\delta$ it follows from points (a), (b) and (c) that the associated martingale M^Q satisfies the conditions listed in the definition of $\mathcal{D}_*^{\delta,\epsilon}$, and we conclude $Q \in \mathcal{D}_*^{\delta,\epsilon}$ for all δ sufficiently small.

(ii) (Proof of the inclusion $\mathcal{D}_*^{\delta,-\epsilon} \subset \mathcal{D}^\delta$) Let Q be an arbitrary element of $\mathcal{D}_2^{\delta,-\epsilon}$. To prove that Q is element of $\mathcal{D}^\delta$ for all δ sufficiently small, it suffices to verify that, there exists a $\delta_0 > 0$ such that $q(A) \leq \Psi(p(A))$ for all sets $A \subset h\mathbb{Z}$ and all $\delta \in (0, \delta_0)$. We show this in three steps: (a) for sets $A \subset h\mathbb{Z}\backslash\{-h, h\}$, (b) for sets $A = \{h\}, \{-h\}$ and $\{-h, h\}$ and (c) for all sets $A \subset h\mathbb{Z}$

(a) In view of the uniform limits in Eqn. (5.12) in the definition of scaling family of distortions and Eqns. (8.12)—(8.14) it follows that there exists a δ_0 such that for all $0 < \delta < \delta_0$ the following inequalities are valid uniformly across closed sets $A \in \mathbb{R}\backslash\{0\}$ satisfying $\Gamma_+(\Lambda(A)) > \Lambda(A)$ and $\Gamma_-(\Lambda(A)) < \Lambda(A)$, respectively:

$$q(A_h) - p(A_h) \leq \delta(1-\epsilon/2)[\Gamma_+(\Lambda(A)) - \Lambda(A)] \leq (\Psi(p(A_h)) - p(A_h)), \tag{8.18}$$

$$q(A_h) - p(A_h) \geq \delta(1-\epsilon/2)[\Gamma_-(\Lambda(A)) - \Lambda(A)] \geq (\widehat{\Psi}(p(A_h)) - p(A_h)), \tag{8.19}$$

where as before we denote $A_h = A \cap h\mathbb{Z}$. In the case $\Gamma_+(\Lambda(A)) = \Lambda(A)$ we note $q(A_h) \leq p(A_h) \leq \Psi(p(A_h))$ by concavity Ψ and the definition of the set $\mathcal{D}_*^{\delta,-\epsilon}$. Similarly, we find $q(A_h) \geq p(A_h) \geq \widehat{\Psi}(p(A_h))$ in the case $\Gamma_-(\Lambda(A)) = \Lambda(A)$.

(b) Consider the case $\sigma^2 > 0$ and when $\overline{\xi}^{(\pm 1)}$ and $\underline{\xi}^{(\pm 1)}$ are strictly positive. By combining Lemma 6 and Definition 9, the definition of $\mathcal{D}^{\delta,-\epsilon}$ and Eqn. (8.15) we find the estimates for $j = \pm 1$,

$$\begin{aligned} q_j^i &\leq p_j + (1-\epsilon)\frac{1}{1-p_j}\frac{5\overline{\xi}^{(j)}\sigma^2}{12}\frac{\delta}{h} \leq p_j + (1-\epsilon)\frac{\overline{\xi}^{(j)}\sigma}{2\sqrt{3}}\sqrt{\delta}\cdot\frac{\sigma\sqrt{\delta}\sqrt{3}}{h} + o(\sqrt{\delta}) \leq \Psi(p_j) \\ q_j^i &\geq p_j - (1-\epsilon)\frac{\underline{\xi}^{(j)}\sigma}{2\sqrt{3}}\sqrt{\delta} + o(\sqrt{\delta}) \geq \widehat{\Psi}(p_j) \end{aligned}$$

for all δ sufficiently small. By a similar line of reasoning, using instead Eqn. (8.16), it follows, in the case σ^2, $\overline{\xi}^{(0)}$ and $\underline{\xi}^{(0)}$ are strictly positive,

$$\begin{aligned} q_1^i + q_{-1}^i &\leq p_1 + p_{-1} + (1-\epsilon)\frac{\overline{\xi}^{(0)}\sigma}{2\sqrt{3}}\sqrt{\delta} + o(\sqrt{\delta}) \leq \Psi(p(\{-1,1\})) \\ &\geq \widehat{\Psi}(p(\{-1,1\})) = 1 - \Psi(p(\{-1,1\}^c)). \end{aligned}$$

In the degenerate cases that σ^2, $\overline{\xi}^{(0)}$ or $\underline{\xi}^{(0)}$ are zero we find as before $q(A_h) \leq p(A_h) \leq \Psi(p(A_h))$ and $q(A_h) \geq p(A_h) \geq \widehat{\Psi}(p(A_h))$ for sets $A_h = \{h\}, \{-h\}, \{-h, h\}$, by concavity of Ψ.

(c) We claim that the above line of reasoning implies that $q(A) \leq \Psi(p(A))$ for any subset $A \subset h\mathbb{Z}$. That this is the case can be seen as follows. Since Ψ is concave with $\Psi(0) = 0$, it follows that Ψ is sublinear, in the sense that $\Psi(x+y) - \Psi(y) \leq \Psi(x)$ for all $x, y \in [0,1]$ with $x + y \leq 1$. Analogously it follows that $\widehat{\Psi}$ is super-linear, as $\widehat{\Psi}$ is convex. Hence, in view of the bounds in Eqn. (8.18), it follows that we have

$$\begin{cases} q(A) \leq \Psi(p(A)) & \text{for any set } A \text{ of the form } A = J \cup B \\ q(A) \geq \widehat{\Psi}(p(A)) \Leftrightarrow q(A^c) \leq \Psi(p(A^c)) & \text{with } J = \emptyset, \{h\}, \{-h\} \text{ or } \{-h, h\} \text{ and } B \subset \mathbb{Z}\backslash\{\pm 1, 0\}. \end{cases}$$

Since any subset of $h\mathbb{Z}$ is such that it is either of the form $J \cup B$ or of the form $(J \cup B)^c$ for sets J and B as mentioned in the previous line, the claim follows, and the proof is complete. □

9. Proof of convergence: Upper bound

The proof of Eqn. (8.3) in Proposition 3 is based on two auxiliary results.

Lemma 14. *Let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence of time steps that tends to zero, and let $(h_n)_{n\in\mathbb{N}}$ be the corresponding series of grid-sizes. Let $(Q_n)_{n\in\mathbb{N}}$ be a sequence of measures with $Q_n \in \mathcal{D}^{\delta_n}$ and Radon-Nikodym derivatives Z^{Q_n}, and let $M^{(n)} = M^{Q_n}$ be such that $Z^{Q_n} = \mathbf{Exp}(M^{(n)})$ is the stochastic exponential of $M^{(n)}$. Denote by $\widetilde{M}^{(n)}$ and $\widetilde{Z}^{(n)}$ the stochastic processes on Ω defined by $\widetilde{M}^{(n)}_t = M^{(n)}_{\lfloor t/\delta_n \rfloor}$ and $\widetilde{Z}^{(n)}_t = Z^{(n)}_{\lfloor t/\delta_n \rfloor}$. Then the following hold true:*

(i) *The sequence $(\widetilde{M}^{(n)}, Y^{\delta_n})_{n\in\mathbb{N}}$ is tight.*
(ii) *If we have $\widetilde{M}^{(n)} \Rightarrow \widetilde{M}_\infty$ then it holds $\widetilde{Z}^{(n)} \Rightarrow \widetilde{Z}_\infty$.*

Lemma 15. *Let the sequences $(\delta_n, h_n)_{n\in\mathbb{N}}$ and $(Z^{(n)})_{n\in\mathbb{N}}$ be as in Lemma 14. Any limit point of $(\widetilde{Z}^{(n)}, Y^{(\delta_n)})_{n\in\mathbb{N}}$ is of the form (Z_∞, X) with the measure Q_∞ that has Radon-Nikodym derivatve $\frac{\mathrm{d}Q_\infty}{\mathrm{d}P} = Z_\infty$ satisfying $Q_\infty \in \mathcal{D}_{D,G}$.*

Proof of the identity in Eqn. (8.3) *in Proposition 3.* Let $\epsilon > 0$. For each $n \in \mathbb{N}$ there exists an ϵ-optimal solution $Q^n \in \mathcal{D}^{\delta_n}$, that is,

$$E^{Q^n}\left[F_{\delta_n}\left(R^{\delta_n}\right)\right] \geq \sup_{Q\in\mathcal{D}^{\delta_n}} E^{Q^n}\left[F_{\delta_n}\left(R^{\delta}_n\right)\right] - \epsilon. \tag{9.1}$$

Let Z^n denote the martingale associated to the Radon-Nikodym derivative $\frac{\mathrm{d}Q^n}{\mathrm{d}P}$ given by $Z^n_k = \mathbb{E}[\frac{\mathrm{d}Q^n}{\mathrm{d}P}|\mathcal{G}_k]$ for $k \in \mathbb{N}\cup\{0\}$, and denote by $\widetilde{Z}^n = \{\widetilde{Z}^n_t, t \in [0,T]\}$ the embedding of Z^n into $D([0,T])$ defined by $\widetilde{Z}^n_t = Z^n_{\lfloor t/\delta \rfloor}$.

By Lemma 15 the sequence $(\widetilde{Z}^n, Y^n)_{n\in\mathbb{N}}$ is tight and any limit point is of the form $(\widetilde{Z}_\infty, X)$ where the measure Q_∞ with Radon-Nikodym derivative $\widetilde{Z}_\infty$ belongs to $\mathcal{D}_{D,G}$. The continuity of F implies that also the sequence $(F(Y^n))_{n\in\mathbb{N}}$ is tight with $Y^n = Y^{\delta_n}$. For any random variable Y denote by $Y_m := (-m) \vee Y \wedge m$ the truncation of Y by $m \in \mathbb{R}_+$. Since the collection $(\widetilde{Z}^n_T(F(Y^n))_m)_{n\in\mathbb{N}}$ lies in a ball in L^2 in view of Lemma 12 it follows

$$\limsup_{n\to\infty} \mathbb{E}\left[(\widetilde{Z}^{Q_n}_T F(Y^n))_m\right] \leq \sup_{Q\in\mathcal{D}_{D,G}} \mathbb{E}_Q[(F(X))_m]. \tag{9.2}$$

The triangle and Cauchy-Schwarz inequalities and the estimate $|Y - Y_m| \leq |Y| I_{\{|Y|>m\}}$ for any random variable Y imply

$$\begin{aligned} \left| \sup_{Q\in\mathcal{D}_{D,G}} \mathbb{E}^Q[F(X)] - \sup_{Q\in\mathcal{D}_{D,G}} \mathbb{E}[F(X)_m] \right| &\leq \sup_{Q\in\mathcal{D}_{D,G}} \mathbb{E}^Q[|F(X)| I_{\{|F(X)|>m\}}] \\ &\leq \sup_{Q\in\mathcal{D}_{D,G}} \mathbb{E}[(Z^Q_T)^2]^{1/2}\, \mathbb{E}[F(X)^2 I_{\{|F(X)|>m\}}]^{1/2}, \end{aligned} \tag{9.3}$$

which tends to zero as m tends to infinity by Lebesgue's dominated convergence theorem, since the collection $(Z_T^Q)_{Q\in\mathcal{D}_{D,G}}$ lies in a ball in L^2 by Lemma 12 and $F(X)$ is square-integrable by Lemma 8. Deploying Lemmas 8 and 12 it follows that the expression in Eqn. (9.3) also tends to zero as $m\to\infty$ in the case that the set $\mathcal{D}_{D,G}$ is replaced by $\cup_n\mathcal{D}_1^{(\delta_n)}$ and $F(X)$ by $F_{\delta_n}(R^{\delta_n})$.

These observations imply that the limit of m tending to infinity and the limsup of n tending to infinity in Eqn. (9.2) can be interchanged, yielding $\limsup_{n\to\infty}\mathbb{E}^{Q^n}[F(Y^n)]\leq\sup_{Q\in\mathcal{D}_{D,G}}\mathbb{E}^Q[F(X)]$. Hence, it follows from Eqn. (9.1)

$$\limsup_{n\to\infty}\sup_{Q\in\mathcal{D}^{\delta_n}}E^Q[F_{\delta_n}(R^{\delta_n})]\leq\sup_{Q\in\mathcal{D}_{D,G}}\mathbb{E}^Q[F(X)]+\epsilon.$$

Since $\epsilon>0$ was arbitrary the proof is finished. □

Proof of Lemma 14. (i) In view of Lemma 12 and the construction of Y^δ, the sequence of martingales $(\widetilde{M}^{(n)},Y^{\delta_n})_{n\in\mathbb{N}}$ satisfies the following moment conditions: for any $\epsilon>0$, s,t_1,t_2, with $t_1<t_2$ and $s\in(t_1,t_2)$ and $n\geq\lceil\frac{1}{t_2-t_1}\rceil$ we have

$$\begin{aligned}
&\mathbb{E}\left[\left\|\left(\widetilde{M}^{(n)}(s),Y^{\delta_n}(s)\right)-\left(\widetilde{M}^{(n)}(t_1),Y^{\delta_n}(t_1)\right)\right\|^2\left\|\left(\widetilde{M}^{(n)}(t_2),Y^{\delta_n}(t_2)\right)-\left(\widetilde{M}^{(n)}(s),Y^{\delta_n}(s)\right)\right|^2\right]\\
&=\ \mathbb{E}\left[\left\|\left(\widetilde{M}^{(n)}(s)-\widetilde{M}^{(n)}(t_1),Y^{\delta_n}(s)-Y^{\delta_n}(t_1)\right)\right\|^2E\left[\left\|\left(\widetilde{M}^{(n)}(t_2)-\widetilde{M}^{(n)}(s),Y^{\delta_n}(t_2)-Y^{\delta_n}(s)\right)\right\|^2\middle|\mathcal{F}_s\right]\right]\\
&\leq\ C\frac{(\lfloor ns\rfloor-\lfloor nt_1\rfloor)(\lfloor nt_2\rfloor-\lfloor ns\rfloor)}{n^2}\leq 3(C+\epsilon)^2(t_2-t_1)^2,
\end{aligned}$$

where C is some constant. The moment-criterion for processes in $D[0,T]$ (see [3, Thm. 13.5]) implies that the sequence $(\widetilde{M}^{(n)})_{n\in\mathbb{N}}$ is tight, and as a consequence, also the sequence of pairs $(\widetilde{M}^{(n)},Y^{\delta_n})_{n\in\mathbb{N}}$ is tight.

(ii) The expectation $E\left[\sup_{t\leq T}|\Delta\widetilde{M}_t^{(n)}|\right]$ is finite since we have by the Cauchy-Schwarz inequality

$$\begin{aligned}
E\left[\sup_{t\leq T}|\Delta\widetilde{M}_t^{(n)}|\right]\ &\leq\ E\left[\sup_{t\leq T}|\Delta\widetilde{M}_t^{(n)}|^2\right]^{1/2}\leq\left[\sum_{k=1}^n E[|\Delta\widetilde{M}_k^{(n)}|^2]\right]^{1/2}\\
&\leq\ \left[\sum_{k=1}^n\delta T\widetilde{c}^2\right]^{1/2}=T^{1/2}\widetilde{c}<\infty,
\end{aligned}$$

where in the final equality we deployed the bound in Lemma 12. Hence, in case $\widetilde{M}^{(n)}\Rightarrow\widetilde{M}_\infty$, the stability of stochastic integrals (Thm. 4.4 in [21]) yields the convergence of the corresponding stochastic exponentials: $\widetilde{Z}^{(n)}\Rightarrow\widetilde{Z}_\infty$ as n tends to infinity, where $\widetilde{Z}_\infty$ is the stochastic exponential of $\widetilde{M}_\infty$. □

Proof of Lemma 15. Let $\epsilon>0$ and denote by $Q^{(n)}$ the measure with Radon-Nikodym derivative $Z^{(n)}$. For any stochastic process U and any $s,t\geq 0$ we denote the increment of U over $[s,t]$ by $\Delta_{s,t}U:=U_t-U_s$. Also denote $Y_t^{c(n)}=Y_t^{\delta_n(c)}$ and $Z_t^{A(n)}=Z^{\delta_n,A}$.

Consider the case $\sigma^2>0$. Since the measure $Q^{(n)}$ is contained in the set $\mathcal{D}^{\delta_n}$ which is itself contained in $D_*^{\delta,\epsilon}$ for all n sufficiently large (Lemma 13), it follows from the definition of $\mathcal{D}_*^{\delta_n,\epsilon}$ that we have

$$\mathbb{E}[\Delta_{s,t}\widetilde{Z}^{(n)}\Delta_{s,t}Y^{c(n)}|\widetilde{\mathcal{G}}_{s-\epsilon}]\ \leq\ (t-s+\delta_n)\overline{\Delta}[1+\epsilon]\widetilde{Z}_{s-\epsilon}^{(n)},\tag{9.4}$$

$$\mathbb{E}[\Delta_{s,t}\widetilde{Z}^{(n)}\Delta_{s,t}Y^{c(n)}|\widetilde{\mathcal{G}}_{s-\epsilon}]\ \geq\ -(t-s+\delta_n)\underline{\Delta}[1+\epsilon]\widetilde{Z}_{s-\epsilon}^{(n)}.\tag{9.5}$$

for all n sufficiently large and for all $s,t\in[0,T]$ with $s\leq t$, where $\overline{\Delta}=\frac{10}{12}\sigma^2\Delta_+$ and $\underline{\Delta}=\frac{10}{12}\sigma^2\Delta_-$ and $\widetilde{\mathcal{G}}_s$ denotes the sigma-algebra generated by $\{\widetilde{Z}_u^{(n)}\}_{u\leq s}$.

Let A be a closed subset of $\mathbb{R}\backslash\{0\}$. Then it follows by an analogous line of reasoning that we have

$$\mathbb{E}[\Delta_{s,t}\widetilde{Z}^{(n)}\Delta_{s,t}Z^{A(n)}|\widetilde{\mathcal{G}}_{s-\epsilon}] \leq (t-s+\delta_n)\overline{\gamma}(A)[1+\epsilon]\widetilde{Z}^{(n)}_{s-\epsilon}, \tag{9.6}$$

$$\mathbb{E}[\Delta_{s,t}\widetilde{Z}^{(n)}\Delta_{s,t}Z^{A(n)}|\widetilde{\mathcal{G}}_{s-\epsilon}] \geq -(t-s+\delta_n)\underline{\gamma}(A)[1+\epsilon]\widetilde{Z}^{(n)}_{s-\epsilon}, \tag{9.7}$$

in the cases $\Gamma_+ \neq \mathrm{id}$ and $\Gamma_- \neq \mathrm{id}$ respectively, with $\overline{\gamma}(A) = (\Gamma_+ - \mathrm{id})(\Lambda(A))$ and $\underline{\gamma}(A) = (\Gamma_- - \mathrm{id})(\Lambda(A))$.

Fatou's lemma applied to Eqn. (9.4) implies that for any continuous adapted function $(t,\omega) \mapsto H(t,\omega)$ mapping $[0,T]\times\Omega$ to $[0,1]$ we have

$$\limsup_{n\to\infty}\mathbb{E}[H(s-\epsilon,Y^{(n)})\{\Delta_{s,t}\widetilde{Z}^{(n)}\Delta_{s,t}Y^{c(n)} - (t-s+\delta_n)\overline{\Delta}[1+\epsilon]\widetilde{Z}^{(n)}_{s-\epsilon}\}] \leq 0.$$

Let $\widetilde{M}_\infty$ be any limit point of the sequence of martingales $(\widetilde{M}^{(n)})_{n\in\mathbb{N}}$. Then, by Lemma 14(ii), $\widetilde{Z}_\infty = \mathbf{Exp}(\widetilde{M}_\infty)$ is a limit point of the sequence of stochastic exponentials $(\widetilde{Z}^{(n)})_{n\in\mathbb{N}}$. Since this sequence is contained in a ball in $L^2(\Omega,\mathcal{F},\mathbb{P})$ by Lemma 12, it follows that it is UI, and as a consequence we have that $\widetilde{Z}_\infty$ is a martingale. Recall from Lemma 7 that the sequences $(Y^{c(n)})_{n\in\mathbb{N}}$ and $(Z^{A(n)})_{n\in\mathbb{N}}$ admit limits in the sense of weak convergence. Furthermore, since the collections $\{(\Delta_{s,t}\widetilde{Z}^{(n)})\}_{n\in\mathbb{N}}$, $\{(\Delta_{s,t}Y^{c(n)})^p\}_{n\in\mathbb{N}}$, for any $p>0$, are contained in a ball in $L^2(\Omega,\mathcal{F},\mathbb{P})$ (by Lemmas 8 and 12), it follows that also the collection $\{\Delta_{s,t}\widetilde{Z}^{(n)}\Delta_{s,t}Y^{c(n)}, n\in\mathbb{N}\}$ is UI. Hence we have

$$\mathbb{E}[H(s-\epsilon,X)\{(\Delta_{s,t}\widetilde{Z}_\infty,\Delta_{s,t}X^c) - \widetilde{Z}_\infty(s-\epsilon)\overline{\Delta}(1+\epsilon)(t-s)\}] \leq 0.$$

Since $\epsilon>0$ was arbitrary and $H(s-\epsilon,X)$ is $\mathcal{F}_s$-measurable for any $s\in[0,T]$, and we have $\mathbb{E}[\Delta_{s,t}\widetilde{Z}_\infty\Delta_{s,t}X^c|\mathcal{F}_s] = \mathbb{E}[\langle\widetilde{Z}_\infty,X^c\rangle_t - \langle\widetilde{Z}_\infty,X^c\rangle_s|\mathcal{F}_s]$, the dominated convergence theorem implies that

$$\mathbb{E}[H(s,X)(\langle\widetilde{Z}_\infty,X^c\rangle_t - \langle\widetilde{Z}_\infty,X^c\rangle_s)] \leq E[\widetilde{Z}_\infty(s)H(s,X)\overline{\Delta}(t-s)].$$

It thus follows that

$$E\left[\int_0^T H(s,X)\mathrm{d}\langle\widetilde{Z}_\infty,X^c\rangle_s\right] \leq E\left[\int_0^T \widetilde{Z}_\infty(s)H(s,X)\overline{\Delta}\mathrm{d}s\right].$$

By an approximation argument it follows that the previous identity is valid for any bounded predictable function H. As a consequence, it follows that $s\mapsto\langle\widetilde{Z}_\infty,X\rangle_s$ is absolutely continuous with respect to the Lebesgue measure, and the Radon-Nikodym derivative satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\widetilde{Z}_\infty,X^c\rangle_t \leq \widetilde{Z}_\infty(t)\overline{\Delta}, \qquad P\times\mathrm{d}t \text{ a.e.} \tag{9.8}$$

An analogous line of reasoning, starting from the identities in Eqn. (9.5), (9.6) and (9.7) leads to the following identities, $P\times\mathrm{d}t$ a.e:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\widetilde{Z}_\infty,X^c\rangle_t \quad\geq\quad -\widetilde{Z}_\infty(t)\underline{\Delta}, \tag{9.9}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\widetilde{Z}_\infty,\overline{\mu}(A,\cdot)\rangle_t \quad\leq\quad \widetilde{Z}_\infty(t)[\Gamma_+(\Lambda(A)) - \Lambda(A)], \tag{9.10}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\widetilde{Z}_\infty,\overline{\mu}(A,\cdot)\rangle_t \quad\geq\quad \widetilde{Z}_\infty(t)[\Gamma_-(\Lambda(A)) - \Lambda(A)], \tag{9.11}$$

for any closed subset $A\subset\mathbb{R}\backslash\{0\}$, where $\overline{\mu}$ denotes the compensated Poisson random measure associated to the jumps of X.

The martingale representation theorem implies that $\widetilde{M}_\infty$ admits a representing pair $(H_\infty, U_\infty - 1)$. Combining this representation with Eqns. (9.8)–(9.11) yields that $\sigma^2 H_\infty \in [\sigma^2\Delta_-, \sigma^2\Delta_+]$ and $\int_A(U_\infty(s,x)-1)\Lambda(\mathrm{d}x) \in$

$[\Gamma_-(\Lambda(A))-\Lambda(A),\Gamma_+(\Lambda(A))-\Lambda(A)]$, $P\times \mathrm{d}t$-a.s. By the monotone convergence theorem these identities remain valid for any Borel set $A\subset\mathbb{R}\backslash\{0\}$ with $\Lambda(A)<\infty$.

Thus, we have that Q_∞, the measure with Radon-Nikodym derivative Z_∞ with respect to P, is contained in $\mathcal{D}_{\Delta,\Gamma}$. ☐

10. Proof of convergence: lower bound

To establish the limit in Eqn. (8.4) in Proposition 3 the convergence is first proved for the supremum over the "nice" subset $\widetilde{\mathcal{D}}_{\Delta,\Gamma}$ of $\mathcal{D}_{\Delta,\Gamma}$ given by

$$\widetilde{\mathcal{D}}_{\Delta,\Gamma}=\left\{Q\in\mathcal{D}_{\Delta,\Gamma}:\begin{array}{l}\frac{\mathrm{d}Q}{\mathrm{d}P}|_{\mathcal{F}_T}=\mathbf{Exp}(M^Q)_T\text{ where}\\ M^Q\text{ has representing pair }(H^Q,U^Q-1)\\ \text{with }H^Q\in C([0,T]\times\Omega,[-\Delta_-,\Delta_+]),\text{ and}\\ U^Q\in C([0,T]\times\mathbb{R}_*\times\Omega,\mathbb{R}_+),\text{ such that}\\ (\Gamma_--\mathrm{id})(\Lambda(A))\le\int_A(U_t^Q(x)-1)\Lambda(\mathrm{d}x)\le(\Gamma_+-\mathrm{id})(\Lambda(A))\\ \text{for all }A\in\mathcal{B}^\Lambda(\mathbb{R})\qquad(**)\end{array}\right\},\tag{10.1}$$

where as before $\mathbf{Exp}(M^Q)$ denotes the stochastic exponential of M^Q and we recall that $\mathcal{B}^\Lambda(\mathbb{R})=\{B\in\mathcal{B}(\mathbb{R}):\Lambda(B)<\infty\}$. For any $\eta\in[0,1)$ denote by $\widetilde{\mathcal{D}}^\eta_{\Delta,\Gamma}$ the related set defined by the right-hand side of Eqn. (10.1) with $\Delta_+,\Delta_-,\Gamma_+-\mathrm{id},\Gamma_--\mathrm{id}$ replaced by $\Delta_+(1-\eta)$, $\Delta_-(1-\eta)$, $(\Gamma_+-\mathrm{id})(1-\eta)$, $(\Gamma_--\mathrm{id})(1-\eta))$, respectively.

A key step in the proof is the fact that $\widetilde{\mathcal{D}}_{\Delta,\Gamma}$ forms a dense subset of $\mathcal{D}_{\Delta,\Gamma}$:

Lemma 16. *The convex hull of $\widetilde{\mathcal{D}}_{\Delta,\Gamma}$ is dense in $\mathcal{D}_{\Delta,\Gamma}$, in the sense of weak convergence. In particular, we have*

$$\sup_{Q\in\widetilde{\mathcal{D}}_{\Delta,\Gamma}}\mathbb{E}^Q[F(X)]=\sup_{Q\in\mathcal{D}_{\Delta,\Gamma}}\mathbb{E}^Q[F(X)].$$

Lemma 16 is proved by a standard randomisation argument that is an adaptation to the current setting of the ones that have been used in [20, 30] to derive a corresponding result in the setting of the Wiener space, and is reported in the Appendix.

Proof of the identity in Eqn. (8.4) *in Proposition 3.* Note that the supremum over $\widetilde{\mathcal{D}}_{\Delta,\Gamma}$ on the right-hand side of (8.4) is equal to the supremum over all $\mathcal{D}_{\Delta,\Gamma}$ in view of Lemma 16. Hence, to establish the identity it suffices to prove the inequality in Eqn. (8.4) with $\mathcal{D}_{\Delta,\Gamma}$ replaced by $\widetilde{\mathcal{D}}_{\Delta,\Gamma}$.

Let $\epsilon>0$ and denote by $Q_\epsilon\in\widetilde{\mathcal{D}}_{\Delta,\Gamma}$ an ϵ-optimal solution,

$$\sup_{Q\in\widetilde{\mathcal{D}}_{\Delta,\Gamma}}\mathbb{E}^Q[F(X)]\le\mathbb{E}^{Q_\epsilon}[F(X)]+\epsilon.$$

Let Z_ϵ denote the martingale given by $Z_\epsilon(t)=\mathbb{E}[Z^{Q_\epsilon}_\infty|\mathcal{F}_t]$ where $Z^{Q_\epsilon}_\infty$ denotes the Radon-Nikodym derivative of Q_ϵ with respect to $\mathbb{P}$. Define $H^{\epsilon,\eta}:=\min\{\max\{H^{Q_\epsilon}\wedge(-\Delta_-)(1-\eta)\},\Delta_+(1-\eta)\}$ and

$$U^{\epsilon,\eta}:=U^{Q_\epsilon}-(U^{Q_\epsilon}-1)[\eta\mathbf{1}_{\{U^{Q_\epsilon}>1/(1-\eta)\}}+\mathbf{1}_{\{1\le U^{Q_\epsilon}\le1/(1-\eta)\}}]-\eta(U^{Q_\epsilon}-1)\mathbf{1}_{\{U^{Q_\epsilon}<1\}}.$$

It is straightforward to check that $U^{\epsilon,\eta}$ is non-negative (using that U^{Q_ϵ} is nonnegative), and that we have $H^{\epsilon,\eta}\in\mathcal{L}^2$ and $U^{\epsilon,\eta}\in\widetilde{\mathcal{L}^2}$. Let $Q^{\epsilon,\eta}$ denote the probability measure wih Radon-Nikodym derivative $Z^{\epsilon,\eta}$ given by the stochastic exponential of the martingale $M^{\epsilon,\eta}$ given by

$$M_t^{\epsilon,\eta} = \int_0^t H_s^{\epsilon,\eta} \mathrm{d}X_s^c + \int_{[0,t]\times\mathbb{R}} (U^{\epsilon,\eta}(s,x,X)-1)\widetilde{\mu}^X(\mathrm{d}s\times \mathrm{d}y), \qquad t\in[0,T].$$

Also consider the sequence of martingales $M^{\delta,\epsilon,\eta}$ given by $M_0^{\delta,\epsilon,\eta}=0$ and

$$\Delta M_k^{\delta,\eta} = H^{\epsilon,\eta}_{(k-1)\delta}\Delta R_k^c + \sum_{z_l\in h\mathbb{Z}} (U^{\epsilon,\eta}((k-1)\delta, z_l, R)-1)[I_l^k - p_l], \qquad k\in\mathbb{N},$$

where $\Delta M_k^{\delta,\eta} = M_k^{\delta,\eta} - M_{k-1}^{\delta,\eta}$, $R=R^\delta$ and $R^c = R^{\delta,c}$, and denote the embedding $\widetilde{M}^{\delta,\epsilon,\eta}$ of $M^{\delta,\epsilon,\eta}$ into the Skorokhod space $D[0,T]$ by $\widetilde{M}_t^{\delta,\epsilon,\eta} = M^{\delta,\epsilon,\eta}_{\lfloor t/\delta\rfloor}$ for $t\in[0,T]$.

To complete the proof we will use the following auxiliary result (the proof of which is provided at the end of the section):

Lemma 17. *Let F be as in Theorem 1. The following hold true:*

(i) *The measure $Q^{\epsilon,\eta}$ is element of $\widetilde{\mathcal{D}}^\eta_{\Delta,\Gamma}$, and*

$$\mathbb{E}^{Q^{\epsilon,\eta}}[F(X)] \to \mathbb{E}^{Q^\epsilon}[F(X)], \qquad \text{when } \eta\searrow 0. \tag{10.2}$$

(ii) *The measure $Q^{\delta,\epsilon,\eta}$ with Radon-Nikodym derivative $\mathbf{Exp}(M^{\delta,\epsilon,\eta})$ is element of $\mathcal{D}^\delta$ for all δ sufficiently small.*

(iii) *$\widetilde{M}^{\delta,\epsilon,\eta}$ weakly converges to $M^{\epsilon,\eta}$ in the Skorokhod topology, as δ tends to zero.*

Given Lemma 17 the proof is completed as follows. By a line of reasoning that is analogous to the one used in the proof of Lemma 3 we have, as $n\to\infty$

$$E^{Q^{\delta_n,\eta}}[F_{\delta_n}(R^{\delta_n})] \to \mathbb{E}^{Q^{\eta,\epsilon}}[F(X)],$$

so that

$$\liminf_{n\to\infty} \sup_{Q\in\mathcal{D}^{\delta_n}} E^Q[F_{\delta_n}(R^{\delta_n})] \geq \mathbb{E}^{Q^{\eta,\epsilon}}[F(X)].$$

Letting $\eta\to 0$ we find in view of Eqn. (10.2)

$$\liminf_{n\to\infty} \sup_{Q\in\mathcal{D}^{\delta_n}} E^Q[F_{\delta_n}(R^{\delta_n})] \geq \mathbb{E}^{Q^\epsilon}[F(X)] \geq \sup_{Q\in\widetilde{\mathcal{D}}_{\Delta,\Gamma}} \mathbb{E}^Q[F(X)] - \epsilon.$$

Since $\epsilon>0$ was arbitrary, the proof is complete. □

Proof of Lemma 17. (i) The measure $Q^{\epsilon,\eta}$ is contained in the set $\widetilde{\mathcal{D}}^\eta$ in view of (a) the definition of $H^{\epsilon,\eta}$ and (b) the observation (from the definition of $U^{\epsilon,\eta}$)

$$\begin{aligned}\Lambda^{Q_{\epsilon,\eta}}(A) - \Lambda(A) &\leq (1-\eta)[\Lambda^{Q_\epsilon}-\Lambda](A\cap\{U^{Q_\epsilon}<1\}) + [\Lambda^{Q_\epsilon}-\Lambda](A\cap\{U^{Q_\epsilon}>1\})\\ &\geq (1-\eta)[\Lambda^{Q_\epsilon}-\Lambda](A\cap\{U^{Q_\epsilon}>1\}) + [\Lambda^{Q_\epsilon}-\Lambda](A\cap\{U^{Q_\epsilon}<1\}),\end{aligned}$$

for any $A\in\mathcal{B}(\mathbb{R})$ with $\Lambda(A)<\infty$, from which we deduce

$$(1-\eta)(\Gamma_- - \mathrm{id})(\Lambda(A)) \leq \Lambda^{Q_{\epsilon,\eta}}(A) - \Lambda(A) \leq (1-\eta)(\Gamma_+ - \mathrm{id})(\Lambda(A)).$$

When η tends to zero, $M^{\epsilon,\eta}$ converges to M^ϵ in L^2:

$$\begin{aligned}
&\mathbb{E}\left[(M^\epsilon_T - M^{\epsilon,\eta}_T)^2\right]\\
&= \mathbb{E}\left[\int_0^T (H^{Q^\epsilon}(s,X) - H^{\epsilon,\eta}(s,X))^2\sigma^2 \mathrm{d}s + \int_{[0,T]\times\mathbb{R}} (U^{Q^\epsilon}(s,x,X) - U^{\epsilon,\eta}(s,x,X))^2 \Lambda(\mathrm{d}x)\mathrm{d}s\right]\\
&\le \eta^2\sigma^2 T(\Delta_+ \vee \Delta_-)^2 + \mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} (U^{Q_\epsilon}_s(x) - 1)^2[\eta^2 + \mathbf{1}_{\{0\le U^\epsilon_s(x)-1\le\eta/(1-\eta)\}}]\Lambda(\mathrm{d}x)\mathrm{d}s\right],
\end{aligned}$$

which tends to zero as η tends to zero, by an application of Lebesgue's Dominated Convergence Theorem. Since the Radon-Nikodym derivatives $Z^{\epsilon,\eta}_T$ that are equal to the stochastic exponentials of the martingales $M^{\epsilon,\eta}$ converge weakly to a limit and are contained in a ball in L^2 (by Lemma 5) and $F(X)$ is square integrable (by Lemma 8) it follows $\mathbb{E}^{Q^{\epsilon,\eta}}[F(X)] \to \mathbb{E}^{Q^\epsilon}[F(X)]$ as $\eta \to 0$.

(ii) We will show that $Q^{\delta,\eta} \in \mathcal{D}^{\delta,-\eta}_*$, which establishes the assertion since $\mathcal{D}^{\delta,-\eta}_*$ is contained in $\mathcal{D}^\delta$ for all δ sufficiently small (by Lemma 13).

Let $\sigma^2 > 0$. In view of the relation between h and δ, the definition of the set $\widetilde{\mathcal{D}}^\eta_{\Delta,\Gamma}$ and the value of $p_{\pm1}$ we have for $i = \pm1$

$$\begin{aligned}
E[(\Delta M^{\delta,\epsilon,\eta}_m)(\Delta R^{(i)}_m)|\mathcal{G}_{m-1}] &= H^{\epsilon,\eta}_{(m-1)\delta} E[(\Delta R^{(i)}_m)(\Delta R^c_m)|\mathcal{G}_{m-1}]\\
&\le (1-\eta)h^2 p_i(1-p_i)\cdot\Delta_+\\
&= (1-\eta)\Delta_+[\frac{5}{6}\cdot\frac{1}{6}\cdot 3\delta\sigma^2 + o(\delta)] \qquad \text{when } \delta \searrow 0,
\end{aligned}$$

which is bounded above by $(1-\eta/2)\Delta_+[\frac{5}{12}\cdot\delta\sigma^2]$ for all δ sufficiently small. Analogously, it follows that

$$E[(\Delta M^{\delta,\epsilon,\eta}_m)(\Delta R^{(i)}_m)|\mathcal{G}_{m-1}] \ge -(1-\eta/2)\Delta_-[\frac{5}{12}\cdot\delta\sigma^2 + o(\delta)], \qquad \text{when } \delta \searrow 0.$$

Let $A \in \mathcal{B}(\mathbb{R}\backslash\{0\})$ be compact. In view of the definitions of $M^{\delta,\epsilon,\eta}$ and Z^A we find

$$\begin{aligned}
E[\Delta M^{\delta,\epsilon,\eta}_m \Delta Z^A_m|\mathcal{G}_{m-1}] &= \sum_{z_l\in A}(U^\epsilon_{(m-1)\delta}(z_l) - 1)E\left[(I^m_l - p_l)^2|\mathcal{G}_{m-1}\right]\\
&= \sum_{z_l\in A}(U^\epsilon_{(m-1)\delta}(z_l) - 1)p_l(1-p_l)\\
&= \delta\sum_{z_l\in A, z_l>0}(U^\epsilon_{(m-1)\delta}(z_l) - 1)\Lambda([z_l, z_{l+1})) + \delta\sum_{z_l\in A, z_l<0}(U^\epsilon_{(m-1)\delta}(z_l) - 1)\Lambda((z_{l-1}, z_l]) + o(\delta)\\
&= \delta\int_A (U^\epsilon_{(m-1)\delta}(x) - 1)\Lambda(\mathrm{d}x) + o(\delta), \qquad (10.3)\\
&\le \delta(1-\eta)(\Gamma_+ - \mathrm{id})(\Lambda(A)) + o(\delta), \qquad \text{when } \delta \searrow 0
\end{aligned}$$

which is bounded above by $\delta\,(1-\eta/2)(\Gamma_+ - \mathrm{id})(\Lambda(A))$ for all δ sufficiently small. In the one but last line we used that since $x \mapsto U^\epsilon_{(m-1)\delta}(x)$ is uniformly continuous on the compact set A, it follows that the Riemann-Stieltjes sum converges to the integral as δ (and hence the spatial mesh h) converge to zero.

Analogously it follows that for all δ sufficiently small

$$E[\Delta M^{\delta,\epsilon,\eta}_m \Delta Z^A_m|\mathcal{G}_{m-1}] \ge \delta(1-\eta/2)(\Gamma_- - \mathrm{id})(\Lambda(A)).$$

Combining the previous displays, we see that $Q^{\epsilon,\eta}$ is contained in $\mathcal{D}^{\delta,-\eta}_*$.

(iii) The convergence follows by comparison of the semi-martingale characteristics $(B_t^\delta, \widetilde{C}_t^\delta, \Lambda^\delta)$ of $\widetilde{M}^{\delta,\epsilon,\eta}$ with the characteristics $(B_t, \widetilde{C}_t, \Lambda)$ of M^ϵ, analogously as in Lemma 7. We have weak convergence of $\widetilde{M}^{\delta_n,\epsilon,\eta}$ to $M^{\epsilon,\eta}$ if the conditions (a)–(c) in Eqn. (C.1) hold for any $t \in [0,T]$, which are verified as follows:

(a) The first characteristics of either martingale are given by

$$\begin{aligned} B_t &= -\int_{[0,t]\times\mathbb{R}\setminus[-1,1]} x(U^Q(s,x,X)-1)\Lambda(\mathrm{d}x)\mathrm{d}s, \\ B_t^{\delta_n} &= -\sum_{k\geq 1: t_k\leq t}\sum_{|l|>1} z_l(U^Q(t_{k-1},z_l,Y^{\delta_n})-1)p_l, \end{aligned}$$

where $p_l = \delta\Lambda([z_l, z_{l+1}))$ if $l > 1$ and $p_l = \delta\Lambda((z_{l-1}, z_l])$ if $l < -1$. The continuity of $(s,x,\omega) \mapsto U_Q(s,x,\omega)$ implies that the sum $B_t^{\delta_n}$ converges to B_t as $n \to \infty$, uniformly in t, that is, $\sup_{s\in[0,T]} |B_s - B_s^{\delta_n}| \to 0$ if $n \to \infty$.

(b) The continuity of the functions $(s,x,\omega) \mapsto U_Q(s,x,\omega)$ and $(s,\omega) \mapsto H_Q(s,\omega)$ and the fact that $(p_1 + p_{-1})h^2$ tends to σ^2 in the case $\delta \searrow 0$ yield that the sum $\widetilde{C}_t^{\delta_n}$ tends to the integral $\widetilde{C}_t$ as $n \to \infty$, where

$$\begin{aligned} \widetilde{C}_t &= \int_0^t \left\{\sigma^2 H^Q(s,X)^2 + \int_{\mathbb{R}} (U^Q(s,x,X)-1)^2\Lambda(\mathrm{d}x)\right\}\mathrm{d}s, \\ \widetilde{C}_t^{\delta_n} &= \sum_{k\geq 1: t_k\leq t}\left\{(p_1+p_{-1})h^2 H^Q(t_k,Y^{\delta_n})^2 + \sum_{z_l}(U^Q(t_k,z_l,Y^{\delta_n})-1)^2 p_l\right\}. \end{aligned}$$

(c) Finally, for any $g \in C_0(\mathbb{R})$ we have that the sum

$$\sum_{k\geq 1: t_k\leq t}\left\{\sum_{|l|>1} g(z_l)U^Q(t_k,z_l,Y^{\delta_n})p_l + g(x_1)p_1U^Q(t_k,z_1,Y^{\delta_n}) + g(x_{-1})p_{-1}U^Q(t_k,z_{-1},Y^{\delta_n})\right\}$$

converges to the integral $\int_{[0,t]\times\mathbb{R}} g(x)(U^Q(s,x,X)-1)\Lambda(\mathrm{d}x)\mathrm{d}s$ as $n \to \infty$, using again the continuity of U_Q and the fact that the final to terms tend to zero since we have $x_{\pm 1} \to 0$ and that g is zero in a neighbourhood of zero.

The proof of (iii) is complete by an application of [25, Thm. VII.2.17]. □

Appendix A. Proof of representation of Choquet expectation

The proof of the representation rests on the identification of the absolutely continuous measure that attains the maximum in Eqn. (2.8). For any $\mathcal{X} \in L^2_+(\mu)$ denote by $m_\mathcal{X}$ the measure $m_\mathcal{X} \in \mathcal{M}^{ac}_{1,\mu}$ with Radon-Nikodym derivative $\frac{\mathrm{d}m_\mathcal{X}}{\mathrm{d}\mu} = \phi^D(\mathcal{X})$ where $\phi^D : \mathbb{R}_+ \to \mathbb{R}_+$ is given by

$$\phi^D(x) = \begin{cases} \frac{D(\overline{F}(x))-D(\overline{F}(x-))}{\overline{F}(x)-\overline{F}(x-)}, & \text{in the case that } x \text{ is such that } \overline{F}(x)-\overline{F}(x-)>0 \\ & \quad \text{and either } x>0 \text{ or } F(0-)<\infty \text{ and } x=0, \\ D'(\overline{F}(x)), & \text{in the case that } x \text{ is such that } \overline{F}(x)-\overline{F}(x-)=0 \text{ and } x>0, \\ 0, & \text{otherwise,} \end{cases} \tag{A.1}$$

with $\overline{F}(x) = \mu(u \in \mathbb{R} : \mathcal{X}(u) > x)$, $\overline{F}(0-) = \mu(\mathbb{R}_+)$, $\overline{F}(x-)$ denoting the left-limit of $\overline{F}$ at $x > 0$ and D' the right-derivative of D (where $D'(1)$ denotes the left-derivative at $x = 1$ in case D is a probability distortion). We next verify that the measure $m_\mathcal{X}$ is element of $\mathcal{M}^D$.

Lemma 18. *(i) The measure $m_{\mathcal{X}}$ satisfies*

$$m_{\mathcal{X}}(A) = -\int_B \mathrm{d}D(\overline{F}(y)) \qquad \textit{for } A \in \mathcal{B}^{\mu}(\mathbb{R}),\ A = \mathcal{X}^{-1}(B),\ B \in \mathcal{B}(\mathbb{R}). \tag{A.2}$$

(ii) We have $m_{\mathcal{X}}(A) \le D(\mu(A))$ for all $A \in \mathcal{B}(\mathbb{R})$. In particular, $m_{\mathcal{X}} \in \mathcal{M}_2^D$.

Proof of Lemma 18. (i) For any set $A \in \mathcal{B}(\mathbb{R})$ of the form $A = \mathcal{X}^{-1}(B)$ with $B \in \mathcal{B}(\mathbb{R})$, the definitions of $m_{\mathcal{X}}$ and $\phi = \phi^D$ imply

$$m_{\mathcal{X}}(A) = \mu(\phi(\mathcal{X})I_A) = \int_B \phi(x)(\mu \circ \mathcal{X}^{-1})(\mathrm{d}x) = -\int_B \phi(x)\mathrm{d}\overline{F}(x) = -\int_B \mathrm{d}D(\overline{F}(x)).$$

(ii) The proof of the stated inequality consists of two steps. The first step is to show that the inequality holds for all $A \in \mathcal{B}(\mathbb{R})$ of the form $A = \mathcal{X}^{-1}(B)$ for some $B \in \mathcal{B}(\mathbb{R})$. If D is piecewise linear, then it is straightforward to verify that the stated inequality holds, as a consequence of Eqn. (A.2) and the fact that the right-derivative D' is decreasing (as D is concave). The general case follows by first approximating D by piecewise linear concave distortions and subsequently passing to the limit, using the Monotone Convergence Theorem.

The second step consists in first associating to any Borel set A with $m_{\mathcal{X}}(A) > 0$ the set $A' = \mathcal{X}^{-1}(\mathcal{X}(A))$ and the ratio $\theta(A) = m_{\mathcal{X}}(A)/m_{\mathcal{X}}(A')$, and subsequently observing that we have

$$m_{\mathcal{X}}(A) = \theta(A)m_{\mathcal{X}}(A') \le \theta(A)D(\mu(A')) \le D(\mu(A)),$$

where the first inequality follows from the first step and the second equality from the concavity of D and the fact $\theta(A') \le 1$ as we have $A \subset A'$. This completes the proof of the first line of part (ii). We conclude $m_{\mathcal{X}} \in \mathcal{M}_2^D$, by noting that (a) Lemma 3 implies $m_{\mathcal{X}} \in L^2(\mu)$ and (b) $m_{\mathcal{X}}$ satisfies the bound stated in the first part of (ii). □

Proof of Proposition 1. Consider first the case $\mathcal{X} \in L^2_+(\mu)$. Finiteness of $\mathcal{C}^D[\mathcal{X}]$ follows from Lemma 1. The definitions in Eqns. (2.4), (2.5) and (2.9) of the Choquet integral $\mathcal{C}^D(\mathcal{X})$ and the set $\mathcal{M}^D$ imply that for any $m \in \mathcal{M}^D$ we have

$$\mathcal{C}^D[\mathcal{X}] \ge \int_0^\infty m(\mathcal{X} > x)\mathrm{d}x, \tag{A.3}$$

so that $\mathcal{C}^D[\mathcal{X}] \ge \sup_{m \in \mathcal{M}^D} m(\mathcal{X})$. Equality in Eqn. (A.3) is obtained by the measure $m = m_{\mathcal{X}}$, as we show next: The definitions of $m_{\mathcal{X}}$ and ϕ and a change of the order of integration show

$$m_{\mathcal{X}}[\mathcal{X}] = \mu(\phi(\mathcal{X})\mathcal{X}) = -\int_{[0,\infty)} x\mathrm{d}D(\overline{F}(x)) = \mathcal{C}^D[\mathcal{X}].$$

The proof of the case $\mathcal{X} \in L^2_+(\mu)$ is complete noting $m_{\mathcal{X}} \in \mathcal{M}_2^D$ by Lemma 18. □

Appendix B. Proofs of the construction of the random walk

Proof of Lemma 6: (i) The first assertion follows from the fact that under the conditions on the triplet (δ, h, a) in Eqns. (5.4)–(5.5) there exists a solution (p_{-1}, p_0, p_1) in the set $\mathbb{S} = \{(p_{-1}, p_0, p_1) \in [0,1]^3 : p_{-1} + p_0 + p_1 \le 1\}$ of the following system of three linear equations:

$$p_0 + p_{-1} + p_1 = \alpha(h) \tag{B.1}$$

$$-p_{-1}h + p_1 h = \beta(h) \tag{B.2}$$

$$p_{-1}h^2 + p_1 h^2 = \gamma(h), \tag{B.3}$$

where

$$\alpha(h) := 1 - \sum_{k:|k|\geq a} p_k = 1 - \delta\Lambda(\mathbb{R}\backslash(-ah, ah)) \tag{B.4}$$

$$\beta(h) := -\delta\left(\sum_{k\geq a} kh\Lambda([kh,(k+1)h)) + \sum_{k\leq -a} kh\Lambda(((k-1)h, kh])\right) \tag{B.5}$$

$$\gamma(h) := \delta(\sigma^2 + \Sigma^2(\mathbb{R})) - \delta\left(\sum_{k\geq a}(kh)^2\Lambda([kh,(k+1)h)) + \sum_{k\leq -a}(kh)^2\Lambda(((k-1)h, kh])\right), \tag{B.6}$$

where we used the form of p_k for $|k| \geq a$ given in Eqn. (5.1) .

It is easily verified that this system admits a unique solution in $\mathbb{R}^3$ which is given by

$$p_{\pm 1}(h) = \frac{1}{2}\left(\frac{\gamma(h) \pm \beta(h)h}{h^2}\right), \qquad p_0(h) = \alpha(h) - p_1(h) - p_{-1}(h). \tag{B.7}$$

We see from Eqn. (B.7) that the vector (p_{-1}, p_0, p_1) lies in the unit cube $\mathbb{S}$ if and only if we have the bounds $|\beta(h)|/h \leq \gamma(h)/h^2 \leq \alpha(h) \leq 1$. That these three inequalities are satisfied can be verified from the following estimates:

Lemma 19. *Under the assumptions of Lemma 6 the following estimates hold:*

$$1 - \frac{1}{3a^2}\frac{\Sigma^2(\mathbb{R})}{\widetilde{\Sigma}^2(a)} \leq \alpha(h) \leq 1, \qquad \frac{|\beta(h)|}{h} \leq \frac{1}{3a}\frac{\Sigma^2(\mathbb{R})}{\widetilde{\Sigma}^2(a)}, \tag{B.8}$$

$$0 \leq \frac{\gamma(h)}{h^2} - \frac{\sigma^2 + \Sigma^2(-ah, ah)}{3\widetilde{\Sigma}^2(a)} \leq \frac{\Sigma^2(\mathbb{R})(a^{-2} + 2a^{-1})}{3\widetilde{\Sigma}^2(a)}, \tag{B.9}$$

where $\widetilde{\Sigma}^2(a) = \sigma^2 + \Sigma^2(\mathbb{R})[a^{-1}1_{\{\sigma^2>0\}} + 1_{\{\sigma^2=0\}}]$. *In particular,* $|\beta(h)|/h \leq \gamma(h)/h^2 \leq \alpha(h)$.

Moreover, in the case $\sigma^2 > 0$ *we have*

$$-\frac{\Sigma^2(\mathbb{R})}{3a\sigma^2} \leq \frac{\gamma(h)}{h^2} - \frac{1}{3} \leq \frac{\Sigma^2(-ah, ah) + 2a^{-1}\Sigma^2(\mathbb{R})}{3a\sigma^2}, \qquad \frac{|\beta(h)|}{h} \leq \frac{\Sigma^2(\mathbb{R})}{3a\sigma^2}. \tag{B.10}$$

The proof of this lemma is given at the end of this section.

(ii) To verify the second assertion, we note that, in view of Eqn. (B.7), it suffices to show that, as $h \to 0+$, we have (i') $\alpha(h) \to 1$, (ii') $\beta(h)/h \to 0$ and (iii') $\gamma(h)/h^2$ tends to $1/3$ in the case $\sigma^2 > 0$ and to zero otherwise. These three facts are verified as follows:

(i') Since we have $a(h) \to \infty$ as $h \to 0$ by assumption, Eqn. (B.8) yields $\lim_{h\searrow 0}\alpha(h) = 1$.

(ii') Since $a = a(h) \to \infty$ as $h \to 0$ and $\widetilde{\Sigma}^2(a)$ tends to a non-zero limit as $h \to 0$, namely,

$$\widetilde{\Sigma}^2(a) \equiv \Sigma^2(\mathbb{R}) \text{ if } \sigma^2 = 0 \text{ and } \widetilde{\Sigma}^2(a) \to \sigma^2 \text{ if } \sigma^2 > 0, \tag{B.11}$$

it follows from Eqn. (B.8) that $\lim_{h\searrow 0}\beta(h)/h = 0$.

(iii') Observing that a tends to infinity when h tends to 0, the right-hand side of Eqn. (B.9) converges to zero. Since by assumption $ah \to 0$ as $h \to 0$ it follows that $\Sigma^2(-ah, ah) \to 0$, and we have that the middle fraction in Eqn. (B.9) converges to $1/3$ in the case $\sigma^2 > 0$ and to zero otherwise, in view of Eqn. (B.11).

(iii) For the third assertion we note that, in view of the bound in Eqn. (B.8), it follows $\alpha(h) \geq 1 - (3a^2)^{-1}c_\sigma$, which yields the statement in Eqn. (5.9) by definition of $\alpha(h)$ given in Eqns. (B.1) and (B.4). From Eqns. (B.9) and (B.10) in Lemma 19 we have

$$p_1 \wedge p_{-1} \geq \frac{\gamma(h) - |\beta(h)|h}{2h^2} \geq \left(\frac{1}{6} - \frac{1}{6a}\frac{\Sigma^2(\mathbb{R})}{\sigma^2}\right)\frac{\sigma^2}{\widetilde{\Sigma}^2(a)}.$$

This completes the proof. □

Proof of Lemma 19. The proof rests on the following three observations:

(a) In view of the relation in Eqn. (5.4) and the fact $ah \leq 1$. it holds (with $a \in \mathbb{N}$, $a \geq 2$)

$$\begin{aligned} 0 &\leq \delta\Lambda(\mathbb{R}\backslash(-ah,ah)) \leq \frac{\delta}{(ah)^2}\int_{(-1,1)\backslash(-ah,ah)} x^2\Lambda(\mathrm{d}x) + \delta\int_{\mathbb{R}\backslash(-1,1)} x^2\Lambda(\mathrm{d}x) \\ &\leq \frac{a^{-2}\Sigma^2((-1,1)) + h^2\Sigma^2(\mathbb{R}\backslash(-1,1))}{3\widetilde{\Sigma}^2(a)} \leq \frac{1}{3a^2}\cdot\frac{\Sigma^2(\mathbb{R})}{\widetilde{\Sigma}^2(a)}. \end{aligned} \tag{B.12}$$

(b) Recalling $ah \leq 1$ and denoting $M(I) = \int_I x\Lambda(\mathrm{d}x)$ for any interval I, we have

$$\begin{aligned} |\beta(h)| &\leq \delta M(\mathbb{R}\backslash(-ah,ah)) \leq \frac{h^2 M(\mathbb{R}\backslash(-1,1))}{3\widetilde{\Sigma}^2(a)} + \frac{\delta}{ah}\int_{(-1,1)\backslash(-ah,ah)} x^2\Lambda(\mathrm{d}x) \\ &\leq h^2\cdot\frac{\Sigma^2(\mathbb{R}\backslash(-1,1))}{3\widetilde{\Sigma}^2(a)} + \frac{h}{3a}\cdot\frac{\Sigma^2((-1,1))}{\widetilde{\Sigma}^2(a)} \leq \frac{h}{3a}\cdot\frac{\Sigma^2(\mathbb{R})}{\widetilde{\Sigma}^2(a)}. \end{aligned} \tag{B.13}$$

(c) Since x^2 is increasing, the difference

$$D_h := \int_{\mathbb{R}\backslash(-ah,ah)} x^2\Lambda(\mathrm{d}x) - \left[\sum_{h\geq a}((kh)^2\Lambda([hk,h(k+1)) + \sum_{h\leq -a}((kh)^2\Lambda((h(k-1),hk])\right]$$

between integral and sum is positive and can be estimated by

$$\begin{aligned} |D_h| &\leq \sum_{k\geq a}((k+1)^2-k^2)h^2\Lambda([hk,h(k+1)) + \sum_{k\leq -a}((k-1)^2-k^2)h^2\Lambda((h(k-1),hk])] \\ &= h^2\Lambda(\mathbb{R}\backslash(-ah,ah)) + 2h\sum_{k\geq a} kh\Lambda([hk,h(k+1))\cup((-k-1)h,-hk]). \end{aligned} \tag{B.14}$$

Deploying the bounds in Eqns. (B.12), (B.13) and (B.14) we find the upper bounds

$$h^{-2}\gamma(h) \leq h^{-2}\delta\left(\sigma^2 + \Sigma^2(-ah,ah)\right) + \delta\Lambda\left(\mathbb{R}\backslash(-ah,ah)\right) + 2\frac{|\beta(h)|}{h}$$

$$\leq \frac{\sigma^2 + \Sigma^2(-ah,ah) + (2a^{-1}+a^{-2})\Sigma^2(\mathbb{R})}{3\widetilde{\Sigma}^2(a)} \tag{B.15}$$

$$= \frac{1}{3} + \frac{\Sigma^2(-ah,ah) + (a^{-1}+a^{-2})\Sigma^2(\mathbb{R})}{3\widetilde{\Sigma}^2(a)}. \tag{B.16}$$

Moreover, we have the lower bound

$$\gamma(h) \geq \delta\left\{\sigma^2 + \Sigma^2((-ah,ah))\right\} = h^2\frac{\sigma^2 + \Sigma^2((-ah,ah))}{3\widetilde{\Sigma}^2(a)}. \tag{B.17}$$

In particular, in the case $\sigma^2 > 0$, we find

$$\gamma(h) \geq \frac{h^2}{3} + h^2\frac{\Sigma^2((-ah,ah)) - a^{-1}\Sigma^2(\mathbb{R})}{3(\sigma^2 + a^{-1}\Sigma^2(\mathbb{R})}. \tag{B.18}$$

Eqn. (B.8) follows from (a) and (b), and Eqn. (B.9) from Eqns. (B.15) and (B.17) in (c). The bounds $|\beta(h)|/h \leq \gamma(h)/h^2 \leq \alpha(h)$ follow by combining the condition in Eqn (5.5) with the estimate in Eqn. (B.9). The bounds on $\beta(h)$ and $\gamma(h)$ in Eqn. (B.10) follow from the observation $\widetilde{\Sigma}^2(a) \geq \sigma^2$ (in the case $\sigma^2 > 0$) and the esimate in (b), and from Eqns. (B.16) and (B.17), respectively. □

Appendix C. Proof of weak convergence (Lemma 7)

As the weak convergence of each of the three sequences of processes follows by similar arguments, we shall detail only the proof of the weak convergence of Y^δ. Denote by $(B^\delta, \widetilde{C}^\delta, \nu^\delta)$ and $(B, \widetilde{C}, \nu)$ the semi-martingale characteristics of Y^δ and X respectively (see [25] for a definition). According to [25, Thm. VII.2.17], weak convergence of Y^{δ_n} to X in Skorokhod topology as $n \to \infty$ follows when the semi-martingale characteristics of Y^{δ_n} converge to those of X in the sense that the following limits hold for any $t \in [0, T]$ when $n \to \infty$:

$$(a)\ \sup_{s\leq t} |B_s^{\delta_n} - B_s| \to 0, \quad (b)\ \widetilde{C}_t^{\delta_n} \to \widetilde{C}_t, \tag{C.1}$$

$$(c) \int_{\mathbb{R}} g(x)\nu_t^{\delta_n}(\mathrm{d}x) \to t\int_{\mathbb{R}} g(x)\Lambda(\mathrm{d}x) \text{ for all } g \in C_0(\mathbb{R}), \tag{C.2}$$

where we used the fact $\nu(\mathrm{d}x \times \mathrm{d}t) = \Lambda(\mathrm{d}x)\mathrm{d}t$ and denoted $\nu_t^{\delta_n}(\mathrm{d}x) = \int_{[0,t]} \nu^{\delta_n}(\mathrm{d}s \times \mathrm{d}x)$, and used $C_0(\mathbb{R})$ to denote the set of real-valued continuous functions on $\mathbb{R}$ that vanish in a neighbourhood of 0 and have a limit at infinity. Next we show that the conditions (a)–(c) are satisfied.

(a) Observe that $B_s = s(\psi'(0) - \gamma')$ where $\gamma' = \int_{|x|>1} x\nu(\mathrm{d}x)$ while by definition

$$B_s^\delta = \mathtt{d}s + \left\lfloor \frac{s}{\delta} \right\rfloor (E[Z_1] - E[Z_1 1_{|Z_1|>1}]) = \mathtt{d}s + \left\lfloor \frac{s}{\delta} \right\rfloor \left(E[X_\delta - \mathtt{d}\delta] - \sum_{k:|z_k|>1} \nu[z_k, z_{k+1}) z_k \right) + cs,$$

tends to $B_s = s(\psi'(0) - \gamma')$ as $\delta \searrow 0$, uniformly in s.

(b) $\widetilde{C}_t^\delta$ tends to $t\Sigma^2 = \widetilde{C}_t$ as $\delta \searrow 0$ since we have

$$\widetilde{C}_t^\delta = \left\lfloor \frac{t}{\delta} \right\rfloor (E[Z_1^2] - E[Z_1]^2) = \left\lfloor \frac{t}{\delta} \right\rfloor \mathrm{Var}[Z_1] = \left\lfloor \frac{t}{\delta} \right\rfloor \mathrm{Var}[X_\delta] = \left\lfloor \frac{t}{\delta} \right\rfloor \delta(\sigma^2 + \Sigma^2(\mathbb{R})).$$

(c) Observing that $g(h)$ and $g(-h)$ are equal to zero for h (and hence δ) sufficiently small, we have that

$$\int_{\mathbb{R}} g(x)\nu_t^\delta(\mathrm{d}x) = \left\lfloor \frac{t}{\delta} \right\rfloor \left(\sum_{k:|z_k|>ah} g(z_k)\delta\Lambda[z_k, z_{k+1}) + p_1 g(h) + p_{-1} g(-h) \right)$$

converges to $t\int_{\mathbb{R}} g(x)\Lambda(\mathrm{d}x)$ as $\delta \searrow 0$. □

Appendix D. Proof of the form of the driver g of the g-expectation

In the following it is verified that the function $g_{\Delta,\Gamma}$ satisfies the conditions of a driver function.

Lemma 20. *Let D and G be a drift-shift and a jump-distortion. Then $g_{\Delta,\Gamma}$ is a driver function, that is, (i) $g_{\Delta,\Gamma}(0,0) = 0$ and (ii) $g_{\Delta,\Gamma}$ is Lipschitz continuous: there exists a $K > 0$ such that we have*

$$|g_{\Delta,\Gamma}(z_1, v_1) - g_{\Delta,\Gamma}(z_2, v_2)|^2 \leq K \left\{ |z_1 - z_2|^2 + \int_{\mathbb{R}} |v_1(x) - v_2(x)|^2 \Lambda(\mathrm{d}x) \right\}$$

uniformly for all $z_1, z_2 \in \mathbb{R}$ and all $v_1, v_2 \in L^2(\Lambda)$.

Proof. Item (i) follows from the fact $\mathcal{C}^{\Gamma_+-\mathrm{id}}(0) = \mathcal{C}^{\mathrm{id}-\Gamma_-}(0) = 0$. Item (ii) is a consequence of the form of $g_{\Delta,\Gamma}$ in Eqn. (4.6) in combination with the observation (from the representation in Proposition 1) that we have for a measure distortion D and any $v_1, v_2 \in L^2(\lambda)$

$$|\mathcal{C}^D(v_1^+) - \mathcal{C}^D(v_2^+)|^2 \le \mathcal{C}^D(|v_1 - v_2|)^2 \le K_D \int_{\mathbb{R}_+} |v_1(x) - v_2(x)|^2 \Lambda(\mathrm{d}x), \tag{D.1}$$

where $K_D = \int_0^\infty D(y) \frac{\mathrm{d}y}{y\sqrt{y}}$ which is finite in view of the integrability conditions in Eqn. (2.6)–(**??**). The first and second inequalities in Eqn. (D.1) in turn follow from the representation in Proposition 1 and the estimate in Eqn. (2.7). □

The existence and uniqueness theorem in [4] implies that there exists a unique triplet $(Y^{\mathcal{X}}, Z^{\mathcal{X}}, V^{\mathcal{X}})$ with $Y^{\mathcal{X}} \in \mathcal{L}^2$, $Z^{\mathcal{X}} \in \mathcal{L}^2$ and $V^{\mathcal{X}} \in \widetilde{\mathcal{L}}^2$ that solves the backward stochastic differential equation given by

$$-\,\mathrm{d}Y_t^{\mathcal{X}} = g_{D,G}(Z_t^{\mathcal{X}}, V_t^{\mathcal{X}})\mathrm{d}t - Z_t^{\mathcal{X}}\mathrm{d}X_t^c - \int_{\mathbb{R}} V_t^{\mathcal{X}}(x)\widetilde{\mu}^X(\mathrm{d}t \times \mathrm{d}x), \qquad t \in [0,T), \tag{D.2}$$

$$Y_T^{\mathcal{X}} = \mathcal{X}. \tag{D.3}$$

The value $Y_0^{\mathcal{X}}$ (and the random variable $Y_t^{\mathcal{X}}$) are called the $g_{\Delta,\Gamma}$-expectation (and $\mathcal{F}_t$-conditional $g_{\Delta,\Gamma}$-expectation) of the random variable $\mathcal{X}$.

Proposition 4. *Let D and G be a drift-shift and a jump-distortion, and $\mathcal{X} \in L^2(\Omega, \mathcal{F}_T, P)$. Then $\mathcal{E}_{\Delta,\Gamma}(\mathcal{X})$ is a g-expectation with driver function $g_{\Delta,\Gamma}$. In particular, $Y^{\mathcal{X}}(t) = \mathcal{E}_{\Delta,\Gamma}(\mathcal{X}|\mathcal{F}_t)$ for $t \in [0,T]$.*

Proof. Eqns. (D.2) and (D.3) imply that $\mathcal{X}$ admits the representation in terms of $(Z^{\mathcal{X}}, V^{\mathcal{X}})$

$$\mathcal{X} = Y_t^{\mathcal{X}} - \int_t^T g_{\Delta,\Gamma}(Z_s^{\mathcal{X}}, V_s^{\mathcal{X}})\mathrm{d}s + \int_t^T Z_s^{\mathcal{X}}\mathrm{d}X_s^c + \int_t^T \int_{\mathbb{R}} V_s^{\mathcal{X}}(x)\widetilde{\mu}^X(\mathrm{d}x, \mathrm{d}s), \qquad t \in [0,T]. \tag{D.4}$$

For any $Q \in \mathcal{D}_{\Delta,\Gamma}$ the representation in Eqn. (D.4) can also be re-arranged as

$$\mathcal{X} = Y_t^{\mathcal{X}} + \int_t^T L_s^Q \mathrm{d}s + M_T^Q - M_t^Q \qquad t \in [0,T], \tag{D.5}$$

where $L_s^Q = \sigma Z_s^{\mathcal{X}} H_s^Q + \int_{\mathbb{R}} V_s^{\mathcal{X}}(x)(U_s^Q(x) - 1)\Lambda(\mathrm{d}x) - g_{\Delta,\Gamma}(Z_s^{\mathcal{X}}, V_s^{\mathcal{X}})$ and $M^Q = \{M_t^Q, t \in [0,T]\}$ is given by

$$M_t^Q = \int_0^t Z_s^{\mathcal{X}}\mathrm{d}X_s^{c,Q} + \int_0^t \int_{\mathbb{R}} V_s^{\mathcal{X}}(x)\widetilde{\mu}^{X,Q}(\mathrm{d}x, \mathrm{d}s),$$

which is a square-integrable $\mathbf{F}$-martingale with respect to Q, by virtue of Girsanov's theorem. Taking the $\mathcal{F}_t$-conditional expectation under Q of $\mathcal{X}$ using the representation in Eqn. (D.5) and subsequently taking the essential supremum over all $Q \in \mathcal{D}_{\Delta,\Gamma}$ shows $\mathcal{E}_{\Delta,\Gamma}(\mathcal{X}|\mathcal{F}_t) = Y_t^{\mathcal{X}} + K_t$, where

$$K_t = \operatorname{ess.sup}_{Q \in \mathcal{D}_{\Delta,\Gamma}} \mathbb{E}^Q\left[\int_t^T L_s^Q \middle| \mathcal{F}_t\right].$$

We show next in two steps that $K_t = 0$ for all $t \in [0,T]$, which will complete the proof of the assertion. Firstly, in view of the definition of $\mathcal{D}_{\Delta,\Gamma}$ observe

$$\sigma \int_t^T Z_s^{\mathcal{X}} H_s^Q \mathrm{d}s \le \sigma^2 \int_t^T [(Z_s^{\mathcal{X}})^+\Delta_+ + (Z_s^{\mathcal{X}})^-\Delta_-]\mathrm{d}s \tag{D.6}$$

for all $Q \in \mathcal{D}_{\Delta,\Gamma}$, where the inequality is attained if H^Q is chosen equal to $H^* \in \mathcal{L}^2$ given by

$$H_t^* = \sigma[\Delta_+ \mathbf{1}_{\{Z_t \ge 0\}} + \Delta_- \mathbf{1}_{\{Z_t < 0\}}] \qquad \text{for } t \in [0,T]. \tag{D.7}$$

Secondly, for any $Q \in \mathcal{D}_{\Delta,\Gamma}$ we have

$$\int_{\mathbb{R}} V_s^{\mathcal{X}}(x)(U_s^Q(x)-1)\Lambda(\mathrm{d}x) \leq \int_{\mathbb{R}} V_s^{\mathcal{X}}(x)^+\Lambda_+^Q(\mathrm{d}x) + \int_{\mathbb{R}} V_s^{\mathcal{X}}(x)^-\Lambda_-^Q(\mathrm{d}x), \tag{D.8}$$

where $\Lambda_\pm^Q \in \mathcal{M}_{2,\Lambda}^{ac}$ are the random measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ given by $\Lambda_\pm^Q(\mathrm{d}x) = (U_s^Q(x)-1)^\pm \mathbf{1}_{\{\pm V_s^{\mathcal{X}}(x)>0\}}\Lambda(\mathrm{d}x)$. Note that we have equality in Eqn. (D.8) if U^Q satisfies the relation $U^Q(s,x)-1 = U^Q(s,x)^+\mathbf{1}_{\{V_s^{\mathcal{X}}(x)>0\}} + U^Q(s,x)^-\mathbf{1}_{\{V_s^{\mathcal{X}}(x)<0\}}$ for Λ-a.e. $x \in \mathbb{R}$. Taking the supremum in Eqn. (D.8) over all pairs of measures $(\Lambda_+^Q, \Lambda_-^Q)$ with $\Lambda_\pm^Q \in \mathcal{M}_{2,\Lambda}^{ac}$ satisfying the inequalities $\Lambda_+^Q(A) \leq (\Gamma_+ - \mathrm{id})(\Lambda(A))$ and $\Lambda_-^Q(A) \leq (\mathrm{id} - \Gamma_-)(\Lambda(A))$ for all sets $A \in \mathcal{B}^\Lambda(\mathbb{R})$ yields, in view of the representation in Proposition 1, the inequality

$$\int_{\mathbb{R}} V_s^{\mathcal{X}}(x)(U_s^Q(x)-1)\Lambda(\mathrm{d}x) \leq \mathcal{C}^{\Gamma_+-\mathrm{id}}((V_s^{\mathcal{X}})^+) + \mathcal{C}^{\mathrm{id}-\Gamma_-}((V_s^{\mathcal{X}})^-). \tag{D.9}$$

Equality in Eqns. (D.8) and (D.9) is attained if we take U^Q equal to $U_s^*(x) = \phi_s^*(V_s^{\mathcal{X}}(x))$ with

$$\phi_s^*(x) = \phi_s^{\Gamma_+-\mathrm{id}}(x) I_{A_s^+}(x) + \phi_s^{\mathrm{id}-\Gamma_-}(x) I_{A_s^-}(x)$$

with $A_s^+ = \{x : V_s^{\mathcal{X}}(x) > 0\}$ and $A_s^- = \{x : V_s^{\mathcal{X}}(x) < 0\}$ where $\phi_s^{\Gamma_+-\mathrm{id}}$ and $\phi_s^{\mathrm{id}-\Gamma_-}$ are given by the expression in Eqn. (A.1) with $\overline{F}(x)$ replaced by $\overline{F}_s^+(x) = \Lambda(\{y \in A_s^+ : V_s^{\mathcal{X}}(y) > x\})$ and $\underline{F}_s^-(x) = \Lambda(\{y \in A_s^- : V_s^{\mathcal{X}}(y) < x\})$ respectively. It follows from Lemma 2.7 and the fact $V_s^{\mathcal{X}} \in \widetilde{\mathcal{L}}^2$ that we have $U^* \in \widetilde{\mathcal{L}}^2$.

Finally, observe that for the probability measure $Q^* \in \mathcal{D}_{\Delta,\Gamma}$ that has representing pair (H^*, U^*-1) we have (a) $L_s^{Q^*} = 0$, which can be seen to hold true by observing that for the choice $H^{Q^*} = H^*$ and $U^{Q^*} = U^*$ equality is attained in Eqns. (D.9) and (D.6)] and (b) Q^* is contained in the set $\mathcal{D}_{\Delta,\Gamma}$, which follows in view of the definition of the set $\mathcal{D}_{\Delta,\Gamma}$ which states $Q \in \mathcal{D}_{\Delta,\Gamma}$ and since $H^* \in \mathcal{L}^2$, $U^* \in \widetilde{\mathcal{L}}^2$ and we have

$$H^* \in [-\sigma\Delta_-, \sigma\Delta_+] \quad \text{and} \quad \int_A (U^*-1)(x)\Lambda(\mathrm{d}x) \in [(\Gamma_- - \mathrm{id})(\Lambda(A)), (\Gamma_+ - \mathrm{id})(\Lambda(A))] \text{ for all } A \in \mathcal{B}^\Lambda(\mathbb{R})..$$

Hence we deduce $K_t = 0$ for all $t \in [0,T]$, and the proof is complete. □

References

[1] Anger, B. (1977) Representation of capacities. *Mathematische Annalen* 229, pp. 245–258.

[2] Artzner P., Delbaen F., Eber J.M. & Heath D. (1999) Coherent measures of risk, *Mathematical Finance* 9, pp. 203–228.

[3] Billingsley, P. (1999) *Convergence of Probability Measures*, 2nd ed., Wiley.

[4] Barles, G., Buckdahn, R., Pardoux, E. (1997) Backward stochastic differential equations and integral-partial differential equations. *Stochastics and Stochastic Reports* 60, pp. 57-83.

[5] Boyarchenko, S. I. and Levendorskii, S. Z. (2002) *Non-Gaussian Merton-Black-Scholes theory.* World Scientific.

[6] Carlier, G., Dana, R. A. (2003) Core of convex distortions of a probability *Journal of Economic Theory* 113, pp. 199–222.

[7] Carr, P. and Madan, D. (1999) Option Pricing and the Fast Fourier Transform *Journal of Computational Finance* 2, pp. 61-73.

[8] Kupper, M. and P. Cheridito (2011) Composition of Time-Consistent Dynamic Monetary Risk Measures in Discrete Time *International Journal of Theoretical and Applied Finance* 14, pp. 137–162.

[9] Cherny, A.S. (2007) Pricing and hedging European options with discrete-time coherent risk. *Finance and Stochastics* 11, pp. 537–569.

[10] Cherny, A. and D. Madan (2009), New Measures for Performance Evaluation, *Review of Financial Studies* 22, pp. 2571–2606.

[11] Cohen, S. and Elliot, R. (2010) Comparisons for Backward Stochastic Differential equations on Markov chains and related no-arbitrage conditions. *Annals of Applied Probability* 20, pp. 267–311.

[12] Cohen, S. and Elliot, R. (2010) A general theory of finite state backward stochastic difference equations. Forthcoming, `arxiv: 0810.4957`

[13] Cont, R. and Tankov, P. (2004) *Financial Modelling with Jump Processes.* CRC Press/Chapman & Hall.

[14] Coquet, F., Hu, Y., Memin, J. and Peng, S. (2002) Filtration-consistent nonlinear expectations and related g-expectations. *Probability Theory and Related Fields* 123, pp. 1–27.

[15] Dana R. A. (2005) A representation result for concave Schur concave functions, *Mathematical Finance* 15, pp. 613–634.

[16] Dellacherie, C. (1970) *Quelques commentaires sur les prolongements de capacités.* Séminaire de Probabilités, V. Strasbourg. Lecture Notes in Math. 191, pp. 77–81.

[17] Dellacherie, C. and Meyer, P.A. (1980) *Probabilités et potentiel* (Ch. V–VIII), Hermann, Paris.

[18] Denis, L., Hu, M. and Peng, S. (2011) Function Spaces and Capacity Related to a Sublinear Expectation: Application to G-Brownian Motion Paths. *Potential Analysis* 34, pp. 139–161.

[19] Denneberg, D. (1994) *Non-additive measure and intgral* Theory Decis. Lib. Ser. B Math. Statist. Meth. 27, Kluwer Academic Publishers.

[20] Dolinsky, Y., Nutz, M. and Mete Soner, H. (2012) Weak Approximation of G-expectation. *Stochastic Processes and Their Applications* 122, pp. 664–675.

[21] Duffie, D. and Protter, P. (1992) From discrete- to continuous-time finance: weak convergence of the financial gain process. *Mathematical Finance* 2, pp. 1–15.

[22] El Karoui, N., Peng, S. and Quenez, M.C. (1997) Backward Stochastic Differential Equations in Finance. *Mathematical Finance* 7, pp. 1–71.

[23] Föllmer, H. and Schied, A. (2011) *Stochastic Finance.* De Gruyter, 3rd edition.

[24] Huber, P.J. (1981) *Robust statistics.* John Wiley & Sons.

[25] Jacod, J. and Shiryaev, A.N. (1987) *Limit Theorems for Stochastic Processes*, Springer.

[26] Jobert, A. and Rogers, L. (2008) Valuations and dynamic convex risk measures. *Mathematical Finance* 18, pp. 1–22.

[27] Kupper, M. and Schachermayer, W. (2009) Representation results for law-invariant consistent functions. *Mathematics and Financial Economics* 2, pp. 189–210.

[28] Kunita, H. and Watanabe, S. (1967) On square integrable martingales. *Nagoya Mathematical Journal*, 1967.

[29] Kusuoka, S. (2001) On law invariant coherent risk measures. *Advances in Mathematical Economics* 3, pp. 83–95.

[30] Kusuoka, S. (1995) Limit Theorem on Option Replication Cost with Transaction Costs. *Annals of Applied Probability* 5, pp. 198-221.

[31] Madan, D.B. and Schoutens, W. (2012) Tenor specific pricing *International Journal of Theoretical and Applied Finance* 15.

[32] Pardoux, E., Peng, S. (1990) Adapted solution of a backward stochastic differential equation, *Systems and Control Letters* 14, pp. 55–61.

[33] Peng, S. (1992) A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation *Stochastics and stochastics reports* 38, pp. 119–134.

[34] Protter, P. (2005) *Stochastic Integration and Differential Equations*, Springer.

[35] Riedel, F. (2009) Optimal stopping with multiple priors. *Econometrica*

[36] Royer, M. (2006) Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic Processes and their Applications* 116, pp. 1358–1376.

[37] Roorda, B., Schumacher, H. and Engwerda, J. (2005) Coherent acceptability measures in multiperiod models, *Mathematical Finance* 15, pp. 589–612.

[38] Sarin, R. and Wakker, P. (1992) A simple axiomatization of nonadditive expected utility. *Econometrica* 60, pp. 1255–1272.

[39] Sato, K. (1999) *Lévy processes and infinitely divisible distributions*, Cambridge University Press, Cambridge.

[40] Schied, A. (2004) Risk measures and robust optimization problems. *Lecture notes of a minicourse held at the 8th Symposium on Probability and Stochastic Processes, Puebla.*

[41] Schmeidler, D. (1989) Subjective probability and expected utility without additivity. *Econometrica* 57, pp. 571–587.

[42] Stadje, M. (2010) Extending time-consistent risk measures from discrete time to continuous time: a convergence approach. *Insurance: Mathematics and Economics* 47, pp. 391-404.

[43] Wakker, P. (2001) Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle. *Econometrica* 69, pp. 1039–1059.

[44] Wang, S., Young V. R. and Panjer, H. H. (1997) Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics* 21, pp. 173 - 183.

[45] Wang, S. (2000), A Class of Distortion Operators for Pricing Financial and Insurance Risks, *Journal of Risk and Insurance* 36, pp. 15-36.

[46] Yaari, M. E. (1987) The Dual Theory of Choice Under Risk, *Econometrica* 55, pp. 95-115.

Robert H. Smith School of Business, University of Maryland
E-mail address: dbm@rhsmith.umd.edu

Department of Mathematics, Imperial College London, UK
and Korteweg-de Vries Institute for Mathematics, University of Amsterdam
E-mail address: m.pistorius@imperial.ac.uk

Department of Econometrics and Operations Research, University of Tilburg
E-mail address: m.a.stadje@uvt.nl